sciendo

JAISCR, 2023, Vol. 13, No. 4, pp. 229 – 245
10.2478/jaiscr-2023-0017

FAST ATTACK DETECTION METHOD FOR IMBALANCED DATA IN INDUSTRIAL CYBER-PHYSICAL SYSTEMS

Meng Huang[1,2], Tao Li[1], Beibei Li[1,*], Nian Zhang[3], Hanyuan Huang[1]

[1]*School of Cyber Science and Engineering, Sichuan University, Chengdu, 610065, China*

[2]*College of Computer Science and Engineering, Chongqing Three Gorges University, Wanzhou, Chongqing, 404120, China*

[3]*Departmnt of Electrical and Computer Engineering, University of the District of Columbia, Washington, DC 20008, USA*

[*]*E-mail: libeibei@scu.edu.cn*

Submitted: 5th May 2023; Accepted: 11th September 2023

Abstract

Integrating industrial cyber-physical systems (ICPSs) with modern information technologies (5G, artificial intelligence, and big data analytics) has led to the development of industrial intelligence. Still, it has increased the vulnerability of such systems regarding cybersecurity. Traditional network intrusion detection methods for ICPSs are limited in identifying minority attack categories and suffer from high time complexity. To address these issues, this paper proposes a network intrusion detection scheme, which includes an information-theoretic hybrid feature selection method to reduce data dimensionality and the ALLKNN-LightGBM intrusion detection framework. Experimental results on three industrial datasets demonstrate that the proposed method outperforms four mainstream machine learning methods and other advanced intrusion detection techniques regarding accuracy, F-score, and run time complexity.

Keywords: intrusion detection system, industrial cyber-physical Systems, imbalanced data, all k-nearest neighbor, LightGBM.

1 Introduction

Industrial cyber-physical systems (ICPSs) integrate the physical and information worlds in the industrial field and build a controlled, trusted, scalable, secure, and efficient system [1]. These systems make the industrial field more intelligent, improve industrial productivity, and promote the progress of human industrial technology. With the development of technologies, such as Made in China 2025 [2], Industry 4.0 [3], Industrial Internet of Things [4], and big data technology [5, 6], applying ICPS has been accelerated in numerous industrial domains, such as manufacturing, energy, and medicine.

Although the benefits of ICPSs are obvious, there are also risks. Indeed, ICPSs shift the traditional industrial scene from a "closed" environment to an interconnected network, thereby increasing the attack risk in the industrial field. Additionally,

cybersecurity problems suffer from greater susceptibility ranging from information leakage to physical equipment damage. In recent years, typical network attack incidents include the attack on Iran's nuclear facilities by the Stuxnet virus [7] and unauthorized intrusion into the Maroochy sewage treatment plant in Australia [8]. These incidents indicate that ICPSs will continue to be targets of interest for attackers in the near future. In the NIST ICS Security Guide, the US Department of Commerce highlights the importance of cybersecurity to the secure and dependable functioning of modern industrial operation processes [9]. Therefore, the network security of ICPSs cannot be ignored, and intrusion detection for ICPSs is crucial.

In recent years, the literature on intrusion detection for ICPSs has been increasing yearly. For example, Wang *et al.* [10] designed a region segmentation-based anomaly detection approach for ICPSs. In order to deal with attacks against networks in industrial control systems, Moustafa *et al.* [11] introduced a deep learning-based anomaly detection system. Zolanvari *et al.* [12] utilized classic machine learning approaches, such as random forest, support vector machine, and decision tree, to detect network threats in the industrial Internet of Things. Besides, Ma *et al.* [13] proposed a layered and distributed detection approach based on the system structure and attack categories of each tier of the ICPSs to provide them security protection. Yu *et al.* [14] suggested R-print, a threat detection fingerprinting method based on system residuals for network attacks at the system control layer in ICPSs. Chang *et al.* [15] introduced a forensic-based deep learning method (named Deep-IFS) for industrial Internet-of-Things network detection. Awotunde *et al.* [16] developed a deep learning-based attack detection method for industrial Internet of Things applications that combines deep feedforward neural network with rule-based feature selection. Sampalli *et al.* [17] developed a feature selection and majority voting integration strategy based on RFE-XGBoost (recursive feature elimination-extreme gradient enhancement) for detecting attacks for power grids relying on SCADA systems. Lu *et al.* [18] introduced a new self-study spatial distribution method and a hybrid cognitive computing-based boundary SMOTE and random forest-based intrusion detection method. Li *et al.* [19] suggested a knowledge distillation model based on triadic convolutional neural networks to reduce the time complexity in ICPSs anomaly detection. Shi *et al.* [20] proposed a class of generalized learning system (OCBLS) and stacked OCBLS (ST-OCBLS) detection method for unknown attacks on worker ICPSs. Additionally, Yang *et al.* [21] suggested a mixed statistical-machine learning method for industrial control network anomalous events. This method combines the dynamic threshold method of the seasonal autoregressive moving average (SARIMA) with the long short-term memory (LSTM).

Unfortunately, most machine learning methods for ICPSs intrusion detection ignore data imbalance or do not focus on feature lightweight for high-dimensional data. Data imbalance affects the detection performance of machine learning methods, especially for attack categories with few samples. Besides, the high dimensionality of data affects the computational burden and detection performance of machine learning methods.

To bridge this gap, we propose a novel intrusion detection scheme that comprises feature selection based on information theory and an intrusion detection framework based on ALLKNN-LightGBM. Specifically, first, we propose a hybrid feature selection method based on Joint Mutual Information (JMI), Conditional Mutual Information Maximization (CMIM), and Double Input Symmetrical Relevance (DISR) to reduce data dimensionality. Second, we use the ALLKNN undersampling technique to adjust the imbalanced samples, and then we use LightGBM to train and test the data. The experimental validation is performed on the gas pipeline, SWaT, and ToN_IoT datasets. The main contributions of our work are summarized as follows:

First, we propose a hybrid feature selection method based on information theory. This strategy reduces feature redundancy and obtains a more robust feature subset.

Second, we analyze the imbalanced characteristics of three mainstream network traffic datasets in ICPSs. Additionally, we use the ALLKNN undersampling technique to reduce the samples and balance the dataset, which facilitates the training and detection of the subsequent detection model.

Third, we utilize the LightGBM to quickly train and test the data after processing the imbalance. The simulation experiments on three mainstream ICPS intrusion detection datasets prove that the proposed method has high efficiency in detection accuracy, F-score, and detection time.

The remainder of this paper is organized as follows. Section II explains the basic principles of industrial information physical systems and imbalanced data. Section III describes the proposed architecture, and Section IV evaluates our method's performance. Finally, Section V concludes this work.

2 Preliminaries

2.1 Introduction of industrial cyber-physical systems

A cyber-physical system (CPS) [22, 23] is a multidimensional complex system that integrates the physical and information spaces. Through organic fusion and deep collaboration of computer, communication, and control technologies, it achieves state awareness, data collection, data processing and analysis, scientific decision-making, and precise control. CPS realizes an integrated design of computer, communication, and physical systems, improving the system's reliability, efficiency, and real-time collaboration ability and thus has a wide range of application prospects.

ICPSs usually refer to CPSs in industrial environments, such as smart grids [24], automated water treatment systems [25], and gas pipeline systems [26]. ICPSs are closely linked to existing industrial control systems (ICSs), IoT, and wireless sensor networks. ICPSs enable advanced smart manufacturing and smart services by encapsulating new information technology, such as software-defined networking (SDN), 5G mobile and wireless communications, cloud computing and services, and big data analytics.

Figure 1 illustrates the general framework of ICPSs, which usually involves the physical and information spaces [27]. The physical space, i.e., the physical layer, generally refers to sensors and specific industrial production equipment. In contrast, the information space includes the network, computing and control, application, and security layers. The function of the network layer is to realize communication transmission, and the computing and control layer aims to store, calculate, process and make decisions on information. The application layer is responsible for specific industrial applications, such as smart manufacturing, smart factories, and smart industrial supply chains. The security layer guarantees the security and reliability of ICPSs, including firewalls, intrusion detection systems (IDSs), intrusion prevention systems (IPSs), security certification, and other security measures.

2.2 Imbalanced data

Imbalanced data mainly refers to samples of one category in the original data that are significantly higher or lower in number than those of other categories [28]. If a category imbalance exists in binary classification data, the dataset has majority and minority categories. The minority category means that the sample size is a small percentage of the total sample size, and the majority category means that the sample size is a large percentage of the total sample size. Each sample in dataset D is denoted by (x,y), with x being the feature and y denoting the label, where $y \in \{0,1\}$. Then, in the binary classification model, y=1 shows that the sample belongs to the minority category, and $y = 0$ shows that the sample belongs to the majority category.

As a result, the minority category M is defined as follows:

$$M = \{(x,y)\,|y = 1\}, (x,y) \in D \tag{1}$$

and the majority category N is defined as:

$$N = \{(x,y)\,|x = 0\}, (x,y) \in D \tag{2}$$

where M and N meet the following constraints:

$$M \cap N = \varnothing \tag{3}$$

$$M \cup N = D \tag{4}$$

The imbalance ratio (IR) is the proportion of the number of samples in the majority category to those in the minority category. It describes the extent of imbalance in different data sets uniformly. IR expressed as:

$$ImbalanceRatio(IR) = \frac{Majority\ category}{Minority\ category} = \frac{N}{M} \tag{5}$$

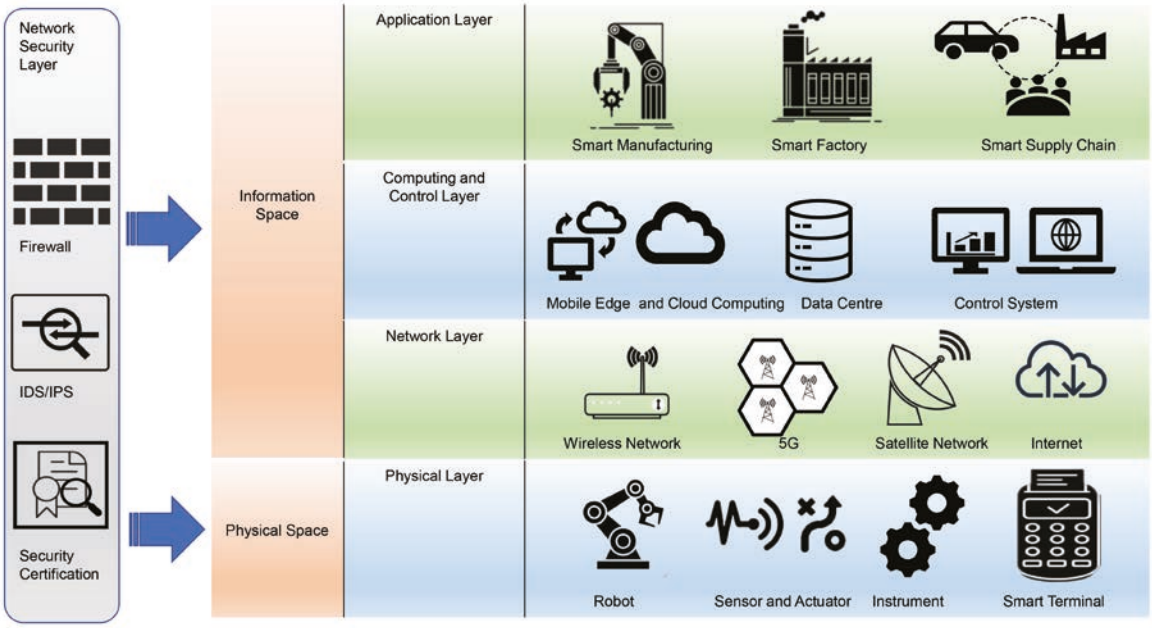


Figure 1. General framework of ICPSs.

Figure 2 depicts a typical example of imbalanced data, where the blue dots indicate the majority category samples in the primary data, which is category A, and the red dots are the minority category samples in the primary data, which is category B. In the figure, the difference in the number of blue and red dots is significant, i.e., category B is the minority category relative to category A.

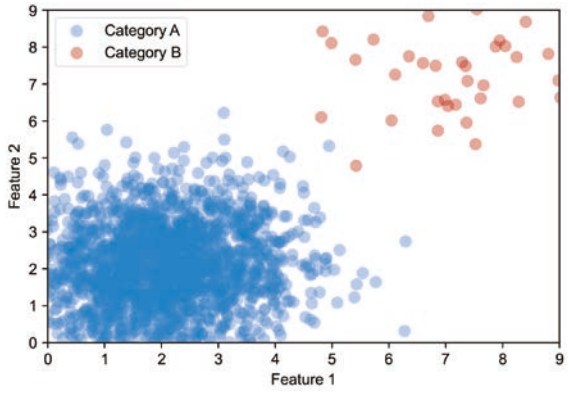


Figure 2. Diagram of imbalanced data.

Data-level approaches, or resampling methods, are used in classification tasks for imbalanced data [29, 30]. These methods modify the sample size in the dataset for each category before employing learning algorithms for training and testing. These approaches improve the classification performance of certain algorithms while reducing the computational overhead during their training.

Ensemble learning approaches are also used in classification tasks for imbalanced data. They focus on combining a data level with it to obtain powerful ensemble classifiers. One approach relies on a specific ensemble learning method, such as the eXtreme gradient boosting tree (XGBoost) [31]. Another category is to embed an alternative imbalance learning method in the ensemble, such as combining SMOTE with Adaboost [32].

3 The proposed method

3.1 The architecture of the proposed method

This paper proposes a network intrusion detection scheme for ICPSs, with Figure 3 presenting the corresponding flowchart.

In Figure 3, the network traffic data in ICPSs includes both normal and malicious data. Extracting features from the network data generally involves high-dimensional characteristics. Hence, to address the high-dimensional data problem, we perform feature selection on the data and propose a hy-

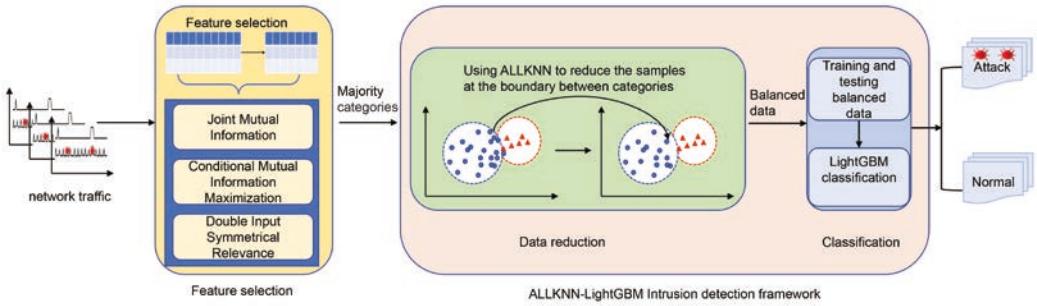


Figure 3. General flowchart of the proposed scheme for ICPSs intrusion detection.

brid feature selection method based on information theory. This method includes Joint Mutual Information (JMI), Conditional Mutual Information Maximization (CMIM), and Double Input Symmetric Relevance (DISR). Then, we design an intrusion detection framework based on ALLKNN-LightGBM. In this framework, we analyze the imbalanced nature of the data and use ALLKNN to reduce the majority sample class. We use LightGBM to train and detect balanced network traffic data in the classification stage.

3.2 Feature selection

Since the high dimensionality of the original imbalanced network traffic data affects the resampling and classification performance, we use an information-theoretic-based feature selection method, which is a filtering method. According to the selection criteria given by Brown [33], we use JMI, CMIM, and DISR to enhance the robustness of the feature selection process. The JMI, CMIM, and DISR methods are described as follows: JMI [34] states the best compromise regarding accuracy, stability, and flexibility. The advantage of JMI is its ability to eliminate redundancy in features. Let X_i, X_j be the i-th and j-th features, respectively. The JMI score value of feature X_i is:

$$J_{jmi}(X_i) = \sum_{X_j \in S} I(X_i X_j; Y), \tag{6}$$

where Y is a label, X_i and X_j are joint random variables, and S is the set of currently selected features.

CMIM [35] is an appealing method for filtering feature selection. CMIM correctly distinguishes between redundant and noisy characteristics, and prioritizes information-rich, irrelevant features [36]. The value of CMIM for each feature X_i is:

$$J_{cmim}(X_i) = \min_{X_j \in S} \left(I(X_i : Y | X_j) \right). \tag{7}$$

DISR [37] is a variant of JMI. DISR considers variable complementarity and the lower bound of mutual information. The DISR method encourages the selection of complementary variables that have already been selected with a higher probability. The feature X_i is calculated as follows:

$$J_{disr}(X_i) = \sum_{X_j \in S} \frac{I(X_i X_j; Y)}{H(X_i X_j Y)}. \tag{8}$$

3.3 Data reduction using the ALLKNN method

The Edited Nearest Neighbor Algorithm (ENN) [38] uses the nearest neighbor algorithm to edit the data, find samples that are noisy or close to the boundary, and remove them while keeping those that belong to the same category of nearest neighbor samples. ALLKNN [39] is an undersampling method and an extension of the ENN. ALLKNN mainly uses the nearest neighbor algorithm (KNN) to modify the dataset by removing samples that are inconsistent with their nearest neighbors. This method calculates the nearest neighbors for each sample in the undersampled categories and removes the samples if they do not satisfy the selection criterion iteratively. Ultimately, the majority and minority categories of the samples tend to balance. The Algorithm 1 based on ALLKNN is as follows.

Algorithm 1 Undersampling algorithm based on ALLKNN

Input: D:the set of samples after feature selection; k:the value of the nearest neighbor.

Output: S: sample subset.

```
j = 1
mark(x) = 1
while j ≤ k do
    NN(j,x) =Solve for the j nearest neighbors of x
    if the majority categories of NN(j,x) are incorrectly classified then
        mark(x) = 0
    end if
    j++
end while
```

3.4 Classification using LightGBM

Ke *et al*. [40] proposed LightGBM, which has been extensively utilized in machine learning. LightGBM is an iteratively trained gradient enhancement system based on a decision tree algorithm. Specifically, LightGBM mainly employs two new algorithms: gradient-based one-sided sampling (GOSS) and mutually exclusive feature bundling (EFB). LightGBM has advantages such as being fast, distributed, and high performance. It improves the classic GBDT algorithm by employing a Histogram-based decision tree algorithm, Gradient-based One-Side Sampling (GOSS), Exclusive Feature Bundling (EFB), and leaf-wise algorithm, and supports efficient parallelism. In industrial practice, LightGBM is faster to train and consumes less memory than other algorithms while maintaining high accuracy.

Besides, LightGBM employs GOSS to reduce small gradient samples while retaining the large ones. The GOSS algorithm mainly trains samples with larger training errors. For GBDT models, each node is usually split where the information gain is greatest, with the information gain defined as:

$$V_{j|O}(d) = \frac{1}{n_O}\left(\frac{\left(\sum_{\{x_i\in O:x_{ij}\le d\}} g_i\right)^2}{n^j_{l|O}(d)} + \frac{\left(\sum_{\{x_i\in O:x_{ij}>d\}} g_i\right)^2}{n^j_{r|O}(d)}\right), \tag{9}$$

where O is the training dataset, g is the negative gradient of the loss function, and x is the sample. $n_O = \sum I[x_i \in O]$, $n^j_{l|O}(d) = \sum I[x_i \in O : x_{ij} \le d]$and $n^j_{r|O}(d) = \sum I[x_i \in O : x_{ij} > d]$.

In the GOSS algorithm, the samples are first sorted in descending order based on the absolute value of their gradients. Next, a sample subset A is obtained by retaining the top-a100% instances with larger gradients, and the remaining subset is denoted as A^c. A subset B is randomly sampled from A^c. Finally, the GOSS algorithm splits the samples based on the estimated variance gain, calculated as follows:

$$\tilde{V}_j(d) = \frac{1}{n}\left(\frac{\left(\sum_{x_i\in A_l} g_i + \frac{1-a}{b}\sum_{x_i\in B_l} g_i\right)^2}{n^j_l(d)} + \frac{\left(\sum_{X_i\in A_r} g_i + \frac{1-a}{b}\sum_{X_i\in B_r} g_i\right)^2}{n^j_r(d)}\right), \tag{10}$$

where $A_l = \{x_i \in A : x_{ij} \le d\}$, $A_r = \{x_i \in A : x_{ij} > d\}$, $B_l = \{x_i \in B : x_{ij} \le d\}$, $B_r = \{x_i \in B : x_{ij} > d\}$ is a coefficient. During training, the Exclusive Feature Bundling (EFB) technique solves the sparsity problem of high-dimensional data. EFB reduces the feature dimension of data without losing information, avoids unnecessary calculations of zero values, and improves the algorithm's execution speed. Specifically, first, EFB creates a graph containing weighted edges, where the weight corresponds to the total conflict between features. Second, it sorts the features in descending order according to the degree of the graph. Finally, it examines each feature in the ranked list and assigns it to an existing bundle with minor conflicts or produces a new one. LightGBM adopts a histogram algorithm, which buckets the original data of features. This strategy reduces the model's complexity, as only discrete values of the feature "bucket" are saved after the feature is "bucketed", significantly reducing memory usage and improving the model's efficiency during training and prediction. Algorithm 2 is a histogram-based method adopting LightGBM, which is presented below.

The Leaf-wise algorithm adopts a more efficient calculation strategy, where the algorithm splits from the leaf with the maximum split gain each time and iterates to improve accuracy. Leaf-wise suffers from producing deeper decision trees that are prone to overfitting. Therefore, LightGBM enhances the decision tree's maximum depth limitation to maintain efficiency while preventing algorithm overfitting.

Algorithm 2 Histogram-based algorithm

Input: T: training data in the experiment; d:max tree depth; F:feature dimension of data.

```
nodeSet = [0]
rowSet =[0,1,2,...]
for n = 1 to d do
    for node in nodeSet do
        usedRows = rowSet[node]
        for i = 1 to m do
            H =new Histogram()
            for j in usedRows do
                bin =I.f[i][j].bin
                H[bin].y = H[bin].y + I.y[j]
                H[bin].n = H[bin].n + 1
            end for
            Determine the best split point for histogram H
        end for
        RowSet and nodeSet are updated using the optimal split points
    end for
end for
```

After undersampling the samples in the datasets using the ALLKNN, we use the LightGBM for training and detection. Regarding LightGBM, this paper uses the gradient boosting decision Tree (GBDT) as the weak classifier. LightGBM is a decision tree-based ensemble method, while gradient Boosting is an algorithm belonging to the boosting ensemble family, which iterates new learners through gradient descent. Assuming a GBDT model comprising K categorical regression trees, the model's detection results are expressed as follows:

$$y_m(x) = \sum_{k=1}^{K} T(x;\omega). \tag{11}$$

where $T(x;\omega)$ denotes the decision tree, denotes the parameter of the decision tree, and K is the number of decision trees. During training, $y_0(x) = 0$ is the first set, and the m-th step is the model representation presented below:

$$y_m(x) = y_{m-1}(x) + T(x;\omega), \tag{12}$$

$$\hat{\omega} = \arg min \sum_{i=1}^{N} L\left(y_{i,} y_{m-1}\left(x_i\right) + T\left(x_i;\omega\right)\right). \tag{13}$$

where $y_{m-1}(x)$ is the current decision tree model, $\hat{\omega}$ is the model parameter for the next decision tree, and $L(\varphi)$ represents the loss function of the model.

4 Experimental results and performance analysis

This section comprehensively evaluates the performance of our proposed intrusion detection scheme. Sections 4.1-4.3 describe the performance metrics used in the experiments, the experimental parameter settings, the data resources, and their imbalance characteristics. Section 4.4 describes the unbalanced processing of experimental data using the ALLKNN method. Section 4.5 challenges LightGBM with 8 mainstream machine learning methods. Section 4.6 describes the classification performance of LightGBM under different sampling methods. Section 4.7 compares our method with 4 mainstream machine learning methods. Section 4.8 compares the proposed method's performance against other advanced intrusion detection methods.

4.1 Performance indicators

In machine learning classification tasks, overall accuracy is an important performance indicator for classification. However, the overall accuracy can be misleading for unbalanced data classification tasks. Therefore, the performance evaluation metrics for classification tasks involving unbalanced data [44] are overall accuracy, G-measure, and F-score.

This paper uses the overall accuracy, F-score, and run time as the core evaluation metrics. These performance metrics are calculated based on the obfuscation matrix of the attack detection. Table 1 presents the confusion matrix of the attack detection, where TP denotes the number of samples that are true attacks and are predicted by the model as attacks. TN denotes the number of benign samples predicted as benign. FP denotes the number of samples that are benign and are predicted as attacks, and FN denotes the number of samples that are attacks and are predicted as benign.

Table 1. Confusion matrix of attack detection

	Predicted attack Category	Predicted benign Category
Actual attack Category	TP	FN
Actual benign Category	FP	TN

1. The overall accuracy rate represents the proportion of cyberattack classifications that are successfully predicted by the model, divided by the entire data sample:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+TP} \tag{14}$$

2. Recall rate represents the ratio of cyberattacks detected by the model to the real cyberattacks:

$$Recall = \frac{TP}{TP+FN} \tag{15}$$

3. Specificity (true negative rate) represents the classification accuracy of the benign samples in the network:

$$Specificity(True\ negative\ rate) = \frac{TN}{TN+FP} \tag{16}$$

4. Accuracy indicates the model's accuracy in predicting real cyberattacks:

$$Precision = \frac{TP}{TP+FP} \tag{17}$$

5. The F-score represents the summed average of recall and specificity:

$$F-score = 2 * \frac{Recall \times Precision}{Recall + Precision} \tag{18}$$

4.2 Parameter settings

In order to validate the experimental results, two groups of comparative experiments are conducted, and the performance is compared in terms of both learning methods and undersampling. The experiments are validated on three dominant datasets in intrusion detection for ICPSs: gas pipeline, SWaT, and ToN_IoT. Additionally, to assess the method's performance more thoroughly, we compare it against various existing advanced intrusion detection methods. All experiments are conducted in Windows 11 using Python 3.9 on an i5-8250 CPU with 16G ARM. Table 2 reports the primary parameters of LightGBM:

Table 2. Key parameter settings of LightGBM

Parameters		
boosting_type	gbdt	Specify the type of weak learner
num_leaves	31	Maximum leaves from trees for base learners
learning_rate	0.1	Boosting learning rate
n_estimators	50	number of decision trees
max_depth	25	Maximum tree depth of base learners

4.3 Dataset and its imbalance analysis

We select three mainstream benchmark datasets in ICPSs for experimentation: gas pipeline, water treatment (SWaT), and ToN_IoT. The gas pipeline [41] dataset was collected from real production, including normal and attack network behavior data. This dataset involves seven types of attacks: NMRI, CMRI, MSCI, MPCI, MFCI, DoS, and Reconnaissance. The dataset contains 97019 samples and 27 features. The SWaT [42] contains simulated data collected on the water treatment platform for 11 days, including 946722 samples and 51 features. The label categories are normal and attack, and we randomly select 10% of the dataset for the experiment. The ToN_IoT [43] dataset was generated from genuine and large-scale testbed network data provided by the Canberra Network IoT Lab at the University of New South Wales. The dataset has 44 features, where the categories are represented as 0 and 1, where 0 indicates normal, and 1 indicates an attack.

We also randomly chose 10% of the data to be the experimental data. Table 3 reports the statistical information for the three benchmark datasets:

Table 3. Information on the three benchmark datasets utilized in the experiment

Datasets	instances	features	Number of categories
Gas pipeline	97019	27	8
SWaT	94672	51	2
TON_IoT	46104	44	2

As shown in Table 3, we analyze the imbalance of the categories in the three experimental datasets. The normal category is used as the majority category in the three datasets. The gas pipeline

dataset is a multi-category dataset with more normal samples than the other categories, accounting for 63.04% of the total sample size. In the gas pipeline dataset, the imbalance ratio (IR) between the normal category and the other categories are: NMRI is 22.3, CMRI is 3.95, MSCI is 78.20, MPCI is 8.01, MFCI is 106.73, DoS is 33.29, and Reconnaissance is 8.99. The SWaT and TON_IoT datasets include only 2 categories, normal and attack, with the corresponding IR being 7.38 and 4.71, respectively.

Table 4. Imbalance ratios for the categories in the three datasets

Dataset	Category	Imbalance ratio (IR)
Gas pipeline	Normal	-
	NMRI	22.3
	CMRI	3.95
	MSCI	78.2
	MPCI	8.01
	MFCI	106.73
	DoS	33.29
	Reconnaissance	8.99
SWaT	Normal	-
	Attack	7.38
ToN_IoT	Normal	-
	Attack	4.71

4.4 ALLKNN processing imbalanced data

The ALLKNN undersampling method mainly processes the sample points of the majority category. The removed sample points of the majority category are concentrated near the boundary of the category, making the minority category samples near the boundary balance with the majority sample points. This increases the detection accuracy of the forthcoming detection method for the minority categories. We use the ALLKNN undersampling method on all experiments and set the K value of KNN to 3. The undersampled Normal, CMRI, MPCI, Reconnaissance, and NMRI are undersampled in the gas pipeline dataset. Normal is undersampled in the SWaT and ToN_IoT datasets. Table 5 compares the samples before and after sampling for each category in the three datasets.

4.5 Performance comparison of different learning methods

Currently, the main learning methods used in network intrusion detection scenarios for ICPSs can be classified into two categories: classic machine learning methods and deep learning methods. We use nine machine learning methods, which are logistic regression (LR), naive Bayes (NB), k-nearest neighbor (KNN), multilayer perceptron (MLP), support vector machine (SVM), random forest (RF), Adaboost, convolutional neural network (CNN), and LightGBM.

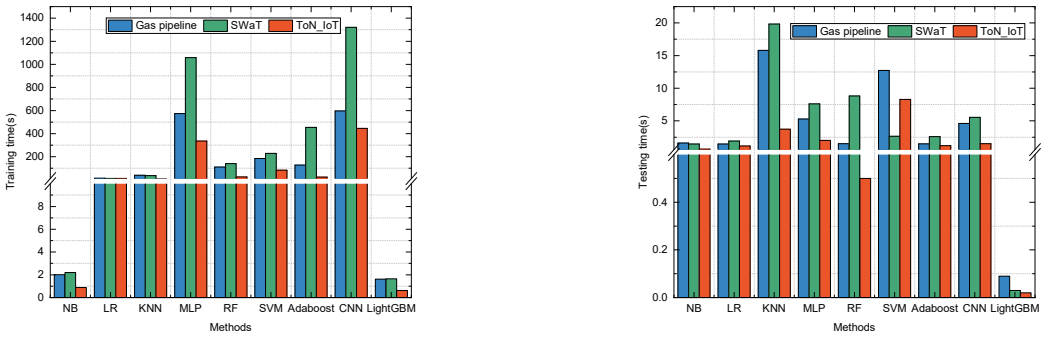
The performance of each learning method on the gas pipeline, SWaT, and ToN_IoT datasets are reported in Tables 6-8. Figure 4 compares the training and testing time for each learning method, and Tables 6-8 present the total time statistics.

Table 6. Performance of various learning methods on the gas pipeline dataset

Methods	Accuracy (%)	F-score (%)	Total time (s)
NB	94.97	93.8	3.63
LR	94.83	93.41	14.96
KNN	94.61	93.62	55.3
MLP	95.03	93.71	579.37
RF	98.89	**98.9**	112.59
SVM	97.73	97.43	197.01
Adaboost	98.84	98.81	129.2
CNN	95.49	94.32	601.3
LightGBM	**99.08**	98.11	**1.71**

Table 5. Comparison of the samples before and after sampling for each category in the three datasets

Data	Categories	Number of original samples	Number of samples after ALLKNN
Gas pipeline	Normal	61156	60450
	CMRI	15466	15411
	MPCI	7637	7235
	Reconnaissance	6805	6786
	NMRI	2763	2566
	DoS	1837	1837
	MSCI	782	782
	MFCI	573	573
SWaT	Normal	83368	82116
	Attack	11304	11304
ToN_IoT	Normal	29958	29808
	Attack	16146	16146



(a) (b)

Figure 4. Training time and testing time comparison of various methods

Table 7. Performance of various learning methods on the SWaT dataset

Methods	Accuracy (%)	F-score (%)	Total time (s)
NB	95.51	95.1	3.8
LR	97.31	97.2	13.92
KNN	99.3	98.35	56.64
MLP	98.97	98.02	1065.8
RF	99.1	**98.95**	148.96
SVM	96.55	96.33	231.2
Adaboost	99.15	98.23	456.8
CNN	96.5	95.2	1325.7
LightGBM	**99.36**	98.54	**1.68**

Table 8. Performance of various learning methods on the ToN_IoT dataset

Methods	Accuracy (%)	F-score (%)	Total time (s)
NB	82.9	79.05	18.71
LR	88.8	89.01	13.52
KNN	99.39	99.4	14.92
MLP	97.61	97.6	338.32
RF	99.93	99.41	25.88
SVM	88.35	88.61	91.2
Adaboost	99.92	99.32	25.24
CNN	98.66	98.7	446.8
LightGBM	**99.94**	**99.56**	**0.64**

Table 6 compares LightGBM with 8 other methods on the gas pipeline dataset, highlighting that LightGBM has the best accuracy and total running time (99.08% and 1.71s, respectively). Although the accuracy values of RF and Adaboost are close to LightGBM, LightGBM is 65 times faster than RF and 75 times faster than Adaboost. Among the F-score indicators, RF is the best, with a value of 98.90%. The F-score value of LightGBM is second only to RF and Adaboost, and its value is 98.11%.

According to Table 7, LightGBM attains the best accuracy and total running time on the SWaT datasets (99.36% and 1.68s, respectively). KNN, RF, and Adaboost are closer in accuracy compared to LightGBM. However, in total time, LightGBM outperforms KNN by 33 times, RF by 88 times, and Adaboost by 272 times. Regarding the F-score metric, RF is the best, with a value of 98.95%. The F-score value of LightGBM is second only to RF, and its value is 98.54%.

According to Table 8, LightGBM has the highest accuracy, F-score, and total running time on the ToN_IoT dataset (99.56%, 99.94%, and 0.64s, respectively), and its performance is closer to that of KNN, RF, and Adaboost. However, in terms of total time, LightGBM is significantly faster than KNN, RF, and Adaboost with a value of 0.64s, which outperforms KNN by 23 times, RF by 40 times, and Adaboost by 39 times.

Figure 4 reveals that LightGBM is faster than the competitor methods during training, requiring 1.62s, 1.65s, and 0.62s on the gas pipeline, SWaT, and ToN_IoT datasets, respectively. Similarly, LightGBM is faster than the competitor methods during testing requiring 0.09s, 0.03s, and 0.02s on the three datasets.

4.6 Performance comparison of different resampling methods

In order to validate the ALLKNN method's performance in coping with data imbalance, this paper compares it with various resampling methods. The comparative methods include oversampling and undersampling methods. The oversampling methods are adaptive synthetic sampling (ADASYN), borderline SMOTE, random oversampling, and synthetic minority over-sampling technique (SMOTE). The undersampling methods include nearMiss, Tomek's links, and random undersampling. The performance of these methods on the three datasets is reported in Table 9.

Table 9 reveals that after being processed by the ALLKNN undersampling method, the detection accuracy of the LightGBM during training and testing is better than the other resampling methods, attaining 99.77%, 99.65%, and 99.96% accuracy, respectively.

4.7 Performance analysis of the proposed method

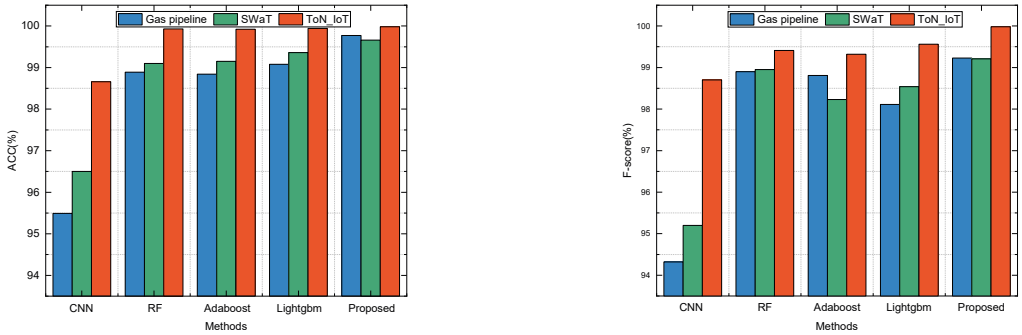
To assess the effectiveness of the learning methods in classification tasks on imbalanced data, solely relying on classification accuracy to evaluate a method's performance is not objective. Hence, this paper employs the F-score and classification accuracy to assess the effectiveness of the methods. The proposed method is contrasted and analyzed with LightGBM, RF, Adaboost, and CNN, and the

Table 9. Accuracy of LightGBM method before and after various data resampling methods

Methods	Gas pipeline	SWaT	TON_IoT
	Accuracy (%)	Accuracy (%)	Accuracy (%)
Original imbalance data	99.08	99.36	99.94
ADASYN	99.18	99.39	99.83
Borderline_SMOTE	99.13	99.25	99.91
SMOTE	98.7	99.3	99.93
Random over-sampling	98.21	98.63	99.1
TomeksLinks	99.45	99.26	99.78
NearMiss	98.14	98.69	99.61
Random under-sampling	98.25	99.05	99.22
ALLKNN	**99.77**	**99.65**	**99.98**

Table 10. Comparative analysis of the proposed method and other methods

Methods	Accuracy (%)			F-score (%)		
	Gas pipeline	SWaT	ToN_IoT	Gas pipeline	SWaT	ToN_IoT
CNN	95.49	96.5	98.66	94.32	95.2	98.7
RF	98.89	99.1	99.93	98.9	98.95	99.41
Adaboost	98.84	99.15	99.92	98.81	98.23	99.32
LightGBM	99.08	99.36	99.94	98.11	98.54	99.56
Proposed	**99.77**	**99.65**	**99.98**	**99.23**	**99.21**	**99.98**



(a) Accuracy (b) F-score

Figure 5. Evaluation of the proposed technique to known methods in terms of accuracy and F-score

Table 11. Comparison results with advanced detection methods on the gas pipeline dataset

Methods	Accuracy (%)	Pecision (%)	Recall (%)	F-score (%)
stacked Long Short Term Memory (LSTM)	92.00	94.00	78.00	85.00
BiSRU	96.23	94.09	97.28	-
GID	97.2	98.00	90.00	93.00
Proposed	**99.77**	**99.34**	**99.13**	**95.39**

Table 12. Comparison results with advanced detection methods on the SWaT dataset

Methods	Accuracy (%)	Pecision (%)	Recall (%)	F-score (%)
DAICS	-	91.85	86.16	88.92
AADS	-	86.6	86.1	86.3
TABOR	-	86.1	78.8	82.3
Proposed	**99.65**	**99.6**	**98.79**	**99.21**

Table 13. Comparison results with advanced detection methods in the ToN_IoT dataset

Methods	Accuracy (%)	Pecision (%)	Recall (%)	F-score (%)
Ensemble-ADS system	99.35	90.54	99.98	95.03
Deep Belief Network (DBN)	97.63	97.63	-	96.65
Proposed	**99.98**	**99.98**	**99.97**	**99.98**

corresponding performance is presented in Table 10 and Figure 5.

As shown in Figure 5, compared with the LightGBM, RF, CNN, and Adaboost methods, the proposed approach provides the best detection accuracy, which is 99.77%, 99.65%, and 99.98%, on the gas pipeline, SWaT, and ToN_IoT datasets, respectively. The proposed approach also performs best regarding the F-score metric, attaining 99.23%, 99.21%, and 99.98% on the gas pipeline, SWaT, and ToN_IoT datasets.

4.8 Performance comparison with advanced detection methods

The developed solution is compared to other advanced methods to analyze further the efficacy of the proposed intrusion detection technique, mainly based on ALLKNN and LightGBM. The comparative results on the three datasets are listed in Tables 11-13.

Table 11 reports the comparative results on the gas pipeline dataset. The comparison methods include stacked Long Short Term Memory (LSTM) [45], BiSRU [46], and GID [47]. Compared with these three methods, the proposed method is optimal in terms of accuracy, recall, precision, and F-score, attaining 99.77%, 99.34%, 99.13%, and 99.23%, respectively.

Table 12 presents the comparative results on the SWaT dataset. The competitor methods include DAICS [48], AADS [49], and TABOR [50]. Compared with these three models, our proposed method is superior in precision, recall, and F-score, attaining 99.60%, 98.79%, and 99.21%, respectively. Additionally, our model's accuracy is 99.65%.

Table 13 lists the comparative results on the TON_IoT dataset. The comparison methods include the Ensemble-ADS system [51] and Deep Belief Network (DBN) [52]. Compared to these methods, our method affords better precision, recall, and F-score, 99.98%, 99.97%, and 99.98%, respectively.

The above comparative results reveal that the proposed intrusion detection method outperforms current advanced intrusion detection methods in terms of accuracy, precision, recall, and F-score.

5 Conclusion

This paper provides a network intrusion detection scheme for ICPSs, which is based on a hybrid feature selection method and the ALLKNN-LightGBM intrusion detection framework. Unlike most existing cyber intrusion detection methods for ICPSs, first, we design an information-theoretic-based feature selection scheme to reduce the data dimensionality. Second, we handle the problem of data categories imbalance on three datasets: gas pipeline, SWaT, and ToN_IoT through the ALLKNN undersampling method. Finally, the LightGBM method deals with the problem of long run time and low detection accuracy. Through experimental verification of the three datasets, compar-

ing LightGBM with eight learning methods proves that LightGBM has a greater detection rate and a lower run time complexity. Additionally, our proposed method is compared with eight other resampling methods and the results show that our method is optimal in terms of accuracy. Moreover, compared with CNN, RF, Adboost, LightGBM, and existing advanced intrusion detection techniques, our method is the best in terms of accuracy and F-score.

This work proposes an intrusion detection scheme against known network attacks in ICPSs. Future work will investigate detecting unknown network attacks in ICPSs.

Acknowledgements

This work is partially supported by the National Key Research and Development Program of China under Grant No. 2020YFB1805400; the National Natural Science Foundation of China under Grant No. 62002248; the Sichuan Youth Science and Technology Innovation Team under Grant No. 2022JDTD0014; the Sichuan Science and Technology Program under Grant No. 2022YFG0193 and No. 2023YFG0113.

References

[1] H. Kayan, M. Nunes, O. Rana, P. Burnap, C. Perera, Cybersecurity of industrial cyber-physical systems: a review, ACM Computing Surveys (CSUR), 54(11s), 2022, 1-35.

[2] F. Tao, Q. Qi, New it driven service-oriented smart manufacturing: framework and characteristics, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(1), 2019, 81-91.

[3] L.D. Xu, E.L. Xu, L. Li, Industry 4.0: state of the art and future trends, International Journal of Production Research, 56(8), 2018, 2941-2962.

[4] A. Corallo, M. Lazoi, M. Lezzi, A. Luperto, Cybersecurity awareness in the context of the Industrial Internet of Things: A systematic literature review, Computers in Industry, 137, 2022, 103614.

[5] Leung M F, Jawaid A, Ip S W, et al, A portfolio recommendation system based on machine learning and big data analytics, Data Science in Finance and Economics, 3(2), 2023, 152-165.

[6] Li C, Chen Y, Shang Y, A review of industrial big data for decision making in intelligent manufacturing, Engineering Science and Technology-an International Journal, 29, 2022, 101021.

[7] D. Kushner, The real story of Stuxnet, IEEE Spectrum, 50(3), 2013, 48-53.

[8] N. Sayfayn and S. Madnick, Cybersafety analysis of the maroochy shire sewage spill, MIT Interdisciplinary Consortium for Improving Critical Infrastructure Cybersecurity, 2017, 1-29.

[9] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, A. Hahn, Guide to industrial control systems (ICS) security. NIST special publication, 800(82), 2011, 16-16.

[10] J. Yang, C. Zhou, S. Yang, H. Xu, B. Hu, Anomaly detection based on zone partition for security protection of industrial cyber-physical systems, IEEE Transactions on Industrial Electronics, 65(5), 2018, 4257-4267.

[11] M. AL-Hawawreh, N. Moustafa, E. Sitnikova, Identification of malicious activities in industrial internet of things based on deep learning models, Journal of Information Security and Applications, 41, 2018, 1-11.

[12] M. Zolanvari, M.A. Teixeira, L. Gupta, K.M. Khan, R. Jain, Machine learning-based network vulnerability analysis of industrial internet of things, IEEE Internet of Things Journal, 6(4), 2019, 6822-6834.

[13] J. Liu, W. Zhang, T. Ma, Z. Tang, Y. Xie, W. Gui, J.P. Niyoyita, Toward security monitoring of industrial cyber-physical systems via hierarchically distributed intrusion detection, Expert Systems with Applications, 158, 2020, 113578.

[14] Z. Hong, C. Yang, L. Yu, R-Print: a system residuals-based fingerprinting for attack detection in industrial cyber-physical systems, IEEE Transactions on Industrial Electronics, 68(11), 2021, 11458-11469.

[15] M. Abdel-Basset, V. Chang, H. Hawash, R.K. Chakrabortty, M. Ryan, Deep-IFS: intrusion detection approach for industrial internet of things traffic in fog Environment, IEEE Transactions on Industrial Informatics. 17(11), 2021, 7704-7715.

[16] J.B. Awotunde, C. Chakraborty, A.E. Adeniyi, Intrusion detection in industrial internet of things network-based on deep learning model with rule-based feature selection, Wireless Communications and Mobile Computing, 2021, 7154587.

[17] D. Upadhyay, J. Manero, M. Zaman, S. Sampalli, Intrusion Detection in SCADA based power grids: recursive feature elimination model with majority vote ensemble algorithm, IEEE Transactions on Network Science and Engineering, 8(3), 2021, 2559-2574.

[18] Y. Gao, J. Chen, H. Miao, B. Song, Y. Lu, W. Pan, Self-learning spatial distribution-based intrusion detection for industrial cyber-physical systems, IEEE Transactions on Computational Social Systems, 9(6), 2022, 1693-1702.

[19] Z. Wang, Z. Li, D. He, S. Chan, A lightweight approach for network intrusion detection in industrial cyber-physical systems based on knowledge distillation and deep metric learning, Expert Systems with Applications, 206, 2022, 117671.

[20] K. Yang, Y. Shi, Z. Yu, Q. Yang, A.K. Sangaiah, H. Zeng, Stacked one-class broad learning system for intrusion detection in industry 4.0, IEEE Transactions on Industrial Informatics, 19(1), 2023, 251-260.

[21] W. Hao, T. Yang, Q. Yang, Hybrid statistical-machine learning for real-time anomaly detection in industrial cyber-physical systems, IEEE Transactions on Automation Science and Engineering, 20(1), 2023, 32-46.

[22] Y. Liu, Y. Peng, B. Wang, S. Yao, Z. Liu, Review on cyber-physical systems, IEEE/CAA Journal of Automatica Sinica, 4(1), 2017, 27-40.

[23] D.G.S. Pivoto, L.F.F. de Almeida, R. da R. Righi, J.J.P.C. Rodrigues, A.B. Lugli, A.M. Alberti, Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: A literature review, Journal of Manufacturing Systems, 58, 2021, 176-192.

[24] R.V. Yohanandhan, R.M. Elavarasan, P. Manoharan, L. Mihet-Popa, Cyber-physical power system (CPPS): a review on modeling, simulation, and analysis with cyber security applications, IEEE Access, 8, 2020, 151019-151064.

[25] R. Qadeer, C. Murguia, C.M. Ahmed, J. Ruths, Multistage Downstream Attack Detection in a Cyber Physical System, In: S.K. Katsikas, F. Cuppens, N. Cuppens, C. Lambrinoudakis, C. Kalloniatis, J. Mylopoulos, A. Antn, S. Gritzalis (Eds.), Computer Security, Springer International Publishing, Cham, 2018, pp. 177-185.

[26] Y. Wadhawan, C. Neuman, Evaluating resilience of gas pipeline systems under cyber-physical attacks: a function-based methodology, in: Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy, Association for Computing Machinery, New York, NY, USA, 2016, pp. 71-80.

[27] J. Yao, X. Xu, X. Liu, MixCPS: Mixed time/event-triggered architecture of cyber-physical systems, Proceedings of the IEEE, 104(5), 2016, 923-937.

[28] Haibo He, E.A. Garcia, Learning from imbalanced data, IEEE Transactions on Knowledge and Data Engineering, 21(9), 2009, 1263-1284.

[29] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, G. Bing, Learning from class-imbalanced data: review of methods and applications, Expert Systems with Applications, 73, 2017, 220-239.

[30] H. Kaur, H.S. Pannu, A.K. Malhi, A systematic review on imbalanced data challenges in machine learning: applications and solutions, ACM Computing Surveys (CSUR), 52(4), 2019, 1-36.

[31] Y.-C. Chang, K.-H. Chang, G.-J. Wu, Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions, Applied Soft Computing, 73, 2018, 914-920.

[32] J. Sun, H. Li, H. Fujita, B. Fu, W. Ai, Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting, Information Fusion, 54, 2020, 128-144.

[33] G. Brown, A. Pocock, M.-J. Zhao, and M. Lujan, Conditional likelihood maximisation: a unifying framework for information theoretic feature selection, The Journal of Machine Learning Research, 13(1), 2012, 27-66.

[34] Yang H, Moody J, Feature selection based on joint mutual information, In: Proceedings of international ICSC symposium on advances in intelligent data analysis. Proceedings of international ICSC symposium on advances in intelligent data analysis. Rochester, NY: Citeseer, 1999, 23.

[35] F. Fleuret, Fast binary feature selection with conditional mutual information, Journal of Machine Learning Research, 5(9), 2004, 1531-1555.

[36] Souza F, Premebida C, Araujo R, High-order conditional mutual information maximization for dealing with high-order dependencies in feature selection, Pattern Recognition, 131, 2022, 108895.

[37] Meyer P E, Bontempi G, On the use of variable complementarity for feature selection in cancer classification, in: Applications of Evolutionary Computing: EvoWorkshops 2006: EvoBIO, EvoCOMNET, EvoHOT, EvoIASP, EvoINTERACTION, EvoMUSART, and EvoSTOC, Budapest, Hungary, 2006, pp. 91-102.

[38] D.L. Wilson, Asymptotic properties of nearest neighbor rules using edited data, IEEE Transactions on Systems, Man, and Cybernetics, SMC-2(3), 1972, 408-421.

[39] I. Tomek, An experiment with the edited nearest-neighbor rule, IEEE Transactions on Systems, Man, and Cyberneticsn, SMC-6 (1976) 448-452.

[40] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye. Q. W, Liu, T. Y., LightGBM: a highly efficient gradient boosting decision tree, Advances in neural information processing systems, 30, 2017, 3149-3157.

[41] Morris T, Gao W, Industrial control system traffic data sets for intrusion detection research, in: Bayro-Corrochano E, Hancock E (eds) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Springer International Publishing, Cham, 2014, pp. 65-78.

[42] Goh J, Adepu S, Junejo KN, Mathur A, A dataset to support research in the design of secure water treatment systems, in: Havarneanu G, Setola R, Nassopoulos H, Wolthusen S (eds) Critical Information Infrastructures Security, 2017, 88-99.

[43] Booij, T. M., Chiscop, I., Meeuwissen, E., Moustafa, N., den Hartog, F. T, ToN_IoT: the role of heterogeneity and the need for standardization of features and attack types in IoT network intrusion data sets, IEEE Internet of Things Journal, 9(1) 2022, 485-496.

[44] W. Pei, B. Xue, M. Zhang, L. Shang, X. Yao, Q. Zhang, A survey on unbalanced classification: how can evolutionary computation help?, IEEE Transactions on Evolutionary Computation, 2023, 1-1.

[45] Feng C, Li T, Chana D, Multi-level anomaly detection in industrial control systems via package signatures and LSTM networks, In: 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), IEEE, 2017, pp.261-272.

[46] J. Ling, Z. Zhu, Y. Luo, H. Wang, An intrusion detection method for industrial control systems based on bidirectional simple recurrent unit, Computers & Electrical Engineering, 91, 2021, 107049.

[47] Y. Zhang, C. Yang, K. Huang, Y. Li, Intrusion detection of industrial internet-of-things based on reconstructed graph neural networks, IEEE Transactions on Network Science and Engineering, 2022, 1-12.

[48] M. Abdelaty, R. Doriguzzi-Corin, D. Siracusa, DAICS: A deep learning solution for anomaly detection in industrial control systems, IEEE Transactions on Emerging Topics in Computing, 10(2), 2022, 1117-1129.

[49] Abdelaty M, Doriguzzi-Corin R, Siracusa D, AADS: A noise-robust anomaly detection framework for industrial control systems, In: Information and Communications Security: 21st International Conference, Springer International Publishing, 2020, pp. 53-70.

[50] Lin. Q, Adepu S, Verwer. S, Mathur. A, TABOR: A graphical model-based approach for anomaly detection in industrial control systems. In: Proceedings of the 2018 on Asia conference on computer and communications security, Association for Computing Machinery, 2018, pp. 525-536.

[51] P. Kumar, G.P. Gupta, R. Tripathi, An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks, Computer Communications, 166, 2021, 110-124.

[52] M.S. Ahmad, S.M. Shah, Unsupervised ensemble based deep learning approach for attack detection in IoT network, Concurrency and Computation: Practice and Experience, 34(27), 2022, e7338.

Meng Huang received the M.S. degree in computer science and technology from Chongqing University, Chongqing, P.R. China, in 2011. He is currently a Ph.D. student with the School of Cyber Science and Engineering, Sichuan University, P.R. China. His current research interests include intrusion detection techniques, artificial immune theory, machine learning, and knowledge graph.
https://orcid.org/0009-0000-9538-5932

Tao Li received his Ph.D. degree in computer science from the University of Electronic Science and Technology of China, in 1994. He is currently a Professor with the School of Cyber Science and Engineering, Sichuan University, China. He is the Chief Scientist of the National Key Research and Development Plan for Cyberspace Security. He is also an editorial board member of Immune Computation and several other international academic journals. His main research interests include network security, artificial immune systems, cloud computing, and cloud storage. He has published nearly 300 papers in IEEE, ACM, Chinese Science, Science Bulletin, Natural Science Progress and other important journals and academic conferences.
https://orcid.org/0000-0002-5302-3180

Beibei Li received the B.E. degree (awarded Outstanding Graduate) in communication engineering from Beijing University of Posts and Telecommunications, P.R. China, in 2014 and the Ph.D. degree (awarded Full Research Scholarship) from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2019. He is curren-

tly an associate professor (doctoral advisor) with the School of Cyber Science and Engineering, Sichuan University, P.R. China. His current research interests include several areas in security and privacy issues on cyber-physical systems (e.g., smart grids, industrial control systems, IoT, etc.), with a focus on intrusion detection techniques, artificial intelligence, and applied cryptography. He has authored or co-authored works in IEEE Transactions on Information Forensics and Security, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Power Systems, IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Network and Service Management, ACM Transactions on Cyber-Physical Systems, IEEE Internet of Things Journal, etc. He won the Best Paper Award in IEEE ISCC 2021.
https://orcid.org/0000-0002-0485-1975

Nian Zhang is a Professor in the Department of Electrical and Computer Engineering at the University of the District of Columbia (UDC), Washington, D.C., USA. She received her Ph.D. degree in Computer Engineering from Missouri University of Science & Technology, USA, and Master's degree in Automatic Control from Huazhong University of Science and Technology, China. Her research interests include machine learning, deep learning, classification, clustering, and optimization. Dr. Zhang serves as an Associate Editor for the IEEE Transactions on Cybernetics, IEEE Transactions on Neural Networks and Learning Systems, Knowledge-Based Systems, and IEEE/CAA Journal of Automatica Sinica. She also serves on the Editorial Board of the Complex & Intelligent Systems. In addition, Dr. Zhang serves as the Chair of the IEEE Computational Intelligence Society (CIS) Task Force on „Interdisciplinary Emergent Technologies" and the Vice Chair of the IEEE CIS' Adaptive Dynamic Programming and Reinforcement Learning Technical Committee.
https://orcid.org/0000-0003-1916-7719

Hanyuan Huang received the M.S. degree in information and communication engineering from Beijing University of Posts and Telecommunications, China in 2019. She is currently pursuing the Ph.D. degree with the School of Cyber Science and Engineering, Sichuan University, China. Her main research interests include cybersecurity, evolutionary computation, and artificial immune systems.
https://orcid.org/0000-0002-9805-7560