Computers & Security 135 (2023) 103531

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Industrial cyber-physical systems protection: A methodological review

Roberto Canonico, Giancarlo Sperlì *

Department of Electrical Engineering and Information Technology (DIETI), University of Naples "Federico II", Via Claudio 21, Naples, Italy

ARTICLE INFO

Keywords:
Cyber-physical systems
Industrial control systems
Security
Artificial intelligence
Cyber-physical attacks
Countermeasures

ABSTRACT

Ubiquitous utilization of Information and Communication Technologies in modern manufacturing plants has transformed them into Cyber-Physical Systems (CPSs), making them susceptible to cyber-attacks, which can have huge economic and social impact.

In this paper, we focus on the security issues that may affect the Industrial Control System (ICS) supervising an industrial establishment. The purpose of this work is to provide readers with an up-to-date view of the methodologies that current literature suggests as the most appropriate to defend an ICS against cyber-attacks. We firstly provide a classification of existing attacks according to the methodology used by the attacker. Subsequently, we propose a classification of defensive countermeasures in *Model* and *Artificial Intelligence*-based approaches by analyzing most recent research contributions. Furthermore, we describe the most used datasets in literature that are available to researchers and practitioners. We conclude the paper by discussing today's open issues, which need to be addressed to cope with the increasing complexity of ICS security.

1. Introduction

The adoption of heavily automated production processes in industrial plants dates back to the 1960s. In the last few years, the incredible advances of ICT have pushed forward a deep transformation of modern industrial factories, with the emergence of groundbreaking paradigms such as *Smart Manufacturing* and *Industry 4.0* (Thoben et al. (2017)). Ubiquitous utilization of ICT technologies in modern manufacturing plants has transformed them into *cyber-physical systems* (CPSs), i.e. "hardware–software systems that tightly couple the physical world and the digitalized (virtual) world" (Colombo et al. (2017); Sinha and Roy (2020)).

The term Smart Manufacturing refers to manufacturing processes in which real-time transmission and analysis of data extracted from the entire product life-cycle, combined with model-based simulation, create intelligence that is used to improve and optimize all aspects of manufacturing. A smart factory, hence, is a fully integrated, collaborative system which is able to adapt to changes of the factory environment, supply network, customer demands, and market scenario. The interconnection of organizational systems advocated by these new paradigms, however, significantly increases the exposure of industrial plants to unprecedented security risks. Malicious hackers may exploit software vulnerabilities in the system components to disrupt the whole production chain, potentially for long periods of time if their attacks are capable of damaging physical assets of the industrial facility. The pervasive diffusion of Cyber-Physical Systems in every domain made industrial plants and facilities susceptible to a plethora of cyber-attacks, which can have both economic and social impact (Corallo et al. (2021)) (Alwan et al. (2022)). To make the situation even more scary, these attacks can be quite easily activated, as it was shown by the cyber-attack perpetrated in May 2021 to the largest fuel pipeline in the U.S. (which eventually led to fuel shortages across the East Coast for some days) (Turton and Mehrotra (2021)).

In this paper, we focus on the security issues that may affect the *Industrial Control System* supervising an industrial establishment. In the past few years, such issues have been addressed in a huge number of research papers. Our research objective is to provide readers with an up-to-date view of the methodologies that current literature suggests as the most appropriate to defend Industrial Control Systems against the different types of attacks that have been reported. More specifically, we proceed by first classifying attacks against industrial CPSs according to the methodology used by the attacker, distinguishing between *data-driven* and *network-driven* attacks, where the former inject deceptive data into the system while the latter create, manipulate or suppress network packets in the communication infrastructure.

We then propose a classification of defensive countermeasures, based on an analysis of the most recent research contributions. At the first level, we classify them in *model-based* and *AI-based* approaches. The

* Corresponding author.
E-mail addresses: roberto.canonico@unina.it (R. Canonico), giancarlo.sperli@unina.it (G. Sperlì).

https://doi.org/10.1016/j.cose.2023.103531
Received 29 November 2022; Received in revised form 12 September 2023; Accepted 4 October 2023
Available online 12 October 2023

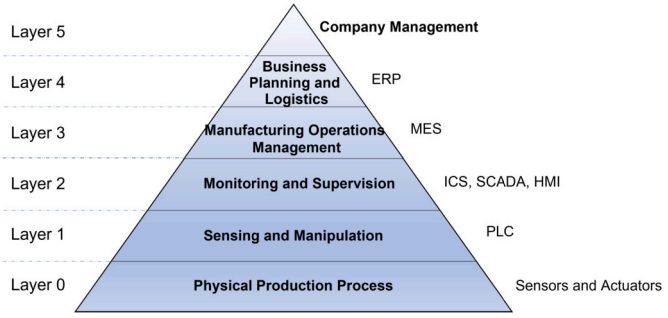

Fig. 1. The ISA-95 automation pyramid.

former type comprised methodologies whose aim is to detect anomalous behaviors through accurate system models. The latter, in turn, relies on AI-based models that can detect attacks in dynamic and complex scenarios that are hard to capture in a single complete model.

Performing a comparative evaluation of proposals has become harder and harder due to the difficulty of reproducing the complex scenarios involved in industrial CPS. Hence, in this paper we also present a discussion about the datasets most used by researchers and practitioners to support their analyses.

Finally, we conclude the paper by analyzing and discussing several open issues, which need to be addressed by the community of researchers to cope with the increasing complexity of systems and technologies and increase the effectiveness of future research efforts.

Hence, we have structured this paper as follows. In Section 2 we provide a high-level description of the role of an ICS in modern industry. In Section 3, having selected the most recent papers that survey the research literature addressing ICS security issues, we present a categorization of these surveys according to different possible classification criteria. In Section 4 we review the known security attacks against an ICS and classify them according to the vulnerabilities they exploit. In Section 5 we classify the methodologies that have been adopted in the literature to devise effective security countermeasures. In Section 6 we report on the most popular datasets that the research community can count on. The availability of properly labeled datasets, in fact, is crucial to support comparative evaluation of experimental studies (Apruzzese et al. (2022)). In Section 7 we analyze the research issues that still appear not to be completely solved and hence call for further research efforts by the community. Finally, in Section 8 we provide a concluding balance of the ongoing research efforts in this field by highlighting the most promising trends.

2. Industrial control systems

In the mid-1990s, the necessity to manage the growing complexity of industrial automation solutions drove the *International Society of Automation* (ISA) to establish ISA-95 (ISA (2008)), a series of standards that defines enterprise and control system integration for manufacturing. These standards have been developed for all kinds of manufacturing environments and all sorts of processes, such as batch, continuous, and repetitive or discrete processes.

The ISA-95 specifications define a hierarchical functional model of an industrial system, usually represented as a pyramid, as shown in Fig. 1. In a top-down view, this model consists of six different levels: *Company Management*, *Business Planning and Logistics*, *Manufacturing Operations Management*, *Monitoring and Supervision*, *Sensing and Manipulation*, and *Physical Production Process*. Such levels describe functions operating at different timescales and with different objectives.

The Monitoring and Supervision layer of an industrial plant is centered around an *Industrial Control System* (ICS), which performs its functions by interacting with sensors, actuators, *Programmable Logic Controllers* (PLCs) and *Human Machine Interfaces* (HMIs). In recent years, ICSs have passed through a significant transformation from proprietary, isolated systems to open systems based on standard technologies and highly interconnected with corporate networks and the public Internet.

An ICS typically consists of a SCADA (*Supervisory Control and Data Acquisition*) system that collects and processes data generated by *Remote Terminal Units* (RTUs) and PLCs, and allows field engineers to monitor the status of the ICS and modify its configuration parameters through an HMI. The control functions of a SCADA system are centralized in a supervisory controller called *Master Terminal Unit* (MTU) which communicates with PLCs and RTUs through a specialized network infrastructure. Industrial networks are based on communication protocols such as Modbus, IEC 61850, IEC 60870, DNP3, and Profinet.

Initially, SCADA systems were used in power transmission, gas pipeline and water distribution control systems. Nowadays, they can be found in manufacturing factories and in all sorts of industrial settings.

2.1. Cyber-security requirements for an ICS

Design and operational management of ICS/SCADA systems have always taken into account non-functional requirements such as RAMS (Reliability, Availability, Maintainability, and Safety). On the contrary, security requirements have been neglected for many years.

In the past, security of a SCADA system was primarily achieved by controlling physical access to its components. Since communications were based on isolated network infrastructures relying on proprietary communication protocols, this led ICS administrators to underestimate security risks originating from network connectivity. To make things worse, until recently it was quite common that SCADA software components ran in outdated Operating Systems whose vulnerabilities are well-known and may be easily exploited by potential attackers. According to the Global ICS and IIoT Risk Report released in October 2017 by CyberX,[1] at the time about 50% of the industrial control networks lacked anti-virus protection and used easily hackable plain-text passwords in their core functions. More than three-quarters of MTUs were running obsolete Windows systems no more supported with security patches. This situation has progressively changed, and today the update of software components of SCADA systems is one of the main security concerns for ICS technical administrators.

Nowadays, it is clear that security needs to be considered as a first-class requirement for the success of the Industry 4.0 paradigm (Mullet et al. (2021)), in which the ICS shifts from a stand-alone plant to a cloud-based environment. Cyber-physical systems have become more and more pervasive within industrial processes with the aim of reducing the human factor burden and achieving new forms of optimization across all layers of the automation pyramid. As a consequence of this evolution, new plants adopting the Industry 4.0 paradigm are more exposed to cyber-attacks (see (Selim et al. (2021))), whose consequences may be detrimental to the operation of the whole enterprise and can lead to financial losses, environmental damages and loss of human lives.

2.2. Cyber-security risk assessment of an ICS

In a complex system such as an industrial plant, that is subject to a wide range of security threats, it is important to evaluate in advance the impact of a security attack in order to plan proper responses to identified risks.

In a broader sense, the National Institute of Standards (NIST) defines risk assessment as *the process of identifying risks to organizational operations (including mission, functions, image, and reputation), organizational assets, individuals, other organizations, and the Nation, resulting from the operation of a system* (National Institute of Standards and Technology (2020)).

[1] https://www.automation.com/en-us/articles/2017/cyberx-releases-global-ics-iiot-risk-report.

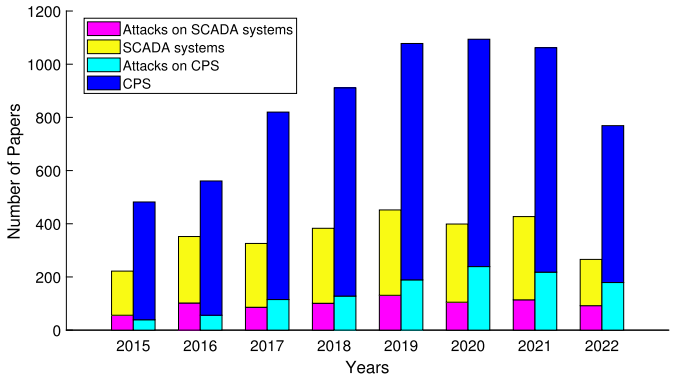

Fig. 2. Number of papers addressing CPS attacks [cyan] (and SCADA attacks [purple]) compared to the total number of papers covering CPS [blue](and SCADA [yellow]) in Computer Science Journals and Conferences. Publication data for 2022 have been gathered in October 2022. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Evaluation of the business impact of security attacks in the context of Industry 4.0 requires new methodologies for risk assessment. In the literature, several risk assessment methods have been proposed. In general, they can be classified as qualitative, quantitative, and semi-quantitative. In (Cherdantseva et al. (2016)), an overview of twenty-four of such methods is presented. In (Corallo et al. (2021)), the authors present the application of the *impact assessment methodology*, following NIST's asset/impact-oriented approach, to the evaluation of the business impact of cybersecurity breaches in an industrial context with networked manufacturing machines. In particular, this latter paper investigates possible differences in terms of cybersecurity risks between factories employing either subtractive or additive manufacturing technologies.

3. A categorization of currently available surveys

Fig. 2 compares the number of papers addressing CPS attacks (and, more specifically, SCADA attacks) against the total number of scientific papers published in Computer Science journals and conferences between 2015 and 2022 on the CPS subject (and SCADA). Figures had been obtained by performing the following queries on the SCOPUS[2] system: i) *TITLE-ABS-KEY ("SCADA" AND "ATTACK") AND SUBJAREA (comp) AND PUBYEAR AFT 2014* ii) *TITLE-ABS-KEY ("SCADA") AND SUBJAREA (comp) AND PUBYEAR AFT 2014* iii) *TITLE-ABS-KEY ("CPS" AND "ATTACK") AND SUBJAREA (comp) AND PUBYEAR AFT 2014* and iv) *TITLE-ABS-KEY ("CPS") AND SUBJAREA (comp) AND PUBYEAR AFT 2014*. It is worth noting that the percentage of published paper focusing on CPS attacks (and, more specifically, SCADA attacks) has increased from 8% in 2015 to 23% 2022 (from 25% to 35% for SCADA). This growth is also justified by the increase in the number of papers addressing the broader field of CPS security, which increased from 13% to 33% over the same period.[3]

Current literature is hence quite rich of papers and surveys addressing CPS security issues from different points of view.

In (Corallo et al. (2021)), the authors assessed the business impact level on critical assets, also investigating attacks that compromise quality information and attempt to steal strategic information. A literature review between 1999 and 2019 has been further discussed by (Engström and Lagerström (2022)) about cyber-attack simulations, focused on 11 key contributions in terms of aims, contributions and problem statements. In turn, Alladi et al. (2020) reviewed most of real major attacks on ICS in the last 20 years. The shift from stand-alone to cloud-based CPS systems has been discussed in (Bhamare et al. (2020)), also underlying the applicability of machine learning techniques to deal with the challenge of industrial processes in migrating to cloud environments. In turn, (Alwan et al. (2022)) summarized data quality management challenges, investigating data quality issues and how it is possible to mitigate errors in sensor nodes' measurement.

Other surveys are focused on IIoT CPSs from the security requirements (Tange et al. (2020)) by reviewing literature between 2011 and 2019 and from the potential mechanism and protection strategies on the basis of nine real-world cybersecurity incidents targeting IoT devices deployed in the consumer, commercial and industrial sectors.

This section aims at highlighting the major differences of this work with respect to previous surveys and to point out the novel contribution of our work. Despite CPSs and SCADA systems have been sparsely investigated from different points of view, all research papers can be classified according to the main focus of their investigation which may be one of: attack methodologies, countermeasure methodologies or application-specific issues.

Attack-based surveys Several surveys have been drawn up analyzing the means used to perform attacks to CPSs.

Different researchers have focused their analysis on network-based attacks on CPSs (Cao et al. (2020)), mainly dealing with the intrusion detection task (Rakas et al. (2020); Franco et al. (2021)). In particular, (Franco et al. (2021)) investigated SCADA-based Intrusion Detection Systems (IDS) through an assessment of their strengths and weaknesses while the latter analyzed honeypots and Honeynets, whose aim is to deceive an intruder in accessing to a real system. A paper from (Luo et al. (2021)) investigated attacks on the network communication layer (Denial of Service, Man-In-The-Middle, Packet Injection) and control system (Malware, false control signals). Other researchers investigated network-based attacks over control schemes for power systems (Yohanandhan et al. (2020)) or feedback control loops in a CPS (Kim et al. (2022)). In particular, the latter classified these attacks according to three different dimensions: i) *attack space*, representing requirements to design an attack, ii) *attack location*, describing network-based data used for making an attack and iii) *stealthiness*, expressing if an anomaly detector can recognize an attack or not.

Other surveys (Aoufi et al. (2020); Alimi et al. (2020); Cui et al. (2020)) reviewed attacks for deceiving the control system of CPS by injecting false data into measurement from sensors. (Aoufi et al. (2020)) analyzed False Data Injection attacks targeting different components of the online power system security. Different classifications are further designed on the basis of targeted components and attack impact, which can be physical and/or economic. False data injection attacks are further investigated in (Cui et al. (2020)), whose authors provide an overview of security concerns generated by these attacks in smart grids.

The last category of surveys (Li et al. (2020); Olowononi et al. (2021); Liu et al. (2021)) investigated adversarial attacks on machine learning models for deceiving their internal architectural. (Li et al. (2020)) investigated a general working flow for adversarial attacks on input sensors (textual interface, audio speakers, surveillance and inertial measurement unit sensors) of CPSs, also providing a taxonomy for organizing existing attacks. Another survey about adversarial attacks on machine learning models has been discussed by (Olowononi et al. (2021)), whose aim is to provide some recommendations to design resilient CPSs against these attacks. In (Liu et al. (2021)), the authors investigated malicious behaviors of adversaries on the Deep Reinforcement Learning controller by designing function-based and performance-based attacks, that can be performed during and after the training phase, respectively.

[2] https://www.scopus.com/.

[3] These results are gathered by performing the following query on SCOPUS system: *TITLE-ABS-KEY ("CPS" AND "SECURITY) AND SUBJAREA (comp) AND PUBYEAR AFT 2014*.

Countermeasures-based surveys Since a CPS is often a critical infrastructure, the need of countermeasures in Industrial Control systems has been discussed in both (Alladi et al. (2020); Ahmed and Zhou (2020)). The former discussed some possible countermeasures against attacks damaging physical equipment while the latter summarized the challenges and the proposed solutions for securing CPS from a physics-based perspective. In (Huang et al. (2022b)), the authors investigated countermeasures against attacks to the power grid by designing a linear programming algorithm for recovering within the attacked area under unknown post-attack power injections.

Other surveys have been designed for investigating the most suitable defensive countermeasures in CPSs, which can be classified according to designed AI models (Ahanger et al. (2022); Aldweesh et al. (2020)) (deep or machine learning) or methodology (Suaboot et al. (2020); Franco et al. (2021); Sengupta et al. (2020)) (IDS, Honeypot or moving target defense).

Machine learning models have been widely applied to CPS security for dealing with CPS vulnerabilities. In (Ahanger et al. (2022)), the authors investigated IoT vulnerabilities by exploring different Machine and Deep Learning techniques to mitigate attack effects. With a specific reference to power systems, (Alimi et al. (2020)) provided a survey of cyber-attack detection through machine learning-based techniques and their effects on power quality (PQ), also discussing and assessing transient stability challenges.

Other machine learning models have been reviewed by (Cui et al. (2020); Alimi et al. (2020); Selim et al. (2021)) to deal with different tasks for improving the reliability of CPSs. (Cui et al. (2020)) dealt with the false data attacks by using machine learning techniques in smart grids, also providing a taxonomy of machine learning-based countermeasures. Machine learning models have been further investigated by (Alimi et al. (2020)) resulting in effective decision-making and control actions in the secured and stable operations of the power system. In turn, (Selim et al. (2021)) reviewed machine learning methods for anomaly events classification and detection in CPSs.

Other surveys (Aldweesh et al. (2020); Luo et al. (2021)) are focused on using deep learning models for dealing with anomaly detection task.

Other countermeasures aimed at detecting intrusion attacks through intrusion detection systems (Suaboot et al. (2020); Rakas et al. (2020); Bashendy et al. (2022)) or at mitigating them by designing honeypots (Franco et al. (2021); Maesschalck et al. (2022)) or moving target defense (Sengupta et al. (2020); Zhou et al. (2022c)).

Finally, a few recent surveys (Ghimire and Rawat (2022); Ferrag et al. (2021)) discussed issues on the designing of Federated Learning architectures, which aim to satisfy the privacy and security requirements of modern CPS to which a large number of devices are connected to.

Application areas-based surveys CPSs have been reviewed from different points of view as shown in (Alladi et al. (2020)), where the authors summarized the major industrial attacks, in terms of economic loss and physical equipment damage, in the last 20 years. In the following, we classify the reviews according to the application area (i.e., smart grid, power system and water system) that they analyzed.

In the last years, the power systems have evolved in CPSs by integrating more and more Information and Communication Technologies, becoming vulnerable to cybersecurity risks and attacks. In (Yohanandhan et al. (2020)), several modeling control schemes have been reviewed for power systems, also providing a taxonomy of network-based attack models. (Amin et al. (2021)) reviewed CPS attacks mitigation on Power Electronic Systems, also classifying possible threats on the smart grid. A further review has been done by (Alimi et al. (2020)) for security and stability in Power System infrastructure by analyzing well-known machine learning models (e.g., Artificial Neural Networks, Decision Trees, Support Vector Machines).

Smart grids are becoming increasingly complex systems due to the number of interconnected devices, becoming susceptible to a growing threat of attacks. In (Nafees et al. (2022)), the authors reviewed the security issues in smart grids by investigating attacks and their impacts as well as the detection schema (e.g., intrusion detection systems, moving target defense and co-simulation techniques) through situational awareness and power system metrics. (Gunduz and Das (2020)) designed a further review of threats and the potential solution in the smart grid, focusing on network-based vulnerabilities. (Zhang et al. (2021a)) classified attacks to smart grid systems in terms of targeted components and discussed several operational and informational defense approaches. Machine learning models have been evaluated by (Cui et al. (2020)) for detection of false data injection attacks in smart grids. Water distribution systems adopt modern technologies in the entire water process (starting from the water supply until to recycling), being the target of different types of attacks. (Addeen et al. (2021)) investigated cyber-physical attacks and possible mitigating detection approaches for water distribution systems, also evaluating their impacts. In turn, (Selim et al. (2021)) provided a study aimed at detecting anomalous activities on water systems infrastructure by classifying the anomaly events through machine learning algorithms.

In spite of the huge number of currently available research surveys, we felt that a review of existing contributions was needed to compare research proposals from a broader methodological perspective, as shown in Table 1. Hence, our work was guided by the following intentions, which represent an innovative point of view with respect to the state-of-the-art analysis:

- instead of classifying attacks on the basis of ICS vulnerabilities (Rakas et al. (2020)) or their specific features (Kim et al. (2022)), we propose an attack classification according to the methodology used by the attacker;
- instead of focusing on a specific application area or category of task (Luo et al. (2021)), attacks (Cao et al. (2020); Cui et al. (2020)) and/or countermeasures (Sengupta et al. (2020); Ahanger et al. (2022)), we classify the state-of-the-art approaches by comparing homogeneous research contributions adopting similar methodologies;
- we cover a broader set of attacks and countermeasures, as shown in Table 1;
- to encourage readers to perform a comparative evaluation of novel proposals, we review the most used openly available datasets, which can be used to reproduce general application scenarios;
- to guide future research efforts of communities, we provide a discussion about open issues, which need to be further addressed from different viewpoints (e.g., data quality, methodology of analysis and technological constraints).

4. Attack classification

In this section, we propose a classification of attacks to ICSs.

In a broader sense, the term attack refers to a deliberate unauthorized action on a system or asset. An attack may be carried out with different malicious objectives, including:

- denial of service,
- physical damages to the system components,
- physical damages to the system users and/or operators,
- information theft.

An attack may be carried out thanks to the knowledge of some security vulnerabilities of the target system.

From a general perspective, attacks against cyber-physical systems can be classified in two broad categories: *physical attacks* and *cyber attacks* depending on the first system weakness that is exploited to start it. Physical attacks require physical access to produce a perturbation of the system by means of tangible malicious activities (e.g. cable disconnection, physical destruction of components, unauthorized access to

Table 1
Comparison between our survey and state-of-the-art ones.

Paper	Attacks										Countermeasure								
	False Data Injection	Topology	Load Redistribution	Stealthy	Adversarial	Man-in-the-Middle	Man-on-the-Side	Spoofing	Denial of Service	Replay	Optimization	Matrix	Statistical	Fuzzy Logic	Game Theory	Reinforcement	Machine Learning	Deep Learning	Federated Learning
(Cao et al. (2020))				✓		✓			✓	✓									
(Rakas et al. (2020))	✓					✓			✓	✓									
(Franco et al. (2021))						✓			✓										
(Luo et al. (2021))	✓					✓			✓								✓	✓	
(Aoufi et al. (2020))	✓	✓											✓				✓		
(Cui et al. (2020))	✓																✓		
(Li et al. (2020))					✓												✓	✓	
(Olowononi et al. (2021))					✓											✓	✓	✓	
(Liu et al. (2021))	✓															✓			
(Huang et al. (2022b))		✓											✓						
(Aldweesh et al. (2020))																		✓	
(Suaboot et al. (2020))																	✓		
(Selim et al. (2021))																	✓		
(Bashendy et al. (2022))									✓							✓			
(Ghimire and Rawat (2022))																			✓
(Ferrag et al. (2021))																			✓
(Amin et al. (2021))		✓																	
(Gunduz and Das (2020))	✓					✓				✓									
(Zhang et al. (2021a))	✓	✓	✓																
Our Contribution	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

machinery control panels, etc.) and can be avoided by adopting proper physical security countermeasures, such as surveillance systems, locks, protective barriers, electronic access control systems, and physical intrusion detection systems. On the other hand, we use the cyber attack term to refer to malicious activities that can be perpetrated by means of networked electronic devices and exploit hardware or software vulnerabilities of the system.

In the last few years, many research works have provided a survey of the different kinds of attacks that have targeted Industrial Control Systems (Alladi et al. (2020); Bout et al. (2022)). Paper (Alladi et al. (2020)), published in 2020, lists twelve attacks that were perpetrated against the ICS of critical infrastructures since 2003. In particular, for each of the presented attacks, the authors describe the attack methodology used and suggest possible solutions to prevent such attacks. The same paper also presents a short survey of a selection of other eleven research papers, published between 2004 and 2018, that investigated security issues in Industrial Control Systems. In (Bout et al. (2022)), the authors investigated Artificial Intelligence models for generating more effective attacks on IoT networks on the basis of different learning strategies (Supervised, Unsupervised and Reinforcement). Furthermore, attacks based on the integration of ML schemes (also named *smart attacks*) have been classified into four categories according to the type of attack on IoT network: Data Analysis, Behavioral deduction, Data Generation and Behavioral Diversion. In turn, other surveys are more focused on a specific class of attacks as well as: false data injection (Aoufi et al. (2020); Cui et al. (2020)), network-based (Yohanandhan et al. (2020); Aldweesh et al. (2020)) and intrusion attacks (Suaboot et al. (2020); Rakas et al. (2020)).

Several attacks classification (Rakas et al. (2020); Xenofontos et al. (2022); Kim et al. (2022)) have been provided in literature from different points of view. One possible classification has been proposed by (Rakas et al. (2020)) according to CPS vulnerabilities: *policy and procedure*, *architecture and design*, *configuration and maintenance*, *physical*, *software development*, and *communication/network*. (Kim et al. (2022)) provided a further classification according to three different perspectives: space, location and stealthiness of an attack. In turn, (Xenofontos et al. (2022)) classified IoT attacks into four categories: *Device*, *Infrastructure*, *Communication* and *Service* attacks. Other classifications focused on a particular class of attacks or tasks in a specific ICS, as shown in (Aoufi et al. (2020)) and (Selim et al. (2021) for False Data Injection in Smart Grid and Anomaly events in the Internet of Thing infrastructure, respectively.

To provide attacks' classification, other two parameters have to be further considered: attacker's goal and attack impact.

Concerning the first issue, it is possible to classify the goals of an attacker into three different categories according to (Gollmann et al. (2015)): i) *damage to production*, whose goal is to jointly compromise product quality and increase operating costs; (ii) *damage to equipment*, which affects the effectiveness of equipment by stressing it, reducing its Remaining Useful Life, and violating safety limits; (iii) *compliance violations*, which interfere with government regulations to be met, such as increasing environmental pollution.

In turn, the second issue is mainly related to the analysis of the severity of damage inflicted when the attacker's goal has been achieved, as shown in (Urbina et al. (2016); Umsonst et al. (2017)). It often requires performing a risk assessment with the goal of classifying vulnerabilities that can be exploited by an attacker to achieve its goal (see (Lanotte et al. (2018, 2021)) for more details). For instance, (Huang et al. (2018)) designed a Bayesian network-based approach to model the attack propagation process in order to infer the probability of compromise of sensors and actuators.

In this section, we investigate the security issues of ICSs by discussing different categories of attacks. For the aim of this survey, we propose to classify attacks as shown in Fig. 3 according to the methodology used to perform the attack. From a broader perspective, we distinguish between *data-driven attacks*, i.e. those that are perpetrated by injecting deceptive data into the system, and *network-driven attacks*, i.e. those that are performed by injecting or suppressing network packets or, more generally, altering the communication flows in the communication infrastructure.

4.1. Deception attack

The pervasive diffusion of ICS in industrial processes has driven a tight interconnection between distributed physical components of critical infrastructures. Such components include field sensors, whose data

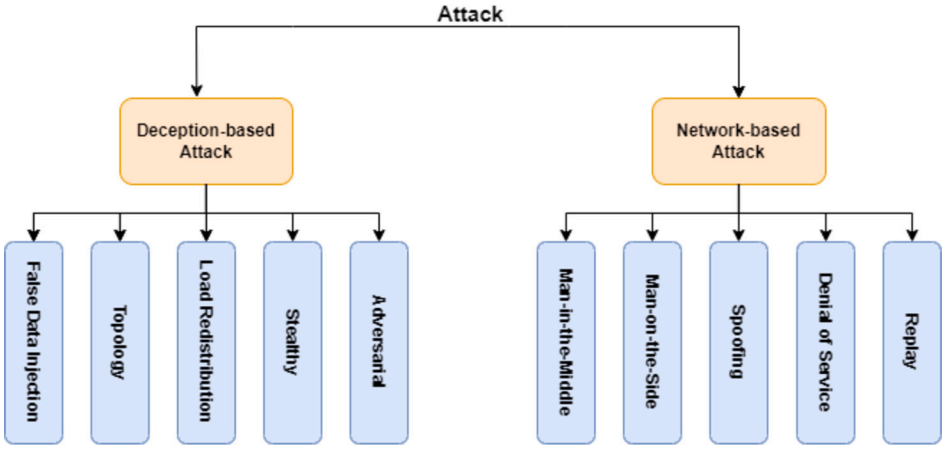

Fig. 3. Attacks classification in ICSs according to the methodology used to perform the attack.

are collected to control and optimize the entire industrial cycle. Nevertheless, the collected measurements can be maliciously altered by injecting deceptive values. Examples of such attacks have been shown in (Kravchik et al. (2021)), in which the authors investigated poisoning attacks (*interpolation* and *back*-gradient) on CPS neural network detectors.

Deception attacks *aim to compromise the security of a CPS by injecting biased information into collected measures from sensors and/or generating false control messages to deceive the CPS control systems.*

Despite different strategies have been proposed to mitigate the impact of these attacks through moving target defense mechanisms (Hu et al. (2021a)), artificial intelligence models Ozay et al. (2016); Camana Acosta et al. (2020); Wu et al. (2021a), statistical methods (Zhao et al. (2020); Jorjani et al. (2021); Huang et al. (2021); Li and Zhao (2021)) and fuzzy-based methodologies (Zhao et al. (2019)) or designing decoys to obfuscate the real components (Cifranic et al. (2020)), several challenges are still open because the attacks are becoming increasingly sophisticated (Cui et al. (2020); Gönen et al. (2020)).

False data injection (FDI) attack is an instance of deception attacks, whose aim is to deceive control systems by injecting false data in the communication line or sensor data (see (Aoufi et al. (2020)) for more details about FDI attacks on power security components). An example of these attacks has been shown in (Liu and Shu (2021)), in which the authors analyzed FDI attacks against ANN-based state estimation by injecting an attack vector into measurements through two different strategies: a population-based and a gradient-based algorithms. (Padhan and Turuk (2022)) further investigated false data injection attacks in CPS, identifying seven different ways used by an attacker for compromising both single components (actuators or sensor) or the entire system. In (Choraria et al. (2022)), the authors designed a false data injection attack on CPS by injecting deceptive values into measurements. In particular, attack's optimal parameters are inferred by combining Lagrange multiplied-based stochastic and gradient descendent optimization methods. In turn, (Tian et al. (2022a)) investigated a data-driven attack strategy based on Robust Linear Regression, that has been evaluated under different conditions of measurement data, with the aim to confound bad data detection algorithms.

Topology attack is a specific class of attacks that tightly coupled cyber and physical layers to deceive the system with the aim to make the system unavailable. In particular, physical attacks require physical access to one of the system components and can be avoided by adopting proper physical security countermeasures, such as surveillance systems, locks, protective barriers, electronic access control systems, and physical intrusion detection systems. Different researchers (Wang et al. (2021c); Zhang et al. (2020); Tu et al. (2020)) have investigated this class of attacks from different points of view. (Wang et al. (2021c)) analyzed coordinated topology attacks in smart grid, whose aim is to deceive the control center by tripping a transmission line masking the signal in the cyber-layer and, successively, creating a fake outage signal for another transmission line. In turn, patterns of sequential cyber-topology attacks have been analyzed by (Zhang et al. (2020)), in which the concept of patterns is defined as minimal attack sequences aimed at causing blackouts in cyber-physical power systems. Another example has been provided by (Tu et al. (2020)), which combines availability and integrity attacks on power grids. (Chung et al. (2019)) designed a further cyber-physical attack targeted at hitting both physical and cyber-layers of the system for deceiving the control system of power system.

Furthermore, different infrastructure solutions have been proposed to deal with this class of attacks, aiming to improve the trustworthiness between CPS components (Homay et al. (2020)) or to unveil and mitigate possible cyber-vulnerabilities in CPS through specific framework (i.e. System-Theoretic Accident Model & Processes (STAMP) (Khan and Madnick (2021))).

Load redistribution (LR) attack is a specific class of attacks that tightly coordinately compromise load and line flow measurement for deceiving the generation dispatch by deceiving the control system. In (Gao et al. (2022)), the authors proposed an approach for quantifying and evaluating the overloading associations[4] among lines under LR. (Kaviani and Hedman (2022)) designed different optimization models for generating random LR attacks to produce an overflow on a transmission line by conveniently modifying the load on each bus. In (Peng et al. (2022)), the authors modeled the game between attacker and operator in smart grid through a bi-level optimization problem, whose aim is to generate LR attack producing system's failures according to a cascading process.

Stealthy attack is distinguished from the previous case because they represent unobservable (stealthy) false data injection attacks (Cao et al. (2020)), that can compromise meters or communication systems by introducing false measurements that evade existing detection algorithms. The relevance of this type of attacks has been investigated in (Rahman et al. (2019)), where the authors provided a formal definition of *Undetected False Data Injection* (UFDI) attack on the grid topology. Different approaches have been proposed to improve the awareness about this class of deception attacks by designing detection methods, mainly based on statistical methods (Sui et al. (2021); Rahman et al. (2019)) and artificial intelligence models (Ashrafuzzaman et al. (2020); Huang and Zhu (2020); Chen et al. (2020)) or through the development of decoy farm

[4] The overloading association is defined as a statistical correlation between two lines according to their susceptibility to simultaneous overloading.

(Ajmal et al. (2021)) for confusing intruders. In particular, (Li and Yang (2022)) designed a stealthy attack model jointly maximizing the remote estimation error and guaranteeing the stealthiness to the detector. (Ren et al. (2022)) designed a stealthy attack strategy for multi-sensor networking systems by corrupting incomplete transmitted data through the scheduling effects of the Round Robin protocol. In (Zhang et al. (2022d)), the authors investigated stealthy attack under the stochastic communication protocol (SCP) with the aim to avoid collisions. Another stealthy attack has been proposed by (Chen et al. (2022a)) by simultaneously considering the attacker's cost reduction and damage production. Furthermore, (Tian et al. (2022b)) investigated a stealthy data injection attack, modeled as a sparse optimization problem, to compute perturbation for evading bad data detection and neural attack detection against state estimation methods in power systems. Finally, an instance of this class concerns the ϵ-stealthy (Li and Yang (2020)), in which the effect of the attacks is characterized by the information-theoretic analysis. A further ϵ-stealthy attack has been designed by (Zhang et al. (2021b)) through the definition of a semi-definite program problem, in which the Kallback-Leibler divergence-based stealthy constraint is converted into a linear matrix inequality.

Adversarial attack aims to compromise CPS through a set of deceptive attacks, that try to counterfeit data coming from different types of sensors (see (Li et al. (2020)) for more details about attacks and possible defenses). In particular, (Kaloudi and Li (2020)) underlined the relevance of the adversarial-based artificial models in conjunction with the conventional attacks to cause greater damage. In (Olowononi et al. (2021)), the authors investigated the resilient machine learning models at CPSs application level (i.e., Smart Cities, building or transportation), focusing on the adversarial attacks to the machine and deep learning models. In (Kravchik et al. (2022)), the authors developed two attacks (*interpolation* and *back-gradient*-based) for poisoning the learning process of the detector without compromising it. The former relies on the assumption that the initial and final stage of poisoning attack is known in advance whilst the latter is an iterative backward computation of parameters by jointly reversing the learning process and performing the second gradients in each iteration. In (Jia et al. (2021)), the authors designed an adversarial attack that simultaneously evades the anomaly detectors and rule checkers of a CPS. In particular, this attack uses gradient-based approaches and genetic algorithm for crafting noise over the sensor and actuator values and deceiving neural network and rule checking system, respectively.

Other attacks are based on the use of Generative Adversarial Network (GAN) to make more effective their effects against the ICSs. An attack based on GAN has been designed by (Chen et al. (2020)) to improve the performance of an attacker in deceiving a machine learning-based Intrusion Detection System. (Zhou et al. (2021b)) designed an attack sample generation algorithm, that, firstly compute the membership distance between dimension on the basis of weight and degree of membership distribution and it, successively, divide the dimension in sub-group on the basis of their distance to make grouping of original data. Finally, a GAN has been developed in order to increase the number of attack samples. Furthermore, the Constrained Adversarial Machine Learning (ConAML) has been proposed by (Li et al. (2021c)) to generate adversarial examples that satisfy the intrinsic constraints of the physical systems.

Furthermore, adversarial networks have been also used for making more effective FDI attacks against ICSs. (Jiao et al. (2021)) proposed a further false injection attack method by designing a Generative Adversarial Network to extract a physical model using historical measurement data, and a self-attention mechanism is integrated to further capture the power flow laws in the data. After offline training, the effective false data can be constructed in a timely fashion without system network information. An adversarial generation method has been defined by (Zeng et al. (2022)) for generating a false detection injection attack on the Deep Reinforcement Learning-based Security Constrained Optimal Power Flow (SCOPF) considering the nonlinear physical constraints in power systems through two main stages of constructor function design and unconstrained optimization problem transformation.

Finally, (Li et al. (2020)) focused their analysis on adversarial attacks on different types of data (i.e., textual content, images and videos), investigating how an intruder can compromise the system by perturbing input data.

4.2. Network-based attack

Modern ICSs implement a tight integration between cyber and physical entities through communication networks that increasingly rely on wireless links to reduce the cost of wires and installation. Unfortunately, these networks have proven to be easy to disrupt and subvert, often becoming target for cyber-attacks, that can lead to compromised system functionality.

Network-based attacks *aim to exploit vulnerabilities of network protocol with the objective of causing data loss or communication delays that compromise system functionality due to the strict real-time requirements of ICSs.*

These attacks affect the performance of different network-based protocols (i.e. Modbus, Ethernet for control automation technology (EtherCAT) and so on). (Parian et al. (2020)) analyzed a weakness in the implementation of Modbus, a protocol commonly used in SCADA systems for remote monitoring, control and acquisition. In particular, the authors developed an implementation of Modbus over TCP/IP, investigating two examples of attacks: infecting the master with a malware and a man-in-the-middle attack. Other analyses (Akpinar and Ozcelik (2021)) have been made for EtherCAT networks by using protocol-specific operations and fields to detect device-level periodicity. Different approaches have been designed to mitigate the effect of the network-based approaches, mainly focused on anomaly detection (Khan and Tomić (2021)) and statistical (Ding et al. (2021)) methods such as the design of run-time security monitor (Khan and Tomić (2021)).

In the following, we present the literature addressing most investigated network-based attacks.

Man-in-the-middle (MIMT) attack concerns the presence of an adversary within the communication process that can read, modify, and inject packets to deceive the controls of a particular system (see (Lanotte et al. (2020)) for more details). An example of this attack has been discussed by (Zhang et al. (2021b)), that analyzed the MITM attack against CPSs under the random access protocol (RAP) scheduling, where an attacker intercepts and modifies the transmitted data and then forwards them on to degrade the system performance.

Man-on-the-side attack (MotS) is a network-based attack that enables an attacker to read and inject packets. This type of attack is a weaker form than *MITM* because it cannot modify packets sent by other hosts. In a CPS, such an attack can be performed on HTTP connections by redirecting a victim to a host controlled by an adversary or by spoofing response commands to the victim (as shown in (Maynard and McLaughlin (2020))).

Spoofing attack is a notoriously well-known network-based attack whose aim is to deceive the CPS monitoring platform by crafting the spoofed response packets, as shown in (Gu et al. (2022)). The aim of an attacker is twofold: on one hand, steal models and configurations of the devices and, on other hand, provide deceptive responses on the basis of collected information. For this reason, (Gu et al. (2022)) investigated an example of the spoofing attack performed by identifying a target device and its configuration with a classification task and, later by training the classifier on the basis of spoofed responses to make the attack even more deceptive.

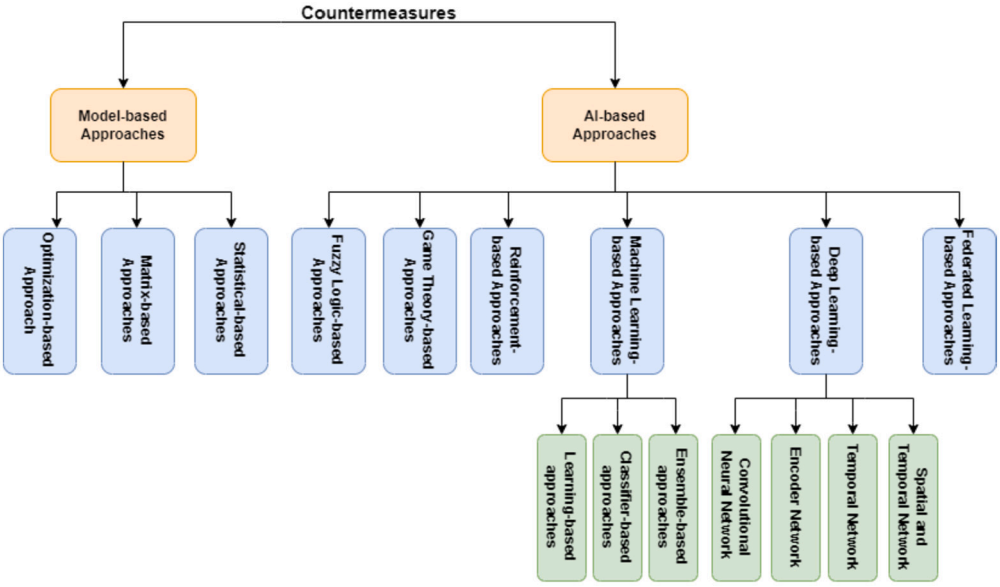

Fig. 4. Countermeasure classification in CPS and SCADA systems.

Denial of service attack (DoS) is another common network-based attack aiming to compromise the availability of a system by flooding it with a large number of network packets (see (Tripathi and Hubballi (2021); Ismail et al. (2021)) for more details). These attacks often affect ICSs due to the lack of knowledge about patterns behavior of an attacker (as shown in (Wang et al. (2021b))) as well as affecting the performance of fault estimation methods due to the measurement transmission over the network is interrupted (as shown in (Yan and Yang (2021))). Coordinated DoS attack has been further discussed in (Dong and Tian (2021)) by considering physical edge removal and overload scenario. In (Huang et al. (2022a)) investigated the DOS attack based on Double Deep Q Network (DDQN) under limited information on power allocation optimization in a multi-process CPS.

Replay attack is a network attack, whose aim is to compromise sensor and actuator by re-transmitting the data packet with malicious action on measurement or signal commands (Li et al. (2021b)). Replay attacks have been investigated by (Palleti et al. (2021)) through a case study about a water distribution plant, in which an attacker deceives sensor's measurement for exploiting the system's leak to bypass water from the main pipeline. (Liu et al. (2020a)) defined a replay attack model, in which an attacker, having access to all the real-time sensory data, generates a replay message for deceiving control system.

5. Countermeasures classification

Despite different security protection methods have been proposed for CPSs security, they were often inefficient to deal with advanced cyber-attacks (i.e. Stuxnet virus[5] or Triton malware,[6] Colonial Pipeline ransomware attack[7] and the SolarWinds[8] etc.), posing several challenges and issues about security requirements in developing countermeasures against cyber-attacks.

For this reason, some studies (Zhou et al. (2021a); Jakaria et al. (2021)) have analyzed the security requirements of individual components as well as infrastructure security schema (Fang et al. (2020)), deployment plans (Zhou et al. (2021a)) and risks assessment (Eckhart et al. (2020)). Furthermore, different metrics (Barrère et al. (2020); Duman et al. (2020)) have been defined to support the security assessment of subsystems in CPSs. The former relies on the AND/OR graphs and hypergraphs to identify the set of critical components while the latter is based on an attack graph model to capture various threats. (Rahman et al. (2019)) designed other metrics by defining the security architecture synthesis model to consider attack vectors derived from a formal model. In (Ivkić et al. (2022)), the Security Cost Modelling Framework (SCMF) has been proposed by using different metrics to measure the overall performance of a CPS.

In turn, several approaches are focused on the definition of honeypots or decoys to confuse the intruder (see (Franco et al. (2021); Zhang and Thing (2021)) for more details). Decepti-SCADA (Cifranic et al. (2020)) is a framework that implements SCADA-specific decoys that can be easily deployed in a critical infrastructure environment. From a broader perspective, (Eckhart et al. (2020)) designed a method based on data representation of CPS aiming to identify security risks, sources and potential consequences to cyber-physical attack graphs. In (Maesschalck et al. (2022)), the authors investigated how honeypots can be integrated within organizations' defensive strategy, also discussing relevant legislation, associated industry-based standards, and guidelines supporting operator compliance.

In this section, we first propose a classification of defensive countermeasures against the attacks on CPSs discussed in section 4. Fig. 4 shows our proposal, in which countermeasures are classified, at a first level, in *model-based* and *AI-based* approaches. The former type comprised methodologies whose aim is to detect anomalous behaviors based on precise models of the system, whilst the latter relies on different AI-based models that are able to detect attacks in dynamic and complex scenarios that are hard to capture in a single complete model.

5.1. Model-based approaches

In this section, we investigate state-of-the-art countermeasures to deal with attacks on ICss. In particular, two main categories are identified on the basis of the applied methodology: *Traditional* and *Statistical* approaches.

Optimization-based approaches This category aims to design approaches, which are summarized in Table 2, for dealing with optimization

[5] https://edition.cnn.com/2011/11/08/tech/iran-stuxnet/index.html.
[6] https://home.treasury.gov/news/press-releases/sm1162.
[7] https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password.
[8] https://www.businessinsider.com/solarwinds-hack-explained-government-agencies-cyber-security-2020-12?r=US&IR=T.

Table 2
Synthesis of the main Optimization-based approaches according to the addressed task and the related optimization strategy.

Paper	Type	Attack	Task
Wu et al. (2020)	Multimode Resource-Constrained project scheduling	Topological	Resource Allocation
Huang et al. (2022b)	Linear Programming-based model	False Data Injection	Resource Allocation
Zhou et al. (2022b)	Linear Programming-based model	Load Redistribution	Resource Allocation
Kaviani and Hedman (2021)	Linear Programming-based model	Load Redistribution	Resource Allocation

Table 3
Synthesis of the main Matrix-based approaches according to the addressed task and the related optimization strategy.

Paper	Type	Attack	Task
Huang et al. (2021)	Sparse Matrix	False Data Injection	Classification
Liu and Labeau (2021)	Sparse Matrix	False Data Injection	Classification

problem under different constraints. In particular, they rely on optimization strategies for improving the resilience of CPSs.

Several studies aimed to build an optimization model for improving resilience of CPS by designing specific allocation resource strategies. A project scheduling problem under resource allocation has been proposed by (Wu et al. (2020)) to incorporate system resilience, also embedding cascading failure model with the aim to evaluate system real-time performance during the recovery process and to determine whether repaired components can be reconnected to the system. In (Huang et al. (2022b)), the authors investigated the task of line state recovery after power injection attack by proposing a linear programming-based model. A corrective dispatch based on dynamic line rating has been proposed by (Zhou et al. (2022b)) in order to mitigate LR attacks by formulating a weighted multi-objective function between identifying optimal line rating and assigning the related mitigation strategy. In (Kaviani and Hedman (2021)), the authors designed a LR detection mechanism based on a linear programming optimization model for defining an attack problem, whose aim is to maximize a branch overflow. In particular, a greedy algorithm has been designed to solve the optimization problem by finding the best attack vector and identifying the most sensitive buses for critical assets.

As shown in Table 2, the linear programming model is the most used strategy for detecting anomalous patterns, that are typically represented in load redistribution (Zhou et al. (2022b); Kaviani and Hedman (2021)) and false data injection (Huang et al. (2022b)) attacks. Despite they seem to achieve good effectiveness performances, the architectural complexity, dynamic interactions among subsystems and heterogeneous information pose different challenges in defining optimization models with the related constraints, also making them NP-hard problems.

Matrix-based approaches This category aims to design approaches, which are summarized in Table 3, for classifying attacks through the definition sparse detection schema. In particular, they are focused on the analysis of sparse measurement/attack matrix for identifying attacks. Some approaches (Huang et al. (2021); Liu and Labeau (2021)) are focused on the analysis of sparse matrix in order to detect false data injection attacks. The former deal with this task by exploring the low-rank feature of the un-attacked measurement matrix whilst the latter is based on a numerical sparsity-based detection scheme to deal with the injection of false data in a wireless sensor network.

As shown in Table 3, these strategies (Kaviani and Hedman (2021); Liu and Labeau (2021)) are mainly applied to deal with false data injection attack although these methodologies may not be feasible with respect to the continuous increase of architectural complexity, stealthiness of deceptive attacks and heterogeneous information.

Statistical approach aims to model the CPS through a model-based approach, which is summarized in Table 4, relying on statistical processes.

An approach based on the Markov decision process has been designed by (Mu et al. (2020)) for estimating the cyber-attack probability, evaluated by cross entropy-based oil simulation method, for each buoy sensor. (Hu et al. (2021b)) designed another approach for dealing with the same task by defining a moving target defense technique. In particular, this approach periodically updates the measurement matrix of state estimation on the basis of device's capability to make perturbations on the transmission line susceptance with the aim to improve the probability of FDI attack detection. (Jorjani et al. (2021)) developed a FDI attack detection system through a graph theory-based approach after applying anomaly detection on the state estimation results. Another counter-measure against FDI attack has been proposed by (Miao et al. (2020)), which defined linear and non-linear attack signal estimators relying on the analysis in the frequency domain. (Sun and Yang (2022)) proposed an event-trigger communication strategy for providing input data to the remote estimator if the estimated error covariance exceeds a given threshold. In particular, the proposed model ensures the stability of the remote estimator by constraining the total attack ratio. A network-based multidimensional moving target has been designed by (Hu et al. (2021a)) for FDI attack detection in the power system by designing the packet random dropping policy for end-to-end communications.

(Li and Zhao (2021)) investigated how uncertain sensor and actuator deception attacks affect the dynamic surface control by designing a dynamic strategy on the basis of a Lyapunov function, whose non-linear terms are approximated by neural networks.

In turn, (Sui et al. (2021)) investigated the effects of stealthy attacks on a CPS, which has been modeled by a stochastic linear system,

Table 4
Synthesis of the main Statistical-based approaches according to the addressed task and the related methodology.

Paper	Type	Attack	Task
Mu et al. (2020)	Markov Decision Process	Deceptive	Classification
Hu et al. (2021b)	Moving Target Defense	False Data Injection	Classification
Jorjani et al. (2021)	Graph Theory	False Data Injection	Classification
Palleti et al. (2021)	Linear Time-Invariant System	Replay	Classification
Naha et al. (2022)	Kullback-Liebler divergence	Replay	Anomaly Detection
Sui et al. (2021)	Stochastic Linear System	Stealthy	Classification
Miao et al. (2020)	Linear and Nonlinear attack estimator	False Data Injection	Classification
Hu et al. (2021a)	Network-based Multidimensional MTD	False Data Injection	Classification
Sun and Yang (2022)	Event-Trigger State Estimator	False Data Injection	Classification
Li and Zhao (2021)	Non-Linear System	Deceptive	Classification
Li et al. (2021b)	State-space modeling	Replay	Classification
Gao and Yang (2022)	Distributed Consensus Filters	Denial of Service	Classification

Table 5
Synthesis of the main Fuzzy Logic-based approaches according to addressed task and the related methodology.

Paper	Type	Attack	Task
Sun et al. (2021)	Fuzzy Testing	Deceptive	Classification
Zhao et al. (2019)	Fuzzy C-Means	False Data Injection	Classification
Yan et al. (2022)	Takagi-Sugeno fuzzy	Denial of Service	Classification
Zhao et al. (2022)	Fuzzy Logic System	False Data Injection	Classification
Alsirhani et al. (2019)	Fuzzy Logic System	Denial of Service	Classification
Han et al. (2021)	Interval Type-2 fuzzy system	Deceptive	Classification

also providing conditions for defining their vulnerability. In (Li et al. (2021b)), a statistical approach based on Kalman filter and Linear-quadratic Gaussian controller has been defined for distinguishing the replay attacks from four different fault scenarios (controller, plant and sensor faults and plant degradation). (Palleti et al. (2021)) designed an approach based on an Input-Output Linear Time-Invariant system for detecting replay attacks in an operational water distribution plant. A statistical approach has been proposed by (Naha et al. (2022)) by inferring the Kullback-Liebler divergence between the distribution before and after the replay attack.

Finally, a distributed consensus approach has been designed by (Gao and Yang (2022)) to deal with DoS attacks when sensors and filters work at different sampling time. In particular, it concatenates sampling samples of both entities for building a novel sequence, that has been analyzed through a distributed multi-rated sampled-data H∞ consensus filter for evaluating its resilience against DoS attack.

As shown in Table 4, statistical approaches are mainly used for dealing with two attacks: network-based and deceptive. The former is often investigated by using linear systems (Li et al. (2021b); Palleti et al. (2021); Naha et al. (2022)) and distributed consensus strategies (Gao and Yang (2022)) with the aim to unveil anomalous patterns from the modeled normal ones. In turn, classifying deceptive attacks require more sophisticated strategies; in fact, both linear and non-linear systems (Miao et al. (2020); Li and Zhao (2021)) are investigated as well as the analysis of these attacks from a multi-dimensional point of view (Hu et al. (2021a)). Nevertheless, the complexity of ICS systems and information heterogeneity pose several challenges in defining linear and non-linear systems, requiring novel techniques that can handle a large amount of heterogeneous data designing properly fusion strategies or multimodal analysis.

5.2. Artificial intelligence-based models

Despite the continuous effort in defining and using Artificial Intelligence models in ICSs (Tsang and Lee (2022); Jagatheesaperumal et al. (2022); Ahanger et al. (2022)), the security requirements are becoming more and more strict; in fact, it is expected that the market for AI-based cyber-security solutions will grow up from $ 16.48 billion in 2022 to $ 93.75 billion net worth by 2030.[9]

For this reason, different researchers are using machine learning models in order to define methodologies or frameworks with the aim to deal with cyber-attacks on ICSs.

5.2.1. Fuzzy logic models

This class of approaches, that are summarized in Table 5, relies on fuzzy logic models (see (Zheng et al. (2022)) for more details) to deal with the uncertainty in identifying deceptive data by handling the truth value ranging among 0 and 1.

In (Han et al. (2021)), the authors investigated the issues concerning the stability of type-2 fuzzy systems for proposing an adaptive control on the basis of these systems to deal with deception attacks. (Zhao et al. (2019)) designed the Enhanced Gravitation Search-fuzzy c-Mean (*EGSA-FCM*) algorithm to cope with the false data injection in the power system. It works in two steps: firstly, an initial solution is computed through an optimization algorithm for searching measurement data from the SCADA system and, successively, the fuzzy c-means has been used for binary classification task by selecting the optimal clustering number through the COS clustering validity judgment index. Another approach for false data injection detection has been proposed by (Zhao et al. (2022)) by designing a fuzzy logic-based system to model the uncertain or non-linear functions for controlling wind turbine.

Other approaches have been proposed in order to design countermeasures against DoS attacks. In (Yan et al. (2022)), the authors designed a non-linear observer for different attack patterns by combining multi-gains switching mechanism and the Takagi-Sugeno fuzzy models with the aim to unveil DoS attacks. (Alsirhani et al. (2019)) developed a system for dynamic DDoS attack detection, in which an ensemble approach driven by the fuzzy logic strategy is used to select a suitable set of classification models to unveil different attack's pattern.

Finally, (Sun et al. (2021)) designed a fuzzy testing approach, whose aim is to identify honeypot by analyzing error handling. In particular, it combines mutation and security rules in the multi-objective function to generate probe packets for scanning and identification purpose.

Despite different fuzzy logic-based approaches have been proposed to deal with different attacks, as shown in Table 5, several challenges are still open. In particular, the stealthiness of deceptive attacks requires more and more sophisticated approaches for their detection due to the increasingly softwarized industrial networking in ICSs. For this reason, approaches have been developed to consider the uncertainty in detection introduced by deceptive attacks through clustering strategies (Zhao et al. (2019)), which aim to unveil anomalous patterns during the classification process, and fuzzy-testing techniques (Sun et al. (2021)) or non-linear systems (Han et al. (2021)), which use the difference in error handling between devices and honeypots or interval type-2 fuzzy defined by a set of membership functions to identify features for classifying attacks, respectively. Nevertheless, the context information and their heterogeneous nature as well as the rate at which they are produced make it difficult and costly to generate features through fuzzy techniques, requiring the definition of novel techniques that consider fusion strategies or the multimodality analysis of information extracted from ICSs. In turn, ensemble classifiers managed by Fuzzy Logic system (Alsirhani et al. (2019)) and Takagi–Sugeno fuzzy models (Yan et al. (2022)) have been designed to investigate DoS statistical patterns although the dynamic behaviors of these attacks and the increasingly softwarized industrial networking in ICSs pose different and yet open challenges.

5.2.2. Game theory-based models

This class of countermeasures relies on game theory models (see (Do et al. (2017)) for more details about their use in privacy settings), in which several games between attacker and defender have been defined according to different types of equilibrium.

In the last years, different games have been defined according to their type and the required equilibrium. In (Liu and Wang (2021)), the authors designed a *FlipIt* game model based on Montecarlo simulation to compute the probability distribution of the time-to-compromise the system of the attacks. In particular, the authors defined a Semi-Markov process, whose aim is to investigate the defense and attack

[9] https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-cybersecurity-market-report.

Table 6
Synthesis of the main Game Theory-based approaches according to the addressed task and the designed game.

Paper	Game	Attack	Task
Liu and Wang (2021)	Stackelberg game	Deceptive	Anomaly Detection
Huang and Zhu (2020)	Multi-stage Bayesian game	Deceptive	Anomaly Detection
Lau et al. (2020)	Stackelberg game	Deceptive	Anomaly Detection
Bakker et al. (2020)	Hypergame	Deceptive	Anomaly Detection
Huang and Zhu (2021)	Duplicity Game	Deceptive	Anomaly Detection
Lakshminarayana et al. (2021)	Zero-sum game	Topological	Anomaly Detection
Dai et al. (2020)	Stochastic game	Denial of Service	Anomaly Detection

strategies of the defender and attacker against the SCADA systems. A dynamic game framework has been designed by (Huang and Zhu (2020)) for deceptive attacks detection by modeling the interaction between a stealthy attacker and a proactive defender. In particular, an incomplete information game has been defined to model that each player has only information about its private information and in each stage, anyone plays its strategy according to its belief until to reach a perfect Bayesian Nash equilibrium, where no players benefit from unilateral deviations from the equilibrium. In turn, (Lau et al. (2020)) designed a Stackelberg Security Game (SSG) to allocate the defense resources on multiple targets subject to cyberattacks, modeled by Semi-Markov Process (SMP) kernel against the SCADA system. The aim of this strategy is to improve the system reliability by increasing the buffered residence time before substation failures. In (Bakker et al. (2020)), another game, named hypergame, has been defined to deal with deceptive attacks by investigating perturbation of sensor readings and calibrated parameters to deceive optimal control. Furthermore, a duplicity game framework has been designed by (Huang and Zhu (2021)) to design proactive defense deceptive strategies relying on three different modules: i) *generator*, using private information and security constraints of ICS to generate security policy, ii) *incentive modulator*, that builds constrained utility transfers between two players by reshaping incentive of users, and *trust manipulator*, in which prior belief of user are distorted over the unknowns. A moving target defense strategy based on game theory has been proposed by (Lakshminarayana et al. (2021)) for detecting topological attacks in CPS. In particular, the authors defined a zero-sum game for identifying the best subset of transmission line reactance to perturb for improving the resilience of the system. The Nash equilibrium has been achieved by observing the payoff at each iteration without requiring game complete knowledge. Finally, (Dai et al. (2020)) designed a game between sensors and attacker under two different perspectives to deal with DoS attack: open-loop, in which both types of players cannot observe others' behaviors and closed-loop case, where players can observe the others' behaviors casually. The Nash equilibrium has been computed on the basis of deep reinforcement learning algorithms for sensors and attackers, respectively.

Table 6 summarizes game theory-based approaches according to the addressed task and the designed game.

The majority of approaches in Table 6 have been proposed to deal with the uncertainty conditions in identifying deceptive attacks. In particular, Bayesian games (Huang and Zhu (2020)) are, firstly, introduced to model the uncertainty as partial information that each player holds with respect to other ones while Stackelberg (Liu and Wang (2021); Lau et al. (2020)) and Duplicity (Huang and Zhu (2021)) games have been, successively, designed to consider possible relationships between players and to investigate possible trust manipulator strategies, respectively. Nevertheless, the stealthiness of the deceptive attacks, information heterogeneity and complexity of ICS architectures pose several challenges in the design of player's actions and strategies and in the definition of the game.

Despite other approaches (Lakshminarayana et al. (2021); Dai et al. (2020)) have been designed to address other attacks in ICSs, they are mainly focused on the anomalous behaviors in ICSs, which are often difficult to identify due to the complexity of their architectures and the interactions among subsystems as well as the heterogeneity of the data.

5.2.3. Reinforcement learning

These approaches rely on the learning strategy, whose aim is to maximize the reward obtained by an agent in achieving a particular goal through an iterative procedure (see (Padakandla (2021)) for more details).

Despite the reinforcement learning strategies have been applied in different industrial environments (Wu et al. (2021b)), (Zeng et al. (2022)) investigated the possible vulnerabilities of deep reinforcement learning-based power system operation and control under the physics-constrained point of view through an assessment methodological framework. Furthermore, a survey about the Deep Reinforcement Learning (DRL) approaches for cyber-security has been designed by (Nguyen and Reddi (2021)), also discussing their application for CPS and intrusion detection as well as multi-agent DRL-based game theory simulations for dealing with cyber-attacks.

(Lei et al. (2022)) developed a reinforcement learning-based approach, whose aim is to identify the most suitable defense of chain, representing a selected number of branches due to the limited defensive resources of the system operator, by using a Q-learning search strategy to deal with the load redistribution attacks. An approach combining reinforcement learning and tree search in action space has been proposed by (Zhang et al. (2022c)) to deal with optimal phasor measurement units placements task, whose aim is to improve the resilience of the entire system by improving the complete system observability of smart grids. A DRL-based approach has been defined by (Zhang et al. (2022b)) to design a resilient defense strategy against FDI attack by inferring the optimal recovery scheme to correct false data.

In turn, other approaches have been developed to cope with DoS attacks in ICSs.(Dai et al. (2020)) designed distributed reinforcement learning algorithms for sensors and attackers to find Nash equilibrium policies under a game between sensors and attackers to deal with denial of service attack. Another DoS approach has been designed by (Chen et al. (2022b)) by developing an asynchronous schema based on a multi-agent deep reinforcement algorithm, whose global reward is maximized by centralized offline learning with shared Convolutional Neural Networks.

Table 7 summarizes reinforcement learning-based approaches according to the addressed task and the related methodology.

As shown in Table 7, three categories of Reinforcement Learning approaches (Reinforcement, deep and distributed learning) have been proposed for supporting defense and recovery strategies by improving model actions through an incentive mechanism. In particular, network-based attacks are classified through distributed strategies by considering different agents (Chen et al. (2022b)) and distributed learning strategy (Zhang et al. (2022c)) while the latter is the main investigated approach (Dai et al. (2020); Zhang et al. (2022b)) to deal with deceptive attacks for unveiling anomalous patterns with respect to normal system behavior. Despite different approaches have been designed using reinforcement learning, the lacks of available labeled datasets as well as the context dynamicity and architectural complexity pose open challenges and several issues in their use.

5.2.4. Machine learning-based models

These approaches aim to design countermeasures by focusing on the definition of strategies for learning main features and/or proper or

Table 7
Synthesis of the main Reinforcement Learning-based approaches according to the addressed task and the related methodology.

Paper	Methodology	Attack	Task
Zhang et al. (2022b)	Distributed Deep Reinforcement Learning	False Data Injection	Classification
Chen et al. (2022b)	Multi-Agent Reinforcement Learning	Denial of Service	Classification
Lei et al. (2022)	Reinforcement Learning	Topological	Anomaly Detection
Zhang et al. (2022c)	Reinforcement Learning	Deceptive	Classification
Dai et al. (2020)	Distributed Deep Reinforcement Learning	Denial of Service	Anomaly Detection

Table 8
Synthesis of the main Machine Learning-based approaches according to the addressed task and related learning strategy.

Paper	Category	Attack	Task
Liu et al. (2020b)	Learning Strategy	Deceptive	Classification
Gumaei et al. (2020)	Learning Strategy	Deceptive	Classification
Wu et al. (2021a)	Learning Strategy	False Data Injection	Classification
Farajzadeh-Zanjani et al. (2021)	Learning Strategy	Network-based attacks	Classification
Chakraborty et al. (2021)	Classifiers-based strategy	Deceptive	Classification
Shlomo et al. (2021)	Classifiers-based strategy	Network-based	Classification
Kalech (2019)	Classifiers-based strategy	Network-based	Classification
Camana Acosta et al. (2020)	Classifiers-based strategy	Stealthy	Classification
Lu et al. (2021)	Classifiers-based strategy	Network-based/Deception	Classification
Ashrafuzzaman et al. (2020)	Ensemble strategy	Stealthy	Classification
Ahmadi et al. (2022)	Ensemble strategy	False Data Injection	Classification
Upadhyay et al. (2021)	Ensemble strategy	Network-based	Classification
Jin et al. (2021)	Ensemble strategy	Deceptive	Anomaly Detection

ensemble classifiers for attack detection (see (Olowononi et al. (2021)) for more details).

Learning strategy A two-stage Bayesian learning-based approach has been designed in (Liu et al. (2020b)), whose aim is the detection of the security status in real-time and to inform the system operators with the confidence of the prediction to deal with online dynamic security assessment and preventive control in power systems. In turn, (Gumaei et al. (2020)) designed an efficient and effective security control approach, combining both feature reduction and detection techniques to reduce the extremely large number of features and achieve an improved detection rate, to detect cyber-attacks on smart grid. A correlation-based feature selection (CFS) method is used to remove irrelevant features, improving detection efficiency. An instance-based learning (IBL) algorithm classifies normal and cyber-attack events using the selected optimal features. In (Wu et al. (2021a)), the authors developed a classifier by aggregating a series of extreme learning machines to deal with FDI attacks to detect anomaly states caused by FDI attacks. Successively, a state forecasting-based bad data identification approach is proposed by exploiting the consistency between the forecast and the received measurements to identify the exact locations of the compromised measurements. Finally, an effective state recovery algorithm applies the quasi-Newton method and Armijo line search to address the possible system unobservable problem due to the removal of attacked measurements. An Adversarial Class-Imbalance Learning (ACIL) scheme has been proposed by (Farajzadeh-Zanjani et al. (2021)) by defining a loss function that iteratively adjusts weights of a multi-layer perceptron in order to learn the minority class distributions along with the majority class distribution.

ML-based classifier A machine learning approach has been designed by (Chakraborty et al. (2021)) by using Industrial IoT sensor readings for accurately tracking down Industrial IoT attacks in real-time. In (Shlomo et al. (2021)), the authors designed two machine learning algorithms to deal with dynamic malicious activities by developing two machine learning models: the former is a supervised strategy by identifying frequent temporal patterns in data payload of the SCADA communication protocols to train a classifier and the latter is an unsupervised method by learning an automaton representing system temporal behavior. Other two ML-based classifiers (Hidden Markov Models (HMM) and Artificial Neural Networks (ANN)) based on temporal patterns have been proposed by (Kalech (2019)) to identify anomalies caused by deceptive commands or misleading measurements. Finally, other two approaches (Camana Acosta et al. (2020); Lu et al. (2021)) have been proposed for detecting cyber-attacks by combining optimization strategies and machine learning-based classifiers. The former relies on the extreme randomized trees algorithm and kernel principal component analysis for dimensionality reduction while the latter is focused on a Population Extremal Optimization-based Deep Belief Network (PEO-DBN) method, in which the PEO algorithm has been used for identifying parameters for DBN, whose aim is to identify cyber-attacks.

Ensemble An ensemble approach has been defined in Ashrafuzzaman et al. (2020), where multiple classifiers are used to make decisions on the combination of individual classifiers. In particular, this schema uses two different ensemble strategies: one is based on supervised classifiers while the other one relies on unsupervised classifiers. A data-driven methodology has been designed by (Ahmadi et al. (2022)) for FDI attack detection in energy forecasting systems by using an ensemble approach based on cross-validation R^2 and z-score metric with the aim to improve its resiliency. In (Upadhyay et al. (2021)), the authors proposed an approach combining Recursive Feature Elimination-eXtreme Gradient Boosting (RFE-XGBoost)-based feature selection with a majority vote ensemble method based on Weighted Feature Importance (WFI) scores with a majority vote ensemble method using nine heterogeneous classifiers. Another ensemble strategy has been designed by (Jin et al. (2021)) for identifying anomalies and diagnosing faults in wind turbines. In this approach, normal behavior is defined on the basis of historical SCADA data and build a Mahalanobis space as a reference space, that has been used to train a model to detect anomalies. Finally, wind turbine faults are diagnosed through the analysis of the distributions and correlations of their SCADA data.

Table 8 summarizes machine learning-based approaches according to the addressed task and the related learning strategy.

These models use well-known machine learning strategies and models for attacks classification. In particular, researchers are firstly focused on the analysis of learning (Liu et al. (2020b); Gumaei et al. (2020); Wu et al. (2021a); Farajzadeh-Zanjani et al. (2021)) strategies with the

aim to deal with representation learning issues. Despite different correlation and feature selection strategies have been proposed, one of the main challenges concern the representation learning of the heterogeneous information coming from several sensors and how they can be merged for improving the knowledge of the system. Successively, different machine learning classification schemes (Kalech (2019); Lu et al. (2021); Ashrafuzzaman et al. (2020); Ahmadi et al. (2022); Upadhyay et al. (2021); Jin et al. (2021) have been designed for mainly classifying the attacks to ICSs although they are becoming more and more stealthy and sophisticated with the aim to deceive the signature on which the classification based approaches rely on.

5.2.5. Deep learning-based models

These approaches rely on models composed by different layers of artificial neural networks that manage complex data structures and feature extraction (see (Macas et al. (2022)) for more details).

Convolutional neural network-based approaches A deep Convolutional Neural Network has been designed by (Zhang et al. (2022a)) for identifying false data injection attacks by extracting temporal and spatial correlation through Kalman filter and Gaussian process regression, respectively. The FDI attack has been addressed by two approaches (Chen et al. (2022a); Zhang et al. (2022a)) by using different types of deep neural networks. The former relies on an autoencoding Gaussian Mixture Model for attack detection according to an unsupervised strategy and the latter, firstly, uses a Kalman filter for extracting spatial and temporal correlation, whose correlation with respect to the output is investigated by a deep Convolutional Neural Network.

In turn, several studies (Raman and Mathur (2022); Nedeljkovic and Jakovljevic (2022)) have been designed for using CNN to deal with anomaly detection task on ICSs. In (Raman and Mathur (2022)), the authors proposed an anomaly-based detector, which aim is to analyze the interaction among the physical components through a deep CNN. A semi-supervised learning-based approach has been developed by (Nedeljkovic and Jakovljevic (2022)), whose aim is to detect cyber-attacks on communication links between smart devices. In particular, the aim of the proposal concerns the selection of suitable CNN architecture and thresholds for intrusion detection starting from a predefined set of network hyper-parameters and normal data behavior.

Finally, (Zhou et al. (2021c)) dealt with the over-fitting issue for anomaly detection in CPS by developing a Few-Shot Learning model with Siamese Convolutional Neural Network (FSL-SCNN). In particular, a distance among input samples is computed by Siamese CNN encoding network on the basis of their optimized feature representation. Using these distance measures, anomaly detection has been developed based on a single cost function (the challenges during the few-shot learning) by combining three different losses, taking into account relative-feature representation loss, encoding loss and prediction loss, respectively.

Encoder network In (Siniosoglou et al. (2021)), the authors proposed the anoMaly dEtection aNd claSsificAtion (MENSA) IDS based on Autoencoder-Generative Adversarial Network (GAN), whose aim is twofold: unveiling anomalies in operational data and classifying attacks based on MODBUS and DNP3. (Sharmeen et al. (2022)) designed an autoencoder-based network whose aim is to learn dynamic patterns of attacks from unlabeled data to improve protection control in Smart Water Network. A stacked autoencoder has been designed by (Yang et al. (2022b)) for analyzing network traffic in industrial IoT to deal with the anomaly detection task. In turn, the FDI attack has been addressed by (Chen et al. (2022a)) through an autoencoding Gaussian Mixture Model for attack detection according to an unsupervised strategy

Temporal network (Kravchik and Shabtai (2021)) developed a deep neural network by combining 1D convolutional neural networks and autoencoders for cyber-attacks detection. In particular, this network is applied to the frequency and temporal representation of input data, also investigating the use of Principal Component Analysis (PCA) in conjunction with specific data pre-processing and feature selection. In (Alguliyev et al. (2021)), the authors designed a deep hybrid model by combining a 1D convolutional neural network, a gated recurrent unit and a long short-term memory neural network based on Scaled Polynomial Constant Unit (SPCU) activation layer. A data-driven schema has been proposed by (Chen et al. (2021)) for making resilient control under FDI attack by correlating control signal and system operational conditions by designing a LSTM-based regression predictor. Alsaedi et al. (2022) developed the UnSupervised Misbehaviour Detection (USMD) framework by integrating Temporal Attention Unit in an LSTM for learning a unified representation of the expected behaviors from multiple-sensor data on the basis of a Maximum Correntropy Criterion (MCC) as the training objective function. In (Musleh et al. (2022)), the authors proposed a deep learning-based approach for FDI attacks detection by long short-term memory autoencoder (LSTM-AE) to learn the normal dynamics utilizing this unsupervised learned model in detecting the various possibilities of FDI attacks affecting the system by evaluating the reconstruction residual of each measurements sample.

Temporal and spatial network An unsupervised anomaly detection method, named *MAD-SGCN*, has been designed by (Qi et al. (2022)) by combining LSTM and spectral-based Graph Convolutional Neural Networks for jointly extracting temporal and spatial features from time series. Spatio-temporal multiscale neural network (STMNN) has been proposed by (He et al. (2021)) to unveil spatio-temporal correlations in SCADA data. Temporal and spatial features are extracted by using a multiscale deep echo state network module and multiscale residual network module, respectively.

Table 9 summarizes deep learning-based approaches according to the addressed task and the chosen deep learning model.

The widespread use of deep learning models has led to the development of methodologies and approaches, which have mainly focused on autoencoder models (Sharmeen et al. (2022); Yang et al. (2022b)) to identify anomalies based on reconstruction error and models (Musleh et al. (2022); Chen et al. (2021); Qi et al. (2022)) to analyze the temporal pattern of sensors. More recent approaches seek to learn both temporal and spatial patterns to identify the attack signature with the aim of improving system performance. Nevertheless, the multimodality of attacks combined with the heterogeneity of information as well as the complexity of architectures still pose open challenges in identifying attacks in ICSs.

5.2.6. Federated learning-based models

In the last years, the architectural complexity, on which heterogeneity devices are connected through several network protocols, focused attention on strictly security requirements of the Industry and Supply Chain 4.0 (Bécue et al. (2021); Asante et al. (2021); Ghimire and Rawat (2022); Ferrag et al. (2021)), leading to the rise of Federated Learning architecture (see (Agrawal et al. (2022)) for more details about their application in intrusion detection).

(Li et al. (2021a)) designed a federated deep learning schema (*DeepFed*) based on Convolutional Neural Network and Gated Recurrent Unit for unveiling cyber threats in industrial CPSs. In turn, (Huong et al. (2021)) deployed an anomaly detection approach over federated learning architecture to identify anomalies in time series data produced by Industrial IoT-based manufacturing systems. In particular, the authors investigate the trade-off between a reduction in consumed bandwidth due to computation on the edges and the resulting resource consumption. A FDI attacks detection approach has been proposed by (Li et al. (2022)) by using Transformer, deployed as a detector on each edge node, for sending data to a central node through Pailler cryptosystem for making secure communication.

(Aouedi et al. (2022)) developed a federated learning scheme by using a semi-supervised strategy to combine the advantages of both unlabeled and labeled data. The representative low-dimensional fea-

Table 9
Synthesis of the main Deep Learning-based approaches according to the addressed task and the related deep learning model.

Paper	DL-model	Attack	Task
Nedeljkovic and Jakovljevic (2022)	CNN	Deceptive	Classification
Kravchik and Shabtai (2021)	1D CNN + AE	Deceptive	Classification
Alguliyev et al. (2021)	CNN+GRU+LSTM	Deceptive	Classification
He et al. (2021)	Multiscale deep echo state and residual networks	Deceptive	Classification
Chen et al. (2022a)	DAGMM	False Data Injection	Classification
Zhou et al. (2021c)	Siamese Neural Network	Network-based	Anomaly Detection
Zhang et al. (2022a)	CNN	False Data Injection	Classification
Siniosoglou et al. (2021)	Autoencoder + GAN	Network-based	Anomaly Detection Classification
Alsaedi et al. (2022)	LSTM+Attention	Deceptive	Classification
Musleh et al. (2022)	LSTM-AE	False Data Injection	Classification
Chen et al. (2021)	LSTM	False Data Injection	Classification
Qi et al. (2022)	LSTM + GCN	Deceptive	Anomaly Detection
Raman and Mathur (2022)	CNN	Stealthy	Anomaly Detection
Sharmeen et al. (2022)	Auto-Encoder	Deceptive	Anomaly Detection
Yang et al. (2022b)	Stacked Auto-Encoder	Network-based Attack	Anomaly Detection

Table 10
Synthesis of the main Federated Learning-based approaches according to the addressed task and the related Machine Learning-based schema.

Paper	ML-schema	Attack	Task
Li et al. (2021a)	CNN+GRU	Deceptive	Intrusion detection
Huong et al. (2021)	VAE-LSTM	Deceptive	Anomaly Detection
Li et al. (2022)	Multi-head self-attention-based transformer network	False Data Injection	Classification
Ahmed Khan et al. (2021)	GRU	Network-based	Classification
Aouedi et al. (2022)	AE on edge GE+FC on central server	Network-based	Classification
Truong et al. (2022)	AE+Transformer+Fourier mixing sublayer	Network-based	Anomaly Detection

tures are learned from unlabeled data by using an AutoEncoder on each device, which is aggregated on a cloud server on which a supervised network, composed of fully connected layers to the global encoder, is trained on publicly available labeled data. In turn, (Ahmed Khan et al. (2021)) designed a federated learning detection model (*DFF-SC4N*) for intrusion detection by training Gated Recurrent Unit on local data and sharing only the learned parameters with the central server. In (Truong et al. (2022)), a federated learning approach has been defined for distributed anomaly detection in ICS context by combining AutoEncoder, Transformer and Fourier sublayers on each device that sends data to a central server that aggregates them.

Table 10 summarizes federated learning-based approaches according to the addressed task and the related Machine Learning-based schema.

As shown in Table 10, approaches based on federated learning are mostly applied to deal with two attacks: Deceptive and network-based. Classifying deceptive attacks under federated architecture poses challenges about privacy-preserving issues, which are investigated by (Li et al. (2021a, 2022)) through a Pailler cryptosystem-based secure communication protocol, and detection methodologies, which are mainly based on deep learning models ((Li et al. (2021a)) and (Li et al. (2022)) use CNN+GRU and VAE+LSTM, respectively). Furthermore, approaches for network attack classification are mainly focused on the efficiency and effectiveness strategies by using GRU (Ahmed Khan et al. (2021); Li et al. (2021a)) and autoencoder (Aouedi et al. (2022); Truong et al. (2022)), respectively. Nevertheless, the heterogeneity of the information and the ICSs architectural complexity requires to design proper fusion strategies as well as architectural design for satisfying federated requirements.

6. Datasets

Since the relevance and needs of security test-beds have been underlined by (Yamin et al. (2020); Conti et al. (2021)), in this section we discuss about the available datasets for evaluating cyber-attacks and countermeasures on ICSs.

In order to assess the proposed approaches, researchers and practitioners have used different datasets, whose main information are summarized in Table 11, that can be classified in three main categories: (i) public dataset, (ii) case study and (iii) numerical analysis. Many studies evaluated the proposed approaches in the context of specific case studies Li et al. (2021b); Hu et al. (2021a); Farajzadeh-Zanjani et al. (2021), test-beds Rahman et al. (2019); Tian et al. (2022b); Wang et al. (2021c) or numerical analysis Li et al. (2021b); Han et al. (2021); Mu et al. (2020), whose details are summarized in the papers although generated data are not publicly available.

Power system dataset[10] It provided by Mississippi State University and Oak Ridge National Laboratory, contains different types of attacks on power system (see (Borges Hink et al. (2014)) for more details), under the assumption that an attacker already gained access to the system. In particular, different scenarios have been investigated on the basis of operational activities or attacks:

- *Short-circuit fault*: representing a short circuit at one of the power line points;
- *Line maintenance*: in which the relay trip command is used to open one or more breakers;
- *Remote tripping command injection attack*: opening a breaker through a relay command;
- *Replay setting change attack*: changing the protection scheme by disabling the relay function;
- *Data injection attack*: corresponding to deceive the system by simulating a fault.

The dataset has been labeled according to three different classification schemes: (i) multi-class, composed of 37 classes, (ii) three-class, aiming to assign each sample to one of three classes (attack, natural o no events) and (iii) binary, unveiling a scenario is an attack or not.

[10] https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets.

Table 11
Summary of datasets with the related classification schema and the state-of-the-art approaches evaluated on them.

Dataset	Classification Schema	Papers
Power System	Binary, Three and Multi-class	Hassan et al. (2020); Aouedi et al. (2022)
Wind Turbine	Multi-class	He et al. (2021)
SWaT	Binary	Kravchik et al. (2021) Li et al. (2021c) Chakraborty et al. (2021) Ashrafuzzaman et al. (2020) Kravchik and Shabtai (2021) Nedeljkovic and Jakovljevic (2022) Truong et al. (2022) Alguliyev et al. (2021) Alsaedi et al. (2022) Qi et al. (2022)
WADI	Multi-class	Kravchik and Shabtai (2021) Alsaedi et al. (2022) Qi et al. (2022)
Gasoil heating loop	Multi-Class	Alguliyev et al. (2021) Alsaedi et al. (2022) Nedeljkovic and Jakovljevic (2022)
Modbus	Binary	Shlomo et al. (2021); Kalech (2019)
Liquid distribution	Multi-Class	Selim et al. (2021)

Wind turbine dataset (Li et al. (2019)) It contains 49,027 samples collected from a 3 MW direct-drive Wind Turbine (WT) during one year. It has been split into three different subsets (*SCADA operational*, *Status* and *Warning* datasets), which contain data related to WT operating parameters, status and warning information, respectively. The first subset has been labeled by combining the others according to a multi-task strategy, obtaining 32,623 faulty and 32,056 normal samples. In particular, the former is, further, assigned to five different labels:

1. *Feeding faults*: indicating faults in the power feeder cables of the WT;
2. *Excitation errors*: being mainly due to problems with the generator excitation system;
3. *malfunction air cooling*: representing to problems in the air circulation and internal temperature circulation in the WT;
4. *mains failure*, corresponding to issues with mains electricity supply to the WT
5. *generator heating faults* referring to the generator overheating.

Secure water treatment (SWaT) dataset[11] *(Goh et al. (2016))* It is composed by different samples, each one having 51 attributes, to represent sensors and actuators in a secure water treatment test-bed for seven and four days under normal and attacks conditions, respectively.

Water distribution (WADI) dataset[12] *(Ahmed et al. (2017))* It contains data from an urban water distribution test-bed, that has been collected by 16, 14 and 2 days under operation and attack scenario, respectively. In particular, fifteen attacks have been considered on different components on the basis of their objective under the assumption that an attacker gains remote access to the ICS.

Gasoil heating loop (GHL) dataset (Filonov et al. (2016)) It contains data, having 19 attributes, from a testbed composed by a receiving (RT), heating (HT) and collected (HT) tanks, respectively. Furthermore, the normal process logic is affected by four different types of cyber-attacks (unauthorized change of max RT level or HT temperature as well as pump frequency and system relaxing time value).

Modbus dataset It is composed by more than 20 million packets direct to port 502 have been collected from SCADA system of Ben-Gurion University of the Negev in Israel (Kalech (2019)) to deal with a binary task (a Modbus packet is anomalous or not); in particular, the aim of this system is to control different academic subsystems (i.e., lights, heating and security).

Liquid distribution dataset[13] It contains data from a real-world scenario, whose description can be found in (Laso et al. (2017)), regarding activity about water system. In particular, 15 different anomaly situations have been represented according to 5 operational scenarios, that can affect sensor, underlying network or whole subsystem.

IEEE 14-bus test system dataset[14] It is a set of data generated on the basis of an approximation of the American Electric Power system, composed by 14 buses, 5 generators and 11 loads.

IEEE 30-bus system dataset[15] It is a set of data generated on the basis of an approximation of the American Electric Power system, composed by 30 buses, 41 transmission and different transformers and synchronous condensers.

7. Open issues

In this section, we discuss about open issues and challenges in developing counter-measures to protect ICS, which play a key role in industrial infrastructures. Despite several attacks and countermeasures have been investigated in the literature, different challenges are still open. In particular, we classify them into five main categories: i) *data availability*, whose main issues concern the availability of the datasets, mainly unbalanced and unlabeled, ii) *data quality*, relying on privacy constraints and data quality issues, iii) *data processing*, which represents issues in designing of multimodal model and fusion strategies, iv) *system complexity* and v) *explanation outcome*, that concerns challenges arise due to interconnection of heterogeneous equipment in ICS architecture and intrinsic hidden complexity of AI models, respectively.

Lacks of datasets The majority of investigated datasets is mainly proprietary, which does not allow for a fair evaluation among the different approaches. For instance, *CyberGym*[16] is a joint venture that provides training courses about SCADA cyber-defense by building a proprietary system through dedicated labs (see (Kalech (2019)) for more details). In turn, *WADI* dataset[12] has been specifically designed for investigating attacks developed by the research team that collected it. Hence, the lack of datasets limits to obtain more generalized results of the classification performance while other ones are generated with purpose-built test-beds to investigate a specific type of attack or countermeasures.

Data labeling Assigning suitable labels for each sample representing a different type of attack is the aim of data labeling, a time-consuming and expensive task that is often done manually. For instance, (Li et al. (2019)) manually defined six labels (Fault-free, Feeding, Mains, Air-cooling, Excitation, Generator) for their dataset whilst the attacks in the WADI dataset have been designed by (Goh et al. (2016)) based on 41 attacks during 4 days. Nevertheless, maintaining consistency can be fairly challenging due to the subjectivity bias of human evaluators, which has been introduced during this manual annotation process, as shown in (Guerra et al. (2022)).

[11] https://itrust.sutd.edu.sg/itrust-labs_datasets/.
[12] https://itrust.sutd.edu.sg/itrust-labs_datasets/.
[13] https://ars.els-cdn.com/content/image/1-s2.0-S2352340917303402-mmc2.zip.
[14] http://labs.ece.uw.edu/pstca/pf14/pg_tca14bus.htm.
[15] https://labs.ece.uw.edu/pstca/pf30/pg_tca30bus.htm.
[16] https://www.cybergym.com/.

Unbalanced datasets The number of samples corresponding to possible attacks on cyber-physical systems is significantly lower than the ones representing normal behavior. The dataset collected in (Li et al. (2019)) exhibits an imbalance ratio close to 57 : 1, arising a serious imbalance issue in the data. This point poses different challenges in detecting attacks, that are becoming more sophisticated and deceptive; in particular, they feign normal behavior in most cases to conceal their malicious intent for targeting a particular sub-component or ICS, as shown in (Luo et al. (2021); Li et al. (2020)).

Data quality Modern ICSs are becoming increasingly complex and integrated infrastructures, having high computation and communication capabilities and operating often under strictly real-time constraints. Hence, they focus on data, whose life cycle plays a key role in satisfying their operational requirements. Nevertheless, data issues (i.e., incorrect or missing data) can affect the security and reliability of ICSs, which often operate under dynamic and noisy conditions, making data quality yet open challenges, as shown in (Alwan et al. (2022)).

Privacy constraints The ICS systems are mainly based on proprietary industrial networks, that are often designed for protecting and managing the infrastructure of a company. The availability of these data could harm the security of the underlying infrastructure since the eventual intruder or attacker could learn the normal behavior of the network and its topology to make attacks more and more deceptive for the monitoring platform.

Design of multimodal model Modern attacks are becoming more sophisticated (i.e., Triton[6] or Stuxnet[5]) with the aim to deceive monitoring platforms, posing novel and challenging issues. Despite multimedia data analysis have shown high performance in detection task (see for instance (Di Fiore et al. (2022))), the state-of-the-art approaches often do not combine time-series data from different sensors with multimedia data (i.e., sounds generated by industrial equipment or images/video representing infrastructure plant), that can unveil symptom of possible attacks. Therefore, the heterogeneity of the modalities and their representation learning besides the lack of multi-modal datasets can affect the performance of the designed frameworks in detecting more deceptive attacks.

Design of fusion strategies In recent years, attacks have affected several sensors leading to the need of designing fusion techniques to analyze their data at the same time interval (Chen et al. (2022c); Zhou et al. (2022a)) with the aim to improve attack detection. Despite many efforts have been made for developing sensor fusion-based approaches (Gao et al. (2020); Jin et al. (2023)), modern attacks may involve different layers (network or sensor), that can be sampled at different rates, and/or require knowledge fusion techniques (see Yang et al. (2023) for more details), posing several challenges in attack detection.

System complexity In recent years, ICSs are based on increasingly complex architectures, that interconnect heterogeneous equipment via Software Defined Networks (SDN) to manage the system complexity as shown in (Wang et al. (2021a); Ray and Kumar (2021)). In fact, SDN requires a deployment plan satisfying the resiliency requirements of the ICS network according to different constraints (i.e., resiliency threats and measure or topology), as shown in (Jakaria et al. (2021)). This virtualization process posed new challenges to deal with possible attacks (see (Azab et al. (2022); Duy et al. (2021)) for instance) as well as the attacker's capability of using any breaches in the system as entry points for malware spreading, as shown in (Yang et al. (2022a)).

Explanation outcome While AI-based systems may be effective in detecting anomalous unprecedented situations, taking tough decisions, such as the stopping of a production plant, may need to be supported by a human-comprehensible explanation of the AI-based decision process. It has already been proved in the literature that eXplanable-AI (XAI) may bring useful contributions to the optimization of industrial processes (Ahmed et al. (2022)). The strict timing requirements associated with the triggering of security countermeasures lead to new research challenges that have not yet found adequate coverage in the literature.

8. Conclusion

In the last years, ICSs have become increasingly pervasive in different application areas, integrating communications networks, sensors, actuators, and human-machine interfaces within a physical system for monitoring and controlling the entire production process (Humayed et al. (2017)). The application of several technologies within modern ICSs has made them vulnerable to a plethora of attacks, exploiting different kinds of vulnerabilities that affect these systems. Nowadays, more and more sophisticated attacks are designed to provoke system unavailability or operational delays, which can cause both economic and social impact (Corallo et al. (2021)) if not properly and promptly identified or mitigated. While different studies have been proposed in the literature to survey these attacks, in this paper we provided a methodological classification of attacks and countermeasures by comparing homogeneous research contributions. Therefore, our analysis of recent proposals in literature led us to distinguish between *data* and *network*-based driven attacks. These attacks, which may affect the quality of a product or service and potentially cause significant economic losses, are also more complex to detect. In particular, we have observed that a huge effort from researchers is focusing on the characterization of deception attacks, whose goal is to deceive the control system through false data injection inside the system. To cope with these attacks, we propose a first-level classification of defense countermeasures in *model* and *AI*-based approaches. A large body of recent proposals fits in the latter category with the aim to design methodologies and frameworks that jointly use modern computational resources and the increasing amount of data produced by modern equipment.

In spite of the huge number of countermeasures summarized in Section 5, it is more and more complex for researchers to perform a comparative evaluation due to the difficulty in reproducing dynamic and complex real-world scenarios. Hence, we described the most used datasets in order to support and guide researchers and practitioners in evaluating their approaches. Although different analyses have been proposed in the literature for investigating security in ICSs, there are still several open issues in developing countermeasures, that we investigated from different points of view. We first focused on datasets availability, also covering data quality and labeling issues. Successively, we discussed challenges to deal with novel attacks deriving from the increasing architectural complexity of ICSs. Finally, we identified the challenges concerning the methodology used to analyze data from heterogeneous sensors, which may require the integration of different data modalities. And the ability to produce human-comprehensible explanations of AI-based decisions.

CRediT authorship contribution statement

Roberto Canonico: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Giancarlo Sperlì:** Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

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

Data availability

No data was used for the research described in the article.

References

Addeen, H.H., Xiao, Y., Li, J., Guizani, M., 2021. A survey of cyber-physical attacks and detection methods in smart water distribution systems. IEEE Access 9, 99905–99921. https://doi.org/10.1109/ACCESS.2021.3095713.

Agrawal, S., Sarkar, S., Aouedi, O., Yenduri, G., Piamrat, K., Alazab, M., Bhattacharya, S., Maddikunta, P.K.R., Gadekallu, T.R., 2022. Federated learning for intrusion detection system: concepts, challenges and future directions. Comput. Commun. 195, 346–361. https://doi.org/10.1016/j.comcom.2022.09.012.

Ahanger, T.A., Aljumah, A., Atiquzzaman, M., 2022. State-of-the-art survey of artificial intelligent techniques for IoT security. Comput. Netw. 206, 108771. https://doi.org/10.1016/j.comnet.2022.108771.

Ahmadi, A., Nabipour, M., Taheri, S., Mohammadi-Ivatloo, B., Vahidinasab, V., 2022. A new false data injection attack detection model for cyberattack resilient energy forecasting. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2022.3151748.

Ahmed, C.M., Zhou, J., 2020. Challenges and opportunities in cyberphysical systems security: a physics-based perspective. IEEE Secur. Priv. 18 (6), 14–22. https://doi.org/10.1109/MSEC.2020.3002851.

Ahmed, C.M., Palleti, V.R., Mathur, A.P., 2017. WADI: a water distribution testbed for research in the design of secure cyber physical systems. In: Proceedings of the 3rd International Workshop on Cyber-Physical Systems for Smart Water Networks. New York, NY, USA: Association for Computing Machinery; CySWATER '17, pp. 25–28.

Ahmed, I., Jeon, G., Piccialli, F., 2022. From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. IEEE Trans. Ind. Inform. 18 (8), 5031–5042. https://doi.org/10.1109/TII.2022.3146552.

Ahmed Khan, I., Nour Moustafa, D., Pi, D., Hussain, Y., Khan, N.A., 2021. DFF-SC4N: a deep federated defence framework for protecting supply chain 4.0 networks. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2021.3108811.

Ajmal, A.B., Alam, M., Khaliq, A.A., Khan, S., Qadir, Z., Mahmud, M.A.P., 2021. Last line of defense: reliability through inducing cyber threat hunting with deception in SCADA networks. IEEE Access 9, 126789–126800. https://doi.org/10.1109/ACCESS.2021.3111420.

Akpinar, K.O., Ozcelik, I., 2021. Methodology to determine the device-level periodicity for anomaly detection in EtherCAT-based industrial control network. IEEE Trans. Netw. Serv. Manag. 18 (2), 2308–2319. https://doi.org/10.1109/TNSM.2020.3037050.

Aldweesh, A., Derhab, A., Emam, A.Z., 2020. Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues. Knowl.-Based Syst. 189, 105124. https://doi.org/10.1016/j.knosys.2019.105124.

Alguliyev, R., Imamverdiyev, Y., Sukhostat, L., 2021. Hybrid DeepGCL model for cyber-attacks detection on cyber-physical systems. Neural Comput. Appl., 1–16. https://doi.org/10.1007/s00521-021-05785-2.

Alimi, O.A., Ouahada, K., Abu-Mahfouz, A.M., 2020. A review of machine learning approaches to power system security and stability. IEEE Access 8, 113512–113531. https://doi.org/10.1109/ACCESS.2020.3003568.

Alladi, T., Chamola, V., Zeadally, S., 2020. Industrial control systems: cyberattack trends and countermeasures. Comput. Commun. 155, 1–8. https://doi.org/10.1016/j.comcom.2020.03.007.

Alsaedi, A., Tari, Z., Mahmud, R., Moustafa, N., Mahmood, A.N., Anwar, A., 2022. USMD: UnSupervised misbehaviour detection for multi-sensor data. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2022.3143493.

Alsirhani, A., Sampalli, S., Bodorik, P., 2019. DDoS detection system: using a set of classification algorithms controlled by fuzzy logic system in apache spark. IEEE Trans. Netw. Serv. Manag. 16 (3), 936–949. https://doi.org/10.1109/TNSM.2019.2929425.

Alwan, A.A., Ciupala, M.A., Brimicombe, A.J., Ghorashi, S.A., Baravalle, A., Falcarin, P., 2022. Data quality challenges in large-scale cyber-physical systems: a systematic review. Inf. Sci. 105, 101951. https://doi.org/10.1016/j.is.2021.101951.

Amin, M., El-Sousy, F.F.M., Aziz, G.A.A., Gaber, K., Mohammed, O.A., 2021. CPS attacks mitigation approaches on power electronic systems with security challenges for smart grid applications: a review. IEEE Access 9, 38571–38601. https://doi.org/10.1109/ACCESS.2021.3063229.

Aouedi, O., Piamrat, K., Muller, G., Singh, K., 2022. Federated semi-supervised learning for attack detection in industrial Internet of things. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2022.3156642.

Aoufi, S., Derhab, A., Guerroumi, M., 2020. Survey of false data injection in smart power grid: attacks, countermeasures and challenges. J. Inf. Secur. Appl. 54, 102518. https://doi.org/10.1016/j.jisa.2020.102518.

Apruzzese, G., Pajola, L., Conti, M., 2022. The cross-evaluation of machine learning-based network intrusion detection systems. IEEE Trans. Netw. Serv. Manag., 1. https://doi.org/10.1109/TNSM.2022.3157344.

Asante, M., Epiphaniou, G., Maple, C., Al-Khateeb, H., Bottarelli, M., Ghafoor, K.Z., 2021. Distributed ledger technologies in supply chain security management: a comprehensive survey. IEEE Trans. Eng. Manag., 1–27. https://doi.org/10.1109/TEM.2021.3053655.

Ashrafuzzaman, M., Das, S., Chakhchoukh, Y., Shiva, S., Sheldon, F.T., 2020. Detecting stealthy false data injection attacks in the smart grid using ensemble-based machine learning. Comput. Secur. 97, 101994. https://doi.org/10.1016/j.cose.2020.101994.

Azab, M., Samir, M., Samir, E., 2022. "MystifY": a proactive moving-target defense for a resilient SDN controller in software defined CPS. Comput. Commun. 189, 205–220. https://doi.org/10.1016/j.comcom.2022.03.019.

Bakker, C., Bhattacharya, A., Chatterjee, S., Vrabie, D.L., 2020. Hypergames and cyber-physical security for control systems. ACM Trans. Cyber-Phys. Syst. 4 (4). https://doi.org/10.1145/3384676.

Barrère, M., Hankin, C., Nicolaou, N., Eliades, D.G., Parisini, T., 2020. Measuring cyber-physical security in industrial control systems via minimum-effort attack strategies. J. Inf. Secur. Appl. 52, 102471. https://doi.org/10.1016/j.jisa.2020.102471.

Bashendy, M., Tantawy, A., Erradi, A., 2022. Intrusion response systems for cyber-physical systems: a comprehensive survey. Comput. Secur., 102984. https://doi.org/10.1016/j.cose.2022.102984.

Bécue, A., Praça, I., Gama, J., 2021. Artificial intelligence, cyber-threats and industry 4.0: challenges and opportunities. Artif. Intell. Rev. 54 (5), 3849–3886. https://doi.org/10.1007/s10462-020-09942-2.

Bhamare, D., Zolanvari, M., Erbad, A., Jain, R., Khan, K., Meskin, N., 2020. Cybersecurity for industrial control systems: a survey. Comput. Secur. 89, 101677. https://doi.org/10.1016/j.cose.2019.101677.

Borges Hink, R.C., Beaver, J.M., Buckner, M.A., Morris, T., Adhikari, U., Pan, S., 2014. Machine learning for power system disturbance and cyber-attack discrimination. In: 2014 7th International Symposium on Resilient Control Systems (ISRCS), pp. 1–8.

Bout, E., Loscri, V., Gallais, A., 2022. How machine learning changes the nature of cyberattacks on IoT networks: a survey. IEEE Commun. Surv. Tutor. 24 (1), 248–279. https://doi.org/10.1109/COMST.2021.3127267.

Camana Acosta, M.R., Ahmed, S., Garcia, C.E., Koo, I., 2020. Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks. IEEE Access 8, 19921–19933. https://doi.org/10.1109/ACCESS.2020.2968934.

Cao, L., Jiang, X., Zhao, Y., Wang, S., You, D., Xu, X., 2020. A survey of network attacks on cyber-physical systems. IEEE Access 8, 44219–44227. https://doi.org/10.1109/ACCESS.2020.2977423.

Chakraborty, S., Onuchowska, A., Samtani, S., Jank, W., Wolfram, B., 2021. Machine learning for automated industrial IoT attack detection: an efficiency-complexity trade-off. ACM Trans. Manag. Inf. Syst. 12 (4). https://doi.org/10.1145/3460822.

Chen, C., Chen, Y., Zhao, J., Zhang, K., Ni, M., Ren, B., 2021. Data-driven resilient automatic generation control against false data injection attacks. IEEE Trans. Ind. Inform. 17 (12), 8092–8101. https://doi.org/10.1109/TII.2021.3058413.

Chen, C., Wang, Y., Cui, M., Zhao, J., Bi, W., Chen, Y., Zhang, X., 2022a. Data-Driven detection of stealthy false data injection attack against power system state estimation. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2022.3149106.

Chen, J., Gao, X., Deng, R., He, Y., Fang, C., Cheng, P., 2020. Generating adversarial examples against machine learning based intrusion detector in industrial control systems. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2020.3037500.

Chen, P., Liu, S., Chen, B., Yu, L., 2022b. Multi-agent reinforcement learning for decentralized resilient secondary control of energy storage systems against DoS attacks. IEEE Trans. Smart Grid 13 (3), 1739–1750. https://doi.org/10.1109/TSG.2022.3142087.

Chen, Y., Zhang, T., Kong, F., Zhang, L., Deng, Q., 2022c. Attack-resilient fusion of sensor data with uncertain delays. ACM Trans. Embed. Comput. Syst. 21 (4). https://doi.org/10.1145/3532181.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., Stoddart, K., 2016. A review of cyber security risk assessment methods for SCADA systems. Comput. Secur. 56, 1–27. https://doi.org/10.1016/j.cose.2015.09.009.

Choraria, M., Chattopadhyay, A., Mitra, U., Ström, E.G., 2022. Design of false data injection attack on distributed process estimation. IEEE Trans. Inf. Forensics Secur. 17, 670–683. https://doi.org/10.1109/TIFS.2022.3146078.

Chung, H.M., Li, W.T., Yuen, C., Chung, W.H., Zhang, Y., Wen, C.K., 2019. Local cyber-physical attack for masking line outage and topology attack in smart grid. IEEE Trans. Smart Grid 10 (4), 4577–4588. https://doi.org/10.1109/TSG.2018.2865316.

Cifranic, N., Hallman, R.A., Romero-Mariona, J., Souza, B., Calton, T., Coca, G., 2020. Decepti-SCADA: a cyber deception framework for active defense of networked critical infrastructures. Int. Things 12, 100320. https://doi.org/10.1016/j.iot.2020.100320.

Colombo, A.W., Karnouskos, S., Kaynak, O., Shi, Y., Yin, S., 2017. Industrial cyberphysical systems: a backbone of the fourth industrial revolution. IEEE Ind. Electron. Mag. 11 (1), 6–16. https://doi.org/10.1109/MIE.2017.2648857.

Conti, M., Donadel, D., Turrin, F., 2021. A survey on industrial control system testbeds and datasets for security research. IEEE Commun. Surv. Tutor. 23 (4), 2248–2294. https://doi.org/10.1109/COMST.2021.3094360.

Corallo, A., Lazoi, M., Lezzi, M., Pontrandolfo, P., 2021. Cybersecurity challenges for manufacturing systems 4.0: assessment of the business impact level. IEEE Trans. Eng. Manag., 1–21. https://doi.org/10.1109/TEM.2021.3084687.

Cui, L., Qu, Y., Gao, L., Xie, G., Yu, S., 2020. Detecting false data attacks using machine learning techniques in smart grid: a survey. J. Netw. Comput. Appl. 170, 102808. https://doi.org/10.1016/j.jnca.2020.102808.

Dai, P., Yu, W., Wang, H., Wen, G., Lv, Y., 2020. Distributed reinforcement learning for cyber-physical system with multiple remote state estimation under DoS attacker. IEEE Trans. Netw. Sci. Eng. 7 (4), 3212–3222. https://doi.org/10.1109/TNSE.2020.3018871.

Di Fiore, E., Ferraro, A., Galli, A., Moscato, V., Sperlì, G., 2022. An anomalous sound detection methodology for predictive maintenance. Expert Syst. Appl. 209, 118324. https://doi.org/10.1016/j.eswa.2022.118324.

Ding, D., Han, Q.L., Wang, Z., Ge, X., 2021. Recursive filtering of distributed cyber-physical systems with attack detection. IEEE Trans. Syst. Man Cybern. Syst. 51 (10), 6466–6476. https://doi.org/10.1109/TSMC.2019.2960541.

Do, C.T., Tran, N.H., Hong, C., Kamhoua, C.A., Kwiat, K.A., Blasch, E., Ren, S., Pissinou, N., Iyengar, S.S., 2017. Game theory for cyber security and privacy. ACM Comput. Surv. 50 (2). https://doi.org/10.1145/3057268.

Dong, Z., Tian, M., 2021. Modeling and vulnerability analysis of spatially embedded heterogeneous cyber-physical systems with functional dependency. IEEE Trans. Netw. Sci. Eng., 1. https://doi.org/10.1109/TNSE.2021.3114332.

Duman, O., Zhang, M., Wang, L., Debbabi, M., Atallah, R., Lebel, B., 2020. Factor of security (FoS): quantifying the security effectiveness of redundant smart grid subsystems. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2020.3009931.

Duy, P.T., Tien, L.K., Khoa, N.H., Hien, D.T.T., Nguyen, A.G.T., Pham, V.H., 2021. DIG-FuPAS: Deceive IDS with GAN and function-preserving on adversarial samples in SDN-enabled networks. Comput. Secur. 109, 102367. https://doi.org/10.1016/j.cose.2021.102367.

Eckhart, M., Ekelhart, A., Weippl, E.R., 2020. Automated security risk identification using AutomationML-based engineering data. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2020.3033150.

Engström, V., Lagerström, R., 2022. Two decades of cyberattack simulations: a systematic literature review. Comput. Secur. 116, 102681. https://doi.org/10.1016/j.cose.2022.102681.

Fang, L., Zhang, H., Li, M., Ge, C., Liu, L., Liu, Z., 2020. A secure and fine-grained scheme for data security in industrial IoT platforms for smart city. IEEE Int. Things J. 7 (9), 7982–7990. https://doi.org/10.1109/JIOT.2020.2996664.

Farajzadeh-Zanjani, M., Hallaji, E., Razavi-Far, R., Saif, M., 2021. Generative-adversarial class-imbalance learning for classifying cyber-attacks and faults - a cyber-physical power system. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2021.3118636.

Ferrag, M.A., Friha, O., Maglaras, L., Janicke, H., Shu, L., 2021. Federated deep learning for cyber security in the Internet of things: concepts, applications, and experimental analysis. IEEE Access 9, 138509–138542. https://doi.org/10.1109/ACCESS.2021.3118642.

Filonov, P., Lavrentyev, A., Vorontsov, A., 2016. Multivariate industrial time series with cyber-attack simulation: fault detection using an lstm-based predictive data model. arXiv preprint. arXiv:1612.06676.

Franco, J., Aris, A., Canberk, B., Uluagac, A.S., 2021. A survey of honeypots and honeynets for Internet of things, industrial Internet of things, and cyber-physical systems. IEEE Commun. Surv. Tutor., 1. https://doi.org/10.1109/COMST.2021.3106669.

Gao, L., Chen, B., Yu, L., 2020. Fusion-based FDI attack detection in cyber-physical systems. IEEE Trans. Circuits Syst. II, Express Briefs 67 (8), 1487–1491. https://doi.org/10.1109/TCSII.2019.2939276.

Gao, R., Yang, G.H., 2022. Distributed multi-rate sampled-data H∞ consensus filtering for cyber-physical systems under denial-of-service attacks. Inf. Sci. 587, 607–625. https://doi.org/10.1016/j.ins.2021.12.046.

Gao, S., Lei, J., Shi, J., Wei, X., Dong, M., Han, Z., 2022. Assessment of overloading correlations among transmission lines under load redistribution attacks. IEEE Trans. Smart Grid 13 (2), 1570–1581. https://doi.org/10.1109/TSG.2021.3134306.

Ghimire, B., Rawat, D.B., 2022. Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for Internet of things. IEEE Int. Things J. 9 (11), 8229–8249. https://doi.org/10.1109/JIOT.2022.3150363.

Goh, J., Adepu, S., Junejo, K.N., Mathur, A., 2016. A dataset to support research in the design of secure water treatment systems. In: International Conference on Critical Information Infrastructures Security. Springer, pp. 88–99.

Gollmann, D., Gurikov, P., Isakov, A., Krotofil, M., Larsen, J., Winnicki, A., 2015. Cyber-physical systems security: experimental analysis of a vinyl acetate monomer plant. In: Proceedings of the 1st ACM Workshop on Cyber-Physical System Security. New York, NY, USA: Association for Computing Machinery; CPSS '15, pp. 1–12.

Gönen, S., Sayan, H.H., Yılmaz, E.N., Üstünsoy, F., Karacayılmaz, G., 2020. False data injection attacks and the insider threat in smart systems. Comput. Secur. 97, 101955. https://doi.org/10.1016/j.cose.2020.101955.

Gu, Q., Formby, D., Ji, S., Saltaformaggio, B., Bourgeois, A., Beyah, R., 2022. This hacker knows physics: device physics aware mimicry attacks in cyber-physical systems. IEEE Trans. Dependable Secure Comput. 19 (5), 3218–3230. https://doi.org/10.1109/TDSC.2021.3089163.

Guerra, J.L., Catania, C., Veas, E., 2022. Datasets are not enough: challenges in labeling network traffic. Comput. Secur. 120, 102810. https://doi.org/10.1016/j.cose.2022.102810.

Gumaei, A., Hassan, M.M., Huda, S., Hassan, M.R., Camacho, D., Del Ser, J., Fortino, G., 2020. A robust cyberattack detection approach using optimal features of SCADA power systems in smart grids. Appl. Soft Comput. 96, 106658. https://doi.org/10.1016/j.asoc.2020.106658.

Gunduz, M.Z., Das, R., 2020. Cyber-security on smart grid: threats and potential solutions. Comput. Netw. 169, 107094. https://doi.org/10.1016/j.comnet.2019.107094.

Han, S., Kommuri, S.K., Lee, S., 2021. Affine transformed IT2 fuzzy event-triggered control under deception attacks. IEEE Trans. Fuzzy Syst. 29 (2), 322–335. https://doi.org/10.1109/TFUZZ.2020.2999779.

Hassan, M.M., Gumaei, A., Huda, S., Almogren, A., 2020. Increasing the trustworthiness in the industrial IoT networks through a reliable cyberattack detection model. IEEE Trans. Ind. Inform. 16 (9), 6154–6162. https://doi.org/10.1109/TII.2020.2970074.

He, Q., Pang, Y., Jiang, G., Xie, P., 2021. A spatio-temporal multiscale neural network approach for wind turbine fault diagnosis with imbalanced SCADA data. IEEE Trans. Ind. Inform. 17 (10), 6875–6884. https://doi.org/10.1109/TII.2020.3041114.

Homay, A., Chrysoulas, C., Boudani, B.E., de Sousa, M., Wollschlaeger, M., 2020. A security and authentication layer for SCADA/DCS applications. Microprocess. Microsyst., 103479. https://doi.org/10.1016/j.micpro.2020.103479.

Hu, Y., Xun, P., Zhu, P., Xiong, Y., Zhu, Y., Shi, W., Hu, C., 2021a. Network-based multidimensional moving target defense against false data injection attack in power system. Comput. Secur. 107, 102283. https://doi.org/10.1016/j.cose.2021.102283.

Hu, Y., Zhu, P., Xun, P., Liu, B., Kang, W., Xiong, Y., Shi, W., 2021b. CPMTD: cyber-physical moving target defense for hardening the security of power system against false data injected attack. Comput. Secur. 111, 102465. https://doi.org/10.1016/j.cose.2021.102465.

Huang, K., Zhou, C., Tian, Y.C., Yang, S., Qin, Y., 2018. Assessing the physical impact of cyberattacks on industrial cyber-physical systems. IEEE Trans. Ind. Electron. 65 (10), 8153–8162. https://doi.org/10.1109/TIE.2018.2798605.

Huang, K., Xiang, Z., Deng, W., Yang, C., Wang, Z., 2021. False data injection attacks detection in smart grid: a structural sparse matrix separation method. IEEE Trans. Netw. Sci. Eng. 8 (3), 2545–2558. https://doi.org/10.1109/TNSE.2021.3098738.

Huang, L., Zhu, Q., 2020. A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems. Comput. Secur. 89, 101660. https://doi.org/10.1016/j.cose.2019.101660.

Huang, L., Zhu, Q., 2021. Duplicity games for deception design with an application to insider threat mitigation. IEEE Trans. Inf. Forensics Secur. 16, 4843–4856. https://doi.org/10.1109/TIFS.2021.3118886.

Huang, M., Ding, K., Dey, S., Li, Y., Shi, L., 2022a. Learning-based DoS attack power allocation in multiprocess systems. IEEE Trans. Neural Netw. Learn. Syst., 1–14. https://doi.org/10.1109/TNNLS.2022.3148924.

Huang, Y., He, T., Chaudhuri, N.R., La Porta, T.F., 2022b. Link state estimation under cyber-physical attacks: theory and algorithms. IEEE Trans. Smart Grid 13 (5), 3760–3773. https://doi.org/10.1109/TSG.2022.3171169.

Humayed, A., Lin, J., Li, F., Luo, B., 2017. Cyber-physical systems security—a survey. IEEE Int. Things J. 4 (6), 1802–1831. https://doi.org/10.1109/JIOT.2017.2703172.

Huong, T.T., Bac, T.P., Long, D.M., Luong, T.D., Dan, N.M., Quang, L.A., Cong, L.T., Thang, B.D., Tran, K.P., 2021. Detecting cyberattacks using anomaly detection in industrial control systems: a federated learning approach. Comput. Ind. 132, 103509. https://doi.org/10.1016/j.compind.2021.103509.

ISA, I., 2008. Hackers breached colonial pipeline using compromised password. https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95. (Accessed 1 February 2022).

Ismail, S., Hassen, H.R., Just, M., Zantout, H., 2021. A review of amplification-based distributed denial of service attacks and their mitigation. Comput. Secur. 109, 102380. https://doi.org/10.1016/j.cose.2021.102380.

Ivkić, I., Sailer, P., Gouglidis, A., Mauthe, A., Tauber, M., 2022. A security cost modelling framework for cyber-physical systems. ACM Trans. Internet Technol. 22 (2). https://doi.org/10.1145/3450752.

Jagatheesaperumal, S.K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., Guizani, M., 2022. The duo of artificial intelligence and big data for industry 4.0: applications, techniques, challenges, and future research directions. IEEE Int. Things J. 9 (15), 12861–12885. https://doi.org/10.1109/JIOT.2021.3139827.

Jakaria, A.H.M., Rahman, M.A., Gokhale, A., 2021. Resiliency-aware deployment of SDN in smart grid SCADA: a formal synthesis model. IEEE Trans. Netw. Serv. Manag. 18 (2), 1430–1444. https://doi.org/10.1109/TNSM.2021.3050148.

Jia, Y., Wang, J., Poskitt, C.M., Chattopadhyay, S., Sun, J., Chen, Y., 2021. Adversarial attacks and mitigation for anomaly detectors of cyber-physical systems. Int. J. Crit. Infrastruct. Protect. 34, 100452. https://doi.org/10.1016/j.ijcip.2021.100452.

Jiao, R., Xun, G., Liu, X., Yan, G., 2021. A new AC false data injection attack method without network information. IEEE Trans. Smart Grid 12 (6), 5280–5289. https://doi.org/10.1109/TSG.2021.3102329.

Jin, X., Xu, Z., Qiao, W., 2021. Condition monitoring of wind turbine generators using SCADA data analysis. IEEE Trans. Sustain. Energy 12 (1), 202–210. https://doi.org/10.1109/TSTE.2020.2989220.

Jin, Y., Wang, S., Liu, F., Fan, H., Hu, Y., Li, X., Liu, S., 2023. Deep temporal state perception towards artificial cyber-physical systems. IEEE Int. Things J., 1. https://doi.org/10.1109/JIOT.2023.3239413.

Jorjani, M., Seifi, H., Varjani, A.Y., 2021. A graph theory-based approach to detect false data injection attacks in power system AC state estimation. IEEE Trans. Ind. Inform. 17 (4), 2465–2475. https://doi.org/10.1109/TII.2020.2999571.

Kalech, M., 2019. Cyber-attack detection in SCADA systems using temporal pattern recognition techniques. Comput. Secur. 84, 225–238. https://doi.org/10.1016/j.cose.2019.03.007.

Kaloudi, N., Li, J., 2020. The AI-based cyber threat landscape: a survey. ACM Comput. Surv. 53 (1). https://doi.org/10.1145/3372823.

Kaviani, R., Hedman, K.W., 2021. A detection mechanism against load-redistribution attacks in smart grids. IEEE Trans. Smart Grid 12 (1), 704–714. https://doi.org/10.1109/TSG.2020.3017562.

Kaviani, R., Hedman, K.W., 2022. An enhanced energy management system including a real-time load-redistribution threat analysis tool and cyber-physical sced. IEEE Trans. Power Syst. 37 (5), 3346–3358. https://doi.org/10.1109/TPWRS.2021.3135357.

Khan, M.T., Tomić, I., 2021. Securing industrial cyber–physical systems: a run-time multilayer monitoring. IEEE Trans. Ind. Inform. 17 (9), 6251–6259. https://doi.org/10.1109/TII.2020.3032968.

Khan, S., Madnick, S.E., 2021. Cybersafety: a system-theoretic approach to identify cyber-vulnerabilities amp; mitigation requirements in industrial control systems. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2021.3093214.

Kim, S., Park, K.J., Lu, C., 2022. A survey on network security for cyber–physical systems: from threats to resilient design. IEEE Commun. Surv. Tutor. 24 (3), 1534–1573. https://doi.org/10.1109/COMST.2022.3187531.

Kravchik, M., Shabtai, A., 2021. Efficient cyber attack detection in industrial control systems using lightweight neural networks and PCA. IEEE Trans. Dependable Secure Comput., 1. https://doi.org/10.1109/TDSC.2021.3050101.

Kravchik, M., Biggio, B., Shabtai, A., 2021. Poisoning attacks on cyber attack detectors for industrial control systems. In: Proceedings of the 36th Annual ACM Symposium on Applied Computing. New York, NY, USA: Association for Computing Machinery; SAC '21, pp. 116–125.

Kravchik, M., Demetrio, L., Biggio, B., Shabtai, A., 2022. Practical evaluation of poisoning attacks on online anomaly detectors in industrial control systems. Comput. Secur. 122, 102901. https://doi.org/10.1016/j.cose.2022.102901.

Lakshminarayana, S., Belmega, E.V., Poor, H.V., 2021. Moving-target defense against cyber-physical attacks in power grids via game theory. IEEE Trans. Smart Grid 12 (6), 5244–5257. https://doi.org/10.1109/TSG.2021.3095083.

Lanotte, R., Merro, M., Tini, S., 2018. Towards a formal notion of impact metric for cyber-physical attacks. In: Integrated Formal Methods: 14th International Conference. IFM 2018, Maynooth, Ireland, September 5-7, 2018, Proceedings 14. Springer, pp. 296–315.

Lanotte, R., Merro, M., Munteanu, A., Viganò, L., 2020. A formal approach to physics-based attacks in cyber-physical systems. ACM Trans. Trans. Priv. Secur. 23 (1). https://doi.org/10.1145/3373270.

Lanotte, R., Merro, M., Munteanu, A., Tini, S., 2021. Formal impact metrics for cyber-physical attacks. In: 2021 IEEE 34th Computer Security Foundations Symposium (CSF), pp. 1–16.

Laso, P.M., Brosset, D., Puentes, J., 2017. Dataset of anomalies and malicious acts in a cyber-physical subsystem. Data Brief 14, 186–191. https://doi.org/10.1016/j.dib.2017.07.038.

Lau, P., Wei, W., Wang, L., Liu, Z., Ten, C.W., 2020. A cybersecurity insurance model for power system reliability considering optimal defense resource allocation. IEEE Trans. Smart Grid 11 (5), 4403–4414. https://doi.org/10.1109/TSG.2020.2992782.

Lei, J., Gao, S., Shi, J., Wei, X., Dong, M., Wang, W., Han, Z., 2022. A reinforcement learning approach for defending against multi-scenario load redistribution attacks. IEEE Trans. Smart Grid 13 (5), 3711–3722. https://doi.org/10.1109/TSG.2022.3175470.

Li, B., Wu, Y., Song, J., Lu, R., Li, T., Zhao, L., 2021a. DeepFed: federated deep learning for intrusion detection in industrial cyber–physical systems. IEEE Trans. Ind. Inform. 17 (8), 5615–5624. https://doi.org/10.1109/TII.2020.3023430.

Li, D., Gebraeel, N., Paynabar, K., 2021b. Detection and differentiation of replay attack and equipment faults in SCADA systems. IEEE Trans. Autom. Sci. Eng. 18 (4), 1626–1639. https://doi.org/10.1109/TASE.2020.3013760.

Li, J., Liu, Y., Chen, T., Xiao, Z., Li, Z., Wang, J., 2020. Adversarial attacks and defenses on cyber–physical systems: a survey. IEEE Int. Things J. 7 (6), 5103–5115. https://doi.org/10.1109/JIOT.2020.2975654.

Li, J., Yang, Y., Sun, J.S., Tomsovic, K., Qi, H., 2021c. ConAML: constrained adversarial machine learning for cyber-physical systems. In: Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security. New York, NY, USA: Association for Computing Machinery; ASIA CCS '21, pp. 52–66.

Li, Y., Liu, S., Shu, L., 2019. Wind turbine fault diagnosis based on Gaussian process classifiers applied to operational data. Renew. Energy 134, 357–366. https://doi.org/10.1016/j.renene.2018.10.088.

Li, Y., Wei, X., Li, Y., Dong, Z., Shahidehpour, M., 2022. Detection of false data injection attacks in smart grid: a secure federated deep learning approach. IEEE Trans. Smart Grid, 1. https://doi.org/10.1109/TSG.2022.3204796.

Li, Y.G., Yang, G.H., 2020. Worst-case ϵ-stealthy false data injection attacks in cyber-physical systems. Inf. Sci. 515, 352–364. https://doi.org/10.1016/j.ins.2019.12.029.

Li, Y.G., Yang, G.H., 2022. Optimal completely stealthy attacks against remote estimation in cyber-physical systems. Inf. Sci. 590, 15–28. https://doi.org/10.1016/j.ins.2022.01.014.

Li, Z., Zhao, J., 2021. Resilient adaptive control of switched nonlinear cyber-physical systems under uncertain deception attacks. Inf. Sci. 543, 398–409. https://doi.org/10.1016/j.ins.2020.07.022.

Liu, H., Mo, Y., Yan, J., Xie, L., Johansson, K.H., 2020a. An online approach to physical watermark design. IEEE Trans. Autom. Control 65 (9), 3895–3902. https://doi.org/10.1109/TAC.2020.2971994.

Liu, J., Labeau, F., 2021. Detection of false data injection attacks in industrial wireless sensor networks exploiting network numerical sparsity. IEEE Trans. Signal Inf. Process. Netw. 7, 676–688. https://doi.org/10.1109/TSIPN.2021.3122289.

Liu, T., Shu, T., 2021. On the security of ANN-based AC state estimation in smart grid. Comput. Secur. 105, 102265. https://doi.org/10.1016/j.cose.2021.102265.

Liu, T., Liu, Y., Liu, J., Wang, L., Xu, L., Qiu, G., Gao, H., 2020b. A Bayesian learning based scheme for online dynamic security assessment and preventive control. IEEE Trans. Power Syst. 35 (5), 4088–4099. https://doi.org/10.1109/TPWRS.2020.2983477.

Liu, X., Yu, W., Liang, F., Griffith, D., Golmie, N., 2021. On deep reinforcement learning security for industrial Internet of things. Comput. Commun. 168, 20–32. https://doi.org/10.1016/j.comcom.2020.12.013.

Liu, Z., Wang, L., 2021. FlipIt game model-based defense strategy against cyberattacks on SCADA systems considering insider assistance. IEEE Trans. Inf. Forensics Secur. 16, 2791–2804. https://doi.org/10.1109/TIFS.2021.3065504.

Lu, K.D., Zeng, G.Q., Luo, X., Weng, J., Luo, W., Wu, Y., 2021. Evolutionary deep belief network for cyber-attack detection in industrial automation and control system. IEEE Trans. Ind. Inform. 17 (11), 7618–7627. https://doi.org/10.1109/TII.2021.3053304.

Luo, Y., Xiao, Y., Cheng, L., Peng, G., Yao, D.D., 2021. Deep learning-based anomaly detection in cyber-physical systems: progress and opportunities. ACM Comput. Surv. 54 (5). https://doi.org/10.1145/3453155.

Macas, M., Wu, C., Fuertes, W., 2022. A survey on deep learning for cybersecurity: progress, challenges, and opportunities. Comput. Netw. 212, 109032. https://doi.org/10.1016/j.comnet.2022.109032.

Maesschalck, S., Giotsas, V., Green, B., Race, N., 2022. Don't get stung, cover your ICS in honey: how do honeypots fit within industrial control system security. Comput. Secur. 114, 102598. https://doi.org/10.1016/j.cose.2021.102598.

Maynard, P., McLaughlin, K., 2020. Towards understanding man-on-the-side attacks (MotS) in SCADA networks. arXiv preprint. arXiv:2004.14334.

Miao, K., Shi, X., Zhang, W.A., 2020. Attack signal estimation for intrusion detection in industrial control system. Comput. Secur. 96, 101926. https://doi.org/10.1016/j.cose.2020.101926.

Mu, L., Zhao, E., Wang, Y., Zomaya, A.Y., 2020. Buoy Sensor cyberattack detection in offshore petroleum cyber-physical systems. IEEE Trans. Serv. Comput. 13 (4), 653–662. https://doi.org/10.1109/TSC.2020.2964548.

Mullet, V., Sondi, P., Ramat, E., 2021. A review of cybersecurity guidelines for manufacturing factories in industry 4.0. IEEE Access 9, 23235–23263. https://doi.org/10.1109/ACCESS.2021.3056650.

Musleh, A.S., Chen, G., Dong, Z.Y., Wang, C., Chen, S., 2022. Attack detection in automatic generation control systems using LSTM-based stacked autoencoders. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2022.3178418.

Nafees, M.N., Saxena, N., Cardenas, A., Grijalva, S., Burnap, P., 2022. Smart grid cyber-physical situational awareness of complex operational technology attacks: a review. ACM Comput. Surv. https://doi.org/10.1145/3565570.

Naha, A., Teixeira, A.M.H., Ahlen, A., Dey, S., 2022. Sequential detection of replay attacks. IEEE Trans. Autom. Control, 1. https://doi.org/10.1109/TAC.2022.3174004.

National Institute of Standards and Technology, 2020. Security and Privacy Controls for Information Systems and Organizations. Technical Report NIST Special Publication (SP) 800-53, Rev. 5. Includes updates as of October 10, 2020. U.S. Department of Commerce, Gaithersburg, MD.

Nedeljkovic, D., Jakovljevic, Z., 2022. CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. Comput. Secur. 114, 102585. https://doi.org/10.1016/j.cose.2021.102585.

Nguyen, T.T., Reddi, V.J., 2021. Deep reinforcement learning for cyber security. IEEE Trans. Neural Netw. Learn. Syst., 1–17. https://doi.org/10.1109/TNNLS.2021.3121870.

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

Ozay, M., Esnaola, I., Yarman Vural, F.T., Kulkarni, S.R., Poor, H.V., 2016. Machine learning methods for attack detection in the smart grid. IEEE Trans. Neural Netw. Learn. Syst. 27 (8), 1773–1786. https://doi.org/10.1109/TNNLS.2015.2404803.

Padakandla, S., 2021. A survey of reinforcement learning algorithms for dynamically varying environments. ACM Comput. Surv. 54 (6). https://doi.org/10.1145/3459991.

Padhan, S., Turuk, A.K., 2022. Design of false data injection attacks in cyber-physical systems. Inf. Sci. 608, 825–843. https://doi.org/10.1016/j.ins.2022.06.082.

Palleti, V.R., Mishra, V.K., Ahmed, C.M., Mathur, A., 2021. Can replay attacks designed to steal water from water distribution systems remain undetected? ACM Trans. Cyber-Phys. Syst. 5 (1). https://doi.org/10.1145/3406764.

Parian, C., Guldimann, T., Bhatia, S., 2020. Fooling the master: exploiting weaknesses in the Modbus protocol. In: Third International Conference on Computing and Network Communications (CoCoNet'19). Proc. Comput. Sci. 171, 2453–2458. https://doi.org/10.1016/j.procs.2020.04.265.

Peng, D.T., Dong, J., Yang, J., Peng, Q., 2022. Dynamical failures driven by false load injection attacks against smart grid. IEEE Trans. Inf. Forensics Secur. 17, 2213–2226. https://doi.org/10.1109/TIFS.2022.3181860.

Qi, P., Li, D., Ng, S.K., 2022. MAD-SGCN: multivariate anomaly detection with self-learning graph convolutional networks. In: 2022 IEEE 38th International Conference on Data Engineering (ICDE), pp. 1232–1244.

Rahman, M.A., Datta, A., Al-Shaer, E., 2019. Security design against stealthy attacks on power system state estimation: a formal approach. Comput. Secur. 84, 301–317. https://doi.org/10.1016/j.cose.2019.03.022.

Rakas, S.V.B., Stojanović, M.D., Marković-Petrović, J.D., 2020. A review of research work on network-based SCADA intrusion detection systems. IEEE Access 8, 93083–93108. https://doi.org/10.1109/ACCESS.2020.2994961.

Raman, M.R.G., Mathur, A.P., 2022. A hybrid physics-based data-driven framework for anomaly detection in industrial control systems. IEEE Trans. Syst. Man Cybern. Syst. 52 (9), 6003–6014. https://doi.org/10.1109/TSMC.2021.3131662.

Ray, P.P., Kumar, N., 2021. SDN/NFV architectures for edge-cloud oriented IoT: a systematic review. Comput. Commun. 169, 129–153. https://doi.org/10.1016/j.comcom.2021.01.018.

Ren, X.X., Yang, G., Zhang, X.G., 2022. Protocol-based optimal stealthy data-injection attacks via compromised sensors in cyber-physical systems. IEEE Trans. Ind. Electron., 1. https://doi.org/10.1109/TIE.2022.3169829.

Selim, G.E.I., Hemdan, E.E.D., Shehata, A.M., El-Fishawy, N.A., 2021. Anomaly events classification and detection system in critical industrial Internet of things infrastructure using machine learning algorithms. Multimed. Tools Appl. 80 (8), 12619–12640. https://doi.org/10.1007/s11042-020-10354-1.

Sengupta, S., Chowdhary, A., Sabur, A., Alshamrani, A., Huang, D., Kambhampati, S., 2020. A survey of moving target defenses for network security. IEEE Commun. Surv. Tutor. 22 (3), 1909–1941. https://doi.org/10.1109/COMST.2020.2982955.

Sharmeen, S., Huda, S., Abawajy, J., Ahmed, C.M., Hassan, M.M., Fortino, G., 2022. An advanced boundary protection control for the smart water network using semisupervised and deep learning approaches. IEEE Int. Things J. 9 (10), 7298–7310. https://doi.org/10.1109/JIOT.2021.3100461.

Shlomo, A., Kalech, M., Moskovitch, R., 2021. Temporal pattern-based malicious activity detection in SCADA systems. Comput. Secur. 102, 102153. https://doi.org/10.1016/j.cose.2020.102153.

Sinha, D., Roy, R., 2020. Reviewing cyber-physical system as a part of smart factory in industry 4.0. IEEE Eng. Manag. Rev. 48 (2), 103–117. https://doi.org/10.1109/EMR.2020.2992606.

Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P., Sarigiannidis, P., 2021. A unified deep learning anomaly detection and classification approach for smart grid environments. IEEE Trans. Netw. Serv. Manag. 18 (2), 1137–1151. https://doi.org/10.1109/TNSM.2021.3078381.

Suaboot, J., Fahad, A., Tari, Z., Grundy, J., Mahmood, A.N., Almalawi, A., Zomaya, A.Y., Drira, K., 2020. A taxonomy of supervised learning for IDSs in SCADA environments. ACM Comput. Surv. 53 (2). https://doi.org/10.1145/3379499.

Sui, T., Mo, Y., Marelli, D., Sun, X., Fu, M., 2021. The vulnerability of cyber-physical system under stealthy attacks. IEEE Trans. Autom. Control 66 (2), 637–650. https://doi.org/10.1109/TAC.2020.2987307.

Sun, Y., Tian, Z., Li, M., Su, S., Du, X., Guizani, M., 2021. Honeypot identification in softwarized industrial cyber–physical systems. IEEE Trans. Ind. Inform. 17 (8), 5542–5551. https://doi.org/10.1109/TII.2020.3044576.

Sun, Y.C., Yang, G.H., 2022. Event-triggered remote state estimation for cyber-physical systems under malicious DoS attacks. Inf. Sci. 602, 43–56. https://doi.org/10.1016/j.ins.2022.04.033.

Tange, K., De Donno, M., Fafoutis, X., Dragoni, N., 2020. A systematic survey of industrial Internet of things security: requirements and fog computing opportunities. IEEE Commun. Surv. Tutor. 22 (4), 2489–2520. https://doi.org/10.1109/COMST.2020.3011208.

Thoben, K.D., Wiesner, S., Wuest, T., 2017. "Industrie 4.0" and smart manufacturing – a review of research issues and application examples. Int. J. Autom. Technol. 11 (1), 4–16. https://doi.org/10.20965/ijat.2017.p0004.

Tian, J., Wang, B., Li, J., Konstantinou, C., 2022a. Datadriven false data injection attacks against cyber-physical power systems. Comput. Secur. 121, 102836. https://doi.org/10.1016/j.cose.2022.102836.

Tian, J., Wang, B., Li, J., Wang, Z., Ma, B., Ozay, M., 2022b. Exploring targeted and stealthy false data injection attacks via adversarial machine learning. IEEE Int. Things J. 9 (15), 14116–14125. https://doi.org/10.1109/JIOT.2022.3147040.

Tripathi, N., Hubballi, N., 2021. Application layer denial-of-service attacks and defense mechanisms: a survey. ACM Comput. Surv. 54 (4). https://doi.org/10.1145/3448291.

Truong, H.T., Ta, B.P., Le, Q.A., Nguyen, D.M., Le, C.T., Nguyen, H.X., Do, H.T., Nguyen, H.T., Tran, K.P., 2022. Light-weight federated learning-based anomaly detection for time-series data in industrial control systems. Comput. Ind. 140, 103692. https://doi.org/10.1016/j.compind.2022.103692.

Tsang, Y., Lee, C., 2022. Artificial intelligence in industrial design: a semi-automated literature survey. Eng. Appl. Artif. Intell. 112, 104884. https://doi.org/10.1016/j.engappai.2022.104884.

Tu, H., Xia, Y., Tse, C.K., Chen, X., 2020. A hybrid cyber attack model for cyber-physical power systems. IEEE Access 8, 114876–114883. https://doi.org/10.1109/ACCESS.2020.3003323.

Turton, W., Mehrotra, K., 2021. ISA95, Enterprise-Control System Integration. (Accessed 2022-02-01).

Umsonst, D., Sandberg, H., Cárdenas, A.A., 2017. Security analysis of control system anomaly detectors. In: 2017 American Control Conference (ACC), pp. 5500–5506.

Upadhyay, D., Manero, J., Zaman, M., Sampalli, S., 2021. Intrusion detection in SCADA based power grids: recursive feature elimination model with majority vote ensemble algorithm. IEEE Trans. Netw. Sci. Eng. 8 (3), 2559–2574. https://doi.org/10.1109/TNSE.2021.3099371.

Urbina, D.I., Giraldo, J.A., Cardenas, A.A., Tippenhauer, N.O., Valente, J., Faisal, M., Ruths, J., Candell, R., Sandberg, H., 2016. Limiting the impact of stealthy attacks on industrial control systems. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York, NY, USA: Association for Computing Machinery; CCS '16, pp. 1092–1105.

Wang, D., Zhao, N., Song, B., Lin, P., Yu, F.R., 2021a. Resource management for secure computation offloading in softwarized cyber–physical systems. IEEE Int. Things J. 8 (11), 9294–9304. https://doi.org/10.1109/JIOT.2021.3057594.

Wang, P.B., Ren, X.M., Zheng, D.D., 2021b. Event-triggered resilient control for cyber-physical systems under periodic DoS jamming attacks. Inf. Sci. 577, 541–556. https://doi.org/10.1016/j.ins.2021.07.002.

Wang, Z., He, H., Wan, Z., Sun, Y., 2021c. Coordinated topology attacks in smart grid using deep reinforcement learning. IEEE Trans. Ind. Inform. 17 (2), 1407–1415. https://doi.org/10.1109/TII.2020.2994977.

Wu, G., Li, M., Li, Z.S., 2020. Resilience-based optimal recovery strategy for cyber–physical power systems considering component multistate failures. IEEE Trans. Reliab., 1–15. https://doi.org/10.1109/TR.2020.3025179.

Wu, T., Xue, W., Wang, H., Chung, C.Y., Wang, G., Peng, J., Yang, Q., 2021a. Extreme learning machine-based state reconstruction for automatic attack filtering in cyber physical power system. IEEE Trans. Ind. Inform. 17 (3), 1892–1904. https://doi.org/10.1109/TII.2020.2984315.

Wu, Y., Wang, Z., Ma, Y., Leung, V.C., 2021b. Deep reinforcement learning for blockchain in industrial IoT: a survey. Comput. Netw. 191, 108004. https://doi.org/10.1016/j.comnet.2021.108004.

Xenofontos, C., Zografopoulos, I., Konstantinou, C., Jolfaei, A., Khan, M.K., Choo, K.K.R., 2022. Consumer, commercial, and industrial IoT (in)security: attack taxonomy and case studies. IEEE Int. Things J. 9 (1), 199–221. https://doi.org/10.1109/JIOT.2021.3079916.

Yamin, M.M., Katt, B., Gkioulos, V., 2020. Cyber ranges and security testbeds: scenarios, functions, tools and architecture. Comput. Secur. 88, 101636. https://doi.org/10.1016/j.cose.2019.101636.

Yan, J.J., Yang, G.H., 2021. Adaptive fault estimation for cyber-physical systems with intermittent DoS attacks. Inf. Sci. 547, 746–762. https://doi.org/10.1016/j.ins.2020.08.086.

Yan, J.J., Yang, G.H., Liu, X.X., 2022. A multi-gain switching mechanism-based secure estimation scheme against DoS attacks for nonlinear industrial cyber-physical systems. IEEE Trans. Ind. Electron., 1–10. https://doi.org/10.1109/TIE.2022.3186379.

Yang, B., Yu, Z., Cai, Y., 2022a. Malicious software spread modeling and control in cyber–physical systems. Knowl.-Based Syst. 248, 108913. https://doi.org/10.1016/j.knosys.2022.108913.

Yang, J., Yang, L.T., Wang, H., Gao, Y., Zhao, Y., Xie, X., Lu, Y., 2023. Representation learning for knowledge fusion and reasoning in cyber–physical–social systems: survey and perspectives. Inf. Fusion 90, 59–73. https://doi.org/10.1016/j.inffus.2022.09.003.

Yang, K., Shi, Y., Yu, Z., Yang, Q., Sangaiah, A.K., Zeng, H., 2022b. Stacked one-class broad learning system for intrusion detection in industry 4.0. IEEE Trans. Ind. Inform., 1. https://doi.org/10.1109/TII.2022.3157727.

Yohanandhan, R.V., Elavarasan, R.M., Manoharan, P., Mihet-Popa, L., 2020. Cyber-physical power system (CPPS): a review on modeling, simulation, and analysis with cyber security applications. IEEE Access 8, 151019–151064. https://doi.org/10.1109/ACCESS.2020.3016826.

Zeng, L., Sun, M., Wan, X., Zhang, Z., Deng, R., Xu, Y., 2022. Physics-constrained vulnerability assessment of deep reinforcement learning-based SCOPF. IEEE Trans. Power Syst., 1–15. https://doi.org/10.1109/TPWRS.2022.3192558.

Zhang, G., Li, J., Bamisile, O., Cai, D., Hu, W., Huang, Q., 2022a. Spatio-temporal correlation-based false data injection attack detection using deep convolutional neural network. IEEE Trans. Smart Grid 13 (1), 750–761. https://doi.org/10.1109/TSG.2021.3109628.

Zhang, H., Liu, B., Wu, H., 2021a. Smart grid cyber-physical attack and defense: a review. IEEE Access 9, 29641–29659. https://doi.org/10.1109/ACCESS.2021.3058628.

Zhang, H., Yue, D., Dou, C., Hancke, G.P., 2022b. Resilient optimal defensive strategy of micro-grids system via distributed deep reinforcement learning approach against FDI attack. IEEE Trans. Neural Netw. Learn. Syst., 1–11. https://doi.org/10.1109/TNNLS.2022.3175917.

Zhang, L., Thing, V., 2021. Three decades of deception techniques in active cyber defense - retrospect and outlook. Comput. Secur. 106, 102288. https://doi.org/10.1016/j.cose.2021.102288.

Zhang, M., Wu, Z., Yan, J., Lu, R., Guan, X., 2022c. Attack-resilient optimal PMU placement via reinforcement learning guided tree search in smart grids. IEEE Trans. Inf. Forensics Secur. 17, 1919–1929. https://doi.org/10.1109/TIFS.2022.3173728.

Zhang, X.G., Yang, G.H., Wasly, S., 2021b. Man-in-the-middle attack against cyber-physical systems under random access protocol. Inf. Sci. 576, 708–724. https://doi.org/10.1016/j.ins.2021.07.083.

Zhang, X.G., Yang, G.H., Ren, X.X., 2022d. Optimal stealthy attack on cyber-physical systems and its application to a networked PMSM system. IEEE Trans. Ind. Electron., 1–9. https://doi.org/10.1109/TIE.2022.3196363.

Zhang, Z., Huang, S., Liu, F., Mei, S., 2020. Pattern analysis of topological attacks in cyber-physical power systems considering cascading outages. IEEE Access 8, 134257–134267. https://doi.org/10.1109/ACCESS.2020.3006555.

Zhao, S., Yang, Q., Cheng, P., Deng, R., Xia, J., 2022. Adaptive resilient control for variable-speed wind turbines against false data injection attacks. IEEE Trans. Sustain. Energy 13 (2), 971–985. https://doi.org/10.1109/TSTE.2022.3141766.

Zhao, Y., Xu, J., Wu, J., 2019. A new method for bad data identification of oilfield system based on enhanced gravitational search-fuzzy C-means algorithm. IEEE Trans. Ind. Inform. 15 (11), 5963–5970. https://doi.org/10.1109/TII.2019.2935749.

Zhao, Z., Huang, Y., Zhen, Z., Li, Y., 2020. Data-Driven false data-injection attack design and detection in cyber-physical systems. IEEE Trans. Cybern., 1–9. https://doi.org/10.1109/TCYB.2020.2969320.

Zheng, Y., Xu, Z., Wang, X., 2022. The fusion of deep learning and fuzzy systems: a state-of-the-art survey. IEEE Trans. Fuzzy Syst. 30 (8), 2783–2799. https://doi.org/10.1109/TFUZZ.2021.3062899.

Zhou, C., Hu, B., Shi, Y., Tian, Y.C., Li, X., Zhao, Y., 2021a. A unified architectural approach for cyberattack-resilient industrial control systems. Proc. IEEE 109 (4), 517–541. https://doi.org/10.1109/JPROC.2020.3034595.

Zhou, J., Ding, W., Yang, W., 2022a. A secure encoding mechanism against deception attacks on multisensor remote state estimation. IEEE Trans. Inf. Forensics Secur. 17, 1959–1969. https://doi.org/10.1109/TIFS.2022.3175617.

Zhou, M., Wu, J., Long, C., Liu, C., Kundur, D., 2022b. Dynamic-line-rating-based robust corrective dispatch against load redistribution attacks with unknown objectives. IEEE Int. Things J. 9 (18), 17756–17766. https://doi.org/10.1109/JIOT.2022.3160864.

Zhou, W., min Kong, X., li Li, K., ming Li, X., lin Ren, L., Yan, Y., Sha, Y., ying Cao, X., jun Liu, X., 2021b. Attack sample generation algorithm based on data association group by GAN in industrial control dataset. Comput. Commun. 173, 206–213. https://doi.org/10.1016/j.comcom.2021.04.014.

Zhou, X., Liang, W., Shimizu, S., Ma, J., Jin, Q., 2021c. Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems. IEEE Trans. Ind. Inform. 17 (8), 5790–5798. https://doi.org/10.1109/TII.2020.3047675.

Zhou, Y., Cheng, G., Zhao, Y., Chen, Z., Jiang, S., 2022c. Toward proactive and efficient DDoS mitigation in IIoT systems: a moving target defense approach. IEEE Trans. Ind. Inform. 18 (4), 2734–2744. https://doi.org/10.1109/TII.2021.3090719.

Roberto Canonico is Associate Professor at University of Naples Federico II since 2005. He received the Laurea degree (cum laude) in Electronic Engineering from University of Naples Federico II in 1995, and a Ph.D. in Computer Engineering from the same University in 2000. In 2000, he was a visiting Research Associate at Lancaster University, UK. His current research interests include Software Defined Networking, applications of AI to networking, green networking, resource management in cloud-native network architectures.

Giancarlo Sperlì is an assistant professor at the Department of Electrical Engineering and Information Technology of the University of Naples Federico II. He obtained his PhD in Information Technology and Electrical Engineering at the same University defending his thesis: “Multimedia Social Networks”. He is a member of the Pattern analysis and Intelligent Computation for mUltimedia Systems (PICUS) departmental research groups. His main research interests are in the area of Cybersecurity, Multimedia data analysis and Social Networks Analysis. He served as guest editor of different special issues on International Journals. Finally, he has authored about 98 publications in international journals, conference proceedings and book chapters.