Received November 24, 2019, accepted December 5, 2019, date of publication December 13, 2019, date of current version December 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2959699

Power System Anomaly Detection Based on OCSVM Optimized by Improved Particle Swarm Optimization

ZHONGFENG WANG [1,2,3], YATONG FU [1,2,3,4], CHUNHE SONG [1,2,3], PENG ZENG [1,2,3], AND LIN QIAO [5]

[1] Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China
[2] Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China
[3] Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China
[5] State Grid Liaoning Electric Power Company, Ltd., Shenyang 110000, China

Corresponding author: Chunhe Song (songchunhe@sia.cn)

This work was supported in part by the State Grid Corporation Science and Technology Project under Contract SGLNXT00YJJS1800110.

ABSTRACT This paper tries to solve anomaly detection, a very important issue in ensuring the safe and stable operation of power system. As the proportion of abnormal data in the operation of power system is very small, a one-class support vector machine (OCSVM) is adopted in this paper, which is suitable for classification of unbalanced data. However, the performance of OCSVM is sensitive to its parameters, and an unsuitable choice will decrease the classification accuracy and generalization ability of it. In this paper, particle swarm optimization (PSO) is used to optimize the parameters of OCSVM. The original PSO algorithm converges slowly and easily falls into local optimum. To overcome this issue, this paper proposes an improved PSO algorithm for parameters optimization, in which adaptive speed weighting and adaptive population splitting are introduced to improve the convergence speed of the algorithm and help the algorithm jump out of the local optimal position. Experiments on standard benchmarks and real power system experimental data sets demonstrate the effectiveness of the proposed algorithm.

INDEX TERMS Power system anomaly detection, one-class support vector machine, particle swarm optimization, adaptive speed weighting, adaptive population splitting.

I. INTRODUCTION

With the rapid development of information technology, the amount of data generated by various industries has increased dramatically [1]. As society enters the era of big data, data becomes an important asset and competitive resource for the industry. The massive data generated by the power system has promoted the intelligent development of the power grid industry. The concept of ''Smart Grid'' and ''Internet of Things'' and so on has brought unprecedented opportunities and challenges to the power system. In order to obtain the industry gain brought by massive power data, data processing technologies such as data analysis and data mining are introduced into the power system. Anomaly detection is an important part of the data processing, which is a crucial step in data analysis and mining, and has attracted wide attention of researchers [2]. The data set of abnormal points have great research value, play an important role in areas such as equipment maintenance, credit fraud, cyber-attacks and etc. By detecting abnormal data, faulty equipment can be identified, network attacks can be prevented, operating costs can be saved, operational risks of the power systems can be reduced, and power system safety and reliability can be improved.

In the field of anomaly detection, due to the immature development of the industry in the early stage, anomaly detection of the power system mostly uses manual detection, which wastes a lot of time and human resources. With the continuous development of data processing technology, many efficient and available anomaly detection algorithms have emerged, which improves the detection accuracy in the power systems and reduces the economic loss of power systems.

The associate editor coordinating the review of this manuscript and approving it for publication was Min Xia.

At present, the common anomaly detection algorithms can be roughly divided into four categories: statistics-based anomaly detection algorithms, density-based anomaly detection algorithms, distance-based anomaly detection algorithm, and the model-based anomaly detection algorithm. Each algorithm has different application fields according to different algorithm principles, and should be flexibly select different anomaly detection algorithms according to different application scenarios. At present, in the field of anomaly detection, it is a cutting-edge research method to use machine learning related algorithms to model data. However, the selection of the anomaly detection algorithm does not mean that the anomaly detection problem can be completely solved. The parameters in the algorithm often affect the time and accuracy of anomaly detection.

Different from general anomaly detection, the proportion of abnormal data in the operation of power system is very small, making it an issue of unbalanced data classification. In order to improve the stability of the power system operation and reduce the operating cost of the system, and improve the accuracy and efficiency of anomaly detection, this paper adopts one-class support vector machine (OCSVM) as anomaly detection model, which is suitable for classification of unbalanced data. Meanwhile, to solve the sensitivity problem of parameters selection, an improved particle swarm optimization algorithm (PSO) is adopted to optimize OVSVM.

The main contributions of this paper are as follows:

(1) The principles and characteristics of SVM and OCSVM are analyzed, and the issue of unbalanced data classification is analyzed.
(2) An improved PSO algorithm is proposed for parameter optimization, in which adaptive speed weighting and adaptive population splitting are introduced to improve the convergence speed of the algorithm and help the algorithm jump out of the local optimum.
(3) A novel algorithm of OCSVM optimized by PSO is proposed, which improves the classification accuracy of OCSVM.
(4) Experiments on standard benchmarks and real power system experimental data sets demonstrate the effectiveness of the proposed algorithm.

The rest of the paper is organized as follows. Section II briefly reviews related works. Section III describes the details of one-class support vector machines. The improved particle swarm optimization algorithm is shown in Section IV. The experiments and results are shown in Section V. conclusion and future works are given in Section VI.

II. POWER SYSTEM ANOMALY DETECTION BASED ON OCSVM

In the field of anomaly detection, researchers have done a lot of work and achieved fruitful results. Currently, there are many algorithms that can be used to anomaly detection. Liu et al. proposed an anomaly detector deployment awareness detection framework based on multi-dimensional resources balance, and experiments show that this framework achieves a higher scalability and detection accuracy [3]. Shakya et al. show a weighted hybrid model utilizing Support Vector Machine and Naive Bayes for anomaly discovery [4]. Liu et al. proposed a dynamic and adaptive anomaly detection algorithm based on Self-Organizing Maps (SOM) for virtual machines [5]. Qin and Lou proposed a hydrological time series anomaly pattern detection algorithm based on isolation forest [6].

In this paper, support vector machine (SVM) is used as the anomaly detection model. SVM is a commonly used unsupervised classification method in machine learning. Its basic model is a linear classifier with the largest interval defined in feature space. The learning strategy is to maximize the interval so that data sets can be classified correctly. Because of its rigorous theoretical basis and mathematical deduction process, support vector machine can transform the binary classification problem into convex quadratic programming optimization problem, which is widely used in classification and regression problems.

Since Cortes and Vapnik officially proposed support vector machine in 1995, SVM has witnessed rapid development and many research results have been obtained. In 2018, Dai proposed a weighted Euclidean distance, radial integral kernel function SVM and dimensionality reduction algorithm for large data packet classification, which can effectively shorten the modeling time and improve the classification accuracy [7]. In the field of one-class support vector machine, Scholkopf et al. first proposed the concept of one-class SVM in [8], which laid the foundation for the application and research of one-class SVM. Huang improves the one-class support vector machine, and propose an online version of v-OeSVM [9]. Experiments on toy and real datasets show that v-OeSVM is a good mean to target a given false alarm rate while the AUC increases slowly as the number of new samples increases. Shahbudin et al. applied one-class SVM to weed classification, in which feature vectors of weed images were extracted using Gabor Wavelet and Fast Fourier Transform [10].

Since the detection of abnormal points in real life is an unsupervised problem, and the ratio of abnormal data to normal data is extremely unbalanced, OCSVM has a special advantage in solving such unsupervised sample imbalance problems. For SVM, its purpose is to find a separate hyper-plane that is capable of correctly classifying the sample set and having the largest geometric spacing. For a one-class SVM, there is only one type of point in the data set, so the origin in the feature space is used as the abnormal point, its purpose is to find a hyper-plane that wraps the sample points tightly and keeps them away from the origin as far as possible. As shown in the figure below, the white point in the figure is the normal point, and the black point is the abnormal point. In order to maximize the support vector set interval, the classification hyper-plane found by SVM is shown by the

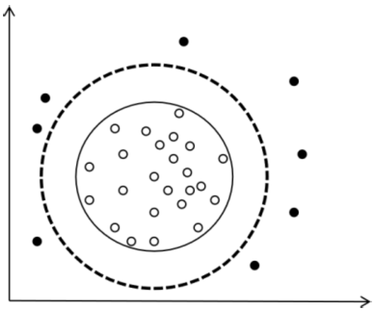

FIGURE 1. SVM and OCSVM classification problem.

black dotted line. For one-class SVM, the hyper-plane found by the one-class SVM is shown by the solid black line.

There are n sample points in the space $\{x_1, x_2 \ldots x_n\}$, and there is a decision function $f(x) = sign(\omega \cdot \Phi(x) - \rho)$ to construct a hyper-plane to correctly classify the data set. ω is a normal vector to the hyper plane, $\Phi(x)$ map points on the sample space to the feature space, ρ is the compensation vector. Then the objective function and constraint conditions of OCSVM are as follows:

$$\min: \frac{1}{2}\|\omega\|^2 + \frac{1}{v \cdot n}\sum_{i=1}^{n}\xi_i - \rho \tag{1}$$

$$st: \omega \cdot \Phi(x_i) \geq \rho - \xi_i, \quad \xi_i > 0 \tag{2}$$

where ρ is the trade-off parameter.

Define the Lagrangian function:

$$\frac{1}{2}\|\omega\|^2 + \frac{1}{v \cdot n}\sum_{i=1}^{n}\xi_i - \rho - \sum_{i=1}^{n}(\omega \cdot \Phi(x_i) - \rho + \xi_i)\alpha_i - \sum_{i=1}^{n}\xi_i\beta_i \tag{3}$$

α, β is a Lagrange multiplier vector:

$$\alpha = (\alpha_1, \alpha_2 \ldots \alpha_n)^T, \quad \beta = (\beta_1, \beta_2 \ldots \beta_n)^T.$$

The above problem is transformed into its dual problem, and the kernel function $K(x_i, x_j)$ is introduced:

$$\min: \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i \cdot \alpha_j K(x_i, x_j) \tag{4}$$

$$st: 0 \leq \alpha_i \leq \frac{1}{v \cdot n}, \quad \sum_{i=1}^{n}\alpha_i = 1 \tag{5}$$

Finally:

$$\rho = \sum_{i=1}^{n}\alpha_i \cdot K(x_i, x_j) \tag{6}$$

$$f(x) = sign(\sum_{i=1}^{n}\alpha_i \cdot K(x_i, x_j) - \rho) \tag{7}$$

Commonly used kernel functions are linear kernel functions, polynomial kernel functions and Gaussian kernel functions. The support vector machines generally use Gaussian kernel functions to map data sets to high-dimensional spaces:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \tag{8}$$

Therefore, for one-class support vector machine, parameters v and σ need to be selected to determine the one-class support vector machine model.

III. PARAMETER OPTIMIZATION FOR OCSVM

Currently there are many algorithms such as the genetic algorithm (GA), the grid method, and PSO can be used to optimize the parameters in OCSVM. GA was first proposed by John Hollander of the University of Michigan and his colleagues in the research of cellular automata in the 1960s. In recent years, genetic algorithm has been greatly developed. Wei et al. studied the mutation operator in depth, and discussed how to solve the maximum problem by genetic algorithm [11]. The particle swarm algorithm was first proposed by Kennedy and Eberhart in 1995 in [12]. In 2018, Duan et al. introduced particle swarm optimization to solve the parameter selection problem of support vector machine prediction model [13]. The optimal prediction model is obtained by choosing the appropriate learning parameters using the swarm algorithm. In the improvement of particle swarm optimization, Xue et al. introduces a Time-Varying Inertia Weighting (TVIW) strategy based on a combination of Gravitation Search Algorithm and Particle Swarm Optimization called the TVIW-PSO-GSA. It effectively improves the classification accuracy and efficiency of the classifier [14]. In addition to the above classical algorithms, some swarm intelligence algorithms can also be used for parameter optimization. Gao proposed an improved hybrid swarm intelligence algorithm and carried out simulation experiments [15], compared with particle swarm optimization and artificial bee colony algorithm, the convergence speed and accuracy of hybrid algorithm are better than standard algorithm. Xiao et al. put forward a kind of salp swarm algorithm based on particle optimization [16], aiming at the problem that the global search ability and local development ability of the PSO algorithm are difficult to coordinate and easy to fall into local optimization. Compared with other algorithms, the algorithm has a great improvement in convergence speed, accuracy and robustness. Deng et al. proposed an improved ant colony algorithm for traveling salesman problem [17]. Experiments show that the algorithm has good performance and robustness. OCSVM is widely used in various fields. Miao et al. is used for anomaly data recognition in wireless sensor networks [18]. Li et al. uses OCSVM for fault detection in closed-loop systems [19]. In the field of parameter optimization of OCSVM, Anaissi *et al.* proposed a new algorithm, called edge support vector [20], to adjust parameters of Gauss model. Deng *et al.* used genetic algorithm to optimize parameters [21].

A. PARTICLE SWARM OPTIMIZATION

In this paper PSO is used to optimize the parameters of OCSVM. The PSO algorithm first randomly initializes the position and velocity of the particles, and then calculates the fitness value of each particle. The speed and position of each particle are updated according to the global optimal value of

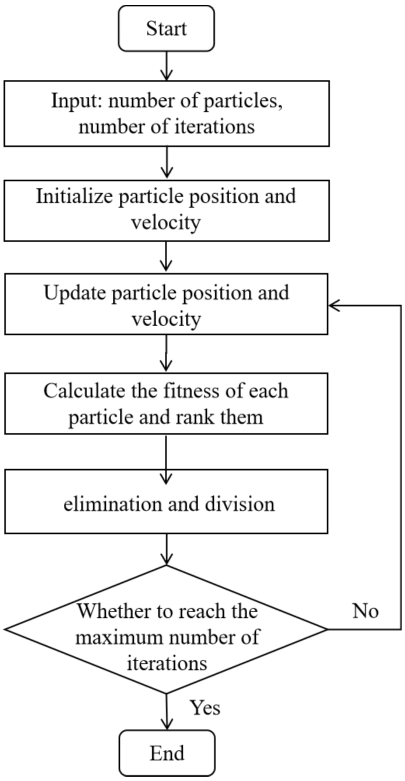

FIGURE 2. The flow chart of the improved particle swarm optimization algorithm.

the population and the historical optimal value of the particle, keep the above process iterating until the stopping condition is satisfied. Suppose there is a search space of n-dimensional. If there are m particles in the particle swarm and the particle set is $D = \{\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3 \cdots \mathrm{x}_\mathrm{m}\}$, the position of the i-th particle in the particle group is recorded as $x_i = (x_{i1}, x_{i2}, x_{i3} \cdots x_{in})$, and the moving speed of the i-th particle in the particle group is recorded as $v_i = (v_{i1}, v_{i2}, v_{i3} \cdots v_{in})$. The optimal position currently searched for the i-th particle in the particle group is denoted as $x_{ei} = (x_{ei1}, x_{ei2}, x_{ei3} \cdots x_{ein})$, and the global optimal position in the particle group is denoted as $x_g = (x_{g1}, x_{g2}, x_{g3} \cdots x_{gn})$. Then the i-th particle in the particle group is updated as follows:

$$v_i = \omega_i v_i + c_1 r_1 (x_{ei} - x_i) + c_2 r_2 (x_g - x_i) \tag{9}$$

$$x_i = x_i + v_i \tag{10}$$

where ω_i is the inertia factor, c_1 c_2 is the acceleration factor, c_1 is the speed of approaching the optimal position of the individual, c_2 is the speed of approaching the global optimal position of the particle, r_1 r_2 is the random number between (0,1).

B. IMPROVED PARTICLE SWARM OPTIMIZATION

The particle swarm optimization algorithm converges slowly in the early stage and easily falls into local extremum in the later stage. In the current research of adaptive particle swarm optimization algorithm, most of the research methods are based on the number of iterations to improve the inertia factor or learning factor longitudinally, and the means of improvement are relatively simple and lack of reasonable comparative experiments. In order to improve the convergence speed and accuracy of the particle swarm optimization algorithm, an improved particle swarm optimization algorithm is proposed, and two novel strategies, adaptive speed weighting and adaptive population splitting, are proposed to improve the convergence speed of the algorithm and help the algorithm jump out of the local optimum.

1) ADAPTIVE SPEED WEIGHTING

In order to solve the low accuracy of the particle swarm optimization algorithm and increase the diversity of the population, different learning factors are assigned to each particle according to the difference of the fitness value. In the original particle swarm optimization algorithm, the particle moves towards the historical optimal position of the particle with velocity c_1 and to the global optimal position with velocity c_2, where c_1 c_2 is a fixed constant and the general value is 2. However, the particle fitness values in the particle swarm are high or low, and the uniqueness of each particle is neglected by moving the fixed value to the historical optimal position and the global optimal position. Considering the difference in the fitness value of each particle, each particle is given a different acceleration factor to move to the historical optimal value and the global optimal value at different speeds. The lower the particle fitness, the better the particle position, the smaller the speed of moving to the global optimal position, and the greater the speed of moving to the historical optimal position. The higher the particle fitness, the worse the particle position, the greater the speed of moving to the global optimal position, and the lower the speed of moving to the historical optimal position.

Firstly, the fitness values of the particles in the particle swarm are sorted from small to large, and the fitness values of i-th particles in the population are ranked as s_i. Then the acceleration factors of the i-th particles are formulated as follows:

$$c_{1i} = {}^{R}\!/_{n} \times s_i \tag{11}$$

$$c_{2i} = R - c_{1i} \tag{12}$$

where R is a constant, and R is taken as 2 in this paper.

The update formula for the i-th particle is:

$$v_i = \omega_i v_i + c_{1i} r_1 (x_{ei} - x_i) + c_{2i} r_2 (x_g - x_i) \tag{13}$$

$$x_i = x_i + v_i \tag{14}$$

2) ADAPTIVE POPULATION SPLITTING

In order to improve the convergence speed of particle swarm optimization and the fine search ability near the optimal fitness value, population splitting and elimination mechanism are introduced.

The higher the fitness of the particle in the population, the poorer the position of the particle, the lower the possibility of searching for the optimal fitness near the particle.

Sorting the population particles from high to low according to the fitness, and eliminating a part of the particles with higher fitness in the population, not only has less influence on the optimal value of the population search, but also the convergence speed can be effectively improved.

Particles with lower fitness value in the population have better location, and the probability of finding the optimal fitness value near them is higher. Therefore, splitting operation should be carried out on these particles, to enhance the fine search ability of particles in the region.

The algorithm steps are as follows.

Step1: Initialize the number of particle groups m, the particle swarm is recorded as $D = \{\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3 \cdots \mathrm{x}_\mathrm{m}\}$, set the number of iterations to k, and for $x_i \in D$, the initial position $x_i = (x_{i1}, x_{i2}, x_{i3} \cdots x_{in})$ and the velocity $v_i = (v_{i1}, v_{i2}, v_{i3} \cdots v_{in})$ of the particle x_i are set;

Step 2: Calculate the fitness value $f(x_i)$ of the particles in the particle swarm and sort the particle fitness values from small to large;

Step3: Carry out population elimination and splitting work.

Step4: Update the optimal position of the individual particles x_{ei} and the optimal position of the particle swarm x_g. Update the particle velocity and position in the particle swarm.

Step 5: Determine whether the maximum number of iterations has been reached.

Step 6: When the number of iterations reaches the maximum number of iterations or the optimal fitness value of the particle group reaches the ideal fitness value, the algorithm ends, otherwise it returns to Step 2.

IV. EXPERIMENTAL SIMULATION

A. EXPERIMENTAL DATA

The simulation data of this experiment comes from ECO data set (Electricity Consumption & Occupancy) [22]. The ECO data set is a comprehensive data set for non-intrusive load monitoring and occupancy detection research. It was collected in 6 Swiss households over a period of 8 months. For each of the households, the ECO data set provides:

(1) 1 Hz aggregate consumption data. Each measurement contains data on current, voltage, and phase shift for each of the three phases in the household.

(2) 1 Hz plug-level data measured from selected appliances.

(3) Occupancy information measured through a tablet computer (manual labeling) and a passive infrared sensor (in some of the households).

In order to facilitate the experimental verification and compare the performance of each algorithm reasonably, this paper modifies the original data and labels the abnormal data manually. 1 represents the normal data, and -1 represents the difference. Descriptions of part of data are given in Table 1 as follows

TABLE 1. partial data of ECO data set.

	1	2	3	4
power_all_phases	113.79	112.64	112.65	113.02
powerl1	74.95	75.15	74.87	75.27
powerl2	28.24	28.16	28.6	29.22
powerl3	9.28	9.51	10.02	9.49
powerl3_change	-1.93	9.51	10.02	9.49
y	-1	1	1	1

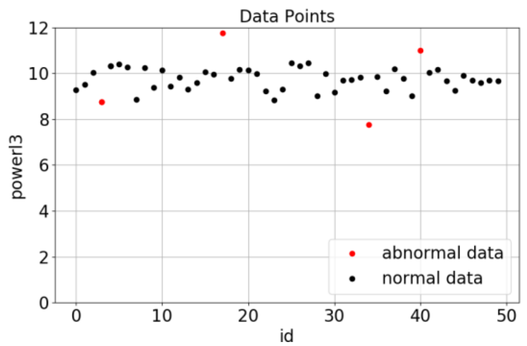

FIGURE 3. Partial data from powerl3_change.

Experimental data is listed as powerl3, and some points are shown as follows, where red points are abnormal points and black points are normal points:

In order to compare the performance of each algorithm in anomaly recognition, three indexes are used to measure the algorithm performance.

The recall rate:

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

The accurate rate is measured as:

$$P = \frac{TP}{TP + FP} \tag{16}$$

$$F1 - measure : f1 = {}^{2P \cdot R}/_{P + R} \tag{17}$$

where TP is the number of positive classes predicted to be positive, FP is the number of negative classes predicted to be positive, and FN is the number of positive classes predicted to be negative.

B. EXPERIMENTAL RESULTS ON BENCHMARKS

In order to verify the effectiveness of the proposed improved particle swarm optimization algorithm, a comparative experiment is performed below. The number of particle swarm $m = 100$, the maximum number of iterations $k = 100$, the optimization interval of parameters A and B are (0,1).

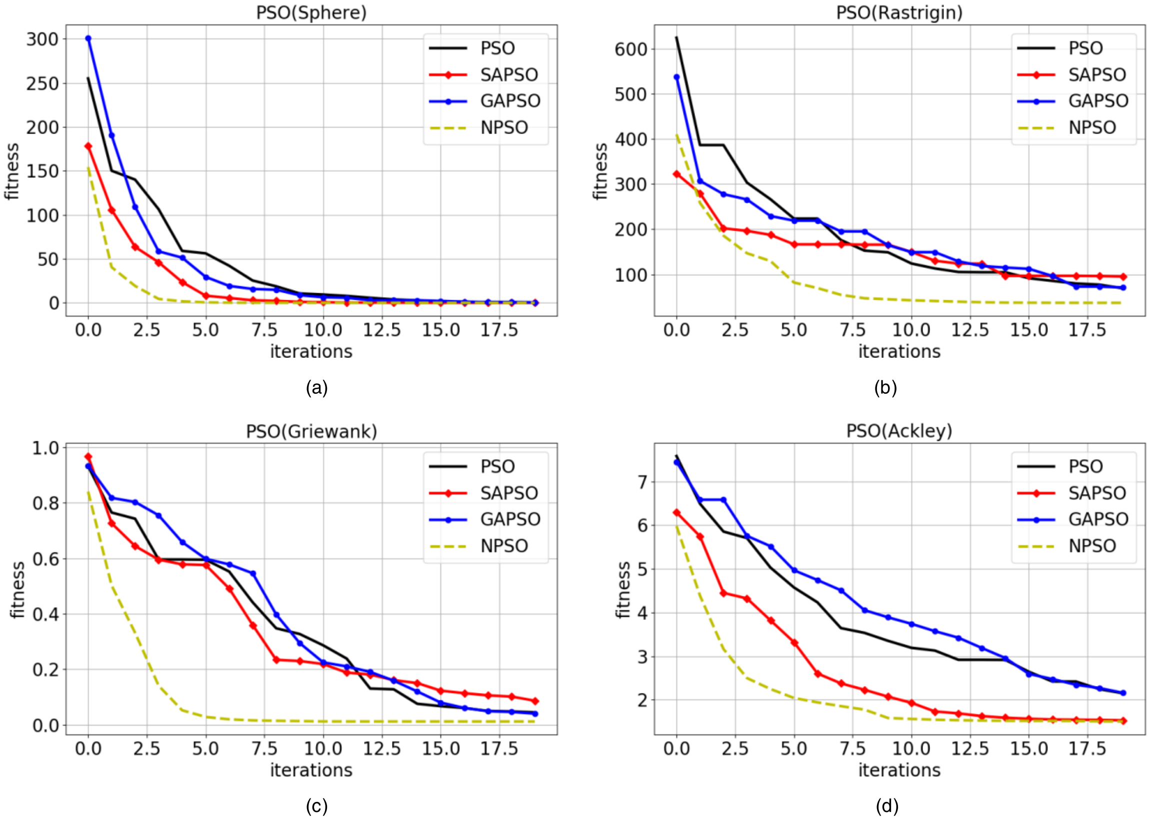

FIGURE 4. **Contrast experiment of improved particle swarm optimization algorithm. (a) Sphere function, (b) Rastrigin function, (c) Griewank function, (d) Ackley function.**

The initial position and velocity of particle swarm optimization are random values, but in order to objectively compare the performance of other particle swarm optimization algorithms and improved particle swarm optimization algorithms, the initial values of particles are the same set of random values. In order to verify the effectiveness of the proposed algorithm, the algorithm is compared with two improved particle swarm optimization algorithms [23], [24]. In this paper, the following four benchmarks are used as the fitness functions. The benchmarks used in Section V.B have multiple local minimums, and the locations of these minimums are close with each other. Therefore, such benchmarks can be used to test the capacity of finding the global minimum of an algorithm.

(1) Sphere function:

$$f(x) = \sum_{i=1}^{i=n} x_i^2 \tag{18}$$

(2) Rastrigin function:

$$f(x) = \sum_{i=1}^{n} (x_i^2 - 10\cos(2\pi x_i) + 10) \tag{19}$$

(3) Griewank function:

$$f(x) = \sum_{i=1}^{n} {x_i^2}\Big/{4000} - \prod_{i=1}^{n} \cos({x_i}\Big/{\sqrt{i}}) + 1 \tag{20}$$

(4) Ackley function:

$$\begin{aligned} g(x) = &-20\exp(-0.2\sqrt{{}^{1}\!/_{n}\sum_{i=1}^{n} x_i^2)} \\ &- \exp({}^{1}\!/_{n}\sum_{i=1}^{n}\cos(2\pi x)) \end{aligned} \tag{21}$$

$$f(x) = g(x) + 20 + e \tag{22}$$

The original algorithm is named PSO, the proposed PSO algorithm in this paper is named NPSO, the improved PSO algorithm proposed in [23] is named SAPSO, and the improved PSO algorithm proposed in [24] is named GAPSO. The experimental results are shown in Figure 4.

It can be seen from the Fig.4a that, although all four PSO algorithms converge to the global minimal location on benchmark Sphere, the proposed NPSO has a much faster convergence speed. Meanwhile, from Fig.4b, Fig.4c and Fig.4d it can be seen that, for benchmarks Rastrigin, Griewank and Ackley, the proposed NPSO does not only has faster convergence speeds, but also achieves the best results among all four PSO algorithms. The main reason of it is that adaptive speed weighting and adaptive population splitting in the proposed NPSO can help the algorithm jump out of the local optimum location and find a better solution while the other three algorithms are trapped in the local optimum location. In summary,

TABLE 2. Anomaly detection algorithms.

statistics-based anomaly detection algorithm	pauta criterion	box_plot
density-based anomaly detection algorithm	density-based spatial clustering of applications with noise(DBSCAN)	local anomaly factor(LOF)
distance-based anomaly detection algorithm	k-nearest neighbor(KNN)	knn-weight(KNNW)
model-based anomaly detection algorithm	Isolation Forest(Iforest)	one-class support vector machine (OCSVM)

TABLE 3. Anomaly detection experiment in power system.

	Accuracy_rate	Recall_rate	F1_measure
3sigma	62.50%	2.70%	5.18%
box_plot	70.42%	72.16%	71.27%
KNN	30.45%	47.57%	37.13%
KNNW	29.33%	47.57%	36.29%
DBSCAN	58.42%	91.89%	71.43%
LOF	47.00%	25.41%	32.98%
OCSVM	66.67%	88.64%	76.10%
Iforest	67.76%	81.73%	74.09%

the proposed algorithm has a faster convergence speed and can overcome the drawback that the general algorithm is vulnerable to the local optimum location.

C. EXPERIMENTAL RESULTS ON THE REAL DATASET

As mentioned above, the anomaly detection algorithms can be roughly divided into four types, and the representative algorithms of each type are summarized in TABLE 2. The above algorithms are used to detect abnormal data of power system. The accuracy, recall rate and F1_measure value of each algorithm are tested. The experimental results are shown in TABLE 3.

From TABLE 3, it can be seen that the accuracy of each anomaly detection algorithm is relatively low. The values of OCSVM, DBSCAN, box_plot and Iforest can be greater than 70%, and OCSVM has the highest recognition accuracy. In order to improve the accuracy of anomaly detection, the following two processes are carried out:

1) SELECT EIGENVALUE

In the experiment of anomaly detection, the original data is input directly as the eigenvalue.

In order to improve the performance of the anomaly detection algorithm, the feature space is reconstructed in this paper. Assume that the data set has n pieces of data, that is $D = \{x_1, x_2, x_3 \cdots x_n\}$, There are m attributes, $x_i = \{x_{i1}, x_{i2}, x_{i3} \cdots x_{im}\}$, the $i-th$ attribute value of the $j-th$ data is recorded as x_{ij}, construct a sub-sample set of sample points x_{ij} about attribute j, that is $N_k(x_{ij}) = \{x_{(i-k)j}, x_{(i-k+1)j} \cdots x_{(i+k-1)j}, x_{(i+k)j}\} - x_{ij}$, Construct a sub-sample set of sample point x_{ij} with respect to item i data, $N(x_{ij}) = \{x_{i1}, x_{i2} \cdots x_{im}\} - x_{ij}$. Construct the prediction equation of sample point x_{ij}, that is $f(N(x_{ij}))$. Construct two features for sample point x_{ij}:

$$x'_{ij} = \frac{\sum_{x_{kj} \in N_k(x_{ij})} x_{kj}}{\left|N_k(x_{ij})\right|} \tag{23}$$

$$x''_{ij} = \left|x_{ij} - f(N(x_{ij}))\right| \tag{24}$$

2) PARAMETER OPTIMIZATION

Using the F1_measure value as the fitness function on the test data set, the optimal parameters obtained by the improved particle swarm optimization algorithm are as follows:

$$v = 0.001, \quad \sigma = 0.047.$$

The parameters were brought into the one-class support vector machine for anomaly detection. Before performing the parameter optimization operation, the accuracy rate, recall rate, and F1_measure value were 66.67%, 88.64%, and 76.10%, respectively. After the parameter optimization operation, the accuracy rate, recall rate, and F1_measure value increased by 26.14%, 2.17% and 15.70%. Finally, the accuracy rate, recall rate, and F1_measure values are 92.81%, 90.81% and 91.80%, respectively. The experimental results show that the proposed algorithm can improve the performance of the algorithm in anomaly detection effectively.

To compare the performance of the two algorithms, the ROC curves of each algorithm can be drawn in the same coordinate to identify the advantages and disadvantages intuitively. The ROC curves near the upper left corner represent the most accurate algorithm. The ROC curve is based on a series of different two-category methods, with a true positive rate as the ordinate and a false positive rate plotted on the abscissa. The area under ROC curve (AUC) of each algorithm can also be calculated and compared. Which algorithm has the largest AUC and which one has the best diagnostic value.

The anomaly detection algorithm of one-class support vector machine optimized by parameters is recorded as NPSO-OCSVM. The ROC curves is shown in the following figure. As can be seen from Figure 5, the ROC curve of NPSO-OCSVM is closer to the upper left corner and more

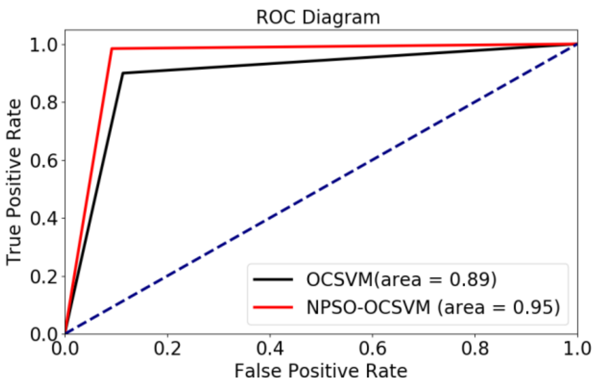

FIGURE 5. **ROC curves.**

deviated from the diagonal than OCSVM. This indicates that NPSO-OCSVM has a higher classification accuracy rate. And the AUE value of OCSVM is 0.89, and that of NPSO-OCSVM is 0.95. It can be seen that the quality of the model has been significantly improved after the improved particle swarm optimization.

V. SUMMARY

In this paper, OCSVM is used to solve the problem of power system anomaly detection. In order to improve the classification accuracy of OCSVM, an improved particle swarm optimization algorithm is proposed, in which the adaptive learning factor is introduced to increase the diversity of the population and improve the searching ability of the population. Furthermore, the splitting and elimination mechanism is introduced to improve the fine searching ability of the population at the optimal fitness value, and to improve the convergence speed of the particle swarm optimization algorithm. Compared with the several other anomaly detection algorithms, the anomaly detection algorithm based on OCSVM optimized by the improved PSO algorithm achieves satisfactory results in recall rate, accuracy rate and F1_measure value.

REFERENCES

[1] X. HongJu, W. Fei, W. FenMei, and W. XiuZhen, "Some key problems of data management in army data engineering based on big data," in *Proc. ICBDA*, Mar. 2017, pp. 149–152.

[2] Z. Guan, T. Ji, X. Qian, Y. Ma, and X. Hong, "A survey on big data preprocessing," in *Proc. ACIT-CSII-BCD*, Jul. 2017, pp. 241–247.

[3] J. Liu, H. Zhang, and G. Xu, "An anomaly detector deployment awareness detection framework based on multi-dimensional resources balancing in cloud platform," *IEEE Access*, vol. 6, pp. 44927–44933, 2018.

[4] S. Shakya and S. Sigdel, "An approach to develop a hybrid algorithm based on support vector machine and naive Bayes for anomaly detection," in *Proc. ICCCA*, May 2017, pp. 323–327.

[5] J. Liu, S. Chen, Z. Zhou, and T. Wu, "An anomaly detection algorithm of cloud platform based on self-organizing maps," *Math. Problems Eng.*, vol. 2016, Mar. 2016, Art. no. 3570305.

[6] Y. Qin and Y. Lou, "Hydrological time series anomaly pattern detection based on isolation forest," in *Proc. ITNEC*, Mar. 2019, pp. 1706–1710.

[7] H. Dai, "Research on SVM improved algorithm for large data classification," in *Proc. ICBDA*, Mar. 2018, pp. 181–185.

[8] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, and J. Smola, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443–1471, 2001.

[9] Y. Xue and P. Beauseroy, "Constant false alarm rate for online one class SVM learning," in *Proc. ICASS*, Apr. 2018, pp. 2821–2825.

[10] S. Shahbudin, M. Zamri, M. Kassim, S. A. C. Abdullah, and S. I. Suliman, "Weed classification using one class support vector machine," in *Proc. ICEESE*, Nov. 2017, pp. 7–10.

[11] P. Wei, Y. Mingming, T. Yuejun, and X. Huixiang, "Micro variation chaos genetic algorithm," in *Proc. CCDC*, Jun. 2018, pp. 4523–4528.

[12] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory,"in *Proc. MHS*, Oct. 1995, pp. 39–43.

[13] M. Duan, "Short-time prediction of traffic flow based on PSO optimized SVM," in *Proc. ICITBS*, Jan. 2018, pp. 41–45.

[14] H. Xue, Y. Bai, H. Hu, T. Xu, and H. Liang, "A novel hybrid model based on TVIW-PSO-GSA algorithm and support vector machine for classification problems," *IEEE Access*, vol. 7, pp. 27789–27801, 2019.

[15] Y. Gao, "An improved hybrid group intelligent algorithm based on artificial bee colony and particle swarm optimization," in *Proc. ICVRIS*, Aug. 2018, pp. 160–163.

[16] B. Xiao, R. Wang, Y. Xu, W. Song, J. Wang, and Y. Wu, "Salp swarm algorithm based on particle-best," in *Proc. ITNEC*, Mar. 2019, pp. 1383–1387.

[17] X.-Y. Deng, W.-L. Yu, and L.-M. Zhang, "A new ant colony optimization with global exploring capability and rapid convergence," in *Proc. WCICA*, Jul. 2012, pp. 579–583.

[18] X. Miao, Y. Liu, H. Zhao, and C. Li, "Distributed online one-class support vector machine for anomaly detection over networks," *IEEE Trans. Cybern.*, vol. 49, no. 4, pp. 1475–1488, Apr. 2019.

[19] Z. Li and X. Li, "Fault detection in the closed-loop system using one-class support vector machine," in *Proc. DDCLS*, May 2018, pp. 251–255.

[20] A. Anaissi, A. Braytee, and M. Naji, "Gaussian Kernel parameter optimization in one-class support vector machines," in *Proc. IJCNN*, Jul. 2018, pp. 1–8.

[21] X. Deng, P. Jiang, X. Peng, and C. Mi, "An intelligent outlier detection method with one class support tucker machine and genetic algorithm toward big sensor data in Internet of Things," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4672–4683, Jun. 2019.

[22] R. Cicchetti, "NILM-Eval: Disaggregation of real-world electricity consumption data," M.S. thesis, ETH Zürich, 2014.

[23] L. Xu, C. Xin, D. Niu, Q. An, C. Yuan, and Y. Xiao, "Power consumption prediction model of least squares support vector machine based on adaptive Particl," *Sci. Technol. Eng.*, vol. 19, no. 6, pp. 136–141, 2019.

[24] Y. Wang, H. Mu, and G. Xu, "Prediction of output power of photovoltaic cells based on the improved particle swarm optimization support vector machine," *Electr. Autom.*, vol. 41, no. 3, pp. 63–65, 2019.

ZHONGFENG WANG received the Ph.D. degree in mechatronic engineering from the Graduate School, Chinese Academy of Sciences, Beijing, China, in 2006. He is currently a Professor with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. His research interests include image processing, wireless sensor networks, and industrial communication technologies.

YATONG FU received the B.E. degree in automation from the University of Jinan, Jinan, China, in 2017. She is currently pursuing the M.S. degree with the Shenyang Institute of Automation, Shenyang, China. Her research interests include smart grid and edge computing.

CHUNHE SONG received the Ph.D. degree in computer science from Northeastern University, Shenyang, China, in 2012. He is currently a Professor with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. His research interests include image processing and edge computing.

PENG ZENG received the Ph.D. degree in mechatronic engineering from the Graduate School, Chinese Academy of Sciences, Beijing, China, in 2005. He is currently the Director of the Industrial Control Network and System Department, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China. He is also the Vice Director of the Network Control System Key Laboratory, Chinese Academy of Sciences. He has authored or coauthored more than 100 conference papers, journal articles, and book chapters related to his research area and five pending/filed U.S. patents. His research interests include image processing, wireless sensor networks, and industrial communication technologies.

LIN QIAO received the master's degree from North China Electric Power University, Beijing, China. He is currently the Director of the Transportation Inspection Center and Control Center, State Grid Liaoning Electric Power Company, Ltd. His research interests include smart grid, image processing, and communication.