Cyber-Physical Integrated Intrusion Detection Scheme in SCADA System of Process Manufacturing Industry

JUNLEI QIAN1,2, XUEQIANG DU2, BO CHEN3, BIN QU3, KAI ZENG3, AND JIANPENG LIU4

1Institute of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China
2Automation Research and Design Institute of Metallurgical Industry, Beijing 100071, China
3College of Electrical Engineering, North China University of Science and Technology, Tangshan 063000, China
4Tang Steel International Engineering Technology Corporation, Tangshan 063000, China

Corresponding author: Junlei Qian (catherineqjl@163.com)

This work was supported in part by the Natural Science Foundation of Hebei Province under Grant F2019209443.

ABSTRACT Most Intrusion Detection System (IDS) used in Supervisory Control and Data Acquisition (SCADA) systems now are focused on the cyber field but ignored the process states in physical field of the plants. Attacks aiming at the protocol traffics can be detected but attacks aiming at the processes are difficult to be detected such as the Man In The Middle (MITM) attacks and the Replay attacks. We propose a scheme in both cyber way and physical way to detect the above attacks. Validation of process states is used to detect malicious behaviors to prevent the physical components from being damaged which can be caused by MITM, Replay, and Zero-day attacks. Nonparallel hyperplane based fuzzy classifier is presented to realize the classification of branching shaped data sets which are difficult to be classified by two parallel hyperplane of Support Vector Machine (SVM) to detect attacks caused by DoS (SYN flood) and other attacks in cyber field. Modbus/TCP traffic data are used to test the algorithm and simulation process states are used to test the validation part and the performance of this hybrid scheme is excellent.

INDEX TERMS Cyber attack, SVM, IDS, ICS, process state, nonparallel hyperplane.

I. INTRODUCTION

There are different aspects between Information Technology System (ITS) and Industrial Control System (ICS) networks: the performance requirements, risk management requirements, and communication etc. [1]. High delay may be acceptable in ITS but the response in ICS is time-critical. Data confidentiality and integrity are the most important and fault tolerance is less important in ITS, while human safety is the most important, followed by the protection of the process, and fault tolerance is essential in ICS. Standard communication protocols are used in ITS, but many proprietary communication protocols without timestamp, ID certification, and encryption supported by different vendors are used in ICS. The above mentioned delay, fault and information reveal caused by protocols can be caused by Denial of Service (DoS), Man In The Middle (MITM), Replay and Zero-day attacks to ICS.

Attacks applied to ITS may cause a congest or an information reveal, but attacks apply to the ICS networks can cause both information reveal and huge damage to physical infrastructure. So cybersecurity is currently a vital part for Supervisory Control and Data Acquisition (SCADA) systems which are widely used in ICS to guarantee the security of a controlled process. The security of a SCADA system including the protection of the physical infrastructures and the controlled processes, communication protocols, asset management, and so on [2], which are the key components of intelligent manufacture and cannot be handled in the same way as ITS counterparts. These key components usually including computers, servers, network equipments, Remote Terminal Units (RTU), Programmable Logical Controllers (PLC), Distributed Control Systems (DCS) and the supporting softwares like Human Machine Interface (HMI).

From a hacker’s perspective, attacking only one equipment such as a machine is easy to be detected by the fault diagnosis system, and the caused damage is not satisfied. It’s more destructive if they attack a manufacturing process by tampering controlled parameters and let the process ruin the
key infrastructure. An investigation showed that about 67% vulnerabilities in ICS are distributed in the above components in 2018, and most of the vulnerabilities are distributed in the mainstream products like PLCs, DCSs and the related softwares of the most famous manufacturers like Schneider and Siemens [3]. All of the vulnerabilities of the above components can be used to attack the controlled processes in plants.

A SCADA system, as shown in Fig.1 includes three levels: the control level, the supervisory level and the management level. Large amount of data are collected from sensors to the PLCs in control level, and the data will be sent to the operator’s station in supervisory level. The data will be compared with the set point in the PLCs, then the PLCs make decisions and the instructions will be sent to the actuators in the control level. Both the data from sensors and the instructions to the actuators can be monitored at the supervisory level.

The three points we marked by exclamation marks in Fig.1 are where the vulnerabilities locate: the PLCs, the HMIs on the operators’ computer, and the switches, and the types and targets of attacks are listed in Table1. Attacks on the PLCs usually tamper the data from sensors which may cause a false decision by the controllers. Attacks on the HMIs may tamper the set point of a process which can cause a false decision too. The above two attacks can be carried out on the switches, and so do DoS which can cause the jam of the networks.

**TABLE 1.** General attacks considered in this scheme.

| Attack Name | Targets |
|-------------|---------|
| 1 | MITM | HMI, PLC |
| 2 | RE | RGF |
| 3 | DoS | HMI, PLC, Switch |
| 4 | Zero-day | HMI, PLC, Switch, Computer |

Many Known vulnerabilities have been confirmed by the PLC suppliers that will threat the control systems [4], [5]. But up to now, there’s little security measure has been taken by neither the suppliers nor the integrators to the three layered SCADA networks beside isolating the three layered networks away from the internet. Many hackers’ attention is focused on devastating the manufacture or even the plants. By using the vulnerabilities of the protocols in SCADA systems the above attacks can tamper the data between the sensors and the controllers to break the control logic of the process which can cause the process out of control and result in physical damage.

Most study is focused on detecting the abnormal data from traffic data like in the ITS. Some of the attacks such as Buffer Overflow are easy to detect by this method and the vulnerability can be fixed by programme, but other attacks like MITM, Replay and Zero-day attacks are difficult to detect because they have the ability to elude the current anomaly-based detection methods [6]. So novel methods should be used in ICS to detect the attacks aiming at SCADA systems. From the perspective of cyber security, methods used in ITS can be referenced to detect Buffer Overflow and DoS, and from the perspective of physical security in plants, we need to do more research on the characteristics of networks and components in the SCADA systems.

Here, we provide a hybrid method in SCADA systems to detect abnormal in traffic data and the broken constrain in the control logic to guarantee normal production processes. The rest of this paper is organized as follows. The typical attacks often used in manufacturing processes are listed and the related works are given in section II. The validation of the realtime process states in control level is presented in section III. The Nonparallel hyperplanes based fuzzy classifier is presented in section IV. And in section V, the result of application is presented. Section VI is the conclusion.

II. PROBLEM FORMULATION AND PRELIMINARIES

The recent attacks on ICS have demolished the thoughtful security of SCADA systems [7]–[9]. The attacks on critical infrastructures make the securing and protecting of ICS systems extremely important.

A. ATTACKS AIMING AT SCADA SYSTEMS IN PROCESS MANUFACTURING INDUSTRY

A typical ICS controls and monitors the functions of key components (e.g. electric motor, pumps, temperature or level of water, pressure of pipelines or tanks) in SCADA systems. The major functions of a SCADA system are data acquisition and control. Sensors measure the parameters of the controlled object, then transmit the values of these parameters to the controller via protocols such as Modbus/TCP, Profinibus or Profinet etc.. After receiving the data the controllers compute and compare the values with the set point, then give a decision which will be sent to the actuators like valves or electric motors to regulate the opening of valves or the speed of the turbine to control the pressure or the flow. The values from sensors and the instructions to the actuators are related according to the control logic, so data in the Modbus/TCP protocol are related too.
Severity of damage from the above mentioned three types of attacks to ICS is significantly different compared with the same attacks to ITS:

The MITM attack and Replay attack are difficult to detect because these two attacks are based on Address Resolution Protocol (ARP) cache spoofing which can disguise themselves as the RTUs by sending fake ARP messages to the PLCs to associate the attacker’s MAC address with the IP address of a real RTU. Then it can communicate with the PLC normally like the correct RTU. They can tamper the data from sensors or the instructions from the PLCs to to break the constraint relationship between the actuators and the controlled parameters which can make the system out of control.

In a Replay attack, hackers even don’t need to tamper any data, the only thing they need to do is to send the data package eavesdropped from the sensors at another time. The package can not be detected to be false because there’s no timestamp in the protocol used in SCADA systems and the data are in the right form. The IP address is disguised to be a correct RTU, so everything is right, and the analysis on traffic data can not detect any abnormal. But because the time changed, and the realtime values of the sensors are not the ones in the package, so the wrong values will cause the controllers to make wrong decisions and further more the wrong decisions may cause a explode to damage the crucial physical infrastructure.

A Zero-day vulnerability is often used by hackers to attack an ICS. The detail of Zero-day attack can only be detected after the attack is confirmed. That is to say, this attack can only be detected by the abnormal results of manufacture, so there’s no other existing methods is effective detect it.

It’s difficult to detect the above attacks by only analyzing the traffic data in ICS. But they have the same characteristic: they all aim at breaking the control logic of the manufacturing process to cause physical damages. With the widely use of fault self-diagnose system and Product Lifecycle Management (PLM), seldom wrong logic is caused by the equipment failure, so if we can detect the abnormal of the control logic we can detect the above attacks.

In this proposed scheme, two ways are used to detect different attacks in ICS. In control level we establish state database for the controlled process by ranked key parameters, then compare the real time states with the history states by calculating the critical scores. By comparing the critical scores with the threshold we can conclude whether the running state is normal or not. DoS (SYN flood ) attacks are the common but difficult to detect attacks in supervisory level because of the branching feature of the traffic data. We use Nonparallel Hyperplanes based Fuzzy Classifier (NHFC) to classify the traffic data in supervisory level which has better performance than the conventional methods such as least squares (LS) based methods.

The manufacturing processes in steel plants are typical complex processes with the networks in Fig.1 as the control system. The above discussed three attacks often aim at steel plants as the object. So we design this scheme with steel plant as the background.

**B. RELATED WORKS**

Some methods used in ICS for intrusion detection are the same as used in ITS. But in ICS, we need more effective methods to detect attacks which are difficult to be detected by the existing methods. Special protocols used in SCADA systems such as Modbus, Modbus/TCP, Profinet, Profinet etc. lack of necessary protections like ID certification, timestamp or encryption. Most intrusions on SCADA systems like MITM attack and Replay attack are rooted from the vulnerabilities of these protocols. So most communications in SCADA systems are running without enough security protection [10].

| Time          | Protocol data | Traffic data | Label |
|---------------|---------------|--------------|-------|
| 2019-5-17 19:28:19:103250 | 5 | 90177 | 17 | 5 | 7 | 3 | 5 | normal |
| 2019-5-17 19:28:19:103250 | 5 | 90177 | 17 | 5 | 7 | 3 | 5 | normal |
| 2019-5-17 19:28:19:103250 | 5 | 90177 | 17 | 5 | 7 | 3 | 5 | normal |
| 2019-5-17 19:28:19:103250 | 5 | 90177 | 17 | 5 | 7 | 3 | 5 | normal |

Some kinds of intrusions can be identified easily by analyzing the traffic data because they showed abnormal in the function code of the protocol, overlength of data or a wrong IP/MAC address as shown in Table 2. But DoS attacks do not rely on particular network protocols or system weaknesses. It’s common for attackers to use legal but wrong instructions to carry out flooding-based DoS attacks in ICS. They simply exploit the huge resource between the different users which attempt to block legitimate user’s system access by reducing system availability [11]. The magnitude of the traffic is enough to jam, so the controller or the server will deny to work. A. Sperotto analyzed in his paper [12] that method based on traffic provided good performance in detecting the DoS attack.

X.Yang used the generalized entropy metric and the information distance metric to detect the low-rate DDoS attacks [13]. Hidden semi-Markov Model (HSMM) is proposed to get high classification accuracy and a mechanism to construct browsing behavior is also proposed in [14]. Ganapathy combined the intelligent agent weighed distance method and the improved intelligent agent support vector machine (SVM) based multiclassifier intrusion detection method to detect the outlier in [15]. W.Shang combined fuzzy C-means clustering (FCM) with SVM to calculate the distance between the cluster center and the industrial control network communication data in [16]. M. Wan used improved one
class support vector machine (OCSVM) to detect intrusion of SCADA systems [17]. F. Zhang constructed multiple-layer cyber-attack detection system to defend different attacks in [18]. H. Lin, and D. Hadzisomnovic use process state to detect malicious behaviors in SCADA systems [19], [20]. In [21]–[24], H. Ghaeini use state-aware anomaly detection in ICS which including hierarchical intrusion detection system (HAMIDS), Cumulative SUM (CUSUM), and Physics-based attestation (PAtt) methods.

Some of the above methods are mature in ITS, then be transplanted to ICS networks. The order of security priority in ICS networks is availability, integrity, and confidentiality, whereas this order in the ITS is confidentiality, integrity, and availability [10]. Availability is the most important factor in the security of ICS, so the transplanted methods can not solve the problem completely. Many attacks to ICS are aiming at the correlation of the parameters and the actuators in the processes. They spend long time to investigate the correlation in the process, and attempt to break the control logic to destroy the key infrastructure. In 2014 a steel plant in German was attacked by Advanced Persistent Threat (APT). The attacker used a spear phishing email to gain access to the corporate network and then sneaked into the plant network. They showed knowledge in ICS and was able to cause multiple components of the system to fail. This specifically impacted critical process components to become unregulated, which resulted in massive physical damage [7]. So only to improve the algorithm of the IDS can not meeting the needs in ICS networks. Algorithms used in [16] and [17] are data-driven to classify the function code in the Modbus protocols data sets. But attacks such as MITM, Replay, Zero-day and DoS are not sensitive to this method. Auto-Associative Kernel Regression (AAKR) is used in [18] and [25] to calculate the distance of the parameter vectors of process data to judge whether the data are tampered or not, but AAKR does not take the state sequence of the digital variables into account. PAtt method only calculate the state of digital parameters in [22]. And CUSUM need to calculate the residual of the process and the computation load depends on the complexity of the controlled process.

C. THE CONTRIBUTIONS OF THIS PAPER

Based on the research in ICS and process control, the contributions of this paper are summarized as follows:

1) Process state database is built in control level of ICS which is used to validate the realtime running states of the controlled process. The validation system is independent from the control system after data acquisition, so it will not occupy control system resources.

2) In control level, ‘Data’ segments of Modbus protocol are extracted to compose the realtime states. Whitelist based validation is used for digital variables, and correlation score based validation is used for analog variables to make sure whether the realtime states are normal or not. The validation can detect malicious behavior of the process to prevent the physical infrastructure from being damaged.

3) In supervisory level, Nonparallel hyperplanes based fuzzy classifier algorithm is proposed which can guarantee good interpretability of fuzzy classifier and solves the problem of classification of data sets with branching feature caused by DoS (SYN flood) attacks.

III. VALIDATION OF THE PROCESS STATES IN CONTROL LEVEL

There are hundreds of entities in a SCADA system, so hundreds of protocol data are sent simultaneously. The correlation coefficients between the realtime state and the historic state can be used to detect malicious nodes [26]. Spearman’s correlation coefficient is used in [2] to calculate the correlation coefficient of the vector of entities. In practice, the Modbus protocol data including IP/MAC address, function code and detecting data from sensors can be summarized as:

\[
\{\text{SourceIP}|\text{DestinationIP}|\text{Port}|\text{Functioncode}|\text{Data}\}
\]

We can extract the digital inputs DI, digital outputs DO, analog inputs AI and analog outputs AO from the ‘Data’ segment of the protocol. In order to increase the computing efficiency, we classify these parameters into two kinds of states: states for digital variables and states for analog variables.

For digital variables, we use the method of whitelists to verify the validity of the realtime states.

For analog variables, we calculate the correlation score between the realtime state vectors and the historic vectors.

A. VALIDATION OF THE PROCESS STATES FOR DIGITAL VALUES

When digital variables are the main variables in a process, the state database should enumerate all of the normal states of the process.

For example, three Hot Blast Stoves (HBSs) are used to supply hot blast to Blast Furnace (BF) in turn as in Fig.2. There are five “on-off” valves as digital variables in each HBS to control the air, gas, hot blast, cold blast and exhaust gas respectively. These valves switch states at a certain sequence, and other valves not tagged in the figure are bypass valves which are used for transitional adjustment, not in
Let \( r \) represent both real-time and historic data of \( m \)s represented as follows:

\[
\text{the real-time key parameter sequence. We construct matrix the number of sets queried from database corresponding to s as p of m process parameters of a process in a SCADA system. analog variables, assume FOR ANALOG VALUES}
\]

Any state can not be found in Table 3 is abnormal such as:

\[
V_{11} = 1 & V_{12} = 1 & V_{13} = 1 & V_{14} = 1 & V_{15} = 1 \rightarrow \text{Abnormal}
\]

### B. VALIDATION OF THE PROCESS STATES FOR ANALOG VALUES

When the main parameters of the controlled process are analog variables, assume \( P = p_1, p_2, p_3, \ldots, p_m \) be a set of \( m \) process parameters of a process in a SCADA system. Each \( p_i, i \in (1, m) \) is a vector of \( m \) related sensor values as \( p_{1,1}, p_{1,2}, p_{1,3}, \ldots, p_{1,m} \) is the real-time vector, while \( n \) is the number of sets queried from database corresponding to the real-time key parameter sequence. We construct matrix \( P \) represent both real-time and historic data of \( m \) sensors in \( P \):

\[
P_{n,m} = \begin{bmatrix}
p_{1,1} & p_{1,2} & \cdots & p_{1,m} \\
p_{2,1} & p_{2,2} & \cdots & p_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
p_{n,1} & p_{n,2} & \cdots & p_{n,m}
\end{bmatrix}
\]

Let \( S = s_1, s_2, \ldots, s_n \) be \( n \) states, in which each state \( s_i \) is represented as follows:

\[
s_1 = \begin{bmatrix}
p_{1,1} \\
p_{1,2} \\
\vdots \\
p_{1,m}
\end{bmatrix} \quad s_2 = \begin{bmatrix}
p_{2,1} \\
p_{2,2} \\
\vdots \\
p_{2,m}
\end{bmatrix} \quad \cdots s_n = \begin{bmatrix}
p_{n,1} \\
p_{n,2} \\
\vdots \\
p_{n,m}
\end{bmatrix}
\]

Calculate the Euclidian distance between the two states \( s_1 \) and \( s_2 \):

\[
d(s_1, s_2) = \sqrt{\sum_{i=1}^{m} (s_{1,i} - s_{2,i})^2}
\]

The correlation score for state \( s_i \) is defined as follows:

\[
\chi(s_i, S, k) = \frac{1}{k} \sum_{j=1}^{k} d(s_j, NN_j(s_i, S/s_j, k))
\]

where \( NN_j(s_i, S/s_j, k) \) denotes the \( j \)th nearest vector from the database of \( s_i \) in \( S \), and \( k \) is the number of queried \( s_i \), and \( 2 \leq k \leq |S| \) [26].

### TABLE 3. The normal states of one HBS.

| V₁₁ | V₁₂ | V₁₃ | V₁₄ | V₁₅ |
|-----|-----|-----|-----|-----|
| 1   | 0   | 0   | 0   | 0   |
| 0   | 1   | 0   | 0   | 0   |
| 0   | 0   | 1   | 0   | 0   |
| 0   | 0   | 0   | 1   | 0   |
| 0   | 0   | 0   | 0   | 1   |

Calculate the mean of the \( n \) critical scores, where \( c(i) \) is the critical score in critical states corresponding to the key parameter, in our example is \( F \). Then we get \( \mu \):

\[
\mu = \frac{1}{n} \sum_{i=1}^{n} c(i)
\]

Calculate the standard deviation of \( c(i) \):

\[
\sigma = \frac{1}{n} \sqrt{\sum_{i=1}^{n} (c(i) - \mu)^2}
\]

Then we get the threshold of the correlation score as:

\[
\gamma = \mu + \beta \sigma
\]

where \( \beta \) is the tuning parameter that trade-off between the detection rate and the false positive rate. \( \beta \) can be calculated by calculating the score of the allowed maximum deviation of the controlled parameter, then divided by \( \sigma \). From the threshold \( \gamma \), the state can be classified into two types: normal or abnormal. If the result is abnormal, the system will send an alert. The complete steps of this phase is in Algorithm 1.

### Algorithm 1 Validate Algorithm for Analog Values

1: Extract the ‘data’ segment from the protocol;
2: Create a ranked sequence of real-time state by the data in step 1;
3: Suppose the last state is normal, calculate the Euclidian distance \( d \) between the real-time state and the last state by (3), and compare \( d \) with the threshold \( \gamma \). If the distance is less than the threshold the process is in normal state.
4: If the distance in step 3 is not less than the threshold, create matrix \( P \) by (1) which including both the real-time and historic analog values in new state;
5: Calculate \( d(s_1, s_2) \) and \( \chi \) by equation (3), (4);
6: Calculate the threshold of the correlation score \( \gamma \) by (5), (6) and (7);
7: Compare \( \chi \) with \( \gamma \) and output the result like (8).

The control process of BF top pressure is the example as shown in Fig. 3.

In Fig. 3, the control system of BF top pressure in steel plant includes two parts: the Pressure Regulators Group (PRG) and Top Gas Recovery Turbine Unit (TRT). Ignore the transitional period of switch, either PRG or TRT is in use the BF top pressure should be in steady state. \( V_1 \) to \( V_4 \) and the turbine are used to control the BF top pressure. \( V_5 \) and \( V_6 \) are used to switch these two parts.

When PRG is in use, to maintain the value of \( P \) in steady state, the opening of \( V_1 \) to \( V_4 \) are related with the flow \( F \). The dash line between the \( P \) sensor and the valves means the control loop. When TRT is in use, the angle of the Stationary Blade (SB) and the rotate speed of the turbine are related with the total output gas flow \( F \). Table 4 shows the values after normalization of the analog variables in different status.
When PRG is working, $P$, $V_1$, $V_2$, $V_3$, and $V_4$ are all analog values, so create a database of the process states according to the control logic of the processes. Extract the data from Modbus protocol at regular time and create ranked vectors for parameters of key infrastructures as (2). In steady states of manufacturing processes, the parameters keep steady too. Suppose the last state is normal, calculate the Euclidian distance between the realtime state and the last state first by (3). If this distance is less than the threshold value, the data is in normal range. Otherwise, query the history states to construct the state matrix $P$ as (1) and calculate the correlation score $\chi$ by (4). Then calculate the threshold of the correlation score $\gamma$ by (5) (6) and (7). At last (8) is the result to judge whether the realtime data is in normal range or not.

Remark 1: Whitelist and correlation score methods are used to the digital and analog parameters to validate the realtime states of the controlled processes to make sure that the control logics are not broken by attackers. Malicious behaving caused by Replay, MITM, and Zero-day attacks can be detected by these methods to guarantee the security of manufacture in plants.

Remark 2: The reason for us to apply the validation in the control level is safety. In many SCADA systems the signal from sensors to the PLCs are transmitted by hard wire. Values from sensors are transmitted by $4-20mA$ currents which is impossible to be tampered. Even these values are transmitted by BUS, it’s difficult to tamper the data from the sensors’ side and we can extract the data from RTU before the data can be tampered in PLCs. So the data in the running database are the true values of the sensors.

IV. THE ANALYSIS OF TRAFFIC DATA IN SUPERVISORY LEVEL

In order to perform a low-consume and time-efficient intrusion detection, the pre-processing of raw date is necessary which including: map symbolic-valued attributes such as date, IP address and connect type into numeric-valued attributes; change the out-scaled false values such as Buffer Overflow values into right but featured values; and extract the feature of the data by T-distributed Stochastic Neighbor Embedding(t-SNE).

A. FEATURE EXTRACTION FOR TRAFFIC DATA

Data used to detect intrusion in ICS include two category: protocol data and traffic data. Table 2 shows the feature of part of the data, however, the complete data are high dimensional data including 63 variables. T-SNE is used here as the visualization method for the data. High dimension data can be reduced to 2-D or 3-D space while respecting relative distances between data samples. Two dimensional scatter plots of 500 samples with DoS attacking is shown in Fig.4 after the application of t-SNE, which we can see the branching feature. This feature is because DoS attack sometimes seems normal in the beginning and its attacking feature will be apparent after some period. The branching feature make it difficult to classify the data sets by one surface or two parallel surfaces.

B. FUZZY CLASSIFIER FOR TRAFFIC DATA

Because of the large amount of data generated in plants, data-driven fuzzy system is widely used in industrial applications such as fault diagnosis or decision support systems. The advantage of fuzzy system is that it’s able to model a given problem to any degree of accuracy [27]. Our proposal to use Takagi-Sugeno (T-S) fuzzy model is to detect the intrusion by dividing the data into two clusters: normal and abnormal.

There are two phases for data driven T-S fuzzy system modeling [28]: in phase one, we determine by what rules and how many regions we will divide the data sets. More regions leads high accuracy but too many regions will lead to a complex fuzzy model as well as overfitting problem, while too less regions will reduce the accuracy of the fuzzy system. In phase two, we determine the linear parameter of each rule which can be viewed as a linear regression problem.
We adopt T-S fuzzy model for d-dimension input variable:

\[ \mathbf{x} = (x_1, x_2, \ldots, x_d)^T \in \mathcal{R}^d \]

the i-th fuzzy rule is presented as:

\[ \mathcal{R}_i^c : IF \ x_1 \ is \ A_{i1}^c, \ \ldots \ and \ x_d \ is \ A_{id}^c \ THEN \ \ y^i = l_i(\mathbf{x}) \]

\(i = 1, 2, \ldots, r\). \(r\) is the number of fuzzy rules, \(y^i\) is the output variable,

\[ l_i(\mathbf{x}) = \omega_{i0} + \omega_{i1}x_1 + \omega_{i2}x_2 + \cdots + \omega_{in}x_n = [1, \mathbf{x}^T]^T \omega_i \]

is the fuzzy rule consequent,

\[ \omega_i = [\omega_{i0}, \omega_{i1}, \omega_{i2}, \ldots, \omega_{in}]^T \]

is the column vector of parameters for the i-th fuzzy rule consequent part. \(A_i^c(j = 1, 2, \ldots, d)\) are linguistic proposition characterized fuzzy membership functions.

\[ \mu_{A_j}(x_j) = \exp[-\frac{(x_j - c_{ij})^2}{\sigma_{ij}^2}] \]

(9)

here \(c_{ij}\) is the mean of the corresponding membership function and \(\sigma_{ij}\) is the variance in each dimension.

The output function of the T-S fuzzy system with weighted-average defuzzifier is derived as:

\[ y(\mathbf{x}) = \sum_{i=1}^{r} \phi_i(\mathbf{x})y^i \]

(10)

here \(\phi_i(\mathbf{x})\) is the firing strength of the i-th fuzzy rule with respect to input variable \(\mathbf{x}\) by:

\[ \phi_i(\mathbf{x}) = \frac{\prod_{j=1}^{d} \mu_{A_j}(x_j)}{\sum_{k=1}^{r} \prod_{j=1}^{d} \mu_{A_j}(x_j)} \]

(11)

Then, the T-S fuzzy classifier predict its label by the following decision function:

\[ \text{sign}(y(\mathbf{x})) = \begin{cases} +1, & \text{if } y(\mathbf{x}) \geq 0 \\ -1, & \text{otherwise} \end{cases} \]

(12)

\section*{C. ANTECEDENT PARAMETERS EXTRACTION}

Fuzzy rule antecedent can be obtained based on prior knowledge. But, in many cases, the priori knowledge is not easy to obtain. For sample sets:

\[ \mathcal{X}^+ = \{(x_i^+, +1)\}_{i=1}^{m_1}, \ \mathcal{X}^- = \{(x_i^-, -1)\}_{i=1}^{m_2} \]

clustering methods can be used to divide the input space into blocks and obtain the Antecedent Membership Function (AMF). Iterative Vector Quantization (VQ) clustering approach is used in this paper to extract the antecedent of the fuzzy rule.

By the iterative VQ clustering algorithm, each fuzzy rule is related with one cluster. The bell-shaped membership function \(A_j^c\) generated by the projection of the clusters onto each variables, i.e., the cluster center is viewed as the mean \(c_{ij}\) for \(i = 1, 2, \ldots, r\) and the variance in each dimension is viewed as \(\sigma_{ij}\) for \(j = 1, 2, \ldots, n\). we set \(\sigma_{ij} = \max(\sigma_{ij}, \epsilon)\) with a small value of \(\epsilon\) avoid the width of bell-shaped membership function is too small.

\section*{D. CONSEQUENT PARAMETERS LEARNING}

Several methods are used to estimate fuzzy consequent parameters such as LS [29], genetic algorithm (GA) [30], and gradient descent [31]. But these methods are all based on empirical risk minimization without considering the test error. So methods based on large margin criterion were used in many literatures [32]. But the shortage of these methods is that they can not handle the cross or branching classification problem well like Fig.4. By using the t-SNE we have seen the traffic data we collected from SCADA system has the branching feature. The accuracy of SVM to solve this problem is 90%, and fuzzy-based SVM classifier with the accuracy of 93.73% [31]. The above two algorithms perform poor for branching classification problem.

\section*{E. NONPARALLEL HYPERPLANE BASED FUZZY CLASSIFIER FOR TRAFFIC DATA}

J. Li used fuzzy based two nonparallel hyperplanes to solve the cross classification problems in [33] with each hyperplane is as close as possible to one of the data set and as far as possible from the other data set. It can solve the cross classification problems well and improved the weakness as SVM, poor interpretability.

From the result of T-S fuzzy classifier:

\[ \mathcal{X} = \{(x_i, y_i)\}_{i=1}^{m} \in \mathcal{X} = \{+1, -1\} \]

clustering methods can be used to divide the input space into some fuzzy regions.

Set class 1 as the positive set and class 2 the negative set:

\[ \mathcal{X}^+ = \{(x_i^+, +1)\}_{i=1}^{m_1}, \ \mathcal{X}^- = \{(x_i^-, -1)\}_{i=1}^{m_2} \]

Then, the T-S fuzzy classifier predict its label by the following decision function:

\[ x^T \mathbf{w} - \gamma = 0 \]

(13)

The first plane is closest to class 1 data set and furthest from class 2 data set, while the second plane is closest to class 2 data set and furthest from class 1 data set.

For the T-S fuzzy model we discussed above, the consequent of the fuzzy rule is defined as:

\[ l_i(\mathbf{x}) = [1, \mathbf{x}^T] \mathbf{w}_i \]

(14)

where \(\mathbf{w}_i = [\omega_{i0}, \omega_{i1}, \ldots, \omega_{id}]^T \in \mathcal{R}^{d+1}\) indicates the column vector of parameters for the i-th fuzzy rule consequent part.

The following definition is given:

\[ \Psi(x) = [\phi_1(x)[1, \mathbf{x}^T], \phi_2(x)[1, \mathbf{x}^T], \ldots, \phi_r(x)[1, \mathbf{x}^T]]^T \]

(16)
The output of T-S fuzzy model is:

$$y(x) = \Psi(x)^T W$$ (17)

Now, the T-S model can be seen as a map: the data set from a cluster to a richer feature space, then a hyperplane can be constructed in the richer feature space. Formally, map the data with:

$$\Psi : \mathbf{x}^1 \mapsto \Psi(x)$$
$$W : \mathbf{w}^1 \mapsto W_1, \quad \mathbf{w}^2 \mapsto W_2$$ (18)

then learn the linear relationship from $$\Psi(x)$$ to $$y$$ as (16).

So we define two nonparallel hyperplanes in the richer feature space corresponding to (13) as:

$$\Psi(x)^T W_1 - \gamma I = 0, \quad \Psi(x)^T W_2 - \gamma I = 0$$ (19)

These two hyperplanes in (19) are used to determine whether a sample point belong to positive class or negative class. In other words, if a sample point is close to one hyperplane and far away from another one, then it belongs to positive class, and vice versa.

To achieve an excellent classification performance of these two hyperplanes, each hyperplane is expected to be as close as possible to one class set, simultaneously as far as possible to the other class set, which can be deduced as the following two optimization problems:

$$\min_{W_1 \neq 0} \frac{\| PW_1 - \gamma I \|^2}{\| W_1 \gamma I \|^2}$$
$$\min_{W_2 \neq 0} \frac{\| PW_2 - \gamma I \|^2}{\| W_2 \gamma I \|^2}$$ (20)

where

$$\begin{bmatrix}
\Psi(x_1^-)^T \\
\Psi(x_2^-)^T \\
\vdots \\
\Psi(x_m^-)^T
\end{bmatrix}, \quad Q = \begin{bmatrix}
\Psi(x_1^+)^T \\
\Psi(x_2^+)^T \\
\vdots \\
\Psi(x_m^+)^T
\end{bmatrix}, \quad P \in \mathbb{R}^{m_1 \times r(d+1)}, \quad Q \in \mathbb{R}^{m_2 \times r(d+1)}$$ (22)

Simplify the optimization equation (20), we have:

$$\min_{W_1 \neq 0} \frac{\| PW_1 - \gamma I \|^2}{\| QW_1 - \gamma I \|^2}$$ (23)

To avoid the overfitting problem, regularization is necessary here, so a nonnegative parameter $$\varepsilon$$ is used as follows:

$$\min_{W_1 \neq 0} \frac{\| PW_1 - \gamma I \|^2 + \varepsilon}{\| QW_1 - \gamma I \|^2}$$ (24)

Let’s define:

$$\begin{align*}
R &= [P - I]^T [P - I] + \varepsilon I \\
S &= [Q - I]^T [Q - I] \\
Z &= [W_1 \gamma]
\end{align*}$$ (25)

where $$R$$ and $$S$$ are symmetric matrices, then the optimization equation (24) becomes:

$$\min_{Z \neq 0} h(Z) := \frac{Z^T R Z}{Z^T S Z}$$ (26)

The function (26) is known as the Rayleigh quotient [34] which has the properties (of Boundedness) as follows:

1) **Boundedness** The Rayleigh quotient ranges over the interval $$[\lambda_1, \lambda_{m+1}]$$ as $$Z$$ ranges over the unit sphere, where $$\lambda_1$$ and $$\lambda_{m+1}$$ are the minimum and maximum eigenvalues of the generalized eigenvalue problem:

$$RZ = \lambda SZ$$ (27)

2) **Stationarity**

$$\nabla r(Z) = 2(\frac{GZ - r(Z)HZ}{ZH}) = 0$$ (28)

Thus, $$r(Z)$$ is stationary at and only at the eigenvectors of the generalized eigenvalue problem.

Assume the eigenvector of (27) corresponding to the minimum eigenvalue is denoted by $$Z^*$$, then the first hyperplane in (19) is:

$$\Psi(x)^T Z^* = 0$$ (29)

By the same process, the second hyperplane in (19) is:

$$\Psi(x)^T O^* = 0$$ (30)

This process is named Nonparallel Hyperplane based Fuzzy Classifier (NHFC), and the steps are in Algorithm 2.

**Algorithm 2 NHFC Algorithm**

1. Cluster the positive and negative class set respectively by T-S fuzzy algorithm to get the result as $$S^+ = \{x_i^+, +1\}_{i=1}^{m_1}, S^- = \{x_i^-, -1\}_{i=1}^{m_2}$$.
2. Constructing matrices $$P$$ and $$Q$$ as (22) based on (16).
3. Constructing matrices $$R$$, $$S$$, $$Z$$, as (25) and (26).
4. Compute the eigenvector of (27) corresponding to the smallest eigenvalue as $$Z^*$$, and by the same process to get $$O^*$$.
5. Output the classifier result by (17) as

$$h(\Psi(x)^T Z^*, \Psi(x)^T O^*) = \begin{cases} 
+1, \quad &\text{if } \Psi(x)^T Z^* \leq \Psi(x)^T O^* \\
-1, \quad &\text{otherwise}
\end{cases}$$

**Remark 3:** The principle of algorithm NHFC is to map the vector $$x$$ to the space $$\Psi(x)$$ and obtain two hyperplanes by $$W_1$$ and $$W_2$$ as defined in (18). The character of two hyperplanes is the difference between NHFC and the traditional classification algorithms. The principle of traditional algorithms is to find one hyperplane, one side of which is the positive class and the other side is the negative class.

**Remark 4:** The characteristic of two nonparallel hyperplanes of NHFC shows more superiority than the traditional classification algorithms with only one hyperplane when used in the classification of data with branching feature.
And the good interpretability of NHFC algorithm is superior to GEPSVM.

V. APPLICATION AND DISCUSSION

In supervisory level, a data set of 41 dimension and including 2200 samples are collected from a SCADA system which are combined with Modbus/TCP protocol data and traffic data. All simulations are implemented in MATLAB 2014a(64-bit) environment and the configure of the computer is inter(R) Core(TM) i5-4460 CPU 3.2GHz.

FIGURE 5. The intrusion detection by NHFC.

For the 2200 sets of effective data, the output of each sample is set to be an binary variable. The 2200 sets of effective data is divided into two groups. 1600 samples of the data sets are used to train the model, and 600 samples are used to test. Fig. 5 presented the performance of our method. The blue ° on axis 1 shows the labeled normal data, and the blue ° on axis 2 shows the labeled abnormal data. The red * on axis 1 shows the predicted normal data, and the red * on axis −1 shows the predicted abnormal data. From Fig. 5 we can see the result of prediction is very good with the accuracy rate of 97.22%.

TABLE 5. Classification accuracy of different algorithms.

| Algorithm | Training (AVG±STD) | Testing (AVG±STD) |
|-----------|---------------------|-------------------|
| NHFC | 93.3±0.32 | 92.7±0.45 |
| SVM | 93.6±0.48 | 94.2±0.33 |
| OCSVM | 90.3±0.53 | 94.4±0.58 |
| FCMSVM | 91.5±0.39 | 94.3±0.55 |
| GENFIS | 91.5±0.39 | 94.3±0.55 |

Table 5 shows the average classification accuracy with 30 independent runs for some representative classification algorithms. The hyper parameter of the algorithms are determined by 10-fold cross-validation. The format is (AVG±STD), here AVG is the average and STD is the standard deviation of the accuracy on 10-fold cross-validation. We can see that our proposed method perform better than SVM, OCSVM, FCMSVM and GENFIS, and this method has a better interpretability.

The nonparallel hyperplanes make an important contribution in the detection of DoS (SYN flood) intrusion which is difficult to be detected because of the feature of the branching shape.

In control level, simulation data are used to validate the effectiveness of the state based method. Table 4 are part of the state data of Fig.3 which are indicated by the values of normalized $4 - 20mA$ current range in (0, 1). If the realtime measures of $F$, and $V_1$-$V_4$ are: 0.75, 0.88, 0.88, 0.86, 0, then from (4)-(7),

$$\chi = 0.135, \quad \mu = 0.385, \quad \sigma = 0.025$$

Let $\beta = 1$, then $y = 0.41$. The result is $\chi < y$, so the realtime state is normal.

The detection accuracy of the proposed approach is based on the analysis of critical states, and the number of critical states of different process are different too. Test should be carried out to each process to choose a appropriate $\beta$ which can minimize the false rate.

VI. CONCLUSION

The validation of process states is used to guarantee the correlation of the parameters in the controlled processes. White list is used to validate the digital variables and correlation scores are used to validate the analog variables in the processes. The validation can detect the data faults caused by MITM attacks, Replay attacks, and Zero-day attacks to protect the control processes. The nonparallel hyperplanes based fuzzy classifier are used to classify the data with branching feature, and attacks like DoS and Buffer Overflow can be detected with high accuracy by this algorithm. The process state based method and the NHFC based algorithm are in ‘or’ relation. Comparisons are provided to the same data sets, and the proposed hybrid method has a high accuracy in detecting Replay attack, MITM attack, Zero-day attack, and DoS intrusion in both cyber and physical fields of the networks of SCADA system.

By using this provided scheme, we can detect whether the above attacks are happening but can not make sure the
location and the type of the attacks. In the future study, we’ll do more research to locate the attacks and identify the type of the attacks. Further more we can do more work to defend the attacks to protect the manufacturing process.

APPENDIX
See Table 6.

REFERENCES

[1] Guide to Industrial Control Systems (ICS) Security, NIST, Gaithersburg, MD, USA, 2015.

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

[3] ICS Vulnerabilities: 2018 in Review. Accessed: 2019. [Online]. Available: https://www.ptsecurity.com/wc-en/analytics/ics-vulnerabilities-2019/

[4] A. Abbasi, T. Scharnowski, and T. Holz. Doors of Durin: The Veiled Gate to Siemens S7 Silicon. Accessed: 2019. [Online]. Available: https://www.blackhat.com/eu-19/briefings/schedule/#doors-of-durin-the-veiled-gate-to-siemens-s7-silicon-18023

[5] K. C. Mahapatra and S. Megash, “Analysis of vulnerabilities in the protocols used in SCADA systems,” Int. J. Adv. Res. Comput. Eng. Technol., vol. 4, no. 3, pp. 1014–1019, 2015.

[6] C. Wang, J. Zheng, and X. Li, “Research on DDoS attacks detection based on RED-SVM,” in Proc. 10th Int. Conf. Intell. Comput. Technol. Autom. (ICICTA), Oct. 2017, pp. 161–165.

[7] R. M. Lee, M. J. Assante, and T. Conway. German Steel Mill Cyber Attack. Accessed: Dec. 30, 2014. [Online]. Available: https://ics.sans.org/media/ICS-CPEP-case-Study-2-German-Steelworks_Facility.pdf

[8] DESMI Hit by Cyber Attack. Accessed: 2020. [Online]. Available: https://www.desmi.com/news-(5)/desmi-hit-by-cyber-attack.aspx

[9] Portuguese Energy Giant Hit by Ransomware Attack. Accessed: 2020. [Online]. Available: https://www.power-eng.com/2020/04/20/portuguese-energy-giant-hit-by-ransomware-attack/

[10] J. Qian, C. Hua, X. Guan, T. Xin, and L. Zhang, “A trusted-ID referenced key scheme for securing SCADA communication in iron and steel plants,” IEEE Access, vol. 7, pp. 46947–46958, 2019.

[11] R. H. Jhaveri, A. D. Patel, and K. J. Dangarwala, “Comprehensive study of various DoS attacks and defense approaches in MANETS,” in Proc. Int. Conf. Emerg. Trends Sci., Eng. Technol. (INCOSET), Dec. 2012, pp. 25–31.

[12] A. Sperotto and R. Sadre, “A labeled data set for flow-based intrusion detection,” in Proc. Int. Workshop IP Oper. Manage., 2009, pp. 29–30.

[13] Y. Xie and S.-Z. Yu, “A large-scale hidden semi-Markov model for anomaly detection on user browsing behaviors,” IEEE/ACM Trans. Netw., vol. 17, no. 1, pp. 54–65, Feb. 2009.

[14] S. Ganapathy, P. Yogesh, and A. Kainan, “Intelligent agent-based intrusion detection system using enhanced multiclass SVM,” Computational Intell. Neurosci., vol. 2012, Sep. 2012, Art. no. 850259, doi: 10.1155/2012/850259.

[15] W. Shang, J. Cui, C. Song, J. Zhao, and P. Zeng, “Research on industrial control anomaly detection based on FCM and SVM,” in Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun./12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE), Aug. 2018, pp. 2324–9013.

[16] M. Wan, W. Shang, and P. Zeng, “Dynamic characteristics for one-class classification anomaly detection in networked control systems,” IEEE Trans. Inf. Forensics Security, vol. 12, no. 12, pp. 3011–3023, Dec. 2017.

[17] F. Zhang, H. A. D. E. Kodituwakku, J. W. Hines, and J. Coble, “Multilayer data-driven cyber-attack detection system for industrial control systems,” IEEE Trans. Ind. Informat., vol. 15, no. 7, pp. 4362–4369, Jul. 2019.

[18] H. Lin, A. Slagell, Z. Kalbarczyk, P. W. Sauer, and R. K. Iyer, “Semantic security analysis of SCADA networks to detect malicious control commands in power grids,” in Proc. 1st ACM Workshop Smart Energy Grid Secur. (SEGS), 2013, pp. 29–34.

[19] D. Hadz’iosmanović, R. Sommer, and E. Zambon, “Through the eye of the PLC: Semantic security monitoring for industrial processes,” in Proc. 30th Annu. Comput. Secur. Appl. Conf., 2014, pp. 126–135.

[20] H. Ghaeini and N. Tippenhauer, “HAMIDS: hierarchical monitoring intrusion detection system for industrial control systems,” in Proc. 2nd ACM Workshop Cyber-Phys. Syst. Secur. Privacy, 2016, pp. 103–111.

[21] H. Ghaeini, M. Chan, R. Bahmani, F. Brasser, L. Garcia, J. Zhou, A. R. Sadeghi, N. Tippenhauer, and S. Zonouz, “Patt: Physics-based attestation of control systems,” in Proc. 22nd Int. Symp. Res. Attacks, Intrusions Defenses (RAID), 2019, pp. 165–180.

[22] H. R. Ghaeini, N. O. Tippenhauer, and J. Zhou, “Zero residual attacks on industrial control systems and stateful countermeasures,” in Proc. 14th Int. Conf. Availability, Rel. Secur. (ARES), 2019, pp. 2236–2246.

[23] B. Jeffries, J. W. Hines, and K. C. Gross, “Behavior-based approach to misuse detection of a simulated SCADA system,” in Proc. 10th Nucl. Plant Instrum. Control Hum., Mach. Interface Technol., San Francisco, CA, USA, Jun. 2017, pp. 1761–1771.

[24] A. Almalawi, A. Fahad, T. Zari, A. Alamri, R. AlGhamdi, and A. Y. Zomaya, “An efficient data-driven clustering technique to detect attacks in SCADA systems,” IEEE Trans. Inf. Forensics Security, vol. 11, no. 5, pp. 893–906, May 2016.

[25] E. Lugofoer and S. Kindermann, “SparseFIS: Data-driven learning of fuzzy systems with sparsity constraints,” IEEE Trans. Fuzzy Syst., vol. 18, no. 2, pp. 396–411, Apr. 2010.

[26] M. Luo, F. Sun, and H. Liu, “Hierarchical structured sparse representation for T–S fuzzy systems identification,” IEEE Trans. Fuzzy Syst., vol. 21, no. 6, pp. 1032–1043, Dec. 2013.

[27] S. Umeyama, “Least-squares estimation of transformation parameters between two point patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 13, no. 4, pp. 376–380, Apr. 1991.

[28] K. Shimojima, T. Fukuda, and H. Hasegawa, “Self-tuning fuzzy modeling with adaptive membership function, rules, and hierarchical structure based on genetic algorithm,” Fuzzy Sets Syst., vol. 71, no. 3, pp. 295–309, May 1995.

[29] Y. Shi and M. Mizumoto, “Some considerations on conventional neuro-fuzzy learning algorithms by gradient descent method,” Fuzzy Sets Syst., vol. 112, no. 1, pp. 51–63, May 2000.

[30] W. Cheng and C. Jiang, “An incremental support vector machine-trained TS-type fuzzy system for online classification problems,” Fuzzy Sets Syst., vol. 163, no. 1, pp. 24–44, 2011.

[31] J. Li, C. Hua, Y. Yang, and X. Guan, “Fuzzy classifier design for development of a high speed sensor,” in Proc. 10th Nucl. Plant Instrum. Control Hum., Mach. Interface Technol., CA, USA, Jun. 2017, pp. 1761–1771.
XUEQIANG DU was born in Tangshan, Hebei, China, in 1972. He received the B.S. degree in metallurgy, in 1995, and the M.S. degree in automation of metallurgy from the Hebei Institute of Technology, Tangshan, in 1998. He is currently pursuing the Ph.D. degree with the Automation Research and Design Institute of Metallurgical Industry, China.

Since 1998, he has been engaged in designing, research and development, and integration of metallurgical automation. His current research interests include simulation and optimization of metallurgical energy systems.

BO CHEN was born in Xinzhou, Shanxi, China, in 1984. He received the B.S. degree in automation from Xidian University, Xi’an, in 2005, and the Ph.D. degree in signal and information processing from the University of Chinese Academy of Sciences, in 2010.

Since 2016, he has been an Associate Professor with the North China University of Science and Technology. His research interests include control technology and its application, and digital signal processing.

BIN QU was born in Baoding, Hebei, China, in 1975. He received the B.S. degree in agricultural electrification automation and the M.S. degree in agricultural mechanization engineering from Hebei Agricultural University, Baoding, in 1998 and 2003, respectively. He is currently pursuing the Ph.D. degree with the Hebei University of Technology.

Since 2003, he has been a Lecturer with the North China University of Science and Technology. His research interests include modeling and simulation of complex systems and new energy generation control technology.

KAI ZENG was born in Tangshan, Hebei, China, in 1990. He received the B.S. degree in electrical engineering and automation and the M.S. degree in control science and engineering from the North China University of Science and Technology, Tangshan, in 2014 and 2017, respectively, where he is currently pursuing the Ph.D. degree.

He had been a Visiting Student with the University of Regina, Regina, SK, Canada, in 2016. Since 2019, he has been a Lecturer with the North China University of Science and Technology. His research interests include data analysis and machine learning.

JIANPENG LIU was born in Tianjin, China, in 1981. He received the B.S. and M.S. degrees in mineral processing engineering from the Wuhan University of Science and Technology, in 2004 and 2008, respectively.

Since 2008, he has been working with Tang Steel International Engineering Technology Corporation, Tangshan. His research interest includes intelligent manufacturing in process industry.

* * *