EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Detecting stealthy attacks against industrial control systems based on residual skewness analysis

Yan Hu[1], Hong Li[2,3]*, Hong Yang[4], Yuyan Sun[2,3], Limin Sun[2,3] and Zhiliang Wang[1]

Abstract

With the integration of the modern industrial control systems (ICS) with the Internet technology, ICS can make full use of the rich resources on the Internet to facilitate remote process control. However, every coin has two sides. More exposure to the outside IT world has made ICS an attractive target for hackers, so it becomes urgent to protect the security of ICS. Skilled attackers can penetrate control networks and then manipulate sensor readings or control signals persistently until the system crashes, while still keeping themselves undetected by following the expected behavior of the system closely. This kind of attacks are referred to as stealthy attacks. As far as we know, many existing intrusion detection techniques only investigate the magnitudes of behavior residuals, so they cannot detect this kind of stealthy attacks. In this paper, we discover that residuals generated during stealthy attacks exhibit significant skewness compared to attack-free residuals. Based on the new observation, we propose an effective and fast technique to detect stealthy attacks against ICS based on residual skewness analysis. Skewness coefficients can distinguish the counterfeited residuals from the attack-free residuals effectively. A larger absolute value of the skewness coefficient generally indicates the occurrence of a more intense stealthy attack. Finally, we conduct comprehensive experiments to verify the effectiveness and efficiency of the proposed stealthy attack detection approach.

Keywords: Industrial control systems, Stealthy attacks, Intrusion detection, Residual skewness analysis

1 Introduction

Nowadays, industrial control systems (ICS) [1] play a very important role in national critical infrastructures, such as smart grids [2–4], water treatment systems [5], chemical processing plants [6], oil and natural gas pipelines [7], or large-scale communication systems [8]. With the rapid development of Internet technology (IT), ICS are also strengthening the connectivity to the Internet so as to make full use of the rich resources on the Internet to support remote process control and intelligent decision-making. However, the growing openness of ICS has made them an attractive target for malicious attackers [9, 10]. In 2010, the notorious cyber worm "Stuxnet" infected the core control program of the Natanz uranium enrichment base in Iran and misled the centrifuge that produces enriched uranium into accelerating unconventionally, and finally caused a severe damage to the centrifuge and the whole nuclear plant was forced to stop. In 2015, the "BlackEnergy3" attacked the Ukrainian power grid. The counterfeited control instructions of relays caused abnormal circuit disconnections, immediately followed by a large-scale blackout. At Black Hat 2017 [11], Dr. Staggs stated that the wind farm vendor design and implementation flaws left the wind turbine programmable automation controllers and OPC (OLE for process control) servers vulnerable to attacks. Additionally, they designed attack tools to exploit wind farm control network design and implementation vulnerabilities. So many ICS security incidents indicate that the security of ICS has become an urgent international issue [12, 13].

Intrusion detection systems (IDS) [14, 15] provide an effective solution to identify malicious attacks against

*Correspondence: lihong@iie.ac.cn
[2]Beijing Key Laboratory of IoT Information Security, Institute of Information Engineering, Chinese Academy of Sciences, A 89 Minzhuang Road, Beijing 100195, China
[3]School of Cyber Security, University of Chinese Academy of Sciences, A 89 Minzhuang Road, Beijing 100195, China
Full list of author information is available at the end of the article

traditional information systems by analyzing network protocols and traffic data. However, when applying IDS to ICS, the real-time process data is another important factor to consider [16]. The evolution of an industrial process generally follows fundamental laws of nature, which is a distinct feature of ICS. Attackers usually attempt to cause fatal physical damages to ICS by manipulating process data (e.g., sensor readings [17, 18] or control signals [19, 20]) maliciously. Therefore, by monitoring and analyzing the "physics" of ICS, we can detect a wide variety of intrusions. IDS generally construct a physical model for the target control system, based on which to forecast its expected behaviors. Once the monitored behaviors deviate from the expected values significantly, an alarm is raised.

However, in recent years, Liu et al. [18] discovered a new kind of stealthy attacks against ICS, which can bypass existing intrusion detection schemes. As we all know, the dynamic behavior of a control system generally does not change significantly within a short time period due to physical constraints. Therefore, the attacker can make the observed behavior of a system follow its expected behavior closely during a stealthy attack, but still inject enough false information into the system after a long period of time [16], and finally cause a fatal damage to the target system. Since then, stealthy attacks against ICS have attracted much attention [21, 22]. Previously, we proposed a detection approach against stealthy attacks based on residual permutation entropy [23].

In this paper, we propose an effective and much faster stealthy attack detection technique based on residual skewness analysis of system behaviors, which is more suitable for the real-time requirement of industrial control systems. Counterfeited residuals generally conform to a skewed distribution, which is different from a normal distribution, if the intruder intends to achieve specific attack goals. The values of the residual skewness coefficient can effectively distinguish a residual sequence generated during a stealthy attack from an attack-free residual sequence. Accordingly, stealthy attacks can be identified successfully. We launch stealthy attacks on two simulated ICS and verify the effectiveness of the proposed stealthy attack detection technique. The key contributions of this work are summarized as follows:

- We investigate the prediction residuals of system behaviors under stealthy attacks and discover that the residual distribution exhibits a significant degree of skewness compared to a normal distribution.
- We make full use of the skewness contained in the prediction residuals and propose a novel detection technique against stealthy attacks based on residual skewness analysis.
- Comprehensive experiments are conducted on simulated ICS to verify the effectiveness and efficiency of the proposed stealthy attack detection approach.

The rest of the paper is organized as follows. Section 2 introduces some research literature about ICS IDS. In Section 3, we present some preliminaries of our approach. In Section 4, we elaborate on the novel detection technique against stealthy attacks based on residual skewness analysis. Experiments are conducted to verify the effectiveness and efficiency of the proposed stealthy attack detection approach in Section 5. Experimental results are discussed in Section 6. Finally, we draw a conclusion in Section 7.

2 Related work

Due to the increasing connectivity between ICS and the outside IT world, cyber attacks against IT systems also endanger ICS. Traditionally, intrusion detection techniques against cyber attacks are mainly divided into two categories: misuse-based and anomaly-based. Misuse-based intrusion detection techniques, also referred as signature-based, rely on a precise definition of malicious system behaviors. If system activities match the known malicious behavior patterns, a potential attack is detected. Anomaly-based intrusion detection techniques exploit a definition of normal behavior and flag any visible deviation from normal behavior as unintentional faults or intentional attacks. In this section, we try to present a new taxonomy of intrusion detection techniques on ICS. Attacks against ICS often cause abnormal network traffics or violate network protocol specifications. Furthermore, due to the close correlation between ICS and physical processes, investigating process data can also help identify malicious intrusions against ICS. Therefore, we introduce the research literature of ICS IDS from three aspects: network traffic mining, network protocol analysis, and process data analysis.

2.1 Intrusion detection based on network traffic mining

ICS have relatively fixed operation objects and business processes, simple and static network topologies, and small numbers of applications, which result in relatively stable traffic patterns under normal conditions. Fluctuation of network traffics generally indicates the status change of ICS, which enables intrusion detection based on network traffic mining.

Traditional IDS based on network traffic analysis [24] generally extract information such as source and destination IP addresses and ports, traffic durations, and average time intervals between adjacent packets, and then apply data mining technologies to these collected information to identify abnormal system behaviors. The commonly used traffic mining techniques include supervised clustering

[25], semi-supervised clustering [26], mixed Gaussian model [27], neural network [28, 29], fuzzy logic [30–32], single-class support vector machine [33], multi-class support vector machine [34], and deep learning [35]. The purpose of these techniques is to establish complex nonlinear relationships between network traffics and system behaviors. The relationships, together with the current network traffic data, are then used to judge the security status of a target system. However, the computation overhead is usually high due to the large number of traffic features. In order to improve detection efficiency, some researchers utilized techniques like the ant colony algorithm [36] and the principal component analysis method [37] to remove redundant traffic features.

2.2 Intrusion detection based on network protocol analysis

Protocol specifications generally define the packet formats and communication modes allowed by the protocol. Intrusion detection rules can be extracted from protocol specifications. Accordingly, malicious behaviors that violate protocol specifications can be identified effectively. Common open protocols in ICS include ModBus, ICCP/TASE.2, and DNP3. These protocols are vulnerable to a variety of network attacks such as theft, tampering, and counterfeiting.

Cheung et al. [38] constructed a protocol specification model based on legal values of different data fields and legal relationships between different fields in a data packet. Additionally, they built normal communication patterns based on the security requirements, the data transmission directions and transmission ports of specific ICS. Anomalies violating the protocol specification model or the desired communication patterns could be detected, which belongs to anomaly-based intrusion detection techniques. Morris et al. [39] used Snort (an intrusion detection software) to generate signatures for ModBus protocol vulnerabilities. These signatures were used to examine communication data in field networks and identify illegal data, which is a typical misuse-based approach. Moreover, in order to achieve rapid development, other researchers modify the traditional IDS to make them suitable for ICS. Lin et al. [40] integrated a packet parser of industrial control protocols (e.g., DNP3) into the famous network intrusion detection system *Bro* developed by the University of Berkeley, to support intrusion detection in ICS.

In addition to open protocols, IDS based on proprietary protocols are also designed. Hong et al. [41] analyzed automatic systems in the substations of smart grids and detected anomalies or malicious behaviors in multicast messages based on the specifications extracted from the IEC 61850 standards (e.g., Generic Object Oriented Substation Event (GOOSE) and Sample Value technology (SV)). Hadeli et al. [42] extracted legal and illegal network traffic models from the protocol specifications of power systems and transformed them into Snort rules for intrusion detection.

The above two categories of IDS build the first security barrier for ICS. However, the close relationship between ICS and the physical world makes ICS different from traditional information systems. Therefore, the above two categories of IDS, originally designed for traditional information systems, are difficult to identify attacks against physical processes, which do not cause abnormal network traffics nor violate network protocol specifications. Therefore, IDS based on process data analysis have emerged.

2.3 Intrusion detection based on process data analysis

Process information is an important factor to consider in ICS IDS. Attackers usually mislead the controller into making wrong decisions [17] by tampering with process information, and finally cause a fatal damage to ICS. Such attacks can be detected by comparing the observed and expected process values in real time. Once the deviation exceeds a predefined threshold significantly, an alarm is raised [43]. Hadžiosmanović et al. [44] classified process variables into three categories: constants, enumeration, and continuous variables. Afterwards, a normal behavior model was built for each process variable. During system operation, once an observed process value deviated from its normal behavior model, the system generated an alarm. Carcano et al. [45] used measurement data from multiple industrial sensors to denote system states and proposed a state distance measurement method. Intrusions could be identified by examining the proximity between the current state and the critical states.

Other researchers use time series forecasting techniques to predict the future outputs of ICS. The predicted outputs are compared with the monitored values to generate residuals. Afterwards, some statistical analysis techniques are performed on the residuals to identify intrusions. If the system operates normally, the residual sequence follows a Gaussian distribution approximately. Once an intrusion occurs, the actual behavior of a system deviates from its expected behavior, i.e., the residuals are different from 0 observably [46]. Cárdenas et al. [47] summarized two categories of intrusion detection methods based on residual analysis: sequential detection and change detection. The former aims to find intrusions as soon as possible, i.e., determining the shortest residual sequence based on which IDS can make a normal/abnormal judgment. The latter detects a possible anomaly at an unknown time point. In other words, the system detects the transition from a normal state to an abnormal state based on whether the residual or the accumulated residual exceeds a certain threshold. The commonly used change detection methods can be classified into two categories: stateless [48] and stateful [16]. The stateless and stateful detection methods raise alarms when the residual and the cumulative residual at the current time point exceed a threshold, respectively.

However, Liu et al. [18] discovered a new kind of data injection attacks against state estimation in power grids in 2011. This attack injects erroneous data into the system persistently until the system crashes, but always keeps the residual magnitudes below the detection threshold, thus to bypass the stateless intrusion detection scheme. This is the first stealthy attack against ICS. Since then, stealthy attacks have emerged in a variety of industrial control scenarios (e.g., chemical process control [47] and industrial waste water treatment [49]). Until 2016, Urbina et al. [16, 50] stated that existing intrusion detection technology still cannot detect stealthy attacks effectively, so they proposed a new method to measure the negative impacts of stealthy attacks on ICS and tried to limit the negative impacts by configuring detection schemes and metrics properly. Since then, some researchers have conducted further research on stealthy attacks, but they mainly focused on how to perform stealthy attacks on specific ICS [21] or exploring the impacts of stealthy attacks on some more complex systems [22]. As a result, detecting stealthy attacks against ICS becomes an urgent issue. In our previous work [23], we proposed a detection technique against stealthy attacks based on the analysis of residual permutation entropy. This technique was effective but not very fast. In this paper, we propose an effective and much faster technique to detect stealthy attacks based on residual skewness analysis, which utilizes the residual distribution skewness to identify abnormal system behaviors.

3 Preliminaries

The approach proposed in this paper belongs to the category of IDS based on process data analysis. Intrusion detection based on process data analysis mainly includes three steps. First, build a physical model for the target system in order to predict its expected outputs $\hat{y}_k$ in the future. Second, compute the residuals r_k between the observed outputs y_k and the predicted values $\hat{y}_k$ during system operation. Third, perform statistical analysis on the residual sequence to detect intrusions. In this section, we introduce physical models of ICS, prediction techniques, and intrusion detection statistics.

3.1 Physical models of ICS

Physical models generally characterize time-varying behaviors of ICS, so a reasonable model can predict the expected behavior of a system accurately. We can derive physical models from first principles (e.g., Newton's laws, electromagnetic laws, and fluid dynamics) or from historical data of ICS using system identification technology. There are two commonly used models in system identification: auto-regressive integrated moving average (ARIMA) [51] and linear dynamical state-space (LDS) [52]. The ARIMA model of a time series $\{y_k\}$ is formalized as follows:

$$y_k = \sum_{i=1}^{p} \phi_i y_{k-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{k-j} + \varepsilon_k, \tag{1}$$

where y_k and y_{k-i} $(i = 1, 2, \ldots, p)$ are the current and last p output values of a system, ε_k and ε_{k-j} $(j = 1, 2, \ldots, q)$ are the current and last q prediction errors, which are Gaussian noises with a zero mean and a non-zero variance, ϕ_i and θ_j are model parameters, which should be estimated from the time series $\{y_k\}$ [53].

ARIMA models just build relationships between system outputs, but cannot relate system inputs to system outputs. If both the control signals (inputs) and the sensor readings (outputs) are available, we can construct the LDS model as follows:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{K}\boldsymbol{\varepsilon}_k, \tag{2}$$

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{D}\boldsymbol{u}_k + \boldsymbol{e}_k, \tag{3}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and $\boldsymbol{K}$ are system matrices characterizing the dynamics of a physical system, and $\boldsymbol{\varepsilon}_k$ and $\boldsymbol{e}_k$ are process and sensor noises following Gaussian distributions. $\boldsymbol{D}$ is generally equal to $\boldsymbol{0}$ owing to the strict causality of most physical systems. The LDS model indicates that the next state $\boldsymbol{x}_{k+1} \in \mathbb{R}^n$ of a system is determined by the current state $\boldsymbol{x}_k \in \mathbb{R}^n$ and the current control signal $\boldsymbol{u}_k \in \mathbb{R}^p$. Additionally, as shown in Eq. (3), the expected output $\boldsymbol{y}_k \in \mathbb{R}^q$ of the system is a linear combination of system states $\boldsymbol{x}_k$.

3.2 Kalman filtering for process forecasting

Kalman filtering (KF) [54] is a well-known technique to forecast the future behavior of a LDS model. The KF algorithm performs two operations recursively: prediction and update. The prediction step projects forward the current posteriori state to the next priori state, along with uncertainties. Once the system output (inevitably corrupted with some errors and noises) of the next step is measured, the update step computes the posteriori state of the next step as a weighted average of its priori estimate and the sensor measurement. A greater weight is assigned to a priori state estimate with higher certainty.

We respectively use $\boldsymbol{x}_k^-$ and $\boldsymbol{x}_k$ to denote the priori and posteriori states at step k before and after the k-th system output $\boldsymbol{y}_k$ is observed. The prediction step is denoted by:

$$\boldsymbol{x}_{k+1}^- = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k, \tag{4}$$

$$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{K}\boldsymbol{Q}_k\boldsymbol{K}^{\mathrm{T}}, \tag{5}$$

where $\boldsymbol{P}_{k+1}^{-}$ and $\boldsymbol{P}_k$ denote the priori and posteriori covariance matrices of prediction errors at step $k+1$ and k, respectively, and $\boldsymbol{Q}_k$ is the covariance matrix of the process noise $\boldsymbol{\varepsilon}_k$ at step k. Accordingly, KF predicts the next expected output $\hat{\boldsymbol{y}}_{k+1}$ of the system as follows:

$$\hat{\boldsymbol{y}}_{k+1} = \boldsymbol{C}\boldsymbol{x}_{k+1}^{-}. \tag{6}$$

Once the next system output $\boldsymbol{y}_{k+1}$ is measured, the update step is performed as follows:

$$\boldsymbol{KAL}_{k+1} = \boldsymbol{P}_{k+1}^{-}\boldsymbol{C}^{\mathrm{T}}\left[\boldsymbol{C}\boldsymbol{P}_{k+1}^{-}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{R}_k\right]^{-1}, \tag{7}$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_{k+1}^{-} + \boldsymbol{KAL}_{k+1}\left[\boldsymbol{y}_{k+1} - \boldsymbol{C}\boldsymbol{x}_{k+1}^{-}\right], \tag{8}$$

$$\boldsymbol{P}_{k+1} = [\boldsymbol{I} - \boldsymbol{KAL}_{k+1}\boldsymbol{C}]\,\boldsymbol{P}_{k+1}^{-}, \tag{9}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{R}_k$ denotes the covariance matrix of the measurement noise $\boldsymbol{e}_k$, the Kalman gain matrix $\boldsymbol{KAL}_{k+1}$ is estimated by minimizing $\boldsymbol{P}_{k+1}$. $\boldsymbol{P}_{k+1}$ in Eq. (9) is the consequent minimized posteriori covariance matrix. As shown in Eq. (8), the posteriori state $\boldsymbol{x}_{k+1}$ is computed as a weighted average of the priori state estimate $\boldsymbol{x}_{k+1}^{-}$ and the deviation between the new sensor measurement $\boldsymbol{y}_{k+1}$ and its forecast $\boldsymbol{C}\boldsymbol{x}_{k+1}^{-}$. $\boldsymbol{KAL}_{k+1}$ determines how much the new sensor measurement contributes to the posteriori state estimation. If the past prediction is with higher certainty (i.e., $\boldsymbol{P}_k$ smaller and accordingly $\boldsymbol{P}_{k+1}^{-}$ smaller), the contribution of the new sensor measurement $\boldsymbol{y}_{k+1}$ should be less ($\boldsymbol{KAL}_{k+1}$ smaller).

3.3 Detection statistics

After building the physical model for the target control system and performing the process forecasting procedure, IDS perform statistical analysis on the forecasting residuals to detect potential attacks. Generally, there are two kinds of residual testing techniques: stateless and stateful [50].

The stateless test raises an alarm for each observable deviation, i.e., $|y_k - \hat{y}_k| = |r_k| \geq \tau_1$ $(k > 0)$, where y_k and $\hat{y}_k$ are the measured system output and its forecast at step k, and τ_1 is a pre-defined detection threshold. In the stateful test, the change (no matter how small) of r_k is tracked using another statistic S_k. The non-parametric CUmulative SUM (CUSUM) is one of the most popular stateful detection statistic. It is a variable defined recursively as $S_0 = 0$ and $S_{k+1} = (S_k + |r_k| - \delta)^+$, where $(x)^+$ denotes $\max(0, x)$, and δ is a small positive value used to keep S_k from increasing persistently when the system operates normally. Once S_k exceeds the detection threshold τ (τ is defined based on a tolerable false alarm rate), in other words, there exists a persistent deviation across multiple time steps, an alarm is generated and S_{k+1} is reset to 0 when the detection procedure restarts. The intrusion detection procedure based on process data analysis is summarized in Fig. 1.

4 Detecting stealthy attacks

In this section, we present the novel detection approach against stealthy attacks based on residual skewness analysis. We first take a water level control system as an example to describe the stealthy attack model. Then, we present the detection strategies against stealthy attacks.

4.1 The stealthy attack model

We take a water level control system as a motivating example to describe the stealthy attack model against ICS. The architecture of the system is shown in Fig. 2. The water level in the tank should be maintained below 0.8 m (the high level) and above 0.2 m (the low level) by turning on or off the inlet and outlet pumps at proper moments. Water spill occurs at 1.1 m.

Suppose that each pump has only two states: on and off. A water level sensor is used to monitor the water level in the tank and transmits measurement data to the controller (PLC). The PLC generates appropriate control commands according to the real-time sensor measurements. For simplicity, the outlet pump is assumed to keep working when the system operates normally. As a result, only the inlet pump needs to be controlled to maintain the water level in

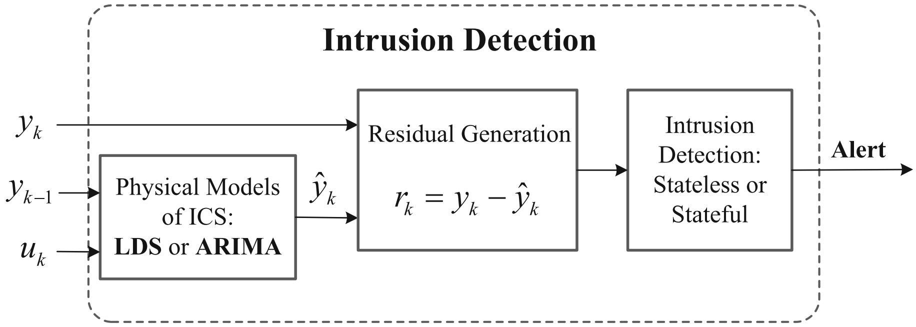

Fig. 1 Intrusion detection based on process data analysis

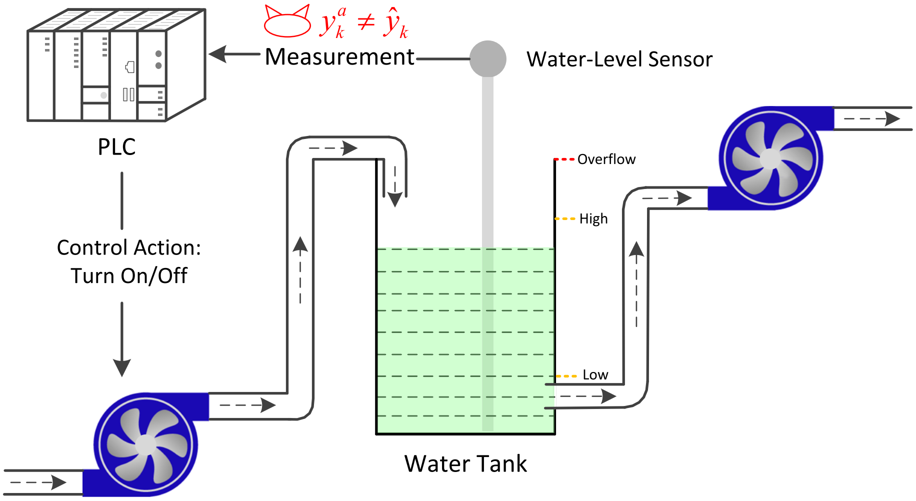

Fig. 2 A water level control system

the tank. Moreover, we assume that the amount of water coming in is greater than the amount of water going out per unit time while the two pumps are both working. The inlet pump should be turned off once the water level exceeds the high level, and should be turned on again once the water level goes down below the low level.

We assume that the adversary is able to gain knowledge of the physical model of the target ICS, the process forecasting and intrusion detection techniques, and can tamper with the sensor measurements secretly. Thus the adversary can launch a successful stealthy attack. The physical model of the system can be derived from the mass balance equation, which relates the water level h with the volume of water coming in Q^{in} and the volume of water going out Q^{out} per unit time as follows:

$$\text{Area}\frac{\mathrm{d}h}{\mathrm{d}t} = Q^{\text{in}} - Q^{\text{out}}, \tag{10}$$

where Area denotes the cross-sectional area of the tank, and Q^{in} and Q^{out} are positive constants when the two pumps are both working, and zero otherwise. Assuming that the discrete time interval is 1 s, the LDS model is derived as follows:

$$h_{k+1} = h_k + \frac{Q_k^{\text{in}} - Q_k^{\text{out}}}{\text{Area}}, \tag{11}$$

where h_{k+1} and h_k are the water heights at step $k+1$ and k, and $Q_k^{\text{in}} - Q_k^{\text{out}}$ is the control input at step k. In this example, we assume that Q_k^{out} keeps constant when the system operates normally and Q_k^{in} changes over time according to the control instructions issued by the controller. As a result, this equation is not an ARIMA model but a LDS model with $x_k = h_k$, $u_k = \left[Q_k^{\text{in}}, Q_k^{\text{out}}\right]^T$, $B = \left[\frac{1}{\text{Area}}, -\frac{1}{\text{Area}}\right]$, $A = 1$, and $C = 1$.

The adversary attempts to manipulate the water level in the tank maliciously by tampering with the sensor measurements persistently but remain undetected until water spill occurs. Specially, during a surge stealthy attack [47], the goal of the adversary is to cause maximum damage to the system as quickly as possible. Suppose that the stateful test is adopted by IDS due to its stronger detection ability compared to the stateless test. Once the detection threshold τ is reached, the stateful statistic S_k should stay at the threshold until the water overflows. Otherwise, the attack can be easily identified by IDS. Accordingly, the adversary needs to solve the following equation:

$$S_k + |y_k^a - \hat{y}_k| - \delta = \tau, \tag{12}$$

where y_k^a and $\hat{y}_k$ denote the observed and forecasted water levels during a stealthy attack, respectively. By solving this equation, the adversary can get the following attack model:

$$y_k^a = \begin{cases} \hat{y}_k - (\tau + \delta), & k = 1 \\ \hat{y}_k - \delta, & k > 1 \end{cases} \tag{13}$$

The model means that the fake water levels that are lower than their forecasts should be sent to the PLC persistently until the water spill occurs. In the first step of the attack, the residual between the fake water level and its forecast is $-(\tau + \delta)$. In the following steps, the residuals should be kept at $-\delta$. In another word, the adversary should increase the observed water levels at a lower rate than the forecasts. The attack goal is achieved when the controller receives a high water-level measurement from the sensor and issues a "turn-off" control command to the inlet pump, but the deviation (Δ) between the observed sensor measurement and the real water level exceeds overflow-high. Figure 3 illustrates three attacks with different slopes from the low level to the high level.

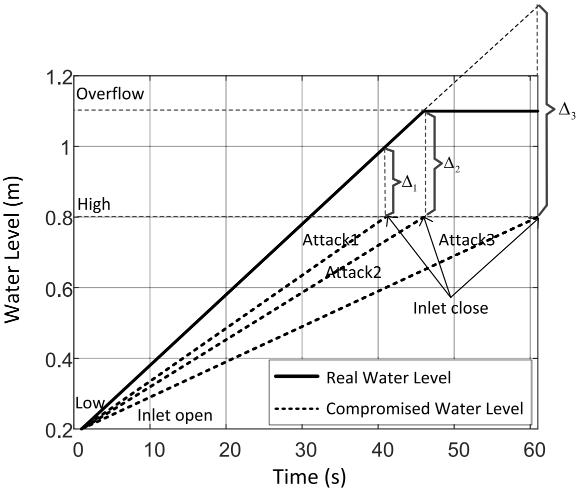

Fig. 3 Different attacks on the water level control system

According to the maximum deviations (Δ) caused by the three attacks, we can draw a conclusion that only a_2 and a_3 can make the tank overflow, and only a_3 achieves a water spill. a_1 is not a successful attack since it yields a smaller deviation $\Delta_1 <$ overflow-high.

This example verifies that the state-of-the-art stateless or stateful statistics cannot identify this kind of stealthy attacks, since only the residual magnitudes ($|y_k^a - \hat{y}_k|$) are investigated but the residual signs are ignored. In order to achieve a successful stealthy attack, the adversary has to make the residual signs follow certain regularities. In this example, the residuals generated during a surge stealthy attack are denoted by:

$$r_k = y_k^a - \hat{y}_k = \begin{cases} -(\tau + \delta), & k = 1 \\ -\delta, & k > 1 \end{cases} \tag{14}$$

Negative signs of residuals enable the adversary to inject enough false data into the system until it crashes. Moveover, in order to complete a successful stealthy attack as quickly as possible, the adversary keeps the residual magnitudes as large as possible under the premise of not being detected. The two features make the residuals generated during a stealthy attack exhibit significant skewness when compared to Gaussian noises. Based on the new discovery, we propose a novel stealthy attack detection technique based on residual skewness analysis.

4.2 Detecting Stealthy Attacks Based on Residual Skewness Analysis

The proposed stealthy attack detection approach mainly includes three steps as follows:

(1) Estimate parameters of the normal residual distribution. Suppose that the attack-free forecasting residuals follow a normal distribution. A priori residual distribution is helpful to stealthy attack detection. Therefore, we first collect a series of attack-free residuals by operating the target ICS in "air-gapped" separation for a period of time and then estimate the two parameters (mean μ and variance σ^2) of the normal residual distribution using the maximum likelihood estimation (MLE) method as follows:

$$\mu = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{15}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \tag{16}$$

where x_i is the ith value of the attack-free residual sequence, and $\bar{x}$ denotes the mean value.

(2) Compute the skewness coefficients of the residuals to be tested. During the stealthy attack detection, we first generate an artificial random sequence $\boldsymbol{r}_{\text{rand}}$ following the normal distribution estimated above (i.e., $\boldsymbol{r}_{\text{rand}} \sim \mathcal{N}(\mu, \sigma^2)$). After that, we replace a small proportion of entries in the original residual sequence $\boldsymbol{r}_o$ to be tested with $\boldsymbol{r}_{\text{rand}}$ and generate a new sequence $\boldsymbol{r}_{\text{test}}$ for testing. Here, we define an new operator ⊎ to denote the sequence replacement operation as follows:

$$\boldsymbol{r}_{\text{test}} = \boldsymbol{r}_o \uplus \boldsymbol{r}_{\text{rand}}, \tag{17}$$

where $L(\boldsymbol{r}_{\text{rand}})/L(\boldsymbol{r}_o) \approx \theta$, and $L(\cdot)$ denotes the length of a sequence and θ is a positive real value around 5%. The procedure of the sequence replacement is shown in Fig. 4. Afterwards, we compute the skewness coefficient (SC) of the new residual sequence $\boldsymbol{r}_{\text{test}}$ as follows:

$$\text{SC} = \frac{\sum_{i=1}^{l}(r_i - \bar{r})^3}{{\sigma_r}^3}, \tag{18}$$

where l is the length of $\boldsymbol{r}_{\text{test}}$, r_i is the ith entry in $\boldsymbol{r}_{\text{test}}$, $\bar{r}$ and σ_r are the mean value and standard deviation of $\boldsymbol{r}_{\text{test}}$, respectively. If the residuals are set equal to $-\delta$ or δ by the adversary during a stealthy attack, and a small portion of residuals are replaced with normal residuals, the residual distribution becomes right-skewed or left-skewed (i.e., the tail is on the right or left side of the distribution), as shown in Fig. 5. This feature can help us identify the counterfeited residuals and further detect stealthy attacks.

(3) Detecting stealthy attacks according to the skewness coefficients of residuals. Generally, there are two kinds of industrial control scenarios: a larger or a smaller value of a process variable indicates a more dangerous system state. In the first scenario, the attacker attempts to counterfeit negative residuals persistently. In order to eliminate the negative residuals, the controller generates commands to increase the value until the system crashes. However, in this case, the skewness coefficient of the observed residuals is greater than 0, since the residual distribution is right-skewed as shown in Fig. 5a, indicating the occurrence of a stealthy attack. The second scenario is just the

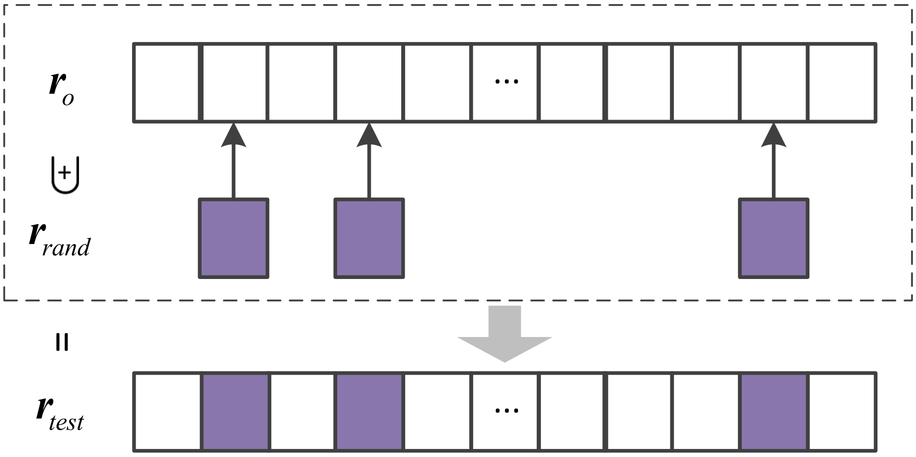

Fig. 4 The sequence replacement procedure

opposite. The attacker tries to counterfeit positive residuals, making the real value of the target process variable decrease over time until the system crashes. In this scenario, the skewness coefficient of the observed residuals is negative, since the residual distribution is left-skewed as illustrated in Fig. 5b.

Therefore, we should fully understand the characteristics of the target ICS before intrusion detection, i.e., which scenario the system belongs to. During attack detection, the skewness coefficients of residuals are computed and investigated over time. If the sign of the skewness efficient conforms to the current control scenario and its absolute value exceeds a predefined positive threshold ϵ (i.e., $|SC| > \epsilon$), a stealthy attack is detected and an alarm is raised. For simplicity, we can only investigate the absolute value of the skewness coefficient for attack detection. However, its sign can help the system operator better understand the adversary's intentions and then make appropriate strategies for system recovery. The entire procedure of the Detecting Stealthy Attacks based on Residual Skewness Analysis algorithm, or "DSARSA" for short, is summarized in Algorithm 1.

In this algorithm, lines 1 and 2 estimate the state-space model and the normal distribution parameters of the attack-free residuals. Line 3 defines a counter used in attack detection. Lines 4 to 26 perform the stealthy attack detection procedure. Lines 5 to 7 present the prediction procedure of Kalman Filtering, and the updating procedure of Kalman Filtering is described by Lines 22 to 24. Lines 8 and 9 compute the current forecasting residual. The skewness coefficient of the residual sequence to be tested is computed by lines 10 to 21. If the absolute value of the skewness coefficient exceeds the detection threshold ϵ, the detection procedure is terminated, and a flag F indicating the occurrence of a stealthy attack is returned by the algorithm and triggers an alarm(lines 17 to 20, 27). Once the alarm is handled properly and the system goes back to safety, the detection procedure restarts.

5 Experimental

In this section, we study the effectiveness of the stealthy attack detection approach based on residual skewness analysis by conducting experiments in a Matlab-Simulink environment.

A water level control system and a water's pH value control system are simulated in our experiment. Both of them are typical ICS as discussed in [16]. Note that the proposed approach can apply to a variety of ICS in addition to the two experimental systems as long as the state-space model of the system can be constructed.

The first system has been discussed as a motivating example in Section 4.1. The dynamics of the water level in the tank can be described by a well-known LDS model derived from the mass balance equation. For simplicity, we assume that the cross-sectional area of the tank is 1 m^2, and the outlet pump keeps working when the system operates normally. The inlet pump should be turned off when the water level exceeds 0.8 m and be turned on again when the water level drops below 0.2 m. Water spill occurs at 1.1 m.

The water's pH value control system is a more complex non-linear system as presented in [16]. The HCl dosage determines the pH value of the water. The HCl pump starts to dose HCl into the water if the pH value exceeds 7.05, and the pump is turned off if the pH value drops below 6.95. Figure 6 depicts the actions (ON/OFF) of the HCl pump and the water's pH values responding to it. The time-delay feature of the system causes the wide oscillations of the pH response curve. The nonlinearity and high latency make it difficult to drive a LDS model from

Algorithm 1: *DSARSA* Algorithm

```
Input: attack-free residual sequence $\boldsymbol{r}$, detection
       threshold $\epsilon$, length $l$ of the residual sequence
       to be tested, ratio $\theta$ of residuals to be
       replaced, initial parameters $\boldsymbol{x}_0$, $\boldsymbol{P}_0$, $\boldsymbol{Q}_0$ and $\boldsymbol{R}_0$
Output: detection result $F$
Estimate the system state-space model ($\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{K}, \boldsymbol{C}$
and $\boldsymbol{D}$);
Estimate the normal distribution parameters $\mu$ and
$\sigma^2$ of the attack-free residuals $\boldsymbol{r}$;
$k \leftarrow 1$;
while true do
    $\boldsymbol{x}_k^- \leftarrow \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{u}_{k-1}$;
    $\boldsymbol{P}_k^- \leftarrow \boldsymbol{A}\boldsymbol{P}_{k-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{K}\boldsymbol{Q}_{k-1}\boldsymbol{K}^{\mathrm{T}}$;
    $\hat{y}_k \leftarrow \boldsymbol{C}\boldsymbol{x}_k^-$;
    get the observed value of the target process
    variable $y_k$;
    $r_k \leftarrow y_k - \hat{y}_k$;
    if $k \geq l$ then
        $\boldsymbol{r}_o \leftarrow \{r_{k-l+1}, \ldots, r_k\}$;
        generate a random residual sequence of
        length $\lfloor l \times \theta \rfloor$: $\boldsymbol{r}_{rand} \sim \mathcal{N}(\mu, \sigma^2)$;
        $\boldsymbol{r}_{test} \leftarrow \boldsymbol{r} \uplus \boldsymbol{r}_{rand}$ ;
        $\bar{r} \leftarrow \frac{1}{l}\sum_{i=k-l+1}^{k} r_i$;
        $\sigma_r \leftarrow \sqrt{\frac{1}{l}\sum_{i=k-l+1}^{k}(r_i - \bar{r})^2}$;
        $SC(\boldsymbol{r}_{test}) \leftarrow [\sum_{i=k-l+1}^{k}(r_i - \bar{r})^3]/{\sigma_r}^3$;
        if $|SC(\boldsymbol{r}_{test})| > \epsilon$ then
            $F \leftarrow TRUE$;
            break;
        end
    end
    $\boldsymbol{KAL}_k \leftarrow \boldsymbol{P}_k^- \boldsymbol{C}^{\mathrm{T}}[\boldsymbol{C}\boldsymbol{P}_k^-\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{R}]^{-1}$;
    $\boldsymbol{x}_k \leftarrow \boldsymbol{x}_k^- + \boldsymbol{KAL}_k[\boldsymbol{y}_k - \boldsymbol{C}\boldsymbol{x}_k^-]$;
    $\boldsymbol{P}_k \leftarrow [\boldsymbol{I} - \boldsymbol{KAL}_k\boldsymbol{C}]\,\boldsymbol{P}_k^-$;
    $k \leftarrow k + 1$;
end
return $F$;
```

first principles. Therefore, we use system identification techniques to build a high-order LDS model to simulate the system dynamics approximately.

6 Results and discussion

On the two simulated ICS, we launch surge stealthy attacks. During attack detection, we set the length of the residual sequence for testing equal to 100 and the parameter θ equal to 5%. Then, we investigate the residual sequence $\{r_{k-99}, \ldots, r_{k-1}, r_k\}$ at each step $k \geq 100$.

In the water level control system, the simulated surge stealthy attack starts from 201 s, as illustrated in Fig. 7a. After that, the deviation between the sensor reading and the real water level in the tank increases persistently until the water spill occurs at 286 s. Figure 7b shows the residuals between the forecasted and measured water levels. It can be seen from Fig. 7c that the skewness coefficient curve stays close to 0 from 1 s to 200 s, but starts to rise significantly after 200 s, indicating the occurrence of the stealthy attack. Additionally, the positive skewness coefficients indicate a right-skewed residual distribution. In other words, there is a small number of large values in the right-hand tail of the distribution, which comes from the artificial random sequence $\boldsymbol{r}_{\mathrm{rand}}$, and a large number of small values in the left hand, which comes from the original residual sequence $\boldsymbol{r}_o$ for testing. As a result,

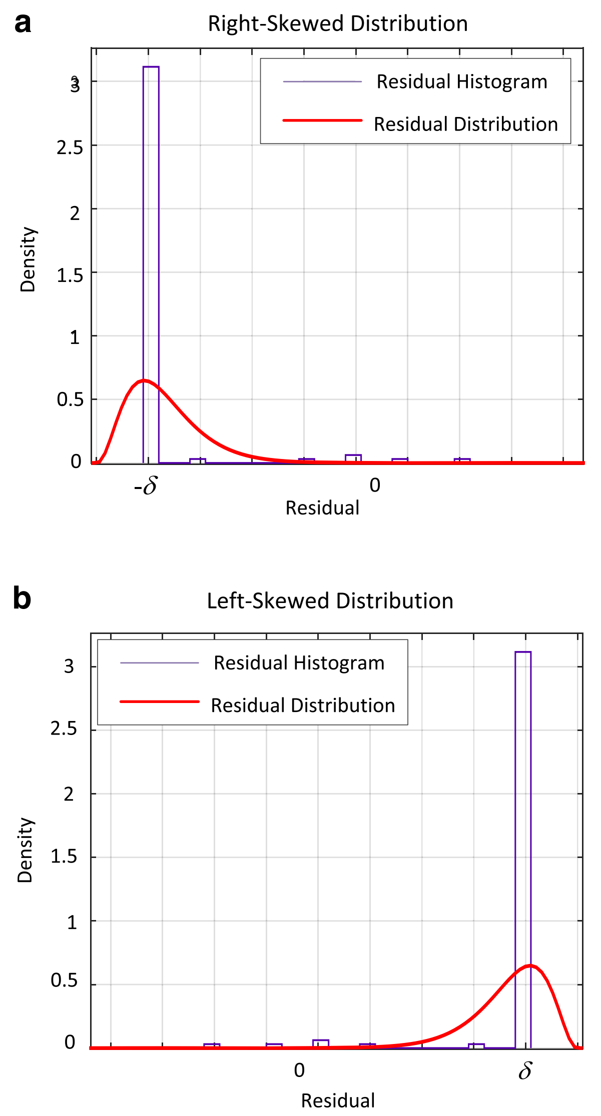

Fig. 5 The right-skewed (**a**) and left-skewed (**b**) residual distributions. The purple line and red line denote the residual histogram and the fitted residual distribution, respectively

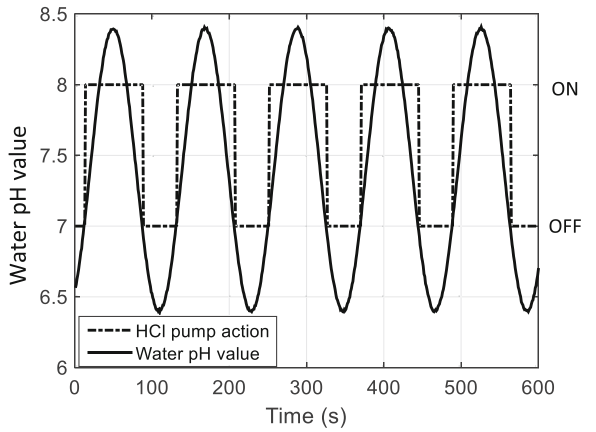

Fig. 6 The relationship between the water's pH value and the actions of the HCl pump

we can draw a conclusion that the attacker attempts to deceive the controller with the fake negative residuals and mislead the controller into making opposite decisions until the tank overflows. Figure 7d to f show the intrusion process, the compromised residuals and the detection result on the water's pH value control system. The stealthy attack starts from 301 s and the skewness coefficient curve starts to decline near 301 s, which indicates a left-skewed residual distribution, i.e., the tail is in the left hand. In this scenario, the attacker tries to counterfeit positive residuals. Accordingly, the deceived controller keeps increasing the HCl dosage into the water until the water container is corroded. Figure 8 shows that the counterfeited residuals fluctuate randomly in a small range above $-\delta$ or under δ, and our detection scheme can still detect this variant of surge stealthy attacks successfully. The experimental results indicate that the residual skewness coefficient is sensitive to the occurrence of stealthy attacks and verify the excellent detection ability of the proposed approach.

Additionally, skilled attackers may replace some entries in the residual sequence with a series of random values (i.e., $\{r_i\} \sim \mathcal{N}(\mu, \sigma^2)$), trying to bypass the intrusion detection system. Figure 9 illustrates that the attacker replace 10% of entries in the residual sequence with random values. In this case, the proposed detection scheme is still capable of identifying this kind of advanced stealthy attacks effectively (i.e., the skewness coefficient curve starts to rise or decline sharply from a certain time point), although the convergent absolute values of skewness coefficients are smaller than those in the above two attack scenarios shown in Figs. 7 and 8. However, it is more difficult for the adversary to achieve his goal if the ratio of the random values becomes higher, so we study the impacts of the ratio of random values on the time to achieve attack goals and the detection ability of our approach.

Figure 10a and c show the impacts of the ratio of the random residuals on the time to achieve attack goals on the water level control system and the water's pH value control system, respectively. We can see that the time to achieve attack goals increases quickly as the ratio of the random residuals rises, especially when the ratio exceeds 60%. Figure 10b and d show that the ratio of the random residuals can also weaken the detection ability of

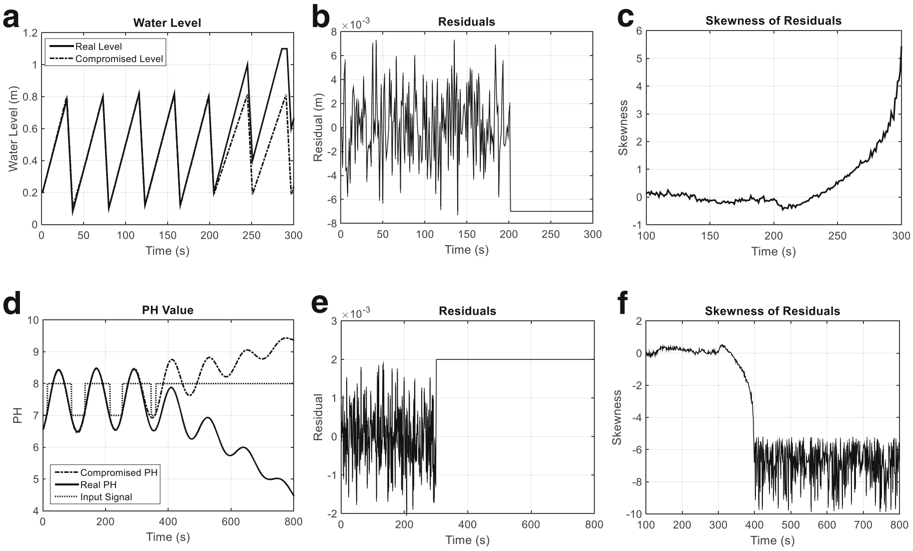

Fig. 7 Detection of surge stealthy attacks. When the surge stealthy attacks in (**b**) and (**e**) are conducted on the water level control system and the pH value control system, respectively, the real water level shown in (**a**) deviates significantly from the seemingly normal but compromised water level, and the real pH value shown in (**d**) also deviates observably from the seemingly normal but compromised pH value. The significant changes of skewness coefficients of residuals shown in (**c**) and (**f**) indicate that the approach proposed can identity surge stealthy attacks effectively

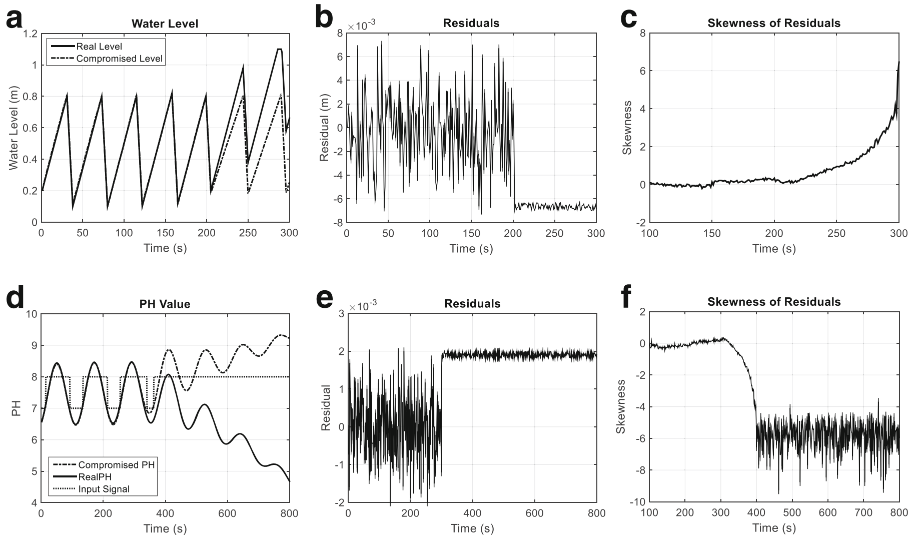

Fig. 8 Detection of the variant of surge stealthy attacks. When the variant of surge stealthy attacks in (**b**) and (**e**) are conducted on the water level control system and the pH value control system, respectively, the real water level shown in (**a**) deviates significantly from the seemingly normal but compromised water level, and the real pH value shown in (**d**) also deviates observably from the seemingly normal but compromised pH value. The significant changes of skewness coefficients of residuals shown in (**c**) and (**f**) indicate that the approach proposed can identity the variant of surge stealthy attacks effectively

our approach. When the ratio is less than 70%, the convergent values of skewness coefficients are significantly different from 0 (i.e., greater than 0 in the water level control system and less than 0 in the water's pH value control system). However, when the ratio reaches or exceeds 80%, it is not easy for our detection scheme to identify the stealthy attack. Additionally, when the ratio exceeds 80%, the stealthy attack detection technique based on residual permutation entropy [23] cannot work well either. Therefore, the detection abilities of the technique proposed and the technique proposed previously are nearly equal. Fortunately, in this case, it takes a much longer time to achieve the attack goals, so attackers are generally unwilling to counterfeit so many random residuals during an

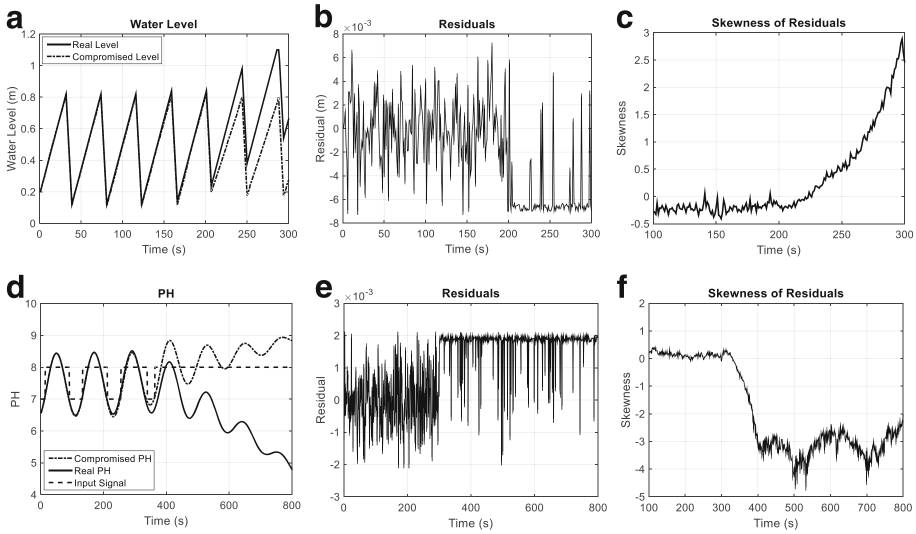

Fig. 9 Detection of stealthy attacks with a certain percentage of random residuals. When stealthy attacks with a certain percentage of random residuals in (**b**) and (**e**) are conducted on the water level control system and the pH value control system, respectively, the real water level shown in (**a**) deviates significantly from the seemingly normal but compromised water level, and the real pH value shown in (**d**) also deviates observably from the seemingly normal but compromised pH value. The significant changes of skewness coefficients of residuals shown in (**c**) and (**f**) indicate that the approach proposed can identity stealthy attacks with a certain percentage of random residuals effectively

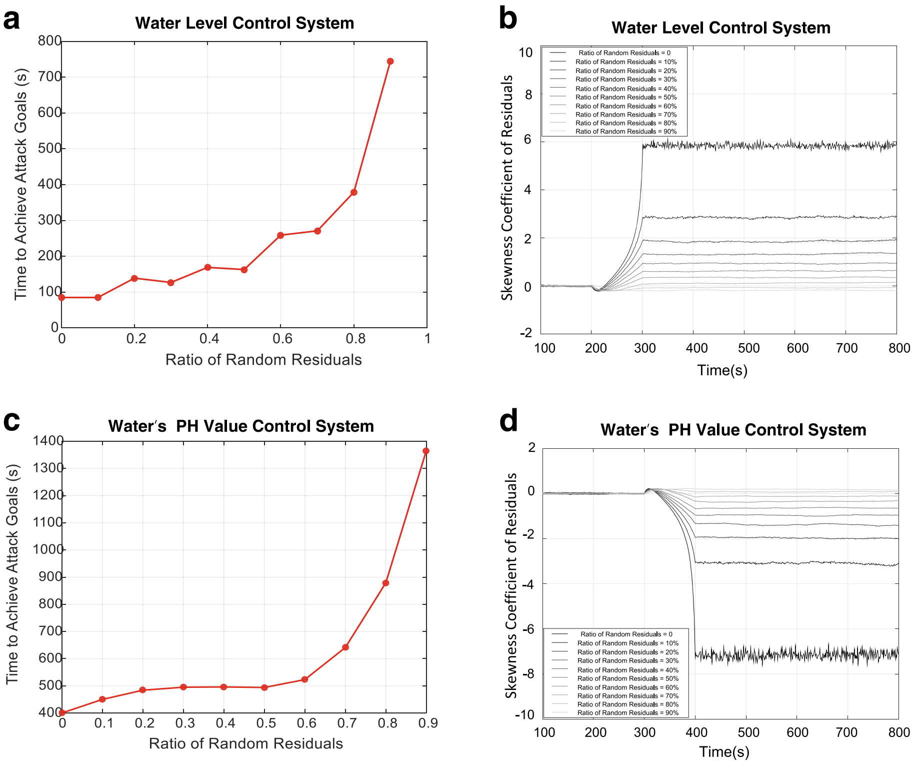

Fig. 10 Impacts of the ratio of random residuals on attack time and detection ability. Time to achieve attack goals becomes longer as the ratio of random residuals increases in the water level control system (**a**) and the pH value control system (**c**). Additionally, when the ratio of random residuals increases, the change of skewness Coefficient of residuals becomes more subtle, as shown in (**b**) and (**d**)

attack. Hence, the proposed residual skewness analysis-based technique is able to detect stealthy attacks against ICS effectively in most cases.

It is worth noting that there exists an interesting phenomenon in Fig. 10. It can be seen from Fig. 10b that the skewness coefficient curve drops slightly at the beginning of the stealthy attack, and then rises significantly. This phenomenon is caused by a transition from a left-skewed residual distribution to a right-skewed residual distribution, since we investigate a set of time sliding windows of residuals during intrusion detection. At the beginning of a stealthy attack, most of the residuals in the current sliding window are Gaussian noises and only a small portion of counterfeited negative residuals, which results in a left-skewed distribution, so the skewness coefficient is less than 0. As time goes on, the sliding window contains more counterfeited negative residuals and only a small portion of gaussian noises, so the left-skewed

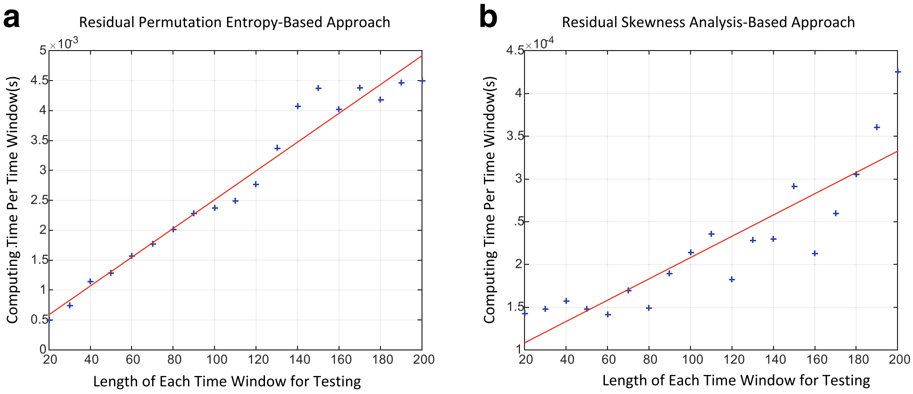

Fig. 11 Detection time comparison of two different approaches, (**a**) the residual permutation entropy-based approach and (**b**) the residual skewness analysis-based approach. The blue points are observed computing time and the red lines are fitted lines for the blue points

distribution turns into a right-skewed distribution, and the skewness coefficient becomes greater than 0. A similar phenomenon occurs in the water's pH value control system as shown in Fig. 10d. A right-skewed distribution turns into a left-skewed distribution.

Additionally, we study the impacts of the length of time windows for testing on the computing time of the detection algorithm, and compare the computing time of the proposed approach with that of the residual permutation entropy-based approach proposed in our previous work [23]. Figure 11 shows that the detection approach proposed in this paper is about ten times faster than the approach proposed previously. Therefore, we can conclude that the residual skewness analysis-based approach is more efficient and more suitable for industrial control systems, which requires low latency and high reliability [1].

7 Conclusions

In this paper, we propose an effective and efficient detection technique against stealthy attacks on ICS. This approach makes full use of the distribution skewness of the forecasting residuals generated during stealthy attacks, which can effectively distinguish the counterfeited residuals from the attack-free residuals. As a result, the occurrence of stealthy attacks can be identified effectively. Comprehensive experimental results verify the effectiveness and efficiency of the proposed approach.

However, this method proposed in this paper still has some shortcomings. The values of the algorithm parameters (e.g., the detection threshold ϵ, the length l of the residual sequence for testing, the ratio θ of residuals to be replaced) should be set manually. Overdependence on human experience may weaken the detection ability of our approach. In the future, we will try to study and model the relationships between the algorithm parameters and the detection performance, based on which to devise an automatic and real-time parameter updating technique, to accomplish the adaptive updating of the parameter values according to the changing detection performance, and evaluate the proposed techniques on larger industrial control systems.

Abbreviations

ARIMA: Auto-regressive integrated moving average; DNP3: Distributed network protocol version 3; CUSUM: Cumulative sum; GOOSE: Generic Object Oriented Substation Event; ICCP: Inter-control center communications protocol; ICS: Industrial control systems; IEC: International electro-technical commission; IDS: Intrusion detection systems; IT: Internet technology; KF: Kalman filtering; LDS: Linear dynamical state-space; MLE: Maximum likelihood estimation; OPC: OLE for process control; PLC: Programmable logic controller; SV: Sample value

Acknowledgements

The authors appreciate the reviewers for their helpful comments and suggestions for the improvement of this paper.

Availability of data and materials

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Funding

This work is supported by the Fundamental Research Funds for the Central Universities (FRF-BD-18-016A), the National Natural Science Foundation of China (61802016 and 61702506), and the National Social Science Foundation of China (17ZDA331), and Project funded by China Postdoctoral Science Foundation (2018M641198).

Authors' contributions

YH contributed to the main idea and designed the mathematical model. HL and HY designed and carried out the simulation and wrote the code of the simulation program. YS analyzed the results. ZW and LS verified the correctness of the proposed technique. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]School of Computer and Communication Engineering, University of Science and Technology Beijing, 30 Xueyuan Road, Beijing 100083, China. [2]Beijing Key Laboratory of IoT Information Security, Institute of Information Engineering, Chinese Academy of Sciences, A 89 Minzhuang Road, Beijing 100195, China. [3]School of Cyber Security, University of Chinese Academy of Sciences, A 89 Minzhuang Road, Beijing 100195, China. [4]Beijing ZKWA Technology CO. LTD., E-Park, Xingshikou Road, Haidian District, Beijing 100093, China.

Received: 14 September 2018 Accepted: 28 February 2019
Published online: 19 March 2019

References

1. K. Stouffer, J. Falco, K. Scarfone, Guide to industrial control systems (ICS) security. NIST Spec. Publ. **800**(82), 16–16 (2011)
2. J. Tian, R. Tan, X. Guan, T. Liu, Enhanced hidden moving target defense in smart grids. IEEE Trans. Smart Grid. **10**(2), 2208–2223 (2019)
3. Y. Mo, T. H.-J. Kim, K. Brancik, D. Dickinson, H. Lee, A. Perrig, B. Sinopoli, Cyber-physical security of a smart grid infrastructure. Proc. IEEE. **100**(1), 195–209 (2012)
4. T. Liu, Y. Liu, Y. Mao, Y. Sun, X. Guan, W. Gong, S. Xiao, A dynamic secret-based encryption scheme for smart grid wireless communication. IEEE Trans. Smart Grid. **5**(3), 1175–1182 (2014)
5. J. Weiss, in *Securing Water and Wastewater Systems*. Industrial control system (ICS) cyber security for water and wastewater systems (Springer, Berlin, 2014), pp. 87–105
6. S. Yin, S. X. Ding, A. Haghani, H. Hao, P. Zhang, A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process. J. Process Control. **22**(9), 1567–1581 (2012)
7. M. reza Akhondi, A. Talevski, S. Carlsen, S. Petersen, in *Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications*. Applications of wireless sensor networks in the oil, gas and resources industries (IEEE, Piscataway, 2010), pp. 941–948
8. A. D. Papadopoulos, A. Tanzman, R. A. Baker Jr, R. G. Belliardi, D. J. Dube, System for remotely accessing an industrial control system over a commercial communications network. U.S. Patent No. 6,061,603 (2000). https://patents.google.com/patent/US6061603A/en
9. L. A. Maglaras, L. H. Kim, H. Janicke, M. A. Ferrag, S. Rallis, P. Fragkou, A. Maglarasf, T. J. Cruz, Cyber security of critical infrastructures. ICT Express. **4**, 42–45 (2018)
10. R. K. Koehler, When the lights go out: vulnerabilities to US critical infrastructure, the Russian cyber threat, and a new way forward. Georgetown Secur. Stud. Rev. **7**(1), 27–36 (2018)
11. J. Staggs, in *Black Hat 2017*. Adventures in attacking wind farm control networks (Black Hat, Las Vegas, 2017). https://www.blackhat.com/docs/us-17/wednesday/us-17-Staggs-Adventures-In-Attacking-Wind-Farm-Control-Networks.pdf
12. D. Ding, Q.-L. Han, Y. Xiang, X. Ge, X.-M. Zhang, A survey on security control and attack detection for industrial cyber-physical systems. Neurocomputing. **275**, 1674–1683 (2018)

13. Z. Ling, K. Liu, Y. Xu, Y. Jin, X. Fu, in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. An end-to-end view of iot security and privacy (IEEE, Piscataway, 2017), pp. 1–7
14. P. Haller, B. Genge, Using sensitivity analysis and cross-association for the design of intrusion detection systems in industrial cyber-physical systems. IEEE Access. **5**, 9336–9347 (2017)
15. Z. Zhang, H. Zhu, S. Luo, Y. Xin, X. Liu, Intrusion detection based on state context and hierarchical trust in wireless sensor networks. IEEE Access. **5**, 12088–12102 (2017)
16. D. I. Urbina, J. A. Giraldo, A. A. Cardenas, N. O. Tippenhauer, J. Valente, M. Faisal, J. Ruths, R. Candell, H. Sandberg, in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. Limiting the impact of stealthy attacks on industrial control systems (ACM, New York, 2016), pp. 1092–1105
17. M. Krotofil, J. Larsen, D. Gollmann, in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*. The process matters: ensuring data veracity in cyber-physical systems (ACM, New York, 2015), pp. 133–144
18. Y. Liu, P. Ning, M. K. Reiter, False data injection attacks against state estimation in electric power grids. ACM Trans. Inf. Syst. Secur. **14**(1), 13 (2011)
19. R. M. Gerdes, C. Winstead, K. Heaslip, in *Proceedings of the 29th Annual Computer Security Applications Conference*. Cps: an efficiency-motivated attack against autonomous vehicular transportation (ACM, New York, 2013), pp. 99–108
20. R. Tan, V. Badrinath Krishna, D. K. Yau, Z. Kalbarczyk, in *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*. Impact of integrity attacks on real-time pricing in smart grids (ACM, New York, 2013), pp. 439–450
21. A. Kleinmann, O. Amichay, A. Wool, D. Tenenbaum, O. Bar, L. Lev, in *Computer Security*. Stealthy deception attacks against SCADA systems (Springer, Berlin, 2017), pp. 93–109
22. E. Kung, S. Dey, L. Shi, The performance and limitations of $epsilon$-stealthy attacks on higher order systems. IEEE rans. Autom. Control. **62**(2), 941–947 (2017)
23. Y. Hu, H. Li, T. H. Luan, A. Yang, L. Sun, Z. Wang, R. Wang, Detecting stealthy attacks on industrial control systems using a permutation entropy-based method. Futur. Gener. Comput. Syst. (2018). ISSN 0167-739X, https://doi.org/10.1016/j.future.2018.07.027
24. P. Stavroulakis, M. Stamp, *Handbook of Information and Communication Security*. (Springer, Berlin, 2010)
25. C.-H. Tsang, S. Kwong, in *Proceedings of 2005 IEEE International Conference on Industrial Technology*. Multi-agent intrusion detection system in industrial network using ant colony clustering approach and unsupervised feature extraction (IEEE, Piscataway, 2005), pp. 51–56
26. H. Wang, *On anomaly detection and defense resource allocation of industrial control networks*. (Zhejiang University, China, 2014)
27. I. Kiss, B. Genge, P. Haller, in *Proceedings of 2015 IEEE 13th International Conference on Industrial Informatics*. A clustering-based approach to detect cyber attacks in process control systems (IEEE, Piscataway, 2015), pp. 142–148
28. T. Vollmer, M. Manic, in *Proceedings of 2009 2nd International Symposium on Resilient Control Systems*. Computationally efficient neural network intrusion security awareness (IEEE, Piscataway, 2009), pp. 25–30
29. O. Linda, T. Vollmer, M. Manic, in *Proceedings of 2009 International Joint Conference on Neural Networks*. Neural network based intrusion detection system for critical infrastructures (IEEE, Piscataway, 2009), pp. 1827–1834
30. O. Linda, M. Manic, T. Vollmer, J. Wright, in *Proceedings of 2011 IEEE Symposium on Computational Intelligence in Cyber Security*. Fuzzy logic based anomaly detection for embedded network security cyber sensor (IEEE, Piscataway, 2011), pp. 202–209
31. O. Linda, M. Manic, J. Alves-Foss, T. Vollmer, in *Proceedings of 2011 4th International Symposium on Resilient Control Systems*. Towards resilient critical infrastructures: application of type-2 fuzzy logic in embedded network security cyber sensor (IEEE, Piscataway, 2011), pp. 26–32
32. O. Linda, M. Manic, T. Vollmer, in *Proceedings of 2012 5th International Symposium on Resilient Control Systems*. Improving cyber-security of smart grid systems via anomaly detection and linguistic domain knowledge (IEEE, Piscataway, 2012), pp. 48–54
33. L. A. Maglaras, J. Jiang, in *Science and Information Conference*. Intrusion detection in SCADA systems using machine learning techniques (IEEE, Piscataway, 2014), pp. 626–631
34. Y. Luo, *Reasearch and design on intrusion detection methods for industrial control system. PhD thesis*. (Zhejiang University, China, 2013)
35. A. Javaid, Q. Niyaz, W. Sun, M. Alam, in *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*. A deep learning approach for network intrusion detection system (ICST, Boston, 2016), pp. 21–26
36. M. H. Aghdam, P. Kabiri, Feature selection for intrusion detection system using ant colony optimization. IJ Netw. Secur. **18**(3), 420–432 (2016)
37. C. Hou, J. Hanhong, W. Rui, L. Liu, A probabilistic principal component analysis approach for detecting traffic anomaly in industrial networks. J. Xi'an Jiaotong Univ. **46**(2), 78–83 (2012)
38. S. Cheung, B. Dutertre, M. Fong, U. Lindqvist, K. Skinner, A. Valdes, in *Proceedings of the SCADA Security Scientific Symposium, vol. 46*. Using model-based intrusion detection for SCADA networks (Citeseer, Princeton, 2007), pp. 1–12
39. T. Morris, R. Vaughn, Y. Dandass, in *Proceedings of the 45th Hawaii International Conference on System Science*. A retrofit network intrusion detection system for MODBUS RTU and ASCII industrial control systems (IEEE, Piscataway, 2012), pp. 2338–2345
40. H. Lin, A. Slagell, C. Di Martino, Z. Kalbarczyk, R. K. Iyer, in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*. Adapting Bro into SCADA: building a specification-based intrusion detection system for the dnp3 protocol (ACM, New York, 2013), p. 5
41. J. Hong, C.-C. Liu, M. Govindarasu, in *Innovative Smart Grid Technologies Conference (ISGT), 2014 IEEE PES*. Detection of cyber intrusions using network-based multicast messages for substation automation (IEEE, Piscataway, 2014), pp. 1–5
42. H. Hadeli, R. Schierholz, M. Braendle, C. Tuduce, in *Proceedings of 2009 IEEE Conference on Emerging Technologies & Factory Automation*. Leveraging determinism in industrial control systems for advanced anomaly detection and reliable security configuration (IEEE, Piscataway, 2009), pp. 1–8
43. E. Colbert, D. Sullivan, S. Hutchinson, K. Renard, S. Smith, in *Proceedings of the 11th International Conference on Cyber Warfare and Security*. A process-oriented intrusion detection method for industrial control systems (Academic Conferences International Limited, England, 2016), p. 497
44. D. Hadžiosmanović, R. Sommer, E. Zambon, P. H. Hartel, in *Proceedings of the 30th Annual Computer Security Applications Conference*. Through the eye of the PLC: semantic security monitoring for industrial processes (ACM, New York, 2014), pp. 126–135
45. A. Carcano, A. Coletta, M. Guglielmi, M. Masera, I. N. Fovino, A. Trombetta, A multidimensional critical state analysis for detecting intrusions in SCADA systems. IEEE Trans. Ind. Inform. **7**(2), 179–186 (2011)
46. R. J. Patton, Robustness in model-based fault diagnosis: the 1995 situation. Annu. Rev. Control. **21**, 103–123 (1997)
47. A. A. Cárdenas, S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang, S. Sastry, in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*. Attacks against process control systems: risk assessment, detection, and response (ACM, New York, 2011), pp. 355–366
48. S. Sridhar, M. Govindarasu, Model-based attack detection and mitigation for automatic generation control. IEEE Trans. Smart Grid. **5**(2), 580–591 (2014)
49. S. Amin, X. Litrico, S. Sastry, A. M. Bayen, Cyber security of water SCADA systems–Part I: Analysis and experimentation of stealthy deception attacks. IEEE Trans. Control Syst. Technol. **21**(5), 1963–1970 (2013)
50. D. I. Urbina, J. Giraldo, A. A. Cardenas, J. Valente, M. Faisal, N. O. Tippenhauer, J. Ruths, R. Candell, H. Sandberg, *Survey and New Directions for Physics-based Attack Detection in Control Systems*. (US Department of Commerce, National Institute of Standards and Technology, Gaithersburg, 2016)
51. D. W. Clarke, Application of generalized predictive control to industrial processes. IEEE Control. Syst. Mag. **8**(2), 49–55 (1988)
52. S. J. Qin, T. A. Badgwell, A survey of industrial model predictive control technology. Control. Eng. Pract. **11**(7), 733–764 (2003)
53. W. S. Levine, *The Control Handbook*. (CRC press, Boca Raton, 1996)
54. M. S. Grewal, in *International Encyclopedia of Statistical Science*. Kalman filtering (Springer, Berlin, 2011), pp. 705–708