IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 14, 2020, accepted May 6, 2020, date of publication May 15, 2020, date of current version June 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994961

A Review of Research Work on Network-Based SCADA Intrusion Detection Systems

SLAVICA V. BOŠTJANČIČ RAKAS[1], (Member, IEEE), MIRJANA D. STOJANOVIĆ[2], (Member, IEEE), AND JASNA D. MARKOVIĆ-PETROVIĆ[3]

[1]Mihailo Pupin Institute, University of Belgrade, 11060 Belgrade, Serbia
[2]Faculty of Transport and Traffic Engineering, University of Belgrade, 11000 Belgrade, Serbia
[3]CE Djerdap Hydroelectric Power Plants Ltd., 19300 Negotin, Serbia

Corresponding author: Slavica V. Boštjančič Rakas (slavica.bostjancic@pupin.rs)

This work was supported in part by the Ministry of Education, Science and Technological Development of Serbia.

ABSTRACT Specific intrusion detection systems (IDSs) are needed to secure modern supervisory control and data acquisition (SCADA) systems due to their architecture, stringent real-time requirements, network traffic features and specific application layer protocols. This article aims to contribute to assess the state-of-the-art, identify the open issues and provide an insight for future study areas. To achieve these objectives, we start from the factors that impact the design of dedicated intrusion detection systems in SCADA networks and focus on network-based IDS solutions. We propose a structured evaluation methodology that encompasses detection techniques, protected protocols, implementation tools, test environments and IDS performance. Special attention is focused on assessing implementation maturity as well as the applicability of each surveyed solution in the Future Internet environment. Based on that, we provide a brief description and evaluation of 26 selected research papers, published in the period 2015–2019. Results of our analysis indicate considerable progress regarding the development of machine learning-based detection methods, implementation platforms, and to some extent, sophisticated testbeds. We also identify research gaps and conclude the analysis with a list of the most important directions for further research.

INDEX TERMS Anomaly-based detection, network security, SCADA, signature-based detection, specification-based detection.

I. INTRODUCTION

Supervisory control and data acquisition (SCADA) systems control and monitor geographically dispersed process equipment on multiple sites, often spread over large distances, where centralized data acquisition and control are essential to system operation. They are one of the most widespread types of industrial control systems (ICS) and are commonly used in the industrial sectors like electric power generation, transmission and distribution, oil refineries and natural gas distribution, water and wastewater treatment, and transportation systems. Failures and malfunctions of such systems may have serious consequences due to their strategic importance for national critical infrastructures.

SCADA systems have a rich and long history: from monolithic systems of the first generation, through distributed systems of the second generation that used proprietary network technologies, to present systems of the third generation that are fully networked and make use of the Internet technologies [1]. The upcoming fourth-generation SCADA systems adopt Industrial Internet of Things (IIoT) and the Future Internet (FIN) technologies such as cloud/fog computing, big data analytics, mobile computing, etc. At present, SCADA physical and cyber security are converging; however, that is a relatively recent phenomenon that appeared with third-generation SCADA systems, when Internet technologies started to be gradually introduced to them.

Several successful attacks on worldwide industrial control systems were notified in the past decades [2]–[5]. The SANS Institute publishes biennially the reports on attacks against SCADA and other control systems. The main conclusion of the 2017 report was that the amount of external threats affecting vital, mission-critical systems was growing annually [6]. Another report by the same source, in 2019, focused more broadly on securing the operational technology domain inside organizations [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Mamoun Alazab.

While security products in public and enterprise IT networks have reached a high level of maturity, introducing the same approach to industrial networks is not straightforward, and in many cases new solutions are needed that are tailored to the control environment. Security issue of a continual real-time system assumes a comprehensive analysis and holistic understanding of network security, control theory, and physical systems. Hurd and McCarty identified existing tools that could be used to prevent, detect, mitigate, or investigate cyber attacks in the ICS environment [8]. Their research did not include evaluation of the tools and verification of vendors' claims about tool capabilities.

The importance of protecting critical infrastructures is proved through intensive efforts of standardization bodies to provide standards and guidelines for increasing their security. Relevant standards and recommendations comprise general IT security standards, common standards and directions for protecting SCADA and industrial control systems, and specific directions concerning particular industrial sectors. A comprehensive review of SCADA security standards and recommendations can be found in the literature [2], [9]–[11].

Securing critical infrastructure received a similar level of attention from many research initiatives including European projects such as SAFEGUARD, CRUTIAL, CRISALIS, MICIE/CockpitCI and their successor ATENA [11]; CRISP in the United States; CIGRÉ D2 working groups, as well as regional and national projects. Work is continuing to assess the state-of-the-art and to identify the challenges for future research, and this article aims to contribute to these efforts.

Intrusion detection is defined as ''the process of monitoring the events occurring in a computer system or network and analyzing them for signs of possible incidents, which are violations or imminent threats of violation of computer security policies, acceptable use policies, or standard security practices'' [12].

Intrusion detection technologies are traditionally classified on the basis of recognized events and the methodology used for identification of incidents. According to the recognized event type, IDSs are most often classified to network-based (NIDS) and host-based (HIDS). Basic methodologies for incidents detection comprise signature-based detection, anomaly-based detection and specification-based detection. Besides events monitoring and analysis, IDS typically records information about the events, notifies the administrator about important events through warnings and alarms, and generates appropriate reports.

Research and development of intrusion detection systems for SCADA networks has gained a strong momentum in the past decade. Stringent requirements for real-time operation and data integrity, regular traffic patterns and a limited set of telecommunication protocols stipulate the design and implementation of dedicated, sophisticated intrusion detection systems.

This article provides a comprehensive comparison and evaluation of recent research related to SCADA-specific network-based intrusion detection systems, as well as identification of research gaps and recommendations for future research.

The rest of the article is organized as follows. In Section II, we explain factors that affect the design of SCADA-specific IDS, i.e., the hierarchical SCADA architecture, traffic properties and specific protocols, and cyber vulnerabilities and attacks. Section III surveys the related work and explains our motivation for this study. In Section IV, we explain the review methodology in terms of papers selection and IDS evaluation methodology. Section V contains a brief description of selected solutions. Section VI presents evaluation and comparison of solutions in terms of general features, test environments and performance evaluation. This section ends with the concluding evaluation of surveyed solutions. Section VII summarizes the most important results of the analysis and identifies directions for further research in the area. Section VIII concludes the article.

II. FACTORS THAT AFFECT THE DESIGN OF SCADA-SPECIFIC IDS

Although the incorporation of the Internet technologies in industrial networking has reduced the boundary between SCADA and enterprise networks, basically, they still have different requirements, which naturally cause differences in network design as well as security objectives and solutions [13]. In general, a secure enterprise IT system should provide the following, by order of priority: confidentiality, integrity and availability. In SCADA systems, the order of priorities is reversed, and availability is the most important requirement. While the primary objective in the Internet is to protect vital central servers, in process control an edge device may be of the same importance like a central database server. The ultimate goal is to achieve required performance of a real-time system, operating on the 24/7 basis under conditions in which regular behavior coexists with system failures, environmental conditions, human errors, and cyber attacks. The three groups of factors that affect design of SCADA-specific IDSs are hierarchical architecture, network traffic properties, and cyber vulnerabilities and attacks.

A. HIERARCHICAL SCADA ARCHITECTURE

Industrial control networks are characterized by deep and functionally separated hierarchies with different protocols and physical standards [13], [14]. Figure 1 represents hierarchical architecture of a SCADA system with common components and configuration [15]–[17].

The lowest level 0 represents physical devices that interact directly with industrial hardware, interconnected via fieldbus. Controllers at level 1 process signals from field devices and generate appropriate commands for these devices. They include remote terminal units (RTUs), programmable logic controllers (PLCs) and intelligent electronic devices (IEDs) that perform local control of actuators and sensor monitoring. Processing results are forwarded to control center at level 2 for further analysis and response control.

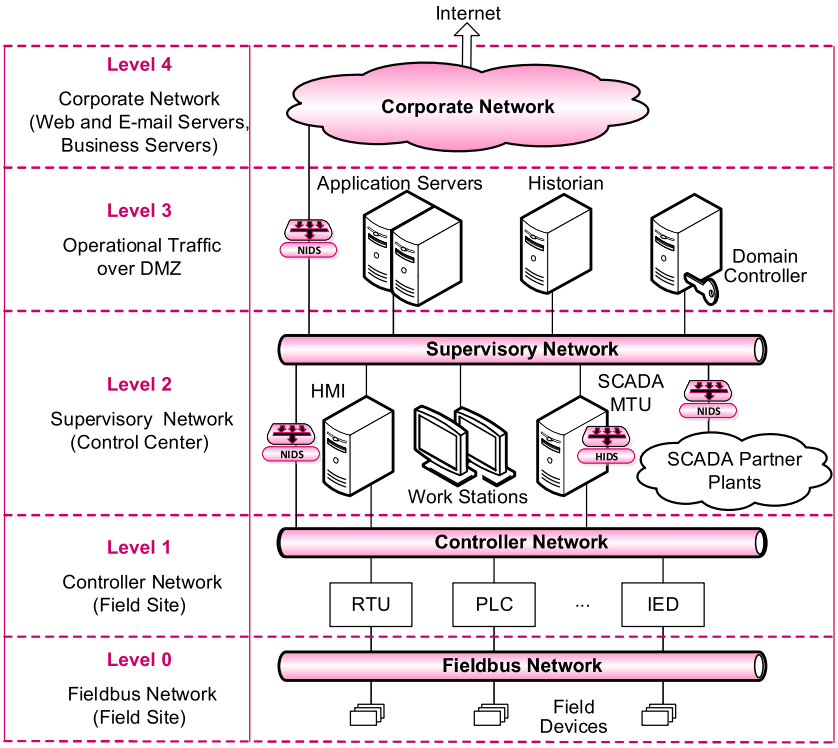

FIGURE 1. General layout of a SCADA system.

Supervisory network connects SCADA server (master terminal unit, MTU), historian server, engineering work stations, human machine interface (HMI) server and consoles, as well as communication devices. Control center collects and analyzes information from field sites, presents them on the HMI consoles, and generates appropriate actions. Control center is also responsible for general alarms, analysis of trends and generating the reports. Communication subsystem connects control center with field sites and allows operators remote access to field sites for diagnostic and failures repairing purposes. Level 3 represents demilitarized zone (DMZ), where application servers, historian server and domain controller are located. Level 4 corresponds to the corporate IT network, which is connected to the Internet.

B. TRAFFIC PROPERTIES IN SCADA NETWORKS

SCADA networks are characterized by regular traffic patterns and a limited set of telecommunication protocols. The number of connections is mainly permanent, while connectivity of particular nodes depends on their functions in the network. Such features are inherently suitable for development and implementation of anomaly-based intrusion detection techniques. We further address traffic features, quality of service (QoS) requirements, updates and order of events, protocols, and addressing principles.

1) TRAFFIC FEATURES

SCADA network traffic is characterized by throughput stability, periodic patterns, clear statistics of packet size, predictable flow direction, and expected connection lifetime [13], [18].

Throughput stability is one of the main indicators of regular operation. Significant increase of throughput points to the events that cause high traffic intensity, e.g., some forms of cyber attacks, scanning, failures, or operating errors. Periodic traffic prevails in SCADA networks due to transmission of data samples in regular intervals, that is, a fixed number of packets being transmitted at fixed intervals. The sampling period depends on the device and control requirements. Aperiodic events may occur at any time due to change of state or alarm conditions, but may also indicate some forms of attacks.

Most of the systems at the fieldbus level send packets without additional buffering due to severe delay requirements. Thus, a clear packet size statistics is created, and the average packet size represents a good indicator of regular behavior or anomaly.

Flow direction indicates which system initiates the connection, and it is known in a typical protocol operation. After connection establishment, data amount sent from one system to another is predictable with large probability, especially for the known service. Deviation from such a behavior usually indicates an anomaly. In addition, Transmission Control Protocol (TCP) connections have expected lifetimes, and connection duration usually shows a very small variance.

2) QoS REQUIREMENTS

Packet transfer delay and packet inter-arrival times from all network nodes are meaningful data for intrusion detection in

SCADA networks. This is a consequence of real-time operating requirements, especially at the fieldbus level. Packet timing and the associated statistics are regular, but different from the traffic generated by typical applications in enterprise networks and providers' networks. Response time should typically be less than the sample time of collected data. The most stringent requirements are at the fieldbus and controller network levels. Most of applications pose response time requirements in the range of 250 μs to 1 ms whereas less stringent processes require response times in the range of 1 ms to 10 ms [13]. Higher levels tend to have progressively lower delay requirements, typically up to 1 second.

3) UPDATES AND ORDER OF EVENTS

In addition to timely delivery of all relevant data, updates should be performed on a regular basis, because the data is only valid in its assigned time period [13], [18]. The order of updating is particularly important for sensor data concerning monitoring of the same process or correlated processes. The order of data arrival to the control center plays an important role in presentation of process dynamics and influences decision making, by either a control algorithm (software) or a human operator who monitors the industrial process.

4) PROTOCOLS AND THEIR CONFIGURATION

Each SCADA network implements a precisely defined set of protocols. The appearance of new protocols indicates serious changes in the network. Protocols are typically configured statically, in a way that guarantees the best network performance. Monitoring of protocols configuration parameters enables detection of poorly configured services as well as malicious activities. Packet payloads, originating from SCADA applications, are also precisely defined. Changes of payload format may indicate anomalies in system's behavior. Similarly, the observed anomalies in payload contents may point to erroneous system configuration or the presence of malicious activities in the network.

5) ADDRESSING

SCADA systems that apply the Internet protocol suite implement appropriate addressing mechanisms at different protocol layers. Many networks use static allocation of IP addresses and transport layer port numbers; hence it is expected that sockets (pairs ''IP address: Port'') remain constant. Mapping of Medium Access Control (MAC) addresses to IP addresses can be used to detect changes in hardware components. Although MAC addresses can be forged, they are still useful to detect impersonation. They also help administrator to keep evidence of legitimate system hardware.

C. CYBER VULNERABILITIES AND ATTACKS

According to [2] vulnerabilities of industrial control systems are broadly classified into following groups: policy and procedure, architecture and design, configuration and maintenance, physical, software development, and communication/network. Important factors that may affect SCADA vulnerabilities include human errors, resource limitations of physical devices, unsecure legacy systems and proprietary protocols, equipment failures and other accidents caused by negligence, and natural disasters.

Attacks on SCADA systems can be launched by external sources, e.g., terrorists, hackers, competitors, industrial espionage, or by internal sources, such as disgruntled employees, third-party vendors, or site engineers. Different taxonomies of attacks on ICS/SCADA systems can be found in the literature [11].

Control actions in SCADA systems are performed on the basis of the data received from RTUs. If an intruder wants to jeopardize process control, the attack will be focused on modifying control data or completely blocking the data transfer. According to that, four classes of ICS attacks are recognized in [19], namely reconnaissance, response and measurement injection, command injection and denial of service (DoS).

Reconnaissance attacks aim to discover information about a network and to identify the equipment characteristics. Response injection attacks insert false responses into a control system, and subsequently cause control algorithms to make incorrect decisions. Therefore, it is important to protect the integrity of the sensor measurements from the physical process. Command injection attacks insert false control commands into a control system. This may happen due to human intervention, which results in false control action, or by injecting false commands that cause overwriting RTU programs and remote terminal register settings. DoS attacks tend to disrupt the communication link between the RTU and MTU or HMI, which makes process monitoring and control impossible. There are different forms of DoS attacks, and they can be launched at any layer of the protocol stack causing physical jamming, disconnection, and malfunction of network protocols. Their common property is the aim to cause the unresponsiveness of targeted hardware or software. The attacks are more dangerous and harder to prevent if a group of attackers coordinate in DoS (distributed DoS, DDoS) [20]. In DDoS attacks, each individual attacker can generate traffic similar to the legitimate one, but the attack strength is increasing by using multiple coordinated sources. This property makes intrusion detection more difficult.

Maglaras *et al.* [21] classify ICS cyber attacks into the following four categories: key-based attacks, data-based attacks, impersonation-based attacks and physical-based attacks. Key-based attacks try to capture secret keys that are used by consumers and suppliers for registration and authentication. Data-based attacks include a number of attacks that try to change data without authorization, e.g., modification attack, data integrity attack, repudiation attack, etc. Impersonation-based attacks try to impersonate a trusted individual or company to gain access to sensitive data. Examples of such attacks are man-in-the-middle (MITM) attack, eavesdropping attack, replay attack and redirection attack. Physical-based attacks manipulate the physical properties of devices to cause sensors and embedded devices

to malfunction, e.g., differential attack, malware attack, collusion attack, and inference attack.

Classification presented in [10] starts from the fact that attacks can occur at all levels of SCADA network and assumes attacks on hardware, attacks on software, and attacks on network connections.

III. MOTIVATION AND RELATED WORK

The motivation for this work came from the fact that only a few studies dealt with systematic and comprehensive review of research work in SCADA-specific intrusion detection systems.

Zhu and Sastry made the first systematic and thorough effort in investigating and assessing the SCADA-specific intrusion detection techniques and systems [22]. They explained the necessity of designing SCADA-specific IDS and presented a comprehensive comparative analysis of the nine prototypes designed in the period 2004–2008. Results of the analysis pointed out the most critical shortcomings concerning the lack of well-considered threat models, inadequate addressing of intrusion detection accuracy and the need to develop IDSs based on the knowledge of how SCADA systems operate in practice.

Mitchell and Chen surveyed intrusion detection techniques for cyber physical systems such as distributed control systems, networked control systems, sensor actuator networks, wireless industrial sensor networks, and SCADA systems [23]. In particular, they proposed a classification of the existing IDS techniques for cyber physical systems, discussed their advantages and drawbacks, and presented a brief overview and comparison of 28 systems, developed in the period 2003–2013, and comprising aerospace, automotive, medical, and SCADA IDSs. They identified a number of open issues in the research area such as the lack of clear definition of IDS performance metrics, the need to focus audits on application layer data, the lack of proper attacker models, the need to further develop and validate certain detection techniques, the need to develop federated IDSs, etc.

Garitano *et al.* provided an overview of anomaly-based detection systems for SCADA networks [24]. Their study encompassed nine representative IDS solutions, most of them developed in the period 2005–2010. The evaluation indicated that although a variety of anomaly detection techniques were proposed, almost all of them used simulated traffic for learning and testing purposes, and were not verified in realistic environments.

Nazir and Patel provided a broad overview of techniques and tools needed to find out SCADA system vulnerabilities, including network-based intrusion detection systems [16]. They identified signature-based and anomaly-based approaches and briefly described eight NIDS solutions, developed in the period 2008–2014. Rubio *et al.* followed a similar approach [25]. They reviewed the threats that affect industrial control systems and analyzed the state, evolution and applicability of both academic and industrial intrusion detection mechanisms. Ghosh and Sampalli distinguished rule-based and anomaly-based approaches, and provided a brief overview of six techniques, developed in the period 2013–2017 [10].

Hu *et al.* provided a survey of intrusion detection on industrial control systems [26]. They proposed IDS taxonomy based on the following techniques: protocol analysis, traffic mining, and control process analysis, and analyzed the advantages and disadvantages of different IDS categories. Similarly, Murray *et al.* outlined a number of publicly disclosed SCADA vulnerabilities, in addition to approaches for detecting attacks in operational networks [27].

A comprehensive compilation of intrusion detection and prevention systems intended to secure smart grids is presented in [28]. The authors provided analysis of 37 cases, published in the period 2010–2018, and classified them according to applicability domain, i.e., the entire ecosystem, advanced metering infrastructure, SCADA, substations and synchrophasors.

Our work differs from previous approaches in the following aspects.

First, we selected research papers according to the recommendations from [29] and encompassed the five-year period 2015–2019. To the best of our knowledge, none of previous review papers on SCADA intrusion detection systems followed similar methodology.

Second, in order to conduct meaningful analysis we limited the set of surveyed papers to obtain fairly comparable works. Hence, we focus on original research papers dealing with network-based SCADA-specific IDSs.

Third, we try to establish correspondence between different approaches to the classification of intrusion detection methodologies used in SCADA networks. Unlike [22], [23] and [26], we do not attempt to propose SCADA-specific classification trees, but rather follow a well-established IDS classification for general-purpose IT systems.

Fourth, we pay special attention to comprehensive analysis of test environments as well as IDS performance evaluation in terms of accuracy, timeliness, response to incidents and efficiency.

Finally, bearing in mind evolution towards fourth generation SCADA systems, we assess applicability of each solution to the FIN environment.

The main objectives of this work are:

- To propose a systematic and comprehensive evaluation methodology for SCADA-specific IDSs;
- To perform a critical evaluation of recent IDS solutions, and to assess their strengths, weaknesses, implementation maturity, as well as suitability to FIN environment;
- To identify gaps in current research and to propose relevant research priorities for future work in the area.

IV. REVIEW METHODOLOGY

A. SELECTION OF PAPERS

Selection of papers was performed following the recommendations from [29]. The initial set of research papers was created by searching the three well known databases,

TABLE 1. Classifications of intrusion detection methodologies: A comparative overview.

General classification [30]	Adopted classification [12], [31], [32]		Classification from [22]	Classification from [23]	Alternative terms used in the literature
Blacklist	Signature-based		Knowledge-based	Knowledge-based	Rule-based Misuse detection
Whitelist	Anomaly-based	Statistical-based	Behavior-based	Behavior-based	Behavior-based
		Knowledge-based	Knowledge-based	Behavior-based	Rule-based (expert systems)
		Machine learning-based	Behavior-based	Behavior-based	Behavior-based
	Specification-based		Hybrid	Behavior-specification-based	Stateful protocol analysis Deep packet inspection Model-based

namely the IEEE Xplore, SCOPUS and Web of Science (WoS), with the following keywords: ''SCADA'' and ''intrusion detection''. The search was performed in January 2020 and we considered the 5-year period from 2015 to 2019. As result, we obtained 310 papers in total: 71 papers from IEEE Xplore, 131 papers from SCOPUS and 105 papers from WoS. After manual inspection and exclusion of the replicated papers, the resulting set consisted of 168 papers.

Further, we selected candidate papers based on abstract and title, focusing on papers that contained original proposals for SCADA-specific NIDS solutions. The remaining set of papers consisted of 86 papers.

After inspecting the full versions of candidate papers, the objective was to obtain the set of unique and comparable solutions. Thus, we eliminated similar papers by the same authors or different authors but describing the results of the same projects, as well as papers that did not contain IDS evaluation results. The final selection resulted in 26 papers.

B. IDS EVALUATION METHODOLOGY

We propose the SCADA IDS evaluation methodology that identifies general features of the proposed solution, analyzes system's performance, and assesses strengths and weaknesses of the system. The evaluation encompasses the following aspects: (1) detection methodology; (2) protected protocols; (3) implementation tools; (4) test environment and (5) performance evaluation. Overall assessment is performed based on the previous five evaluation properties.

1) DETECTION METHODOLOGY

From the collected material, we have found significant diversity in terminology regarding classification of intrusion detection methodologies in SCADA systems. To help resolve the confusion in terminology, we will further explain several approaches to the classification and try to establish correspondence between those approaches and the one that we adopt in this study. Table 1 summarizes classifications of intrusion detection methodologies.

The most general classification of intrusion detection methodologies is to blacklist and whitelist approaches [30]. Blacklist approaches assume that all processes/requests are approved unless they are explicitly mentioned on the blacklist. Whitelist approaches profile ''normal behavior'' so that deviations can be reported.

Attempts were made to define classification trees for IDSs in SCADA and other control systems, in order to facilitate systematization of the existing detection techniques. Both approaches [22], [23] distinguish knowledge-based and behavioral-based techniques as the two main categories of real-time intrusion detection techniques. Knowledge-based methods rely on primary evidence such as semantic definitions, predefined policies, model of legitimate data flow, and abstraction of known illegal patterns. Behavioral-based methods need secondary evidence to make contextual analysis.

After literature review, we believe that there is no need for dedicated classification of SCADA-specific intrusion detection methods, because they can all be categorized according to the well-established classification proposed in [12], [31] and [32]. The benefit of such an approach is confusion avoidance and facilitation of reading and understanding for both SCADA experts and IT experts. We adopt that basic detection methodologies comprise signature-based detection, anomaly-based detection and specification-based detection.

Signature-based detection is a blacklist approach, which encompasses techniques that compare monitored events with patterns that correspond to known threats (signatures) in order to identify possible incidents. Signature-based methods are very efficient in detection of known threats, but completely inefficient when new or unknown threats or modified attacks appear. Besides, there are issues related to maintaining a signature database due to need to continuously update signatures.

Anomaly-based detection is a whitelist approach, which includes techniques that compare monitored events with the list of activities, which were predefined as normal to

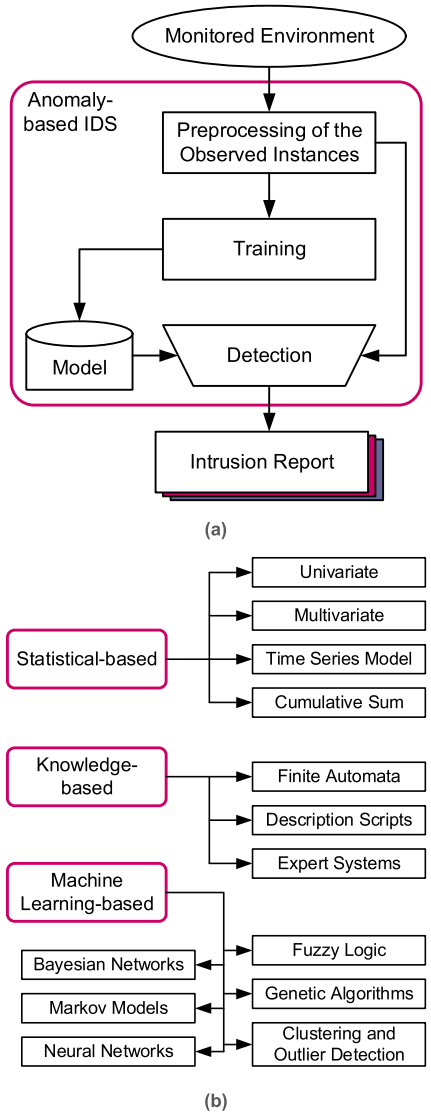

FIGURE 2. **Anomaly-based IDS: (*a*) functional architecture; (*b*) classification tree (adapted from [31] and [32]).**

identify significant deviations. The general advantage of anomaly-based techniques refers to efficient detection of unknown threats. However, erroneous inclusion of malicious activities in profiles is a typical problem of these techniques. Another issue concerns accuracy of generated profiles and appears because of complex activities in the network.

The generic functional architecture of anomaly-based IDS is depicted in Fig. 2(a). In the preprocessing phase, the observed instances are represented in a predefined form. IDS creates static or dynamic models (profiles) representing normal behavior of users, hosts, network connections, and applications. During a training period the initial profile is generated which can be done in different ways, depending on the IDS type.

According to the nature of processing involved in the behavioral model, anomaly-based techniques can be classified into three main categories: statistical-based, knowledge-based, and machine learning-based, as illustrated in Fig. 2(b).

Statistical-based techniques use statistical properties and tests to determine whether the observed behavior deviates significantly from the expected behavior. They include a number of techniques based on univariate, multivariate, time-series models and cumulative sums (CUSUM). The advantages of statistical-based techniques include the ability to learn the expected behavior of the system (without prior knowledge about its normal activity) and the ability to provide accurate long-term detection of malicious activities. Their main disadvantage refers to possibility that the attacker trains the system in such a way that the malicious traffic is considered as normal.

Knowledge-based techniques try to capture the claimed behavior from the available system data. They involve techniques based on finite automata, description languages and expert systems. The most widespread are expert systems, which classify the observed data according to a set of rules. This set of rules is obtained from different attributes and classes that are identified from the training data. The advantages of knowledge-based techniques include their robustness and flexibility. Their main disadvantage refers to difficult and time-consuming task of acquiring high-quality knowledge.

Machine learning-based techniques establish an explicit or implicit model that allows classification of analyzed patterns. The principles and applicability of machine learning techniques are similar to the statistical techniques. However, machine learning-based techniques enable NIDS to change its reaction as it acquires new information. Well-known machine learning-based techniques are Bayesian networks, Markov models, neural networks, fuzzy logic, genetic algorithms, and clustering and outlier detection algorithms. There are two main types of machine learning tasks: supervised and unsupervised. Supervised learning uses a ground truth, which assumes prior knowledge of the output values for given data samples. Therefore, the goal is to learn a function that maps an input to an output based on example input-output pairs. Unsupervised learning does not have labeled outputs; hence, its task is to infer a function that describes the structure of unlabeled data (i.e., data that are not classified). The advantages of machine learning-based techniques include flexibility, adaptability as well as the ability to capture interdependencies of the observed instances. Their main disadvantage is high resource consumption.

Data mining and Knowledge Discovery Database (KDD) are widely used to discover patterns and correlations in large datasets. Many different IDS processing approaches use the term ‘‘data mining’’, as a generic concept related to preprocessing of the observed instances. In such a context, almost every machine learning scheme includes some data mining technique.

Specification-based detection is a whitelist approach, which includes techniques that compare monitored events with predefined profiles. These profiles are generated from definitions of protocol activities for each state of the protocol machine. They use universal profiles defined by standardization bodies and/or software manufacturers. These methods

can identify irregular sequences of messages like replication of the same command, or issuing of a command, which is not preceded by a command predicted in protocol specification. Their main disadvantage refers to the intensive usage of processor and memory resources due to recording the states of a large number of simultaneous sessions and complex analysis of those states. Another disadvantage is the need to develop a dedicated profile for each protocol.

2) PROTECTED PROTOCOLS

Among a range of standard and vendor-specific SCADA communication protocols the most widespread are Modbus, IEC 60870-5 series, Distributed Network Protocol (DNP3), IEC 61850 series, and EtherNet/IP. The majority of protocols are created or extended to operate over TCP/IP networks. In addition, most of the current fieldbus protocols are Ethernet-based. A comprehensive survey and taxonomy of SCADA protocols can be found in the literature [13], [33], [34].

Modbus is a proprietary protocol, based on the client-server paradigm. Modbus client is a master device, which generates and sends requests to the Modbus server (slave). Modbus request contains function code, which identifies the required service and the list of associated parameters with detailed description of the service request. Modbus client functions encompass reading of discrete inputs/coils (at the bit level) or registers (16-bit words), writing of appropriate output values and diagnostic functions for the server. After receiving and processing the request, the server generates a response and sends it to the client. Modbus messages are encapsulated either to the serial protocol at the data link layer (Modbus RTU/ASCII) or to the TCP/IP stack (Modbus TCP).

The IEC 60870-5-104 protocol enables TCP/IP based network access for the IEC 60870-5-101 protocol, which is used for basic telecontrol tasks between control centers and substations. This protocol is widely applied to SCADA systems in Europe, China and some other non-US countries.

The DNP3 is intended for communication between substation computers, RTUs, IEDs and master stations. It is also applied in water treatment systems. The protocol is widely used in North America in place of IEC 60870-5 series. The hierarchical protocol architecture consists of the application layer, the transport layer, and the data link layer. The protocol supports three communication modes between a master and remote stations: peer-to-peer transaction, broadcast transaction, and unsolicited response. In 2000, the DNP Technical Committee defined a specification for carrying DNP3 over TCP/IP stack.

IEC 61850 is a global standard that is applicable to electrical substation automation systems, and defines the communication between IEDs in the substation and the related system requirements. Automated system consists of IEDs that are interconnected to perform automation, protection, monitoring, metering, and control of substations. IEC 61850 defines requirements for interoperability between multi-vendor IEDs, and establishes standards for object models, device behavior, naming conventions, and services. The abstract data models defined in IEC 61850 can be mapped to a number of protocols including Manufacturing Message Specification (MMS), Generic Object Oriented Substation Event (GOOSE), Sampled Measured Values (SMV), etc. These protocols can run over TCP/IP networks or substation LANs. Security mechanisms for IEC 61850 are described in the IEC 62351 series of standards, and include requirements for intrusion detection.

EtherNet/IP is a set of industrial network protocols that allows using the Common Industrial Protocol (CIP) over standard Ethernet and TCP/IP. EtherNet/IP is used in Ethernet industrial modules from different vendors. The CIP is a media independent application layer protocol designed to support automation applications. It encompasses a set of communication services such as control, safety, synchronization, motion, configuration and information. Transfer of automation data between two devices relies on request/reply paradigm. CIP is object-oriented protocol, which means that every network device contains a series of objects. Each object has well-defined attributes (data), services (commands) and behaviors (reactions to events). Device profiles specify different sets of CIP objects that must be implemented, configuration options and data formats for different device types.

3) IMPLEMENTATION TOOLS

There are a number of open-source network intrusion detection systems that are used for development of SCADA-specific solutions [8]. The most widespread systems are Snort, Suricata and Bro.

Snort is the most widely deployed IDS worldwide. It relies on a relatively simple language for specification of misuses and attack signatures. In most cases, signatures are encoded by a single Snort rule, which defines connection endpoints and packet attributes.

Suricata is a newer network threat detection engine that is capable of real-time intrusion detection, inline intrusion prevention, network security monitoring and offline processing of captured packets. It is backward compatible with Snort rule sets. Suricata is designed as a multi-threaded system, which can take advantage of multi-core processors. Therefore, it can examine large volumes of traffic without reducing the number of rules.

Bro (known as Zeek, since January 2020) is slightly alternative compared to Snort and Suricata. It is a passive, open-source network traffic analyzer, which is organized into two major components: event engine and policy script interpreter. The event engine reduces the incoming packet stream into a series of higher-level events, while policy script interpreter executes a set of event handlers written in a custom scripting language.

General-purpose programming languages, including C, C++, C#, Perl, Python and Java, are also used to develop SCADA-specific IDS applications. Proprietary software platforms are being developed using those languages. Besides, some of general-purpose open-source tools

Confusion matrix			Derived evaluation metrics
Actual	Predicted		1. $FPR = FP/(FP + TN)$
	Attack	Normal	2. $FNR = FN/(TP + FN)$
Attack	*TP*	*FN*	3. $DR = TPR = Recall = TP/(TP + FN) = 1 - FNR$
Normal	*FP*	*TN*	4. $TNR = TN/(FP + TN) = 1 - FPR$
			5. $Accuracy = (TP + TN)/(TP + TN + FP + FN)$
			6. $Precision = TP/(TP + FP)$
			7. $F\text{-}measure = 2/(1/Precision + 1/Recall) = 2TP/(2TP + FP + FN)$

FIGURE 3. Confusion matrix and derived evaluation metrics.

(WEKA, TensorFlow, LIBSVM, Anaconda) are used to build SCADA-specific solutions.

4) TEST ENVIRONMENT

Pen-testing activities, which are typical for non-industrial environments, are unacceptable for SCADA and other industrial control systems, because they must maintain the operational continuity. Hence, new security solutions need reliable test environments that meet the requirements regarding fidelity, repeatability, measurement accuracy and safe execution [35]. In the broad sense, test environment encompasses testbed, datasets and simulated attacks.

Testbed is a platform for conducting exhaustive, transparent, and replicable testing of algorithms, methods, prototypes, etc. It may include software, hardware and networking components. According to [16] and [35], SCADA security testbed can be implemented in one of the following ways:

- **Cyber physical system (CPS) testbed:** uses real hardware and software to pursue lines of experimentation and exploration.
- **Emulation-based testbed:** may use different combinations of physical devices and software to simulate the control network and the physical process.
- **Software simulation testbed:** can be simple simulation-based (assumes a single software simulation package for testing purposes) or federated simulation-based (may have several interacting simulations such as plant, network, etc.).
- **Virtualization-based testbed:** uses virtualization technology to build a low-cost, high-fidelity, reusable, and easy-to-maintain testbed.

Capturing and preprocessing SCADA network traffic is needed before intrusion detection. Due to confidentiality of real SCADA network data, researchers often use synthetic datasets or experimental datasets obtained from CPS testbeds. The analysis provided in [36] indicates that the existing models used to describe the Internet traffic cannot be easily applied to SCADA traffic for several reasons, such as different diurnal patterns, absence of self-similar correlations in the time series, and different distribution of connection sizes. Hence, widespread public datasets such as KDD99 and UNSW-NB15 might not be representative for SCADA traffic. On the other side, there are several publicly available SCADA-specific datasets intended for power, gas pipeline and water storage tank systems [37].

Simulation of attacks is the prevalent method in test scenarios. A comprehensive review of attack simulation tools can be found in [16]. Typical simulated attacks on SCADA systems include malware attacks, network attacks, communication protocol attacks, DoS/MITM, false data injection, false sequential logic attacks, and data integrity attacks. Some of them are included in the aforementioned public datasets. An alternative to simulated attacks is to obtain real attacker data from honeypots, but this might increase risk of providing information to potential intruders.

5) PERFORMANCE EVALUATION

To the best of our knowledge, there are no dedicated performance evaluation techniques for SCADA-specific IDSs. Instead, general techniques are used that are developed for IDS evaluation in public and enterprise IT networks. A comprehensive survey of IDS performance evaluation techniques and the associated metrics can be found in [38]. We focus on the following criteria: detection accuracy, timeliness, response to incidents and efficiency.

a: DETECTION ACCURACY

Detection accuracy (also known as classification accuracy or effectiveness) represents the ability of the system to distinguish between intrusive and non-intrusive activities. It is represented by a set of measures that determine how correctly an IDS works. Further, we focus on the set of most common metrics.

Confusion matrix represents true and false classification results, as indicated in Fig. 3. The variables of confusion matrix are:

- True positive (*TP*) – number of successfully detected malicious activities;
- True negative (*TN*) – number of normal activities that are successfully labeled as non-intrusive;
- False negative (*FN*) – number of malicious activities that are not detected, but considered as normal;
- False positive (*FP*) or false alarm (*FA*) – number of normal activities that are detected as malicious.

Figure 3 represents different evaluation metrics that are derived as functions of the confusion matrix variables. Those metrics are as follows:

1. False positive rate (*FPR*) measures the ratio between the number of normal instances detected as attacks and the total number of normal activities.
2. False negative rate (*FNR*) measures the ratio between number of malicious activities that are not detected and the total number of malicious activities.
3. Detection rate (*DR*), also known as True Positive Rate (*TPR*) or *Recall*, measures the fraction of anomalies that are successfully identified.
4. True Negative Rate (*TNR*) measures the ratio between the number of normal instances detected as non-intrusive and the total number of normal activities.
5. *Accuracy* measures the fraction of instances that are correctly classified.
6. *Precision* denotes the probability that a detected anomaly is correct.
7. *F-measure* represents the weighted harmonic mean of *Precision* and *Recall*.

In addition, Receiver Operating Characteristics (*ROC*) curves can be used to assess performance of a classification model, by visualizing the relation between *TPR* and *FPR* metrics. Area under the *ROC* curve (known as *AUC*) is a performance measurement for classification problem at various thresholds settings. *AUC* takes values between 0 and 1; higher *AUC* denotes better IDS performance.

b: TIMELINESS

Timeliness refers to the system's ability to perform its analysis as quickly as possible. The objective is to enable prompt response to incident to minimize the damage within a specific time period. Timeliness is usually estimated concerning the time needed to process the unit of analysis, i.e., packet, group of packets, traffic flow, communication session or dataset instance. Detection latency represents the time between the attack detection and the actual moment of the attack. Total delay represents the time between the response of the system and the actual moment of the attack.

c: RESPONSE TO INCIDENTS

In general, response to incidents can be passive and active. Passive response assumes alert generation after detection of an incident. Active response encompasses prevention capabilities and/or integration with the other security mechanisms. Intrusion prevention system (IPS) is a tool that generates response to detected threats by attempt to preclude their realization [12], [31].

Both IDS and IPS are integral parts of the overall security management system. Efficient solution typically assumes combination of different technologies. For example, security information and event management (SIEM) software imports information from various security-related logs and performs correlation of the corresponding events. Some SIEM products can initiate preventive responses to certain events.

d: EFFICIENCY

Efficiency refers to the resources needed to be allocated to the system including CPU and memory usage. IDS can collect and analyze data continually as the data is acquired or in blocks, after an event has occurred. Continuous mode is also known as real-time processing, and provides the opportunity for administrator to take action while the intrusion is in progress. A very important evaluation criterion for continuous mode is the system's ability to process traffic on a high speed network with minimum packet loss. Finally, the performance of any NIDS depends on its configuration, monitored network properties, and the system's placement in that network.

V. A BRIEF DESCRIPTION OF SELECTED SOLUTIONS

Table 2 contains general information, concerning authors, year of publishing, technique title, application domain, country of the first author and number of citations. Number of citations refers to Google Scholar Citation Index as on 19 March 2020.

A. SIGNATURE-BASED TECHNIQUES

In [50], Wong *et al.* present their contribution to Suricata IDS by including the support for detecting cyber attacks in EtherNet/IP-based SCADA. Design and implementation of the EtherNet/IP support in Suricata encompasses the following phases: (1) definition of rules for the protocol; (2) parsing the rules and storing them in the appropriate matching data structure and (3) developing an EtherNet/IP packet parser. Two design solutions are considered and implemented – one based on the examination of individual packets and the other based on examining packet streams. The latter solution is adopted and integrated into Suricata Release 3.2beta1 in October 2016.

B. STATISTICAL-BASED TECHNIQUES

Kwon *et al.* propose an intrusion detection technique for IEC 61850 substations, which focuses on GOOSE and MMS protocols, taking into account specification-based metrics and multivariate analysis of network features [40]. To detect malicious traffic, the proposed technique uses static and dynamic features. With static features it verifies the syntax correctness of the protocol. Dynamic features depend on the network environment. Anomaly detection represents a function of the three weighted input parameters, i.e., network metric, GOOSE metric and MMS metric.

Time series models were considered in [47] and [57]. The model described in [47] is based on the assumption that the majority of traffic in SCADA networks follows periodic patterns. Therefore, the main module, called Periodicity Analyzer, can detect anomalies in traffic periodicity. To detect the response injection attacks, an auxiliary module is introduced, called Telemetry Analyzer. The model

TABLE 2. Overview of selected papers: General information.

No.	Authors, year & reference	Technique title	Application domain	Country	Citations
1.	Erez & Wool, 2015 [39]	Domain-aware anomaly detection system for Modbus TCP	Power supply system	Israel	77
2.	Kwon et al., 2015 [40]	Behavior-based intrusion detection technique for smart grid infrastructure	IEC 61850 substation	Korea	24
3.	Al Balushi et al., 2016 [41]	Ontology based SCADA intrusion detection framework	Some SCADA systems	UK	1
4.	Almalawi et al., 2016 [42]	Data-driven clustering SCADA intrusion detection technique	Water distribution system	Saudi Arabia	51
5.	Cruz et al., 2016 [43]	SCADA cybersecurity detection framework	Electrical distribution grid	Portugal	66
6.	Da Silva et al., 2016 [44]	One-class IDS for SDN-based SCADA systems	Power grid	Brazil	14
7.	Ghaeini & Tippen-hauer, 2016 [45]	Hierarchical monitoring IDS for ICS	Water treatment system	Singapore	36
8.	Udd et al., 2016 [46]	Bro based intrusion detection in a SCADA system	Electrical substation	Sweden	34
9.	Zhang et al., 2016 [47]	Traffic periodicity and telemetry based IDS for SCADA systems	Some SCADA systems	China	14
10.	Feng et al., 2017 [48]	Multi-level anomaly detection in ICS	Some SCADA systems	UK	59
11.	Wan et al., 2017 [49]	One-class classification anomaly detection in SCADA systems	Some SCADA systems	China	27
12.	Wong et al., 2017 [50]	Enhanced Suricata based IDS for SCADA networks	Some SCADA systems	Canada	16
13.	Y. Yang et al., 2017 [51]	Multidimensional IDS for IEC 61850-based SCADA networks	IEC 61850 substation	China	52
14.	Adepu & Mathur, 2018 [52]	Distributed attack detection in a water treatment plant	Water treatment system	Singapore	14
15.	Ghazi & Doustmoham-madi, 2018 [53]	SCADA intrusion detection based on Petri Net	Some SCADA systems	Iran	3
16.	Hijazi & Flaus, 2019 [54]	A deep learning approach for IDS in industry network	Some SCADA systems	France	3
17.	Lin et al., 2018 [55]	Detection of control-related attacks in power grids	Power grid	USA	61
18.	Myers et al. 2018 [56]	Anomaly detection for ICS using process mining	Some SCADA systems	Australia	10
19.	Wang and Feng, 2018 [57]	SCADA intrusion detection using graphical features	Power system	China	0
20.	Wressnegger et al., 2018 [58]	ZOE: Content-based anomaly detection for ICS	Some SCADA systems	Germany	7
21.	Benisha & Raja Ratna, 2019 [59]	Detection of bias injection attacks in SCADA systems	Some SCADA systems	India	0
22.	Derhab et al., 2019 [60]	IDS for SDN-enabled industrial IoT security	IoT-based ICS	Saudi Arabia	6

TABLE 2. *(Continued.)* **Overview of selected papers: General information.**

23.	Keshk et al., 2019 [61]	Privacy-preserving based anomaly detection for cyber-physical systems	Power system	Australia	2
24.	Khan et al., 2019 [62]	Hybrid-multilevel anomaly prediction for SCADA intrusion detection	Some SCADA systems	China	3
25.	Lai et al., 2019 [63]	Industrial anomaly detection based on convolutional neural network	Some SCADA systems	China	2
26.	H. Yang et al., 2019 [64]	Deep-learning-based SCADA network intrusion detection	Power system	USA	3

presented in [57] relies on graphical features. The time series of IEC 60870-5-104 protocol messages transmission were visualized, using the visual coordinate point or polygon representation. Further, features of the graph were extracted to form a new feature dataset.

C. KNOWLEDGE-BASED TECHNIQUES

Erez and Wool propose a domain-aware anomaly detection system that detects irregular changes in Modbus TCP register values [39]. After manual classification of registers, the procedure for automated classification is activated, based on observing the values recorded during a short time-frame (classification window). Two types of classification algorithms are developed: the single-window classification and the incremental classification. The concept of finite state machines is used to include register class in the previous classification window as well as other relevant information from the entire classification process.

Expert systems were considered in [41], [55] and [58]. The system proposed in [41] uses ontology for extraction of semantic relations between attacks and detection of intrusions. Ontology is used to define the logical relationships between packet and attack instances, cyber attacks and the Modbus TCP communications. A semantic analysis framework that integrates NIDS with analysis of AC power flow capable of estimating consequences of control commands execution is proposed in [55]. The parameters of the power flow analysis algorithm are dynamically adjustable to provide a balance between detection accuracy and latency. A robust anomaly detection method that can be applied in environments with high entropy data is presented in [58]. The method can model the content of unknown binary protocols, by deriving prototype models that are specific to individual types of messages. These protocol models can be used to eliminate irrelevant (noisy) data features in binary protocols.

D. MACHINE LEARNING-BASED TECHNIQUES

1) NEURAL NETWORKS

Hijazi and Flaus present an intrusion detection system based on deep learning with artificial neural network (ANN) [54]. The proposed technique uses multi-layer perceptron algorithm with binary classification, trains high-dimensional Modbus data and labels the data as normal or malicious.

Solutions based on the deep learning with convolutional neural network (CNN) are presented in [63] and [64]. In [63], the original one-dimensional data are mapped using the feature mapping method based on Mahalanobis distance to obtain a two-dimensional matrix that represents CNN input. The approach proposed in [64] uses a CNN to characterize relevant temporal patterns of SCADA traffic and to identify time windows where network attacks are present.

2) CLUSTERING AND OUTLIER DETECTION

Almalawi *et al.* address IDS for detecting SCADA-specific attacks by monitoring the states of process parameters, which indicate the criticality of the underlying system [42]. The system works in two phases: (1) a data-driven clustering technique of process parameters identifies normal and critical states of the target system and activates a criticality scoring mechanism and (2) a detection rule technique extracts a set of proximity-based detection rules that fully represent all identified states; it actually groups normal and critical states into micro-clusters, each of which is used to extract unique proximity-based detection rules to monitor the criticality degree of underlying system.

Da Silva *et al.* propose IDS for software-defined networking (SDN) based SCADA that can be used to monitor small and large-scale systems, and to promptly and accurately manipulate large datasets [44]. It relies on the OpenFlow protocol, to periodically gather information from network devices and to generate samples from gathered statistics. These samples are processed by one-class classification (OCC) algorithms based on support vector machines (SVMs), i.e., one-class SVM (OCSVM) and support vector data description (SVDD).

Wan *et al.* adopt an approach that correlates industrial communication characteristics with the time sequence, and further extracts function control behavior and process data behavior [49]. Further, OCSVM is applied to detect the corresponding anomalies. Besides, reconstruction error based on kernel principal component analysis (RE-KPCA) is used to improve classification performance.

Myers *et al.* present a process mining anomaly detection method that uses ICS data logs and the conformance checking analysis technique [56]. A conformance checking analysis

determines the extent to which real behaviors (captured in the logs) matches the expected behaviors (captured in the process model).

Benisha and Raja Ratna propose intrusion detection and prevention system for bias injection attacks [59]. First, modified grey wolf optimization is used to extract the features needed for classification. Second, entropy-based extreme learning machine extracts the features and detects the malicious data with corresponding intrusion time, file location, and date. Finally, the data are encrypted to prevent further attacks.

Derhab *et al.* present a security architecture that integrates the blockchain and the SDN technologies [60]. The intrusion detection method includes the random subspace learning and k-nearest neighbors (KNN) algorithms to defend against the forged commands. In addition, a blockchain-based integrity checking system is applied to protect the OpenFlow protocol in SDN-enabled industrial IoT systems.

Khan *et al.* adopt an automated multi-level intrusion detection approach that combines Bloom filter and KNN [62]. After preprocessing and dimensionality reduction of the observed data, Bloom filter is applied at the first level to create a signature database for legitimate network packets. At the second level, KNN-based algorithm is activated to detect potential zero day attacks.

E. SPECIFICATION-BASED TECHNIQUES

Ghaeini and Tippenhauer propose intrusion detection framework called HAMIDS (HierArchical Monitoring Intrusion Detection System) [45]. It is implemented on Bro with support for EtherNet/IP and CIP traffic parsing, as Bro extensions. Its main features encompass: (1) the capability to detect anomalies on both the fieldbus and controller network levels, and (2) the possibility to detect anomalies that have distributed impact on the cyber-physical process by locating several Bro IDSs at different network levels.

Y. Yang *et al.* present a multidimensional intrusion detection system for IEC 61850 based substations, which integrates physical knowledge, protocol specifications, and logical behaviors [51]. Their approach comprises access control detection, protocol whitelisting, model-based detection, and multiparameter-based detection. The protocol whitelist detection deals with various protocols including MMS, Simple Network Time Protocol (SNTP), GOOSE, SMV, etc. The system consists of five modules: IDS configuration module, network traffic capture module, IDS process core module, IDS rule module and IDS result module.

F. HYBRID TECHNIQUES

Cruz *et al.* propose a distributed detection framework, which integrates detection strategies such as detection agents, correlators, and topology and system-specific detection mechanisms [43]. The role of detection agents is to gather the information about suspicious activities from the underlying ICS infrastructure elements and to normalize and send it to the local correlator (for its network domain). Detection agents implement signature-based tools to detect known threats. The distributed multilevel correlation structure is responsible for processing events, collected by the detection agents, with OCSVM adaptive anomaly detection modules. Information from topology databases or asset management systems is also used in the analysis.

Udd *et al.* present a system designed to protect IEC 60870-5-104 protocol, which combines two anomaly detection mechanisms, namely automatic whitelisting and timing analysis [46]. The system operates on electrical substation's internal network and detects zero-day malicious threats as well as benign incidents and misconfigurations. The three main parts of the proposed system are IEC 60870-5-104 parser, learning component and detection component, and they are implemented in the Bro framework.

Feng *et al.* propose anomaly detection framework based on network packet signatures and machine learning techniques [48]. First, a signature database of normal traffic behavior is created by observing communication patterns between the field devices. This database is stored in a Bloom filter, which compares each incoming packet signature to the database signatures. If there is no match, a packet is classified as an anomaly. Otherwise, it is forwarded to the time-series level detector. This detector is represented by a stacked long short term memory (LSTM) neural network-based classifier, a deep learning technique used to address the time dependency between successive packets and to predict packet signatures from previously observed network packets.

Adepu and Mathur present the Distributed Attack Detection (DAD) method to detect attacks in real-time by identifying anomalies in behavior of the physical process in the plant [52]. DAD introduces the notion of state entanglement to derive invariants from the plant design; anomalies are then identified by using monitors that are implementations of invariants. In addition, DAD also uses a simplified version of the CUSUM algorithm, which is based on continuous state variables in the physical process.

Ghazi and Doustmohammadi describe an algorithm based on Petri Net that simultaneously detects misuse and anomaly behavior of the cyber physical system [53]. Anomaly detection is based on Neural First Order Hybrid Petri Net (NFOHPN) with online fast independent component analysis (multivariate statistical method).

Keshk *et al.* propose a privacy-preserving anomaly detection framework for protecting confidential information and discovering malicious activities in power systems networks [61]. To preserve privacy, the data pre-processing module filters and transforms original data into a new format using Pearson correlation coefficient technique. The anomaly detection module is then activated using a Gaussian mixture model (GMM) and Kalman filter to precisely estimate the posterior probabilities of legitimate and anomalous events.

VI. EVALUATION AND COMPARISON OF SOLUTIONS

This section contains evaluation and comparison of selected solutions in terms of general features, test environments and performance evaluation.

A. DETECTION METHODOLOGY, PROTECTED PROTOCOLS AND IMPLEMENTATION TOOLS

We first evaluate and compare general features of presented IDS architectures regarding detection methodology, protected protocols and implementation tools. Table 3 summarizes those general features.

Figure 4 shows statistics of the surveyed papers in terms of detection methodology. Anomaly-based methods prevail (almost two thirds) and among them machine learning techniques based on deep learning as well as clustering and outlier detection are predominant. Apart from inherent suitability for SCADA systems in terms of identifying traffic patterns, driving force for their expansion is probably capability to support FIN technologies, i.e., big data analytics, high level of automation and continuous detection improvement. For the same reasons, signature-based techniques are practically being abandoned.

Specification-based techniques are applied for specific protocols, standalone or in combination with other methods [45], [46], [51]. More comprehensive solutions use a number of integrated techniques to detect a variety of attacks more accurately [43], [46], [48], [52], [53], [61].

Regarding protected protocols, the most widespread SCADA protocols are comprised in surveyed studies, including Modbus, IEC 60870-5 series, DNP3, IEC 61850 and EtherNet/IP. About 69% of surveyed papers consider only one protocol, while 12% deal with multiprotocol environments. The information about SCADA protocol is not available in 19% of surveyed papers.

Various implementation tools are being applied, as indicated in Fig. 5. Open-source NIDS such as Snort, Suricata and Bro are predominantly used for signature-based and hybrid techniques. MATLAB is used for implementation of algorithms rather than mature solutions. In other cases, general-purpose programming languages like C/C++, Java, Python are used either for developing proprietary applications/platforms or in general-purpose open-source platforms. With the growing trend of machine learning techniques, the latter are increasingly used to build SCADA-specific solutions [44], [54], [56], [60], [62], [63]. The information about implementation tool was not available in six papers.

Different test environments were used in presented studies, as summarized in Table 4. Consequently, different testing scenarios were developed.

B. TEST ENVIRONMENT

Regarding testbeds, seven solutions have been verified in powerful CPS testbeds. Software simulation testbeds prevail, with majority of simple simulation based testbeds. Four testbeds are virtualization-based, while one testbed is emulation-based. Figure 6 shows statistics of the surveyed papers in terms of testbeds.

Regarding datasets, only five papers include tests with real SCADA network data [39], [40], [51], [58], [59]. The other 21 papers include one or more experimental and/or synthetic datasets; among them 15 datasets are publicly available. Only two of public datasets are not SCADA-specific, namely KDD99 and UNSW-NB15.

The most diverse situation is with simulated attacks. In some cases, system's behavior under attacks was not analyzed because tests were performed on a real system and limited to suspicious messages and events [39] or the tests were focused only on system's efficiency [50]. Among the other 24 papers the most common simulated attacks comprise attacks on general Internet protocols (11), command/response injection or modification (10), DoS (10) and attacks on SCADA protocols (7). The other simulated attacks were: reconnaissance (5), MITM (3), unauthorized access (1) and probing (1). Six studies present thorough specification and simulation of a number of realistic attacks, which were intended to jeopardize the particular control process [40], [43], [45], [51], [52], [56]. Publicly available datasets contain a number of subsets with simulated attacks, which were selectively used in corresponding studies. Several solutions need more comprehensive testing to different forms of attacks [42], [44], [49], [54], [55], [57]. It should be noted that only two studies included independent validation, performed by invited hackers [45] and six independent teams [52].

C. PERFORMANCE EVALUATION

As mentioned in Section IV.B, we evaluate IDS performance regarding detection accuracy, timeliness, response to incidents and efficiency. Table 5 contains survey of selected papers in terms of IDS performance evaluation completeness. We further provide detailed assessment of selected papers in terms of each evaluation criterion.

1) DETECTION ACCURACY

Table 6 contains summary of SCADA intrusion detection techniques in terms of detection accuracy. The information is based on the data reported in surveyed studies. The depth of analysis varies between studies; hence, the question arises as to whether the results are completely relevant.

Grade in Table 6 denotes summary assessment of the detection accuracy analysis. Low-grade stands for six papers where analysis is not presented or the results are given in a descriptive way; for such systems, more thorough testing is needed to determine real detection capabilities. For example, if the tests were performed without attack simulations, only *FPR* could be calculated, and what was tested is in fact the accuracy of normal data classification [39]. Similarly, if detection accuracy analysis is not present, the statement that ''accuracy is 100%'' should be accepted with caution [45], [51], [54]. Some papers contain descriptive results rather than well-defined evaluation metrics [52], [56].

TABLE 3. Summary of SCADA intrusion detection techniques: Detection methodology, protected protocols and implementation tools.

Reference	Detection methodology	Protected protocols	Implementation tools
Erez & Wool [39]	Finite automata	Modbus TCP	Not published
Kwon et al. [40]	Multivariate	IEC 61850: MMS, GOOSE	Not published
Al Balushi et al. [41]	Expert system: Ontology-based	Modbus TCP	Java
Almalawi et al.[42]	Clustering and outlier detection	Modbus TCP	Not published
Cruz et al. [43]	Hybrid: Signature-based + clustering and outlier detection (OCSVM)	Modbus TCP	Snort
Da Silva et al. [44]	Clustering and outlier detection: OCSVM and SVDD	Modbus TCP	LIBSVM tool (C/C++)
Ghaeini & Tippenhauer [45]	Specification-based	EtherNet/IP	Bro
Udd et al. [46]	Hybrid: Specification-based + time series	IEC 60870-5-104	Bro
Zhang et al. [47]	Time series	Modbus TCP	Not published
Feng et al. [48]	Hybrid: Clustering and outlier detection (Bloom filter) + neural network	Modbus application layer protocol	Not published
Wan et al. [49]	Clustering and outlier detection: OCSVM + RE-KPCA	Modbus TCP	Proprietary Linux C software and MATLAB
Wong et al. [50]	Signature-based	EtherNet/IP	Suricata
Y. Yang et al. [51]	Specification-based	IEC 61850: MMS, GOOSE, SMV	ITACA platform (C/C++)
Adepu & Mathur [52]	Hybrid: Finite automata (state condition graphs) + simplified version of CUSUM	EtherNet/IP	Structured Text and Python
Ghazi & Doustmohammadi [53]	Hybrid: Neural networks (NFOHPN) + multivariate	Not published	MATLAB
Hijazi & Flaus [54]	Deep learning based on ANN	Modbus TCP	TensorFlow and Keras
Lin et al. [55]	Expert system: Semantic analysis framework	DNP3	Bro
Myers et al. [56]	Clustering and outlier detection: Process mining	Siemens S7 (process level)	Python and Process Mining toolkit (ProM)
Wang and Feng [57]	Time series (visualized)	IEC 61870-5-104	MATLAB
Wressnegger et al. [58]	Expert system: Building prototype models for network messages + noise-resilient anomaly detection	Six industrial protocols	Self-developed tool
Benisha & Raja Ratna [59]	Clustering and outlier detection: Entropy-based extreme learning machine	Not published	MATLAB
Derhab et al. [60]	Clustering and outlier detection: Random subspace learning + KNN	Not published	Weka
Keshk et al. [61]	Hybrid: Clustering and outlier detection (GMM) + time series (Kalman filter)	Not published	Programming language R

TABLE 3. *(Continued.)* **Summary of SCADA intrusion detection techniques: Detection methodology, protected protocols and implementation tools.**

Khan et al. [62]	Clustering and outlier detection: Bloom filter + KNN	Modbus TCP	Anaconda (Python 3.7) and Spyder
Lai et al. [63]	Deep learning based on CNN	Not specified	TensorFlow
H. Yang et al. [64]	Deep learning based on CNN	DNP3	Not published

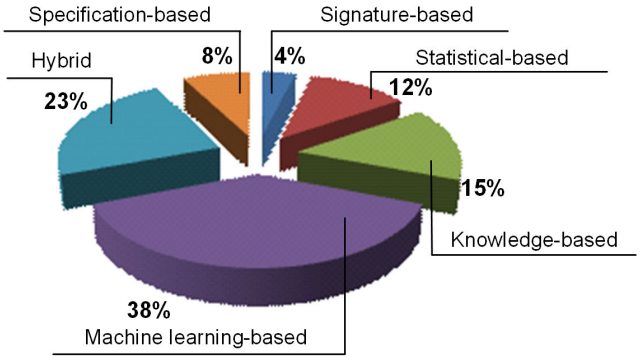

FIGURE 4. Statistics of the surveyed papers regarding detection methodology.

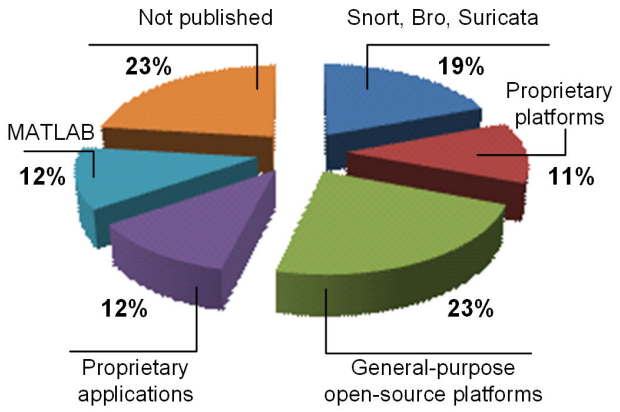

FIGURE 5. Statistics of the surveyed papers regarding implementation tools.

Eight papers are rated as medium-grade, which means that the results are presented through smaller number of evaluation metrics (typically *Accuracy* and *FPR*). High-grade denotes comprehensive detection accuracy analysis with results in terms of evaluation metrics listed in Section IV.B; eleven papers fulfill that criterion. We further present observations related to high and medium-graded studies.

Statistical-based techniques provide high accuracy, with low *FPR* and *FNR* rates [40], [47], [57]. Similarly, knowledge-based techniques provide good overall accuracy [41], [55], [58].

Among machine learning-based techniques, deep learning based on CNN [63], [64] outperforms techniques based on clustering and outlier detection[42], [44], [49], [59], [60], [62].

Detection accuracy of hybrid methods [43], [46], [48], [53], [61] depends on combined techniques.

2) TIMELINESS

Timeliness analysis is available in 11 studies, as indicated in Table 7. Among the results concerning packet as a unit of analysis, deep packet inspection applied in [51] outperforms other techniques [41], [43], [46] at least for an order of magnitude. In the cases where dataset instance is observed as a unit of analysis, deep learning method presented in [63] performs much worse than clustering and outlier detection [60] and hybrid method presented in [61]. This is not surprising since deep learning inherently requires large datasets to obtain high accuracy.

The results concerning detection latency are hardly comparable probably due to different experimentation platforms; thus, detection latency seems lower in test scenarios with simple simulation.

3) RESPONSE TO INCIDENTS

Only four systems provide active responses to detected attacks, as indicated in Table 8. The IDS described in [43] is linked with the hierarchically organized set of correlators in order to react accordingly to the detected intrusions. In [44], the proposed NIDS can act proactively or it can generate alarms for SCADA operators, through the appropriate management interface, to define response policies. Intrusion response mechanism presented in [55] is designed for a specific attack scenario in which an intruder injects malicious commands to disconnect multiple transmission lines simultaneously. Technique presented in [59] supposes data encryption and selection of trusted path, after attack detection.

4) EFFICIENCY

Only five papers provide the results of efficiency evaluation, as summarized in Table 9. It can be observed that the results concerning memory usage are comparable for [43], [50] and [56], while the systems presented in [45] and [62] use less memory (of an order or two of magnitude). Results concerning CPU usage are hardly comparable due to different processor platforms; however, the results presented in [43] and [50] confirm that CPU usage increases for higher traffic load. Packet loss and/or alert loss under high traffic load are addressed in [50].

D. CONCLUDING EVALUATION

For each surveyed IDS solution, the concluding evaluation comprises the following:

TABLE 4. Summary of SCADA intrusion detection techniques in terms of test environments.

Reference	Testbed	Institution	Datasets	Simulated attacks
Erez & Wool [39]	CPS testbed: Real system that controls electric power supply	Tel Aviv University, Israel	Datasets obtained by recording real traffic	Only suspicious messages and events
Kwon et al. [40]	CPS testbed: Based on real grid-connected IDS sensor equipment in a digital substation	Smart grid testbed in Korea	Experimental datasets (collecting network traffic of a digital substation for a week)	Scan attack; DoS, GOOSE attack, MMS attack, general network attacks
Al Balushi et al. [41]	Federated simulation: Modbus server (ConPot), Modbus client (Python), attacker (KALI Linux)	Not published	Two datasets: (1) Digital Bond and (2) synthetic dataset	DoS, reconnaissance, protocol specification violation and attacks on system integrity
Almalawi et al. [42]	Virtualization: EPANET (hydraulic and water quality modeling tool), Modbus/TCP-Salve simulator and VMware Player	Not published	Eight labeled datasets: five of them publicly available and three synthetic datasets	Packet modification performed through MITM attack
Cruz et al. [43]	CPS testbed: Hybrid Environment for Development and Validation (HEDVa)	HEDVa designed by Israeli Electric Company	Four experimental datasets comprising data from normal operation mode and three types of attacks	Layer 2/3/4 attacks, SCADA protocol and process-level attacks, and compromised HMI/PLC
Da Silva et al. [44]	Federated simulation: SDN-based SCADA system of a power grid company	Not published	Synthetic datasets	DoS attack at one of the substations
Ghaeini & Tippenhauer [45]	CPS testbed: SWaT (Secure Water Treatment) industrial control system	Singapore University of Technology and Design	Experimental datasets (traffic captured at ports specified for EtherNet/IP and CIP)	SCADA-specific attacks, general network attacks, invited hackers
Udd et al. [46]	Simple simulation: Open source tools (Wireshark, Tcprewrite, Mergecap, Bittwiste, Scapy)	Not published	Experimental datasets created from KTH Royal Institute of Technology, Sweden	Port scans from unknown and known hosts, MITM, spoofing
Zhang et al. [47]	Virtualization: IDS modules performed on separate virtual machines (Kali Linux 2.0) and deployed under a LAN	Not published	Synthetic datasets	Command injection, response injection, DoS, reconnaissance
Feng et al. [48]	Simple simulation: Analysis of data traces	Not published	Public dataset from laboratory-scale testbed of a gas pipeline system	Command injection, response injection, DoS, reconnaissance
Wan et al. [49]	Federated simulation: Libcap, Linux C, MATLAB	Not published	Synthetic datasets	Minor, medium and major attacks (regarding number of abnormal addresses or values)
Wong et al. [50]	Federated simulation: Traffic generator and the Suricata IDS	Not published	Synthetic datasets from several packet captures (pcap files)	Not published
Y. Yang et al. [51]	CPS testbed: Replication of the SCADA network of a typical 500kV smart substation.	Laboratory of Substation Intelligent Equipment Testing Techno-logy, China	Real SCADA traffic captured from an operating 500 kV substation based on IEC 61850	Malformed packet attack, DoS attack, spoofing attack, MITM attack

TABLE 4. *(Continued.)* **Summary of SCADA intrusion detection techniques in terms of test environments.**

Adepu & Mathur [52]	CPS testbed: SWaT (Secure Water Treatment) industrial control system	Singapore University of Technology and Design	Experimental datasets extracted from SWaT	Static and dynamic attacks, launched on single or multiple stages of the plant
Ghazi & Doustmoham-madi [53]	Simple simulation: MATLAB	Not published	KDD99 dataset (public)	DoS, unauthorized access, probing
Hijazi & Flaus [54]	Virtualization: Three virtual machines simulating network traffic, HMI and PLC.	Not published	Synthetic datasets (Modbus/TCP traffic)	Anomalous packets with different IP addresses, ports, functions, and values combinations
Lin et al. [55]	Virtualization: Network communications (two virtual machines) and MATPOWER (power system simulation)	Not published	Synthetic datasets (simulated DNP3 network traffic)	Control-related attacks
Myers et al. [56]	Two CPS testbeds: (1) power plant system and (2) water tank system	Not published	Two public datasets recorded from industry standard ICS devices	Injection attacks, flooding
Wang and Feng [57]	Simple simulation: Master station simulation software and protocol tester	Not published	Synthetic datasets	Various forms of SCADA-specific attacks
Wressnegger et al. [58]	Simple simulation: Analysis of data traces	Not published	Real SCADA traffic captured from a large European energy producer	Self-developed tool for the automatic generation of network attacks against unknown protocols
Benisha & Raja Ratna [59]	Simple simulation: MATLAB	Not published	Real traffic captured from the SCADA wind turbines	Biased injection attacks
Derhab et al. [60]	Emulation: Experimental platform combining open-source cloud, SDN, blockchain and networking technologies	Not published	Power System (public)	Command injection, relay setting change, data injection
Keshk et al. [61]	Simple simulation: Programming language R	Not published	Two public datasets: Power System and UNSW-NB15	Various forms of SCADA-specific attacks
Khan et al. [62]	Simple simulation: Anaconda (Python 3.7) and Spyder	Not published	Experimental dataset from the gas pipeline system (Mississippi State University's SCADA lab)	Command injection, response injection, code injection, DoS, reconnaissance
Lai et al. [63]	Simple simulation: TensorFlow	Not published	As in [62]	As in [62]
H. Yang et al. [64]	Simple simulation: Analysis of data traces	Not published	Two experimental datasets extracted from energy-delivery systems' testbeds (University of Arkansas and EPRI)	Attacks on common Internet protocols, DNP3-specific attacks

TABLE 5. Summary of selected papers in terms of IDS performance evaluation completeness.

Reference	Detection accuracy	Timeliness	Active response	Efficiency
Erez & Wool [39]	✓			
Kwon et al. [40]	✓			
Al Balushi et al. [41]	✓	✓		
Almalawi et al. [42]	✓			
Cruz et al. [43]	✓	✓	✓	✓
Da Silva et al. [44]	✓		✓	
Ghaeini & Tippenhauer [45]	✓			✓
Udd et al. [46]	✓	✓		
Zhang et al. [47]	✓			
Feng et al. [48]	✓			
Wan et al. [49]	✓	✓		
Wong et al. [50]				✓
Y. Yang et al. [51]	✓	✓		
Adepu & Mathur [52]	✓			
Ghazi & Doustmohammadi [53]	✓			
Hijazi & Flaus [54]	✓			
Lin et al. [55]	✓	✓	✓	
Myers et al. [56]	✓			✓
Wang and Feng [57]	✓			
Wressnegger et al. [58]	✓			
Benisha & Raja Ratna [59]	✓		✓	
Derhab et al. [60]	✓	✓	✓	
Keshk et al. [61]	✓	✓		
Khan et al. [62]	✓	✓		✓
Lai et al. [63]	✓	✓		
H. Yang et al. [64]	✓			

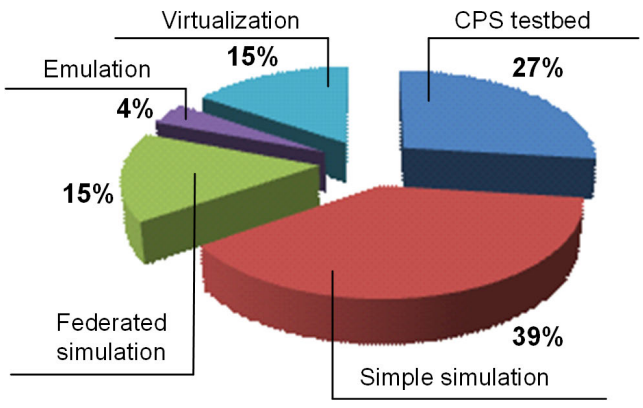

FIGURE 6. Statistics of the surveyed papers regarding testbeds.

- Strengths and weaknesses in terms of detection methodology, protected protocols, implementation tools, test environment and performance evaluation;
- Maturity stage – refers to the level of implementation readiness, as explained in Table 10. We adopt maturity stage dimension from the model for software process improvement proposed in [65].
- FIN – refers to the applicability in the Future Internet environment, and assumes standards-based 1–4 grading scale, as explained in Table 11.

Table 12 summarizes the concluding evaluation of surveyed SCADA IDS solutions.

VII. SUMMARY OF REVIEW FINDINGS AND FUTURE RESEARCH DIRECTIONS

Regarding previous surveys of intrusion detection systems in ICS, and particularly SCADA environment, we identify a progress in some areas, but also some open issues remain. The main results of our analysis are as follows.

1. It is recognized that signature-based techniques are insufficient to secure SCADA systems due to their inherent drawbacks regarding inability to cope with new or unknown threats and the need to continuously update signatures. In the future, they will probably be used in hybrid techniques, i.e., as a complement to anomaly-based and specification-based techniques.
2. A variety of novel anomaly-based detection techniques for SCADA were proposed and tested. Among them, machine learning-based techniques have gained a strong

TABLE 6. Summary of SCADA intrusion detection techniques in terms of detection accuracy.

Reference	Detection accuracy	Grade
Erez & Wool [39]	*FPR* = 0.86%; *Accuracy* = 93%	Low
Kwon et al. [40]	*FPR* = 0; *FNR* = 0.011; *DR* = 0.9889; *Accuracy* = 99%; *Precision* = 1; *F-measure* = 0.914	High
Al Balushi et al. [41]	*FPR* = 1.09%; *DR* ~ 97.80%; *Accuracy* ~ 96.70%;	Medium
Almalawi et al. [42]	$0 \leq FPR \leq 19.92\%$; $92.98\% \leq DR \leq 100\%$; $92\% \leq Precision \leq 100\%$	High
Cruz et al. [43]	$1.18\% \leq FPR \leq 3.25\%$; $94.6\% \leq Accuracy \leq 98.8\%$	Medium
Da Silva et al. [44]	OCSVM: *FPR* = 0; *FNR* ~ 1.565%; *DR* ~ 98.435%; *TNR* = 100%; *Accuracy* ~ 99.18%; SVDD: *FPR* = 0; *FNR* ~ 4.281%; *DR* ~ 95.718%; *TNR* = 100%; *Accuracy* ~ 97.755%	High
Ghaeini & Tippenhauer [45]	Detection of network-based attacks on Historian, PLC and HMI/SCADA, with *FPR* = 0; Failed to detect insider command injection attacks	Low
Udd et al. [46]	For all attacks: $0.01\% \leq FPR < 0.4\%$ (depends on the predefined timing threshold value); For port scans and MITM: *DR* = 100%; for spoofing: *DR* = 0	Medium
Zhang et al. [47]	*FPR* = 0.72%; *FNR* = 0.37%	Medium
Feng et al. [48]	*DR* = 78%; *Accuracy* = 92%; *Precision* = 94%; *F-measure* = 85%	High
Wan et al. [49]	Function control behavior: $57.55\% \leq Accuracy \leq 91.37\%$ (depending on kernel function) Process data behavior: $72.92\% \leq Accuracy \leq 94.44\%$ (depending on attack type)	High
Y. Yang et al. [51]	*Accuracy* = 100%	Low
Adepu & Mathur [52]	$0 \leq FNR \leq 38.89\%$	Low
Ghazi & Doustmohammadi [53]	$0.4\% \leq FPR \leq 2.9\%$; $96.2\% \leq DR \leq 100\%$	Medium
Hijazi & Flaus [54]	*Accuracy* = 99.9%	Low
Lin et al. [55]	$0 \leq FPR \leq 0.78\%$; $0.00048\% \leq FNR \leq 0.013\%$	Medium
Myers et al. [56]	Results obtained by ProM "Trace fitness" First dataset: *TP*=16; *FN*=7; *FP*=2; *TN*=13; Second dataset: *TP*=9; *FN*=3; *FP*=0; *TN*=8	Low
Wang and Feng [57]	95,8%; ≤ *Accuracy* ≤ 00%; *Precision* = 100%	Medium
Wressnegger et al. [58]	$0.8463 \leq AUC \leq 0.9984$ (for $0.0001 \leq FPR \leq 0.01$)	Medium
Benisha & Raja Ratna [59]	*DR* = 81,81%; *Accuracy* = 98,79%; *Precision* = 97,28%; *F-measure* = 56,21%	High
Derhab et al. [60]	Binary classification: $5.3\% \leq FPR \leq 8.5\%$: $94.86\% \leq Accuracy \leq 96.74\%$; Multi-class classification: $0.3\% \leq FPR \leq 0.4\%$; $88.32\% \leq Accuracy \leq 91.08\%$	High
Keshk et al. [61]	Power System dataset: $3.47\% \leq FPR \leq 5.66\%$; $94.04\% \leq DR \leq 96.24\%$; $94.56\% \leq Accuracy \leq 97.27\%$ UNSW-NB15 dataset: $6.23\% \leq FPR \leq 8.81\%$; $90.19\% \leq DR \leq 92.07\%$; $90.70\% \leq Accuracy \leq 93.92\%$	High

TABLE 6. *(Continued.)* **Summary of SCADA intrusion detection techniques in terms of detection accuracy.**

Khan et al. [62]	*DR* = *92%; Accuracy* = 97%; *Precision* = 98%; *F-measure* = 95%; *ROC* analysis also provided	High
Lai et al. [63]	Binary classification: *DR* = 98.76%; *Accuracy* = 99.46%; *Precision* = 99.58%; *F-measure* = 99.17% Multi-class classification: 93.81% ≤ *DR* ≤ 99.11%; *Accuracy* = 99.32%; 94.88% ≤ *Precision* ≤ 100%; 94.88% ≤ *F-measure* ≤ 99.29%	High
H. Yang et al. [64]	93.90% ≤ *DR* ≤ 100%; *Accuracy* = 99.38%; 94% ≤ *Precision* ≤ 100%	High

TABLE 7. **Summary of SCADA intrusion detection techniques in terms of timeliness.**

Reference	Unit of analysis	Per-unit processing time	Detection latency
Al Balushi et al. [41]	Packet	~ 1 ms	Packet feature extraction: 492.77 ms Attack detection (15 rules): 21.86 ms
Cruz et al. [43]	Packet (event message)	~ 2 ms for 1 kB message < 10 ms for 20 kB message	
Udd et al. [46]	Packet	~ 4 ms	
Wan et al. [49]			~ 1 s per each attack sample
Y. Yang et al. [51]	Packet	< 0.3 ms	
Ghazi and Doustmohammadi [53]			122.8 ms
Lin et al. [55]			Large-scale system: ~ 600 ms Small-scale system: ~ 5 ms
Derhab et al. [60]	Dataset instance (128 features)	2.3 ms for binary classification 9.4 ms for multi-class classification	
Keshk et al. [61]	Dataset instance (8 features)	5.9 ms	
Khan et al. [62]			0.109 ms
Lai et al. [63]	Dataset instance	253 ms for feature mapping 192 ms for detection	

TABLE 8. **Summary of SCADA intrusion detection techniques in terms of active response.**

Reference	Type of active response
Cruz et al. [43]	Linked with the distributed multilevel correlation structure
Da Silva et al. [44]	(1) Alarm generation; (2) Automatic response – redirecting of anomalous flow to Honeypot, dropping malicious packets
Lin et al. [55]	Exploits reclose logic in relays to prevent physical damage caused by an attempt to disconnect multiple transmission lines
Benisha and Ratna [59]	Data encryption using the Hybrid Elliptical Curve Cryptography

momentum in the past few years, either stand-alone or in combination with other techniques. Our research shows that additional work is needed to improve their overall detection accuracy; this particularly stands for clustering and outlier detection. Knowledge-based techniques perform better in terms of detection accuracy, but on the count of deteriorated timeliness, especially for large-scale systems. Statistical-based techniques are most useful in hybrid techniques because of their high detection accuracy.

3. Specification-based techniques gain in importance for SCADA application layer protocols. They perform well in terms of both detection accuracy and per-unit processing time.
4. Integration of two or more detection methods may contribute to improvement of the IDS scope and detection accuracy. Particular attention should be focused on hybrid techniques that ensure high level of reusability, in multiprotocol environments.
5. The most widespread SCADA protocols are addressed in the recent works. Still, Modbus TCP prevails, while additional research efforts are needed towards environments such as digital substations and smart grids. In such a context, securing SCADA systems that use advanced networking technologies, e.g., SDN, deserves more attention.

TABLE 9. Summary of SCADA intrusion detection techniques in terms of efficiency.

Reference	Efficiency
Cruz et al. [43]	Memory usage: 810 MB CPU usage (six-core processor): ~ 70% for 1, 2, and 5 kB event bursts; ~ 80% for 10 and 20 kB event bursts
Ghaeini & Tippenhauer [45]	Memory usage: ~ 8.8% (of 1 GB RAM); CPU usage (quad-core processor): ~ 25% for SCADA specific attacks and general network attacks
Wong et al. [50]	Two offline tests: each test was run 5 times at different throughput rates and different limitation of memory usage by Suricata (750 MB and 1.5 GB) CPU usage (dual-core processor): 20% at throughput of 18 Mbps; 120% at throughput of 130 Mbps Packet loss: Begins for throughput > 37 Mbps; For throughput ≥ 50 Mbps more packets are lost with lower memory usage
Myers et al. [56]	Pre-processing (transforming device log to event log): Memory usage: 130 MB; CPU time: 8.56 s
Khan et al. [62]	Memory usage: 1040 kB

TABLE 10. Maturity Stage (MS) dimension.

MS	Description
1 – Initial	Simulation phase or information about software tool is not published
2 – Aware	Awareness to implementation process is gained
3 – Defined	Focused on implementation process on a well-established software platform
4 – Optimizing	The solution is implemented and the focus lies on continuous improvement

TABLE 11. Applicability to the Future Internet (FIN) environment.

FIN	Description
1	Little or no applicability
2	Partially applicable, e.g., designed for digital substations, unsupervised learning, etc.
3	Applicable in terms of big data analytics, distributed detection, large scale of devices
4	Fully applicable

6. Open-source NIDS implementation tools such as Snort, Bro and Suricata are being superseded by open-source and proprietary IDS platforms, developed in some of general-purpose programming languages and equipped with appropriate application programming interfaces. To enable development of superior performance IDS, attempts to build open-source platforms dedicated to industrial environment should be enforced.
7. It is recognized that realistic and comprehensive cyber physical system testbeds are needed to allow for experimentation with different solutions. Although they were used in 27% of surveyed papers, it is well-known that building such testbeds requires significant research efforts and financial resources; moreover, they are often part of national strategies for critical infrastructure protection. If they are unavailable, sophisticated simulation/emulation testbeds should be developed. Virtualization may help to provide inexpensive, credible and reusable testbeds. Particularly, simple simulation-based testbeds should be avoided due to their low fidelity and poor reusability.
8. There is a strong need to use datasets from real SCADA networks; hence, national strategies for critical infrastructure protection should find a way to make them available to research community. Besides, good strategy is to reuse datasets, either publicly available or obtained from CPS testbeds.
9. Still, there is a lack of proper attack models and scenarios in which the attackers try to exploit vulnerabilities in SCADA systems. Consequently, the reports on IDS performance evaluation might be insufficiently reliable and hardly comparable. Hence, efforts are needed to improve frameworks for modeling cyber attacks and procedures to apply them in the appropriate testbeds. Work on free and open source SCADA-specific adversary emulation tools should also be encouraged; such tools are available for general-purpose IT networks.
10. Performance evaluation remains the most critical issue. The work is needed on identification and specification of requirements for IDSs in SCADA networks, and establishing a common set of performance metrics. This should include at least detection accuracy, timeliness, response to incidents and efficiency. Procedures for IDS performance testing (and reporting results) should be established in accordance with the predefined set of requirements. In surveyed papers, detection accuracy was considered with very different level of details; however, less than half of them contained comprehensive analysis. The situation is much worse with timeliness and efficiency analysis. Timeliness analysis should be presented in each new proposal, since it is

TABLE 12. Concluding evaluation of surveyed SCADA intrusion detection systems.

Reference	Strengths	Weaknesses	MS	FIN
Erez & Wool [39]	Accurate modeling of Modbus TCP for intrusion detection purpose Tested on a real SCADA system	Implementation tool not specified Lack of simulated attacks Classification accuracy needs improvement	1	1
Kwon et al. [40]	Designed for IEC 61850 substations Capability for big data analytics Powerful testbed Wide spectrum of simulated attacks	Implementation tool not specified	1	3
Al Balushi et al. [41]	Modular structure Wide spectrum of simulated attacks Good timeliness analysis Comparison with a similar method	Needs extension to other industrial protocols, besides Modbus TCP	2	1
Almalawi et al. [42]	Unsupervised learning task Multiple datasets	Implementation tool not specified Tested for MITM attack only	1	2
Cruz et al. [43]	Hybrid detection methodology Distributed and modular structure Powerful testbed Wide spectrum of simulated attacks Extensive efficiency analysis Active response	Lack of detailed explanation about detection accuracy evaluation	3	3
Da Silva et al. [44]	Intended for SDN-based SCADA Suitable for large scale systems Active response	Tested for DoS attack only	3	4
Ghaeini & Tippenhauer [45]	Hierarchical and distributed IDS Web-based user interface Powerful testbed (SWaT) Tested to wide spectrum of attacks, including invited hackers	Incapable of detecting all insider attacks Lack of detailed explanation about detection accuracy evaluation	3	3
Udd et al. [46]	Hybrid detection methodology One of the two systems concerning IEC 60870-5-104 protocol	Incapable of detecting spoofing attacks Lack of detailed explanation on detection accuracy and timeliness evaluation	3	1
Zhang et al. [47]	Modular structure Capable of effectively resisting response injection and DoS attacks	Implementation tool not specified	1	1
Feng et al. [48]	Hybrid detection methodology Comparison with six similar methods	Implementation tool not specified *DR* metric needs improvement	1	1
Wan et al. [49]	Observation of function control and process data behaviors Comparison with two similar methods	Needs more detailed specification of simulated attacks	2	1
Wong et al. [50]	Extension of Suricata signature-based systems with EtherNet/IP Became a part of Suricata release Extensive efficiency analysis	Lack of simulated attacks Verification of detection capabilities not published	4	1
Y. Yang et al. [51]	Designed for IEC 61850 substations Mature software platform (ITACA) Powerful testbed Comparison with four similar methods	Lack of detailed explanation about detection accuracy evaluation	3	2
Adepu & Mathur [52]	Hybrid and distributed detection method, included in SWaT testbed One of the experiments was carried out by six independent teams	Use of unconventional metrics in detection accuracy evaluation	4	3
Ghazi & Doustmohammadi [53]	Hybrid detection methodology Comparison with four similar methods	Needs testing on SCADA-specific datasets	1	1

crucial parameter to assess system's ability to respond to incident in real time. In addition, it is very important to provide results of efficiency analysis (under heavy traffic load, if possible) since they represent indirect

TABLE 12. *(Continued.)* **Concluding evaluation of surveyed SCADA intrusion detection systems.**

Hijazi & Flaus [54]	Capability for big data analytics	Needs testing with SCADA-specific attacks	2	3
Lin et al. [55]	Semantic analysis framework Active response	Needs more realistic simulation of control related attacks	3	1
Myers et al. [56]	Process mining method Two powerful testbeds	Presents only confusion matrix obtained by ProM toolkit	2	2
Wang and Feng [57]	Use of graphical features One of the two systems concerning IEC 60870-5-104 protocol	Needs more detailed specification of simulated attacks	1	1
Wressnegger et al. [58]	Analysis of six industrial protocols Large scale evaluation (92700 devices) Comparison with a similar method	Lack of specifications for the self-developed implementation and testing tools	2	3
Benisha & Raja Ratna [59]	Modular structure Active response	*DR* metric needs improvement	1	1
Derhab et al. [60]	Intended for IoT-based ICS Blockchain-based integrity checking Active response	*FPR* metric needs improvement in binary classification *Accuracy* needs improvement in multi-class classification	3	4
Keshk et al. [61]	Hybrid detection methodology Evaluation of privacy Comparison with 7 similar methods	Detection accuracy needs improvement	1	1
Khan et al. [62]	Multilevel anomaly detection Comparison with five similar methods	Detection accuracy might be improved	2	1
Lai et al. [63]	Capability for big data analytics Comparison with three similar methods	Detection accuracy achieved on the count of timeliness	2	3
H. Yang et al. [64]	Capability for big data analytics	Implementation tool not specified	1	3

measures that take into account the time and space complexities of intrusion detection algorithm.

11. Only four of surveyed papers discussed active responses to detected attacks. There is a strong need to perceive the overall SCADA security system architecture and to define procedures for real-time interaction of the IDS and other components of the security system like correlators, SIEM software, etc. Particularly, work on IPS capabilities should be strongly encouraged.
12. Since SCADA systems are complex environments that are highly dependent on the specific industrial process and the expertise of responsible professionals and decision makers, it is complicated to develop universal guidelines that should be applicable to all cases, due to diverse technical, socio-economic and regulatory conditions. For this reason, reports on best practices remain valuable for most professionals; in this regard, IDSs surveyed in this paper (particularly, the ones with high maturity score) can assist other practitioners in developing their own solutions.
13. Finally, evolution towards fourth-generation SCADA brings new research challenges related to security in industrial IoT environment that assumes the use of FIN technologies. Apart from evolving threats, factors that affect security include heterogeneity of physical objects, networking technologies and applications that should be able to communicate and collaboratively provide immutable and verifiable data. Bearing that in mind, we have analyzed applicability of each surveyed solution in the future environment. Prerequisites for applicability include capability for big data analytics, distributed intrusion detection, involving large scale of devices, etc. In such a context, the prevalence of machine learning-based techniques is quite reasonable. Particularly, unsupervised machine learning algorithms should be better understood and explored for use in future SCADA networks. FIN environment strengthens the need for CPS testbeds with large-scale industrial devices, massive datasets and realistic attack models for IIoT.

VIII. CONCLUSION

Growth of solutions for SCADA IDS gains in importance with proliferation of advanced networking technologies and the ongoing evolution towards fourth-generation SCADA systems. In this work, we focused on research work on network-based SCADA intrusion detection systems in the period 2015–2019.

We proposed the evaluation methodology that encompassed identification of general IDS features and analysis of system's characteristics regarding detection technique, protected protocols, implementation tools, test environment and performance evaluation. Final assessment is then performed based on the previous analysis, including strengths, weaknesses, maturity stage, as well as portability to FIN environment. In comparison with related works, results of our study point to significant progress in developing new intrusion detection methods (particularly machine learning-based), using open-source implementation tools, and creating sophisticated security testbeds.

The most important future research directions include development of proper attack models, establishment of procedures for IDS performance evaluation, and integration of IDS with other components of ICS security system, particularly bearing in mind the migration towards Future Internet environment.

REFERENCES

[1] N. Tariq, M. Asim, and F. A. Khan, "Securing SCADA-based critical infrastructures: Challenges and open issues," *Procedia Comput. Sci.*, vol. 155, pp. 612–617, 2019.

[2] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "Guide to industrial control systems (ICS) security," NIST, Gaithersburg, MD, USA, NIST Special Publication 800-82, 2015, vol. 2. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r2.pdf

[3] R. I. Ogie, "Cyber security incidents on critical infrastructure and industrial networks," in *Proc. 9th Int. Conf. Comput. Autom. Eng. (ICCAE)*, Sydney, NSW, Australia, 2017, pp. 254–258.

[4] R. Derbyshire, B. Green, D. Prince, A. Mauthe, and D. Hutchison, "An analysis of cyber security attack taxonomies," in *Proc. IEEE Eur. Symp. Secur. Privacy Workshops (EuroS&PW)*, New York, NY, USA, Apr. 2018, pp. 153–161.

[5] K. E. Hemsley and R. E. Fisher, "History of industrial control system cyber incidents," Idaho Nat. Lab., Idaho Falls, ID, USA, INL/CON-18-44411-Revision-2, 2018. [Online]. Available: https://www.osti.gov/servlets/purl/1505628

[6] B. Gregory-Brown. Securing Industrial Control Systems-2017. A SANS Survey, SANS Institute, 2017. [Online]. Available: https://www.sans.org/reading-room/whitepapers/analyst/securing-industrial-control-systems-2017-37860

[7] B. Filkins and D. Wylie. State of OT/ICS Cybersecurity Survey. SANS Institute & OWL Cyber Defense, 2019. [Online]. Available: https://owlcyberdefense.com/resource/sans-institute-2019-state-of-ot-ics-cybersecurity-survey/

[8] C. M. Hurd and M. V. McCarty. A Survey of Security Tools for the Industrial Control System Environment. DARPA Public Release Center, 2017. [Online]. Available: https://www.osti.gov/servlets/purl/1376870

[9] X. Zhou, Z. Xu, L. Wang, and K. Chen, "What should we do? A structured review of SCADA system cyber security standards," in *Proc. 4th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Barcelona, Spain, 2017, pp. 605–614.

[10] S. Ghosh and S. Sampalli, "A survey of security in SCADA networks: Current issues and future challenges," *IEEE Access*, vol. 7, pp. 135812–135831, Jul. 2019.

[11] M. Stojanović and S. B. Rakas, Eds., *Cyber Security of Industrial Control Systems in the Future Internet Environment*. Hershey, PA, USA: IGI Global, 2020.

[12] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems," NIST, Gaithersburg, MD, USA, NIST Special Publication 800-94, 2007. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-94.pdf

[13] B. Galloway and G. P. Hancke, "Introduction to industrial control networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 860–880, 2nd Quart., 2013.

[14] J. D. Markovic-Petrovic, M. D. Stojanovic, and S. V. B. Rakas, "A fuzzy AHP approach for security risk assessment in SCADA networks," *Adv. Electr. Comput. Eng.*, vol. 19, no. 3, pp. 69–74, 2019.

[15] J. D. Markovic-Petrovic and M. D. Stojanovic, "An improved risk assessment method for SCADA information security," *Elektronika ir Elektrotechnika*, vol. 20, no. 7, pp. 69–72, Sep. 2014.

[16] S. Nazir, S. Patel, and D. Patel, "Assessing and augmenting SCADA cyber security: A survey of techniques," *Comput. Secur.*, vol. 70, pp. 436–454, Sep. 2017.

[17] M. Stojanovic, S. Bostjancic-Rakas, and J. Markovic-Petrovic, "SCADA systems in the cloud and fog environments: Migration scenarios and security issues," *Facta Universitatis-Series, Electron. Energetics*, vol. 32, no. 3, pp. 345–358, 2019.

[18] M. Mantere, M. Sailio, and S. Noponen, "Network traffic features for anomaly detection in specific industrial control system network," *Future Internet*, vol. 5, no. 4, pp. 460–473, 2013.

[19] T. Morris and W. Gao, "Classifications of industrial control system cyber attacks," in *Proc. 1st Int. Symp. ICS SCADA Cyber Secur. Res.*, Leicester, U.K., 2013, pp. 22–29.

[20] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2046–2069, 4th Quart., 2013.

[21] L. Maglaras, M. A. Ferrag, A. Derhab, M. Mukherjee, and H. Janicke, "Cyber security: From regulations and policies to practice," in *Strategic Innovative Marketing and Tourism* (Springer Proceedings in Business and Economics), A. Kavoura, E. Kefallonitis, and A. Giovanis, Eds. Cham, Switzerland: Springer, 2019, pp. 763–770.

[22] B. Zhu and S. Shankar, "SCADA-specific intrusion detection/prevention systems: A survey and taxonomy," in *Proc. 1st Workshop Secure Control Syst. (SCS)*, Stockholm, Sweden, 2010, pp. 1–16.

[23] R. Mitchell and I.-R. Chen, "A survey of intrusion detection techniques for cyber physical systems," *ACM Comput. Surv.*, vol. 46, no. 4, Mar. 2014, Art. no. 55.

[24] I. Garitano, R. Uribeetxeberria, and U. Zurutuza, "A review of SCADA anomaly detection systems," in *Proc. 6th Int. Conf. Soft Comput. Models Ind. Environ. Appl. (SOCO)*, vol. 87, E. Corchado, Eds. Berlin, Germany: Springer, 2011, pp. 357–366.

[25] J. E. Rubio, C. Alcaraz, R. Roman, and J. Lopez, "Analysis of intrusion detection systems in industrial ecosystems," in *Proc. 14th Int. Joint Conf. e-Bus. Telecommun. (SECRYPT)*, Madrid, Spain, 2017, pp. 116–128.

[26] Y. Hu, A. Yang, H. Li, Y. Sun, and L. Sun, "A survey of intrusion detection on industrial control systems," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 8, pp. 1–14, 2018.

[27] G. Murray, M. Peacock, P. Rabadia, and P. Kerai, "Detection techniques in operational technology infrastructure," in *Proc. 16th Austral. Inf. Secur. Manage. Conf.*, Perth, WA, Australia, 2018, pp. 97–105.

[28] P. I. Radoglou-Grammatikis and P. G. Sarigiannidis, "Securing the smart grid: A comprehensive compilation of intrusion detection and prevention systems," *IEEE Access*, vol. 7, pp. 46595–46620, Apr. 2019.

[29] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Inf. Softw. Technol.*, vol. 55, no. 12, pp. 2049–2075, Dec. 2013.

[30] C. W. Johnson, "Barriers to the use of intrusion detection systems in safety-critical applications," in *Computer Safety, Reliability, and Security* (Lecture Notes in Computer Science), vol. 9337, F. Koorneef and C. van Gulijk, Eds. Cham, Switzerland: Springer, 2015, pp. 375–384.

[31] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *Comput. Secur.*, vol. 28, nos. 1–2, pp. 18–28, Feb. 2009.

[32] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, Dec. 2019, Art. no. 20.

[33] K. S. Manoj, *Industrial Automation with SCADA: Concepts, Communications and Security*. Chennai, India: Notion Press, 2019.

[34] Y. Yuan and Y. Yang, *IEC 61850-Based Smart Substations: Principles, Testing, Operation and Maintenance*. Amsterdam, The Netherlands: Elsevier, 2019.

[35] Y. Geng, Y. Wang, W. Liu, Q. Wei, K. Liu, and H. Wu, "A survey of industrial control system testbeds," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 569, no. 4, 2019, Art. no. 042030.

[36] R. R. Barbosa, R. Sadre, and A. Pras, "Difficulties in modeling SCADA traffic: A comparative analysis," in *Passive and Active Measurement* (Lecture Notes in Computer Science), vol. 7192, N. Taft and F. Ricciato, Eds. Berlin, Germany: Springer, 2012, pp. 126–135.

[37] U. Adhikari, S. Pan, T. Morris, R. Borges, and J. Beaver. *Industrial Control System (ICS) Cyber Attack Datasets*. Accessed: Mar. 19, 2020. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

[38] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 303–336, 1st Quart., 2014.

[39] N. Erez and A. Wool, "Control variable classification, modeling and anomaly detection in Modbus/TCP SCADA systems," *Int. J. Crit. Infrastruct. Protection*, vol. 10, pp. 59–70, Sep. 2015.

[40] Y. Kwon, H. K. Kim, Y. H. Lim, and J. I. Lim, "A behavior-based intrusion detection technique for smart grid infrastructure," in *Proc. IEEE Eindhoven PowerTech*, Eindhoven, The Netherlands, Jun. 2015, pp. 1–6.

[41] A. Al Balushi, K. McLaughlin, and S. Sezer, "OSCIDS: An ontology based SCADA intrusion detection framework," in *Proc. 13th Int. Joint Conf. e-Bus. Telecommun.*, Lisbon, Portugal, 2016, pp. 327–335.

[42] A. Almalawi, A. Fahad, Z. Tari, A. Alamri, R. AlGhamdi, and A. Y. Zomaya, "An efficient data-driven clustering technique to detect attacks in SCADA systems," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 5, pp. 893–906, May 2016.
[43] T. Cruz, L. Rosa, J. Proenca, L. Maglaras, M. Aubigny, L. Lev, J. Jiang, and P. Simoes, "A cybersecurity detection framework for supervisory control and data acquisition systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2236–2246, Dec. 2016.
[44] E. G. D. Silva, A. S. D. Silva, J. A. Wickboldt, P. Smith, L. Z. Granville, and A. Schaeffer-Filho, "A one-class NIDS for SDN-based SCADA systems," in *Proc. IEEE 40th Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, Atlanta, GA, USA, Jun. 2016, pp. 303–312.
[45] H. R. Ghaeini and N. O. Tippenhauer, "HAMIDS: Hierarchical monitoring intrusion detection system for industrial control systems," in *Proc. 2nd ACM Workshop Cyber-Phys. Syst. Secur. Privacy (CPS-SPC)*, Vienna, Austria, 2016, pp. 103–111.
[46] R. Udd, M. Asplund, S. Nadjm-Tehrani, M. Kazemtabrizi, and M. Ekstedt, "Exploiting bro for intrusion detection in a SCADA system," in *Proc. 2nd ACM Int. Workshop Cyber-Phys. Syst. Secur. (CPSS)*, Xi'an, China, 2016, pp. 44–51.
[47] J. Zhang, S. Gan, X. Liu, and P. Zhu, "Intrusion detection in SCADA systems by traffic periodicity and telemetry analysis," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Messina, Italy, Jun. 2016, pp. 318–325.
[48] C. Feng, T. Li, and D. Chana, "Multi-level anomaly detection in industrial control systems via package signatures and LSTM networks," in *Proc. 47th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Denver, CO, USA, Jun. 2017, pp. 261–272.
[49] M. Wan, W. Shang, and P. Zeng, "Double behavior characteristics for one-class classification anomaly detection in networked control systems," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 3011–3023, Dec. 2017.
[50] K. Wong, C. Dillabaugh, N. Seddigh, and B. Nandy, "Enhancing suricata intrusion detection system for cyber security in SCADA networks," in *Proc. IEEE 30th Can. Conf. Electr. Comput. Eng. (CCECE)*, Windsor, ON, Canada, Apr. 2017, pp. 1–5.
[51] Y. Yang, H.-Q. Xu, L. Gao, Y.-B. Yuan, K. McLaughlin, and S. Sezer, "Multidimensional intrusion detection system for IEC 61850-based SCADA networks," *IEEE Trans. Power Del.*, vol. 32, no. 2, pp. 1068–1078, Apr. 2017.
[52] S. Adepu and A. Mathur, "Distributed attack detection in a water treatment plant: Method and case study," *IEEE Trans. Dependable Secure Comput.*, to be published, doi: 10.1109/TDSC.2018.2875008.
[53] Z. Ghazi and A. Doustmohammadi, "Intrusion detection in cyber-physical systems based on Petri net," *Inf. Technol. Control*, vol. 47, no. 2, pp. 220–235, 2018.
[54] A. Hijazi, E. A. El Safadi, and J.-M. Flaus, "A deep learning approach for intrusion detection system in industry network," in *Proc. 1st Int. Conf. Big Data Cyber-Secur. Intell.*, Hadath, Lebanon, 2018, pp. 55–62.
[55] H. Lin, A. Slagell, Z. T. Kalbarczyk, P. W. Sauer, and R. K. Iyer, "Runtime semantic security analysis to detect and mitigate control-related attacks in power grids," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 163–178, Jan. 2018.
[56] D. Myers, S. Suriadi, K. Radke, and E. Foo, "Anomaly detection for industrial control systems using process mining," *Comput. Secur.*, vol. 78, pp. 103–125, Sep. 2018.
[57] D. Wang and D. Feng, "Intrusion detection model of SCADA using graphical features," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Chongqing, China, Oct. 2018, pp. 1208–1214.
[58] C. Wressnegger, A. Kellner, and K. Rieck, "ZOE: Content-based anomaly detection for industrial control systems," in *Proc. 48th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Luxembourg City, Luxembourg, Jun. 2018, pp. 127–138.
[59] R. B. Benisha and S. R. Ratna, "Design of intrusion detection and prevention in SCADA system for the detection of bias injection attacks," *Secur. Commun. Netw.*, vol. 2019, Nov. 2019, Art. no. 1082485.
[60] A. Derhab, M. Guerroumi, A. Gumaei, L. Maglaras, M. A. Ferrag, M. Mukherjee, and F. A. Khan, "Blockchain and random subspace learning-based IDS for SDN-enabled industrial IoT security," *Sensors*, vol. 19, no. 14, pp. 1–24, Jul. 2019.
[61] M. Keshk, E. Sitnikova, N. Moustafa, J. Hu, and I. Khalil, "An integrated framework for privacy-preserving based anomaly detection for cyber-physical systems," *IEEE Trans. Sustain. Comput.*, to be published, doi: 10.1109/TSUSC.2019.2906657.
[62] I. A. Khan, D. Pi, Z. U. Khan, Y. Hussain, and A. Nawaz, "HML-IDS: A hybrid-multilevel anomaly prediction approach for intrusion detection in SCADA systems," *IEEE Access*, vol. 7, pp. 89507–89521, Jul. 2019.
[63] Y. Lai, J. Zhang, and Z. Liu, "Industrial anomaly detection and attack classification method based on convolutional neural network," *Secur. Commun. Netw.*, vol. 2019, pp. 1–11, Sep. 2019.
[64] H. Yang, L. Cheng, and M. C. Chuah, "Deep-learning-based network intrusion detection for SCADA systems," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Washington, DC, USA, Jun. 2019, pp. 1–7.
[65] M. Niazi, D. Wilson, and D. Zowghi, "A maturity model for the implementation of software process improvement: An empirical study," *J. Syst. Softw.*, vol. 74, no. 2, pp. 155–172, Jan. 2005.

SLAVICA V. BOŠTJANČIČ RAKAS (Member, IEEE) received the B.Sc. and M.Sc. degrees in traffic engineering and the Ph.D. degree in technical sciences from the University of Belgrade, Serbia, in 2004, 2007, and 2011, respectively. She joined the Mihailo Pupin Institute, Belgrade, in 2005, where she is currently a Research Fellow in the area of telecommunication networks. She has participated in National and International Research Projects and studies concerning design of next generation networks, quality of service, network management systems as well as cyber security of industrial control systems, such as SCADA and dynamic line rating. As an author or coauthor, she published more than 60 articles at national and international journals, books and conferences. She was the Co-Editor of the book on ICS cyber security in the Future Internet environment. Her research interests include quality of service architectures, network management in the future internet, and security issues in industrial control systems.

MIRJANA D. STOJANOVIĆ (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and the Ph.D. degree in technical sciences from the University of Belgrade, Serbia, in 1985, 1993, 2005, respectively. She held research position with the Mihailo Pupin Institute, University of Belgrade, and was involved in developing telecommunication equipment and systems for regional power utilities and major Serbian corporate systems. She is currently a Full Professor of information and communication technologies with the Faculty of Transport and Traffic Engineering, University of Belgrade. She has participated in a number of National and International Research and Development Projects, including technical projects of the International Council on Large Electric Systems, CIGRÉ. As an author or coauthor she published more than 170 book chapters, journal articles, and conference papers in her field. She was the Lead Editor of the book on ICS cyber security in the Future Internet environment. She also published a monograph on teletraffic engineering and two university textbooks (in Serbian). Her research interests include communication protocols, cyber security, service and network management, and the future internet technologies.

JASNA D. MARKOVIĆ-PETROVIĆ received the B.Sc. and M.Sc. degrees in electrical engineering and the Ph.D. degree in technical sciences from the University of Belgrade, Serbia, in 1992, 2011, 2018, respectively. She is currently with the Public Enterprise Electric Power Industry of Serbia for more than 25 years. Her activities involve implementation of the technical information system, participation in projects concerning upgrading the remote control system of the hydropower plant, and implementation of the SCADA security system. She is a member of the Serbian National CIGRÉ Study Committee D2. As an author or coauthor, she published a number of book chapters, journal articles and conference papers in her field. Her main research interests involve smart grids, SCADA and industrial control systems security, and cyber risk management.