Received May 28, 2019, accepted June 20, 2019, date of publication June 24, 2019, date of current version July 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924800

A Software-Defined Security Approach for Securing Field Zones in Industrial Control Systems

JUN YANG[1], CHUNJIE ZHOU[1], YU-CHU TIAN[2], (Member, IEEE), AND SHUANG-HUA YANG[3], (Senior Member, IEEE)

[1]National Key Laboratory of Science and Technology on Multispectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China
[2]School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, QLD 4001, Australia
[3]Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China

Corresponding author: Chunjie Zhou (cjiezhou@hust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61873103, Grant 61433006 and 61873119, and in part by the Australian Research Council (ARC) through the Discovery Project Scheme, Australia, under Grant DP170103305.

ABSTRACT Industrial control systems (ICSs) are facing increasingly severe security threats. Zone isolation, a commonly adopted idea for stopping attack propagation in general information systems, has been investigated for ICS security protection. It is usually implemented through perimeter security techniques. However, anomaly states of the physical processes in a compromised field zone may spread into other zones through the inter-zone information interaction. Due to the coupling of the physical processes between different zones, it is difficult to prevent the propagation of attack impact in ICSs. In this paper, a software-defined security (SDSec) approach is presented to address this problem. It consists of a hybrid anomaly detection module and a multi-level security response module, both of which work together to secure the ICS field zones. The hybrid anomaly detection module inspects anomaly behaviors from the perspectives of network communications and physical process states. The multi-level security response module helps prevent unapproved packets from communications, thus isolating any compromised zone. It also generates attack mitigation strategies to secure physical processes. Hardware-in-the-loop simulations are conducted to demonstrate the effectiveness of the presented approach.

INDEX TERMS Industrial control system, zone protection, software-defined security (SDSec), attack mitigation, anomaly detection.

I. INTRODUCTION

Industrial control systems (ICSs) are facing increasingly severe security threats from cyber-attacks. Due to the tight integration of cyber and physical domains, threats are also introduced into the field systems of ICSs [1]. Security hazards in physical processes may cause considerable asset damages, injuries and casualties, and/or loss of support to critical infrastructure. Therefore, security protection in ICSs becomes emerging and significant [2].

As a commonly used technology for preventing attack propagation in general information systems, zone isolation has been recently investigated for ICS security protection [3]. The international standard "IEC62443" [4] defines the concept of security zones. It also recommends the way of system segmentation for security zones. A diagram of zone-based ICS security protection is shown in Fig. 1. The perimeter protection strategy with a high security level is implemented in each zone, preventing attack propagation through network links from the compromised zone into others. When different security strategies and/or security levels are deployed in these zones, intruding several zones simultaneously by attackers becomes difficult. As a result, the probability that the whole system fails is greatly reduced [5].

However, the field zones are tightly connected to the physical processes of ICSs. Consequently, perimeter-based protection methods are not effective enough to stop the propagation of attack impacts from a compromised zone to

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

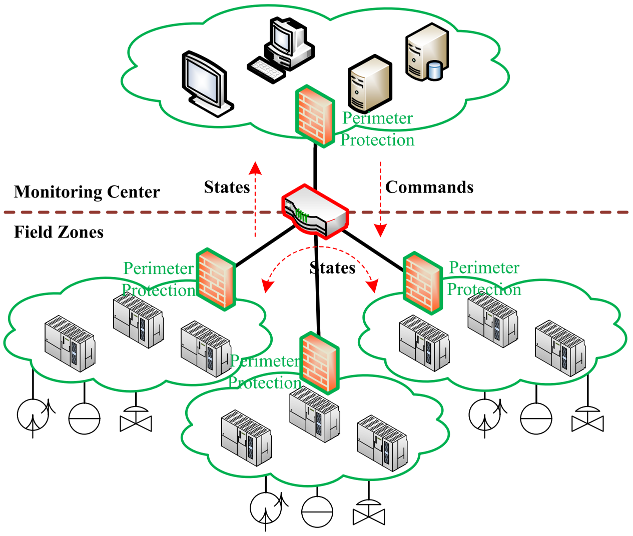


FIGURE 1. Zone-based security perimeter protection in ICSs.

other zones. This physical propagation of damages in the physical domain may cause the whole system to fail [6]. Furthermore, as shown in Fig. 1, abnormal physical states can still spread from a compromised zone to others. This may cause some incorrect adjustments or commands in other normal zones. Thus, critical process states and commands in the communications between the field zones should also be inspected dynamically and in real time. Moreover, the devices between the monitoring center and field zones are prone to be attacked. As a result, malicious commands and false states may spread everywhere. Therefore, with the tight coupling of the cyber and physical domains in ICSs, zone-based security perimeter protection is still challenging in ICSs.

To address this problem, a software-defined security (SDSec) approach is presented in this paper. It consists of a hybrid anomaly detection module and a multi-level security response module, both of which work together to secure the ICS field zones. In the closed loop of ''detection-response'' framework, the hybrid anomaly detection module inspects anomaly behaviors from the perspective of physical process states and network communications between the field zones. The multi-level security response module is used to prevent unapproved packets and mitigate attack impact on the ICS physical processes. In summary, our work in this paper makes the following contributions:

1) Considering both inter-zone communications and intra-zone physical processes, a systematic security solution including anomaly detection and security response is presented for securing the ICS field zones. It is implemented by an SDSec-based protection framework, which enables a bypassed deployment to reduce the need of redesigning or configuring the control laws in the local field zones, as well as modifying the legacy network architecture;
2) An hybrid anomaly detection mechanism that integrates multiple improved detection techniques is proposed to overcome the problem on insufficient comprehensiveness by a single detection method. With this mechanism, diversified types of anomalies can be detected, thereby enabling the security response to be more specific;
3) A multi-level security response strategy is presented to handle the attacks on both inter-zone communications and intra-zone field physical processes. Moreover, dynamic isolating the compromised zones enhances the flexibility of security protection, while adaptive regulating the physical processes enables the attack impact mitigation without the need of prescribing a limit to attack types.

The rest of this paper is organized as follows: Section II reviews related work on securing ICS field zones and introduces the concept of SDSec and its applications. The architecture of our presented approach is described in Section III. Our method for anomaly detection between field zones is presented in Section IV. This is followed by a discussion in Section V on our security response against the detected anomalies. Experimental studies are conducted in Section VI to demonstrate the presented approach. Finally, Section VII concludes the paper.

II. BACKGROUND AND RELATED WORK

This section begins with a brief review of related work on security protection for field zones in ICSs. Then, it introduces some basic concepts of SDSec.

A. SECURITY PROTECTION FOR FIELD ZONES

Efforts have been made on securing field zones in ICSs. The conception of zone-based security protection was introduced by the international standard ''IEC62443'' and the National Institute of Standards and Technology (NIST)'s ''Guide to Industrial Control Systems Security'' [7]. On the basis of this conception, methods on zone partition are designed in recent years. Machii *et al.* [8] presented a dynamical zoning methodology for safety and security protection of ICSs. Genge *et al.* [9] regarded the design requirements and security recommendations (outlined by the ''IEC62443'' standard) of ICS network as an integer linear programming problem, and designed a secure scheme based on ''zones and conduits''. Their main contribution is enhancing the security of ICS installations with the consideration of saving costs on investments. Most recently, an automatic zoning algorithm was designed for the industrial physical processes [5], where the causal model of the processes is used. These methods can be applied to the design stage of ICSs. But since the security protection of ICSs needs to cover their whole life cycle, dynamic defense techniques are required to be developed for the security of the ICSs' run stage.

After analyzing the cause-effect relationships between the physical processes, Hashimoto *et al.* [3] proposed a zoning approach to secure their two-tank testbed. It introduced a qualitative method on the anomaly analysis through the information in other normal zones. Then, this team proposed a quantitative method that was based on principle component analysis (PCA) to detect the anomalies on their testbed [10]. Mechtri *et al.* [11] used the PCA algorithm to reduce data dimension in their anomaly detection method for network communications. But PCA-based method needs to transform the dimensions of the data, it cannot point out which zones are abnormal, thereby, it may be not suitable for the following security response. Later, Yang *et al.* [5] developed a quantitative detection method based on zone partition, such that anomalies happened in the field zones could be detected. But these approaches were specific to the protection of a single zone, without considering the security problems between the different zones. From the communication security perspective, Carcano *et al.* [12] defined an industrial state modeling language to inspect the critical states of physical processes in the communication network. A filtering rule against cyber-attacks that tamper these critical states in Modbus TCP and DNP3 protocols was presented in [6]. In the work [13], an anomaly-based method was presented at the cyber-physical interaction layer to detect the integrity attacks in critical infrastructure systems. But these works were biased towards attack/anomaly detection, and didn't involve how to respond correctly to the detected exceptions.

Secure control that aims at designing resilient control laws against cyber-attacks can be used to secure the field zones in ICSs, where game theory-based and model-based are two commonly used methods [14]. In [15], Jin *et al.* developed an adaptive controller for mitigating time-varying and state-dependent attacks. Hossain *et al.* [16] regarded the attacker and the defender as a sequential Stackelberg stochastic game model, and presented a way of selecting the optimal response strategy. Although these approaches were built on the solid foundation of theories, there was a lack of studying how game-theoretical responses can be applied to real systems, while the model-based strategies in field devices need to consider low-level actions and redesign the algorithm of field controllers. This motivates our work in this paper from the perspective of SDSec protection for ICS field zones.

B. SOFTWARE-DEFINED SECURITY (SDSEC)

SDSec is a new security model in which the information security is controlled and managed by software [17]. In SDSec-based protection systems, most security strategies such as intrusion detection, authentication policy, and access control are monitored and governed through software. SDSec develops a new perspective to design, deploy and implement security measures [18], and it is playing a growing role in information security. Qiu *et al.* [19] proposed a security controller-based SDSec architecture, which could be used to interact with other security components and protect the network devices. In order to solve the problem of resource consumption and security protection efficiency, the work in [20] presented a resource scheduling algorithm to allocate security tasks.

Recently, the concept of SDSec is also introduced to industry systems. For instance, a software-defined network (SDN)-based communication method was presented to enhance the cyber-security and resilience of a campus microgrid [21]. The works in [22] gave a security monitoring and control approach for IEC 61850-based communication systems by using SDN. But there was a lack of studying how to mitigate the attack impacts in physical systems. Genge and Haller [23] developed a hierarchical SDN control plan for addressing the requirements of ICS communication infrastructure. Molina et al. in [24] surveyed the current works of using SDN technique for ICS network protection. But these studies were proactive response approaches, they could not identity what type of field devices were compromised and react as soon as an attack was detected. In [25], Piedrahita *et al.* described a prospect of intrusion response solution in ICSs by leveraging the SDN and network functions virtualization (NFV) techniques, but it did not suggest a specific method.

In summary, SDSec is an attractive tool for security protection. This motivates us to leverage this technique to secure the field zones of ICSs. However, with consideration of the inherent features of the ICSs discussed before, several SDSec functions are required to be well designed when applied to ICSs. Two typical scenarios are: i) Malicious behaviors on both inter-zone communications and physical processes need to be identified and prevented; ii) Suitable securing strategies should be generated to respond to the detected cyber-attacks and/or physical damages.

III. ARCHITECTURE OF SDSEC-BASED PROTECTION APPROACH

Our SDSec-based approach aims at providing a high-level security method to protect the communications between the field zones, as well as securing the physical processes in compromised zones. In the respect of protecting inter-zone communications, it puts efforts into detecting the anomalies of the traffic and/or communication state, thereby isolating the malicious packets between the zones. while in the respect of securing intra-zone physical processes, it makes efforts to inspect the outliers of physical states, thereby mitigating the attack impacts in compromised zones. It consists of two modules: hybrid anomaly detection and multi-level security response, and the approach will be embedded into the communication device over the ICS field zones. The architecture of the presented SDSec protection approach is depicted in Fig. 2. In this architecture, the hybrid anomaly detection module is responsible for inspecting the abnormal behaviors in all the field zones, while the multi-level security response module is in charge of making security strategies to secure the physical processes from the perspective of stopping attack impact propagation.

A. HYBRID ANOMALY DETECTION

In this module, attacks on both the inter-zone communications and intra-zone physical processes are considered. Initially, packages between ICS field zones are mirrored from

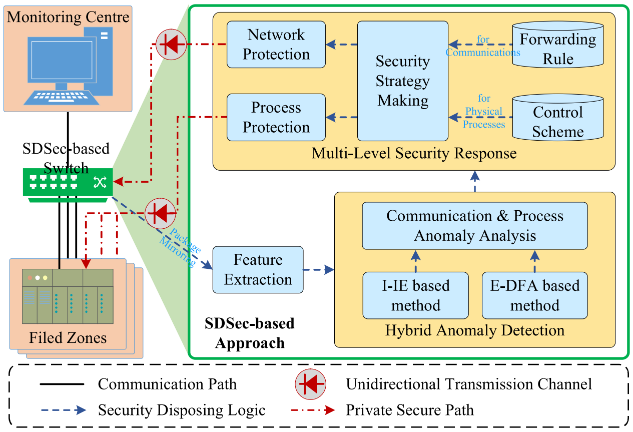


FIGURE 2. Architecture of SDSec-based approach for securing field zones.

the communication device. And then, an improved information entropy (I-IE) based method and an extended deterministic finite automaton (E-DFA) method are proposed, both of which are used to analyze industrial network anomalies (such as illegal access, unexpected packages, and flow burst event etc.) and physical process anomalies (such as malicious commands, anomaly states and so on). If attack behaviors are confirmed, the proposed approach will implement the corresponding protection measures to secure the communication network and mitigate the impact of the attacks on the ICS physical system.

B. MULTI-LEVEL SECURITY RESPONSE

The multi-level security response strategy is proposed for attack mitigation and physical safety protection. When a malicious behavior is detected, a ''packet drop'' rule will be defined for the compromised links. This helps to cut off the corresponding inter-zone communication paths, preventing attack propagation through the cyber domain. If the physical processes are abnormal, control strategies will be sent to the devices in the compromised zones for attack mitigation.

In addition, since the SDSec-based system itself is an important part of the whole ICS architecture, it also needs to be protected. In our approach, two security measures are implemented to secure the system: 1) As the SDSec-based system has no communication task, all the packets whose destination address is the system should be directly dropped. We leverage a packet redirection (mirroring) technology to send ingress packets to the system. 2) A specialized unidirectional transmission channel is built to send the security response strategy to field devices such that malicious packets cannot intrude into the SDSec-based system.

Detailed developments of the components of our SDSec-based approach will be presented in the following sections.

IV. ANOMALY DETECTION BETWEEN FIELD ZONES

This section develops a hybrid anomaly detection method. It begins with an introduction of the framework of the method. This is followed by discussions of two anomaly detection mechanisms.

A. FRAMEWORK OF THE ANOMALY DETECTION METHOD

As most of ICSs are often resource-constrained and time-sensitive, security protection methods for ICSs should be as simple as possible in the case of meeting the protection requirements, especially for industrial field processes [26]. Compared with other statistic based and machine learning based methods, information entropy based method may be simpler in the aspect of resource consumption, implementation complexity, and so on. Moreover, since the mode of industrial communication network is relatively fixed and the traffic is not heavy in most of ICSs, information entropy based method has the ability to detect anomalies of network [27]. Also, as state transition is a key factor to describe the behaviors of industrial communication, building the state machine of the protocol helps detect the anomalies, such as a message appearing out of its position in the normal sequence or a message referring to a single unexpected bit [28]. But it may be difficult for I-IE based method to these anomalies. Therefore, a hybrid anomaly detection method is presented with an extended deterministic finite automaton (E-DFA) metric for detecting the anomalies of communication states and physical process states.

Fig. 3 shows the workflow of our anomaly detection method. It is composed of three modules: feature extraction, I-IE anomaly detection and E-DFA anomaly detection. In the feature extraction module, received packets are grouped firstly, then traffic features and behavior features are respectively extracted. The traffic features include IP address, port number, and MAC address [30], while the behavior features include query/response, function code, and others. In I-IE detection module, the entropy of the received packets is computed. Then, anomaly detection is conducted by comparing the computed result with its threshold. There are two phases for E-DFA based anomaly detection method: training phase and running phase. In the training phase, packets in a normal system are collected to construct an E-DFA reference model. In the running phase, a real-time E-DFA model is matched with the reference model. Anomalies on communication state and/or application data are derived from the matching results. Besides, in order to determine the anomaly zone or link, and reduce the complexity of the model, both the I-IE model and E-DFA model are built for the anomaly detection on the communications between any two zones, and the received packets should be split into separated channels.

B. I-IE ANOMALY DETECTION

Let $\mathscr{N}_\tau$ denote the number of received packets in the interval of τ and $\mathscr{A} = \{a_1, a_2, \ldots, a_m\}$ represent the traffic feature set, where a_i is the i^{th} feature and m indicates the types of the features. A binary symbol $\mathscr{X}_{ij}$ denotes the existence of a_i in the j^{th} packet, where 1 means 'yes' and 0 means 'no'. The proportion $p(a_i)$ of the quantity of a_i in $\mathscr{N}_\tau$ can be calculated by

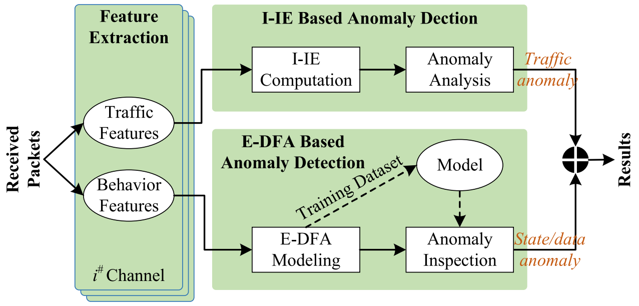


FIGURE 3. Framework of our anomaly detection.

$p(a_i) = \sum_{j=1}^{\mathcal{N}_\tau} \mathcal{X}_{ij}/\mathcal{N}_\tau$. Therefore, the information entropy of the feature set $\mathscr{A}$ is represented as:

$$H(\mathscr{A},\tau) = -\sum_{i=1}^{m} p(a_i)\log p(a_i). \tag{1}$$

This information entropy reflects the distributions of the traffic features. Once an attack against the network and/or devices happens, e.g., a Denial of Service (DoS) attack or an IP/port scan attack, the value of the information entropy will be quite different from that under normal conditions [28]. Therefore, if a bound of $H(\mathscr{A},\tau)$ is confirmed by a training set obtained in normal conditions, it can be used as the anomaly detection threshold [31].

However, there are some shortcomings in Eq. (1), e.g. if an attack just makes the number of some features to be exchanged, i.e. $p(a_i) = \tilde{p}(a_j), p(a_j) = \tilde{p}(a_i)$, this anomaly will not be detected, where $\tilde{p}$ denotes the proportion under an anomaly condition. Moreover, it is also difficult to handle the scenario $\sum_{j=1}^{\mathcal{N}_{\tau-1}} \mathcal{X}_{ij} = k\sum_{j=1}^{\mathcal{N}_\tau} \tilde{\mathcal{X}}_{ij}, (k \neq 0)$, where $\tilde{\mathcal{X}}$ means the binary symbol under a normal condition [32].

The following I-IE is presented to address these issues:

$$H'(\mathscr{A},\tau) = -\sum_{i=1}^{m} \omega_i p(a_i)\log p(a_i) + \alpha\frac{\min\{\mathcal{N}_{\tau-1},\mathcal{N}_\tau\}}{\max\{\mathcal{N}_{\tau-1},\mathcal{N}_\tau\}}, \tag{2}$$

where set $\boldsymbol{\omega} = \{\omega_i | i = 1 \text{ to } m\}$ denotes the weights of the features, and α is an adjustment coefficient. Here, the weights are used for handling the first scenario $p(a_i) = \tilde{p}(a_j), p(a_j) = \tilde{p}(a_i)$. Therefore, the values of the weights should be different, i.e. $\forall i, j,\ \omega_i \neq \omega_j$. In our approach, random values in [0, 1] are used for the weight assignment. The proportion between $\mathcal{N}_{\tau-1}$ and $\mathcal{N}_\tau$ is used for handling the second scenario $\sum_{j=1}^{\mathcal{N}_{\tau-1}} \mathcal{X}_{ij} = k\sum_{j=1}^{\mathcal{N}_\tau} \tilde{\mathcal{X}}_{ij}$. Since that there are no obvious change the traffics in normal conditions, the coefficient α can be set to 1.

In addition, as the bound of $H(\mathscr{A},\tau)$ in Eq. (1) is in $[0, \log m]$, it is usually normalized to vary from 0 to 1, i.e.:

$$H(\mathscr{A},\tau) = -\sum_{i=1}^{m} p(a_i)\log p(a_i)/\log m. \tag{3}$$

In this paper, the concept of normalization is also introduced in Eq. (2), i.e.,

$$H'(\mathscr{A},\tau) = -\sum_{i=1}^{m} \frac{\omega_i p(a_i)\log p(a_i)}{\log m} + \alpha\frac{\min\{\mathcal{N}_{\tau-1},\mathcal{N}_\tau\}}{\max\{\mathcal{N}_{\tau-1},\mathcal{N}_\tau\}}. \tag{4}$$

However, there may be a certain fluctuation of the network traffic, for instance, ARP packets timing triggering, packet resending, network delay and so on. Therefore, the detection threshold of I-IE based method should be set to a certain range. In this paper, a statistical method named exponentially weighted moving average (EWMA) control chart [33] is adopted, which is defined as:

$$E(\mathscr{A},\tau) = (1-\lambda)E(\mathscr{A},\tau-1) + \lambda H'(\mathscr{A},\tau), \tag{5}$$

where λ $(0 < \lambda < 1)$ is the weight factor, its value is usually set between 0.2 to 0.3, and $E(\mathscr{A}, 0) = H(\mathscr{A}, 0)$. The upper control limit (UCL) and lower control limit (LCL) are defined by

$$\text{UCL} \setminus \text{LCL} = \mu \pm \ell\sigma\sqrt{\frac{\lambda}{2-\lambda}\left(1-(1-\lambda)^{2\tau}\right)}, \tag{6}$$

where μ and σ are the expectation and deviation of the residuals respectively, ℓ is the control limit width. As $0 < \lambda < 1$, $(1-\lambda)^{2\tau} \simeq 0$ with τ increasing, Equation (6) can be converted to

$$\text{UCL} \setminus \text{LCL} = \mu \pm \ell\sigma\sqrt{\frac{\lambda}{2-\lambda}}. \tag{7}$$

If the EWMA value $E(\mathscr{A},\tau)$ is larger than UCL or smaller than LCL, the abnormal features would be determined.

C. E-DFA ANOMALY DETECTION

Because general DFA is incapable of checking the physical process states and commands, an E-DFA based approach is proposed, and both the communication states and the application data are inspected. A tuple is defined as follows to describe extended communication behaviors:

$$\Lambda \overset{def}{=} \langle \text{state} : \text{data} \rangle \overset{def}{=} \langle \sigma, \delta : b, D \rangle, \tag{8}$$

where σ denotes an input symbol, δ means a function code, b represents the start address of the application data in memory block of the protocol, and $D = \{\langle \text{addr}^i, d_1^i, d_2^i \rangle,\ i = 1 \text{ to } \ell\}$ represents the lower and upper thresholds (d_1^i, d_2^i) in the address addr^i. More precisely, the input symbol σ denotes that whether the state of the current packet is "query" or "response". The function code δ denotes protocol's function. For example, δ includes "input register", "holding register" and so on in Modbus TCP, while δ includes "data I/O", "PLC control", "Block oriented" and so on in Siemens S7. The E-DFA is defined as:

$$\text{E-DFA} \overset{def}{=} \{\langle \boldsymbol{\Lambda}, \boldsymbol{q} \rangle \mid \langle \Lambda^i, q^i \rangle, i = 1 \text{ to } \mathcal{N}\}, \tag{9}$$

where q^i is the sequence in a set of packets, $\mathcal{N}$ is the number of behaviors. e.g., $\langle \Lambda^6, \{5, 13\} \rangle$ means that the 5th and 13th received packets belong to a same behavior Λ^6.

Algorithm 1 E-DFA Modeling Algorithm.

```
In: $f = \{f^i | i = 1 \text{ to } \kappa\}$.
Out: E-DFA
E-DFA ← ∅
Screen out the industrial protocol traffic from $(f) \rightarrow \mathscr{F}$
for each $\tilde{f}^i$ in $\mathscr{F}$ do
    for each $\Lambda^j$ in $\mathbf{\Lambda}$ do
        if $\tilde{f}^i$.state $= \Lambda^j$.state then
            $\Lambda^j \leftarrow \tilde{f}^i$, $q^j \leftarrow q^j \cup i$
        end if
    end for
    if $\tilde{f}^i$.state $\notin \mathbf{\Lambda}$ then
        E-DFA ← E-DFA $\cup < \tilde{f}^i, i >$
        $\mathscr{N} \leftarrow$ length of E-DFA
        for each $y^k$ in $\tilde{f}^i$.data do
            $\Lambda^{\mathscr{N}}.d_1^k \leftarrow \underline{y}^k$, $\Lambda^{\mathscr{N}}.d_2^k \leftarrow \bar{y}^k$,
        end for
    end if
end for
```

Modeling the E-DFA by manual means is time consuming, especially for a complex system, and thus, an automatic modeling algorithm is designed in our method. The procedure of building an E-DFA model for a group of packets is shown in Algorithm 1, where $\boldsymbol{f}$ denotes the ingress packets, $\mathbf{\Lambda}$ is a temporary set for the communication behaviors and '∅' represents an empty set. The modeling procedures are shown as follows:

Step 1: Remove the packets that do not belong to industrial protocol.

Step 2: If the current communication behavior $\tilde{f}^i$.state has happened before, record the sequence number.

Step 3: If the current communication behavior $\tilde{f}^i$.state does not happen, add a new behavior into the set $\mathbf{\Lambda}$. Moreover, the lower and upper thresholds $\underline{y}^k$, $\bar{y}^k$ for each application data y^k should be added into the set, respectively.

In the training phase, the application data of each packet is inspected firstly, and the algorithm is shown in Algorithm 2, where $\mathbf{\Theta}$ is a temporary set for the application data, and Ψ is a set for the thresholds. In Algorithm 2, the abnormal states of physical processes can be found out.

Since the communication behaviors in ICSs are usually periodic [28], [34], the length of the period can be taken as the E-DFA matching length. But there are some special situations need to be considered in reality, such as:

1) Packet retransmission: a same packet is sent repeatedly at two adjacent instant, i.e. $f^i = f^{i+1}$;
2) New protocol-based packet: this typically occurs in some manual operations.

Therefore, a handful of additional packets should be allowed in E-DFA based anomaly detection module. A maximum allowed number of protocol packets $\mathscr{N}_{\max}$ is defined:

$$\mathscr{N}_{\max} = \mathscr{N}_T + \mathscr{N}_o, \tag{10}$$

Algorithm 2 Data Inspection Algorithm

```
In: $f$, E-DFA reference model.
Out: Result
Result ← ∅
Extract physical states in $f \rightarrow \mathbf{\Theta}$
Find the thresholds of $\mathbf{\Theta}$ in E-DFA $\rightarrow \Psi$
$\ell \leftarrow$ length of $\mathbf{\Theta}$
for $i \leftarrow 1$ to $\ell$ do
    if $\Theta^i < \Psi.d_1^i$ or $\Theta^i > \Psi.d_2^i$ then
        Result ← Result ∪ "data anomaly"(in $\Psi$.addr$^i$)
    else
        Result ← Result ∪ "data normal"(in $\Psi$.addr$^i$)
    end if
end for
```

where $\mathscr{N}_T$ means the length of the reference E-DFA model, $\mathscr{N}_o$ denotes the allowed changes in a period.

The E-DFA matching algorithm is shown in Algorithm 3, where N.state^i means the state in "New E-DFA", while R.state^i denotes the state in "Reference E-DFA". Initially, the length of the "New E-DFA" is analyzed, and if $\mathscr{N}_f > \mathscr{N}_{\max}$, the state will be regarded as "state anomaly". Then, the state sequence is inspected. If the number of unknown state is larger than $\mathscr{N}_o$, the anomaly is determined. The termination condition of the matching loop is that "state anomaly" is found or all packets are analyzed.

In addition, if the operation conditions are changed, the reference E-DFA model should be re-modeled.

V. ZONE SECURITY RESPONSE

Various anomalies may be found out through the presented anomaly detection approach. Thus, a hybrid method is needed to accurately cope with the results. This section presents a multiple response mechanism to synthetically handle compromised zones and the corresponding links. It begins with an overview of the method. Then, an inter-zone communication protection mechanism and an attack impact mitigation method of physical processes are developed.

A. OVERVIEW OF THE MULTI-LEVEL SECURITY RESPONSE

Directly isolating abnormal field zones is not always a good way when some anomalies are detected. This is because some attacks are not against the inter-zone communications, but the physical processes in the field zones. Our approach presented in this paper is able to detect three types of anomalies: i) traffic anomaly, ii) state anomaly, and iii) data anomaly. The first type of anomaly is detected by I-IE based method, while the last two types are detected by E-DFA based model.

For the first two types of anomalies, there is a great possibility that the communication is invaded. Thus, the related zones should be isolated for security protection. However, handling the communication packets may be not an effective way for the "data anomaly", which needs to be treated by sending some suitable control strategies. For the anomalies on system commands, we can design a pre-filter module to

Algorithm 3 E-DFA Matching Algorithm

```
In: $f = \{f^i | i = 1 \text{ to } \kappa\}$, E-DFA reference model.
Out: Result
Build the E-DFA model for $f$ → "New E-DFA"
$\mathcal{N}_f \leftarrow$ length of "New E-DFA"
if $\mathcal{N}_f > \mathcal{N}_{\max}$ then
    Result ← "state anomaly"
else
    $m \leftarrow 1,\ \ell \leftarrow 0$
    for $j \leftarrow 1$ to $\mathcal{N}_T$ do
        for $i \leftarrow m$ to $\mathcal{N}_f$ do
            if $N.\text{state}^i = R.\text{state}^j$ then
                Result ← "state normal", $m \leftarrow i + 1$
                Break this loop
            else if $i = \mathcal{N}_f$ then
                $\ell \leftarrow \ell + 1$
                if $\ell \leq \mathcal{N}_o$ then
                    $m \leftarrow m + 1$
                else
                    Result ← "state anomaly"
                    Stop matching
                end if
            end if
        end for
    end for
end if
```

TABLE 1. Anomaly types and corresponding security strategies.

No.	Anomaly types	Detected by	Security strategies
1	traffic anomaly	I-IE	Isolation
2	state anomaly	E-DFA	Isolation
3	data anomaly	E-DFA	Mitigation

drop these malicious packets and send the original commands to the field devices. Therefore, in the following part for responding the "data anomaly", we will focus on mitigating the attack impacts when the physical state anomalies in a field zone are detected. TABLE 1 summarizes possible anomaly types and corresponding strategies to treat them. In the table, "Isolation" indicates that the malicious packets should be dropped, thus isolating the corresponding devices or zones. "Mitigation" indicates that suitable control strategies need to be sent to the related devices of the field zones.

B. SECURITY MANAGEMENT FOR INTER-ZONE COMMUNICATIONS

This part is designed to handle the first two anomalies shown in TABLE 1. In the SDSec-based Switch, packet-forwarding rules are defined by a security table that consists of a list of items. An item of the security table includes three basic elements: match fields, priority and instructions, which is used to identify the specific packets, as shown in Fig. 4.

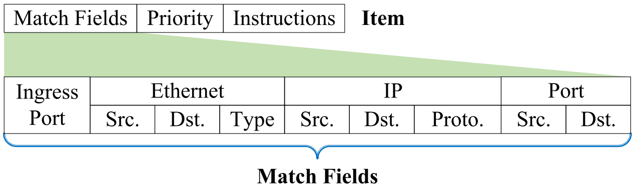


FIGURE 4. The structure of an item.

The detailed explanations of an item in the security table are shown as follows:

1) Match Fields: it defines the contents that need to match against an ingress packet, the matching contents include ingress port, Ethernet source/destination address, Ethernet type, IP source/destination address, IP protocol, source port and destination port etc.;
2) Priority: it defines the matching precedence of the item when the switch receives an ingress packet; and
3) Instructions: it defines a set of instructions that are executed when a packet matches the item, and the execution instruction includes packet apply-action ("forward", "drop" or "sending to anomaly detection module") and write-action (updating the instruction for apply-action).

A blacklist (BL) and a whitelist (WL) are defined in the switch, where the BL denotes the untrusted devices whose transmission must be prohibited, and the WL denotes the trusted devices that can be transferred. Their structures are defined as:

BL- match: in_port = i, address = <Src. / Dst.>; # priority; instruction: drop.

WL- match: in_port = i, address = <Src. & Dst.>; # priority; instruction: out_port = j.

In the BL, the match field only covers with the source (Src.) address or destination (Dst.) address of a packet, which means that illegal devices are unable to send or receive any packets. In this way, the network communications in a compromised zone can be isolated. In the WL, both the Src. address and the Dst. address are required to match, which means that both sides of the link must be confirmed before the packet can be forwarded. Besides, there are two ways for generating the BL and the WL, one is through predefining, and the other is using the anomaly detection results to update. Fig. 5 shows the flow diagram of request packet processing in the SDSec-based Switch.

When a packet f arrives, it successively matches with the BL and the WL, $f \in$ BL represents the packet should be dropped, while $f \in$ WL means it can be forwarded. If no item is matched in neither the BL nor the WL, the packet will be stored in a buffer and wait for anomaly analysis. In each detection cycle T, the stored packets will be analyzed through the proposed detection approach. Once the entropy value of these packets is out of range (by I-IE method), or the states in the E-DFA model are abnormal, the BL is updated after tracing the Src. address of the abnormal packets.

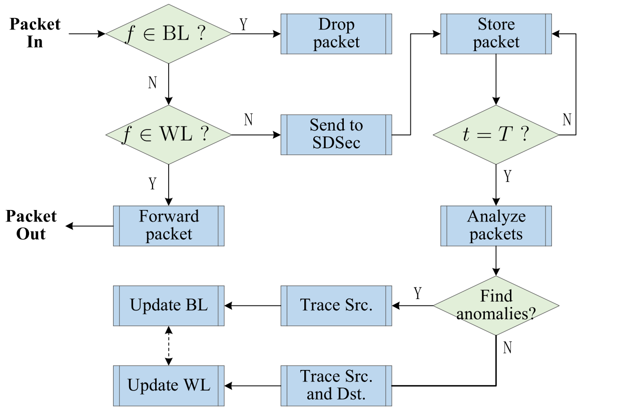


FIGURE 5. Flow diagram of request packet processing in the SDSec-based switch.

Otherwise, the WL will be updated. Meanwhile, it is required to make sure that a same address can not be in the BL and the WL simultaneously. In addition, as both the I-IE model and E-DFA model are built for the communications between each two zones, it facilitates the address tracing when anomalies are found.

Since that some cyber-attacks, such as UDP Flooding, ICMP Flooding, and TCP SYN Flooding, may block communications or make the resource of the switch unavailable [30], in order to prevent such attacks, incoming packets that match with the BL are dropped directly, and the BL matching is prior to the WL.

It should be noted that there is no need to add the packet address into the BL when only "data anomaly" is detected by E-DFA based method, because the outliers may be not caused by cyber-attacks in network communications. Moreover, network isolation also cannot prevent the influence propagation through the coupling of physical processes. Therefore, anomalies on the application data are not considered in Fig. 5, and additional security strategies should be developed.

C. ATTACK MITIGATION OF PHYSICAL PROCESSES

This part is designed to deal with the "data anomaly" of the physical processes. Our mitigation strategy generates and sends a compensation signal into the controller in the field zones to be against these anomalies. Since the attack behaviors may be time-varying and random, the security strategy also requires dynamic regulation ability.

Fig. 6 shows the schematic of the attack mitigation in ICS field zones, where y_a^i and u_a^i represent the attacks in the i^{th} field zone, ξ (ξ^i) denotes the compensation signal. The attack behavior analysis and our countermeasures are shown as follows:

1) ATTACK BEHAVIOR ANALYSIS

In the i^{th} field zone, let y^i denote the sensor measurements, u^i represent the control inputs. In order to break down the physical processes, attacks usually attempt to tamper the sensor

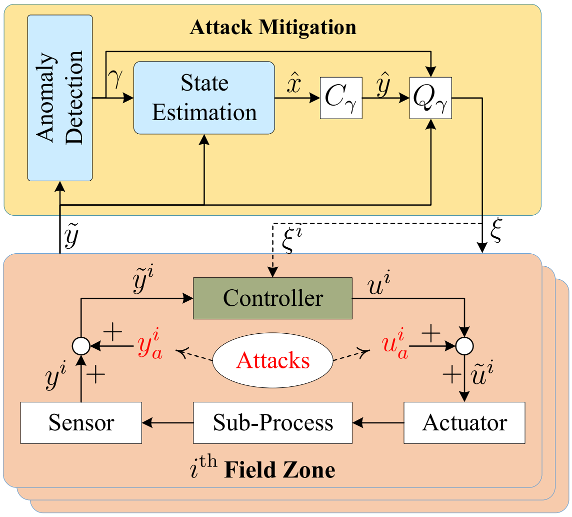


FIGURE 6. Schematic of the attack mitigation of physical processes.

measurements and/or control inputs. Thus, the received sensor signals in the controller are changed to be $\tilde{y}^i$, while the arrived control signals in the actuators are tampered into $\tilde{u}^i$. Since these two attack scenarios can be transformed into a same problem for the closed-loop analysis in the controller (the impact of these two attacks differ by the feedback matrix), only y_a^i is considered here. The method can be easily extended to the attacks on control inputs. The compromised sensor measurements can be described as:

$$\tilde{y}^i(t) = y^i(t) + y_a^i(t), \quad t > t_a, \tag{11}$$

where t_a denotes the attack instant. In order to mitigate the attack impacts and prevent other normal zones from being affected, it needs to give a compensation signal to the controller such that the control inputs are generated for the actual states y^i. Therefore, the compensation signal in the i^{th} field zone should satisfied:

$$\xi^i(t) = -y_a^i(t), \quad t > t_a. \tag{12}$$

2) MITIGATION STRATEGY GENERATION

In our approach, the compensation signal is generated in the SDSec-based system, which has a global perspective of the field zones. Thus, we can use a global model for the mitigation strategy generation. Assume that the physical processes are a linear time-invariant (LTI) system, which is modeled as:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + w(t), \\ y(t) = Cx(t) + v(t), \\ u(t) = Ky(t). \end{cases} \tag{13}$$

where $x(t) \in \mathbb{R}^n$ denotes system states, $u(t) \in \mathbb{R}^m$ represents control inputs, $y(t) \in \mathbb{R}^l$ means the sensor measurements; $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{l\times n}$ are system matrix, input matrix and output matrix respectively; $w(t) \in \mathbb{R}^n$ and $v(t) \in \mathbb{R}^l$ are zero-mean Gaussian white noise, $w(t) \sim \mathcal{N}(0, Q_w)$

and $v(t) \sim \mathcal{N}(0, Q_v)$, respectively; K denotes the output feedback matrix.

Assumption 1: The LTI system (A, B, C) described in Eq. (13) is controllable and observable.

With an attack value $y_a(t)$ and a compensation signal $\xi(t)$, the closed-loop system can be represented as:

$$\begin{cases} \dot{x}(t) = A_{cl}x(t) + B_{cl}\big(y_a(t) + \xi(t)\big) + M \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}, \\ y(t) = Cx(t) + v(t), \end{cases} \tag{14}$$

where $A_{cl} = A + BKC$, $B_{cl} = BK$, and $M = [I, BK]$, I is an identity matrix.

To generate a compensation signal against the attack value, the actual states $y(t)$ should be observed. However, when outliers are found in a zone, we can not directly obtain the observations in this zone from the sensor measurements. To address this problem, a switched Kalman filter [35] is adopted in this paper, in which the healthy sensor measurements are used to predict the actual system states. Initially, the anomaly detection module will generate an abnormal label γ for our state estimation module, where $\gamma = \mathrm{diag}\{\gamma^1, \gamma^2, \ldots \gamma^N\}$ denotes a switching diagonal matrix, $\gamma^i \in \{0, I\}$, $(1 \leq i \leq N)$, N is the zone amount. $\gamma^i = I$ denotes the i^{th} zone is normal, while $\gamma^i = 0$ means the i^{th} zone is abnormal. Then, the switched Kalman filter can be given by:

$$\begin{cases} \dot{\hat{x}}(t+1|t) = A_{cl}\hat{x}(t), \\ \hat{x}(t) = A_{cl}\hat{x}(t|t-1) + K_\gamma(t)\big(y_\gamma(t) - C_\gamma\hat{x}(t|t-1)\big). \end{cases} \tag{15}$$

where $y_\gamma(t) = \gamma\tilde{y}(t)$ denotes the healthy measurements, and C_γ is constructed from the matrix C through removing the rows related to the abnormal measurements, $K_\gamma(t)$ is the Kalman gain.

$$\begin{cases} K_\gamma(t) = P^*_\gamma(t)C^T_\gamma(t) \times \big(C_\gamma P^*_\gamma(t)C^T_\gamma(t) + R_{1\gamma}\big)^{-1}, \\ P_\gamma(t) = (I - K_\gamma(t)C_\gamma)P^*_\gamma(t), \\ P^*_\gamma(t) = A_{cl}P_\gamma(t-1)A^T_{cl} + R_{2\gamma}, \end{cases} \tag{16}$$

where $R_{1\gamma}, R_{2\gamma}$ are the process noise covariance matrix and the measurement noise covariance matrix, respectively. It should be noted that, Eq. (15) can be used to take the prediction step for the system states if and only if the observability matrix

$$O_\gamma = \big[C_\gamma \; C_\gamma A_{cl} \; \ldots \; C_\gamma A^{(n-1)}_{cl}\big]^T$$

has full rank. Therefore, redundant sensors need to be deployed if the observable condition is not satisfied, but this issue is beyond the focus of this paper.

For the prediction state $\hat{x}(t|t-1)$, the compensation signal $\xi^i(t)$ is given by

$$\xi^i(t) = Q_\gamma(t)\big(C\hat{x}(t|t-1) - \tilde{y}(t)\big), \quad \gamma^i = 0. \tag{17}$$

Here, $Q_\gamma(t)$ is also a diagonal matrix based on γ, $Q^i_\gamma(t) = 0$ when $\gamma^i = I$, while $Q^i_\gamma(t) = I$ when $\gamma^i = 0$. In this way, if the sensor measurements in the i^{th} zone is compromised, we can send such $\xi^i(t)$ to correct the attacked measurements. Since the compensation signal $\xi^i(t)$ is adaptive, our approach does not need to prescribe a limit to attack types, and it even has the ability to against time-varying and random attacks. Moreover, the solution over the ICS field zones can reduce the need of redesigning or configuring the control laws of the controller in each field zone.

Let $e(t) = x(t) - \hat{x}(t)$ denote the estimation error. Based on the observation in Eq. (15), the closed-loop system dynamics in Eq. (14) can be represented as:

$$\dot{x}(t) = A_{cl}x(t) - B_{cl}Q_\gamma(t)Ce(t) + B_{cl}\big(I - Q_\gamma(t)\big)y_a(t) + \big(M - B_{cl}Q_\gamma(t)\big)\begin{bmatrix} w(t) \\ v(t) \end{bmatrix}. \tag{18}$$

Since $y^j_a(t) = 0$ $(j \in [1, N],\ j \neq i)$ and $Q^i_\gamma(t) = I$, there exists $\big(I + Q_\gamma(t)\big)y_a(t) = 0$. Moreover, the proof of lemma 4 in literature [35] is shown that the deterministic part of the error $e(t)$ is asymptotically stable. Therefore, according to Assumption 1, the closed-system is bounded and asymptotically stable.

VI. EXPERIMENTAL STUDIES

This section conducts experimental studies to demonstrate our presented SDSes protection approach. It begins with construction of a hardware-in-the-loop (HIL) simulation system and a corresponding E-DFA model. Then, typical attack scenarios are designed on the system. After that, it shows how the proposed approach is used to detect these attacks and manage compromised zones.

A. EXPERIMENTAL DESIGN

In our experimental part, the SDSec-based approach is embedded in an OpenFlow switch. The OpenFlow switch is a software program device which forwards packets in a software-defined networking (SDN) environment. It centralizes network intelligence by decoupling the routing process (control plane) from the forwarding process of network packets (data plane), enabling the network to be programmable dynamically. Thereby, our security approach can be deployed in the control plane, and network security strategy execution can take advantage of the data plane. The HIL simulation system is depicted in Fig. 7.

It consists of an HMI, an SDSec-based System, three CNx and a simplified TEP [36] Simulation Host. The SDSec-based System consists of an OpenFlow controller (named Security Host, SH) and an OpenFlow switch. The SH is simulated by Mininet in Linux system, which is running on an embedded system with ARM Cortex-A8 processor, 512MB RAM and 8GB flash memory, while the switch is simulated by a soft route with 6 LAN port, Core i5 processor, 4GB RAM and 128G ROM. The CNx also uses an embedded processor, which has the same hardware configuration with the SH. The HMI, the three CNxs and the SH are linked to the OpenFlow switch through the LAN port, while the link (named Safety Management Channel, SMC) between the SH and the three CNxs adopts a CAN-bus. Besides, the TEP

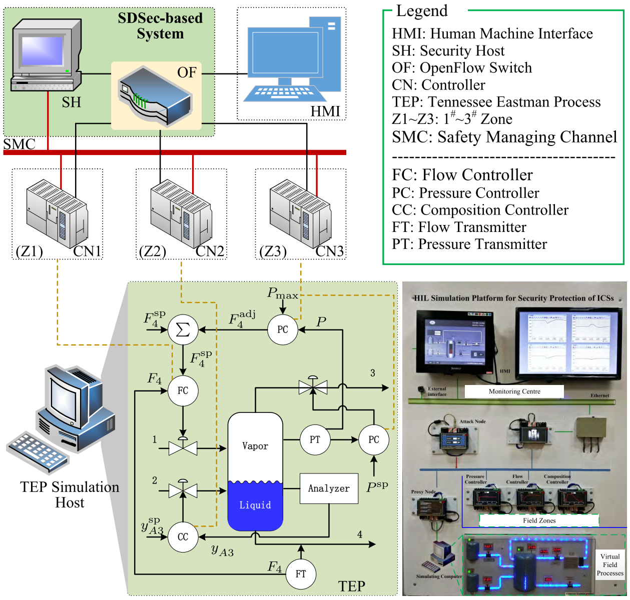


FIGURE 7. A HIL simulation system.

Simulation Host and the three CNxs are linked also through a CAN-bus.

The objective of the control of the simplified TEP is to maintain the product rate F_4 at 100kmolh^{-1} with the pressure P at 2700kPa and the ratio of a reactant in the purge y_{A3} at 47mol%. Three feedback loops for the control of F_4, y_{A3} and P are distributed in Z1, Z2 and Z3, respectively. The HMI is in charge of adjusting the set-points of inputs 1 and 2. It also monitors all states of the simplified TEP. Modbus TCP protocol is used for the communication between HMI and CNx.

The simplified TEP dynamics is given by [37]:

$$A=\begin{bmatrix} -1.333 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & -11 & -2.5 & 0 & 0 & 0 & 0\\ 0 & 4 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & -0.8 & 0 & 0\\ 0 & 0 & 0 & 0.5 & -4.1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & -20.1 & -2\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

$$B=\begin{bmatrix} 2 & 0 & 0 & 0\\ 0 & 0 & 0 & 8\\ 0 & 0 & 0 & 0\\ 4.5 & 0 & -1.125 & 0\\ 12.75 & 0 & -0.75 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$C=\begin{bmatrix} 1.333 & -4.25 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 8 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & -0.075 & 1.5 \end{bmatrix},$$

$$w(t)\sim\mathcal{N}(0,0.02^2),\quad v(t)\sim\mathcal{N}(0,0.05^2).$$

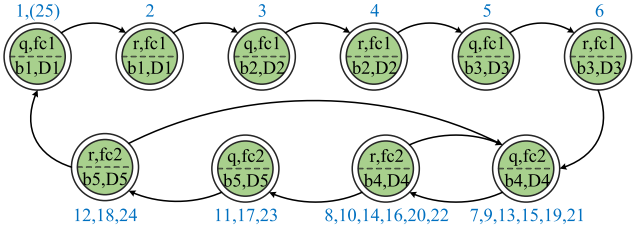
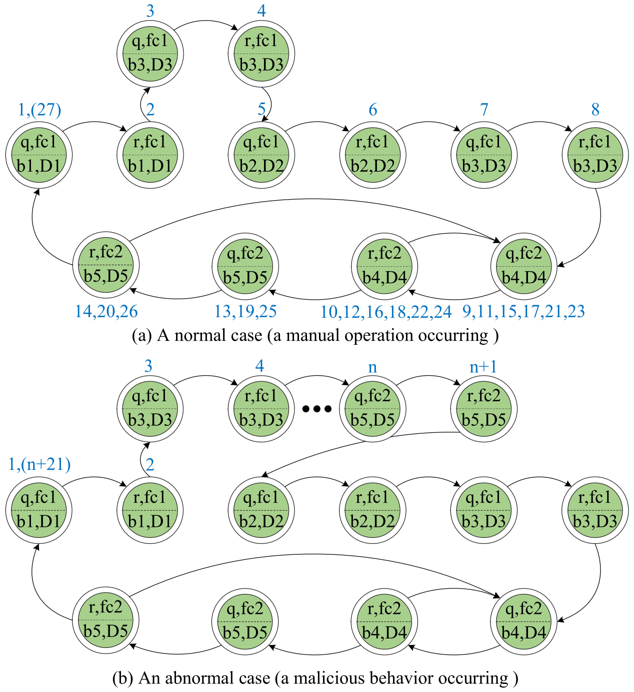


FIGURE 8. The E-DFA model between HMI and CN1 in normal conditions.



(a) A normal case (a manual operation occurring)

(b) An abnormal case (a malicious behavior occurring)

FIGURE 9. Examples on normal condition and abnormal condition.

Initially, communication packets in normal conditions are collected for the training of the detection models. As the communication occurs periodically with Modbus TCP protocol if no additional manual operations are added, the distributions of the traffic features are periodic as well. In this experimental system, there are two functions for the Modbus communication: "write single register" and "read holding registers", and both of which consist of "query - response". Fig. 8 shows the E-DFA model between HMI and CN1 in a period, where "r" and "q" represent the input symbols "response" and "query", respectively; "fc1" and "fc2" represent the function code values "write single register" and "read holding registers", respectively; "D1" to "D5" denote five communication data sets, and "b1" to"b5" denote the corresponding start addresses, and the numbers with blue font indicate the sequences. In this testbed, the period $T = 1$s, $\mathcal{N}_T = 24$ and $\mathcal{N}_{\max} = 28$. Fig. 9 shows two typical examples on operation occurring and malicious behavior occurring.

In I-IE based detection part, we select traffic features including TCP/IP and Modbus protocol, which are shown in

TABLE 2. Selected network features in I-IE.

Protocol	Features	Comments
TCP/IP	Packets	The number of packets
	MAC address	Src. or Dst. MAC address
	IP address	The Src. or Dst. IP address
	Port	The Src. or Dst. port
Modbus	Modbus packets	The number of Modbus packets
	Function code	The function code of Modbus

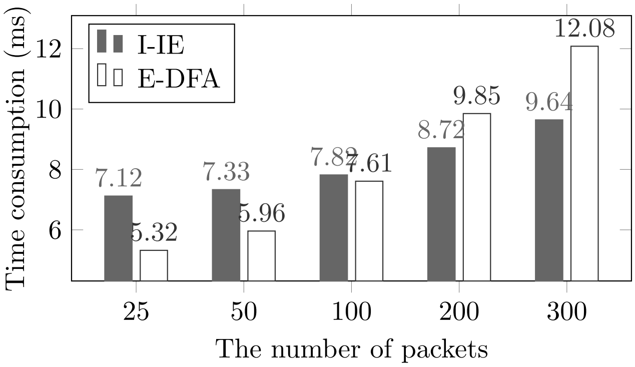


FIGURE 10. Time consumption for anomaly detection from I-IE and E-DFA.

TABLE 2. The detection thresholds of I-IE are determined by the statistical average in several sampling periods.

B. RESULT ON ANOMALY DETECTION

Our first set of experiments tests time consumption of the packet anomaly detection of proposed approach. The results are shown in Fig. 10. It can be seen from the figure that I-IE computation consumes more time than E-DFA matching for a low number of packets. However, with the increase in the number of packets, I-IE computation gradually becomes faster than E-DFA matching. This is because more iterations in E-DFA are required for state matching when the number of packets is big. Nevertheless, the time consumptions for both I-IE and E-DFA are as low as in the order of 10 ms (for 300 packets). They are far shorter than the communication cycle, which is typically 1s or longer in a TEP control system. In addition, the time consumption is expected to be drastically reduced if some improved measures are adopted, such as using an extended character-set in DFA [39], and choosing a fast entropy computation method [40]. By this way, our approach can be fit for other faster protocols.

The second set of experiments evaluates the detection accuracy. Four typical attack scenarios defined in [38] are constructed in our experiments:

1) Distributed Denial of Service (DDoS): DDoS attacks aim to block communication channels;
2) Man-in-the-middle (MITM): an MITM attack replays and possibly tampers communication data;
3) Spoofing: Spoofing attack sends malicious configuration data to the field systems; and
4) False Data Injection (FDI): FDI attack tampers process states from one field zone to others.

The first two attacks are implemented by injecting malicious packet through the LAN port of the OpenFlow switch, the spoofing attack will affect the testbed by tampering the set-points in the HMI, and the FDI attack is implemented by sending a fault value to the TEP Simulation Host.

The success rates (SR) of anomaly detection from I-IE and E-DFA are tabulated in TABLE 3, where SR is defined as the ratio of the normal classified packets to the number of malicious packets. Because MITM attacks need to use an extra device to disturb network communications, the entropy will have a great change. This is also the case for DDoS attacks. Therefore, MITM and DDoS attacks can be easily detected from I-IE. However, as shown in TABLE 3, I-IE is not effective for detection of Spoofing and FDI attacks. In comparison, E-DFA exhibits a high success rate for detection of these two types of attacks though it is not effective for DDoS and MITM attacks.

TABLE 3. SR (%) of anomaly detection of attacks.

	DDoS	MITM	Spoofing	FDI
I-IE	100	100	12.4	0
E-DFA	0	17.5	64.9	91.4
I-IE + E-DFA	100	100	71.8	91.4

A combination of both I-IE and E-DFA gives good success rates of detection for all the four types of attacks. The success rates reach 100%, 100%, 71.8% and 91.4% for DDoS, MITM, Spoofing and FDI attacks, respectively, as shown in TABLE 3.

In order to further illustrate the detection performance, three common criteria: false positive rate (FPR), false negative rate and detection accuracy (DA) are adopted, and their definitions can be found in literature [5]. The experiment results are shown in TABLE 4, where "Normal" denotes that the packets are collected in normal conditions, while "Anomaly" means that they are collected in attack situation. In our approach, the FPR is very low in any situations, but FNR is very different. The results on FPR demonstrates that false alarms appear infrequently in normal conditions, and thus, the detected anomalies have a high credibility, i.e. the DA is high. But there are some anomalies may not be found, especially under the spoofing attacks (FNR $= 28.2\%$). This is because that the attack tampers the set-points by using some random numbers, parts of them may be not out of their allowed ranges. Although the random numbers are also used by the FDI attacks, the tampered variables may cause exceptions to other physical states. This increases the likelihood of the attacks to be detected. Therefore, FNR in the FDI attacks is lower than the spoofing attacks ($8.6\% < 28.2\%$).

Furthermore, quantitative comparisons are also discussed between our approach and other relative works. The results

TABLE 4. The performance of the proposed hybrid anomaly detection method.

	Normal	Anomaly	FPR (%)	FNR (%)	DA (%)
DDoS	2546	2438	1.2	0	98.7
MITM	2487	2503	1.1	0	98.9
Spoofing	2522	2513	1.3	28.2	98.1
FDI	2509	2466	0.9	8.6	98.9

TABLE 5. DA measured in percentage (%) for different detection approaches under different attack types.

Method	DDoS	MITM	Spoofing	FDI
Ours in this paper	98.7	98.9	98.1	98.9
PCA [41]	99.3	36.4	99.0	69.5
SVM [42]	98.9	89.5	91.4	98.1
DBN [43]	99.2	97.4	93.6	97.5

are shown in Table 5, where SVM is short for support vector machine, and DBN is short for deep belief network. It is found that other approaches have a high DA for both DDoS and spoofing attacks. But the DA is much different for other two attacks, and some of them are lower than ours. For example, DA is 36.4% for MITM attack and 69.5% for FDI attack, respectively, in PCA-based method, while it is 89.5% for MITM attack in SVM-based approach. Both our approach and DBN-based approach have a high DA for all the constructed attacks. But in terms of detecting spoofing attack, our approach is better. In addition, the complexity of method implementation of our approach seems to be lower.

Moreover, the detection performance under different values of $\mathcal{N}_o$ in the part of E-DFA based approach is also discussed. The results are illustrated in Fig. 11, where FNR is calculated under a normal condition with some normal manual operations, and FPR is calculated under a spoofing attack scenario. It is observed that FNR is severe when there are no extra state changes allowed in a detection period. FNR decreases when $\mathcal{N}_o$ increases. But the growing of $\mathcal{N}_o$ will make FPR going up. Therefore, in this simulation system, it is appropriate to set $\mathcal{N}_o = 4$.

C. PERFORMANCE OF NETWORK COMMUNICATIONS

To evaluate the performance of network communications of our approach, DoS attacks in Z2 are simulated. Five distributed attack hosts are used, and the bandwidth of the OpenFlow switch is set as 100 Mb/s.

The experimental results are illustrated in Fig. 12. Our first observation is that without deployment of security protection to the network, the communication link will be broken when the rate of attack goes beyond 5000 packets/second. When blacklist is used, a large number of unexpected packets will be filtered out and directly dropped. As a result, the bandwidth performance of the communication is improved significantly. Nevertheless, it is still affected by the attacks from the active communication device. When these attacks are also detected, the communication link is further protected. Consequently, these attacks show very limited impact on the bandwidth performance of the communication link. This is clearly depicted in the upper curve in Fig. 12.

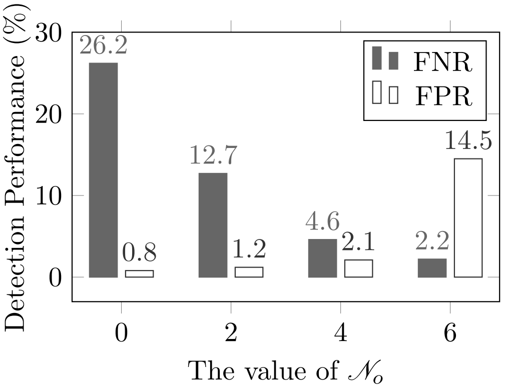


FIGURE 11. The detection performance under different values of $\mathcal{N}_o$.

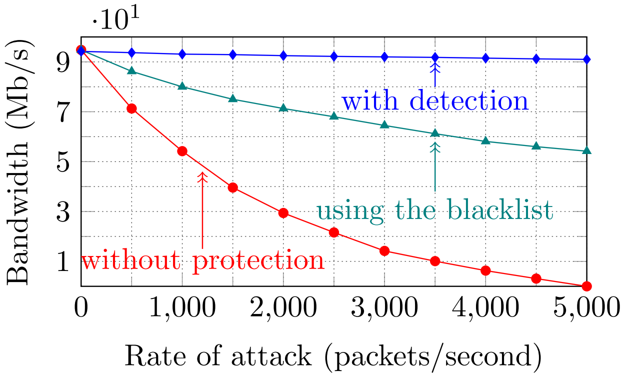


FIGURE 12. The performance of network communication control.

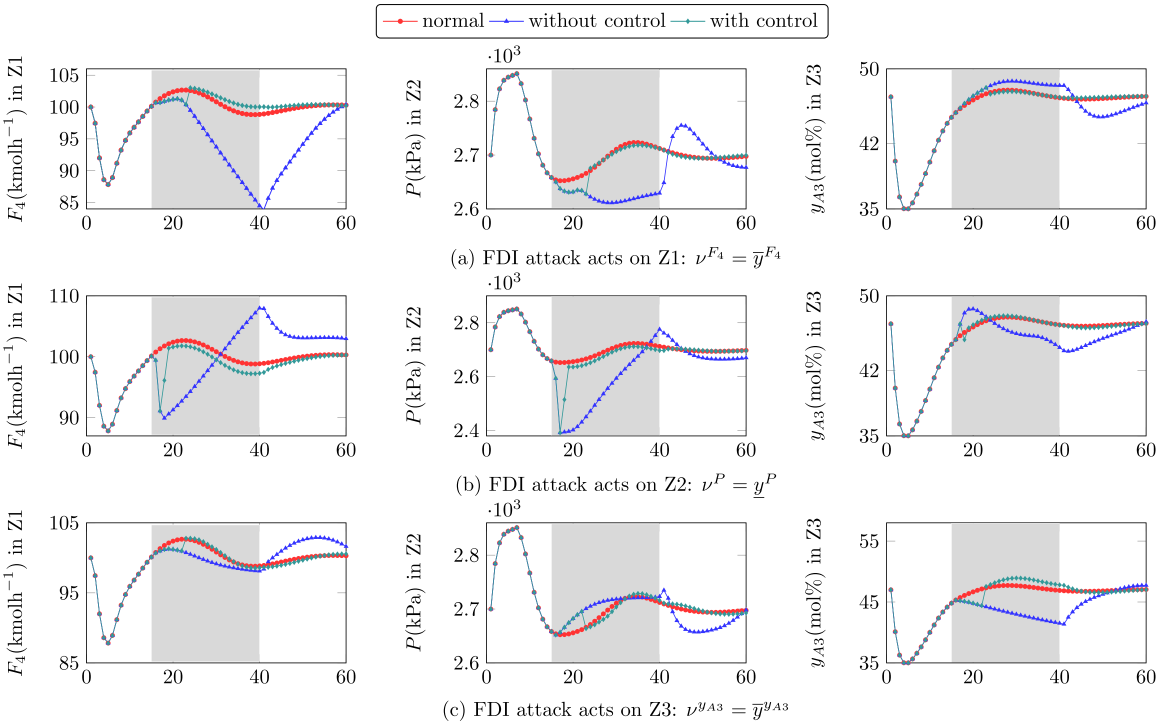
D. ATTACK MITIGATION ON PHYSICAL PROCESSES

In general, an FDI attack model for physical systems can be presented as [44]:

$$\tilde{y}^i(t) = \begin{cases} y^i(t), & \text{if } t \notin \mathbf{\Gamma}_{\nu^i}, \\ \nu^i(t), & \text{if } t \in \mathbf{\Gamma}_{\nu^i},\ \nu^i(t) \in \boldsymbol{V}^i, \end{cases} \tag{19}$$

where $\tilde{y}^i(t)$ and $y^i(t)$ denote the observed value and actual value, respectively; $\nu^i(t)$ represents attacked value, $\mathbf{\Gamma}^i_\nu$ is the duration of ν^i, $\boldsymbol{V}^i$ is an optional method of ν^i. In our experiments, corresponding lower and upper thresholds are used to tamper their observed states, such as $\nu^{F_4} = \bar{y}^{F_4}$ means that we tamper the observed state of product rate F_4 to be its upper threshold.

Fig. 13 shows the system states when different zones are intruded by FDI attacks, where "normal" indicates normal conditions, "with control" indicates that the control strategy acts on the physical system when anomalies are detected, and "without control" means that no control strategy is implemented on the system, and the abscissa represents the sample



(a) FDI attack acts on Z1: $\nu^{F_4} = \overline{y}^{F_4}$

(b) FDI attack acts on Z2: $\nu^P = \underline{y}^P$

(c) FDI attack acts on Z3: $\nu^{y_{A3}} = \overline{y}^{y_{A3}}$

FIGURE 13. The comparison of the system states with and without the control strategy when FDI attacks inject on product rate F_4, pressure P and reactant concentration y_{A3}, respectively.

sequence. In addition, the gray area in Fig. 13 represents the duration of attacks. In these experiments, $\overline{y}^{F_4} = 105$kmolh^{-1}, $\underline{y}^P = 2400$kPa and $\underline{y}^{y_{A3}} = 35$mol% are used by FDI attacks respectively.

As shown in Fig. 13, it can be found that the states $F4$ and P have a strong correlation, once one of them is attacked, these two states will be changed drastically, while the state y_{A3} has less of an effect. But attacking y_{A3} can cause a big influence on the other two states. The curves "with control" illustrate that the proposed control strategy can effectively mitigate the attack impacts in not only the compromised zone but also other zones (See the state y_{A3} under the FDI attack acts on Z1 and Z2). This is due to the fact that there are little influences on their original control loops in the normal zones when the states in the compromised zone have been limited. Thus, the states in these zones can be regulated as usual.

E. DISCUSSION

In summary, our experiments have demonstrated that: 1) Different types of attacks on both networks and physical processes can be detected by our hybrid anomaly detection with a low time consumption and a high accuracy; and 2) Our response strategy for security protection behaves with good performance in regulating the network communications and the physical process of the ICS in the presence of cyber-attacks.

Some characteristics of the proposed method are discussed based on some comparisons with some other existing methods. The comparison results are shown in TABLE 6. It can be found that the current methods mainly focus on either physical processes or network communications, most of them may not stabilize the physical processes. Methods on securing the physical processes need to redesign and/or reprogramming the control strategy in field controller such that it is difficult to be deployed. In addition, few of

TABLE 6. Some comparisons of the proposed method and other existing solutions.

	Our Method	Literature [10]	Literature [12]	Literature [15]	Literature [21]	Literature [22]	Literature [25]
Commun. Inspection	√	×	√	×	√	√	×
Phy. State Inspection	√	√	√	√	×	×	√
Commun. Protection	√	×	×	×	√	√	×
Phy. Process Protection	√	×	×	√	×	×	√
Reactive Response	√	×	×	√	√	√	√
Stabilize Phy. Process	√	×	×	√	×	×	×
Easy to deploy	√	√	×	×	√	√	√

Note: '√' means that the ability could be found or inferred in the literature;
'×' means that the ability was not mentioned and could not be inferred.

the existing methods have all of the abilities of this paper concerned.

VII. CONCLUSION

Security protection is highly desirable in ICSs, especially for physical processes. An SDSec-based approach has been presented in this paper for the field zones of ICSs. Different from existing approaches, the presented approach has considered not only attack propagation through the cyber channel but also the attack impact propagation through physical process interactions. The approach is based on a software-defined security architecture, which realizes a flexible way to secure the field zones. Then, a hybrid detection method is developed from I-IE and E-DFA for detection of abnormal behaviors including attacks on the inter-zone communications and physical processes. The results of the anomaly detection are fed into a security response mechanism for securing zone communications and mitigating attack impacts on the physical processes. We use the SDN environment to implement our experiments, and the results have shown that the SDSec-based approach presented in this paper is effective for protection of the field zones of ICSs.

In this paper, the periodic characteristics of communication traffic in ICSs are considered, and it is fit for many cases. But in some cases, the traffic may exhibit phases in time [45], and thus, the detection model that incorporates multiple phases of the traffic should be considered in the future works. Moreover, current research work only considers security protection between the field zones. An integrated blueprint for securing both the cyber space and physical world of ICSs also needs to be considered.

REFERENCES

[1] P. Haller and B. Genge, "Using sensitivity analysis and cross-association for the design of intrusion detection systems in industrial cyber-physical systems," *IEEE Access*, vol. 5, pp. 9336–9347, 2017.

[2] T. Cruz, L. Rosa, J. Proença, L. Maglaras, M. Aubigny, L. Lev, J. Jiang, and P. Simões, "A cybersecurity detection framework for supervisory control and data acquisition systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2236–2246, Dec. 2016.

[3] Y. Hashimoto, T. Toyoshima, S. Yogo, M. Koike, T. Hamaguchi, S. Jing, and I. Koshijima, "Safety securing approach against cyber-attacks for process control system," *Comput. Chem. Eng.*, vol. 57, pp. 181–186, Oct. 2013.

[4] *Security for Industrial Automation and Control Systems Part 1-1: Terminology, Concepts, and Models*, Standard ISA-62443-3-2, ISA, 2013.

[5] J. Yang, C. Zhou, S. Yang, H. Xu, and B. Hu, "Anomaly detection based on zone partition for security protection of industrial cyber-physical systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4257–4267, May 2018.

[6] I. N. Fovino, A. Coletta, A. Carcano, and M. Masera, "Critical state-based filtering system for securing SCADA network protocols," *IEEE Trans. Ind. Electron.*, vol. 59, no. 10, pp. 3943–3950, Oct. 2012.

[7] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "NIST SP 800-82: Guide to industrial control systems (ICS) security ver_2," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep., 2015. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r2.pdf

[8] W. Machii, I. Kato, M. Koike, M. Matta, T. Aoyama, H. Naruoka, I. Koshijima, and Y. Hashimoto, "Dynamic zoning based on situational activitie for ICS security," in *Proc. 10th Asian Control Conf. (ASCC)*, May/Jun. 2015, pp. 1–5.

[9] B. Genge, P. Haller, and I. Kiss, "Cyber-security-aware network design of industrial control systems," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1384–1397, Sep. 2017.

[10] T. Morita, S. Yogo, M. Koike, T. Hamaguchi, S. Jung, I. Koshijima, and Y. Hashimoto, "Detection of Cyber-attacks with zone dividing and PCA," *Procedia Comput. Sci.*, vol. 22, pp. 727–736, Jan. 2013.

[11] L. Mechtri, F. D. Tolba, and N. Ghoualmi, "Intrusion detection using principal component analysis," in *Proc. 2nd Int. Conf. Eng. Syst. Manage. Appl.*, Mar./Apr. 2010, pp. 1–6.

[12] A. Carcano, A. Coletta, M. Guglielmi, M. Masera, I. N. Fovino, and A. Trombetta, "A multidimensional critical state analysis for detecting intrusions in SCADA systems," *IEEE Trans. Ind. Informat.*, vol. 7, no. 2, pp. 179–186, May 2011.

[13] S. Ntalampiras, "Detection of integrity attacks in cyber-physical critical infrastructures using ensemble modeling," *IEEE Trans Ind. Informat.*, vol. 11, no. 1, pp. 104–111, Feb. 2015.

[14] L. F. Cómbita, J. Giraldo, A. A. Cárdenas, and N. Quijano, "Response and reconfiguration of cyber-physical control systems: A survey," in *Proc. IEEE 2nd Colombian Conf. Autom. Control (CCAC)*, Oct. 2015, pp. 1–6.

[15] X. Jin, W. M. Haddad, and T. Yucelen, "An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 6058–6064, Nov. 2017.

[16] S. Hossain, S. Etigowni, K. Davis, and S. Zonouz, "Towards cyber-physical intrusion tolerance," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2015, pp. 139–144.

[17] X. Luo, "Security protection to industrial control system based on defense-in-depth strategy," *WIT Trans. Eng. Sci.*, vol. 113, pp. 19–27, Feb. 2016.

[18] M. Kalinin, P. Zegzhda, D. Zegzhda, Y. Vasiliev, and V. Belenko, "Software defined security for vehicular ad hoc networks," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 533–537.

[19] X. Qiu, F. Cheng, W. Wang, G. Zhang, and Y. Qiu, "A security controller-based software defined security architecture," in *Proc. 20th Conf. Innov. Clouds, Internet Netw. (ICIN)*, Mar. 2017, pp. 191–195.

[20] G. Zhang, X. Qiu, and W. Chang, "Scheduling of security resources in software defined security architecture," in *Proc. Int. Conf. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, Oct. 2017, pp. 494–503.

[21] D. Jin, Z. Li, C. Hannon, C. Chen, J. Wang, M. Shahidehpour, and C. W. Lee, "Toward a cyber resilient and secure microgrid using software-defined networking," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2494–2504, Sep. 2017.

[22] E. Molina, E. Jacob, J. Matias, N. Moreira, and A. Astarloa, "Using software defined networking to manage and control IEC 61850-based systems," *Comput. Elect. Eng.*, vol. 43, pp. 142–154, Apr. 2015.

[23] B. Genge and P. Haller, "A hierarchical control plane for software-defined networks-based industrial control systems," in *Proc. IFIP Netw. Conf. (IFIP Netw.) Workshops*, May 2016, pp. 73–81.

[24] E. Molina and E. Jacob, "Software-defined networking in cyber-physical systems: A survey," *Comput. Electr. Eng.*., vol. 66, pp. 407–419, Feb. 2018.

[25] A. F. M. Piedrahita, V. Gaur, J. Giraldo, A. A. Cárdenas, and S. J. Rueda, "Leveraging software-defined networking for incident response in industrial control systems," *IEEE Softw.*, vol. 35, no. 1, pp. 44–50, Jan./Feb. 2018.

[26] F. Mallet, E. Villar, and F. Herrera, "MARTE for CPS and CPSoS," in *Cyber-Physical System Design From an Architecture Analysis Viewpoint*, S. Nakajima, J. Talpin, M. Toyoshima, and H. Yu, Eds. Singapore: Springer, 2017, pp. 81–108.

[27] R. H. Dong, D. F. Wu, Q. Y. Zhang, and T. Zhang, "Traffic characteristic map-based intrusion detection model for industrial Internet," *Int. J. Netw. Secur.*, vol. 20, no. 2, pp. 359–370, Mar. 2018.

[28] N. Goldenberg and A. Wool, "Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems," *Int. J. Crit. Infrastruct. Protection*, vol. 6, no. 2, pp. 63–75, Jun. 2013.

[29] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[30] J. Yu, H. Lee, M. S. Kim, and D. Park, "Traffic flooding attack detection with SNMP MIB using SVM," *Comput. Commun.*, vol. 31, no. 17, pp. 4212–4219, 2008.

[31] J. Yang, C. Zhou, and Y. Zhao, "A security protection approach based on software defined network for inter-area communication in industrial control systems," in *Proc. 12th IET Syst. Saf. Cyber-Secur. Conf.*, London, U.K., Oct. 2017, pp. 1–6.

[32] I. Özçelik and R. R. Brooks, "Deceiving entropy based DoS detection," *Comput. Secur.*, vol. 48, pp. 234–245, Feb. 2015.

[33] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.

[34] M. Matta, M. Koike, W. Machii, T. Aoyama, H. Naruoka, I. Koshijima, and Y. Hashimoto, "Industrial control system monitoring based on communication profile," *J. Chem. Eng. Japan*, vol. 48, no. 8, pp. 619–625, 2015.

[35] G. Böker, and J. Lunze, "Stability and performance of switching Kalman filters," *Int. J. Control*, vol. 75, no. 16, pp. 1269–1281, Nov. 2002.

[36] N. L. Ricker, "Model predictive control of a continuous, nonlinear, two-phase reactor," *J. Process Control*, vol. 3, no. 2, pp. 109–123, May 1993.

[37] A. Termehchy and A. Afshar, "A novel design of unknown input observer for fault diagnosis in non-minimum phase Systems," *IFAC Proc. Volumes*, vol. 47, no. 3, pp. 8552–8557, 2014.

[38] Z.-H. Pang, G.-P. Liu, D. Zhou, F. Hou, and D. Sun, "Two-channel false data injection attacks against output tracking control of networked systems," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3242–3251, May 2016.

[39] C. Liu, Y. Pan, A. Chen, and J. Wu, "A DFA with extended character-set for fast deep packet inspection," *IEEE Trans. Comput.*, vol. 63, no. 8, pp. 1925–1937, Aug. 2014.

[40] G. No and I. Ra, "Adaptive DDoS detector design using fast entropy computation method," in *Proc. 5th Int. Conf. Innov. Mobile Internet Services Ubiquitous Comput.*, Seoul, South Korea, Jun./Jul. 2011, pp. 86–93.

[41] S. N. Shirazi, A. Gouglidis, K. N. Syeda, S. Simpson, A. Mauthe, I. M. Stephanakis, and D. Hutchison, "Evaluation of anomaly detection techniques for SCADA communication resilience," in *Proc. IEEE Resilience Week*, Chicago, IL, USA, Aug. 2016, pp. 140–145.

[42] O. A. Wahab, A. Mourad, H. Otrok, and J. Bentahar, "CEAP: SVM-based intelligent detection model for clustered vehicular ad hoc networks," *Expert Syst. Appl.*, vol. 50, pp. 40–54, May 2016.

[43] S. Huda, J. Yearwood, M. M. Hassan, and A. Almogren, "Securing the operations in SCADA-IoT platform based industrial control system using ensemble of deep belief networks," *Appl. Soft Comput.*, vol. 71, pp. 66–77, Oct. 2018.

[44] Y. Huang and A. Cárdenas, S. Amin, Z.-S. Lin, H.-Y. Tsai, and S. Sastry, "Understanding the physical and economic consequences of attacks on control systems," *Int. J. Crit. Infrastruct. Protection*, vol. 3, no. 2, pp. 73–83, Oct. 2009.

[45] C. Markman, A. Wool, and A. Cárdenas, "Temporal phase shifts in SCADA networks," in *Proc. Workshop Cyber-Phys. Syst. Secur. PrivaCy*, New York, NY, USA, Oct. 2018, pp. 84–89.

JUN YANG received the B.S. degree in automation and the M.S. degree in control engineering from the Taiyuan University of Technology, Taiyuan, China, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree in control science and engineering with the School of Automation and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China.

His main research interests include industrial communication, industrial control systems, and intrusion detection.

CHUNJIE ZHOU received the M.S. and Ph.D. degrees in control theory and control engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1991 and 2001, respectively.

He is currently a Professor with the School of Automation, Huazhong University of Science and Technology. His research interests include safety and security control of industrial control systems, theory and application of networked control systems, and artificial intelligence.

YU-CHU TIAN (M'00) received the B.S. degree in electrical engineering from the Wuhan Institute of Engineering, Wuhan, China, in 1982, the M.S. degree in electrical engineering from the East China University of Science and Technology, Shanghai, China, in 1987, and the Ph.D. degree in industrial automation from Zhejiang University, Hangzhou, China, in 1993.

Over the last many years, he was with Zhejiang University; The Hong Kong University of Technology, Hong Kong; the Curtin University of Technology, Perth, WA, USA; and the University of Maryland at College Park, MD, USA. Since 2002, he has been a Professor of computer science with the Queensland University of Technology, Brisbane, QLD, Australia. His current research interests include big data computing, cloud computing, real-time computing, computer networks, and control theory and engineering. He has published a monograph and more than 200 refereed papers and holds a patent. He is the Editor-in-Chief of Springer's book series *Handbook of Real-Time Computing* and an Associate Editor for a few international journals.

SHUANG-HUA YANG (SM'06) received the B.S. degree in instrument and automation and the M.S. degree in process control from the China University of Petroleum (Huadong), Beijing, China, in 1983 and 1986, respectively, and the Ph.D. degree in intelligent systems from Zhejiang University, Hangzhou, China, in 1991.

He is currently a Professor of computer science with Loughborough University, Loughborough, U.K. His current research interests include wireless network-based monitoring and control, safety critical systems, and real-time software maintenance. He is a Fellow of IET and InstMC, U.K. He is an Associate Editor of the *International Journal of Systems Science*, the *Arabian Journal for Science and Engineering* (Springer), and the *International Journal of Computing and Automation*.