A Novel False Data Injection Attack Detection Model of the Cyber-Physical Power System

JIE CAO1,2, DA WANG1, ZHAOYANG QU1,2, (Member, IEEE), MINGSHI CUI3, PENGCHENG XU4, KAI XUE4, AND KEWEI HU4

1School of Computer Science, Northeast Electric Power University, Jilin 132012, China
2Jilin Engineering Technology Research Center of Intelligent Electric Power Big Data Processing, Jilin 132012, China
3EastInner Mongolia Electric Power Company, Hohhot 010000, China
4State Grid Jilin Province Electric Power Supply Company, Changchun 130000, China

Corresponding author: Zhaoyang Qu (qzywww@neepu.edu.cn)

This work was supported in part by The National Natural Science Foundation Project of China under Grant 51437003, in part by The Science and Technology Project of Jilin Provincial Education Department under Grant JJKH20200122KJ, in part by The Science and Technology Development Plan Project of Jilin Province under Grant 20180201092GX, and in part by the Science and Technology Innovation Development Plan Project of Jilin City under Grant 201830817.

ABSTRACT

The False data injection attack (FDIA) against the Cyber-Physical Power System (CPPS) is a kind of data integrity attack. With more and more cyber vulnerabilities detected out, different types of FDIA are emerging as severe threats to the stable operation of CPPS gradually. In this paper, the invasion pathway of the FDIA against CPPS is explored in detail, and a novel FDIA detection model based on ensemble learning is further provided. First, a pseudo-sample database is built to assist the training and evaluation of this model, and it’s more important to update the model in the future. Furthermore, the optimal feature set is extracted to characterize the behavior of the FDIA, which improves the precision of the FDIA detection model. Finally, a focal-loss-lightgbm (FLGB) ensemble classifier is constructed to detect the FDIA behavior automatically and accurately. We illustrated the performance of this model by a fusion of measurement data and power system audit logs. This model utilizes the offline training way, the conclusion shows the high precision and stability of this model, which ensures the stable operation of the smart grid and improves the FDIA resistance ability of the CPPS.

INDEX TERMS

CPPS, FDIA detection model, invasion pathway analysis, ensemble learning.

I. INTRODUCTION

The CPPS fuses computing equipments, communication systems, and the physical power grid into a multidimensional, isomerous, and complex system [1], [2]. Compare to the traditional physical electric power grid, CPPS is more intelligent and steady. However, due to the existence of the vulnerabilities in cyberspace, especially the complex interactive process between electric power flow and information flow, information security has become an important factor that affects the safe and stable operation of the power system. Therefore, the power system is facing serious threats of network attacks continuously. Different from the network attacks in the Internet field, cyber-attacks against CPPS concentrate more on destroying the stability control of the cyber layer over the physical layer, even paralyze the operation of the power system [3]. The blackout in Ukraine is the most representative network attack against the power system, which caused the operator to lose control of the system, lead to a wide range of blackouts and serious economic loss [4].

In the several types of cyber-attacks against CPPS, the FDIA is a type of data integrity attack [5]. To launch the FDIA, the attacker utilizes the vulnerabilities of the communication system in the cyber layer, and then modifies the measurement data of physical devices maliciously [6], [7]. The FDIA leads the control center to make an incorrect state estimate of the current system. Then the control center loses the normal control ability of the physical equipment, which causes the power system to collapse, even leads to a large-scale power failure. In recent years, more kinds of FDIA have emerged, the attacker can use network vulnerabilities to tamper with the measurement data, control data, even the equipment information with the minimum cost, which cause the large-scale chain failures [8]. However, the defense aiming at FDIA become more difficult because of the great concealment of FDIA. Therefore, it is urgent to propose an
FDIA detection method with high precision, which lays the foundation for the defense of FDIA. In order to improve the FDIA resistance ability of CPPS, there are three main FDIA detection methods: 1) The detection methods based on state estimation; 2) The detection methods based on time series prediction; 3) The detection methods based on machine learning.

A. THE DETECTION METHODS BASED ON STATE ESTIMATION

The smart grid adopts the least-squares method to finish the state estimation. If attackers get a sufficient understanding of the information and protection algorithms of the power grid, they can build false data that evades the state estimation algorithm. Build on the characteristics of this kind of FDIA, the traditional state estimation algorithm has been improved, which enhances the ability of CPPS to resist FDIA. The improved methods of state estimation mainly include the Kalman filtering method, residual detection method, and measurement correlation detecting method [9]–[11]. Besides, considering that the setting of the detection threshold is affected by the scale of the power system greatly, the global power system can be divided into sub-systems, and different detection thresholds are set according to the actual parameters of each sub-system [12], [13]. In this way, the FDIA against complex power systems can be detected effectively. The advantage of this kind of method is that it uses mature algorithms, detects FDIA rapidly, and can reflect the characteristics of the power system better. However, the setting of the detection threshold brings empirical error, which causes a bad impact on the detection accuracy. What’s more, when fluctuations happen in the power system, it is prone to cause the false alarm of FDIA [14].

B. THE DETECTION METHODS BASED ON TIME SERIES PREDICTION

Detection methods based on state estimation are used primarily for static analysis, which detects attack behavior at a specific moment. During the continuous dynamic operation of the power system, there is a strong space-time relationship between multiple state variables. Therefore, it’s feasible to use historical data for trajectory analysis to predict the current state of the power grid. Then compare the prediction result with the current actual measurement value, and analyze the areas that may be attacked by the similarity comparison. The detection methods based on time series prediction mainly include the statistical consistency detection, the sequential consistency detection, and sensor trajectory prediction [15]–[17]. This kind of method predicts the distribution of state variables based on the operating law of the system state and the historical database. By matching the running trajectory, various types of false data can be detected effectively, but the computation complexity is high, which causes low detection speed. The power system can’t respond in time when FDIA happens, so it is not suitable for the FDIA detection of complex systems.

C. THE DETECTION METHOD BASED ON MACHINE LEARNING

As the deployment of wide-area measurement systems (WAMS) expands, which provides massive data for data analysis of the power system. Especially the machine learning technology is playing an important role in the FDIA detection against CPPS gradually. The machine learning algorithms establish the complex mapping relationship between input data and output data through training models, and the computation speed is fast. This kind of method doesn’t need to solve complex time-domain equations of the power system, and clear performance indicators can be used to evaluate the performance, it is one of the current main FDIA detection methods. This type of method mainly includes support vector machine (SVM) [18], extreme learning machine [19], fuzzy C-means clustering [20], deep learning [21], ensemble learning, etc., [22]. The detection performance of traditional supervised learning algorithms such as SVM depends on the quality of the data heavily, such as the poor characterization ability of the feature set will cause a low detection rate. Compared to supervised learning algorithms, the detection accuracy of the clustering method is often lower due to its unsupervised training characteristics, and the performance of the model can’t be evaluated clearly for FDIA detection. The deep learning model automatically processes features and the detection accuracy is often higher than traditional machine learning algorithms, but the training of the model relies on large sample datasets heavily, and consumes more training time, so it’s difficult to detect FDIA in real-time [23]. The ensemble learning model is constructed by combining weak classifiers into a strong classifier. Furthermore, the strong classifier adopts the integrated voting mechanism to get higher prediction accuracy. Due to the characteristic of parallel training, the computation consumption is less than deep learning algorithms significantly. So the ensemble learning method balance the accuracy and training time of the model better, is more effective for FDIA detection.

From the above analysis, an FDIA detection model based on ensemble learning is proposed. The main contributions of this work are as follows:

1) This work introduces the imbalance rate of the multi-classification data, a C-KSmote oversampling method is proposed to rebalance the electric power data according to the imbalance rate. The pseudo-data is used to assist in training and evaluating the FDIA detection model. This method generates pseudo-samples that conforms to the characteristics of power data, and decreases the false alarm rate of the FDIA detection model caused by the imbalanced data. The FDIA resistance ability of CPPS is improved obviously by this method.

2) The optimal feature set obtaining method is proposed to characterize the behavior of the FDIA. The optimal feature set contains fewer features than the original data, it not only improves the training precision of the FDIA detection model, but also decreases the model complexity. This method
improves the response speed and precision of CPPS against the FDIA.

3) The focal-loss-lightgbm ensemble classifier is constructed, which improves the detection rate of the anomalous distributed samples. Difficult types of FDIAs can be detected accurately by this classifier, and the fault and FDIA can be distinguished effectively. The ensemble classifier concentrates more on processing the difficult-classification samples and improves the FDIA detection precision obviously, which is beneficial to the stable operation of the CPPS.

The remainder of this paper is organized as follows: Section II is the mathematical modeling of FDIA and the FDIA detection model. Section III performs the theoretical analysis of our proposed FDIA detection model. Section IV is the experimental verifications and analysis of the FDIA detection model. Section V is the conclusion of this article.

II. THE MODELING OF THE FDIA DETECTION MODEL OF CPPS

A. THE DETECTION MODEL OF TRADITIONAL FDIA AGAINST CPPS

The state estimation is the crucial mechanism for maintaining the stability and efficiency of the modern power grids [24]. As is shown in Figure 1, the measurement data of the physical layer, such as bus voltage, bus power flow, branch power flow, and load profiles are collected by the sensors or meters. Then collected data is sent to the control center in the cyber layer by the industrial control system known as the SCADA system. The control center estimates the states of the power system according to the received measurement data, detects the potentials of contingency, and sends the corresponding control signals to the Remote Terminal Units (RTUs) of the physical layer, which ensure the reliable operation of CPPS and finish the closed-loop control of physical power grid.

The state estimator in the control center uses the measurement data reported from the SCADA system and the steady system model to estimate the system state over time. If variable \( z = (z_1, z_2, \ldots, z_m)^T \in \mathbb{R}^m \) represents the measurement vector, \( x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n \) is the state vector, and \( e = (e_1, e_2, \ldots, e_m)^T \in \mathbb{R}^m \) denotes the measurement error vector, the description of the observation model of DC power flow is as shown in equation (1):

\[
\begin{align*}
    z &= h(x) + e \\
    &\text{(1)}
\end{align*}
\]

The observation model of DC power flow can be further formulated by equation (2):

\[
\begin{align*}
    z &= Hx + e \\
    &\text{(2)}
\end{align*}
\]

In equation (2), the \( H \in \mathbb{R}^{m \times n} \) is the Jacobian matrix that represents the topology of the physical grid, and \( e \sim N(0, \sigma^2) \) denotes the environment noise modeled by Additive White Gaussian Noise (AWGN) with the standard deviation \( \sigma \).

According to the statistical estimation criteria (Minimum Mean-Square Error), the estimated system state \( \hat{x} \) can be further formulated by equation (3):

\[
\begin{align*}
    \hat{x} &= (H^T \wedge H)^{-1} H^T \wedge z \\
    &\text{(3)}
\end{align*}
\]

The sign \( \wedge \) is a diagonal matrix whose diagonal element is \( \wedge_{ii} = \sigma^{-2} \).

If attackers have a sufficient knowledge of the system topology that is represented by the Jacobian matrix \( H \), they can launch the FDIA by tampering the state estimation data such as load profiles. The observation model in the presence of the FDIA can be formulated by equation (4):

\[
\begin{align*}
    \hat{z}_a &= Hx + a + e \\
    &\text{(4)}
\end{align*}
\]

The variable \( a \) represents the vector of tampered data, when the FDIA happens, variable \( a \neq 0 \). The traditional FDIA detection model based on machine learning is a binary classification task. That is, if there are a total of \( i \) samples in the dataset \( X = \{X_1, X_2, \ldots, X_i\} \), the detection result is \( y = \{y_1, y_2, \ldots, y_i\} \), then the detection model of FDIA can be
characterized as follows:

\[
y_i = \begin{cases} 
1, & \text{FDIA} \\
0, & \text{noFDIA}
\end{cases}
\]  

(5)

As more and more vulnerabilities of the cyber layer are discovered, multiple types of FDIAs have been proved to have the ability of invading the CPPS potentially. However, the traditional FDIA detection mechanism is only designed to detect a certain type of FDIA, which limits the detecting range.

**B. THE DETECTION MODEL OF NEW EXTENSION OF FDIA**

The CPPS contains all links to the generation, transmission, distribution, and the consumption of electric power. The traditional FDIA only considers the intruding scheme for tampering with the measurement data of a specific power sub-system. With the continuous deep research of FDIA, the definition of FDIA has been further expanded. As shown in Figure 2, in a broad sense, the FDIA may occur at various abstract layers of CPPS, more specifically, the attackers may launch the FDIA against the monitoring, control, and protection devices, which damage the control system or the applications associated with it such as the generation prediction, state estimation, economic dispatch, and energy trading [25]. Therefore, when designing an FDIA detection model, the detection of different kinds of FDIA should be taken into consideration.

In our proposed FDIA detection model, we apply the ensemble learning technique to detect different types of FDIA against CPPS. When FDIA occurs, it often accompanies the fault of electrical equipment in the physical layer after a short time. So when the transient process happens in the power system, the characteristic of measurement data is far similar to that when FDIA happens. To distinguish the fault and FDIA accurately, the detection of the fault of the power system is also taken into consideration especially when designing this model. To reflect the relationship between the input data and the detection results, the multi-classification ensemble classifier is used to modeling this FDIA detection model. For a multi-classification task, if the original dataset is \( D \), there are \( j \) samples in the dataset, the dimension of the dataset is \( n \), then the dataset \( D \) can be described as equation (6):

\[
D = \begin{pmatrix}
D_{11} & D_{12} & \cdots & D_{1n} \\
D_{21} & D_{22} & \cdots & D_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
D_{j1} & D_{j2} & \cdots & D_{jn}
\end{pmatrix}
\]  

(6)

If the detection result is \( y \), the variable \( y \) is a list that contains the number of \( \{0, 1, 2, 3, 4\} \), the length of the variable \( y \) is \( j \). The labeled number and the final detection result of FDIA is as the description in equation (7):

\[
y_j = \begin{cases} 
0, & \text{noFDIA} \\
1, & \text{fault} \\
2, & \text{FDIA type 1} \\
3, & \text{FDIA type 2} \\
4, & \text{FDIA type 3}
\end{cases}
\]  

(7)

As is shown in equation (7), the FDIA type 1, type 2, and type 3 are measurement data tampering attack, remote trip instruction tampering attack, and relay setting tampering attack respectively.

The model we proposed in this paper expands the scope of FDIA detection. In a narrow sense, this model can distinguish if there is FDIA hacking into the CPPS efficiently. Broadly speaking, the specific type of FDIA can be detected accurately, and the false alarm rate is decreased obviously when the transient process happens in the physical layer.
C. THE FRAMEWORK OF THE FDIA DETECTION MODEL

The framework of the model is as shown in Figure 3. The original measurement data is collected by the Phasor Measurement Unit (PMU), and missing values in the data are filled with the mean value of the same column. To reduce the false alarm rate caused by imbalanced data, we proposed a C-KSmote over-sampling method to generate pseudo-data. The pseudo-data was stored into the pseudo-sample database, which assists in training and evaluating the ensemble classifier, and further updating the model. Based on the final balanced data, a new feature set was constructed, this feature set contains 32 electrical variables with real physical meaning. Then the constructed feature set was inserted into the original feature set randomly, which makes the feature space more suitable for characterizing the FDIA behavior.

To reduce the dimension of the data, the JMIM feature was used to extract the optimal feature subset for FDIA detection. Finally, the loss function (Focal Loss) was used to improve the lightgbm ensemble learning algorithm, which increases the weight of samples with the low prediction accuracy, and further enhances the detection precision of FDIA.

III. THE THEORETICAL ANALYSIS OF THE FDIA DETECTION MODEL OF CPPS

A. THE C-KSMOTE OVERSAMPLING METHOD OF THE ELECTRIC POWER DATA

In a quantitative collected dataset, the normal data occupies a larger proportion when the power system operating continuously. On the contrary, the probability of the FDIA happening is lower, so the data size of the FDIA category is smaller. The performance of the classifier tends to bias to the majority of normal samples due to the imbalanced dataset [26]. In the case of minority FDIA samples misclassified, the model may still obtain a high accuracy on the test set. It is unreasonable to use the total accuracy to measure the performance of the FDIA detection model. What’s more, the stability of the FDIA detection model is important for the security of the power system, so the false alarm rate caused by the serious imbalanced dataset is more costly. It is significant to adopt an appropriate oversampling strategy to balance the tampered data.

Aiming at the above problem, a C-KSmote oversampling method was proposed to generate pseudo-data, the pseudo-samples are stored into the pseudo-sample database for further use. First, we define an imbalance rate quantitative method of a multi-class dataset. If the number of minority samples is \( a \), and the average sample size of each class is \( b \), then the imbalance rate of the multi-class dataset is \( r = (b - a)/a \). If the imbalance rate of the data exceeds a predetermined threshold, then oversampling is performed to all minority samples. The principle of the C-KSmote oversampling method is as shown in Figure 4:

For each class of the minority samples, first cluster the minority samples into clusters (isolated samples do not participate in clustering and oversampling), the noise samples \( \{A, B, C, D, E, F\} \) are filtered when clustering,
when finishing oversampling, the noise samples are added into the balanced dataset. Then assign weights to each cluster. If the number of samples in a cluster is small, the weight of this cluster is high. On the contrary, the clusters contain more samples get low weight. For each of the clusters, select a sample from the cluster randomly, and perform a linear interpolation of the cluster center and the selected sample to generate a new sample. Repeat the above steps until all clusters finished oversampling. The detailed principle of the C-KSmote algorithm is as shown in Algorithm 1.

After all clusters finish oversampling, the pseudo-data is stored into the non-relational database to construct the pseudo-sample database. The pseudo-sample database is then used to assist in training and evaluating the classifier, and update the model in the future.

The proposed C-KSmote method manages to avoid the generation of noise by oversampling only in minority sample areas. Moreover, it focus on both the between-class imbalance and within-class imbalance, which generates enough pseudo-data that is more similar to the real power data. Besides, the pseudo-sample database is used not only for training and evaluating the classifier, but also updating the model in the future periodically.

B. THE OPTIMAL FEATURE SET OBTAINING METHOD FOR FDIA DETECTION

In order to obtain the optimal feature set for FDIA detection, we performed feature engineering that combines the feature construction and feature selection (FCS) method. To obtain a comprehensive feature set to characterize the FDIA behavior, we constructed 8 features of each PMU, which are mixed with original features randomly. However, The constructed features increase the dimension of the data, and there may exist redundancy features, so the mixed feature set needs
The framework of the FCS method is as shown in Figure 5:

1) FEATURE CONSTRUCTION FOR PMU MEASUREMENT DATA
In this paper, the feature construction method is used to construct new features with real physical meaning through the non-linear transformation. The constructed features are more conducive for classifiers to detect FDIA. The smart grid uses intelligent electronic devices such as smart meters, sensors to collect state data of the power system. There exist non-linear or linear dependencies between the multi-state variables of electric power data. Therefore, based on the limited original variables, new feature variables with real physical meanings can be further calculated. We constructed 32 new features and inserted these features into the original feature set randomly. To delete the useless or irrelevant features, the mixed feature set should be further processed.

2) THT OPTIMAL FEATURE SET OBTAINING METHOD
In order to obtain the optimal characterization feature set of FDIA behavior and reduce the dimension of the data, a filter feature selection method—Joint Mutual Information Maximization (JMIM) is used in this paper [27]. It’s a feature selection algorithm based on mutual information, which minimizes the data dimension without deleting important features. It’s logical to obtain the optimal feature set for FDIA detection by the JMIM method.

If the mixed feature set is $F = \{f_1, f_2, \ldots, f_N\}$, the data dimension is $N$. The principle of the JMIM algorithm is as follows:

**Definition 1:** The mutual information of variable $a$ and $b$ is defined as follows:

$$I(a, b) = H(b) - H(b|a)$$  (8)

In equation (7), $H(\cdot \cdot \cdot)$ and $H(\cdot \cdot \cdot|\cdot \cdot \cdot)$ denotes the solution of entropy and conditional entropy.

**Definition 2:** The joint mutual information of the variable $A$, $B$ and $C$ is defined as:

$$I(A, B; C) = I(A, B|C) + I(B, C)$$  (9)

Define the currently selected feature set is $S$, if the features $f_i$ and a feature $f_S$ in $S$ are highly correlated, then the mutual information of $f_i$ and $f_S$ is calculated as follows: $I(f_i, C) \approx I(f_S, C) \approx I(f_i, f_S; C)$. The JMIM employs the joint mutual information and the “maximum of the minimum” approach, which selects the most relevant features with the label. The selected features are more suitable to characterize the FDIA behavior. The equation (10) shows the final feature set selected by the JMIM algorithm:

$$f_{JMIM} = \arg \max_{f_i \in F-S}(\min_{f_S \in S}(I(f_i, f_S; C)))$$  (10)

In the feature selection process, a forward greedy search algorithm is used to select related $k$ features in the feature space, the forward greedy search strategy is shown in Algorithm 2:

C. IMPROVED ENSEMBLE LEARNING CLASSIFIER

1) GRADIENT BOOSTING DECISION TREE ALGORITHM
The Gradient Boosted Decision Tree (GBDT) is an ensemble learning framework. This framework can be expressed as an additive model of the decision tree, which uses the steepest descent method to approximate each decision tree. In each iteration, make the new decision tree follow the fastest reduction direction of the loss function (Negative gradient direction) to get higher prediction accuracy. The principle of the GBDT algorithm is as shown in Figure 6.

2) LIGHT GRADIENT BOOSTING MACHINE (LIGHTGBM)
With the geometric growth of the data scale, the GDBT algorithm faces the disadvantages of easy overfitting and slow training speed. Aiming at the above disadvantages,
Algorithm 2 Forward Greedy Search Strategy for JMIM

Input: The mixed feature set
Output: The selected feature set

Begin:
  $F \leftarrow$ The mixed feature set
  $S \leftarrow$ Empty set
For $\forall f_i \in F$ calculate $I(C; f_i)$;
  Search for a feature $f_i$ that has the highest mutual information with label, $F \leftarrow F \backslash \{f_i\}$; $S \leftarrow \{f_i\}$.
  For $1 < i \leq k$:
    $f_i = \arg \max_{f_i \in F - S} (\min_{f_S \in S} I(f_i, f_S; C))$;
    $F \leftarrow F \backslash \{f_i\}$;
    $S \leftarrow S \cup \{f_i\}$.
End for
End for
Return $S$
End

the lightgbm algorithm makes several improvements to the GBDT algorithm. There are two main improvement mechanisms. 1) Integrate the histogram algorithm to process features (histogram); 2) Utilize the deep growth strategy (Leaf-wise) to control the model complexity. The prediction accuracy of lightgbm is higher than the traditional GBDT algorithm, and the time consumption is shorter [28].

(1) Histogram algorithm: The histogram algorithm discretizes the continuous floating feature-values into $K$ bins, and then build a histogram with the width of $K$. The building method of the histogram is as shown in Figure 7. When the decision tree traversing, it uses the discrete feature-value as an index of the cumulative statistics measure. Then it finds the optimal splitting point according to the discrete value to traverse. The histogram algorithm can reduce the memory consumption and decrease training time effectively [29].

(2) Deep growth strategy: Lightgbm traverses all leaf nodes before the feature splitting, then uses the leaf with the largest gain to split. The algorithm repeats the above steps until all leaves finish splitting. Compared to the classifier with traditional level growth strategy, the lightgbm which utilizes the deep growth strategy gets higher accuracy under the same times of splits. At the same time, lightgbm adds the maximum depth limit when the model grows, which prevents the model from overfitting. The deep growth strategy is as shown in Figure 8, where the white nodes represent the nodes with the largest gain.

3) FOCAL-LOSS-LIGHTGBM(FLGB)
If the detection precision of the model can’t satisfy the requirements of the power grid, it is not conducive to process the faults cause by attacks in time. When FDIA happens, the measurement data of the physical layer often contains data with anomalous distributed, these abnormal
samples are difficult to classify by the classifier. To deal with difficult-classification samples, the focus loss function is used to improve the lightgbm classifier, which assigns a higher weight to difficult-classification samples [30].

For the traditional lightgbm classifier, the multi-classification loss function is the cross-entropy loss function. As is shown in equation (11), where \( p_i \) denotes the probability that the sample is easy to classify, \( y_i \) represents the actual label of the data, \( T \) represents the number of classes.

\[
L(p_i, y_i) = - \sum_{i=1}^{T} y_i \log p_i
\]  

(11)

The multi-classification focus-loss function is as shown in equation (12):

\[
L_{\text{focal-loss}} = \sum_{i=1}^{T} (1 - p_i)^\gamma L(p_i, y_i)
\]  

(12)

In the iteration process of focal-loss-lightgbm, when a sample is misclassified, the value of \( p_i \) is small, the regulatory factor \((1 - p_i)\) is equal to 1 approximately, and the loss is not affected. When a sample is easy to classify, the regulatory factor \((1 - p_i)\) is equal to 0 approximately, so the weight of easy-classification samples is reduced. The hyper-parameter \( \gamma \) adjusts the proportion of lower weight smoothly, which belongs to the easy-classification samples. The effect of the regulatory factor can be enhanced by increasing the hyper-parameter \( \gamma \). Based on the above analysis, the focus loss function decreases the weight of easy-classification samples and increases the weight of difficult-classification samples. Therefore, the FLGB classifier pays more attention to difficult-classification samples when training the classifier, which further improves the accuracy of FDIA detection of CPPS.

IV. EXPERIMENTAL ANALYSIS

In this section, we evaluate the performance of our proposed CKS-FCS-FLGB model for FDIA detection. The dataset is from the public Google power system cyber-attack dataset [31]. All classification experiments utilized the offline training method.

A. DATASETS

The dataset is collected by the Oak Ridge National Laboratory. The topology of the power system is shown in Figure 9. L1 and L2 are transmission lines, P1 and P2 are three-phase generators. R1~R4 are four intelligent electronic devices that control the on/off of the four circuit breakers DR1~DR4. The PDC is used to process vector data, and it is responsible for storing asynchronous data and displaying historical data. Besides, the intelligent electronic devices use distance protection mechanism, the circuit breaker opens when a fault occurs. This micro power system contains the characteristic of CPPS, where the FDIA happens in the cyber layer and causes the fault of physical devices.

FIGURE 9. The topology of the power system.

The dataset consists of a total of five classes of samples after processing, which includes normal operation and the line maintenance status of the power system, power system fault status, measurement data tampering attack, remote trip instruction tampering attack, and relay setting tampering attack. These five scenarios are described as the following descriptions:

Class-0: The normal operation and the line maintenance of the power system.

Class-1: The fault status: The small current ground fault takes place in the power system.

Class-2: Measurement data tampering attack: The attacker injects false data from the data acquisition channel that evades the bad data detection mechanisms, then the control center makes an incorrect state estimation of the power system, which makes the physical grid operate unstably.

Class-3: Remote trip instruction tampering attack: The attacker tampers with the trip command of the relay protection device and operates the controller directly.

Class-4: Relay setting tampering attack: The attacker tampers with the relevant settings of the relay so that the controllers cannot switch off in time.

Each sub-dataset contains 128 features, which includes the data collected by 4 Phasor Measurement Units (PMUs), the snort alarms, and system logs. Each PMU measures 29 variables and the remaining 12 features are the log information of the control panel. The feature names and the descriptions are shown in Table 1. Among them, the symbol “#” in symbol “R#” means the index of PMU. For example, “R1-PM1” represents the voltage phase magnitude measured by PMU R1.

B. THE PSEUDO-DATA OBTAINING METHOD OF THE POWER DATA

For the original data, the sample statistics are as shown in Figure 10. The missing values in the data are filled using the mean-value method. The statistics show that there is a serious
imbalance disadvantage in each sub-dataset. The sample size of class 4 (Remote trip instruction tampering attack) is much higher than the other four categories of samples. Therefore, the sample size of the other four categories of data is not enough to train the model precisely.

According to the imbalance rate quantitative method proposed in this paper, the imbalance rate of the 15 sub-datasets is calculated and the results are as shown in Table 2. The sub-dataset with the lowest imbalance rate is sub-dataset 2, the imbalance rate is 74.46%. The sub-dataset with the highest imbalance rate is sub-dataset 13, the imbalance rate is 242.22%. Generally speaking, for a multi-classification task, the imbalance rate that is more than 50% would affect the precision of the classifier seriously. In this paper, the imbalance threshold is set to 50%. So all sub-datasets need to oversample to become balanced.

Table 2. The Imbalance rate statistic of each sub-dataset.

| Sub-dataset | Imbalance rate (%) | Sub-dataset | Imbalance rate (%) | Sub-dataset | Imbalance rate (%) |
|-------------|-------------------|-------------|-------------------|-------------|-------------------|
| 1           | 127.28            | 6           | 135.36            | 11          | 139.73            |
| 2           | 74.46             | 7           | 161.75            | 12          | 106.44            |
| 3           | 213.86            | 8           | 110.87            | 13          | 242.22            |
| 4           | 86.75             | 9           | 87.99             | 14          | 138.41            |
| 5           | 131.91            | 10          | 114.98            | 15          | 142.53            |
number of iterations, the training and test set division ratio is 3:1. The default values, the meanings, and the effects of the parameters of FLGB are as shown in Table 4:

The total accuracy, average recall, average precision, and average F1-Score are used as performance indicators to evaluate the FDIA detection model. The effect of each indicator is described as follows. 1) Accuracy: this measures the proportion of all samples that are classified correctly in the test set. 2) Precision: which measures the proportion of positive examples among the test set that are classified as positive examples. 3) Recall: which measures the proportion of positive samples among the test set that are classified correctly. 4) F1-score: the harmonic mean value of precision and recall. The four indicators are utilized to evaluate the FDIA detection model.

In the case of the large-sample, the parameter “num_boost_round” of lightgbm is set to 5000, if the loss of each iteration doesn’t decrease obviously, then the model stops iterating early. The lightgbm classifier performs 3182 iterations, the final multi-log-loss is 0.341296. CKS-FCS-FLGB performs 3708 iterations, the multi-focal-loss is 0.006144, which decreases the loss more obviously than the multi-log-loss. The total accuracy of

| Feature | Description |
|---------|-------------|
| R1:AP   | Phase A active power |
| R1:AQ   | Phase A reactive power |
| R1:BP   | Phase B active power |
| R1:BQ   | Phase B reactive power |
| R1:CP   | Phase C active power |
| R1:CQ   | Phase C reactive power |
| R1:THR-P| 3-phase active power |
| R1:THR-Q| 3-phase reactive power |
TABLE 4. Default parameter of FLGB.

| Parameter          | Meaning                  | Default values | Effect                      |
|--------------------|--------------------------|----------------|-----------------------------|
| learning_rate      | Learning rate            | 0.1            | Reduce the weight of each value |
| max_depth          | Maximum depth of the tree| -1             | Prevent overfitting          |
| num_boost_round    | Number of iterations     | 100            | Improve accuracy            |
| num_leaves         | Number of leaf nodes     | 31             | Prevent overfitting          |
| n_jobs             | Number of parallel threads| -1             | Reduce the training time     |

FIGURE 13. The classification report of the lightgbm and the CKS-FCS-FLGB model on the large-sample dataset.

The traditional lightgbm classifier and the CKS-FCS-FLGB model are 89.26% and 95.46% respectively. The total accuracy of the CKS-FCS-FLGB model is 5.76% higher than lightgbm classifier. Due to the imbalance problem of the original data, the total accuracy of lightgbm classifier is lower. To reflect the performance of the classifier clearly, the classification report of lightgbm classifier and the CKS-FCS-FLGB model is shown in Figure 13. The average precision, average recall, and average F1-Score of lightgbm classifier are 89.71%, 85.93%, and 87.67% respectively. The average detection precision, average recall, and average F1-Score of the CKS-FCS-FLGB model are 95.23%, 95.97%, and 95.36%. Compared to the traditional lightgbm classifier, these performances of the FDIA detection model are improved by 5.52%, 10.04%, and 7.69% respectively.

In the case of the large-sample, the CKS-FCS-FLGB model proposed in this paper not only improves the detection accuracy of various FDIA significantly, but also reduces the false alarm rate of the FDIA detection model.

Avoiding overfitting is a key task of a machine learning classifier. The characteristic of an overfitting model is that the model obtains outstanding learning ability on the training set.
However, the prediction error of unknown samples is too high. The overfitting model limits the detection performance on unknown data, so the performance is not reliable for FDIA detection. To check whether the CKS-FCS-FLGB model is overfitting, in the case of small-sample data, the learning curve of the model is as shown in Figure 16:

In Figure 16, as the number of training samples increases, the training score and cross-validation score converge to a line gradually. If enough representative samples are used for training the CKS-FCS-FLGB model, then it can detect the unknown FDIA in the test set well, which means this model is reliable for FDIA detection.

To further verify the performance of the CKS-FCS-FLGB model in the case of a small-sample dataset, the performance comparison of the CKS-FCS-FLGB model is compared to lightgbm classifier on each sub-dataset, the comparison results are shown in Figure 17.

As is shown in Figure 17, the total accuracy of the CKS-FCS-FLGB model is 96.78%, the average recall is 96.78%, the average precision is 96.70%, and the average
F1-Score is 96.79% for FDIA detection. When a transient process occurs in the power system, the false alarm rate is only 2%, so the CKS-FCS-FLGB model will not cause a high false alarm of FDIA when the fluctuations happen in the power system. Further analysis, the CKS-FCS-FLGB model distinguishes the FDIA and fluctuations effectively.

2) THE ANTI-NOISE PERFORMANCE OF CKS-FCS-FLGB MODEL

As the regulations by the IEEE C37.118 standard, the measurement error of PMU should be lower than 1% [32]. When the signal-to-noise ratio (SNR) of the data is 20dB, it means that the data contain 1% of noise. To verify the anti-noise performance of the CKS-FCS-FLGB model, we added Gaussian White Noise with the SNR of 40, 30, 20, and 15dB into the data, which is used to simulate the measurement error of the wide-area measurement system. The corresponding measurement errors to each SNR are 0.01%, 0.1%, 1%, and 3%. The anti-noise performance of the CKS-FCS-FLGB model is shown in Figure 18:

In Figure 18, it can be deduced that as the noise intensity increases, the performance of the CKS-FCS-FLGB model does not decrease sharply. Through further analysis, the CKS-FCS-FLGB model is not sensitive to the impact of noise. The CKS-FCS-FLGB model has a stable FDIA detection performances when the measurement data contains serious noise.

3) THE COMPARISON OF CKS-FCS-FLGB AND OTHER ALGORITHMS

In this section, the CKS-FCS-FLGB is compared with other machine learning algorithms on large-sample data. These algorithms include the ensemble learning algorithms and other classical machine learning algorithms. The ensemble learning algorithms mainly include Extreme Gradient Boosting (Xgboost), GBDT, Random Forest (RF), Adaptive Boosting (Adaboost), and Voting Classifier (VC). The classical machine learning algorithms include Decision Tree (DT), Artificial Neural Network (ANN), Support Vector Machine (SVM), Logistic Regression (LR), and K-Nearest Neighbor (KNN). The performance comparisons of these algorithms are as shown in Figure 19:

Under the default parameters, the total accuracy, average precision, average recall, and average F1-Score of the CKS-FCS-FLGB model are 95.46%, 95.23%, 95.97%, and 95.36%, which are 27.25%, 25.42%, 38.92%, and 38.56% higher than other ensemble learning algorithms. At the same time, these performances of the CKS-FCS-FLGB
model are 36.39%, 37.12%, 53.55%, and 55.98% higher than other classic machine learning algorithms respectively.

In summary, compared to other algorithms, the CKS-FCS-FLGB model obtains higher detection precision of different types of FDIAs, which is more appropriate to detect the FDIA against CPPS.

V. CONCLUSION

This paper researches the potential invasion pathway of the FDIA against CPPS, and a novel CKS-FCS-FLGB model is further designed to detect different types of FDIAs. We addressed the problem of a high false alarm rate of the model caused by imbalanced data. Then we proposed a C-KSmote oversampling method, which is used to obtain pseudo-data. The pseudo-data is stored into the pseudo-sample database that assists in training and evaluating the model, and updating the model in the future. We extracted the optimal feature set to characterize the behavior of the FDIA, it reduces the dimension of data and improves the training precision of the model. Based on the above work, the FLGB ensemble classifier is constructed to detect the FDIA, which concentrates more on improving the accuracy of difficult-classification samples, and further improves the detection precision of FDIA. We verified the performance of this model both on large-sample and small-sample dataset. The simulations results show that this model can detect different types of FDIA efficiently even on a small-sample dataset. This model also gets great anti-noise performance. Especially, the model obtains great performance of distinguishing the transient process and FDIA, which improves the disturbance resisting ability of CPPS. In summary, the CKS-FCS-FLGB FDIA detection model is 36.39%, 37.12%, 53.55%, and 55.98% higher than other classic machine learning algorithms respectively.
model is proposed, which improves the detection precision of FDIA and ensures the stable operation of CPPS.

REFERENCES

[1] S. H. Kim and S. Park, “CPS (cyber physical system) based manufacturing system optimization,” Procedia Comput. Sci., vol. 122, pp. 518–524, Jan. 2017.

[2] E. R. Griffio, C. Greer, and D. A. Wollman, “Framework for cyber-physical systems,” vol. 1, 2017.

[3] Y. Li, Z. Li, and L. Chen, “Dynamic state estimation of generators under cyber attacks,” IEEE Access, vol. 7, 2019, Art. no. 125253.

[4] R. Lai, X. Qiu, and J. Wu, “Robustness of asymmetric cyber-physical power systems against cyber attacks,” IEEE Access, vol. 7, pp. 61342–61352, 2019.

[5] J. Liang, L. Sankar, and O. Kosut, “Vulnerability analysis and consequences of false data injection attack on power system state estimation,” IEEE Trans. Power Syst., vol. 31, no. 5, pp. 3864–3872, Sep. 2016.

[6] 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.

[7] R. Tan, H. H. Nguyen, E. Y. S. Foo, X. Dong, D. K. Y. Yau, Z. Kalbarczyk, R. K. Iyer, and H. B. Gooi, “Optimal false data injection attack against automatic generation control in power grids,” in Proc. ACM/IEEE 7th Int. Conf. Cyber-Phys. Syst. (ICCPS), Apr. 2016, pp. 1–10.

[8] W. Dong, Z. Yang, P. Xun, P. Zhu, and B. Wang, “Advanced bad data injection attack and its migration in cyber-physical systems,” Electronics, vol. 8, no. 9, pp. 941, Aug. 2019.

[9] Z. Hu, Y. Wang, X. Tian, X. Yang, D. Meng, and R. Fan, “False data injection attacks identification for smart grids,” in Proc. 3rd Int. Conf. Technol. Adv. Elect., Electron. Comput. Eng. (TAEECE), Apr. 2015, pp. 139–143.

[10] R. B. Bobba, K. M. Rogers, and Q. Wang, “Detecting false data injection attacks on DC state estimation,” in Proc. 1st Workshop Secure Control Syst. (CPSWEEK), 2010.

[11] Y. Chakhchoukh and H. Ishii, “Enhancing robustness to cyber-attacks in power systems through multiple least trimmed squares state estimations,” IEEE Trans. Power Syst., vol. 31, no. 6, pp. 4395–4405, Nov. 2016.

[12] Y. Gu, T. Liu, D. Wang, X. Guan, and Z. Xu, “Bad data detection method for smart grids based on distributed state estimation,” in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2013, pp. 4483–4487.

[13] Y. Guan and X. Ge, “Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks,” IEEE Trans. Signal Inf. Process. Netw., vol. 4, no. 1, pp. 48–59, Mar. 2018.

[14] O. Ahmed, X. Guan, T. Liu, Y. Gu, C. Shen, and Z. Xu, “Extended distributed state estimation: A detection method against tolerable false data injection attacks in smart grids,” Energies, vol. 7, no. 3, pp. 1517–1538, Mar. 2014.

[15] 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.

[16] S. Li, Y. Yiilmaz, and X. Wang, “Quickest detection of false data injection attack in wide-area smart grids,” IEEE Trans. Smart Grid, vol. 6, no. 6, pp. 2725–2735, Nov. 2015.

[17] H. M. Khalid and J. C.-H. Peng, “Immunity toward data-injection attacks using multisensor track fusion-based model prediction,” IEEE Trans. Smart Grid, vol. 8, no. 2, pp. 697–707, Mar. 2017.

[18] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, “Feature selection-based detection of covert cyber deception assaults in smart grid communications networks using machine learning,” IEEE Access, vol. 6, pp. 27518–27529, 2018.

[19] D. Xue, X. Jing, and H. Liu, “Detection of false data injection attacks in smart grid utilizing ELMBased OCON framework,” IEEE Access, vol. 7, pp. 31762–31773, 2019.

[20] M. Mohammaddourfard, A. Sami, and A. R. Seifi, “A statistical unsupervised method against false data injection attacks: A visualization-based approach,” Expert Syst. Appl., vol. 84, pp. 242–261, Oct. 2017.

[21] J. J. Q. Yu, Y. Hou, and V. O. K. Li, “Online false data injection attack detection with wavelet transform and deep neural networks,” IEEE Trans. Ind. Informat., vol. 14, no. 7, pp. 3271–3280, Jul. 2018.

[22] D. Wang, X. Wang, Y. Zhang, and L. Jin, “Detection of power grid disturbances and cyber-attacks based on machine learning,” J. Inf. Secur. Appl., vol. 46, pp. 42–52, Jun. 2019.

[23] Z. J. Zhou and J. Feng, “Deep forest: Towards an alternative to deep neural networks,” in Proc. IJCAI, vol. 17, 2017, pp. 3553–3559.

[24] Y. He, G. J. Mendis, and J. Wei, “Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism,” IEEE Trans. Smart Grid, vol. 8, no. 3, pp. 2505–2516, Sep. 2017.

[25] A. S. Musleh, G. Chen, and Z. Y. Dong, “A survey on the detection algorithms for false data injection attacks in smart grids,” IEEE Trans. Smart Grid, vol. 11, no. 3, pp. 2218–2234, May 2020.

[26] F. Last, G. Douzas, and F. Bacao, “Oversampling for imbalanced learning based on K-means and SMOTE,” 2017, arXiv:1711.00837. [Online]. Available: http://arxiv.org/abs/1711.00837

[27] M. Bennasar, Y. Hicks, and R. Setchi, “Feature selection using joint mutual information maximisation,” Expert Syst. Appl., vol. 42, no. 22, pp. 8520–8532, 2015.

[28] Y. Ju, G. Sun, Q. Chen, M. Zhang, H. Zhu, and M. U. Rehman, “A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting,” IEEE Access, vol. 7, pp. 28309–28318, 2019.

[29] G. Ke, Q. Meng, and T. Finley, “Lightgbm: A highly efficient gradient boosting decision tree,” in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 3146–3154.

[30] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal loss for dense object detection,” in Proc. IEEE Int. Conf. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 2980–2988.

[31] Datasets. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris/uah-ics-data-sets

[32] M. He, V. Vittal, and J. Zhang, “Online dynamic security assessment with missing PMU measurements: A data mining approach,” IEEE Trans. Power Syst., vol. 28, no. 2, pp. 1969–1977, May 2013.

JIE CAO received the Ph.D. degree in computer science and technology from Jilin University, Changchun, China, in 2017. She is currently an Associate Professor and also a Master’s Tutor with the School of Computer Science, Northeast Electric Power University, Jilin. Her research interests include computer network, machine learning, and power grid stability and security.

DA WANG received the B.S. degree from the School of Electrical Engineering, Northeast Electric Power University, Jilin, China, in 2014, where he is currently pursuing the M.S. degree with the School of Computer Engineering. His research interests include computer networks and power cyber-physical system security.

ZHAOYANG QU (Member, IEEE) received the Ph.D. degree in electrical engineering from China Northeast Electric Power University, in 2010, and the M.S. degree from the Dalian University of Technology, in 1988. He is currently a Professor and also a Doctoral Tutor with the School of Information Engineering, Northeast Electric Power University. He is also the Vice President of the Jilin Province Image and Graphics Society, the Head of the Power Big Data Intelligent Processing Engineering Technology Research Center, and the Jilin Governor Baishan Scholar. His interests include network technology, smart grid and power information processing, and virtual reality. He has published more than 46 articles in SCI/EI international conference proceedings and journals. He is a member of the China Electric Engineering Society Power Information Committee. He received the Second Prize of Jilin Province Science and Technology Progress Award. He is also a top-notch innovative talent in Jilin Province and a young and middle-aged professional and technical talent with outstanding contributions. He presided over the completion of two national natural science funds.
MINGSHI CUI received the B.S. degree from North China Electric Power University, in 2010. He is currently an Engineer with State Grid Inner Mongolia Eastern Electric Power Company. His research interest is information processing in the smart grid.

PENGCHENG XU received the B.S. degree from the Changchun University of Technology. He is currently a Senior Engineer with State Grid Jilin Province Electric Power Supply Company.

KAI XUE received the M.S. degree from Northeast Normal University. He is currently a Senior Engineer with State Grid Jilin Power Company, Ltd., Training Center. His research interest is power system automation.

KEWEI HU was born in May 1968. He received the master’s degree in power system and automation from Northeast Electric Power University. He is currently working with the Power Dispatching Control Center, State Grid Jilin Electric Power Company, Ltd. His research direction is power system dispatching automation.

* * *