INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 16, Issue: 4, Month: August, Year: 2021
Article Number: 4232, https://doi.org/10.15837/ijccc.2021.4.4232

CCC Publications

AGORA UNIVERSITY PRESS

# Improved RBF Network Intrusion Detection Model Based on Edge Computing with Multi-algorithm Fusion

Xuejun Liu, Kaili Li, Wenhui Wang, Yong Yan, Yun Sha, Jianping Chen, Jiaojiao Qin

**Xuejun Liu*, Kaili Li, Wenhui Wang, Yong Yan, Yun Sha, Jianping Chen, Jiaojiao Qin**
Department of Information Engineering
Beijing Institute of Petrochemical Technology
19 Qingyuan North Road, Daxing District, Beijing
*Corresponding author: lxj@bipt.edu.cn

**Abstract**

Edge computing is difficult to deploy a complete and reliable security strategy due to its distributed computing architecture and inherent heterogeneity of equipment and limited resources. When malicious attacks occur, the loss will be huge. RBF neural network has strong nonlinear representation ability and fast learning convergence speed, which is suitable for intrusion detection of edge detection industrial control network. In this paper, an improved RBF network intrusion detection model based on multi-algorithm fusion is proposed. Kernel Principal Component Analysis (KPCA) is used to extract data dimensions and simplify data representations. Then Subtractive Clustering algorithm (SCM) and Grey Wolf algorithm (GWO) are used to jointly optimize RBF neural network parameters to avoid falling into local optimum, reduce the calculation of model training and improve the detection accuracy. The algorithm can better adapt to the edge computing platform with weak computing ability and bearing capacity. The experimental results of BATADAL dataset and Gas dataset show that the accuracy of the algorithm is over 99% and the training time of larger samples is shortened by 50 times for BATADAL dataset. The results show that the improved RBF network is effective in improving the convergence speed and accuracy in intrusion detection.

**Keywords:** RBF network, intrusion detection, kernel principal component analysis, grey wolf algorithm, edge computing.

## 1 Introduction

Edge computing is a new technology introduced in the field of industrial Internet in recent years. By sinking computing, storage, network, communication and other resources to the edge of the network (industrial site, data source, etc.), it provides edge intelligent services for applications in close range [1]. It can meet the key requirements of intelligent access, real-time communication and privacy protection in the industrial Internet environment, and effectively reduce network overhead and system resource consumption. However, edge devices are usually difficult to deploy a complete and reliable security policy due to resource constraints, and are prone to single point failure when malicious attacks occur [2]. Destroyers can choose to start from any Internet of Things equipment with security vulnerabilities to destroy and pollute the entire Internet ecology, tamper and falsify data, resulting in paralysis of the entire system security defense. The data gathered by the industrial Internet usually has high privacy and confidentiality [3]. Edge computing can complete some data processing and computing tasks at the edge of the industrial network, avoiding the risk of leakage of sensitive data in the transmission process of long physical links [4]. However, due to its distributed computing architecture and inherent heterogeneity of equipment, it is more likely to become the target of Trojans and malicious software attacks. For resource-constrained edge nodes, the existing security strategy cannot well adapt to the edge computing architecture. Therefore, it is necessary to redesign more secure protection schemes and measures. Intrusion detection classification algorithm is a kind of protection technology that collects equipment and network-related information and classifies them to determine whether there is abnormal behavior in the system [5].

Compared with cloud computing, edge computing has the advantages of low latency and high real-time performance, which can better meet the needs of the construction of industrial Internet in the future. However, due to the edge computing form of intelligent gateways, sensors, industrial computers, boxes and other factories close to equipment, the computing ability and carrying capacity are weak. Due to the limit of business carrying capacity, complex data processing models, edge computing will not be able to adapt to the platform [6]. Therefore, based on the architecture of cloud computing, we combine edge computing with classification algorithm to reduce the amount of calculation and improve the calculation rate, so as to meet the security requirements of edge-side platform applications. Among many classification algorithms, RBF has the advantages of simple network structure, adaptability and strong nonlinear representation ability. It can adapt to different industrial production environments by adjusting the number of neurons in the layer and the calculation rules between the network layers, and can be used as an anomaly detection scheme for industrial control network.

Similar to the hyperparameter training of deep neural network, RBF also needs to face the problem of parameter optimization, for which researchers at home and abroad have made a lot of studies. In order to adjust the network structure of RBF and make it more suitable for complex scenes, literature [7] used Principal Component Analysis (PCA) to reduce the dimension of RBF network input to solve the problem of seismic attribute redundancy. Literature [8] proposed an RBF network learning algorithm based on improved K-Means to eliminate the sensitivity of clustering and reduce the time complexity. However, due to the use of local activation function, RBF network can approach any nonlinear function, but it is easy to fall into the local optimal. Besides, a lot of research has been carried out at home and abroad. Sun Qian et al. use Particle Swarm Optimization (PSO) to optimize neural network parameters to solve the problem of slow convergence speed caused by random selection of initial parameters of RBF network. But the time complexity increases and it is easier to fall into local optimization [9]. Literature [10] put forward Genetic Algorithm (GA) to optimize the structural parameters such as weight and network layer element number of RBF network, which is reliable, but the selection of parameters is subjective, the convergence speed is slow, and it is easy to fall into premature convergence [11]. The experiment of RBF network model based on ant colony algorithm by Liu Chang improves its accuracy and speeds up the convergence. However, the algorithm needs to search the path and has high computational cost [12]. The Grey Wolf algorithm (GWO) has the advantages of simple structure, few parameters to be set and it is easy to implement in experimental coding. It is also suitable for network parameter optimization, such as Swarna, for intrusion detection system, DNN neural network parameters are optimized and adjusted by Grey Wolf algorithm, which reduces the time complexity and improves the accuracy, which proves its feasibility [13].

Therefore, in view of the attribute redundancy and nonlinear separable industrial control network datasets, the variable weight is difficult to determine when the algorithm is trained, and it is easy to fall into local optimum. The paper proposes an improved RBF network intrusion detection model based on multi-algorithm fusion. The algorithm can improve the accuracy of classification, reduce the amount of calculation and improve the calculation speed, so as to better adapt to the edge computing platform, facilitate data processing and analysis, and provide technical support for intelligent decision-making and execution of local business.

## 2 Edge computing platform

In the industrial control application scenario, the problem of computing resource allocation and load balancing is very complex, and the resource management mechanism in cloud computing mode is usually difficult to match the edge computing production environment. Therefore, edge cloud computing architecture is widely used in the industrial field. Edge computing will be close to the network edge side of the source of objects or data, and an open platform that integrates core capabilities of network, computing, storage and application to provide edge intelligent services. Edge computing is closely related to industrial control systems. Industrial control system with industrial Internet interface is essentially an edge computing device, as shown in Figure 1.

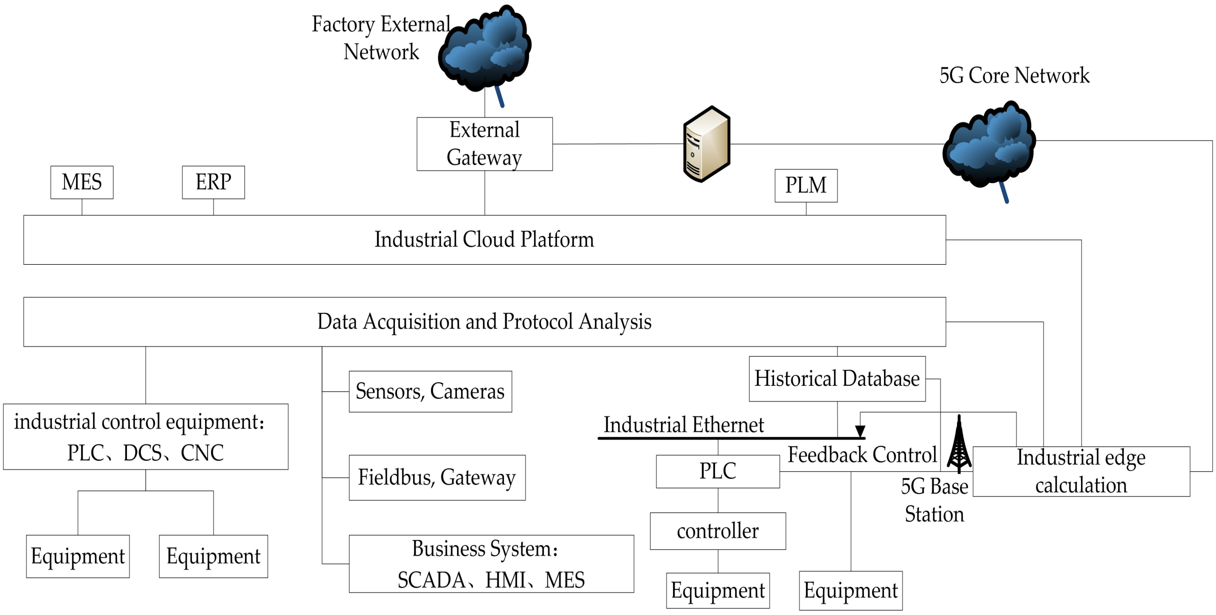

Figure 1: 5G edge computing embedded in industrial architecture

However, in the edge computing environment, due to the limited computing resources within nodes and the differences in computing capabilities between nodes, the edge layer is difficult to deal with computing tasks (such as big data analysis, deep learning model fusion training and historical data management) and storage tasks (such as mirror warehouse). In addition, because the distribution of edge servers and device nodes is usually related to the deployment of industrial environment, the distribution frame is fixed, and it is vulnerable to malicious attacks, resulting in huge losses. Therefore, in view of this situation, this paper proposes an intrusion detection model based on multi-algorithm fusion of RBF neural network to reduce the computational requirements and improve the computational speed, so as to meet the basic needs of data processing and security protection of the edge platform.

## 3 Data processing

In the face of large-scale diversified heterogeneous data on the edge side, the difficulty of extracting useful information from the data will also increase accordingly. How to efficiently clean, aggregate and analyze it has become a key technology. In view of this situation, we use the data preprocessing method to standardize the data, and then use KPCA to extract the nonlinear feature of the data, reduce the dimension of the input data, simplify the data representation, and reduce the calculation amount after entering the neural network.

### 3.1 Preprocessing

The industrial control network dataset is the data collected directly from the industrial control equipment (PLC, frequency converter, etc.) in the manufacturing field. Because of the complex field situation, there will be confusion of data types, redundant attributes, missing values and outliers, which can't be directly applied to the prediction model, so the data must be preprocessed first.

(1) Data cleaning: The probability of data missing in industrial control network is random. For example, power failure of frequency converter or sensor, the data collected in a certain period of time is empty or noise data, which is a small probability event, and the frequency of occurrence is not high, and the proportion of certain records is low, we can simply delete the samples with missing value.

(2) Type conversion: Since the tag value cannot be directly used in the mathematical equation of the network model, it is necessary to convert the tag value to a numerical value.

(3) Data reduction: In order to eliminate the dimensional influence between data features, it is necessary to normalize the features to make the different indexes comparable. When the optimization objective function is used to solve the problem, the update speed of different features tends to be consistent, which can accelerate the convergence speed of the model. Due to the non-linear characteristics of the dataset, Z-score standardization is adopted: $X = (D - \mu)/\delta$, where $D$ is raw dataset with no missing values, $\mu$ is the mean value of all samples in a single dimension, and $\delta$ is the standard deviation of all samples in a single dimension.

### 3.2 KPCA nonlinear feature extraction

Suppose that the centralized sample set after preprocessing is $X = [X_1, X_2, \cdots, X_n]$, and it is a $n \times d$ matrix, which represents that the number of input dimension is $n$ , and the dimension of each sample is $d$. Any vector in the space can be represented linearly by all samples, so the sample set $X$ is used to represent the eigenvectors:

$$\omega_i = \alpha\Phi(x_i) \tag{1}$$

In this paper, $X$ is mapped to a higher dimensional space and a nonlinear mapping function $\Phi(x)$ is introduced [14], that is:

$$\sum\nolimits_{i=1}^{d}(\Phi(x_i)\cdot\Phi(x_i)^\top)\sum\alpha\Phi(x_i) = \lambda\sum\alpha\Phi(x_i) \tag{2}$$

We multiple both sides with $\Phi(x_i)^\top$ ,

$$\Phi(x_i)^\top\sum\nolimits_{i=1}^{d}(\Phi(x_i)\cdot\Phi(x_i)^\top)\sum\alpha\Phi(x_i) = \lambda\Phi(x_i)^\top\sum\alpha\Phi(x_i) \tag{3}$$

Command $k(x_i, x_i) = \Phi(x_i)^\top \cdot \Phi(x_i)$, we get the relation:

$$K\alpha = \lambda\alpha \tag{4}$$

Where $K$ is the kernel matrix corresponding to $k$, $\alpha$ is the eigenvector in the space, and $\lambda$ is the corresponding eigenvalue. Then, a set of bases of the high-dimensional space is obtained by formula (1), which constitutes a subspace of the high-dimensional space $V_r$ . The sample points mapped $\Phi(x_i)$ to the high-dimensional space are projected on $V_r$, and the nonlinear principal component components are obtained as follows:

$$M_n = f_r(x_j) = \Phi(x_j)\cdot V_r = \sum\nolimits_{i=1}^{n}\alpha_r[\Phi(x_i)\cdot\Phi(x_j)] = \sum\nolimits_{i=1}^{n}\alpha_r K(x_i, x_j) \tag{5}$$

## 4 Improved RBF Network Intrusion Detection Model

Due to the inherent distributed characteristics of edge computing, the computing resources within nodes are limited and the computing ability between nodes is reduced, which cannot meet the requirements of big data analysis and deep learning model fusion training. Therefore, this paper proposes an improved RBF network intrusion detection model based on the discrete distribution of computing and storage resources. By optimizing the parameters to adjust the local optimum, the accuracy and calculation rate are improved to adapt to the anomaly detection problem of edge computing platform.

### 4.1 RBF Network Model

RBF network is a simple three-layer forward network. In RBF [15], input vector is directly mapped to the hidden layer, and the mapping from hidden layer space to output space is linear. That is to say, the output of the network is the linear weighted sum of the output of hidden units. The weight here is the network adjustable parameter [16]. In this way, the output is nonlinear for the mapping of network from input to output, while it is linear for adjustable parameters. The weight of the network can be solved directly by linear equations. Thus greatly accelerate the learning speed is very conducive to deployment to the edge side platform, as shown in Figure 2.

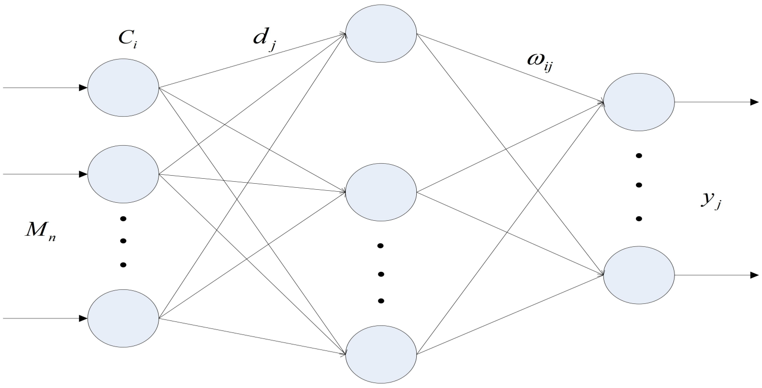

Figure 2: Topological structure of RBF network

The activation function can be expressed as follows [17]:

$$R(M_n - c_i) = exp(-\|M_n - c_i\|^2/2\sigma_i^2) \tag{6}$$

According to the linear weighted sum, the output is obtained as follows:

$$y_j = \sum_{i=1}^{h} \omega_{ij} exp(-\|M_n - c_i\|^2/2{\sigma_i}^2) \quad j = 1, 2, \cdots, n \tag{7}$$

The supervised learning algorithm is used to train all the parameters of the network. It mainly involves gradient descent of the cost function (mean square error), and correcting each parameter. The total error of the network is as follows:

$$E = \sum_{i=1}^{n} E_n = \frac{1}{2}\sum_{i=1}^{n}(d_j - y_j)^2 \tag{8}$$

Where $M_n$ is input samples, $c_i$ is the central point, $\sigma_i$ is the width of the corresponding radial basis function, $h$ is the number of nodes in the hidden layer, $d_j$ is the expected output, $y_j$ is the actual output, and $n$ is the classification number of outputs.

In this paper, RBF network structure is used, and the main parameters are optimized as follows: how to determine the data center and expansion constant of each radial basis function, and how to correct the output weight. This is because the radial basis function has radial symmetry. The farther the input of neurons is away from the data center, the lower the activation degree of neurons is, which is called "local characteristic". In view of the above problems, this paper mainly uses the subtractive clustering algorithm to determine the data center of RBF and the initial value of the expansion constant. The grey Wolf algorithm is used to optimize and update the output weight and modify it. Finally, the network model is trained to achieve the classification effect.

### 4.2 Subtractive clustering optimization algorithm

The subtractive clustering algorithm determines the data center ($c$) according to the density index, and the expansion constant ($\sigma$) according to the distance between the cluster centers. Calculate the distance between each data point of the drop-dimensional data $M(n \times t)$ to get the density index:

$$D_i = \sum_{j=1}^{n} exp[-\alpha d(m_i, m_j)] \quad \alpha = 4/\tau_1^2 \tag{9}$$

Find the data with the largest density index as the first cluster center, and then calculate the density of all the remaining data points:

$$D_i = D_i - D_{c1} exp[-\beta d(m_i, m_{c1})] \quad \beta = 4/\tau_2^2 \tag{10}$$

Where $\tau_1$ is the neighborhood radius of ; $\tau_2$ is a neighborhood whose density index function is significantly reduced. $\tau_2 = \eta\tau_1$ .

Find the maximum density index, and take this point as the cluster center until satisfy the conditions: $D_{ci}/D_{c1} \leq \varepsilon$, then the iteration stops, the number of clusters is recorded as $h$. Calculate the distance between all samples and each cluster center, divide samples according to the principle of minimum distance:

$$I(M_j) = \underset{i}{min} \|M_j - c_i\| \quad i = 1, 2, \cdots, h \tag{11}$$

Calculate the distance between cluster centers:

$$d_i = \underset{i}{min} \|c_j - c_i\| \tag{12}$$

Determine the expansion constant $\sigma_i = \kappa d_i$ according to the distance between cluster centers, which $\kappa$ is called overlap coefficient.

### 4.3 The grey wolf optimization algorithm

The radial basis function has radial symmetry. The farther the input of neurons is from the data center, the lower the activation of neurons, which is called "local optimum". In this paper, the grey wolf algorithm is used to globally converge the weight and ameliorate the local optimal problem. The weight matrix $W$ from the hidden layer to the output layer trained by gradient descent method was used as the initial position variable of adult gray wolves [18]:

$$W = [\omega_{11}, \omega_{12}, \cdots, \omega_{1n}, \omega_{21}, \cdots, \omega_{hn}] \tag{13}$$

Construct the social hierarchy model of grey wolf [19], as shown in Figure 3. The total network error is calculated according to formula (8) and used as the fitness of each individual in the population. The three grey wolves with the best fitness in the wolf pack are marked as $\alpha$,$\beta$,$\sigma$. Respectively, the remaining grey wolves are marked as $\omega$ [20].

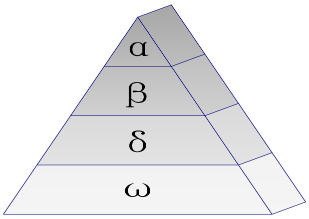

Figure 3: Schematic diagram of social hierarchy of Grey Wolf algorithm

Encircle prey: the wolf will gradually approach the prey and surround it when hunting. The mathematical model of this behavior is as follows [21]:

$$\vec{D} = |\vec{C} \cdot \vec{X}_P(t) - \vec{X}(t) \quad \vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad \vec{A} = 2\vec{\alpha} \cdot \vec{r}_1 - \vec{\alpha} \quad \vec{C} = 2 \cdot \vec{r}_2 \tag{14}$$

$t$ is the current iteration number; $\vec{A}$ and $\vec{C}$ are the synergy coefficient vector; $\vec{X}_P(t)$ represents the position vector of prey; $\vec{X}(t)$ represents the current position vector of grey wolf [22]; $\vec{r}_1$ and $\vec{r}_2$ are the random vector in [0, 1].

Hunt [23]: in each iteration process, the best three grey wolves ( $\alpha$,$\beta$,$\sigma$)in the current population are retained, and the positions of $\omega$ is updated according to their location information [24]. The mathematical model of this behavior can be expressed as follows [25]:

$$\begin{gathered}\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\sigma = |\vec{C}_3 \cdot \vec{X}_\sigma - \vec{X}| \\ \vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\sigma - \vec{A}_3 \cdot (\vec{D}_\sigma) \\ \vec{X}(t+1) = (\vec{X}_1 + \vec{X}_2 + \vec{X}_3)/3\end{gathered} \tag{15}$$

$\vec{X}_\alpha$,$\vec{X}_\beta$,$\vec{X}_\sigma$ respectively represent the position vector of $\alpha$,$\beta$,$\sigma$ in the current population; $\vec{X}$represents the position vector of grey wolf; $\vec{D}_\alpha$,$\vec{D}_\beta$,$\vec{D}_\sigma$ respectively represent the distance between the current candidate grey wolf and the optimal three wolves [26].

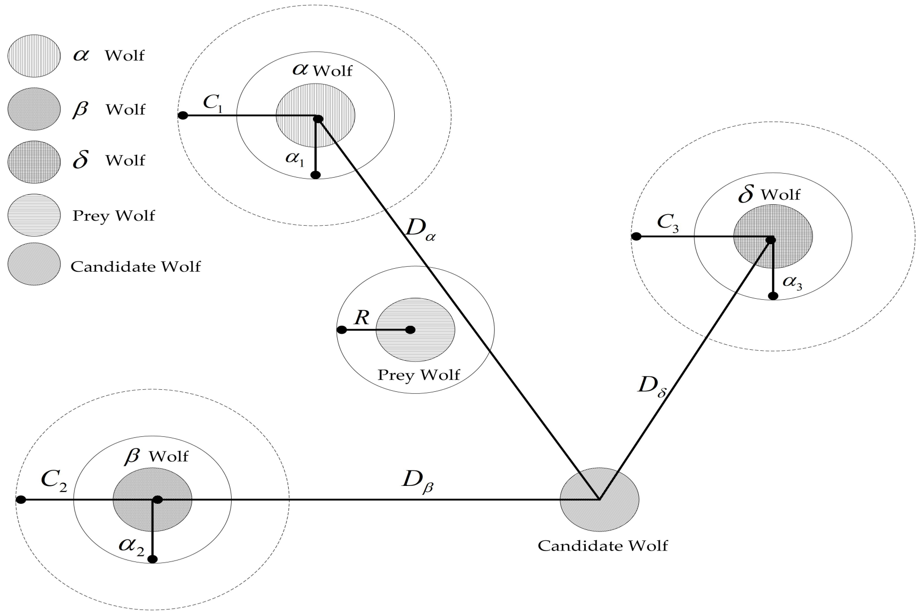

Figure 4: Hunting process of Grey Wolf algorithm

It can be inferred from Figure 4, there is necessary for $\alpha$,$\beta$,$\sigma$ to predict the approximate location of the prey [27]. Then the wolf $\omega$ randomly updates its position near the prey under the guidance of the current optimal three wolves [28].

# 5 Work flow of proposed algorithm

## 5.1 Pseudo-code

**Algorithm 1: Function <KPCA>**

```
Input: industrial control dataset
Output: Dataset with dimension reduction
Data preprocessing        //according to chapter 3.1
The centralized sample set $X = [n \times d]$ after preprocessing
Selected kernel function $k$
Calculate feature vector $(\alpha_1, \alpha_2, \cdots, \alpha_d)$ by formula (4)
Sort the eigenvalues and obtain the adjusted eigenvector $(\alpha_1, \alpha_2, \cdots, \alpha_t)$
Calculate basis vector $\omega_i$ by formula (1) as subspace $V_r$
Project sample points $\Phi(x_i)$ on $V_r$
Calculate nonlinear principal component $f_r(x_j)$ by formula (5)
Obtain new dataset $M_n$ with dimension reduction
```

**Algorithm 2: Function <SCM-RBF>**

```
Input: Dataset with dimension reduction $M_n$
Output: Data center and expansion constant of radial basis function
Calculate density index $D_i$ by formula (9)
Set the sample with the largest density index as the cluster center $c_1$
Calculate density index $D_{ci}$ of other samples by formula (10)
$h = 1$          //initialize the number of cluster centers
$i = 1$
while ($D_{ci}/D_{c1} \leq \varepsilon$) do
        $h = h$
        $c_i(h) = [c_1, c_2, \cdots, c_i]$
        else
                $i = i + 1$
                Update all density index $D_{ci}$ of other samples
                $D_c \leftarrow max(D_{ci})$
                $c_i \leftarrow$ sample with the largest density index
                $h = h + 1$
end while
while ($c_i(h) \neq c_{i-1}(h)$) do
        Calculate the distances between cluster center and samples
        Reclassify each sample by formula (11)
        Reset cluster center as $c_{i-1}(h)$
end while
Calculate the distances between cluster centers $d_i$ by formula (12)
Set expansion factor $\sigma_i = \kappa d_i$
```

**Algorithm 3: Function < GWO-RBF>**

```
Input: Data center $c_i$ and expansion constant of radial basis function $\sigma_i$
Output: The RBF model after training
Initialize weight matrix $W$ randomly
Map $W$ as the position vector of artificial gray wolves by formula (13)
Calculate the total error $E$ of RBF network by formula (6), (7) and (8)
while ($E \leq E_{min}$) do
       // $E$ is the total error of the network, $E_{min}$ is minimum value set
        while ($t \leq t_{max}$) do
                // $t$ is the iteration, $t_{max}$ is the maximum iterations set
                Fitness of each individual in the population $\leftarrow E$
                Mark the three gray wolves with best fitness as $\alpha$,$\beta$,$\sigma$ respectively
                The rest of population is marked as $\omega$
                Predict the position of optimal solution(prey) by formula (14)
                Keep the best three wolves ($\alpha$,$\beta$,$\sigma$) in the current population
                Update the position of $\omega$ by $\alpha$,$\beta$,$\sigma$ by formula (15)
                Update weight matrix $W$
                Update the total error $E$ of RBF network
                else
                    Keep current weight matrix $W$ and start training
        end while
end while
```

### 5.2 Work flow of proposed algorithm

Above all, the specific flow chart of the algorithm is as shown in Figure 5.

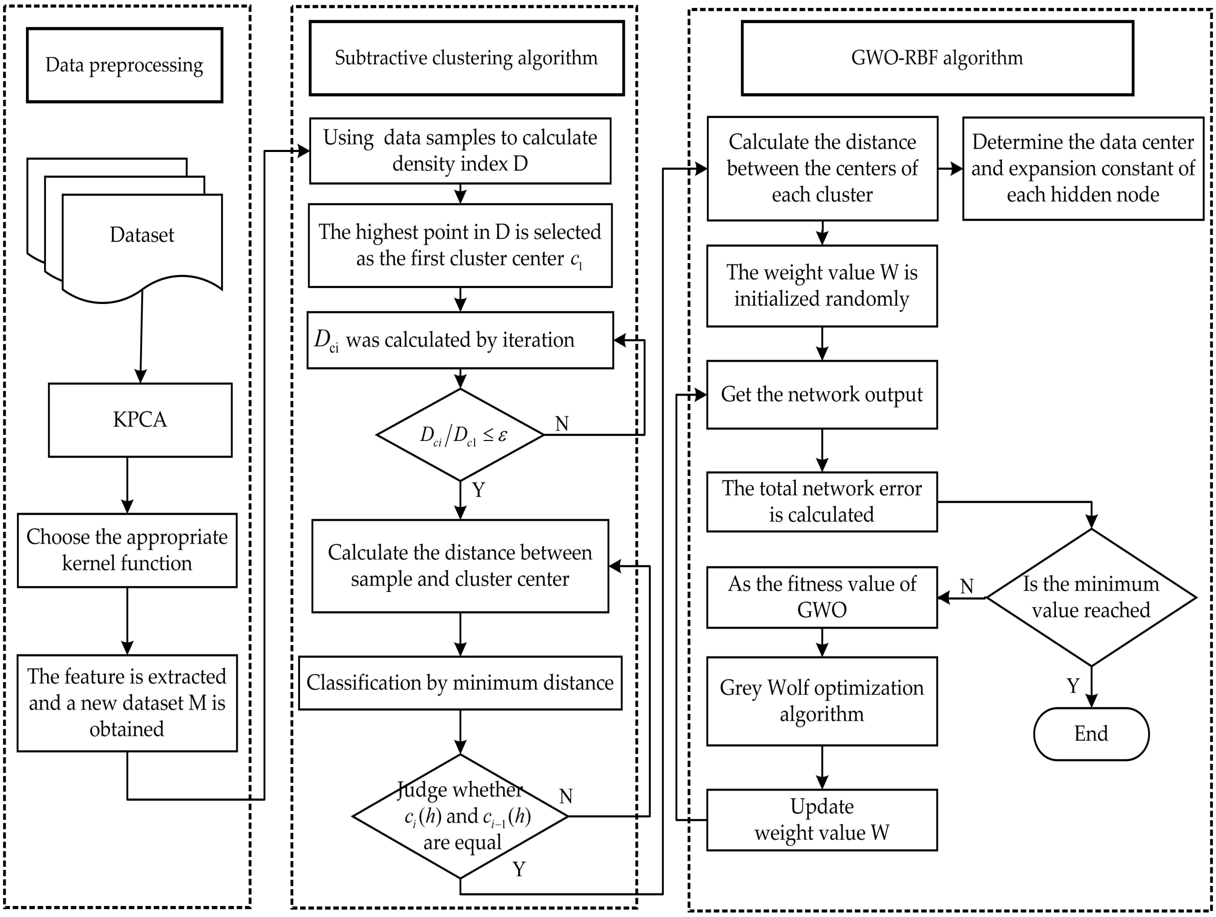
Data preprocessing
Dataset
KPCA
Choose the appropriate kernel function
The feature is extracted and a new dataset M is obtained
Subtractive clustering algorithm
Using data samples to calculate density index D
The highest point in D is selected as the first cluster center $c_1$
$D_{ci}$ was calculated by iteration
$D_{ci}/D_{c1} \le \varepsilon$
N
Y
Calculate the distance between sample and cluster center
Classification by minimum distance
Judge whether $c_i(h)$ and $c_{i-1}(h)$ are equal
N
Y
GWO-RBF algorithm
Calculate the distance between the centers of each cluster
Determine the data center and expansion constant of each hidden node
The weight value W is initialized randomly
Get the network output
The total network error is calculated
Is the minimum value reached
N
Y
As the fitness value of GWO
Grey Wolf optimization algorithm
Update weight value W
End

Figure 5: Work flow of proposed algorithm in this paper

## 6 Experiments

In the experiment, we used two kinds of public industrial control datasets, including BATADAL dataset and Gas dataset. We also use smaller sample dataset (Iris dataset) as controls. The PSO-RBF algorithm and SCM-RBF algorithm were used for comparison and verification.

### 6.1 Dataset

The Battle of the Attack Detection Algorithms (BATADAL) is the most recent competition on planning and management of water networks undertaken within the Water Distribution [29]. BATADAL dataset is mainly the data record of estimated water consumption dynamically established by EPANET simulation system, which covers a time step of 492 hours and contains seven attacks. Physical attacks are used to interfere the operation of system, which could lead to overflow, pump speed reduction, abnormal activation or deactivation of pump, and other abnormal activity. In addition, the attacker changes the signal sent by the sensor to the SCADA by adding an offset that changes over time.

Dataset of a gas pipeline system in Mississippi State University's Critical Infrastructure Protection Center (Gas dataset) [30] is used for intrusion detection and evaluation of industrial control system, which collects information of network affairs between remote terminal unit (RTU) and main control unit (MTU) in SCADA natural gas pipeline. It is used to simulate actual attacks and operator activities on natural gas pipeline.

Table 1: Basic situation of dataset used in experiment

| Name | Number of samples | Number of positive samples | Number of negative samples | Number of training samples | Number of testing samples | Dimension | Number of clusters |
|---|---|---|---|---|---|---|---|
| BATADAL dataset | 12938 | 8762 | 4176 | 10350 | 2588 | 44 | 2 |
| Gas dataset | 97019 | 61156 | 35863 | 77615 | 19404 | 27 | 8 |
| Iris dataset | 150 | - | - | 120 | 30 | 4 | 3 |

Iris dataset is a classic UCI dataset, which collects three types of iris, namely Setosa, Versicolor and Virginia iris. Each record contains four characteristics: calyx length, calyx width, petal length and petal width. See Table 1 for details.

## 6.2 Experimentation

Edge calculation platform parameters:

Processor: Intel(R) Core(TM) i7-3517U CPU @ 1.90GHz 2.39GHz, System: Windows 7, Installing memory: 8.00GB, Compiler environment: Python 3.7.4.

### 6.2.1 KPCA dimension reduction

Select the appropriate kernel function, then determine the number of nonlinear principal components, and then extract the nonlinear principal components.

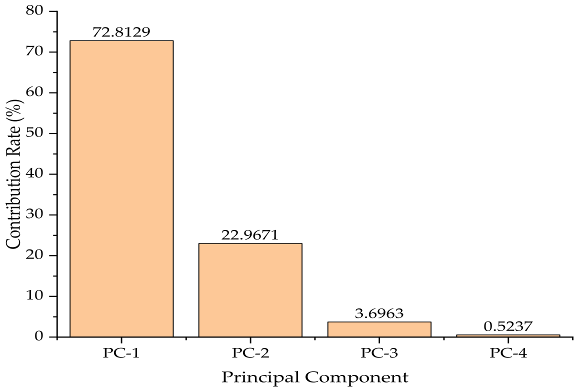

Figure 6: Histogram of contribution rate of eigenvector

According to the set cumulative contribution rate, the eigenvectors corresponding to the first $t(t < d)$ eigenvalues are selected as $\alpha_1, \alpha_2, \cdots, \alpha_t$. Taking the iris dataset as an example,as shown in Figure 6, the contribution rate of the first two eigenvalues reaches about 96% after selecting linear kernel for high-dimensional mapping, so the corresponding eigenvectors are retained.

### 6.2.2 Experimental results

The parameters used in this experiment are as follows: neighborhood radius: $\tau_1 = 2$ , $\tau_2 = 4$; $\eta = 1.5$; $\varepsilon = 0.99$; overlap coefficient: $\kappa = 1$; number of wolves: $n = 20$ ; maximum number of iterations: $t_{max} = 10$; upper and lower bounds of random variables: $ub = 10$ , $lb = 0.01$ .

Due to the unbalanced proportion of positive and negative samples in BATADAL database and Gas dataset, the clustering algorithm is easy to cluster the two data centers to the position of positive samples at the same time. It means that the clustering algorithm ignores the role of negative samples, resulting in poor classification effect. On this basis, particle swarm optimization (PSO) makes it easier to fall into local optimum, increase the operation time and restrain the convergence speed of the network. The grey wolf algorithm establishes a decentralized model and randomly assigns search coefficients, which greatly improves the global search ability, avoids reducing the limitations of the clustering algorithm, and improves the classification accuracy rate.

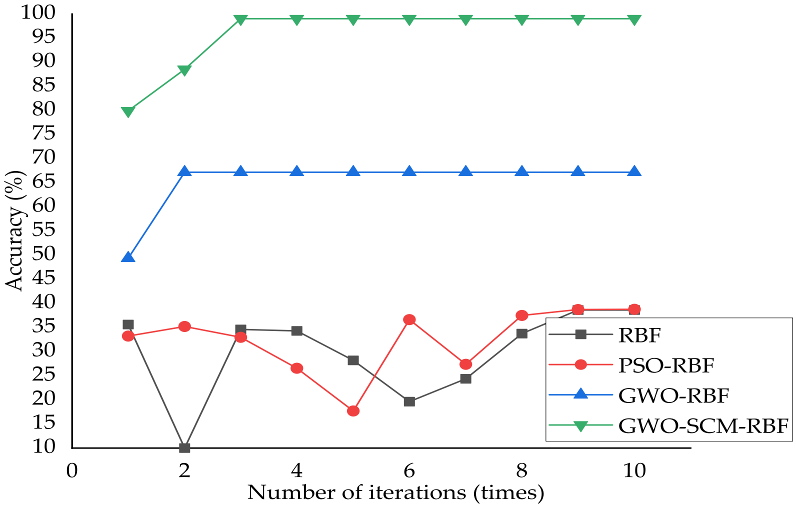

Figure 7: Comparison of algorithm accuracy of BATADAL dataset

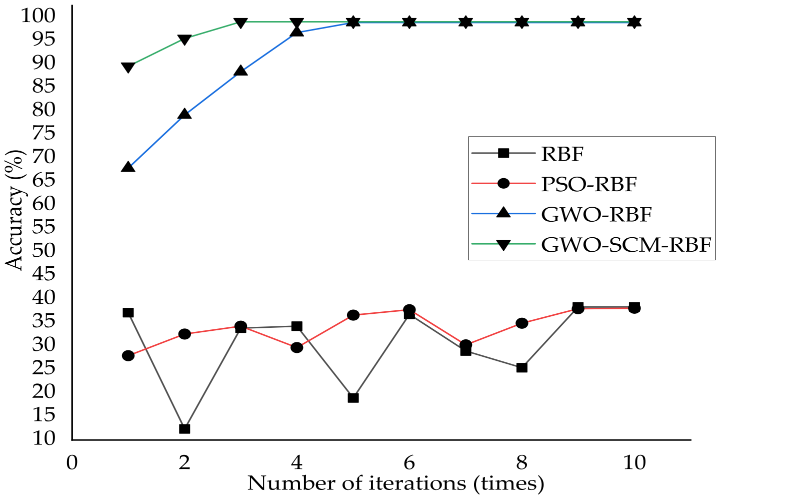

Figure 8: Comparison of the accuracy of algorithms in Gas dataset

As shown in Figure 7 and Figure 8, the prediction accuracy of our algorithm (GWO-SCM-RBF) can be stabilized at about 90% for the industrial control dataset with large sample, which is higher than that of PSO-RBF, GWO-RBF and RBF constructed by traditional learning. At the same time, we can see that the change trend of traditional RBF using gradient descent method is complex and the convergence is very slow from the trend of the graph. Although particle swarm optimization can improve the speed of searching for extreme value, it is easy to fall into the local optimization, and the accuracy rate cannot be improved. The grey wolf algorithm can reach stability after iteration to 2-3

Table 2: Comparison of evaluation indexes of the algorithm

| Name | | RBF | PSO-RBF | GWO-RBF | GWO-SCM-RBF | KPCA-GWO-SCM-RBF |
|---|---|---|---|---|---|---|
| BATADAL dataset | Accuracy (%) | 38.6 | 38.8 | 67.2 | 99.0 | 99.2 |
| | Time Consuming(s) | 927.9 | 912.6 | 12998.6 | 175860.3 | 34598.2 |
| Gas dataset | Accuracy (%) | 38.3 | 38.0 | 98.9 | 99.0 | 99.5 |
| | Time Consuming(s) | 48085.2 | 49228.7 | 12998.7 | 29925.3 | 25326.7 |
| Iris dataset | Accuracy (%) | 96.7 | 96.7 | 90.0 | 99.0 | 99.3 |
| | Time Consuming(s) | 0.17 | 0.42 | 1.01 | 5.24 | 5.05 |

generations. The enhancement of search ability makes the convergence speed faster, the trend is more stable than other algorithms, and the classification effect is enhanced.

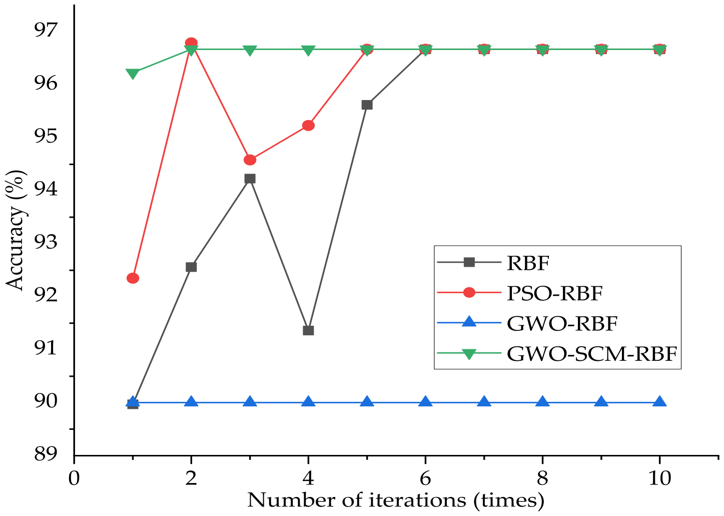

Figure 9: Comparison of the accuracy of each algorithm in Iris dataset

For Iris dataset, the accuracy of all kinds of algorithms is good. However, it can be seen from the Figure 9 that the accuracy of GWO-RBF algorithm is slightly inferior. The RBF algorithm and PSO-RBF algorithm tend to be stable around the sixth generation, while the algorithm in this paper can find the global optimal value in the first two generations, and the convergence speed and stability are significantly better. This shows that this algorithm is also suitable for smaller dataset.

After using KPCA to reduce the dimension, SCM-GWO is used to jointly optimize parameters of RBF network, which can significantly improve the operation efficiency. The higher the dimension of the dataset, the more obvious the time cost savings. For example, the training time of BATADAL dataset is shortened by about 50 times. The specific evaluation indexes of experimental results are shown in Table 2.

It can be seen from the above table that the accuracy of the improved RBF network intrusion detection model based on multi-algorithm fusion can exceed 99% in the three datasets. Compared with the contrast algorithm, the operation efficiency and convergence speed are significantly improved, which can effectively solve two problems. One is that traditional RBF weight training is easy to fall into local optimization, the other is that the width of data center and radial basis function is not suitable due to the sensitivity of initial cluster center. The improved RBF network improves the

Table 3: Comparison of accuracy on different algorithms

| Name | Accuracy (%) | False Positive Rate (%) | False Negative Rate (%) | Time Consuming(s) |
|---|---|---|---|---|
| LSTM | 98.3 | 2.17 | 2.69 | 3249751.4 |
| RNN | 95.3 | 2.56 | 2.92 | 3156245.4 |
| KNN | 93.1 | 2.10 | 4.22 | 17892.3 |
| DNN | 90.4 | 2.03 | 3.22 | 229481.3 |
| OURS | 99.5 | 2.09 | 2.43 | 25326.7 |

classification effect and is more stable, which shows the effectiveness and practicability in intrusion detection.

For the Gas dataset, this paper also refers to the algorithms in the published papers to conduct a comparative study, such as the LSTM algorithm [31] and RNN algorithm [32] based on time series, the traditional machine learning KNN algorithm [33], and the DNN deep learning algorithm [34], so as to more comprehensively investigate the effectiveness of the proposed algorithm in solving the industrial control intrusion problem.

Meanwhile, it can be seen from Table 3 that compared with other classification algorithms, the algorithm in this paper has the highest accuracy. DNN is a traditional deep neural network with more sufficient feature extraction and the lowest false positive rate. However, DNN has poor ability to identify attack data with fewer samples and a high false positive rate. Such algorithms as LSTM and RNN need to spend a long time on an experimental platform with only a CPU, which is not conducive to the real-time monitoring of industrial control. Experimental results show that the proposed algorithm has higher accuracy and lower time cost under the same edge side platform configuration.

## 7 Discussion

In this paper, a detection model based on KPCA nonlinear extraction of main features, SCM and GWO joint optimization of RBF network parameters and other multi-algorithm fusion is proposed. A comparative test is carried out on the datasets of public networks such as Singapore Waterworks and Mississippi Natural Gas, which improves the computational efficiency and detection accuracy of the traditional RBF neural network on the CPU, and adapts to the computational requirements and resource constraints of the edge-side platform. However, although the subtractive clustering algorithm used in the current algorithm is independent of the data dimension, it shows a linear relationship with the number of samples. When the sample size is too large, it will also increase the time complexity, which is a subject that needs further study.

### Funding

The APC was funded by R&D center "Cercetare Dezvoltare Agora" of Agora University.

### Author contributions

The authors contributed equally to this work.

### Conflict of interest

The authors declare no conflict of interest.

## References

[1] XiaoFeng Lu; YuYing Liao; Pietro Lio; Pan Hui. (2020). An Efficient Asynchronous Federated Learning Mechanism for Edge Computing, *Journal of Computer Research and Development*, 57(12), 2571–2582, 2020.

[2] ShuaiQi Shen; Kuan Zhang; Yi Zhou; Song CI. (2020). Security in edge-assisted Internet of Things:Challenges and Solutions, *Science China(Information Sciences)*, 63(12), 27–40, 2020.

[3] ChengXu Zhang; HongWu Li; Yang Qu; JinWu Wei.(2021). Development and Application of 5G Edge Computing for Industrial Internet, *Telecommunications Science*, 37(01), 129–136, 2021.

[4] Hui Li; Xiuhua Li; Qingyu Xiong; Junhao Wen; Luxi Cheng; Bin Xing.(2021). Edge Computing Powering the Industrial Internet: Architecture, Applications, and Challenges, *Computer Science*, 48(01), 1–10, 2021.

[5] WenAn Zhang; Zhen Hong; JunWei Zhu; Chen Bo. (2019). Survey of Network Intrusion Detection Methods in Industrial Control System, *Control and Decision Making*, 34(11):, 2277–2288, 2019.

[6] ShangHong Zhang; GaoFeng Cui; YaTing Long; WeiDong Wang. (2021). Joint Computing and Communication Resource Allocation for Satellite Communication Networks with Edge Computing, *Chinese Communication*, 18(07), 236–252, 2021.

[7] WanLi Jia; Yang Liu; LeLe Zhang. (2019). Reservoir Prediction Based on RBF Neural Network Optimized by PCA, *Proceedings of 2019 China Geoscience joint annual meeting (17)*, 143–144, 2019.

[8] Zhen Pang; WeiHong Xu. (2012). A learning method of RBF Neural Network Based on Improved K-Means Method, *Computer Engineering and Application*, 48(11), 161–163+184, 2012.

[9] Qian Sun; Xin Zhao. (2020). DNA Sequence Classification Based on RBF Neural Network Optimized by Particle Swarm Optimization, *Modern Electronic Technology*, 43(09), 87–91, 2020.

[10] Rui Yang. (2011). *Optimization of RBF Neural Network Controller Based on Genetic Algorithm*, Harbin University of Technology, 2011.

[11] Wei Lou; XingGao Liu. (2007). Prediction Model of Melt Index of Polypropylene Based on PCA-GA-RBF Network, *Journal of Petrochemical Colleges and Universities*, 2007(03), 82–85, 2007.

[12] Chang Liu; ZhiGang Li; LiGuo Wei; JiZhong Wang; Yang Li. (2019). Application of RBF Neural Network Based on Ant Colony Algorithm in Impulse Grain Flow Sensor, *Jiangsu Agricultural Sciences*, 47(15), 259–263, 2019.

[13] Swarna Priya R.M.; Praveen Kumar Reddy Maddikunta; Parimala M.; Srinivas Koppu; Thippa Reddy Gadekallu; Chiranji Lal Chowdhary; Mamoun Alazab. (2020). An Effective Feature Engineering for DNN Using Hybrid PCA-GWO for Intrusion Detection in IOMT Architecture, *Computer Communications*, 160, 2020.

[14] Chen Decheng; Fu Rong; Song Shaoqun; Qin Jie. (2018). Network Security Situation Awareness of Power Dispatching Automation System Based on LDA-RBF, *2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*, 23–25, 2018.

[15] ZhiZheng Zhang; DongJie Wang; YongLiang Zhang. (2020). Research on Fault Monitoring and Diagnosis Method Based on Improved KPCA-SVM Based on PSO Method, *Modern Manufacturing Engineering*, (09), 101–107, 2020.

[16] YongMing Han; ChenYu Fan; ZhiQiang Geng; Bo Ma; Di Cong; Kai Chen; Bin Yu. (2020). Energy Efficient Building Envelope Using Novel RBF Neural Network Integrated Affinity Propagation, *Energy*, 209, 2020.

[17] Shadi; Abpeykar Mehdi; GhateeHadi; ZareEnsemble. (2019). Decision Forest of RBF Networks via Hybrid Feature Clustering Approach for High-Dimensional Data Classification, *Computational Statistics & Data Analysis*, 12–36, 2019.

[18] GuiMei Yao; BaoBin Miao. (2019). Ship Course Keeping Control Design Based on Reduced Clustering and Adaptive Neuro Fuzzy Inference, *Ship Engineering*, 41(04), 82–87+13957, 2019.

[19] Olusegun David Samuel; Modestus O. Okwu; Oluwayomi J. Oyejide; Ebrahim Taghinezhad; Asif Afzal; Mohammad Kaveh. (2020). Optimizing Biodiesel Production from Abundant Waste Oils through Empirical Method and Grey Wolf Optimizer, *Fuel*, 281, 2020.

[20] Akanksha Bhardwaj; Alpesh Kumar. (2020). Numerical Solution of Time Fractional Tricomi-type Equation by An RBF Based Meshless Method, *Engineering Analysis with Boundary Elements*, 118, 2020.

[21] Mathematics. (2020). New Findings from Jiangsu University of Science and Technology Describe Advances in Mathematics (Coordinated Control and Dynamic Optimal Dispatch of Islanded Microgrid System Based on GWO), *Journal of Mathematics*, 20–22, 2020.

[22] Vaclav Skala; Martin Cervenka. (2019). Novel RBF Approximation Method Based on Geometrical Properties for Signal Processing with a New RBF Function: Experimental Comparison, *IEEE 15th International Scientific Conference on Informatics*, 2019.

[23] Vivek Yadav; Girish Parmar; Rajesh Bhatt. (2019). Robustness Analysis with Perturbation for Control System with GWO, *2019 4th International Conference on Information Systems and Computer Networks (ISCON)*, 2019.

[24] YouWei Liu; ShaosSheng Fan; Yong Feng; LiJun Tang. (2019). Stockbridge Damper Identification Of Overhead Power Lines Based On HOG Feature And GWO-SVM, *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, 2019.

[25] Alisha Ahmed; Girish Parmar; Rajeev Gupta. (2018). Application of GWO in Design of Fractional Order PID Controller for Control of DC Motor and Robustness Analysis, *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2018.

[26] Allan Christian Krainski Ferrari; Gideon Villar Leandro; Leandro dos Santos Coelho; Carlos Alexandre Gouvea da Silva. (2019). Tuning of Control Parameters of Grey Wolf Optimizer using Fuzzy Inference, *IEEE Latin America Transactions*, 17(07), 1191–1198, 2019.

[27] Xu Liang; Di Wang; Ming Huang. (2019). Improved Grey Wolf Optimizer and Their Applications, *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, 2019.

[28] Kartikeya Jaiswal; Himanshu Mittal; Sonia Kukreja. (2017). Randomized Grey Wolf Optimizer (RGWO) with Randomly Weighted Coefficients, *2017 Tenth International Conference on Contemporary Computing (IC3)*, 2017.

[29] Riccardo Taormina; Nils Ole Tippenhauer; Stefano Galelli; Elad Salomons. (2018). The Battle of The Attack Detection Algorithms: Disclosing Cyber Attacks On Water Distribution Netwoks, *Water Resources Planning and Management*, 2018.

[30] Siwar Kriaa; Ludovic Pietre-Cambacedes; Marc Bouissou; Yoran Halgand. (2015). A Survey of Approaches Combining Safety and Security for Industrial Control Systems, *Reliability Engineering and System Safety*, 139(5), 156–178, 2015.

[31] Lin Zhang; YanWen Huang; Jie Xuan; Xiong Fu; QiaoMin Lin; RuChuan Wang. (2021). Trust Evaluation Model Based on PSO and LSTM for Huge Information Environments, *Chinese Journal of Electronics*, 30(01), 92–101, 2021.

[32] Yue Hu; YaFeng Wang; HaoCheng Wang. (2020). A Decoding Method Based on RNN for OvTDM, *Chinese Communication*, 17(04), 1–10, 2020.

[33] Bo Han; LiNa Qiao; JingLin Chen; XianDa Zhang; YanXia Zhang; YongHeng Zhao. (2021). Genetic KNN: A Weighted KNN Approach Supported by Genetic Algorithm for Photometric Redshift Estimation of Quasars, *Research in Astronomy and Astrophysics*, 21(01), 167–179, 2021.

[34] Mohammad Hashem Haghighat; Jun Li. (2021). Intrusion Detection System Using Voting-Based Neural Network, *Tsinghua Science and Technology*, 26(04), 484–495, 2021.

*Cite this paper as:*

Xuejun L.; Li K.; Wang W.; Yan Y.; Sha Y.; Chen J.; Qin J. (2021). Improved RBF Network Intrusion Detection Model Based on Edge Computing with Multi-algorithm Fusion, *International Journal of Computers Communications & Control*, 16(4), 4232, 2021.

https://doi.org/10.15837/ijccc.2021.4.4232