Received: 23 August 2023 | Revised: 18 December 2023 | Accepted: 11 January 2024 | Published online: 7 February 2024

REVIEW

Journal of Computational and Cognitive Engineering
2024, Vol. 00(00) 1–14
DOI: 10.47852/bonviewJCCE42021539

A Taxonomy of AI Techniques for Security and Privacy in Cyber–Physical Systems

Ajay Bandi[1],*

[1]*School of Computer Science and Information Systems, Northwest Missouri State University, USA*

Abstract: This research paper addresses the concerns related to security and privacy in cyber–physical systems (CPS) and explores the role of artificial intelligence (AI) in addressing these concerns. This paper presents a comprehensive classification of various security and privacy threats in CPS, providing an organized overview of potential risks, economic loss, and enabling effective risk assessment. This paper highlights how AI can help address the security and privacy concerns in CPS by presenting a detailed flow chart that illustrates the step-by-step process of using AI and machine learning (ML) techniques to detect security and privacy issues. This integrated approach serves as a guide for designing ML-based secure CPS, enabling proactive defense mechanisms and improving incident response and recovery. Furthermore, the research explores the various AI techniques that can be employed to address security and privacy concerns in CPS. A taxonomy of ML techniques specifically relevant to security and privacy issues is provided, offering insights into the potential applications of these techniques. In conclusion, this research emphasizes the significance of addressing security and privacy concerns in CPS and highlights the role of AI in tackling these challenges.

Keywords: Internet of Things (IoT), autonomous systems, control systems, real-time systems, Industry 4.0, adversarial machine learning, risk mitigation

1. Introduction

As cyber–physical systems (CPS) became increasingly integrated into our daily lives, concerns about the security and privacy of data generated by these systems are growing. CPS are complex systems that combine physical and computational components, often with a high degree of connectivity and interaction with the internet. This complexity can make them vulnerable to cyberattacks, which can compromise the security and privacy of the data they generate. Therefore, ensuring the protection of data security and privacy in CPS is essential to prevent unauthorized access, tampering, or theft of sensitive information (Singh et al., 2020; Tahsien et al., 2020; Yampolskiy et al., 2012). According to the report of Zion Market Research (2022), the global market size of CPS is expected to increase from $76.98 billion in 2022 to $177.57 billion by 2030, at a compound annual growth rate of 8.01% during the forecast period. The growth is attributed to the advancements in Internet of Things (IoT) technology, which is driving the demand for more efficient and intelligent CPS solutions across various industries such as healthcare, automotive, and industrial automation (Zion Market Research, 2022). In this context, there is a pressing need to address the challenges associated with securing and protecting data generated by CPS and to develop effective strategies and solutions to mitigate potential risks.

One of the most common security issues in CPS is related to the vulnerabilities in the communication channels between the cyber and physical components (Salau et al., 2022). For instance, an attacker may compromise the communication channel to send false commands to the physical components, causing physical damage or disruption to the system. Another common security issue in CPS is related to the lack of authentication and authorization mechanisms for system components. Without proper authentication and authorization, an attacker can gain unauthorized access to the system and manipulate its components (Karale, 2021; Vlachos et al., 2023). One of the most recent and well-known CPS security issues is the SolarWinds cyberattack that occurred in late 2020. The SolarWinds hack was a sophisticated attack on various US government agencies, critical infrastructure, and private companies. It was executed as a supply chain attack by exploiting a flaw in the popular SolarWinds Orion software, which is widely used for IT management. The attackers inserted harmful code into the Orion software, which was then spread to thousands of SolarWinds customers through software updates. This code enabled the attackers to unlawfully access the victims' networks, steal confidential data, and launch additional attacks (Oladimeji & Kerner, 2023). Additionally, CPS are often subject to software vulnerabilities that can be exploited by attackers to compromise the system. These vulnerabilities can arise due to software bugs, design flaws, or inadequate testing procedures. An attacker can exploit these vulnerabilities to take control of the system, steal sensitive data, or cause physical harm. Therefore, the security of CPS is a critical concern, and it requires a multidisciplinary approach involving cybersecurity experts, engineers, and policymakers to develop robust security mechanisms that can protect these systems from cyber threats (Karale, 2021).

IoT, on the other hand, is a network of physical objects that are endowed with sensors, software, and connectivity, enabling them to exchange data and communicate with each other (Bour et al., 2023).

*Corresponding author: Ajay Bandi, School of Computer Science and Information Systems, Northwest Missouri State University, USA. Email: ajay@nwmissouri.edu

The IoT often involves large numbers of devices that are distributed across a wide area, such as a city or a manufacturing plant. IoT devices may be used for a variety of applications, including environmental monitoring, asset tracking, and smart home automation. While CPS and IoT share some commonalities, there are some key differences between them. CPS typically have more complex computational and control systems than IoT devices, as they need to sense and respond to the physical world in real-time (Bharathi & Kumar, 2022; Vlachos et al., 2023). CPS may also require specialized hardware and software to meet specific performance, reliability, and safety requirements. IoT devices, on the other hand, are often simpler and less specialized than CPS, as they may not require the same level of real-time control or safety features. However, IoT devices may need to operate in a more diverse and dynamic environment than CPS, which can pose challenges for security, privacy, and interoperability (Cheh et al., 2017). Table 1 shows the comparison of requirements in CPS and IoT as some of the security and privacy issues of CPS could extend to IoT.

There are various application areas where efforts are being made to address security and privacy issues in CPS. In the field of industrial control systems (ICS), studies have focused on addressing data theft and intrusion attacks (Colelli et al., 2021; Ulybyshev et al., 2021). Solutions such as secure data containers and intrusion detection systems (IDS) have been proposed to safeguard data integrity and monitor system behavior. The use of realistic simulation frameworks like MiniCPS has enabled the development and validation of new defensive strategies for CPS. Weather and satellite applications (Cali et al., 2021) have also been an area of interest. Researchers have proposed frameworks like the Internet of Predictable Things to enhance the resilience of CPS against cybersecurity risks. Adversarial machine learning (AML) attacks in IoT environments have been addressed through the use of machine learning (ML)-based net load forecasting algorithms and the cyberattack detection algorithm.

In the manufacturing domain (Jamal et al., 2023), researchers have focused on cyber–physical security for electric vehicles and developed metrics to measure performance degradation caused by cyber–physical attacks. ML techniques have been employed to process large amounts of data and detect various types of attacks. In the healthcare domain, ensuring the reliability and security of modeled systems against tampering with sensor data is crucial. ML algorithms have been applied to detect data breaches, improve cloud security, and develop frameworks for IoT cloud deployment. The significance of addressing security breaches in healthcare cyber–physical systems (HCPS) is emphasized to prevent unauthorized access to sensitive health data and potential misdiagnosis or incorrect treatment (Bharathi & Kumar, 2022; Khan et al., 2020; Mboweni et al., 2021).

Water treatment is another major application area where security attacks have occurred in CPS. Researchers have focused on intrusion detection mechanisms (Abbas et al., 2015; Junejo & Goh, 2016; Pordelkhaki et al., 2021), false data injection (Pordelkhaki et al., 2021), denial of service (DoS) attacks (Perrone et al., 2021), spoofing (Balduzzi et al., 2014; Perrone et al., 2021), authentication (Balduzzi et al., 2014), and the identification of anomalies in water systems (Abbas et al., 2015; Feng & Tian, 2021; Mboweni et al., 2021). ML methods have been utilized for anomaly detection and vessel trajectory prediction to improve maritime surveillance (Liu et al., 2021). The power domain has seen studies on ML-based detection of attacks in water and power grids (Jamal et al., 2023; Junejo & Goh, 2016; Li et al., 2021; Sengan et al., 2021). False data attacks in ML systems and physical attacks causing damage, device overheating, and power outages have been addressed. Researchers have developed security models and detection mechanisms to mitigate risks and maintain the reliability of power systems (Ashok et al., 2015; Mohammadhassani et al., 2020). In the transportation domain (Cheh et al., 2017; Joo et al., 2018; Mo & Sinopoli, 2012), researchers have explored methods to protect highly confidential information from cyberattacks and assess the risk using models such as Factor Analysis of Information Risk and Crime Prevention through Environmental Design. Detection mechanisms and security models have been developed to respond to physical attacks and ensure system resilience.

The main objective of this paper is to investigate the security and privacy concerns associated with CPS. The authors conducted a thorough analysis of existing literature and made several valuable contributions, such as categorizing the various security and privacy issues in CPS, identifying different domains where CPS are used, and exploring the use of artificial intelligence (AI) techniques to address such issues.

1.1. Contributions of this research

It is crucial to prioritize data protection in CPS systems, especially considering the integration of AI and CPS, which is expected to bring revolutionary advancements in the next decade, alongside the development of 6G communication technologies.

Table 1
Comparison of requirements in CPS and IoT

Requirements	CPS	IoT
Integration of physical and digital components	Tightly coupled and coordinated integration of physical and digital components	Loosely coupled integration of physical and digital components
Real-time responsiveness	High emphasis on real-time responsiveness to ensure timely and precise control actions	Real-time capabilities are beneficial but not always necessary
Safety and reliability	Stringent safety measures and fault-tolerant mechanisms due to critical consequences of failures	Safety and reliability are important but with potentially less critical consequences
Scalability and heterogeneity	Smaller number of interconnected components with higher complexity. Integration of diverse physical and computational entities	Massive scale with billions of interconnected heterogeneous devices. Management of connectivity and interoperability on a global scale
Security and privacy	Paramount importance with measures to ensure integrity, confidentiality, availability, and physical component safety	Significant concerns with device authentication, data encryption, secure communication, and resource-constrained challenges

This study makes several contributions. Firstly, it presents a comprehensive classification diagram that encompasses various security and privacy threats in CPS. Secondly, it explores the utilization of AI in addressing these security and privacy concerns. Thirdly, it provides a taxonomy of AI techniques employed for securing CPS. These contributions collectively enhance our understanding of the security and privacy landscape in CPS and provide valuable insights for developing robust defense mechanisms. Furthermore, this research sheds light on potential challenges and issues that may arise in the future regarding the implementation of CPS systems in terms of security and privacy.

1.2. Scope of the review

CPS systems use hardware devices such as sensors, actuators, microcontrollers, and robotic components that are embedded with computer systems designed to perform specific functions. However, the data generated from these devices are prone to vulnerabilities. To protect these data, researchers have incorporated AI techniques. Table 2 provides a summary of the merits and demerits of existing surveys. The scope of this research focuses on developing taxonomies of security and privacy issues among various application domains and the AI techniques used to address these issues in CPS.

Table 2
Comparison of this study with existing surveys

Authors/Year	Title	Merits	Limitations
Haque et al. (2021)	A survey of machine learning-based cyber–physical attack generation, detection, and mitigation in smart-grid	The paper concentrates on identifying and reducing attacks by examining the most recent research in the SG (Smart Grid) field	The paper did not put much emphasis on categorizing the security and privacy issues in CPS or discussing other application areas besides the smart grid
Zhang et al. (2022)/2021	Deep learning-based attack detection for cyber–physical system cybersecurity: A survey	The paper presents a six-step DL-driven methodology to summarize and analyze the literature review on using DL methods to identify cyberattacks against CPS systems	The paper does not offer a classification system for categorizing different security and privacy concerns and CPS application areas. Moreover, other AI techniques are not concentrated
Karale (2021)	The challenges of IoT addressing security, ethics, privacy, and laws	This paper provides a comprehensive review of the security, ethical, and privacy issues faced by everyday users of IoT. It also examines the current and developing regulations and standards established by governments globally to address these vulnerabilities	The survey did not include the role of AI, and it did not concentrate on the different application areas of IoT
Hasan and Roy (2021)	Trending machine learning models in cyber–physical building environment: a survey	This review explored the applications of different ML algorithms, such as deep learning, transfer learning, active learning, and reinforcement learning, along with other emerging techniques, to tackle challenges in the building environment of CPS	The paper did not offer a complete categorization of AI techniques that can be used to ensure security and privacy in CPS. Furthermore, the paper did not concentrate on other areas of CPS besides the buildings
Olowononi et al. (2021a)/2020	Resilient machine learning for networked cyber-physical systems: A survey for machine learning security to securing machine learning for CPS	This paper explores the relationship between resilient CPS using ML and resilient ML algorithms when applied to CPS	The paper did not concentrate on organizing the security and privacy problems in CPS into categories or discussing additional application areas in CPS
Tahsien et al. (2020)	Machine learning-based solution for security of Internet of Things (IoT): a survey	This literature review focuses on ML approaches for securing IoT, discussing their importance considering potential attacks and presenting ML-based solutions. The review also considers future challenges that may arise in this field	The survey concentrates exclusively on machine learning (ML) techniques, and other AI methods are not given much attention
Asghar et al. (2019)	Cybersecurity in industrial control systems: issues, technologies, and challenges	This review analyzes possible cyberattacks on industrial control systems (ICSs), typical threats and vulnerabilities, and the shortcomings of current ICS cybersecurity solutions	Neither the role of artificial intelligence (AI) nor the various application areas of IoT were taken into consideration

This survey does not concentrate on blockchain-based security, since the primary focus is on AI techniques.

1.3. Organizing and reading map

The introduction section of this paper presents the need and motivation for conducting the research, highlighting the similarities between CPS and IoT and drawing comparisons with existing surveys. This paper is structured as follows: Section 2 outlines the research questions (RQs) and methodology employed in the survey. The study results are presented in Sections 3–5. Section 3 covers various security issues in CPS, while Section 4 discusses how AI helps address security and privacy concerns in different application areas of CPS. Section 5 illustrates the taxonomy of AI techniques used in CPS. Section 6 delves into the research's significance, limitations, and challenges in implementing future CPS systems. Finally, Section 7 provides the study's conclusion.

2. Methodology

The overall research goal is to investigate the various security and privacy issues encountered in the literature regarding CPS and to identify the different application areas that utilize AI techniques to address data protection issues in CPS. The specific RQs and their objectives are presented below.

RQ1

What are the concerns related to security and privacy in CPS? A classification of various security and privacy concerns is presented to answer this RQ. The purpose is to recognize a variety of attacks that could occur in CPS systems.

RQ2

How can AI help address CPS's security and privacy concerns? A flow chart is presented and explains how AI/ML detects security and privacy issues. The purpose is to help provide an integrated approach to design an ML-based secure CPS.

RQ3

What AI techniques are used to address these concerns in CPS? A taxonomy of various ML techniques concerning security and privacy issues is provided to achieve this.

To achieve the research goal, we searched Google Scholar, Semantic Scholar, ACM Digital Library, and IEEE Xplore using the search string" (security or privacy) AND (AI OR artificial intelligence OR ML OR machine learning) AND CPS OR cyber-physical systems". The search strategy is illustrated in Figure 1. We carefully examined the results of the search string and included the related studies in our review. Our inclusion criteria involved a clear definition of specific attacks in CPS and the use of AI or ML techniques to resolve those attacks in various CPS applications. The articles that were not related to IoT or CPS or not related to our research goals are excluded. The studies related to blockchain-based security defense mechanisms are not included because it was beyond the scope of the research but plan to explore this area in future work.

3. Security and Privacy Issues in CPS

This section presents the answer to the RQ1. What are the concerns related to security and privacy in CPS. To answer this RQ, a classification system for different security and privacy concerns is presented in Table 3. The purpose is to recognize a variety of attacks that could occur in CPS systems. The categorization is based on the nature and primary focus of each security and privacy issue. Network-based threats primarily target the network layer, software-based threats focus on vulnerabilities in software components in CPS systems, and ML-based threats specifically exploit weaknesses in ML models. Table 3 also includes the information of the economic loss to the CPS due to the attacks. Figure 2 shows the further classification of the security threats in CPS.

3.1. Network-based threats

The security issues categorized under network-based threats involve attacks that exploit vulnerabilities in the network infrastructure or communication channels. Intrusion attacks, DoS, spoofing, cyberattacks, and vessel trajectory attacks are all examples of attacks that target the network layer. These threats aim to gain unauthorized access, disrupt services, manipulate data, or compromise the integrity and availability of the network. An

Figure 1
Search strategy

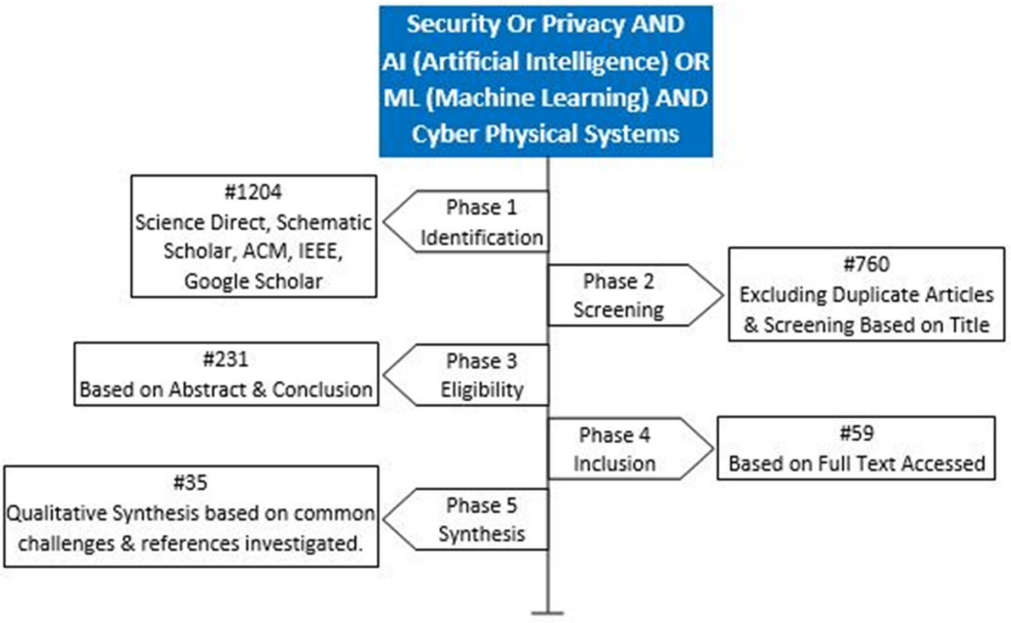


Table 3
Classification of security and privacy issues in CPS

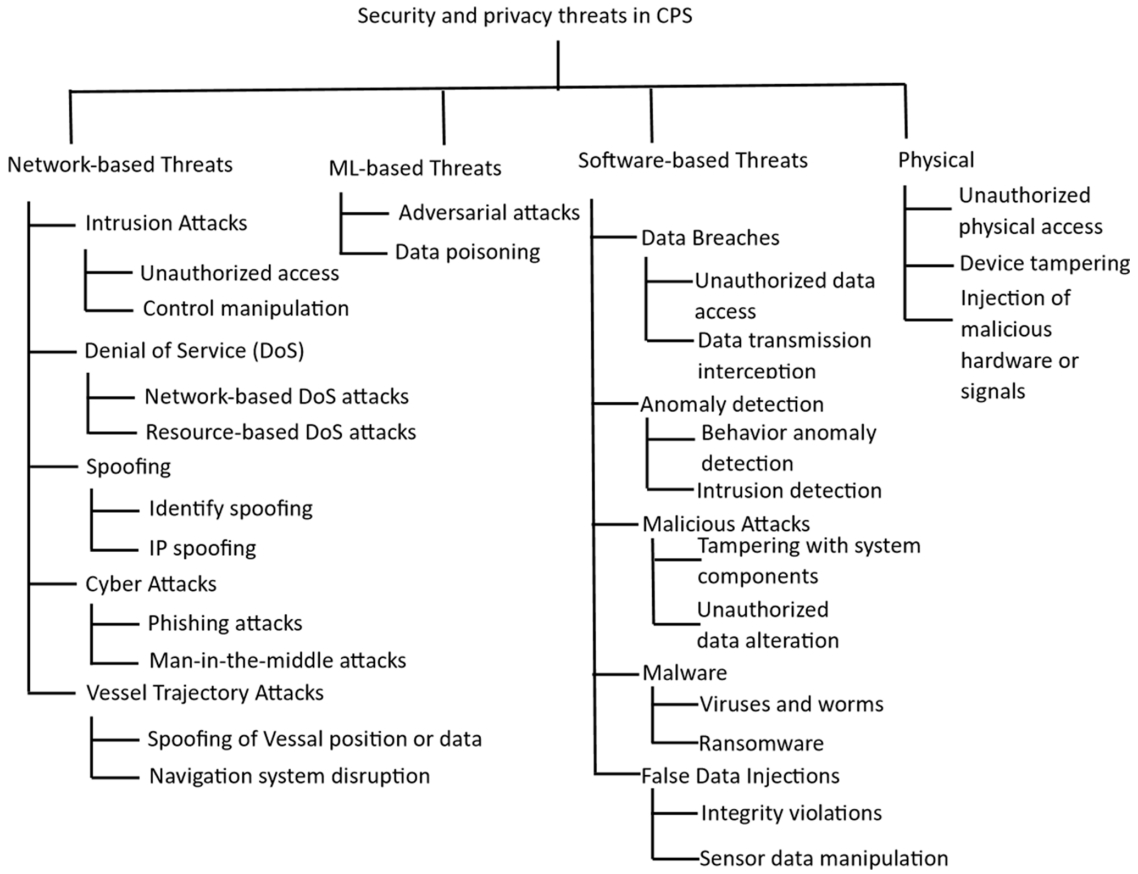
Issue	Subcategories	Explanation	Economic loss
Network-based threats			
Intrusion attacks (Abbas et al., 2015; Colelli et al., 2021; Junejo & Goh, 2016; Pordelkhaki et al., 2021)	Unauthorized access	Gaining unauthorized access to CPS systems or networks	Financial losses from system breaches
	Control manipulation	Unauthorized manipulation or control of CPS system	Operational disruptions
Denial of service (DoS) (Perrone et al., 2021)	Network-based DoS attacks	Overwhelming CPS systems with excessive network traffic, rendering them inaccessible or unusable	Financial losses from system unavailability
	Resource-based DoS attacks	Exhausting system resources (e.g., memory, CPU) to disrupt or degrade the performance of CPS systems	Reduced productivity
Spoofing (Balduzzi et al., 2014; Perrone et al., 2021; Sengan et al., 2021)	Identity spoofing	Falsifying or impersonating identities to gain unauthorized access or deceive CPS systems	Financial losses from unauthorized access
	IP spoofing	Manipulating or forging IP addresses to mask the true source or location of network traffic	Fraudulent activities
Cyberattacks (Balduzzi et al., 2014; Jamal et al., 2023; Joo et al., 2018)	Phishing attacks	Deceiving users through fraudulent communications to obtain sensitive information or access credentials	Financial losses from compromised systems
	Man-in-the-middle attacks	Intercepting and tampering with communications between CPS systems or users, potentially extracting sensitive information	Financial losses from compromised navigation systems
Vessel trajectory attacks (Liu et al., 2021)	Spoofing of vessel position or data	Manipulating or falsifying vessel position or trajectory data within CPS systems, potentially leading to navigational hazards or unauthorized access	Financial losses from compromised navigation systems
	Navigation system disruption	Disrupting or manipulating the navigation systems of vessels within CPS systems, affecting their course or control	Operational disruptions
ML-based threats			
ML-powered attacks (Cali et al., 2021; Li et al., 2021)	Adversarial attacks	Exploiting vulnerabilities or manipulating machine learning models to deceive or compromise CPS systems	Financial losses from compromised systems
	Data poisoning	Introducing malicious or misleading data to bias or manipulate the training process of machine learning models in CPS systems	Loss of data integrity
Software-based threats			
Data breaches (Bharathi & Kumar, 2022; Mo & Sinopoli, 2012; Ulybyshev et al., 2021)	Unauthorized data access	Unauthorized access to sensitive data stored in CPS systems	Legal fines, reputation damage
	Data transmission interception	Intercepting and unauthorized monitoring of data transmission within CPS systems	Loss of customer trust
Anomaly detection (Abbas et al., 2015; Feng & Tian, 2021; Mboweni et al., 2021; Pordelkhaki et al., 2021)	Behavior anomaly detection	Detecting deviations from expected behavior or patterns within CPS systems that may indicate potential security breaches or system malfunctions	Operational disruptions
	Intrusion detection	Identifying and alerting on suspicious activities or attempts to infiltrate CPS systems	Loss of system availability
Malicious attacks (Mboweni et al., 2021; Perrone et al., 2021)	Tampering with system components	Unauthorized alteration, modification, or sabotage of hardware or software components within CPS systems	Damages to hardware or software
	Unauthorized data alteration	Unauthorized modification, deletion, or corruption of data within CPS systems	Loss of data integrity
Malware (Cali et al., 2021; Sengan et al., 2021)	Viruses and worms	Malicious software that can self-replicate and spread across CPS systems, causing damage or disruption	Financial losses from system damage
	Ransomware	Malware that encrypts data or systems, demanding a ransom for their release or restoration	Loss of data, financial losses from ransom
False data injections (Pordelkhaki et al., 2021; Sengan et al., 2021)	Integrity violations	Introducing inaccurate, modified, or fabricated data into CPS systems, compromising their integrity and reliability	Financial losses from incorrect decisions or actions

(*Continued*)

Table 3
(*Continued*)

Issue	Subcategories	Explanation	Economic loss
	Sensor data manipulation	Tampering with sensor data, leading to incorrect decisions or actions based on faulty information	Operational disruptions
Physical-based threats			
Physical attacks (Abbas et al., 2015; Khan et al., 2020; Perrone et al., 2021)	Unauthorized physical access	Unauthorized physical access to CPS components or infrastructure	Potential economic losses from compromised physical security
	Device tampering	Tampering with CPS hardware, sensors, or actuators	Financial losses from compromised device functionality
	Injection of malicious hardware or signals	Introducing malicious hardware components or signals into CPS systems	Risk of economic losses due to compromised system integrity

Figure 2
Classification of security and privacy threats in CPS



IDS is a tool used to identify and alert about any malicious activities that may compromise network security or data stored on connected computers. The system monitors the network continuously and generates alerts when suspicious activity is detected, which can be further investigated by a security analyst or incident responder to mitigate the threat.

Pordelkhaki et al. (2021) explored the use of an ML-based network intrusion detection system (NIDS) for an ICS using a secure water treatment testbed. They combined network traffic data with physical process data from a pre-labeled dataset and evaluated the effectiveness of using privileged information as a supervised learning technique to enhance the detection of network intrusion attacks. They found that this approach was more effective than other ML algorithms that used network traffic data alone. The authors also discussed various other ML algorithms, including the support vector machine plus algorithm (SVM+), decision tree (DT)

algorithm, K-nearest neighbors (KNN), logistic regression, and convolutional neural networks (CNNs), and evaluated their performance in detecting network intrusion attacks. They found that SVM+ outperformed the other algorithms in terms of F1 score, although the dataset used was imbalanced in nature. Colelli et al. (2021) developed a ML tool that detects cyberattacks in CPS to improve their security. They evaluated the performance of three models in classifying normal and anomalous behavior in a water tank system to identify attacks and prevent hazardous conditions. The results were promising, as the ML approach effectively detected and prevented cyberattacks. They also explored the use of supervised ML with the random forest (RF) algorithm to enhance IDS capabilities. They found that this approach had high accuracy and detection rates for both binary and multiclass classification and outperformed other ML algorithms in terms of accuracy, detection rates, and false-positive rates.

DoS attacks aim to make a resource unavailable by disrupting the services of a connected host, while spoofing involves pretending to be something else to gain access to a system for malicious purposes. Perrone et al. (2021) conducted research on using intelligent threat detection to identify malicious activities and anomalous events in water systems that are regulated by SCADA. They compared the
AQ1 effectiveness of different ML techniques such as KNN, NB, SVM, DT, and RF to classify these activities, with RF showing the most reliable performance. In the future, the security of CPS will depend on the use of AI to automate threat identification and countermeasures through SOAR systems, which will enhance situational awareness, emergency response, and crisis management. A cyberattack refers to any effort to gain unauthorized access to a computer, computing system, or computer network with the intention of causing harm. The objective of a cyberattack is to disable, disrupt, destroy, or control computer systems, or to modify, block, delete, manipulate, or steal the data stored within these systems. According to Jamal et al. (2023), ML techniques are crucial for detecting various attacks in CPS, such as replay attacks, DoS attacks, Jamming attacks, time synchronization attacks, and false data injection attacks (FDIAs). Their survey focuses on cyberattacks in different CPS industries, including industrial, construction, cyber manufacturing, and electric power.

AQ2 Liu et al. (2021) suggest that ML algorithms like CNN, LSTM, and hybrid models can effectively predict vessel trajectories by taking into account vessel characteristics, historical movement patterns, and environmental variables, aided by advanced sensor technologies such as AIS and GPS. The SFM-LSTM model combines LSTM with the social force model, providing an accurate and reliable approach for vessel trajectory prediction and enabling smart traffic services in marine transportation systems with the help of AI and IoT technologies. Additionally, data-driven frameworks using LSTM and GRU models have also been used for vessel trajectory prediction.

3.2. ML-based threats

The security issue categorized as an ML-based threat specifically relates to attacks that exploit vulnerabilities in ML models. ML-powered attacks refer to malicious activities that manipulate or deceive ML models within CPS. These threats can include adversarial attacks or techniques that tamper with training data or model outputs, compromising the accuracy, reliability, or robustness of the ML algorithms utilized.

"ML-powered attacks" refer to cyberattacks that utilize ML algorithms to execute malicious activities by identifying and exploiting system vulnerabilities. These types of attacks are becoming more widespread as ML technologies are increasingly adopted across industries. "Adversarial attacks" are one type of ML-powered attack where hackers manipulate ML models by injecting malicious inputs to cause unintended behavior. Organizations need a comprehensive strategy that includes monitoring, detection, and prevention methods, such as anomaly detection, model retraining, and data validation, to safeguard against ML attacks.

To address the potential vulnerabilities of ML in CPS, Li et al. (2021) proposed a defense mechanism called constrained adversarial machine learning (ConAML). ConAML generates adversarial examples that adhere to the intrinsic constraints of physical systems, and a general threat model and the best effort search algorithm were developed to iteratively generate adversarial examples. The authors tested the algorithms on power grids and water treatment systems through simulations, and the results showed that ConAML was effective in generating adversarial examples that reduced the performance of ML models, even under practical constraints. Additionally, the study recommended using techniques such as adversarial detection and retraining to enhance neural networks' resilience against ConAML attacks.

3.3. Software-based threats

The security issues classified as software-based threats primarily focus on vulnerabilities and attacks related to software components in CPS. Data breaches, anomaly detection, malicious attacks, malware, and false data injections are all software-related concerns. These threats target the software layer of CPS and encompass breaches of sensitive data, the detection of anomalous behavior or patterns, the injection of malicious code, and the dissemination of false or manipulated data.

Data breaches occur when sensitive or confidential information is accessed, stolen, or exposed without authorization through various methods like hacking, phishing, physical theft, or human error. Such breaches can result in serious consequences, such as financial losses, reputational damage, legal liabilities, and identity theft. Attacks on cloud-based infrastructure, services, or applications are called cloud security attacks, which pose new security risks and challenges for organizations despite offering flexibility, scalability, and cost savings. To prevent such attacks, organizations need to implement robust security measures like access controls, encryption, firewalls, and intrusion detection and prevention systems, along with regular monitoring and security audits.

Bharathi and Kumar (2022) have proposed a new approach for detecting attacks on HCPS by combining wise greedy routing, agglomeration mean shift maximization clustering, and multi-heuristic cyber ant optimization-based feature extraction. The system employs an ensemble crossover XG Boost classifier to identify attacks and has displayed promising results in terms of accuracy and reducing false positives. In addition, the authors have examined the positive aspects of HCPS compared to the current healthcare system, and the negative effects of cyberattacks on IoT devices and the current limitations of cloud-based security in this context. The authors have also discussed the use of ML-based ensemble crossover XG boost classifiers in healthcare using Matlab simulations, which demonstrated a 99.642% accuracy rate, a 95% precision accuracy, and an F1 score of 98.5%. Anomaly attacks exploit abnormal behavior or patterns in a system or network with the aim of gaining unauthorized access or causing harm. Essentially, they exploit system weaknesses in behavioral patterns to achieve malicious goals.

Feng and Tian (2021) propose a new method called neural system identification and Bayesian filtering (NSIBF) for detecting

anomalies in time series data in CPS. NSIBF uses a customized neural network to identify the system in CPS and a Bayesian filtering algorithm to detect anomalies by monitoring the uncertainty of the system state. The authors evaluated NSIBF on synthetic and real-world datasets, including the PUMP, WADI, and SWAT datasets. They found that NSIBF outperformed existing techniques by 2.9%, 3.7%, and 7.6% at the F1 score on the PUMP, WADI, and SWAT datasets, respectively. Additionally, NSIBF showed significantly better performance compared to NSIBF-RECON and NSIBF-PRED on all three datasets. These results demonstrate the effectiveness of NSIBF for detecting anomalies in complex CPS with noisy sensor data and highlight the advantage of using a neural-identified state-space model and Bayesian filtering to detect anomalies in CPS signals over time. Other researchers, for example, Dhir and Kumar (2020) focused on deep learning (DL) techniques to detect anomalies. The act of intentionally trying to compromise the security, integrity, or availability of a system or network is referred to as a malicious attack, and it can take various forms, such as DoS attacks and social engineering.

Malware refers to software that is installed on a computer without the user's knowledge or consent, and it carries out harmful activities like stealing passwords or money. There are several techniques for identifying malware, but the most common one is to scan the computer for malicious files or programs. Sengan et al. (2021) propose a solution for detecting malware attacks in smart grids by analyzing power system information and signals. Malware can corrupt voltage data, resulting in fraudulent output, and the proposed solution uses an artificial feed-forward network (AFN) with a distance metric cost function to differentiate between secured and malicious data. AFN is capable of handling complex functions and is suitable for the task of identifying malware incidents in smart grids. The solution aims to enhance the security of smart grids by detecting and preventing malware attacks. The alteration of sensor measurements by FDIAs can pose a significant threat to a system's computational capabilities and lead to cyberattacks. Detecting such attacks is crucial to maintain system integrity and security. Sengan et al. (2021) highlighted the importance of detecting these attacks in the smart grid and proposed a true data integrity agent-based model (TDI-ABM) to effectively distinguish between secured data and data generated by intruders. The TDI-ABM can mitigate the effects of FDIA, improve the security of smart grids, and is based on DL applications with various methods and algorithms used to retrieve data from the network.

3.4. Physical threats

While the initial list provided focused on other types of threats, it is important to note that physical attacks can pose significant security risks to CPS. Physical threats target the physical components of a CPS and can have serious consequences. Examples of physical attacks include physical tampering, supply chain attacks, side-channel attacks, and physical destruction. These attacks involve unauthorized manipulation of hardware, compromising the supply chain, exploiting physical information leakage, or causing physical damage to the CPS (Abbas et al., 2015; Khan et al., 2020; Perrone et al., 2021). Ensuring physical security measures are in place is crucial to protect the integrity and functionality of a CPS against these types of attacks.

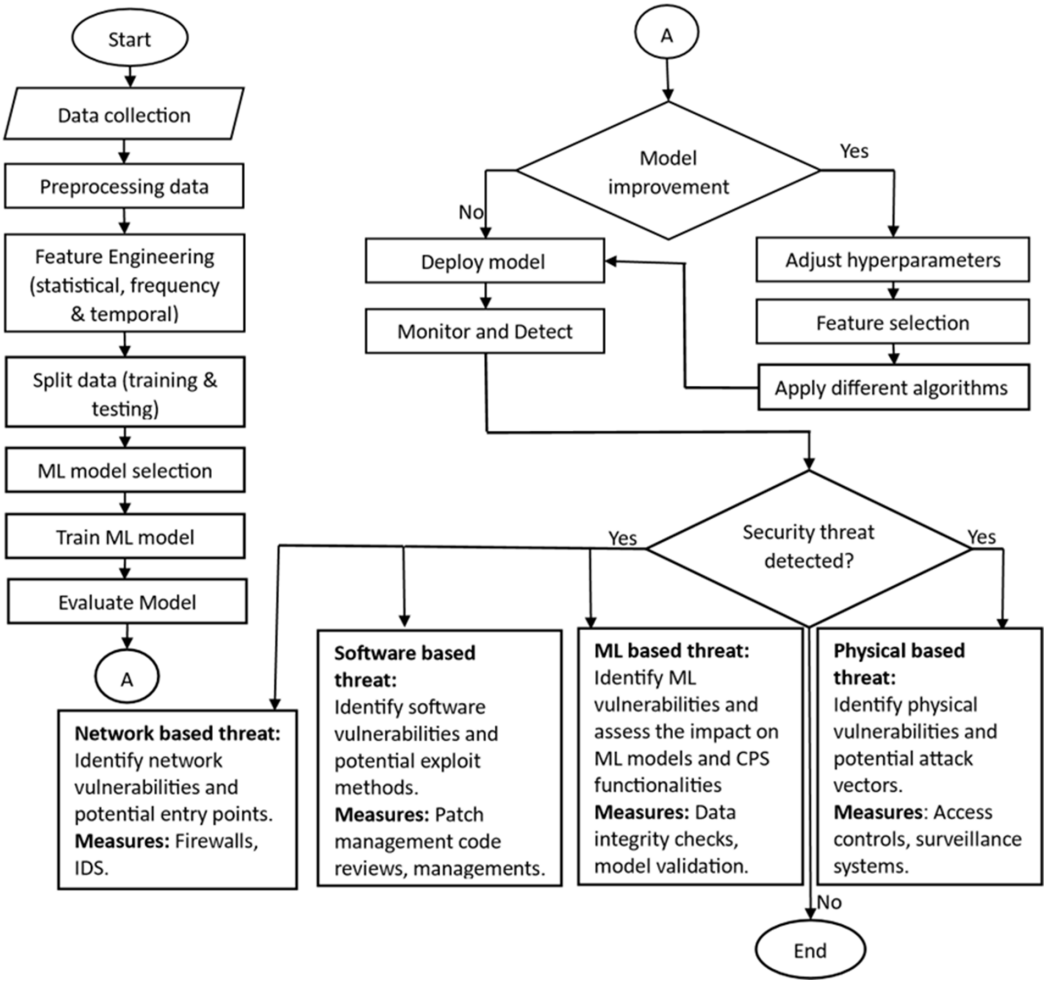
4. How AI Help to Address Security and Privacy Concerns in CPS?

This section answers RQ2. A flow chart is presented and explained how AI/ML is used in detecting security and privacy issues. The purpose is to help present an integrated approach to design an ML-based secure CPS. Figure 3 presents the steps for using ML algorithms in collecting, monitoring, and detecting security threats in CPS. The first step is to collect data from diverse sources within CPS, including sensors, controllers, and network logs. These data are then subjected to preprocessing, where noise is removed, missing values are handled, and it is transformed into a suitable format for ML algorithms. Relevant features are extracted from the preprocessed data, encompassing network traffic patterns, sensor readings, system states, and other pertinent information. The preprocessed data are split into training and testing sets, with the former used to train the selected ML model(s) and the latter to evaluate its performance. ML model selection entails choosing appropriate algorithms, such as anomaly detection, classification algorithms (e.g., DTs, SVMs, neural networks), ensemble methods, or sequence modeling algorithms. The selected model is trained using the training set, learning patterns, and characteristics of normal system behavior. Model performance is evaluated using the testing set, considering metrics like accuracy, precision, recall, and F1 score. If the model's performance is unsatisfactory, iterative improvement is pursued by adjusting hyperparameters, feature selection, or trying different algorithms. Upon achieving satisfactory performance set by the threshold values, the model is deployed in the CPS environment for real-time system monitoring. Continual monitoring and data feeding into the deployed ML model enable the detection of deviations and anomalies that signal possible security threats. If a security threat is detected, three approaches can help identify its nature: network-based threat identification involves analyzing network logs, utilizing NIDS, and conducting packet inspection; software-based threat identification includes reviewing system logs, performing malware analysis, and conducting vulnerability assessments; ML-based threat identification involves analyzing the ML model output, implementing adversarial attack detection techniques, and monitoring model performance. The subsequent paragraphs explain how AI/ML algorithms are used to detect security threats in CPS.

Network-based threats identification involves analyzing network logs and traffic patterns to detect suspicious or malicious activities. By examining network logs, one can look for unusual network behavior, unauthorized access attempts, or unusual data transfers. NIDS play a crucial role in monitoring network traffic and identifying known network-based threats. These systems can detect patterns or signatures of common network attacks, including DDoS attacks, port scanning, or suspicious network connections. Additionally, performing deep packet inspection allows for a thorough examination of network packets. By analyzing packet headers, payloads, and protocols, it becomes possible to identify indicators of network-based threats. This comprehensive analysis helps in identifying malicious activities or anomalies, contributing to an effective network security strategy.

Software-based threat identification involves various techniques to detect and address potential threats originating from software components within a CPS system. Analyzing system logs and event data is crucial in this process, as it allows for the review of activities, error messages, and unauthorized access

Figure 3
Flow chart on how ML is used to detect security threats in CPS



attempts that may indicate a software-based threat. Additionally, conducting malware analysis plays a significant role in identifying potential threats. Suspicious files or programs can be analyzed using antivirus software, sandboxing techniques, or other malware analysis tools to identify any malicious code or behavior. Regular vulnerability assessments and scans are essential to identify known software vulnerabilities that attackers could exploit. By proactively identifying vulnerabilities, it becomes possible to address potential entry points for software-based threats.

ML-based threat identification focuses on detecting and addressing threats that specifically target ML models deployed within a CPS system. Analyzing the output and predictions of the ML model is essential in this process. By examining the model's classifications, false positives or negatives, and instances where the model may be manipulated or attacked, it becomes possible to identify ML-based threats. Implementing techniques to detect and mitigate adversarial attacks is crucial. This can involve monitoring for model evasion attempts, analyzing input data for adversarial perturbations, or employing anomaly detection techniques specifically designed for ML-based threats. Real-time monitoring of the model's performance is also vital. Tracking metrics such as accuracy, precision, recall, and F1 score helps identify sudden drops in performance that could indicate an ML-based attack or model degradation. By actively monitoring and analyzing the ML model's behavior, it becomes possible to identify and mitigate ML-based threats in the CPS system.

AI brings significant benefits to address security and privacy concerns in CPS. It offers capabilities for threat detection and prevention, intrusion detection and response, anomaly detection, vulnerability assessment, predictive maintenance, privacy preservation, behavior analytics, access control, and security analytics. By leveraging these AI-empowered solutions, CPS can strengthen their security posture, detect and respond to threats in real-time, preserve privacy, and ensure the robustness and resilience of their systems. Moreover, AI techniques are invaluable in preserving privacy within CPS environments. Differential privacy is a widely used AI technique (Hassan et al., 2020) and adds noise to data, safeguarding the privacy of individuals or sensitive information while still providing valuable insights. Behavior analytics, powered by AI, enable the detection of suspicious activities or deviations from normal patterns, enabling the identification of potential security breaches or privacy violations. Access control and authentication mechanisms are strengthened through AI, leveraging techniques such as facial recognition, voice recognition, and behavioral biometrics for secure identity verification. AI also plays a vital role in security analytics and incident response. AI-powered security analytics platforms can aggregate and analyze data from various security sources, providing

actionable insights to security teams. This accelerates incident response, allowing for informed decision-making during security incidents. By leveraging AI, CPS environments can enhance their security and privacy safeguards effectively, helping to mitigate risks and protect critical systems and sensitive data.

5. AI Techniques Used to Address the Security and Privacy Issues

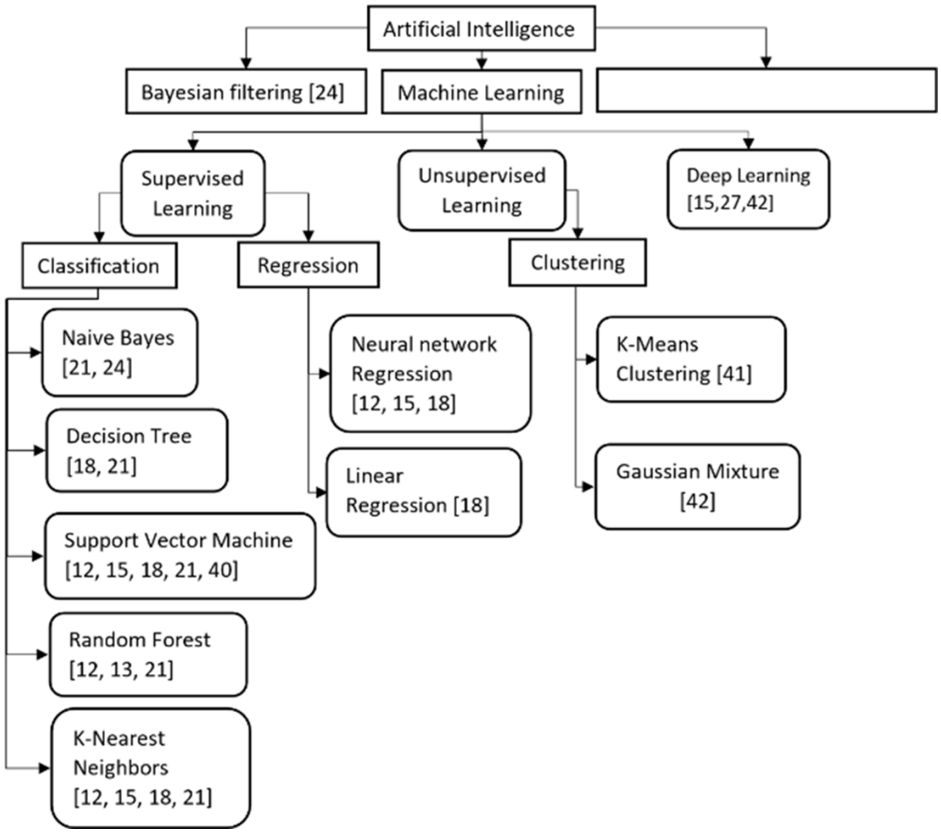
This section answers the third RQ. How can AI help in addressing the security and privacy concerns in CPS? A taxonomy of AI methods is presented to determine which techniques are predominantly utilized and to identify the gaps in identifying security concerns. The AI techniques used to protect data in CPS are shown in Figure 4. The shortcomings of AI are discussed in the next section.

AI is a computer science discipline that enables machines to think and behave like humans using methods such as ML, DL, game theory, optimization theory, and evolutionary algorithms. Bayesian filtering (Feng & Tian, 2021) and robotic automation (Beltrame et al., 2018) are being used to protect sensitive data CPS. Bayesian filtering, which is a statistical technique, is used to detect anomalies and cyberattacks by comparing incoming sensor data to a model of what the data should look like. This technique can identify and filter out corrupted data, thus helping to improve the accuracy of the system. Robotic automation, on the other hand, can be used to create a secure and isolated environment for the data. For example, robots can be used to physically isolate the system, such as by removing external ports, to reduce the risk of unauthorized access. Additionally, robots can be used to monitor and regulate access to sensitive data, ensuring that only authorized personnel have access. The combination of Bayesian filtering with robotic automation can provide a robust solution for protecting data in CPS against cyber threats.

In addressing security and privacy issues in CPS, several AI techniques are employed, with supervised ML algorithms being the most prevalent, particularly classification algorithm (Colelli et al., 2021; Feng & Tian, 2021; Jamal et al., 2023; Perrone et al., 2021; Pordelkhaki et al., 2021; Rahman et al., 2017; Ulybyshev et al., 2021). The classification ML algorithms can be used to train models that can classify data as normal or anomalous. For example, anomaly detection algorithms can be used to identify abnormal network traffic, which could be indicative of a cyberattack. Similarly, classification algorithms can be used to detect malicious software or malware that could compromise the security of the CPS. These algorithms can also be used to protect privacy in CPS by identifying and classifying sensitive data that should not be shared with unauthorized parties. For instance, classification algorithms can be trained to recognize personal information, such as social security numbers or credit card numbers, and prevent them from being transmitted outside a secure network. Therefore, classification ML algorithms can be a powerful tool in detecting and preventing security and privacy issues in CPS. By analyzing data and identifying patterns, these algorithms can help ensure the integrity and safety of critical infrastructure systems. Regression algorithms like neural network regression (Ulybyshev et al., 2021) and linear regression (Pordelkhaki et al., 2021) are also used to detect cyberattacks.

Clustering algorithms are commonly used in the security and privacy of CPS for identifying patterns and grouping similar data

Figure 4
Taxonomy of AI techniques in CPS



points together. Some of the commonly used clustering algorithms in this domain include K-means clustering (Sahin et al., 2022) and Gaussian mixture model (GMM) (Padmajothi & Iqbal, 2022). This algorithm partitions the data points into K distinct clusters based on their similarity. It is commonly used in IDS for identifying anomalous network traffic. GMM is another clustering algorithm used in unsupervised learning. It models the distribution of data as a mixture of several Gaussian distributions and attempts to identify the parameters of each distribution to cluster the data. This helps in identifying the anomalies in the CPS. Reinforcement learning algorithms (Ibrahim & Elhafiz, 2023; Uprety & Rawat, 2021) can be used in CPS security and privacy to develop autonomous decision-making systems that can respond to changing environments and emerging threats. Reinforcement learning algorithms learn from feedback and reinforcement signals generated by the environment to adapt and improve their decision-making over time. For example, in a scenario where a CPS is under attack, reinforcement learning algorithms can be used to automatically adjust security measures to mitigate the effects of the attack. Reinforcement learning can also be used to develop adaptive intrusion detection and response systems that can learn from past attacks and update their responses accordingly. By leveraging the flexibility and adaptability of reinforcement learning, it is possible to develop more efficient and effective security and privacy solutions for CPS.

DL is a subfield of ML, which falls under the category of supervised learning. However, DL models use artificial neural networks that are composed of multiple layers to learn from data, which distinguishes it from traditional ML algorithms (Kaplan et al., 2021). DL can be utilized to detect and prevent cyberattacks (Zhang et al., 2022). By training DL models on large datasets of historical attacks and their corresponding features, such as network traffic patterns and system logs, these models can learn to recognize patterns and anomalies that may indicate an ongoing or potential attack. The use of artificial neural networks with multiple layers allows for complex relationships and dependencies to be captured and learned from the data, potentially leading to more accurate and robust detection capabilities. Additionally, DL models can also be used for anomaly detection in sensor data, helping to identify abnormal behavior that may indicate a physical attack or malfunction in the system. However, it is important to note that DL models can also be vulnerable to adversarial attacks.

AML algorithms are a subfield of ML that aims to detect and defend against attacks on ML models. To protect the data in CPS, AML algorithms are used to identify and mitigate threats to the system. One common type of attack is called an adversarial attack, where an attacker intentionally modifies the input data to mislead the ML model. AML algorithms work by introducing adversarial examples into the training data to improve the model's robustness against attacks. Another approach is to use AML algorithms to identify and classify potential attacks, allowing the system to take appropriate action to defend against them. Therefore, AML algorithms are an important tool for enhancing the security and privacy of CPS and ensuring their resilience against evolving threats.

6. Discussion

6.1. Research significance and limitations

The classification of security threats is significant as it provides an organized overview of potential risks in CPS. It raises awareness, enables risk assessment, and helps in secure CPS design and development. The classification aids in incident response and recovery by guiding targeted actions based on threat categories. It facilitates effective communication and collaboration among stakeholders, fostering a common understanding and knowledge sharing. Eventually, the classification enhances the security and resilience of CPS by guiding proactive measures and promoting a secure environment.

The flow chart explains the step-by-step process of identifying the different security threats using ML algorithms. By exploring the role of AI in CPS security, this research aims to enhance the protection of CPS against potential threats, mitigate privacy risks, enable proactive defense mechanisms, improve incident response and recovery, and promote trust in CPS deployments. The application of AI techniques can contribute to developing advanced security strategies, privacy-preserving mechanisms, and real-time threat detection, ultimately ensuring CPS applications' reliability, resilience, and trustworthiness.

The recommendation is to prioritize research on unsupervised, reinforcement, and DL techniques for CPS applications, as there is currently limited evidence in the literature. Additionally, as ML-based attacks become more prevalent, research is needed to focus on developing robust and secure AI systems. Researchers are advised to define attacks in specific terms instead of general terms. This means that they should provide a detailed description of the attack instead of using broad, vague terms to describe attacks. This research does not include the research related to the blockchain, as the focus was explicitly on using AI techniques to handle security and privacy concerns in the CPS.

6.2. Challenges and implementation issues

This section discusses the challenges and implementation issues to security and privacy in CPS.

6.2.1. Shortcomings of AI

While AI has the potential to improve performance in CPS, there are a few shortcomings that need to be addressed. AI works as a black box, and the user is not always aware of how it works or why a particular decision was made. One of the biggest challenges is the lack of transparency in the decision-making process. AI algorithms can be complex and difficult to understand, which can make it hard to explain why a particular decision was made. In mission-critical systems, it is essential that decisions must be explainable and accountable. Additionally, AI algorithms rely on data to learn and make decisions. In CPS, due to the heterogeneous nature of data, the quality of data can vary and impact the reliability of AI models. AI algorithms are also vulnerable to cyberattacks, which can compromise the safety of the CPS (Li et al., 2021). CPS devices generate huge volumes of data, making it challenging to scale AI algorithms to handle massive amounts of data. Moreover, AI algorithms require significant computational resources.

6.2.2. Federated learning in Edge AI for CPS systems

Federated learning in Edge AI for CPS refers to the use of distributed ML techniques that allow ML models to be trained using data from edge devices in a decentralized and collaborative manner. CPS are systems that integrate physical and computational components, generating vast amounts of data that can be used to improve system performance and reliability (Olowononi et al., 2021b). However, collecting and processing this data can be challenging, especially in large-scale systems that are distributed across multiple locations. Federated learning provides a solution to

this problem by allowing ML models to be trained using data that remain on the local devices where it was generated. In Edge AI for CPS, federated learning can be used to train ML models on data generated by sensors and devices located at the edge of the network. By keeping data local, federated learning can reduce the amount of data that need to be transmitted over the network, which can be important in systems with limited bandwidth or high communication costs. By distributing the learning process across multiple edge devices, federated learning can improve the scalability of ML models, allowing them to adapt to changing conditions and improve system performance in real-time. Therefore, federated learning in Edge AI for CPS provides a flexible and scalable approach to ML that can help improve the performance and reliability of CPS while reducing communication costs and preserving user privacy. However, federated learning in a distributed environment increases the complexity and maintenance of CPS. In heterogeneous distributed CPS, variability in storage capacity, computational power, and energy consumption poses challenges for developing federated models that can effectively execute across multiple devices (Salau et al., 2022).

6.2.3. Beyond 5G technologies

Implementing CPS with beyond 5G technology is an attractive option for many application domains. However, there are several challenges associated with it. The beyond 5G network architectures include the use of more distributed networks, making the system complex and requiring significant investment in network infrastructure (Bandi, 2022; Bandi & Yalamarthi, 2022). Additionally, data management is another challenge as future CPS require efficient systems to collect, process, store, analyze, and visualize data. Developing such systems to handle complex and larger datasets is expensive. Since beyond 5G is in its early stages of development, there is no clear standardization framework for the technology, and ensuring that different systems are compatible and interoperable is essential. Implementing CPS with beyond 5G technology will require a significant investment in research, development, and infrastructure.

6.2.4. Regulatory and legal compliance

CPS applications must comply with a range of regulatory and legal requirements, including safety standards, privacy laws, and data protection regulations. Compliance with these requirements can be complex and time-consuming and can add significant costs to the development and deployment of CPS (Moongilan, 2019). For example, in the healthcare industry, CPS must comply with the Health Insurance Portability and Accountability Act, which sets strict requirements for the protection of patient data. Another example is in the automotive industry, where CPS must comply with safety standards such as the ISO 26262, which provides a framework for the development of safety-critical systems in vehicles. This standard requires a systematic approach to safety engineering, including hazard analysis and risk assessment, as well as extensive testing and verification. In addition to industry-specific regulations, CPS must also comply with more general legal requirements such as data protection regulations and privacy laws. For example, the General Data Protection Regulation in Europe sets strict rules for the collection, use, and storage of personal data, including data generated by CPS. Compliance with these regulations and standards can be challenging, as it requires a deep understanding of the legal and regulatory landscape, as well as significant investment in compliance processes and technologies (Moongilan, 2019).

7. Conclusion and Future Research

In conclusion, the classification of security threats in CPS plays a crucial role in enhancing these systems' overall security and resilience. It provides an organized overview of potential risks, raising awareness and enabling risk assessment. This classification framework aids in designing and developing secure CPS by guiding targeted actions based on threat categories. It facilitates communication and collaboration among stakeholders, fostering a common understanding and promoting knowledge sharing. The research presented here explores the role of AI in CPS security, aiming to enhance the protection of CPS against potential threats and mitigate privacy risks. By utilizing ML algorithms and AI techniques, developing advanced security strategies, privacy-preserving mechanisms, and real-time threat detection is possible. This research contributes to CPS applications' reliability, resilience, and trustworthiness in deploying CPS.

To further advance the field, future research should prioritize investigating unsupervised, reinforcement, and DL techniques for CPS applications, as limited evidence exists in the literature. With the rise of ML-based attacks, developing robust and secure AI systems to safeguard CPS is crucial. Researchers are advised to define attacks in specific terms, providing detailed descriptions rather than broad and vague terms. This approach will lead to a better understanding of attacks and enable the development of effective defense mechanisms. It is important to note that this research does not encompass the study of blockchain concerning CPS. This work focused on utilizing AI techniques to address security and privacy concerns in CPS. Further exploration of blockchain technology and its potential contributions to CPS security would be an opportunity for future investigation. Overall, this study sheds light on CPS's current security and privacy issues and provides insights into potential solutions and areas for further research.

Acknowledgments

The author graduates students Aishwarya Mallela and Sai Kiran Mandapalli for searching articles in the electronic databases.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The author declares that he has no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

Abbas, W., Laszka, A., Vorobeychik, Y., & Koutsoukos, X. (2015). Scheduling intrusion detection systems in resource-bounded cyber-physical systems. In *Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy*, 55–66. https://doi.org/10.1145/2808705.2808711

Asghar, M. R., Hu, Q., & Zeadally, S. (2019). Cybersecurity in industrial control systems: Issues, technologies, and

challenges. *Computer Networks*, *165*, 106946. https://doi.org/10.1016/j.comnet.2019.106946

Ashok, A., Wang, P., Brown, M., & Govindarasu, M. (2015). Experimental evaluation of cyber attacks on automatic generation control using a CPS security testbed. In *IEEE Power & Energy Society General Meeting*, 1–5. https://doi.org/10.1109/PESGM.2015.7286615

Balduzzi, M., Pasta, A., & Wilhoit, K. (2014). A security evaluation of AIS automated identification system. In *Proceedings of the 30th Annual Computer Security Applications Conference*, 436–445. https://doi.org/10.1145/2664243.2664257

Bandi, A. (2022). A review towards AI empowered 6G communication requirements, applications, and technologies in mobile edge computing. In *6th International Conference on Computing Methodologies and Communication*, 12–17. https://doi.org/10.1109/ICCMC53470.2022.9754049

Bandi, A., & Yalamarthi, S. (2022). Towards artificial intelligence empowered security and privacy issues in 6G communications. In *International Conference on Sustainable Computing and Data Communication Systems*, 372–378. https://doi.org/10.1109/ICSCDS53736.2022.9760857

Beltrame, G., Merlo, E., Panerati, J., & Pinciroli, C. (2018). Engineering safety in swarm robotics. In *Proceedings of the 1st International Workshop on Robotics Software Engineering*, 36–39. https://doi.org/10.1145/3196558.3196565

Bharathi, V., & Kumar, C. V. (2022). A real time health care cyber attack detection using ensemble classifier. *Computers and Electrical Engineering*, *101*, 108043. https://doi.org/10.1016/j.compeleceng.2022.108043

Bour, G., Bosco, C., Ugarelli, R., & Jaatun, M. G. (2023). Water-tight IoT–Just add security. *Journal of Cybersecurity and Privacy*, *3*(1), 76–94. https://doi.org/10.3390/jcp3010006

Cali, U., Kuzlu, M., Sharma, V., Pipattanasomporn, M., & Catak, F. O. (2021). Internet of predictable things (IoPT) framework to increase cyber-physical system resiliency. *arXiv Preprint:2101.07816*.

Cheh, C., Keefe, K., Feddersen, B., Chen, B., Temple, W. G., & Sanders, W. H. (2017). Developing models for physical attacks in cyber-physical systems. In *Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and Privacy*, 49–55. https://doi.org/10.1145/3140241.3140249

Colelli, R., Magri, F., Panzieri, S., & Pascucci, F. (2021). Anomaly-based intrusion detection system for cyber-physical system security. In *29th Mediterranean Conference on Control and Automation*, 428–434. https://doi.org/10.1109/MED51440.2021.9480182

Dhir, S., & Kumar, Y. (2020). Study of machine and deep learning classifications in cyber physical system. In *Third International Conference on Smart Systems and Inventive Technology*, 333–338. https://doi.org/10.1109/ICSSIT48917.2020.9214237

Feng, C., & Tian, P. (2021). Time series anomaly detection for cyber-physical systems via neural system identification and bayesian filtering. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2858–2867. https://doi.org/10.1145/3447548.3467137

Haque, N. I., Shahriar, M. H., Dastgir, M. G., Debnath, A., Parvez, I., Sarwat, A., & Rahman, M. A. (2021). A survey of machine learning-based cyber-physical attack generation, detection, and mitigation in smart-grid. In *2020 52nd North American Power Symposium*, 1–6. https://doi.org/10.1109/NAPS50074.2021.9449635

Hasan, Z., & Roy, N. (2021). Trending machine learning models in cyber-physical building environment: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(5), e1422. https://doi.org/10.1002/widm.1422

Hassan, M. U., Rehmani, M. H., & Chen, J. (2020). Differential privacy techniques for cyber physical systems: A survey. *IEEE Communications Surveys & Tutorials*, *22*(1), 746–789. https://doi.org/10.1109/COMST.2019.2944748

Ibrahim, M., & Elhafiz, R. (2023). Security analysis of cyber-physical systems using reinforcement learning. *Sensors*, *23*(3), 1634. https://doi.org/10.3390/s23031634

Jamal, A. A., Majid, A. A. M., Konev, A., Kosachenko, T., & Shelupanov, A. (2023). A review on security analysis of cyber physical systems using machine learning. *Materials Today: Proceedings*, *80*, 2302–2306. https://doi.org/10.1016/j.matpr.2021.06.320

Joo, M., Seo, J., Oh, J., Park, M., & Lee, K. (2018). Situational awareness framework for cyber crime prevention model in cyber physical system. In *Tenth International Conference on Ubiquitous and Future Networks*, 837–842. https://doi.org/10.1109/ICUFN.2018.8436591

Junejo, K. N., & Goh, J. (2016). Behaviour-based attack detection and classification in cyber physical systems using machine learning. In *Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security*, 34–43. https://doi.org/10.1145/2899015.2899016

Kaplan, H., Tehrani, K., & Jamshidi, M. (2021). Fault diagnosis of smart grids based on deep learning approach. In *World Automation Congress*, 164–169. https://doi.org/10.23919/WAC50355.2021.9559474

Karale, A. (2021). The challenges of IoT addressing security, ethics, privacy, and laws. *Internet of Things*, *15*, 100420. https://doi.org/10.1016/j.iot.2021.100420

Khan, M. T., Serpanos, D., Shrobe, H., & Yousuf, M. M. (2020). Rigorous machine learning for secure and autonomous cyber physical systems. In *25th IEEE International Conference on Emerging Technologies and Factory Automation*, *1*, 1815–1819. https://doi.org/10.1109/ETFA46521.2020.9212074

Li, J., Yang, Y., Sun, J. S., Tomsovic, K., & Qi, H. (2021). Conaml: Constrained adversarial machine learning for cyber-physical systems. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, 52–66. https://doi.org/10.1145/3433210.3437513

Liu, R. W., Liang, M., Nie, J., Deng, X., Xiong, Z., Kang, J., . . . , & Zhang, Y. (2021). Intelligent data-driven vessel trajectory prediction in marine transportation cyber-physical system. In *IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, 314–321. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics53846.2021.00058

Mboweni, I. V., Abu-Mahfouz, A. M., & Ramotsoela, D. T. (2021). A machine learning approach to intrusion detection in water distribution systems–A review. In *IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society*, 1–7. https://doi.org/10.1109/IECON48115.2021.9589237

Mo, Y., & Sinopoli, B. (2012). Integrity attacks on cyber-physical systems. In *Proceedings of the 1st International Conference*

on High Confidence Networked Systems, 47–54. https://doi.org/10.1145/2185505.2185514

Mohammadhassani, A., Teymouri, A., Mehrizi-Sani, A., & Tehrani, K. (2020). Performance evaluation of an inverter-based microgrid under cyberattacks. In *IEEE 15th International Conference of System of Systems Engineering*, 211–216. https://doi.org/10.1109/SoSE50414.2020.9130524

Moongilan, D. (2019). 5G Internet of Things (IOT) near and far-fields and regulatory compliance intricacies. In *IEEE 5th World Forum on Internet of Things,* 894–898. https://doi.org/10.1109/WF-IoT.2019.8767334

Oladimeji, S., & Kerner, S. M. (2023). *SolarWinds hack explained: Everything you need to know*. Retrieved from: https://www.techtarget.com/whatis/feature/SolarWinds-hack-explained-Everything-you-need-to-know

Olowononi, F. O., Rawat, D. B., & Liu, C. (2021a). Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for CPS. *IEEE Communications Surveys & Tutorials*, *23*(1), 524–552. https://doi.org/10.1109/COMST.2020.3036778

Olowononi, F. O., Rawat, D. B., & Liu, C. (2021b). Federated learning with differential privacy for resilient vehicular cyber physical systems. In *IEEE 18th Annual Consumer Communications & Networking Conference,* 1–5. https://doi.org/10.1109/CCNC49032.2021.9369480

Padmajothi, V., & Iqbal, J. L. (2022). Review of machine learning and deep learning mechanism in cyber-physical system. *International Journal of Nonlinear Analysis and Applications*, *13*(1), 583–590. https://doi.org/10.22075/ijnaa.2022.5540

Perrone, P., Flammini, F., & Setola, R. (2021). Machine learning for threat recognition in critical cyber-physical systems. In *IEEE International Conference on Cyber Security and Resilience*, 298–303. https://doi.org/10.1109/CSR51186.2021.9527979

Pordelkhaki, M., Fouad, S., & Josephs, M. (2021). Intrusion detection for industrial control systems by machine learning using privileged information. In *IEEE International Conference on Intelligence and Security Informatics*, 1–6. https://doi.org/10.1109/ISI53945.2021.9624757

Rahman, M., Chowdhury, M., Rayamajhi, A., Dey, K., & Martin, J. (2017). Adaptive queue prediction algorithm for an edge centric cyber physical system platform in a connected vehicle environment. *arXiv Preprint:1712.05837*.

Sahin, M. E., Tawalbeh, L., & Muheidat, F. (2022). The security concerns on cyber-physical systems and potential risks analysis using machine learning. *Procedia Computer Science*, *201*, 527–534. https://doi.org/10.1016/j.procs.2022.03.068

Salau, B. A., Rawal, A., & Rawat, D. B. (2022). Recent advances in artificial intelligence for wireless internet of things and cyber–physical systems: A comprehensive survey. *IEEE Internet of Things Journal*, *9*(15), 12916–12930. https://doi.org/10.1109/JIOT.2022.3170449

Sengan, S., Subramaniyaswamy, V., Indragandhi, V., Velayutham, P., & Ravi, L. (2021). Detection of false data cyber-attacks for the assessment of security in smart grid using deep learning. *Computers & Electrical Engineering*, *93*, 107211. https://doi.org/10.1016/j.compeleceng.2021.107211

Singh, S., Yadav, N., & Chuarasia, P. K. (2020). A review on cyber physical system attacks: Issues and challenges. In *International Conference on Communication and Signal Processing,* 1133–1138. https://doi.org/10.1109/ICCSP48568.2020.9182452

Tahsien, S. M., Karimipour, H., & Spachos, P. (2020). Machine learning based solutions for security of Internet of Things (IoT): A survey. *Journal of Network and Computer Applications*, *161*, 102630. https://doi.org/10.1016/j.jnca.2020.102630

Ulybyshev, D., Yilmaz, I., Northern, B., Kholodilo, V., & Rogers, M. (2021). Trustworthy data analysis and sensor data protection in cyber-physical systems. In *Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems*, 13–22. https://doi.org/10.1145/3445969.3450432

Uprety, A., & Rawat, D. B. (2021). Reinforcement learning for IoT security: A comprehensive survey. *IEEE Internet of Things Journal*, *8*(11), 8693–8706. https://doi.org/10.1109/JIOT.2020.3040957

Vlachos, V., Stamatiou, Y. C., & Nikoletseas, S. (2023). The privacy flag observatory: A crowdsourcing tool for real time privacy threats evaluation. *Journal of Cybersecurity and Privacy*, *3*(1), 26–43. https://doi.org/10.3390/jcp3010003

Yampolskiy, M., Horvath, P., Koutsoukos, X. D., Xue, Y., & Sztipanovits, J. (2012). Systematic analysis of cyber-attacks on CPS-evaluating applicability of DFD-based approach. In *5th International Symposium on Resilient Control Systems,* 55–62. https://doi.org/10.1109/ISRCS.2012.6309293

Zhang, J., Pan, L., Han, Q. L., Chen, C., Wen, S., & Xiang, Y. (2022). Deep learning based attack detection for cyber-physical system cybersecurity: A survey. *IEEE/CAA Journal of Automatica Sinica*, *9*(3), 377–391. https://doi.org/10.1109/JAS.2021.1004261

Zion Market Research. (2022). *Cyber-Physical Systems (CPS) market size & industry analysis.* Retrieved from: https://www.zionmarketresearch.com/report/cyber-physical-systems-market

How to Cite: Bandi, A. (2024). A Taxonomy of AI Techniques for Security and Privacy in Cyber–Physical Systems. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE42021539