Received October 31, 2018, accepted November 26, 2018, date of publication December 3, 2018,
date of current version December 27, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2884504

A Novel Data Integrity Attack Detection Algorithm Based on Improved Grey Relational Analysis

ZHENGDAO ZHANG, YUNFEI WANG, AND LINBO XIE
Engineering Research Center of Internet of Things Applied Technology, School of IoT Engineering, Ministry of Education,
Jiangnan University, Wuxi 214122, China

Corresponding author: Zhengdao Zhang (wxzzd@jiangnan.edu.cn)

This work was supported in part by the National Key R&D Program of China under Grant 2018YFD0400900 and in part by the Education Ministry and China Mobile Science Research Foundation under Grant MCM20170204.

ABSTRACT False data injection (FDI) attack is the most common data integrity attack, and it is also one of the most serious threats in industrial control systems (ICSs). Although many detection approaches are developed with burgeoning research interests, the technical capability of existing detection methods is still insufficient because the stealth FDI attacks have been proven to bypass bad data detector. In this paper, a novel data analytical algorithm is proposed to identify the stealth FDI attacks in ICSs according to the correlation analysis. First, we evaluate the correlation between measurements and control variables based on an improved grey relational analysis. Then, SVM is used to classify the FDI attack according to the values of correlation. Through a reliable semi-physical simulation testbed whose virtual plant corresponds to a 330 MW boiler-turbine unit, two FDI attacks that can bypass the detection system are studied. A dataset, which contains the normal data and attack data, is created from the testbed to verify the effectiveness of the proposed algorithm. In addition, the performance of the proposed algorithm is also studied based on the new gas pipeline dataset that is collected by the distributed analytics and security institute in Mississippi State University. Such a novel algorithm, which has better accuracy and reliability, is compared with the state of the art based on the data analysis.

INDEX TERMS Industry control systems, false data injection, attack detection, grey correlation analysis, SVM.

I. INTRODUCTION

With the innovation of information and communication technologies, industrial control systems (ICSs) have changed a lot. Generally, a modern ICS can facilitate the operation of supervisory control and data acquisition (SCADA) system, and improve the production efficiency at the same time. Nowadays, the cybersecurity of ICSs has drawn considerable attentions [1]. In 2010, the Stuxnet worm attacked several industrial workspaces in Iran, including uranium enrichment plant [2]. In 2014, the attacker succeeded in accessing to the ICS of a steel plant in Germany [3]. In 2015, the SCADA system of power gird was remotely accessed by hackers in Ukraine [4].

One of the most common data integrity attacks in an ICS is false data injection (FDI), which was first named in 2009 by Liu *et al.* [5], Liang *et al.* [6]. From the control perspective, the purpose of the FDI attacker is to manipulate the industrial process by modifying control variables. There are two ways for FDI attacker to deploy the attacks on ICSs. One is that the attacker injects malicious data into control variables directly. Another is an indirect way in which the attacker changes the control variables by tampering with sensor measurements. FDI attack is a stealth attack that can inject the attack data into the target system in undetectable ways [7], [8]. Both types of attacks can bypass the bad data detection (BDD), which is a common abnormal detection method in the detection system of an ICS [9], [10].

According to the mechanisms of detection algorithm, the FDI attack detection algorithms can be boiled down to model-based methods and data-based methods. The study of model-based methods has attracted more researchers in the past years [11]–[15]. These methods require an exact model of an ICS, and their performance significantly depend on the modelling accuracy. With the development of information technology, the scale of ICSs has gradually become larger. Generally, it is difficult to get an accurate ICS model and apply model-based methods. Meanwhile, we can get a huge amount of data from ICS to achieve reasonable data-based methods. Up to now, data-based detection methods for FDI attacks include statistics, machine learning, and data mining. Kosut et al. proposed a method, which utilized a Bayesian framework to describe the FDI attacks problem [16]. A heuristic approach was adopted to detect FDI attacks with low computational cost. Valenzuela et al. developed an approach to supervise power flow results and identify anomalies in power system based on principle component analysis (PCA), where PCA is used to separate power flow variability into regular and irregular subspaces [17]. Pan *et al.* [18] presented a hybrid intrusion detection system with data mining method, which can automatically and accurately learn patterns of system scenarios for anomaly detection. Ozay *et al.* [19] showed that the performance of machine learning algorithms for detecting attacks is better than attack detection algorithms employed state vector estimation methods. Wang *et al.* [20] detected FDI attacks by employing the margin setting algorithm (MSA), which is a relatively new machine learning algorithm. MSA achieved high accuracy on detection of the FDI attacks. Foroutan and Salmasi proposed a statistical anomaly detection approach based on Gaussian mixture model (GMM), which was used to cluster the data, and then the FDI attack data was separated from the dataset. Most of the data-based detection methods are focused on the Smart Grid system. These methods only use measurements to detect the measurement injection attack. The model of the Smart Grid system is an auto-regression model. Comparing with Smart Grid system, behaves of ICS rely on the control variables. The variation of measurement can be caused by data injection attack or normal control variation. In contrast with the measurement injection attack, there are control variable tampering attacks in the ICS. Therefore, the attack detection problem in ICS is more complex. The data-based detection algorithms for Smart Grid system are not always suitable for the cyber-attack detection of the ICSs,which contain control variables. The detection algorithm should have the capability of detecting the two FDI attacks by considering the reliability and universality.

The motivation of this work is to design an attack detection algorithm suitable for detecting the two FDI attacks and remains effective when dealing with huge data from the ICS. Therefore, a novel FDI detection algorithm based on the improved grey relational analysis (GRA) is proposed. GRA is a mathematical model for studying the degree of correlation between internal factors of the system [23]. It can be used to analyze data correlation, especially for incomplete and inexact data. GRA has been successfully applied in many fields such as economics, military, transportation, industry, agriculture, and so on. Although the GRA model has been widely used, it still has some inevitable limitations [24]. The GRA calculation formula has no clear geometric meaning. It does not satisfy the normative principle. And, the positive and negative areas in GRA will cancel each other out during the integration process. In this work, we improve GRA for calculating the correlation degree and selecting the correlation combinations of control variables and measurements. First, the control variables and measurements are defined as dependent variables and master variables, respectively. After that, the improved GRA only calculates the correlation degree when the master variable changes. It can be found in our simulations that the improved GRA can extract the difference between attack data and normal data more efficiently than the common GRA does. The main contributions of this paper are summarized as follows:

1) We improve GRA for analyzing the data correlation in ICS. To deal with huge data, we assign the inputs and outputs of ICS as dependent variables and master variables, respectively. Then, we can ignore the calculation about dependent variables.

2) Based on the proposed data analysis algorithm, we further develop an attack detection method to detect two kinds of FDI attacks. The underlying algorithm is applied to a semi-physical simulation testbed to demonstrate its practical potentials. The experimental results indicate that our detection method is more effective than the SVM based method in the sense of $F1$ score.

The rest of this paper is organized as follows. In section II, we first introduce the system model and the model of the two FDI attacks, and then, we analyze the impacts of FDI attacks on system variable dependencies. Section III shows the insufficiency and improvement of the GRA. Section IV describes the detection procedure against the two FDI attacks. Section V demonstrates the experiments. Finally, conclusions are made in section VI.

II. THE SYSTEM MODEL AND TWO FDI ATTACKS

A. THE SYSTEM MODEL

The composition of the ICS can be regarded as a cyber-network with a physical plant, which is schematically represented in Figure 1.

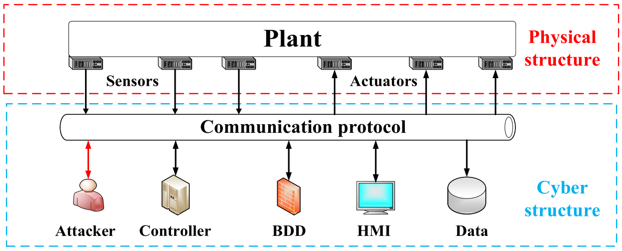


FIGURE 1. The architecture of the ICS.

Specifically, the cyber structure consists of the communication protocols and the components served to supervise and operate the physical plant, while the physical plant encompasses the field devices to be managed. The state equations of the physical plant can be described as

$$x(k+1) = Gx(k) + Bu(k), \quad (1)$$

$$y(k) = Cx(k) + e(k), \quad (2)$$

where $y(k) \in \mathbb{R}^m$ is the measurement, $u(k) \in \mathbb{R}^l$ is the control variable, and $x(k) \in \mathbb{R}^n$ is the state variable at time instant k. $G \in \mathbb{R}^{n\times n}$ is the system matrix, $B \in \mathbb{R}^{n\times l}$ is the input matrix, $C \in \mathbb{R}^{m\times n}$ is the output matrix, and $e \in \mathbb{R}^m$ is the measurement noise. Then, the mathematical model of cyber part can be expressed as

$$u(k+1) = Hy(k), \tag{3}$$

where $H \in \mathbb{R}^{l\times m}$ is the control matrix.

B. BDD AND TWO FDI ATTACKS

BDD is used for detecting the bad data in measurement based on state evaluation. The state estimation problem in ICS is to obtain an estimate $\hat{x}$ that minimizes the weighted least-squares error [25]

$$\hat{x}(k) = \arg\min_{x}\,[y(k) - Cx(k)]^T W[y(k) - Cx(k)], \tag{4}$$

where $W \in \mathbb{R}^{m\times m}$ is a diagonal matrix whose entries are reciprocals of the variances of measurement error e.

The solution of the (4) can be obtained as

$$\hat{x}(k) = (C^T WC)^{-1}C^T Wy(k). \tag{5}$$

Then, the bad data can be recognized when the measurement residual r_0 is as follow

$$r_0 = ||y(k) - C\hat{x}(k)||_2 > \tau, \tag{6}$$

where $||\cdot||_2$ is the L_2-norm and τ is a threshold.

From the view of control, the purpose of the FDI attack is to prevent the target system from proper working by changing control variables. The sign of a successful attack is that the process of injecting attack data is not detected by the detection module [26]. As shown in Figure 2, there are two stealth FDI attacks that can modify the control variables and bypass the BDD. The first one is the measurement injection attack (A1) that changes the control variables by modify the measurement indirectly. The measurement injection attack model is as follow

$$y_a(k) = Cx(k) + \Delta y(k) + e(k), \tag{7}$$

$$u_a(k) = Hy_a(k), \tag{8}$$

where $\Delta y(k)$ is the injection signal.

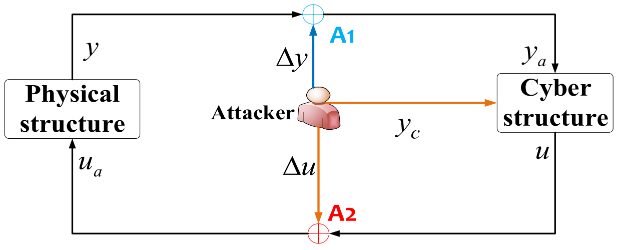


FIGURE 2. Two FDI attacks can bypass the BDD.

We use $Ca(k)$ instead of $\Delta y(k)$, the (7) can be represented as

$$y_a(k) = C(x(k) + a(k)) + e(k). \tag{9}$$

The state estimation $\hat{x}_a(k)$ of $y_a(k)$ can be obtained as

$$\hat{x}_a(k) = (C^T WC)^{-1}C^T Wy_a(k) = \hat{x}(k) + a(k). \tag{10}$$

The measurement residual r_1 of $y_a(k)$ can be got as

$$\begin{aligned} r_1 &= ||y_a(k) - C\hat{x}_a(k)||_2 \\ &= ||C(x(k) + a(k)) + e(k) - C(\hat{x}(k) - a(k))||_2 \\ &= ||y(k) - C\hat{x}(k)||_2. \end{aligned} \tag{11}$$

The measurement residual r_1 is same as the normal system which means the measurement injection attack can invalid the BDD.

The second FDI attack is the control variable tampering attack (A2) that injects the false data into the control variables directly and sends cheating value y_c to mask the change in measurement. The model of control variable tampering attack is as follow

$$x_a(k+1) = Gx(k) + B(u(k) + \Delta u(k)), \tag{12}$$

$$\begin{aligned} y_a(k+1) &= Cx_a(k+1) + e(k+1) \\ &= C(Gx(k) + B(u(k) + \Delta u(k))) + e(k+1) \\ &= C(Gx(k) + Bu(k)) + CB\Delta u(k) + e(k+1) \\ &= C(x(k+1) + B\Delta u(k)) + e(k+1). \end{aligned} \tag{13}$$

The state estimation $\hat{x}_a(k+1)$ of $y_a(k+1)$ can be obtained as

$$\begin{aligned} \hat{x}_a(k+1) &= (C^T WC)^{-1}C^T Wy_a(k+1) \\ &= \hat{x}(k+1) + B\Delta u(k). \end{aligned} \tag{14}$$

The measurement residual r_2 of $y_a(k+1)$ can be obtained as

$$\begin{aligned} r_2 &= ||y_a(k+1) - C\hat{x}_a(k+1)||_2 \\ &= ||C(x(k+1) + B\Delta u(k)) + e(k+1) - C(\hat{x}(k+1) \\ &\quad + B\Delta u(k))||_2 \\ &= ||y(k+1) - C\hat{x}(k+1)||_2. \end{aligned} \tag{15}$$

The measurement residual r_2 is same as the normal system which means control variable tampering attack can bypass the BDD.

C. THE RELATIONSHIP BETWEEN TWO FDI ATTACKS AND THE INHERENT CORRELATION IN THE SYSTEM VARIABLES

Refer to (3), the mathematical model of cyber part can be also expressed as

$$\begin{aligned} u_i &= f_i(y) \\ &= h_{i1}y_1 + h_{i2}y_2, \ldots, +h_{im}y_m + \delta_i, \quad 1 \le i \le l, \end{aligned} \tag{16}$$

where δ_i is the nonlinear part of (16).

Let

$$\rho_{ij} = h_{i1}y_1, \ldots, +h_{i(j-1)}y_{j-1}, h_{i(j+1)}y_{j+1}, \ldots, +h_{im}y_m + \delta_i, \tag{17}$$

the (16) can be reformulated as

$$u_i = h_{ij} y_j + \rho_{ij}, \quad 1 \le j \le m. \tag{18}$$

If the correlation degree $g(u_i, y_j)$ between the u_i and y_j satisfies the following conditions

$$\begin{cases} 0 < g(u_i, y_j) \le 1, & h_{ij} > 0 \\ -1 \le g(u_i, y_j) < 0, & h_{ij} < 0 \end{cases} \tag{19}$$

we define that there is correlation in combination of u_i and y_j, and the correlations of all combinations in system variables are called the inherent correlation of the system. Next, the impacts of two FDI attacks on the inherent correlation of the system are analyzed.

For the measurement injection attack, it can be known that the changes of the control variables will be dominated by Δy when the attack is effective. Therefore, there will be a correlation between the control variables and Δy. We also know that measurement noise e is weak, so it is insufficient to cause changes in control variables. Hence, there is no correlation between sensor measurement noise and control variable. In addition, if the attacker wants to complete the attack during the attack time, the attack signal should not be sparse in frequency [27]. It is also known that measurement noise e does not have sparsity in frequency. Based on the above two points, the Compressed Sensing (CS) can be used to separate the attack signal Δy and measurement noise e from the measurement. Then, the measurement injection attack can be detected by determining whether the correlation between the separation signal and the control variable is similar to the inherent correlation between system variables.

For the control variable tampering attack, the attacker can send the cheating data that is numerically similar with the normal measurement to mask the true measurement. However, it is hard for attacker to make the correlations between cheating data and control variables meet the inherent correlation. Therefore, the control variable tampering attack can be detected by using the inherent correlation in normal system.

From the above analysis, we know that both types of FDI attacks can be detected by using the inherent correlation between systems variables. Therefore, the GRA is employed to calculate the correlation degree between the control variable and measurement, and pick out the combinations that have correlation for FDI attack detection.

III. THE INSUFFICIENCY AND IMPROVEMENT OF THE GRA

A. THE INSUFFICIENCY OF THE GRA

GRA is an important part of the grey system theory and gets the degree of correlation between the sequence curves by comparing the geometry similarity [23], [28]. Its performance for analyzing data correlations, especially for incomplete and inexact data, is superior. Although the GRA model has been widely used in many fields, there are still inevitable limitations [24]. The minimum value of relational degree calculated by the GRA is 0.5 which makes the GRA do not satisfy the normative principle and have poor differentiation degrees. The GRA only considers the area between the sequence curves to analysis the correlation degree, which affects the accuracy and rationality of the calculation results greatly. The most significant effect is that the calculation results of the correlation degree do not match the qualitative analysis, when the positive and negative areas appear to cancel each other out during the integration process [29]. In addition, the change in a control variable may come from the change of one or more measurements at some point. So, it is difficult to compare the similarity of the curve geometry between control variable sequence and measurement sequence for calculating the relevance through the GRA. Hence, the capacity of GRA model for analyzing the data in MIMO system is also insufficient. In order to overcome the above problems, we propose an improved GRA model that is based on the slope of the sequence curve.

B. THE IMPROVED GRA MODEL

Suppose that there are master sequence $Q_M \in \mathbb{R}^v$ and dependent sequence $Q_D \in \mathbb{R}^v$. First, the sequences Q_M and Q_D are done dimensionless processing though the transformation operator ϑ. The transformation operator contains initial transformation, averaging transformation, percentage transformation, and multiple transformation [30]. Then, the initial point zeroing sequences are as follows

$$Q_M^0 = \left[q_M^0(1), \ldots, q_M^0(v)\right] = [q_M(1)\vartheta, \ldots, q_M(v)\vartheta], \tag{20}$$

$$Q_D^0 = \left[q_D^0(1), \ldots, q_D^0(v)\right] = [q_D(1)\vartheta, \ldots, q_D(v)\vartheta]. \tag{21}$$

The transformation operator ϑ based on percentage transformation is used to do dimensionless processing for sequences as follow

$$q^0(\varepsilon) = q(\varepsilon)\vartheta = \frac{q(\varepsilon)}{\max\limits_{1\le\varepsilon\le v} \{q(\varepsilon)\}}. \tag{22}$$

The slope change $\Delta\alpha(\eta)$ of the sequence Q^0 in the interval $[\eta, \eta+1]$ are given as

$$\Delta\alpha(\eta) = q^0(\eta+1) - q^0(\eta), \quad \eta = 1, \ldots, v-1. \tag{23}$$

In order to avoid the influence of other variables in the dependent sequence, the relevance is calculated only when the master sequence changes. The grey relational coefficient $\gamma_{M\to D}(\eta)$ can be defined as

$$\gamma_{M\to D}(\eta) = \operatorname{sgn}\left(\frac{\Delta\alpha_D(\eta)}{\Delta\alpha_M(\eta)}\right) \frac{\min(|\Delta\alpha_M(\eta)|, |\Delta\alpha_D(\eta)|)}{\max(|\Delta\alpha_M(\eta)|, |\Delta\alpha_D(\eta)|)}$$
$$s.t.\ |\Delta\alpha_M(\eta)| > \zeta, \tag{24}$$

where $\zeta > 0$ is the threshold of change amplitude of the master sequence. Suppose that the number of grey relational coefficient that satisfies $|\alpha_M(\eta)| > \zeta$ is N. Then, the grey relational degree $\gamma(Q_M \to Q_D)$ between the master sequence and dependent sequence can be represented as

$$\gamma\left(Q_M \rightarrow Q_D\right) = \frac{1}{N}\left(\sum_{\eta=1}^{N} \gamma_{M\rightarrow D}(\eta)\right), \tag{25}$$

Theorem 1: The improved GRA model meets the normative principle.

Proof: From the definition of the correlation coefficient of the improved GRA model, it can be derived that

$$0 \leq |\gamma_{m\rightarrow d}(\eta)| \leq 1. \tag{26}$$

If and only if the slope changes of the dependent sequence follow the slope changes of the master sequence, and the absolute values of the slope changes of the two sequence curves are the same, there is

$$|\gamma_{m\rightarrow d}(\eta)| = 1. \tag{27}$$

Since the improved GRA model measures the overall similarity by the average of the local correlation degree, the improved GRA model satisfies the normative principle.

Theorem 2: The correlation degree calculated by the improved GRA model is only related to the geometry of the sequence curve and has nothing to do with its spatial relative position and distance.

Proof: It can be known from the definition of the correlation coefficient that the magnitude, positive and negative of the correlation degree are only related to the relative slope change of the sequence curve. The larger the ratio of the slope change amount is, and the greater the degree of correlation is. Therefore, Theorem 2 is established.

IV. THE PROPOSED DETECTION ALGORITHM

A. SELECT THE CORRELATION COMBINATIONS BY USING THE IMPROVED GRA

The windowed sampling is considered to sample each control variable and measurement. The correlation between measurement and control variable is calculated in each time window. Suppose that the size of time window is w. The sampled sequences of measurement y_j and the control variable u_i can be obtained as

$$Y_{j,k} = [y_{j,(k-w+1)}, y_{j,(k-w+2)}, \ldots, y_{j,(k)}], \tag{28}$$
$$U_{i,k} = [u_{i,(k-w+1)}, u_{i,(k-w+2)}, \ldots, u_{i,(k)}]. \tag{29}$$

where $k \in [w, T]$ is the sampling time, and T is total number of samples.

Then, the sequences $Y_{j,k}$ and $U_{i,k}$ are regarded as master sequence and dependent sequence, respectively. All combinations $\omega_{ji} = \left(Y_j \rightarrow U_i\right)$ of measurement sequences and control variable sequences can be expressed as

$$\Omega = \left\{\omega_{ji} | j \in [1, m], i \in [1, l]\right\}. \tag{30}$$

The combination ω_{ji} will be selected when

$$\overline{\gamma}(\omega_{ji}) = \frac{1}{T}\sum_{k=w}^{T} \gamma(Y_{j,k} \rightarrow U_{i,k}) > \theta \tag{31}$$

where $\theta \neq 0$ is a threshold to select the correlation combination for FDI attacks detection. Then, the set Ω' that contains all correlation combinations is selected from Ω as follow

$$\Omega' = \left\{\omega_{ji}^1, \omega_{ji}^2, \ldots, \omega_{ji}^{\mu}\right\}, \quad 1 \leq \mu \leq ml \tag{32}$$

The inherent correlation ξ_{yu} of the system can be expressed as

$$\xi_{yu} = [\gamma(\omega_{ji}^1), \gamma(\omega_{ji}^2), \ldots, \gamma(\omega_{ji}^{\mu})] \tag{33}$$

B. THE DETECTION FOR MEASUREMENT INJECTION ATTACK

Based on the analysis in the section II, we use CS to perform noise separation on measurement sequence contained in Ω'. CS is a new way for signal acquisition and processing which reduces the sampling frequency required for reconstruction of original signals greatly [31]. The CS points out that as long as the signal is sparse or has sparseness on some transform basis, it can be mapped to low dimensional space by using an observation matrix that is not related to the transform basis. Then the accurate reconstructed signal can be obtained by solving L_1-norm minimization problem.

To a measurement sequence $Y_{j,k}$, the measurement matrix $\Phi \in \mathbb{R}^{o\times w}$ is used to resample the $Y_{j,k}$. Then, a sample signal $V_{j,k} \in \mathbb{R}^{o}$, $o \ll w$ can be obtained as

$$V_{j,k} = \Phi Y_{j,k}. \tag{34}$$

We have known that the attack signal and measurement noise do not have sparsity in frequency. Therefore, the discrete Fourier transform is chosen as the sparse transform base $\Psi \in \mathbb{R}^{w\times w}$ to convert the time domain signal to the frequency domain. The $Y_{j,k}$ can be given as Ψs_k, where $s_k \in \mathbb{R}^{w}$ is the sparse representation of $Y_{j,k}$. Then the sample signal can be represented as:

$$V_{j,k} = \Phi\Psi s_k = A s_k, \tag{35}$$

where $A = \Phi\Psi$ is the sensing matrix. Suppose that there is a constant $\beta_\ell \in (0, 1)$ such that

$$(1-\beta_\ell)||s_k||_2^2 \leq ||A_\ell s_k||_2^2 \leq (1+\beta_\ell)||s_k||_2^2, \tag{36}$$

where $A_\ell \in \mathbb{R}^{\ell\times w}$ is the arbitrary row submatrix of A. Then s_k can be obtained by solving the L_1-norm minimization problem

$$\min_{s_k} ||s_k||_1$$
$$s.t.\ V_{j,k} = A s_k. \tag{37}$$

The reconstruction signal $Y_{j,k}^*$ can be expressed as

$$Y_{j,k}^* = \Psi s_k. \tag{38}$$

The residual sequence $\partial_{j,k} \in \mathbb{R}^{w}$ can be obtained as

$$\partial_{j,k} = Y_{j,k} - Y_{j,k}^*. \tag{39}$$

The improved GRA is used to calculate the correlation degree of the combination $\lambda_{ji} = (\partial_{j,k}, U_{j,k})$. When the measurement $Y_{j,k}$ is not attacked, $\partial_{j,k}$ is approximate with measurement noise. There is no correlation between $\partial_{j,k}$ and control variables. If $Y_{j,k}$ is attacked, the residual sequence $\partial_{j,k}$ is mixed with attack signal Δy and measurement noise e. There is correlation between the $\partial_{j,k}$ and the control variable. We calculate the correlation degree between the residual sequence $\partial_{j,k}$ and control variable sequences follow the corresponding combinations in the set Ω'. The calculation results $\xi^k_{\partial u}$ is as follow

$$\xi^k_{\partial u} = [\gamma_k(\lambda^1_{ji}), .., \gamma_k(\lambda^\mu_{ji})]. \tag{40}$$

If the calculation results in $\xi^k_{\partial u}$ are partially or completely similar with the corresponding element in ξ_{yu}, we think that the system is attacked by the measurement injection attack at sample time k.

C. THE DETECTION FOR CONTROL VARIABLE TAMPERING ATTACK

From the analysis in the section II, we have known that when the control variables are attacked, the measurements that are collected from the sensors will be masked by the cheating data y_c. Although the cheating data is numerically similar to the normal measurement, the correlation between the cheating data and the control variable is difficult to maintain the inherent correlation between system variables. We use the improved GRA to calculate the correlation degree between measurements and control variables at sample time k follow the combinations in the set Ω'. The calculation results can be expressed as

$$\xi^k_{y_c u} = [\gamma_k(\omega^1_{ji}), .., \gamma_k(\omega^\mu_{ji})] \tag{41}$$

If the calculation results $\xi^k_{y_c u}$ do not match the inherent correlation ξ_{yu}, we believe that the system is attacked by the control variable tampering attack at sampling time k.

D. THE DETECTION PROCEDURES OF THE PROPOSED ALGORITHM

As shown in Algorithm 1, we designed the detection procedures of the proposed algorithm according to the detection methods about the two FDI attacks. The input of the Algorithm 1 is a raw dataset collected from the ICS. The dataset contains normal data and attack data. The windowed resampling is used to resample the raw dataset. Then, the resampled dataset is divided into training set and test set. The improved GRA is used to select the correlation combinations in normal data of the training set. Next, the correlation degrees between variables in the entire training set are calculated based on the set of correlation combinations. The support vector machine (SVM) is adopted as the classifier to classify the calculation results. The normal data, measurement injection attack data, and control variable tampering attack data are labeled as l_0, l_1, and l_2. Then, the predictive model of SVM will be used to detect the attack data in test set.

Algorithm 1 The Proposed Detection Algorithm

```
Input: The raw dataset contains measurements and control variables.
Initialization: The size of the time window in windowed sampling.
Training:
  Use the window sampling to resample the raw dataset. The 80% of sample results are selected as the training set, and the rest is the test set.
  Select the correlation combinations in normal data of training set by using the improved GRA, and a set $\Omega'$ of the correlation combinations is obtained.
  The CS is used to calculate the residual sequences of measurement sequences that are select from the normal data and measurement injection data according to the set $\Omega'$.
  Through the improved GRA to calculate the correlation degree between the residual sequences and control variable sequences. Then, SVM is used to establish the hyperplane models $f_1(\xi_{\partial u})$ based on the calculation results.
  According to the set $\Omega'$, the correlation degrees in normal data and control variable tampering data are calculated by the improved GRA. Then, the calculation results are used to establish the hyperplane models $f_2(\xi_{y_c u})$ through the SVM.
Testing:
  Calculate the correlation degree $\xi_{y_c u}$ in test set through the improved GRA.
  if $f_2(\xi_{y_c u}) == l_2$ then
    Control variable tampering attack.
  else
    Calculate the correlation degree $\xi_{\partial u}$ in test set through the improved GRA.
    if $f_1(\xi_{\partial u}) == l_1$ then
      Measurement injection attack.
    else
      Normal.
    end if
  end if
```

V. EXPERIMENT

A. TESTBED AND DATA GENERATION

From the architecture of the ICS in Figure 1, we made a semi-physical simulation testbed for verifying the reliability of the proposed algorithm. The virtual plant of the testbed corresponds to a boiler-turbine unit [32] which is widely used in modern power stations. The boiler-turbine unit uses a single boiler to generate steam, then feeds the steam to a single turbine to generate electricity. The boiler-turbine unit, controllers, actuators and sensors are simulated by Raspberry 3B (with the operating system called Raspbian). The communication way employs the Modbus/TCP that is realized by the modbus_tk module in the Python library.

A PC is used to operate these Raspberry 3B and responsible for collecting data, which can be regarded as a simple HMI. In order to inject attacks conveniently and effectively in this testbed, a separate raspberry playing the role of an attacker is connected in this testbed. The communication way among Raspberry 3B employs the physical link, which makes attack simulation more realistic.

The Boiler-turbine unit is a 2×2 system [33]. The throttle valve position Tv and the boiler firing rate Br are control variables. The throttle pressure Pt and the megawatt output Mo are sensor measurements. As shown in Figure 3, the dynamic performances of the boiler-turbine unit are tested at the setpoint of $Mo = 328.86$, $Pt = 17.61$. The close-loop performances of the boiler-turbine unit are stabilization in this testbed.

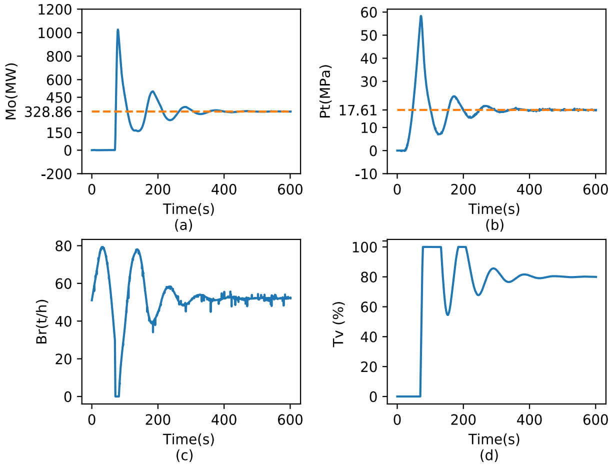


FIGURE 3. **Subfigures (a) and (b) are the dynamic responses of *Mo* and *Pt* respective which show that the outputs of plant follow the setpoints. Subfigures (c) and (d) are the change curves of the control variable *Br* and *Tv*.**

In order to verify the performance of the testbed for simulating attacks, the two FDI attacks are injected in the testbed through the attack model for 100 seconds and the attack results are shown in Figure 4 and 5.

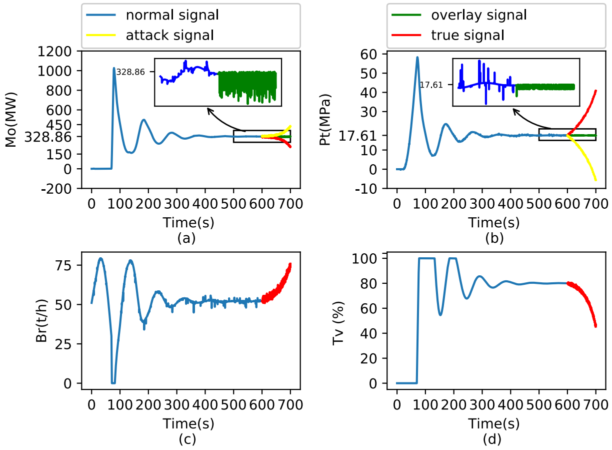


FIGURE 4. **The measurement injection attack is injected into the testbed. (a) The outputs Mo of plant. (b) The pressure Pt of pipeline. (c) The outputs Br of controller. (d) The outputs Tv of controller.**

As shown in Figure 4, the real signals Mo and Pt are masked by the overlay signal which are shown in subfigures (a) and (b). As shown in subfigures (c) and (d), the fuel flow of Br is fast growing and the opening of Tv reduces quickly due to that the overlay signal of Mo and Pt are below the setpoint.

As shown in Figure 5, the ever-rising attack signal of Br and the declining attack signal of Tv are sent to controller simultaneously which are shown in subfigures (a) and (b). The attack signals lead Pt increase rapidly and Mo decline persistently, as schematically represented in subfigures (c) and (d).

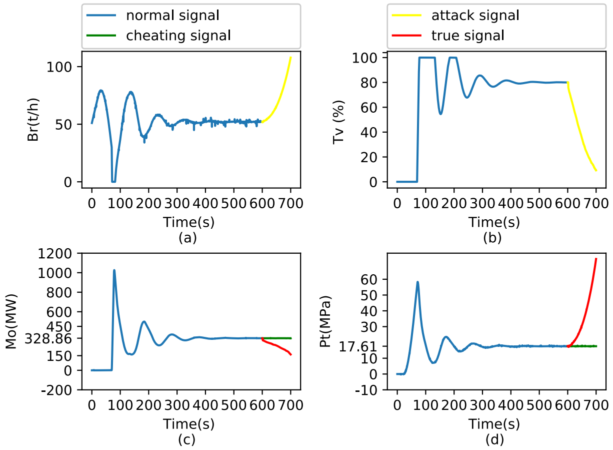


FIGURE 5. **The control variable tampering attack is injected into the testbed. (a) The outputs Br of controller. (b) The outputs Tv of controller. (c) The outputs Mv of plant. (d) The pressure Pt of pipeline.**

From the above analysis, it can be concluded that the testbed is reliable for simulating attacks. A dataset is created through the testbed. The dataset contain five components: Br, Tv, Mo, Pt, and sample time record. The measurements that are collected in measurement injection attack data are overlay signals. For control variable tampering attack, only the cheating signals can be observed.

B. THE CORRELATION ANALYSIS OF VARIABLES IN BOILER-TURBINE UNIT

From the prior knowledge of the Boiler-turbine unit, we can get the correlation between the master variable Pt, Mo and the dependent variable Br, Tv in qualitatively. There are positive correlation in combination ($Pt \rightarrow Tv$) and negative correlation in combinations ($Pt \rightarrow Br$), ($Mo \rightarrow Br$), and ($Mo \rightarrow Tv$).

In this work, the performance of the GRA and the improved GRA is tested in calculating the correlation degree between Boiler-turbine unit variables. First,The correlations are analyzed when the testbed is under measurement injection attack. Then, the CS is used to separate the noise from each measurement. The GRA and improved GRA are used to analyze the relevancy between the control variables and the residual signals. The blue and red dots represent normal and attack data, respectively. From the calculation results which are shown in Figure 6 and 7, it is obvious that GRA does not

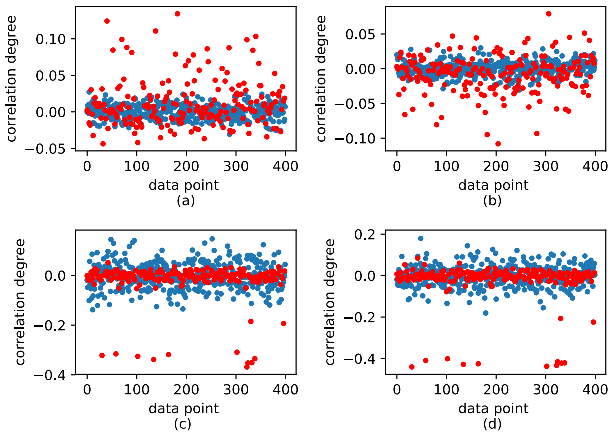


FIGURE 6. The GRA for calculating the correlation degree between the residual signals and control variables(blue dot: normal, red dot: attack). (a) The correlation between Tv and the residual of Pt. (b) The correlation between Br and the residual of Pt. (c) The correlation between Tv and the residual of Mo. (d) The correlation between Br and the residual of Mo.

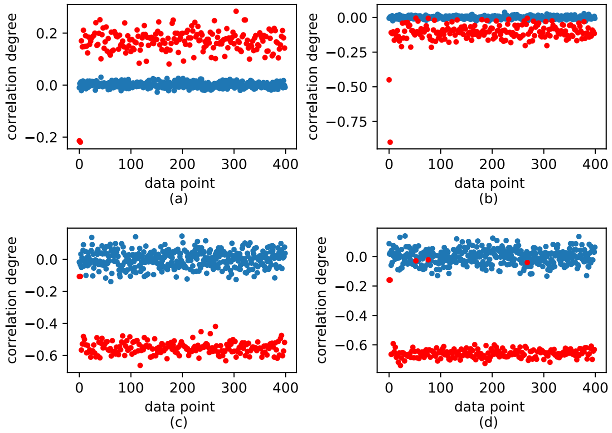


FIGURE 7. The improved GRA for calculating the correlation degree between the residual signals and control variables(blue dot: normal, red dot: attack). (a) The correlation between Tv and the residual of Pt. (b) The correlation between Br and the residual of Pt. (c) The correlation between Tv and the residual of Mo. (d) The correlation between Br and the residual of Mo.

separate attack data from normal data. However, the improved GRA can distinguish the attack data well.

Then, the relevancy between variables is calculated by the GRA and improved GRA when the testbed is attacked by the control variables tampering attack, which is shown in Figure 8 and 9. It can be derived that the improved GRA has better performance on dealing with the control variable tampering attack data than the GRA does.

Through the above quantitative about the correlation in boiler-turbine unit variables, we can derive that there are correlations in the combinations $(Pt \rightarrow Tv)$, $(Pt \rightarrow Br)$, $(Mo \rightarrow Br)$, and $(Mo \rightarrow Tv)$.

C. THE PERFORMANCE EVALUATION OF THE PROPOSED ALGORITHM

For the binary classification problem, the learner prediction results can be divided into true positive (TP), false positive (FP), true negative (TN), and false negative (FN) according to the known true label of the data in the test set [34]. For evaluating the performance of the detection algorithm, the precision P and recall R are defined as

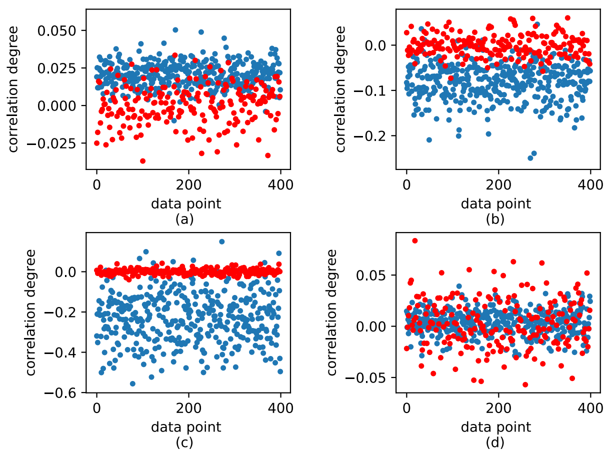


FIGURE 8. The GRA for calculating the correlation degree between variables when control variables are modified(blue dot: normal, red dot: attack). (a) The correlation between Pt and Tv. (b) The correlation between Pt and Br. (c) The correlation between Mo and Tv. (d) The correlation between Mo and Br.

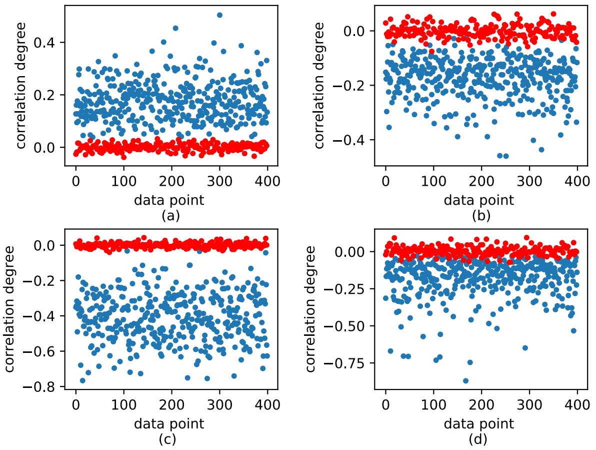


FIGURE 9. The improved GRA for calculating the correlation degree between variables when the control variables are modified(blue dot: normal, red dot: attack). (a) The correlation between Pt and Tv. (b) The correlation between Pt and Br. (c) The correlation between Mo and Tv. (d) The correlation between Mo and Br.

$$P = \frac{TP}{TP + FP}, \tag{42}$$

$$R = \frac{TP}{TP + FN}. \tag{43}$$

In general, when the precision is high, the recall is often low. For comprehensive consideration of precision and recall to evaluate the performance of the proposed algorithm, the $F1$ measure is calculated [34].

$$F1 = 2 \times \frac{P \times R}{P + R}. \tag{44}$$

It is an important step to select the suitable size of the time window to sample system variables in the proposed algorithm. Therefore, the influence of the length of the time window on the performance of the proposed algorithm is studied. As shown in Figure 10, we use different time window sizes to train the detection model through the training set and get the F1 score of the corresponding model by the test set.

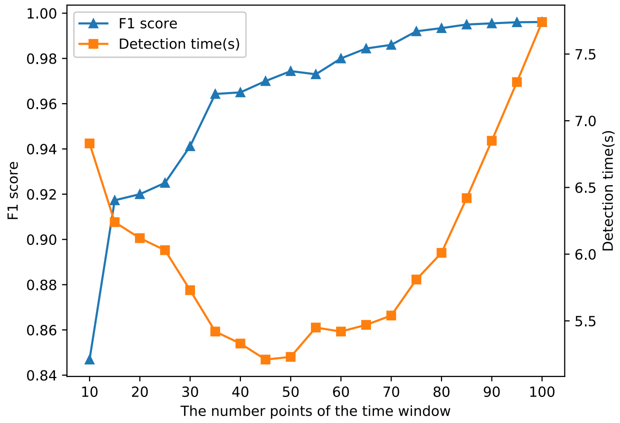


FIGURE 10. The performance of the proposed algorithm with the size of time window rising.

From the blue curve, it is observed that the size of the time window has a significant impact on the $F1$ score of the proposed algorithm. The performance of the proposed algorithm is getting better with the rising of the size of the time window. Besides, we calculate the average detection time of the detection model with different time windows for detecting attacks. The orange curve shows the average detection time with different detection models. In order to comprehensively consider the average detection time and $F1$ score, the time window that has 50 data points is selected as the optimal size of the time window.

D. THE COMPARISON OF ALGORITHM PERFORMANCE

The performance of the proposed algorithm is compared with logistic regression (LR), artificial neural network (ANN) and SVM. Furthermore, the improved GRA can be regard as a new kernel for SVM from the point of view of the kernel function. Hence, the SVM based on the linear kernel and Gaussian kernel are selected. The way for selecting the parameters of LR in [19] is chosen. The default parameters of ANN are chosen from [20] and [35]. The optimal parameters of the SVM based on the linear kernel and Gaussian kernel in [19] and [21] are used in this work. The time window that contains 50 points is used for the comparison in the proposed algorithm. We use each algorithm to detect the measurement injection attack and control variable tampering attack through a dataset that is collected from the proposed testbed. In addition, the F1 score of each algorithm is studied with the change of false data ratio. As the false data ratio changes, the performance of each algorithm is shown in Table 1. It can be seen from the data in Table 1 that the F1 score of the proposed algorithm is always higher than other algorithms, and changes between 0.96 and 0.98 with the change of false data ratio, while the F1 score of other methods are in a downward trend. In order to observe the trend of the F1 score of each algorithm comprehensively, the data in Table 1 is shown in Figure 11 and Figure 12. By observing the changed trend in the F1 score curves of each algorithm, it can be observed that the proposed algorithm has better performance than other algorithms.

TABLE 1. The F1 score of each algorithm based on the proposed dataset.

F1 score		False data ratio					
		0.05	0.10	0.15	0.20	0.25	0.30
A1 attack	Proposed algorithm	0.9855	0.9845	0.9770	0.9725	0.9735	0.9750
	ANN	0.9815	0.9550	0.9469	0.9365	0.9335	0.9149
	LR	0.9495	0.8775	0.8670	0.8505	0.8455	0.8120
	Linear-SVM	0.9544	0.9231	0.8544	0.8526	0.8088	0.7776
	Gaussian-SVM	0.9604	0.9301	0.8805	0.8728	0.8385	0.8075
A2 attack	Proposed algorithm	0.9702	0.9700	0.9629	0.9609	0.9606	0.9604
	ANN	0.9703	0.9570	0.9402	0.9360	0.9286	0.9174
	LR	0.9500	0.9000	0.8499	0.7995	0.7595	0.7160
	Linear-SVM	0.9503	0.9100	0.8553	0.8000	0.7502	0.6999
	Gaussian-SVM	0.9633	0.9202	0.8604	0.8322	0.7633	0.7193

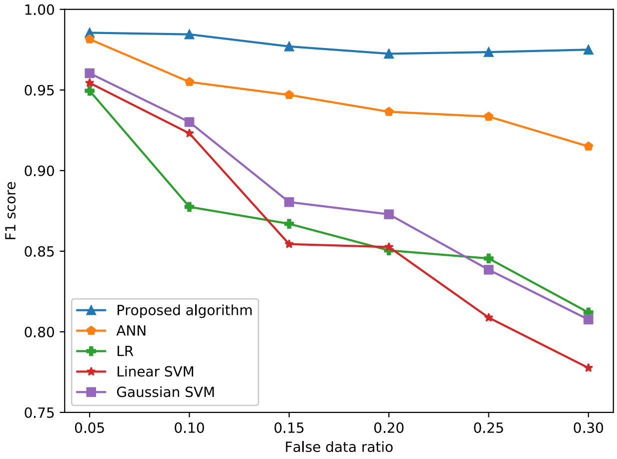


FIGURE 11. The F1 score of each algorithm with the change of false data ratio in the measurement injection attack based on the proposed dataset.

To further illustrate the reliability of the proposed algorithm, the performance of each algorithm is also studied

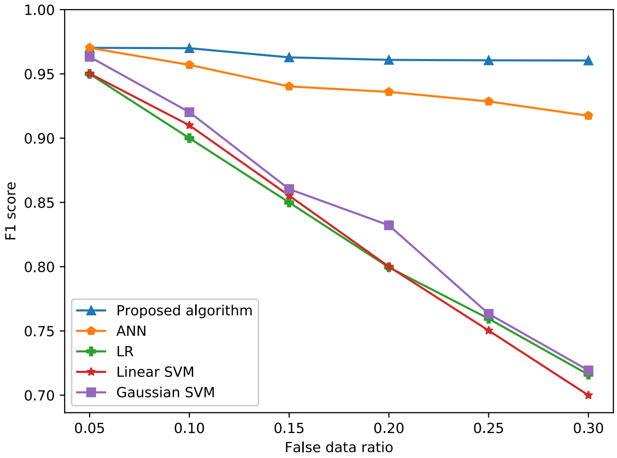


FIGURE 12. The F1 score of each algorithm with the change of false data ratio in the control variable tampering attack based on the proposed dataset.

TABLE 2. The F1 score of each algorithm based on the new gas pipeline dataset.

F1 score		False data ratio					
		0.05	0.10	0.15	0.20	0.25	0.30
A1 attack	Proposed algorithm	0.9755	0.9663	0.9670	0.9633	0.9634	0.9638
	ANN	0.9750	0.9500	0.9059	0.8654	0.8426	0.8310
	LR	0.9560	0.9214	0.8370	0.8032	0.7600	0.7199
	Linear-SVM	0.9693	0.9311	0.8410	0.8290	0.7892	0.7495
	Gaussian-SVM	0.9720	0.9430	0.8569	0.8370	0.7991	0.7678
A2 attack	Proposed algorithm	0.9753	0.9690	0.9646	0.9650	0.9626	0.9664
	ANN	0.9780	0.9541	0.9379	0.9100	0.8747	0.8265
	LR	0.9678	0.9412	0.8679	0.8201	0.7882	0.7399
	Linear-SVM	0.9604	0.9501	0.9124	0.8338	0.8066	0.7796
	Gaussian-SVM	0.9654	0.9620	0.9354	0.8578	0.8288	0.7875

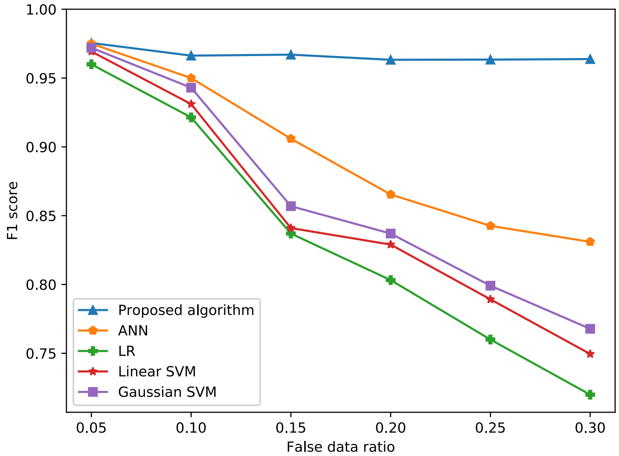


FIGURE 13. The F1 score of each algorithm with the change of false data ratio in the measurement injection attack based on the new gas pipeline dataset.

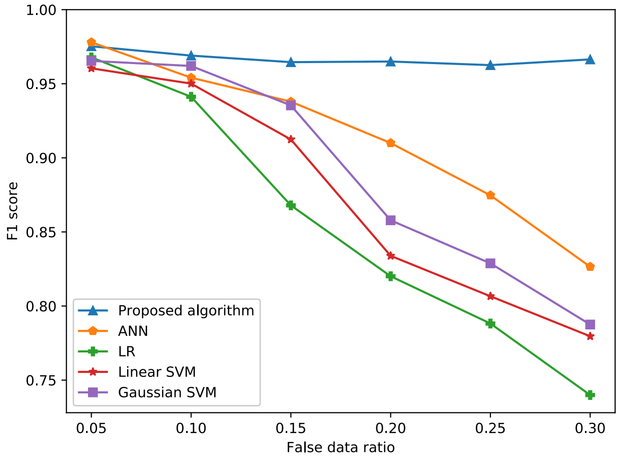


FIGURE 14. The F1 score of each algorithm with the change of false data ratio in the command injection attack based on the new gas pipeline dataset.

based on the new gas pipeline dataset that is collected by the distributed analytic and security institute in Mississippi State University [36]. The measurement injection attack (A1) and command injection attack (A2) contained in the new gas pipeline dataset are studied in this work. The performance of each algorithm is shown in Table 2, Figure 13, and Figure 14. It can be seen that the F1 scores of ANN, LR and SVM are on a downward trend with the change of false data ratio. However, the F1 score of the proposed algorithm is always higher than other algorithms, and changes between 0.96 and 0.97 with the change of false data ratio. It also can be observed that the proposed algorithm has better performance than other algorithms.

VI. CONCLUSION

This paper introduces a novel data analysis algorithm to detect the FDI attacks in ICS. The proposed algorithm distinguishes the attack behavior in ICS by analyzing the correlation between variables with the improved GRA. A dataset that contains normal data and attack data is created from a reliable semi-physical simulation testbed. Based on the dataset, we firstly verify that improved GRA has better performance than GRA does. Then, the impact of the size of the time window sampling on the performance of the proposed algorithm is analyzed. Besides, we compare the F1 score of the proposed algorithm and the state of the art based on the data analysis for cyber-attack detection through the proposed dataset and new gas pipeline dataset. The experiment results show that the proposed algorithm has higher accuracy for detecting the two FDI attacks than other algorithms do. In the future, the proposed method will be applied to process more FDI attacks. Additionally, we hope to employ the proposed algorithm in online detection.

REFERENCES

[1] L.-X. Yang, P. Li, X. Yang, and Y. Y. Tang, "Security evaluation of the cyber networks under advanced persistent threats," *IEEE Access*, vol. 5, pp. 20111–20123, 2017.

[2] J. P. Farwell and R. Rohozinski, "Stuxnet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, 2011.

[3] F. Khnorrami, P. Krishnamurthy, and R. Karri, "Cybersecurity for control systems: A process-aware perspective," *IEEE Des. Test*, vol. 33, no. 5, pp. 75–83, May 2016.

[4] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, Jul. 2017.

[5] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, Chicago, IL, USA, Nov. 2009, pp. 21–32.

[6] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.

[7] M. Esmalifalak, H. Nguyen, R. Zheng, and Z. Han, "Stealth false data injection using independent component analysis in smart grid," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Oct. 2011, pp. 244–248.

[8] X. Liu, Z. Li, X. Liu, and Z. Li, "Masking transmission line outages via false data injection attacks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1592–1602, Jul. 2016.

[9] A. Abur and A. G. Exposito, "Bad data identification when using ampere measurements," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 831–836, May 1997.

[10] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S. Sastry, "Cyber security analysis of state estimators in electric power systems," in *Proc. 49th ACM Conf. Decis. Control*, Dec. 2010, pp. 5991–5998.

[11] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 4, pp. 1396–1407, Jul. 2014.

[12] Y. Mo and B. Sinopoli, "On the performance degradation of cyber-physical systems under stealthy integrity attacks," *IEEE Trans. Autom. Control*, vol. 61, no. 9, pp. 2618–2624, Sep. 2016.

[13] J. Zhao, G. Zhang, M. La Scala, Z. Y. Dong, C. Chen, and J. Wang, "Short-term state forecasting-aided method for detection of smart grid general false data injection attacks," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1580–1590, Jul. 2017.

[14] D. Shi, R. J. Elliott, and T. Chen, "On finite-state stochastic modeling and secure estimation of cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 65–68, Jan. 2017.

[15] L. Hu, Z. Wang, Q.-L. Han, and X. Liu, "State estimation under false data injection attacks: Security analysis and system protection," *Automatica*, vol. 87, pp. 176–183, Jan. 2018.

[16] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[17] J. Valenzuela, J. Wang, and N. Bissinger, "Real-time intrusion detection in power system operations," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1052–1062, May 2013.

[18] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, Nov. 2015.

[19] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773–1786, Aug. 2016.

[20] Y. Wang, M. M. Amin, J. Fu, and H. B. Moussa, "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids," *IEEE Access*, vol. 5, pp. 26022–26033, 2017.

[21] S. A. Foroutan and F. R. Salmasi, "Detection of false data injection attacks against state estimation in smart grids based on a mixture Gaussian distribution learning method," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 2, no. 4, pp. 161–171, Dec. 2017.

[22] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[23] J. L. Deng, "Introduction to grey system theory," *J. Grey Syst.*, vol. 1, no. 1, pp. 1–24, Oct. 1989.

[24] M. Han, R. Q. Zhang, and M. L. Xu, "A variable selection algorithm based on improved grey relational analysis," *Control Decis.*, vol. 32, no. 9, pp. 1647–1652, Sep. 2017.

[25] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*. Berlin, Germany: Springer, 1999.

[26] G. Bernieri, E. E. Miciolino, F. Pascucci, and R. Setola, "Monitoring system reaction in cyber-physical testbed under cyber-attacks," *Comput. Elect. Eng.*, vol. 59, pp. 86–98, Apr. 2017.

[27] A. A. Cárdenas, S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang, and S. Sastry, "Attacks against process control systems: Risk assessment, detection, and response," in *Proc. 6th ACM Symp. Inf., Comput. Commun. Secur.*, Mar. 2011, pp. 355–366.

[28] H. Kuang, M. A. Bashar, K. W. Hipel, and D. M. Kilgour, "Greybased preference in a graph model for conflict resolution with multiple decision makers," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 9, pp. 1254–1267, Sep. 2015.

[29] M. Han, R. Zhang, T. Qiu, M. Xu, and W. Ren, "Multivariate chaotic time series prediction based on improved grey relational analysis," *IEEE Trans. Syst. Man. Cybern. Syst*, to be published, doi: 10.1109/TSMC.2017.2758579.

[30] Y. Dang, S. Liu, B. Liu, and C. Mi, "Improvement on degree of grey slope incidence," *Eng. Sci.*, vol. 6, no. 3, pp. 41–44, Mar. 2004.

[31] E. J. Candés and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[32] L. Tian, D. L. Zeng, J. Z. Liu, and Z. Zhao, "Simplified nonlinear dynamic model for a 330 MW unit," *Proc. CSEE*, vol. 24, no. 8, pp. 180–184, Oct. 2004.

[33] W. Tan, F. Fang, L. Tian, C. Fu, and J. Liu, "Linear control of a boiler–turbine unit: Analysis and design," *ISA Trans.*, vol. 47, no. 2, pp. 189–197, Apr. 2008.

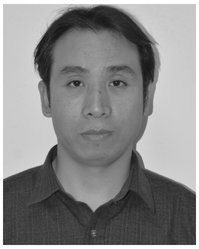
[34] N. Chinchor and B. Sundheim, "MUC-4 evaluation metrics," in *Proc. 5th Conf. Message Understand.*, Aug. 1993, pp. 69–78.

[35] K. Khanna, B. K. Panigrahi, and A. Joshi, "AI-based approach to identify compromised meters in data integrity attacks on smart grid," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 5, pp. 1052–1066, Jan. 2018.

[36] T. H. Morris *et al.* (2015). *Industrial Control System Simulation and Data Logging for Intrusion Detection System Research*. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

ZHENGDAO ZHANG received the Ph.D. degree in control theory and engineering from the Nanjing University of Aeronautics and Astronautics, China, in 2006. From 2010 to 2011, he was a Visiting Scholar with Pennsylvania State University. He is currently an Associate Professor with the Engineering Research Center of Internet of Things Applied Technology, Ministry of Education, Jiangnan University. He published more than 30 papers in journals and conferences. His research interests include industrial control system, security of industrial Internet of Things, and fault prognosis and applications.

YUNFEI WANG received the B.E. degree in automation from the University of Science and Technology at Liaoning, Liaoning, in 2016. He is currently pursuing the M.E. degree in control science and engineering with the Engineering Research Center of Internet of Things Applied Technology, School of IoT Engineering, Jiangnan University, Wuxi, China. His research interests include security of industrial control system.



LINBO XIE received the Ph.D. degree in control theory and control engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2004. He is currently with the School of Internet of Things Engineering, Jiangnan University, Wuxi, China, where he is also a Professor. His interests include analysis and synthesis of networked control systems and advanced process control.