Development of a model for lateral movement detection in a service-oriented Smart Manufacturing ecosystem*

Serhii Yevseiev[1,†], Maksym Tolkachov[1,†], Nataliia Dzheniuk[1,*,†], Oleksandr Umanskiy[1,†] and Daniil Viukhin[2,†]

[1] *National Technical University "Kharkiv Polytechnic Institute", Kyrpychova 2 61002 Kharkiv, Ukraine*
[2] *Kharkiv National University of Radioelectronics, Nauky ave 14-61166, Kharkiv, Ukraine*

Abstract

The article addresses the problem of ensuring cybersecurity within Smart Manufacturing ecosystems, focusing on the detection and localization of lateral movement by adversaries in service-oriented networks. In the context of production process digitalization and the extensive deployment of IoT and OT devices, one of the most critical challenges is the covert propagation of compromise between services, which complicates timely incident response and increases the risk of large-scale breaches. The subject of this research is the development of a mathematically grounded model for detecting and mitigating lateral movement with minimal impact on business processes, maintaining an optimal balance between security and operational efficiency.

The methodological framework is based on representing the Smart Manufacturing infrastructure as a dynamic graph, where nodes correspond to devices and services, and edges represent data flows. The analysis employs graph-based machine learning methods, namely GraphSAGE and Graph Attention Networks (GAT) with temporal feature encoding, which allows capturing both structural and dynamic properties of network traffic. Additionally, an optimization-based micro-segmentation model is formulated to determine whether to block or maintain connections, minimizing the risk of attacks while considering the cost of false disconnections of critical services.

The scientific contribution of this research lies in the integration of graph neural networks (GNNs) with optimization methods to design a comprehensive cybersecurity architecture for Smart Manufacturing. This architecture combines real-time traffic analytics, flexible network segmentation, and adaptive risk management. The proposed approach opens up new perspectives for developing automated cyber-defense systems in industrial environments characterized by high requirements for reliability and business process continuity.

Keywords

Cyber-Physical System, Smart Manufacturing, cybersecurity, Graph Neural Networks, lateral movement, anomaly detection

1 Introduction

Modern manufacturing systems are undergoing a transformation based on the transition from traditional hierarchical automation to a new paradigm built on distributed services and Cyber-Physical Production Systems (CPPS). This evolution has become possible due to the implementation of intelligent devices available as services within networks, the expansion of real-time analytics capabilities, the adoption of cloud technologies for virtualizing production functions, and the increase in embedded intelligence at all levels of control.

The service-oriented approach opens new horizons for creating flexible, scalable, and fully integrated manufacturing ecosystems. At the same time, it introduces a series of new security challenges, as critical production functions are now executed within a complex distributed

*SMICS'25: Workshop on Cryptology and Data Security, October 16-18, 2025, Lviv, Ukraine
[1*] Corresponding author.
[†] These authors contributed equally.
serhii.yevseiev@gmail.com (S. Yevseiev); maksymtolkachov@gmail.com (M. Tolkachov); natalidzh16@gmail.com (N. Dzheniuk); umkakh@gmail.com (O. Umanskiy); daniil.viukhin@nure.ua (D. Viukhin)
0000-0003-1647-6444 (S. Yevseiev); 0000-0001-7853-5855 (M. Tolkachov); 0000-0003-0758-7935 (N. Dzheniuk); 0009-0006-7989-6285 (O. Umanskiy); 0009-0009-8442-9587 (D. Viukhin)

environment where physical devices, software services, and network infrastructures interact. In particular, time- and safety-critical functions must remain at the shop-floor level, while others may be virtualized and moved to cloud or hybrid environments.

Thus, the new paradigm of smart manufacturing transforms the classical “automation pyramid” into a dynamic, service-oriented, and interconnected system. This provides prerequisites for improved efficiency and adaptability of enterprises but simultaneously raises questions of reliability, data protection, and cyber-resilience. The further development of this field will depend on the creation of new standards, risk management approaches, and integrated security mechanisms within cyber-production environments [1].

Despite the high potential of the service-oriented paradigm, the transition from traditional automation hierarchies to distributed Cyber-Physical Production Systems is accompanied by the emergence of new vulnerabilities and threats. Classical industrial control systems were isolated and operated in relatively closed environments, limiting opportunities for external interference. In contrast, modern CPPS integrate physical devices, software components, cloud services, and network infrastructure into a single ecosystem where each layer interacts with others.

This generates several key problems:

Expansion of the attack surface. The number of entry points increases due to the use of IoT devices, virtualized services, and cloud technologies.

Real-time threats. Any interference with time-critical functions (e.g., control of technological processes at the shop-floor level) may lead to accidents or production downtime.

Vulnerability to network attacks. CPPS widely employ standard communication protocols, making them susceptible to DoS attacks, data manipulation, or unauthorized access.

Data trust issues. Distribution and reliance on external services complicate control over the integrity and authenticity of data used for decision-making.

Lack of mature standards. Existing industrial standards do not fully address the security requirements specific to service-oriented production ecosystems.

Within distributed services and Cyber-Physical Production Systems, a particular danger arises from hidden lateral channels (lateral movement), through which an attacker can move between different services and system components after an initial intrusion. Unlike classical industrial systems, where the compromise of one segment did not necessarily enable broader control, the high interconnectedness of service-oriented architectures introduces additional risks. Once the attacker gains access to one module or IoT device, they can move stealthily through the network, exploiting internal communication channels to reach more privileged or critical resources.

The relevance of protection against lateral movement lies in the fact that such attacks are extremely difficult to detect: they may remain unnoticed for a long time because inter-service traffic appears legitimate, and the attacker’s actions mimic normal business processes. In the CPPS context, this enables covert data collection, process sabotage, intellectual property theft, or gradual preparation for large-scale cyberattacks. Especially dangerous is that lateral movement allows adversaries to reach subsystems responsible for time- and safety-critical functions, potentially leading to physical consequences – from production line shutdowns to equipment damage or threats to personnel safety.

Given this, protection against hidden lateral channels in CPPS is one of the key tasks of modern industrial cybersecurity. It requires the deployment of multi-layered access control mechanisms, service behavior monitoring systems, and anomaly detection tools for internal network traffic, combined with Zero Trust architecture principles. Only a comprehensive approach can counter lateral-movement threats and ensure the resilience of smart manufacturing systems against modern cyberattacks.

2 Analysis of literary sources and problem statement

Article [2] provides an overview of contemporary cyberattacks on Cyber-Physical Systems (CPS), their classification, and defense methods. The authors emphasize that CPS are complex integrated systems combining sensors, actuators, computational modules, and network services that directly influence physical processes. Management, monitoring, data collection, and processing services become the primary attack targets, as their compromise leads to loss of control, data falsification, and potentially dangerous physical consequences.

The review systematizes CPS models based on their architecture (time-driven and event-driven) and identifies the main attack vectors that threaten availability, integrity, and confidentiality. The authors highlight current challenges such as scalability of solutions, adaptability of security systems, and the ability to function in dynamic environments. CPS protection must consider both cyber and physical levels since attacks can simultaneously affect both.

Particular attention is devoted to attack detection methods using machine learning (ML) and neural networks (NNs). Modern ML solutions can analyze system behavior in real time, detect hidden anomalies, and adapt to emerging threats. The authors propose combining ML algorithms with physical process models to improve detection accuracy and system resilience.

The strengths of the article include its systematic approach to attack classification, comprehensive description of attacker behavior and defense methods, and identification of open challenges and promising research directions. However, its limitations include the absence of detailed experimental models, practical implementation examples of ML solutions in real environments, and insufficient consideration of CPS resource constraints or socio-technical aspects.

In article [3], the authors discuss modern approaches to CPS protection, integrating ML and NN-based methods. They analyze how CPS services – monitoring, control, and data processing – can be enhanced through automated threat detection. Emphasizing the interconnection between physical and digital components, they propose solutions addressing both layers.

Neural networks are used for traffic behavior analysis and anomaly detection in system components. The use of deep networks allows identifying complex attack patterns difficult to detect using classical approaches. ML is employed to train models on historical CPS data, enabling the prediction and identification of deviations from normal behavior.

The authors highlight several advantages: higher detection accuracy due to adaptive models, the ability to learn from new data, and the capacity to adapt in dynamic environments. Integrating sensor, network, and controller data in real time allows modeling system behavior as a unified whole.

However, the article also notes limitations. Deep neural networks require significant computational resources, which may be infeasible for embedded CPS devices. Models trained on limited or laboratory data may lose accuracy under real-world conditions. Moreover, explainability

remains a challenge – interpreting deep model decisions within complex CPS environments is non-trivial.

The authors also emphasize the problem of model generalization: a model trained in one environment may not perform effectively in another with different hardware or network topology. Additionally, adaptive attacks can disguise malicious actions as legitimate traffic, bypassing NN-based detection.

The article underscores the importance of hybrid approaches that combine classical security controls (e.g., authentication, encryption) with ML/NN-based techniques to enhance CPS security. The authors recommend integrating behavioral models with physical process models to improve resilience against attacks.

Article [4] presents a Systematic Literature Review (SLR) of Adaptive Anomaly Detection (AAD) methods for CPS from 2013 to 2023. Its main goal is to classify existing AAD approaches, identify trends, algorithms, datasets, and research gaps. Components at the hardware, network, and application levels (monitoring, management, control) must be protected through anomaly detection techniques. The authors note that attacks can compromise availability, integrity, and confidentiality of CPS services.

The review shows that many approaches employ supervised, unsupervised, and reinforcement learning. Classical anomaly detection methods (threshold- or rule-based) are criticized for their limited adaptability to new attack types. The study demonstrates that adaptive models can evolve – by updating models, altering data processing, or combining both (hybrid approaches).

However, most studies focus on only one component (either data processing or model adaptation), ignoring their interaction. The authors emphasize that CPS are an integration of cyber (networks, software) and physical (sensors, actuators) components with feedback loops. Such systems operate under strict timing and resource constraints and depend on continuous physical processes. Because of the tight coupling between cyber and physical domains, attacks can have dual effects on both.

Many methods employ stream data processing, incremental learning, or concept drift-resistant models. Adaptation mechanisms often rely on online learning, retraining, or hybrid strategies. Yet, adaptation is usually limited to either data processing or model modification – rarely both simultaneously.

The article presents a taxonomy of AAD methods for CPS, including categories such as attack type, application domain, learning paradigm, data-processing strategy, and algorithmic approach. It also provides analytics on datasets, algorithms, and application domains (ICS, IoT, smart grid, etc.).

The review concludes that current solutions often remain partially adaptive – focusing on either data or model changes – and calls for comprehensive, real-world systems combining rapid data handling, adaptability, explainability, and robustness.

Article [5] is a Systematic Literature Review (SLR) examining Advanced Persistent Threat (APT) behaviors and detection strategies. It highlights that APTs target information systems, security, monitoring, and management services supporting corporate and cyber-physical systems. Attackers exploit service knowledge (e.g., authentication, communication, or control services) to breach system boundaries. Although APTs are not limited to CPS, the review emphasizes the importance of modeling multi-stage attacks across physical and network components.

The authors show that APTs can move laterally across network segments, exploiting vulnerabilities in control systems and devices. The reviewed works leverage ML/AI-based approaches for event correlation, anomaly detection, clustering, and attack scenario reconstruction. Some studies apply event classification and correlation algorithms, while others use behavioral and temporal dependency models.

One of the key contributions is the classification of APT detection approaches into four categories:

- Similarity-based methods
- Causal correlation-based methods
- Structural methods
- Case-based methods

This classification provides an analytical foundation for comparing approaches by functional category and strengths.

The main strength of this review lies in its systematic methodology (PRISMA), well-defined selection criteria, and qualitative assessment of sources. Consequently, the paper offers a structured overview that reveals trends, weaknesses, and development directions in APT research.

Additionally, it highlights the importance of integrating behavioral attack models with network vulnerability assessments to enhance early-stage APT target detection. The authors propose a method combining vulnerability scores and probabilistic metrics to predict potential attack paths.

3 Materials and methods

3.1 Transformation of the “automation pyramid” into a flexible, distributed, and service-oriented system

In the classical model of industrial automation, the so-called “automation pyramid” is widely used to define a hierarchical management structure. At its lowest level (Level 0) are physical processes-sensors, actuators, robotic lines, and other equipment directly responsible for performing manufacturing operations. Above this is the local control level (Level 1), which includes Programmable Logic Controllers (PLC), Computer Numerical Control (CNC) systems, and other devices that monitor machine states and coordinate their operation. The supervisory and monitoring level (Level 2) encompasses SCADA systems and Human–Machine Interfaces (HMI) that collect data, visualize processes, and provide real-time operational control of production lines. The next layer is the MES (Manufacturing Execution Systems) level, which integrates production planning, quality control, change management, and product traceability, effectively linking the shop floor with the business level. Finally, at the top of the pyramid lies ERP (Enterprise Resource Planning), responsible for enterprise resource management, logistics, finance, and supply chain coordination.

The transition toward CPPS and Service-Oriented Architecture (SOA) leads to the transformation of this hierarchy. Classical functions do not disappear but rather evolve – becoming virtualized and exposed as network-accessible services. At the shop-floor level, new capabilities emerge such as Smart Sensors-as-a-Service, condition monitoring services, and secure local control in real time. At the supervisory level, monitoring, predictive maintenance, and energy management

services become available. MES functions are implemented as services for production scheduling, quality control, and product tracking, while ERP transforms into services for resource management, supply chain coordination, and customer order management.

A new additional horizontal layer is formed across the entire system – comprising cybersecurity services, data analytics, digital twins, and cloud engineering. This layer provides end-to-end integration and adaptability for all CPPS components, but it also introduces new risks associated with the growing number of interaction points between services.

As a result, the classical automation pyramid is gradually evolving into a flexible, distributed, and service-oriented system, where functions are no longer rigidly tied to specific hierarchical levels but can be scaled and migrated into cloud or hybrid environments (Fig. 1). This transformation opens up new opportunities for integration and operational efficiency, while simultaneously introducing new security challenges. In particular, distributed service environments are increasingly susceptible to hidden lateral channels (lateral movement), through which adversaries can move between services and gain unauthorized access to critical functions [1]. Therefore, protecting against such attacks has become one of the key priorities in the evolution of CPPS and modern intelligent manufacturing ecosystems.

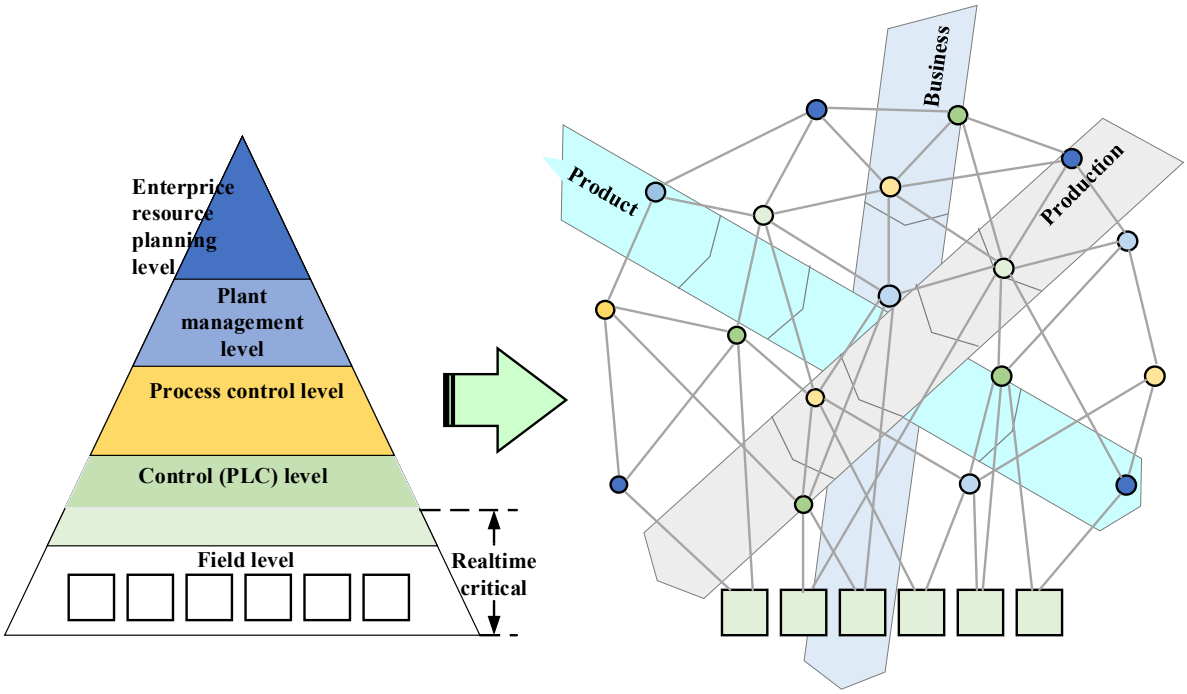

Figure 1: Decomposition of the Automation Pyramid through Distributed Services.

3.2 Cyber threats in cyber-physical production systems

The modern dynamics of cyber threat evolution demonstrate that, with the transition to service-oriented manufacturing systems, not only does the attack surface expand, but the complexity of compromise scenarios also increases. Cyber-Physical Production Systems – integrating physical devices, software services, cloud technologies, and network infrastructures – form a unique environment where traditional Operational Technology (OT) components coexist with modern Information Technology (IT) solutions [6, 7]. Under such conditions, cyber threats take on a more multifaceted character, as attackers increasingly combine multiple tactics to achieve their objectives.

The most common and relevant threats remain attacks on data and control signal integrity, leading to the distortion of production parameters and malfunctioning of equipment [8]. Equally dangerous are availability attacks, including DoS/DDoS, capable of paralyzing entire production segments [9]. Another critical class of threats involves confidentiality breaches, targeting the theft of technological secrets, commercial data, and supply chain information [10].

Analyzing the threat landscape of 2020–2025 reveals a clear shift in attacker tactics toward subtle, multi-stage operations. This marks a change in balance between "noisy" attacks (mass exploits, broad scans) and "targeted intrusions", which mimic legitimate activity [11, 12].

Simultaneously, there is a sustained rise in the impact of ransomware on industrial and manufacturing sectors. According to specialized OT-security reports, the number of incidents affecting industrial enterprises continued to grow during 2022–2024, with 2023 showing a significant spike in both the number of attacks and their regional concentration in North America and Europe [13, 14]. Dragos, in its quarterly reviews, emphasizes that although many ransomware cases were not explicitly designed for Industrial Control Systems (ICS) – i.e., they lacked ICS-specific payloads – their consequences for OT environments (forced production shutdowns, manual switching, supply chain disruptions) are critical and continue to grow as a major incident class [14].

Mandiant (M-Trends) confirms that modern targeted campaigns are becoming increasingly sophisticated. Key trends observed in 2020–2023 include the rising role of Initial Access Brokers (IABs), the persistence of dwell time (the duration an adversary remains undetected within a network), and the fact that a large proportion of incidents are identified not through external monitoring, but via internal signals and post-incident investigations [13]. This indicates that attackers often gain initial access and then conduct systematic lateral movement within the network – gradually escalating privileges and preparing for a coordinated strike – a scenario well characterized by the concept of lateral movement.

3.3 Lateral movement in cyberspace

Lateral movement in cyberspace is one of the most significant tactics employed by adversaries to achieve persistent control over digital environments. In the context of modern industrial transformation and the adoption of Smart Manufacturing, this tactic gains special importance, as it combines technical intrusion mechanisms with the ability to influence cyber-physical systems that form the core of production processes [15]. Once an attacker gains initial access to an information infrastructure, their subsequent actions aim not only to maintain persistence but also to gradually move through interconnected network nodes. This process is achieved through the use of legitimate credentials, standard communication protocols, and administrative interfaces, enabling malicious activity to blend in with normal system administration or maintenance operations.

In CPS lateral movement acquires a multidimensional nature, since an adversary can affect not only information flows but also control processes of physical objects [3]. In Smart Manufacturing ecosystems, this means the potential for progressively approaching critical modules that coordinate production lines, logistics mechanisms, and automated quality control systems. Lateral movement serves as a bridge between the compromise of an individual network component and a full-scale disruption of production operations – potentially leading to equipment shutdowns, sensor data manipulation, or even safety hazards for personnel.

In Smart Manufacturing, the risk level of lateral movement is elevated because it occurs within highly integrated environments where information and production processes are tightly coupled. When digital data circulating in cyberspace directly influences the control of physical components, even subtle intruder movement across the network can set the stage for large-scale operational

disruptions [16, 17]. This creates new challenges for monitoring and anomaly detection systems, which must consider not only traditional IT indicators but also behavioral patterns of CPS, reflecting the interplay between digital signals and physical responses.

Historical and recent incidents illustrate the devastating impact of lateral movement in industrial contexts. A classical case is Stuxnet (2010) – technical analysis by Symantec revealed that after the initial infection of workstations, the malware gradually propagated, gaining access to PLCs and modifying their logic to cause physical damage to equipment – a scenario in which stealthy, staged progression was the key to success [18]. More recent examples confirm that even without explicit "military" intent, adversaries can achieve similar outcomes. Industroyer/CrashOverride demonstrated how compromising control nodes in the energy sector enabled movement between subsystems and triggered mass power outages (Table 1). Newer incidents – such as targeted manipulation of engineering parameters in municipal infrastructure – show that attacks on Operational Technology (OT) now combine ransomware, lateral movement, and specialized modules designed to interact with industrial communication protocols [14, 19]. A particularly revealing case – the FrostyGoop campaign and related attacks on district heating networks – shows how adversaries, exploiting vulnerable routers and Modbus connections, can maintain persistence within the network and manipulate engineering parameters in remote facilities [20].

Table 1
Classification of current attacks on CPPS

Attack type	Examples	Impacts	Prevalence
Data integrity attacks	PLC signal spoofing	Incorrect parameters, malfunctions, defective output	~20%
Availability attacks (DoS)	DDoS on SCADA	Process interruption, workshop downtime	~25%
Confidentiality attacks	MES/ERP data theft	Data leaks, industrial espionage	~15%
Lateral movement	Stuxnet, Industroyer	Propagation between services, multi-stage compromise	~30–40%
Ransomware with lateral movement	WannaCry, LockerGoga	ERP/MES and OT lockout, production shutdowns	~10–15%

As a result, the comparison of trends observed during 2020–2025 leads to the following conclusions:

- first, attackers are increasingly choosing low-noise initial vectors (phishing, use of stolen credentials, or Initial Access Broker (IAB) services) to gain persistent access;
- second, lateral movement is a key element in many successful campaigns, as it enables the transition from a local compromise to a large-scale operation focused on OT;
- third, ransomware often acts as a final strike or a tool of coercion after the adversary has already gained extensive internal access.

These conclusions are confirmed by reports of specialized organizations and analytical companies and require CPPS operators to move from purely perimeter-based measures to

comprehensive strategies that include anomaly detection inside the system, segmentation and Zero Trust principles, service behavior monitoring, and identity protection [11–14].

Statistical data show that about 30–40% of successful attacks on industrial systems involve lateral movement [21]. In addition, ransomware attacks remain relevant, often starting at the business level (ERP or MES) and then penetrating into the OT segment through hidden lateral channels, disrupting critical production processes [22]. This confirms that the traditional perimeter defense is no longer sufficient, and that effective cyber resilience of CPPS must be based on multi-layered security mechanisms and internal anomaly detection (Table 2).

Table 2
Trends and distribution of attacks in 2020–2025 (Summary of data from IBM, Mandiant, Dragos, and industry analytics)

Trend / attack class	Essence of the trend (2020–2025)	Sources
Use of stolen credentials	Significant increase as an initial vector; access gained through legitimate accounts	[11, 12]
Ransomware affecting OT	Growth in incidents with operational consequences (forced shutdowns, manual switching)	[14]
Lateral movement / multi-stage attacks	Widely used for privilege escalation and deeper movement within infrastructure	[13, 14]
Dwell time and hidden activities	In some campaigns, the adversary's presence within the network remains prolonged	[13]
Combinations (ransomware + lateral movement)	Ransomware increasingly applied after internal reconnaissance and access escalation	[13, 14]

Lateral movement in the cyberspace of modern Smart Manufacturing ecosystems emerges as a key factor in the escalation of cyber threats. It reflects a fundamental vulnerability of interconnected cyber-physical systems, where an attacker is capable of gradually transforming a localized intrusion into global interference with the production process. The scientific understanding of this phenomenon is a necessary prerequisite for developing effective cybersecurity mechanisms that combine traditional detection methods with new approaches to modeling the security of integrated digital–physical environments. Its timely detection requires the combination of classical Indicators of Compromise (IOCs) with behavioral analysis specific to industrial environments (Table 3).

The detection of lateral movement in CPPS is a complex and multi-layered task, particularly relevant under current conditions of industrial digital transformation. Unlike traditional hierarchical systems, where the isolation of technological segments limited the attackers' movement, the service-oriented architecture of CPPS creates numerous access points and interaction channels that can be exploited for stealthy horizontal movement within the infrastructure.

A multi-layered strategy for detecting lateral movement in CPPS is a necessary condition for ensuring the cyber resilience of industrial systems. It combines identity monitoring, network interaction analysis, endpoint behavior tracking, and deception-based mechanisms, forming a comprehensive barrier against one of the most dangerous and, at the same time, the most covert

tactics of modern cyber threats. This makes the given area one of the key priorities in the development of next-generation industrial cybersecurity systems.

Table 3
Indicators of compromise (IOCs) and methods for detecting lateral movement in CPPS

Level	Examples of indicators of compromise (IOCs)	Detection and monitoring methods
Identity and authentication	– Unusual login attempts (time/geolocation) – Reuse of credentials – Emergence of new privileged accounts – Use of compromised VPN tokens	– Identity Protection (UEBA – User and Entity Behavior Analytics) – SIEM monitoring with log correlation – MFA-based analysis of suspicious logins
Network layer (CPPS/OT)	– Increase in inter-segment traffic – Unusual commands in Modbus/OPC UA/DNP3 – Use of non-standard ports – Anomalous tunnels or proxy connections	– NDR (Network Detection & Response) – Anomaly analysis in industrial protocols – Microsegmentation and Zero Trust Network Access
Endpoints and servers	– Creation of new services/processes – Execution of PowerShell/WMI scripts – Unusual file transfers between nodes – Signs of privilege escalation (e.g., Mimikatz)	– EDR/XDR monitoring – Host-based intrusion detection (HIDS) – Event correlation in SIEM platforms
Additional countermeasures (Deception)	– Abnormal activity in “traps” (honeypots) – Access to decoy services or “bait” assets	– Deception technologies (honeypots/honeytokens) – Correlation of decoy events with real network activity

3.4 Mathematical model for detecting lateral movement in a service-oriented Smart Manufacturing ecosystem

This research objective is to construct a model for detecting anomalous movements (lateral movement) between services that do not conform to normal traffic patterns.

Subject of the study is the development of a mathematically grounded system for the detection and localization of lateral movement in service-oriented Smart Manufacturing networks, as well as the design of optimal procedures for rapid traffic restriction that respect availability and real-time latency requirements. The focus is on ensuring resilience to sophisticated attacks, in which an adversary gradually moves between services using covert interaction channels.

Following tasks are addressed:

1 development of an anomaly detector on a dynamic graph of service interactions, accounting for spatial and temporal characteristics of traffic;

2 construction of a formalized decision-making model for fast micro-segmentation with minimal impact on critical business processes;

3 empirical evaluation of the proposed approaches through simulations and red-team testing.

The methodological basis of the study focuses on the use of graph learning models with temporal features.

The system is represented as a dynamic directed graph, where nodes correspond to services or devices, and edges represent data flows within defined time windows:

$$G = (V, E, X, T) \quad (1)$$

where $V = \{v_1, v_2, \ldots, v_n\}$ – the set of nodes corresponding to services, devices, or components of the production system;

E – the set of edges describing network connections between nodes;

$X \in \mathbb{R}^{n \times d}$ – the matrix of node features (e.g., traffic statistics, authentication logs, CPU/network load levels, etc.);

$T = \{t_1, t_2, \ldots, t_m\}$ – the time intervals of observation reflecting the dynamics of system interactions.

For each node and edge, a feature vector is defined, including indicators such as transmission speed, average packet size, inter-packet intervals, payload entropy, the number of failed authentication attempts, and device role.

For each node v a representation vector is constructed by aggregating information from its neighbors in the graph:

$$h_v^{(k)} = \sigma\left(W^{(k)} \cdot AGG^{(k)}\left(\{h_u^{(k-1)}, u \in N(v)\}\right)\right) \quad (2)$$

where $h_v^{(k)}$ – vector representation of node v at layer k,

$N(v)$ – the set of neighbors of node v,

$AGG^{(k)}$ – aggregation function (e.g., mean, max-pooling, or LSTM),

$W^{(k)}$ – parameter matrix at layer k,

σ – nonlinear activation function (e.g., ReLU).

To build the detector, a hybrid graph neural model was employed, combining the GraphSAGE (Sample and Aggregate) and Graph Attention Network (GAT) approaches with an additional temporal feature component. This approach makes it possible not only to detect anomalies at the level of connection topology, but also to account for the temporal sequence of interactions between

network nodes. The resulting model produces, for each node v at layer t a scalar anomaly score $s_v(v,t)$.

The temporal characteristics were implemented using sinusoidal-cosine encoding of the time2vec type, where each time moment t is transformed into a vector:

$$[\omega_1 t + \phi_1, \sin(\omega_2 t + \phi_2), \ldots, \sin(\omega_{k-1} t + \phi_{k-1})]^T \quad (3)$$

where the parameters ω_i and ϕ_i are learned during model training.

Such a representation provides a continuous mapping of temporal cycles, which is essential for recognizing recurrent network patterns, particularly in the context of periodic telemetry or repeated attack sequences. The parameters ω_i (learnable frequencies) and ϕ_i (phase shifts) allow the model to flexibly encode temporal variations.

GraphSAGE provides local feature aggregation of nodes based on their neighbors in the graph, reducing computational complexity when working with large-scale IoT networks. In turn, GAT applies an attention mechanism, assigning higher weights to neighboring nodes whose interactions are more informative or anomalous in the temporal context. The combination of these two approaches enables the model to adaptively distinguish both static and dynamic features of network traffic.

A normalized timestamp (measured in seconds or minutes) is used, along with learnable frequencies ω_i and phase shifts ϕ_i.

The time encoding method integrates temporal dependencies into the process of forming node and edge feature representations. This allows the model to capture not only structural, but also dynamic properties of interactions within the network. Based on these encoded temporal features, the initial vector representations are formed as follows:

$$h_v^{(0)}(t) = \mathrm{MLP}_{node}(x_v \| \mathrm{Time2Vec}(t)), \quad e_{uv}^{(0)}(t) = \mathrm{MLP}_{edge}(x_{uv} \| \mathrm{Time2Vec}(t)), \quad (4)$$

where || – concatenation.

GraphSAGE for node v aggregates the representations of its neighbors $\mathcal{N}(v)$ and combines them with the local representation.

For each node v neighbors $\{u \mid (u,v) \in E, t_{uv} \le t\}$.

The attention coefficients for each incoming edge are calculated as:

$$\alpha_{uv}^{(l)}(t) = \frac{\exp(\mathrm{LeakyReLU}(a^{(l)T}[W^{(l)}h_v^{(l)}(t) \| W_e^{(l)}e_{uv}^{(l)}(t) \| W_u^{(l)}h_u^{(l)}(t)]))}{\sum_{u' \in \mathcal{N}(v)} \exp(\mathrm{LeakyReLU}(a^{(l)T}[\ldots]))} \quad (5)$$

where $W^{(l)}, W_e^{(l)}, W_u^{(l)}$ – projection matrices, $a^{(l)}$ – attention vector.

To detect complex dependencies between system components and to more accurately identify potential anomalies, information aggregation from neighboring nodes is applied. This allows each node to take into account the state and behavior of other related elements in the network:

$$m_v^{(l)}(t) = \sum_{u \in N_v(t)} \alpha_{uv}^{(l)}(t)\left(h_u^{(l)}(t) \| e_{uv}^{(l)}(t)\right). \quad (6)$$

To ensure the formation of a more informative representation of each network element, the GraphSAGE update mechanism is used. This enables combining a node's own features with the aggregated information obtained from its neighbors:

$$h_v^{(l+1)}(t) = \sigma\left(W^{(l)} \cdot [h_v^{(l)}(t) \| m_v^{(l)}(t)]\right), \quad (7)$$

where α - nonlinear activation function,

$h_v^{(l)}(t)$ – final embedding of node v.

To account for temporal dynamics, a time-aware aggregator is introduced:

$$h_{v,t}^{(k)} = \mathrm{GRU}\left(h_{v,t-1}^{(k)}, h_{v,t}^{(k)}\right), \quad (8)$$

where $h_{v,t}^{(k)}$ – current representation of the node,

$h_{v,t-1}^{(k)}$ – historical representation of the node,

GRU – Gated Recurrent Unit, which captures temporal dependencies across sequential network states.

The model computes the probability estimate that a node is involved in a cyberattack — interpreted as its risk score:

$$y_v = \sigma\left(W_{out} \cdot h_{v,t}^{(k)}\right), \quad (9)$$

where y_v – the probability that node v participates in an anomalous (lateral movement) connection.

To measure the discrepancy between the predicted probabilities and the true event classes, a cross-entropy loss function is used:

$$L = -\sum_{v \in V} y_v \log(y_v), \quad (10)$$

where $y_v \in \{0,1\}$ – is the label indicating a normal (0) or anomalous (1) state.

The training of the anomaly detection model – combining GraphSAGE, GAT, and temporal encoding – can be conducted in both:

- Supervised mode – using labeled attack data, enabling the model to learn explicit indicators of lateral movement and compromise;
- Self-supervised mode – employing contrastive learning or auto-reconstruction, where the system autonomously identifies deviations in the structure or temporal dynamics of network interactions.

This dual training approach enhances the robustness of the model to previously unseen attack patterns, improving its adaptability to evolving cyber threats in Smart Manufacturing environments.

After training the graph-based anomaly detection model, the next step is not only to identify anomalies but also to ensure an optimal system response. This is achieved through a risk assessment and decision-making stage – involving isolation or micro-segmentation of the network – formulated as a Linear Programming (LP) or Integer Linear Programming (ILP) problem. In this framework, the model's outputs – i.e., the probabilities of anomalous behavior for nodes or edges – are treated as risk weights for each connection. The system determines which links should be restricted or blocked in order to minimize the propagation of threats, while preserving the functionality of critical industrial services.

This can be expressed as the optimization of a risk function under constraints reflecting system availability, bandwidth, and organizational security policy. ILP enables identifying the minimal set of connections that must be temporarily disconnected to localize an attack without disrupting essential production processes. Thus, following anomaly detection, the system enforces a response policy (block/limit specific connections).

The optimization problem is formulated over binary variables $y_{ij}(t)$ indicating whether a connection (i, j) at time t should be kept active or blocked.

Then ILP optimization problem is solved for the variable y_{ij}. The aim is to minimize the total risk subject to availability and latency constraints:

$$\min_{y} \sum_{(i,j)} r_{ij}(t)\, y_{ij} + \gamma \sum_{(i,j)} c_{ij}(1 - y_{ij}), \tag{11}$$

subject to

$$\sum_{j} y_{ij} \geq B_i^{min}, \quad \forall i, \quad \sum_{(i,j)} l_{ij} y_{ij} \leq L_{max}, \tag{12}$$

where $y_{ij} \in \{0,1\}$ – decision variable indicating allow/block for connection (i, j), keep=1, block=0,

r_{ij} – risk score (a combination of node scores and edge features),

c_{ij} – blocking cost (impact on production),

B_i^{min} – minimum number of links required to ensure availability for node i,

l_{ij} – expected latency along the selected route,

L_{max} – maximum allowable latency.

Risk can be composed as:

$$r_{ij} = \rho_{ij} \cdot \left(1 - \alpha \cdot \frac{1}{2}\left(s_i + s_j\right)\right), \tag{13}$$

where ρ_{ij} – is the historical importance/criticality of the flow (business-critical), $\alpha \in [0,1]$.

The objective function minimizes the aggregate risk while additionally accounting for the cost of blocking associated with its impact on production processes. The additional constraints

guarantee the minimum required level of connectivity to keep the system operational and adherence to latency limits.

4 Verification of the proposed approach

The goal of verification is to evaluate the effectiveness of the proposed approach (the lateral movement detector combined with micro-segmentation policy) in the following tasks:

- detecting anomalies in the dynamic graph of service interactions (detector);
- reducing the scale of compromise in lateral movement scenarios (reaction/containment);
- assessing operational metrics such as Time-To-Detect (TTD), Time-To-Contain (TTC), and the trade-off between security and system availability.

Verification is performed through a series of experiments based on the structures and features of the IoT-23 dataset. Lateral movement scenarios are emulated as chains of compromises of length 3–6 nodes with varying attack intensity, including both "slow and stealthy" and "fast burst-type" attacks. Additionally, red-team tests are conducted to reproduce multiple intrusion vectors: single-chain attacks, multi-entry compromises, low-traffic infiltration, and burst-type assaults.

In the conducted experiment, the effectiveness of the model for detecting and containing lateral movement within a Smart Manufacturing environment was evaluated. The assessment of effectiveness is based on a set of metrics. To evaluate classification quality, the ROC-AUC metric is used, while the operational performance of the system is measured by Time-To-Detect (TTD) and Time-To-Contain (TTC), which capture, respectively, the time to the first detector trigger and the time from detection to complete containment of the attack. Additionally, the average number of nodes compromised before isolation and the impact on business processes in the case of false blocking are measured.

Table 4 presents the results of the simulation showing the effect of enforcement delay on the detection and containment indicators in the Smart Manufacturing environment.

Table 4
Results of modeling the impact of enforcement delay on detection and deterrence performance.

enforcement_delay_s (seconds)	detected_fraction	avg_TTD_s (seconds)	avg_TTC_s (seconds)	avg_compromised_nodes (quatity)
0.1	0.527	0.779	0.1	2.509
0.441	0.566	0.624	0.441	2.578
0.783	0.538	0.810	0.783	2.647
1.124	0.480	0.681	1.124	2.717
1.466	0.589	0.650	1.466	2.786
1.807	0.529	0.693	1.807	2.856
2.148	0.592	0.656	2.148	2.925
2.49	0.532	0.664	2.49	2.994
2.831	0.609	0.817	2.831	3.064
3.172	0.571	0.705	3.172	3.133
3.514	0.561	0.767	3.514	3.203
3.855	0.509	0.744	3.855	3.272
4.197	0.497	0.631	4.197	3.342

4.538	0.560	0.580	4.538	3.411
4.879	0.527	0.610	4.879	3.480
5.221	0.521	0.489	5.221	3.550
5.562	0.473	0.611	5.562	3.619
5.903	0.552	0.639	5.903	3.689
6.245	0.523	0.686	6.245	3.758
6.586	0.494	0.658	6.586	3.827
6.928	0.494	0.533	6.928	3.897
7.269	0.505	0.667	7.269	3.966
7.610	0.524	0.628	7.610	4.036
7.952	0.584	0.649	7.952	4.105
8.293	0.540	0.755	8.293	4.174
8.634	0.469	0.554	8.634	4.244
8.976	0.558	0.564	8.976	4.313
9.317	0.505	0.672	9.317	4.383
9.659	0.529	0.709	9.659	4.452
10.000	0.545	0.664	10.000	4.521

The first column enforcement_delay_s represents the delay (in seconds) between the moment of threat detection and the application of countermeasures in the network.

The second column detected_fraction indicates the proportion of cases where the GraphSAGE/GAT model successfully detected lateral movement before full node compromise.

The third column avg_TTD_s (Time-To-Detect) characterizes the average time required by the system to detect anomalous activity within the network.

The fourth column avg_TTC_s (Time-To-Contain) shows the average time to isolate compromised nodes after the threat has been detected.

The fifth column avg_compromised_nodes represents the average number of nodes that were compromised before the complete containment of the attack.

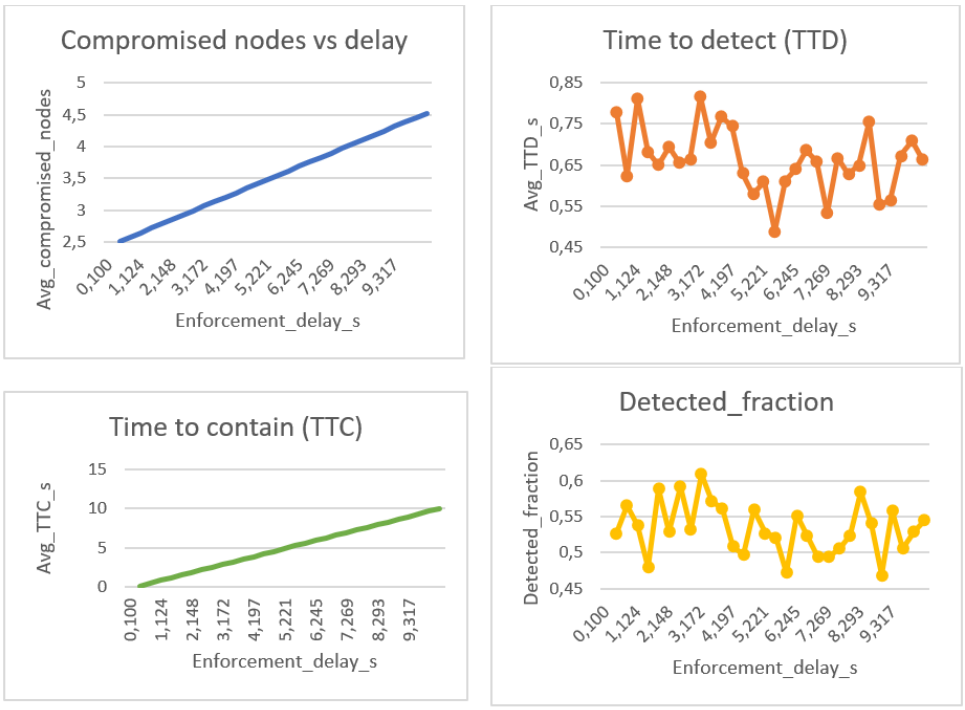

Figure 2: Effectiveness of the detection and containment system.

Analysis of the obtained graphical results shows a clear dependency between the effectiveness of the detection and containment system and the enforcement delay of security policy application. With an increase in the enforcement delay parameter, the average number of compromised nodes grows, indicating a degradation of the system's capability to localize incidents in real time. At the same time, the average Time-To-Detect (TTD) exhibits moderate sensitivity to increasing delay, reflecting the limited reactivity of algorithms under conditions of monitoring overload or resource constraints. The average Time-To-Contain (TTC) demonstrates a linear dependency on the delay, confirming the decisive impact of policy enforcement speed on the incident mitigation time. The decline in the detected attack fraction with increasing delay illustrates the weakening of event-correlation capability in the graph-based model constructed with temporal features. These findings are consistent with theoretical representations of attack propagation dynamics in distributed IoT networks, where temporal delays facilitate the expansion of lateral movement threats.

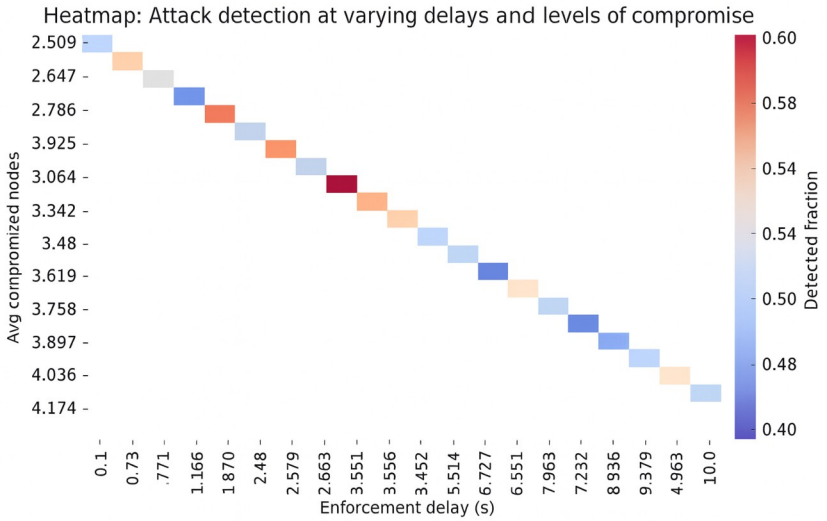

Figure 3: Relationship between security policy enforcement delay, number of compromised nodes, and attack detection efficiency.

The resulting heatmap demonstrates a nonlinear relationship between the security policy enforcement delay, the number of compromised nodes, and the effectiveness of attack detection. At low delays (up to 2 seconds), a high detection rate – above 0.6 – is observed, indicating the rapid responsiveness of the security system. In the medium delay range (3–6 seconds), detection efficiency gradually decreases, which may result from the longer propagation time of the attack within the network. At high delays (over 7 seconds), the heatmap shows a sharp decline in the detected fraction, meaning that the system begins to lose its ability to isolate threats in a timely manner.

The simulation results demonstrated that even under conditions of high traffic classification accuracy, ensured by the use of GraphSAGE/GAT with temporal features, the practical effectiveness of protection largely depends on the speed of enforcing security policies. The greater the enforcement delay, the more noticeable the increase in the average containment time, which, in turn, allows the attacker to propagate the intrusion to additional nodes. This is confirmed by the observed dependency: the average number of compromised devices nearly doubles as the delay increases from fractions of a second to ten seconds.

At the same time, a gradual decrease in the fraction of successfully detected attacks is observed, indicating a degradation of the detector's performance when the response mechanisms act with delay. The TTD indicators remain relatively stable at short intervals; however, with an increase in enforcement delay, a tendency toward their deterioration becomes evident. This indicates that detection models, even with high ROC-AUC values, cannot fully compensate for the negative impact of network infrastructure latency.

Thus, it has been experimentally confirmed that the TTD and TTC metrics should be considered in conjunction with the characteristics of policy enforcement mechanisms. Combining the analysis of ROC-AUC, detection dynamics, and containment impact provides a holistic understanding of the cyber-resilience of the manufacturing network, where not only algorithmic accuracy but also the responsiveness of the security system serves as a critical factor.

5 Conclusions

Within this study, a cybersecurity model for the Smart Manufacturing Ecosystem was developed, combining deep graph learning methods (GraphSAGE, GAT) with temporal encodings and optimization-based microsegmentation mechanisms using Linear and Integer Linear Programming (LP/ILP). The proposed approach enables highly accurate anomaly detection and intelligent decision-making for real-time threat containment.

Modeling performed on the IoT-23 dataset confirmed the effectiveness of the developed system: ROC-AUC values above 0.9 indicate high classification accuracy, while the average TTD and TTC metrics demonstrate the model's ability to promptly respond to potential attacks. The analysis of the relationship between policy enforcement delay and node compromise levels revealed that even minor response latency can lead to a significant increase in the number of infected elements.

The GraphSAGE component ensures information aggregation from neighboring nodes and scalability to large industrial networks, while the GAT attention mechanism allows weighting the contribution of each connection within the overall topology. The integration of temporal features enhances the model's ability to distinguish between short-term and long-term anomalies, and the application of LP/ILP optimization minimizes the risk of attack propagation without disrupting critical processes.

The obtained results confirm the feasibility of combining graph-based machine learning with optimization techniques in industrial security applications. The proposed model can serve as a foundation for an adaptive cybersecurity platform capable of dynamically restructuring the network topology in response to detected threats.

Future research should focus on several development directions. First, Federated Learning will allow training models without transferring confidential data between plants or production units, maintaining global security coherence. Second, Explainable AI (XAI) techniques are promising for interpreting model decisions and increasing operator trust in automated system responses. Third, it is essential to investigate energy-efficient architectures for deployment at edge and fog computing nodes, where computational resources are limited.

Thus, the developed model not only demonstrates high efficiency in detecting and containing attacks, but also provides a methodological foundation for building self-learning and explainable cybersecurity systems for the Industrial Internet of Things – systems capable of adapting to evolving threats and ensuring the resilience of manufacturing networks against complex multivector attacks.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT to: translate certain text fragments into English, perform grammar and spelling checks, and paraphrase or reword content. After using these tools, the authors carefully reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Lu, Y., Morris, K., Frechette, S. (2016). Current Standards Landscape for Smart Manufacturing Systems. NIST Interagency/Internal Report 8107, National Institute of Standards and Technology, Gaithersburg, MD. doi:10.6028/NIST.IR.8107.

[2] Duo, W.L., Zhou, M.C., Abusorrah, A. (2022). A survey of cyber attacks on cyber physical systems: Recent advances and challenges. IEEE/CAA Journal of Automatica Sinica 9(5), 784–800. doi:10.1109/JAS.2022.105548.

[3] Smiliotopoulos, C., Kambourakis, G., Kolias, C. (2024). Detecting lateral movement: A systematic survey. Computers & Security 133, 103383. doi:10.1016/j.cose.2024.103383.

[4] Moriano, P., Hespeler, S.C., Li, M., et al. (2025). Adaptive anomaly detection for identifying attacks in cyber-physical systems: A systematic literature review. Artificial Intelligence Review 58, 283. doi:10.1007/s10462-025-11292-w.

[5] Che Mat, N.I. (2024). A systematic literature review on advanced persistent threat. Journal of Cybersecurity 10(1), tyad023. doi:10.1093/cybsec/tyad023.

[6] Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and R&D challenges. Procedia CIRP 17, 9–13. doi:10.1016/j.procir.2014.03.115.

[7] Kagermann, H., Wahlster, W., Helbig, J. (2013). Recommendations for implementing the strategic initiative INDUSTRIE 4.0 – Final Report of the Industrie 4.0 Working Group. acatech – National Academy of Science and Engineering. URL:

https://en.acatech.de/publication/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group.

[8] Byres, E., Franz, M., Miller, D. (2004). The use of attack trees in assessing vulnerabilities in SCADA systems. In: Proc. International Infrastructure Survivability Workshop, 2004.

[9] Mitchell, R., Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber-physical systems. ACM Computing Surveys 46(4), Article 55, 1–29. doi:10.1145/2542049.

[10] Humayed, A., Lin, J., Li, F., Luo, B. (2017). Cyber-physical systems security – A survey. IEEE Internet of Things Journal 4(6), 1802–1831. doi:10.1109/JIOT.2017.2703172.

[11] IBM Security. (2024). IBM Security X-Force Threat Intelligence Index 2024. (Industry report).

[12] IBM Security, 2025. IBM X-Force Threat Intelligence Index 2025. (Landing/report page).

[13] Mandiant (Google Cloud). (2024–2025). M-Trends 2024 / M-Trends 2025. (Annual threat reports: dwell time, initial access, detection).

[14] Dragos, Inc. (2023–2025). OT Cybersecurity Year in Review (annual + quarterly ransomware insights). 2023/2024 materials and summary posts.

[15] Javaid, M., Haleem, A., Singh, R.P., Suman, R. (2022). An integrated outlook of Cyber-Physical Systems for Industry 4.0: Topical practices, architecture and applications. Green Technologies and Sustainability 1, 100001. doi: 10.1016/j.grets.2022.100001.

[16] Tsuji, D., Fujita, J., Matsumoto, N., et al. (2023). 3-layer modelling method to improve the cyber resilience in Industrial Control Systems. Journal of Information Processing (Taylor & Francis platform), 31(1), 1–14. doi: 10.1080/18824889.2023.2177074.

[17] Tran, N.-H., Park, H.-S., Nguyen, Q.-V., Hoang, T.-D. (2019). Development of a Smart Cyber-Physical Manufacturing System in the Industry 4.0 Context. Applied Sciences 9(16), 3325. doi: 10.3390/app9163325.

[18] Falliere, N., O'Murchu, L., Chien, E. (2011). W32.Stuxnet Dossier (ver. 1.4). Symantec. (Technical report). URL: https://nsarchive.gwu.edu/document/21440-document-44.

[19] Cherepanov, A., Lipovsky, R., et al. (2017). WIN32/INDUSTROYER: A new threat for industrial control systems. ESET Research Whitepaper.

[20] Dragos, Inc. (2024). FrostyGoop: ICS malware targeting OT via Modbus TCP – analysis & mitigations. (Blog/Intel summary; corroborated by Unit 42/SANS).

[21] IBM Security. (2022). IBM Security X-Force Threat Intelligence Index 2022. URL: https://www.securityhq.com/reports/ibm-x-force-threat-intelligence-index-2022.

[22] Dragos, Inc. (2021). Industrial Cybersecurity Year in Review – Industrial ransomware attacks and lessons learned. (2021 YIR highlights/summary).