

Research Article

An Improved Hybrid Feature Selection Algorithm for Electric Charge Recovery Risk

Shenyi Qian, Yongsheng Shi , Huaiguang Wu, and Songtao Shang

School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, China

Correspondence should be addressed to Yongsheng Shi; yongsheng.shi@zzuli.edu.cn

Received 17 April 2020; Revised 30 June 2020; Accepted 7 July 2020; Published 28 August 2020

Academic Editor: Yang Li

Copyright © 2020 Shenyi Qian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to extract more information that affects customer arrears behavior, the feature extraction method is used to extend the low-dimensional features to the high-dimensional features for the warning problem of user arrears risk model of electric charge recovery (ECR). However, there are many irrelevant or redundant features in data, which affect prediction accuracy. In order to reduce the dimension of the feature and improve the prediction result, an improved hybrid feature selection algorithm is proposed, integrating nonlinear inertia weight binary particle swarm optimization with shrinking encircling and exploration mechanism (NBPSOSEE) with sequential backward selection (SBS), namely, NBPSOSEE-SBS, for selecting the optimal feature subset. NBPSOSEE-SBS can not only effectively reduce the redundant or irrelevant features from the feature subset selected by NBPSOSEE but also improve the accuracy of classification. The experimental results show that the proposed NBPSOSEE-SBS can effectively reduce a large number of redundant features and stably improve the prediction results in the case of low execution time, compared with one state-of-the-art optimization algorithm, and seven well-known wrapper-based feature selection approaches for the risk prediction of ECR for power customers.

1. Introduction

With the rapid development of global energy market, smart grid [1] in power industry has been built continuously, and the scale of information data accumulated by power system is becoming larger and larger. As the main income of power companies, electric tariff plays a decisive role in the development of power enterprises. However, in the whole process of power marketing, the risk of arrears' users has always existed, which hinders the development of power enterprises. Electric charge recovery (ECR) has always been a difficult problem that power supply enterprises need to solve urgently. It is also the most important management part of power meter reading, verification, and checking. Once the processing of ECR is delayed, it may have a bad impact on the charging result. It also causes power customers to occupy the funds of power enterprises, which is not conducive to the fund management of power companies. The main operating profit of power grid enterprises comes from ECR. Through the analysis of electric power clients'

paying behavior and customer's characteristics, the risk factors and risk levels of ECR are predicted and evaluated, which contribute to collecting electric charge, formulating effectively preventive measures in time, reducing management risks and safeguarding economic benefits of power enterprises. Therefore, accurate and reliable forecasting of arrear risks is an important reference in terms of determining the management of ECR.

In the past, some people adopted these methods of feature extraction [2–4], information entropy theory [5], combined support vector machine (SVM) with VIKOR method [6], artificial immune algorithm [7], and combined trust evaluation cloud with decision algorithm [8] and established a power arrears risk model, but these results predicted by these methods were not very good. In recent years, some scholars have used quantitative analysis [9], classification [10], clustering [11], ensemble [12], and improved algorithms [13, 14] to analyse arrears behavior of electricity users and improved prediction results. However, with the increase in huge quantities of data and the difficulty

in capital operating, power companies urgently need to adopt faster and more accurate data processing methods to predict the arrears risk of power users in the future.

In order to extract more information that affects customer arrears behavior, low-dimensional features are extended to high-dimensional features via the feature extraction method. However, many irrelevant and redundant features will reduce the accuracy of classification and raise the complexity of dimensions. Therefore, feature selection is a very effective solution [15]. Feature selection and classification methods are widely used in high-dimensional and multiclass data sets [16, 17], which can improve the accuracy of model prediction by removing irrelevant and redundant features. Generally, the feature selection process includes the following stages: selecting feature subsets, evaluating feature subsets, and verifying results. The purpose of this process is to remove unrelated or redundant features and generate a smaller optimal feature subset. Generally, feature selection methods are divided into filter, wrapper, and embedded [18]. For the filter algorithms, the inherent distribution characteristics of the data are used as the basis for feature selection. The process of selecting features by filter approaches is not related to the learner. The filter approaches can be further divided into single variable algorithms and multivariate algorithms [19]. The single variable algorithms are able to evaluate each feature individually, reducing the accuracy of classification. However, multivariate algorithms have the advantage of evaluating the correlation between features. The filter algorithms are independent of the classifier and have fast computation speed. However, there is no interaction between classification algorithm and features in the process of feature selection by filter approaches. The wrapper approaches rely on classification algorithms to evaluate the selected feature subsets, which can achieve a higher classification accuracy than filter methods [20]. However, wrapper algorithms have high computational complexity in feature selection of high-dimensional data sets. In embedded approaches, feature selection is directly integrated into the training process of learners [21], but embedded methods are more complex in concept because it is not easy to obtain better classification results by the improved classification model. In contrast, wrapper approaches can make use of the performance of machine learning algorithm as evaluation criteria for selecting features, making it more flexible and more effective analysis of high-dimensional data. In recent years, wrapper approaches have aroused much attention by solving feature selection problems and seeking global optimal solutions through heuristic algorithms.

In recent years, many researchers have successfully applied wrapper approaches to feature selection. Li and Yang [22] integrated OS-extreme learning machine (EOS-ELM) and binary Jaya-based feature selection to real-time transient stability assessment using the phasor measurement unit (PMU) data. Yang et al. [23] propose a novel binary Jaya optimization algorithm, which is integrated with the lambda iteration method to transform the dual objectives of economy and emission commitment into a single objective problem. Aslan et al. [24] proposed the JayaX binary

optimization algorithm replacing Jaya's solution update rules with XOR operator and compared the results with the latest algorithm, which can produce better quality results in binary optimization problems. The whale optimization algorithm (WOA) [25, 26] was proposed, using the wrapper-based method to reach the optimal subset of features and effectively improve the accuracy of classification. Houssein et al. [27] proposed an S-shaped binary whale optimization algorithm for feature selection. Tawhid and Ibrahim [28] used a feature selection based on binary whale optimization algorithm (BWOA) to solve the problem of feature selection. Rao et al. [29] used the artificial bee colony approach (ABC) based on the boosting decision tree model to improve the quality of the selected features. The method can accelerate the convergence speed and balance exploration and exploitation efficiently. Furthermore, binary artificial bee colony algorithm (BABCO) [30, 31] is used to select feature subset. S. Oreski and G. Oreski [32] used the genetic algorithm-based heuristic for feature selection in credit risk assessment. In the paper, the genetic algorithm (GA) applies the neural network model to select optimal subset of features and increases the results of classification in credit risk assessment. Chen et al. [33] introduced a heuristic feature selection approach into text categorization by using chaos optimization and GA. Shukla et al. [34] proposed a new hybrid feature subset selection framework based on binary genetic algorithm and information theory. The method accelerates the search of important feature subset. Emary et al. [35] used binary grey wolf optimization (BGWO) approaches to select feature subset. Al-Tashi et al. [36] introduced the feature selection method based on GWO for coronary artery disease classification. Several studies used feature selection methods based on particle swarm optimization (PSO) algorithm to search the feature space for optimal solutions [37–39]. PSO performs equally well or better than WOA, ABC, GA, and GWO in solving global optimization problems [40]. Therefore, PSO is a promising approach to many tasks, including feature selection.

Kennedy and Eberhart proposed a particle swarm optimization (PSO) algorithm in 1995 [41]. PSO has been widely applied to many fields due to its superior performance, such as neural network training [42], classifier design [43], clustering analysis [44], and network community [45]. Since many problems are discrete in practice, Kennedy and Eberhart proposed a binary particle swarm optimization (BPSO) algorithm in 1997 [46], which uses binary encoding form to solve the discrete optimal combination problem. The BPSO has also been widely used in many fields, such as knapsack problems [46–48], power systems [49, 50], data mining [51–55], and image processing [56]. Because the search range of particles cannot be dynamically adjusted, BPSO can easily fall into the local optimum and premature convergence with the decline of population diversity. In view of the shortcomings of BPSO, many scholars have proposed various improved methods. First, Chuang and other scholars used chaotic binary particle swarm optimization (CBPSO) algorithm [57–59] to chaotic mapping the inertia weight improving the performance of BPSO. However, the behavior of chaotic mapping does not exist at fixed points, periodic or

quasiperiodic orbits. Liu et al. [60] proposed BPSO with linear adaptive inertia weight, which improved the search performance of particles, but it often ignored the nearby optimal solution in the process of linear transformation of inertia weight. Wu et al. [61] proposed a feature selection algorithm based on hybrid improved binary quantum particle swarm optimization (HI-BQPSO) for feature selection. The proposed HI-BQPSO method effectively and efficiently improves the classification accuracy and introduces the strategies of crossover and mutation, compared with other feature selection approaches on nine gene expression datasets and 36 UCI datasets. The sequential backward selection (SBS) [62] is a heuristic search algorithm, that is, simple to implement, but the amount of computation is greatly affected by the initial feature set.

The traditional BWOA, BABC, BGA, BGWO, BPSO, and CBPSO algorithms have simple structure and few parameters. These algorithms are proved to be effective by using a binary mechanism to select feature subsets, but they are difficult to effectively jump out of the local optimum. For the state-of-the-art metaheuristic optimization algorithm, the HI-BQPSO algorithm enhances the diversity in the search process and can effectively search the optimal feature subset, but HI-BQPSO lacks stability in balance exploration and exploitation for searching global optimal solution. In order to effectively balance exploration and exploitation in the search process, a hybrid nonlinear inertia weight binary particle swarm optimization with shrinking encircling and exploration mechanism is proposed, and then, the SBS method is introduced, namely, NBPSOSEE-SBS, for solving feature selection tasks. NBPSOSEE-SBS can effectively reduce the number of features while maintaining the best classification effect. In order to prove the effectiveness and superiority of NBPSOSEE-SBS, the two groups of comparative experiments are set up, using the logistic regression method, to realize the risk prediction of ECR for power customers in June, July, and August 2018. NBPSOSEE-SBS can significantly reduce the feature dimension, improve the accuracy of classification, and effectively enhance the global search capability. The main contributions of this study are summarized as follows. Firstly, the proposed algorithm achieves the balance between local search and global search by nonlinearly updating the inertia weights and enhances the diversity of particles for searching the optimal solutions. Secondly, two dynamic contraction factors are introduced into the update of velocity and position, which can not only effectively enhance the inheritance ability, self-recognition ability, and social cognitive ability of the particles but also improve the quality of the particle position. Furthermore, a novel position updating approach is proposed to get rid of the trap of local optimum and shrinking encircling, and exploration mechanism is introduced. Finally, the SBS algorithm is used to remove individual redundant features separately and help to find the potential optimal solutions.

2. Methodology

2.1. Binary Particle Swarm Optimization. PSO is a random search algorithm based on group cooperation developed by simulating the foraging behavior of birds [41]. Assuming that the feature dimension of a target search space is D , the

size of the population is m , where $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ is a D -dimensional vector and represents the i^{th} particle, $i = (1, 2, \dots, m)$, and m is the number of particles. x_i is the position of particle i in the target search space. $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ is the velocity determining the direction and distance of each particle flight in dimension D . $pbest_i$ and $gbest$ are the optimal position of the i^{th} particle and the optimal position for the whole population, respectively. The velocity and position of particles are updated by the following equations, respectively:

$$V_{id}^{t+1} = w * V_{id}^t + c_1 * r_1^t * (pbest_{id}^t - X_{id}^t) + c_2 * r_2^t * (gbest_d^t - X_{id}^t), \quad (1)$$

$$X_{id}^{t+1} = X_{id}^t + V_{id}^{t+1}, \quad (2)$$

where $d = (1, 2, \dots, D)$ and $i = (1, 2, \dots, m)$. m is the number of particles. c_1 and c_2 are learning factors. The range of r_1^t and r_2^t is $[0, 1]$. t is the number of iterations.

The PSO is only suitable for continuous function solving. Kennedy and Eberhart proposed the BPSO [46] based on binary encoding form. In the BPSO, the position of each particle is represented by binary strings, and the velocity vector is not. The positions of particles are updated according to the following equation [63]:

$$X_{id} = \begin{cases} 1, & \text{rand} < S(v_{id}), \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where $S(v_{id}) = 1/(1 + e^{-v_{id}})$.

2.2. Fitness Function. The purpose of feature selection is to get the best classification results with the least features. The fitness function is shown in the following equation:

$$\text{fitness} = \frac{\text{F1_High} + \text{F1_Med} + \text{F1_Low}}{3}. \quad (4)$$

The fitness value is the F1-measure average value of predicting the high-risk, medium-risk, and low-risk levels of customer ECR, and its range is between $[0, 1]$. In equation (4), F1_High, F1_Med, and F1_Low represent the F1-measure value of predicting the high-risk, medium-risk, and low-risk levels of customer ECR, respectively. In order to make an objective evaluation of the performance of the model, this paper introduces four evaluation criteria: accuracy, precision, recall, and F1-measure. The definitions are shown in the following equations:

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \quad (5)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (6)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (7)$$

$$F1 - \text{measure} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (8)$$

TP, FP, FN, and TN represent true positive, false positive, false negative, and true negative, respectively, in equations (5)–(8). In theory, the higher the values of accuracy, precision, recall, and F1-measure, the higher the fitness value, and the better the predictive performance of the model.

3. Improved Hybrid Feature Selection Algorithm Based on NBPSOSEE-SBS

The framework for the risk prediction of ECR includes three processes: data preprocessing, selecting the optimal feature subset based on NBPSOSEE-SBS, and classification and evaluation using the logistic regression method. The framework for feature selection of ECR risk based on NBPSOSEE-SBS algorithm is shown in Figure 1.

3.1. Data Preprocessing. The characteristic dimension of the original power data is low, which cannot adequately express the arrears behavior of power users. In order to solve this problem, data are expanded from low-dimensional features to high-dimensional features by feature extraction. According to the electricity data with 21 consecutive months, the features of training set and test set are extracted, respectively. The training set contains the power data for 20 months, and the test set contains the power data for one month. For the data of each month, firstly, the categorical features are transformed based on the method of one-hot encoding and, secondly, adding characteristics of the last six months to the encoded data and calculating the maximum, minimum, average, median, variance, and standard deviation; finally, the original 34-dimensional features are extended to 748 dimensions.

The processing of feature extraction is shown in Algorithm 1, where m is the total number of months, $data_month_k$ is the power data of the k^{th} month, and $dataSet$ is the data set after the features have been extracted.

After all the features have been calculated, the min-max standardization method is used to transform the features and map the values to [0, 1]. The min-max standardization function is shown in the following equation:

$$x^* = \frac{x - \min}{\max - \min} \quad (9)$$

In equation (9), x presents the feature be used for transforming, and \max and \min are the maximum and minimum of each feature, respectively.

The expansion of features can reflect the information of users' historical electricity consumption behavior, as shown in Table 1 (only one of the features is taken as an example due to too many features).

These 26 features of electricity consumption behavior in the current month are expanded by this method. Then, the statistical features such as maximum value, minimum value, mean value, variance, and standard deviation of these

historical electricity consumption features are calculated. Take the expansion of "jfjsl" (payment timeliness rate) as an example, and the specific statistical analysis is shown in Table 2 (only one of the features is taken as an example).

In addition, one-hot coding method is used to transform the category features into numerical features. Finally, the original 34-dimensional features are extended to 748 dimensions.

4. Improved NBPSOSEE Algorithm

Inertial weight w is a very important parameter in the adjustable parameters of the BPSO algorithm. The value of w plays an important role in the performance of the algorithm. The small value w is convenient for local search of the current search area, and a more accurate solution can be obtained to facilitate the convergence of the algorithm, but it is not easy to jump out of the local extremum point; a large value of w is convenient for global search, but it is not easy to get an accurate solution. Literature [60] shows that linear optimization of inertia weight can improve the performance of the algorithm, but this strategy cannot effectively satisfy the optimization process of the algorithm. Therefore, in order to get closer to the actual evolutionary state of the algorithm, this paper performs nonlinear incremental optimization of inertia weight. In each iteration, the inertia weight w is calculated as follows:

$$w = \begin{cases} w_{\min} + \frac{2}{\pi} \arctan\left(\pi * \frac{t(w_{\max} - w_{\min})}{T}\right), & 0.4 < w \leq 0.9, \\ w_{\max}, & w > 0.9, \end{cases} \quad (10)$$

where t and T represent the current iteration number and the maximum number of iterations, respectively. As the number of iterations increases, the inertia weight exhibits a nonlinear incremental state. It can be known that the improved algorithm has a smaller w in the early stage of searching for the optimal solution, and the particles have a stronger local search capability. In the later stage of finding the optimal solution, there is a large w , and the particles have strong global search ability.

Furthermore, in order to enhance the optimal performance of the PSO, two new contraction factors are introduced into the position updating equation in the NBPSOSEE algorithm. Clerc and Kennedy [64] proposed a particle swarm optimization algorithm with contraction factor (CFPSO) in 2002, which is intended to improve the convergence speed while getting rid of the local optimal value. The algorithm flow of CFPSO is similar to the original PSO, but the velocity updating formula of particles is different. CFPSO uses the contraction factor to compress the particle velocity after each update, which not only changes the influence ability of its own historical velocity but also changes the influence ability of the historical optimal position on the particle velocity, so as to improve the convergence speed of the population. However, CFPSO has some drawbacks. Too large value of the contraction factor results in poor

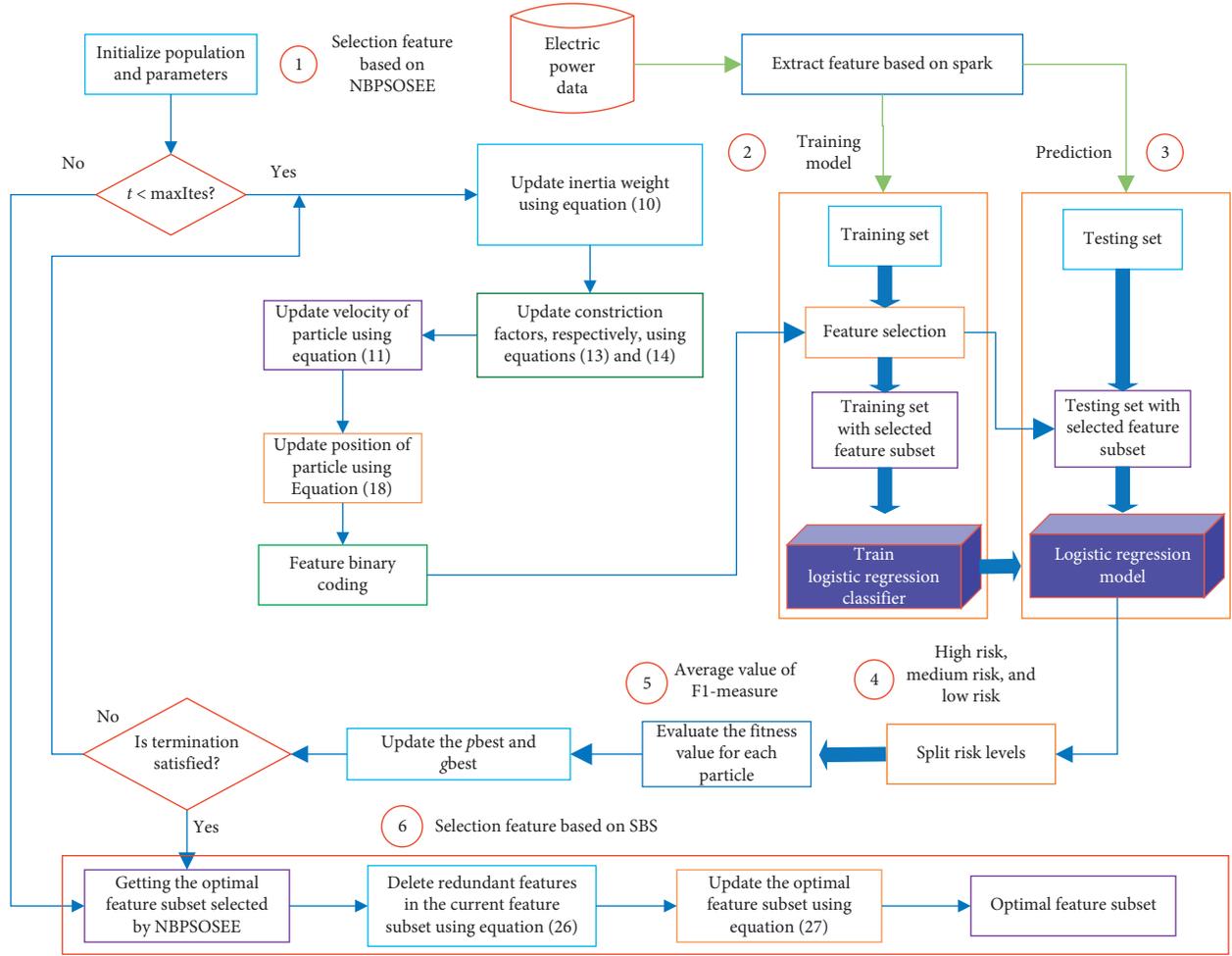


FIGURE 1: Framework for feature selection of ECR risk based on the NBPSOSEE-SBS algorithm.

- (1) **Input:** the original data set
- (2) **Output:** the data set after feature extracted
- (3) Get training set;
- (4) Get test set;
- (5) Create an empty new data table *dataSet*;
- (6) **for** $k=7$ to m/m is the total number of months
- (7) $data_month_k \leftarrow$ Get power data for the k^{th} month;
- (8) One-hot encoding of categorical features in $data_month_k$;
- (9) **for** $j=1$ to $6 // k-6^{\text{th}}$ to $k-1^{\text{th}}$ month
- (10) $data_month_{k-j} \leftarrow$ Get power data for the $k-j^{\text{th}}$ month;
- (11) $data_month_k \leftarrow$ Add the numeric features in $data_month_{k-j}$ to $data_month_k$;
- (12) **end for**
- (13) $data_month_k \leftarrow$ Calculate the maximum, minimum, mean, median, variance, and standard deviation;
- (14) $dataSet \leftarrow dataSet \cup data_month_k //$ Merge the extended datasets to genera
- (15) **end for**

ALGORITHM 1: Feature extraction.

convergence performance and makes the PSO close to random search optimization. If the value of the contraction factor is too small, the PSO will easily converge earlier and reduce the accuracy of classification. In order to solve this problem, the two dynamic contraction factors ω_1 and ω_2 are

used to compress the velocity and position of each update, respectively, in the equations (11) and (12). On adding a contraction factor to the velocity, NBPSOSEE can improve the inheritance ability, self-recognition ability, and social cognitive ability of the particles. NBPSOSEE introduces a

TABLE 1: Expanding the features of historical electricity consumption.

The name of feature	The name of extended feature	Describe
jfjslj (payment timeliness rate)	jfjsl_1	jfjsl of the first month before the forecast month
	jfjsl_2	jfjsl of the second month before the forecast month
	jfjsl_3	jfjsl of the third month before the forecast month
	jfjsl_4	jfjsl of the fourth month before the forecast month
	jfjsl_5	jfjsl of the fifth month before the forecast month
	jfjsl_6	jfjsl of the sixth month before the forecast month

TABLE 2: Calculating the statistical features of historical electricity consumption.

The name of feature	The name of extended feature	Describe
Jfjsl (payment timeliness rate)	jfjsl_min	The minimum value of jfjsl in the first six months of history
	jfjsl_max	The maximum value of jfjsl in the first six months of history
	jfjsl_mean	The average value of jfjsl in the first six months of history
	jfjsl_var	The variance of jfjsl in the first six months of history
	jfjsl_med	The median of jfjsl in the first six months of history
	jfjsl_std	The standard deviation of jfjsl in the first six months of history

contraction factor, which will improve the quality of the particle position when updating the position. In NBPSOSSE, the two dynamic contraction factors can enhance the performance of exploration and exploitation and improve the convergence speed:

$$V_{id}^{t+1} = \omega_1^t * (w * V_{id}^t + c_1 * r_1^t * (pbest_{id}^t - X_{id}^t) + c_2 * r_2^t * (gbest_{id}^t - X_{id}^t)), \quad (11)$$

$$X_{id}^{t+1} = \omega_2^t * X_{id}^t + V_{id}^{t+1}. \quad (12)$$

These two dynamic contraction factors reveal a non-linear convert of the particle velocity and position. Parameters ω_1 and ω_2 are calculated as follows:

$$\omega_1^t = \begin{cases} \frac{1}{(1 + \exp(\tau * ((-PV_{avg}) / (PV_{max} - PV_{min}))))^t}, & \text{if } PV_{avg} < PV_{max} - PV_{min}, \\ \frac{1}{(1 + \exp(\tau * ((-PV_{min}) / (PV_{max}))))^t}, & \text{else,} \end{cases} \quad (13)$$

$$\omega_2^t = 1 - \omega_1^t, \quad (14)$$

$$PV_i^t = x_{id}^{t-1} * [x_{id}^{t-1}]^T, \quad (15)$$

where PV is position value of the particle, t is the current iteration, and τ is a constant ($\tau = 2$).

Then, in NBPSOSEE, using the two mechanisms of shrinking encircling and exploration can improve the search ability of the population. Firstly, the moving position of the particle is determined between the current position and the target position via using shrinking encircling operation, which can shorten the search range of the particles and achieve the purpose of enhancing the local search ability of the population. In addition to shrinking encircling strategy, NBPSOSEE refers to random search mechanism to improve the diversity of particles. When updating the particle position, it is based on the change of coefficient A . If A exceeds

the range of $[-1, 1]$, distance coefficient D will be updated randomly. In order to find the target, the particles will traverse the original target direction, making the population have the performance of global search. In NBPSOSEE, the particle position is updated by equation (18). NBPSOSEE adopts dynamic contraction strategy, shrinking encircling operation and exploration mechanism with some probability, which not only can get rid of the local optimal solution but also can accelerate the convergence speed:

$$D_{id}^t = |C * gbest_{id}^t - X_{id}^t|, \quad (16)$$

$$D_{jd}^t = |C * X_{jd}^t - X_{jd}^t|, \quad (17)$$

$$X_{id}^{t+1} = \begin{cases} gbest_d^t - A * D_{id}^t, & \text{if } p < 0.5 \text{ and } |A| < 1, \\ X_{jd}^t - A * D_{jd}^t, & \text{if } p < 0.5 \text{ and } |A| > 1, \\ \bar{\omega}_2^t * X_{id}^t + V_{id}^{t+1}, & \text{else.} \end{cases} \quad (18)$$

In equations (16) and (17), C represents a coefficient variable, and j is a target that is randomly searched ($j \neq i$). In equation (14), t represents the current number of iterations, A represents a coefficient variable, $gbest_d^t$ is defined as the optimal position of the current target d , and p represents a random value between $[0, 1]$.

The coefficient variables, namely, A and C , are calculated separately as follows:

$$A = 2 * a * r - a, \quad (19)$$

$$C = 2 * r. \quad (20)$$

In the above formulas, a is a variable with a range of values between $[0, 2]$ and presents a linear decreasing trend. It is updated in the form of $a = 2 - 2 * (t/T)$, where t is the current number of iteration, T represents the maximum iteration number, and r is the random value between $[0, 1]$.

4.1. SBS Algorithm. Since the feature subset selected by NBPSOSEE still contains redundant features, the importance of features is different, and the order of arrangement is confusing. Therefore, this paper firstly uses the feature selection method of random forest to sort the importance degree of the features selected by NBPSOSEE and then uses SBS algorithm to delete features with low importance in turn.

For each node on the random forest decision tree, \sqrt{d} features are usually randomly extracted from the d -dimensional feature set. Then, according to the Gini gain maximization principle [65], a feature is selected to divide the data on the node into two left and right child nodes, which means that the data on the parent node is divided into its child nodes n_l and n_r . The Gini gain maximization is to maximize the following equation:

$$\Delta I_G = I_G(n_p) - p_l * I_G(n_l) - p_r * I_G(n_r), \quad (21)$$

$$I_G(n) = 1 - \sum_{c=1}^2 p_c^2. \quad (22)$$

Here, $I_G(n)$ is the Gini index of node n , p_c is the proportion of class c samples in node n , while p_l and p_r are the rates of data divided into left and right child nodes n_l and n_r by parent node n_p , respectively. The importance of features on nodes is shown in the following equation:

$$\text{IMP}_{i,n}^{\text{Gini}} = I_G(n) - I_G(n_l) - I_G(n_r). \quad (23)$$

If the set of nodes in which the feature f_i appears as a node partitioning attribute in the k^{th} decision tree is N , the importance of the feature on the decision tree is calculated as shown in the following equation:

$$\text{IMP}_{i,k}^{\text{Gini}} = \sum_{n \in N} \text{IMP}_{i,n}^{\text{Gini}}. \quad (24)$$

Assuming that there are K trees in the random forest, the importance of the feature f_i in the random forest can be calculated as follows:

$$\text{IMP}_i^{\text{Gini}} = \frac{1}{K} \sum_{k=1}^K \text{IMP}_{i,k}^{\text{Gini}}. \quad (25)$$

Here, K is the number of decision trees in the random forest. After the NBPSOSEE obtains the optimal feature subset, the SBS starts searching, calculates the fitness value corresponding to each feature when it is deleted separately, and then selects the feature subset with the best fitness value to enter the next iteration. The iterative steps of the algorithm in the SBS stage are as follows:

Step 1: determine the optimal feature subset bestSubFs of the SBS stage.

Step 2: delete a feature f in the current feature subset and f satisfies the following equation:

$$f = \arg \max_{f \in \text{bestSubFs}_t} \text{fitness}(\text{bestSubFs}_t - f), \quad (26)$$

where $\text{bestSubFs}_t - f$ denotes the feature subset after bestSubFs_t removes feature f , t is the number of iterations, and fitness is the fitness value. The larger the fitness value in this paper, the better the selected feature subset.

Step 3: update the optimal feature subset and the number of iterations:

$$\text{bestSubFs}_{t+1} = \text{bestSubFs}_t - f, \quad (27)$$

$$t = t + 1. \quad (28)$$

Step 4: repeat steps 2 and 3 until the termination condition is met.

The feature selection processing of NBPSOSEE-SBS is shown in Algorithm 2. Where maxIterations is the maximum number of iterations, swarmSize is the number of particles in the population, dimension is the dimension of each particle, and fitness is the fitness value.

5. Computational Complexity of NBPSOSEE-SBS

The time computational complexity of the BPSO algorithm can be expressed as $O(M \times N \times P)$, where M indicates the number of iterations, N is the population size, and P represents the computing time for the logistic regression model training and predicting. The time computational complexity of the CBPSO algorithm can be expressed as $O((M + R_1) \times N \times P)$, where R_1 is the computing time of logistic map sequence. The time computational complexity of the CBPSO algorithm can be expressed as $O((M + R_1) \times N \times P)$, where R_1 is the computing time of

```

(1) Input: the data set after feature extracted
(2) Output: best feature subset
(3) Get training set and test set after the features have been extracted;
(4) Initialize population, such as initial position, velocity and fitness values of particles, as well as local optimal  $pbest$  and global optimal  $gbest$  of particles;
(5) for  $k = 1 \rightarrow maxIterations$  do
(6)   for  $i = 1 \rightarrow swarmSize$  do
(7)     Update inertia weight using equation (10);
(8)     for  $j = 1 \rightarrow dimension$  do
(9)       Update  $a$  and  $p$ ;
(10)      Update  $A$  and  $C$ , respectively, using equations (19) and (20);
(11)      Update  $\hat{\omega}_1^t$  and  $\hat{\omega}_2^t$ , respectively, using equations (13) and (14);
(12)      Update velocity of particle using equation (11);
(13)      Update position of particle using equation (18);
(14)    end for
(15)  end for
(16)  for  $i = 1 \rightarrow swarmSize$  do
(17)    Update fitness value  $fitness_i$ ;
(18)    Update local optimal  $pbest_i$ ;
(19)  end for
(20)  Update global optimal  $gbest$ ;
(21) end for
(22) Getting the optimal feature subset selected by NBPSOSEE;
(23) Delete a feature  $f$  in the current feature subset using equation (26);
(24) Update the optimal feature subset using equation (27);
(25) Repeat steps 23 and 24 until the termination condition is met;

```

ALGORITHM 2: NBPSOSEE-SBS feature selection.

logistic map sequence. The time computational complexity of the HI-BQPSO algorithm can be expressed as $O((M + R_1) \times (N + R_2) \times P)$, where R_2 is the computing time for correction factors.

The time computational complexity of the proposed NBPSOSEE algorithm can be expressed as $O((M + R_4) \times (N + R_5 + R_6) \times P)$, where R_4 is the computing time for calculating inertia weight; R_5 is the computing time for updating coefficient variables, namely, A and C ; and R_6 is the computing time for dynamic contraction factors $\hat{\omega}_1$ and $\hat{\omega}_2$. The time computational complexity of the proposed NBPSOSEE-SBS algorithm can be expressed as $O((M + R_4) \times (N + R_5 + R_6) \times P + T)$, where T is the computing time of the SBS method. It can be seen that the computational complexity of NBPSOSEE-SBS algorithm is obviously highly compared with the BPSO and CBPSO algorithms. However, R_4 , R_5 , and R_6 are the only simple numerical operations according to equations (10), (13), (14), (19), and (20). Furthermore, a large number of redundant or irrelevant features are deleted by the NBPSOSEE algorithm. SBS method will spend less time deleting the remaining unimportant features. Therefore, the proposed NBPSOSEE-SBS algorithm does not significantly improve the computational complexity.

6. Results

6.1. Data. In this paper, the data set from January 2017 to September 2018 is provided by a power enterprise, including the electricity consumption data of 11,860 high-voltage users

who have had arrears. According to the users' past payment behavior, the power enterprise divides the users into high-risk, medium-risk, or low-risk arrears ones. A total of 34 features are used for data processing and model training. There are 8 category features that represent basic information for these users, while 26 statistical features describe these users consuming electric information.

Experimental environment: operating system Centos 7.0, CPU Intel Xeon E5-2620 V3, memory 128G, development language Python 3.5.

6.2. Experimental Results. In order to prove the effectiveness and superiority of the proposed algorithm, two groups of comparative experiments are set up and use logistic regression model [66, 67] to realize the risk prediction of ECR for power customers in June, July, and August 2018. The first group of experiments verifies the effectiveness of the NBPSOSEE. The second group of experiments proves the superiority of the proposed hybrid feature selection algorithm based on NBPSOSEE with SBS, called NBPSOSEE-SBS.

The relevant parameters selected for NBPSOSEE-SBS are listed in Table 3. Population size 20 to 40 is the optimal size for optimization problems [68]. Generally, population size is set to 20 [46, 58, 60]. Hence, 20 particles are selected to form a particle swarm in this paper. The maximum iteration number is 100; maximum value w_{\max} is 0.9, and minimum value w_{\min} is 0.4 for the inertia weight; learning factor $c_1 = c_2 = 2$.

The logistic regression model is trained by training set with the optimal feature subset selected by the NBPSOSEE-SBS algorithm. The logistic regression model outputs the default probability p of users through the test set, $p \in [0, 1]$. Then, setting the appropriate threshold value θ , according to the threshold value θ , users are divided into three levels: high risk, medium risk, and low risk. The specific division principles are shown in Table 4.

The users with the default probability of $p \geq 70\%$ are divided into the high risk level; the users with the arrears probability of $p \geq 40\%$ and $p < 70\%$ are divided into the medium risk level; and the users with the arrears probability of $p < 40\%$ are divided into the low risk level.

6.3. Experimental Results of NBPSOSEE. Figures 2–4 show the fitness values and the number of features calculated by BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, LBPSOSEE, and proposed NBPSOSEE algorithms for the risk prediction of ECR in June, July, and August 2018. In these pictures, a represents the number of iterations vs. fitness value, and b represents the number of iterations vs. the number of selected features.

Figure 2 shows the test results of the improved NBPSOSEE for ECR risk in June 2018. Compared with the other eight feature selection algorithms, NBPSOSEE gets the maximum fitness value with the least number of features. In terms of number of features and fitness value, the performance of HI-BQPSO, LBPSOSEE, and NBPSOSEE is significantly better than the other six algorithms, of which NBPSOSEE is the best, LBPSOSEE is the second, and HI-BQPSO is the third. In the process of calculating fitness, BABC obtains the least fitness value. The fitness values calculated by HI-BQPSO, LBPSOSEE, and NBPSOSEE are 5.77%, 7.34%, and 11.61% higher than BABC, respectively. However, NBPSOSEE increased by 5.52% and 3.98% in HI-BQPSO and LBPSOSEE, respectively. Furthermore, in the process of selecting feature subset, BABC selects the most features; HI-BQPSO, LBPSOSEE, and NBPSOSEE select 242, 216, and 205 features, respectively. NBPSOSEE selects the fewest features and deletes 190, 179, and 153 fewer features in HI-BQPSO, LBPSOSEE, and BABC, respectively. NBPSOSEE selects approximately one-third of the features from the original feature set, removing a total of 543 redundant or unrelated features. In addition, BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, and LBPSOSEE search for optimal fitness values after 97, 37, 61, 84, 68, 35, 48, and 43 iterations, respectively. The NBPSOSEE obtains the optimal fitness value until 15 iterations and can still search for the best fitness value after 17, 20, and 98 iterations. In the initial search, NBPSOSEE quickly increases the initial fitness value to a very high level, which reflects the convergence performance of the proposed algorithm. The search results show that the proposed NBPSOSEE still has effective global search ability after falling into the local optimal state and can continue to search for a better feature subset in later search.

Figure 3 shows the test results of the proposed NBPSOSEE and other feature selection algorithms for ECR

TABLE 3: Parameters of NBPSOSEE-SBS.

NBPSOSEE-SBS parameters	Values
Particle size (N)	20
Number of iterations ($maxIterations$)	100
Dimension (D)	748
Learning factor (c_1)	2
Learning factor (c_2)	2
Inertia weight (w)	$w_{max} = 0.9; w_{min} = 0.4$
Termination criteria	Maximum number of iterations

TABLE 4: The division principles of risk level.

Threshold value θ	Risk level
$p \geq 70\%$	High risk
$p \geq 40\%$ and $p < 70\%$	Medium risk
$p < 40\%$	Low risk

risk in July. The fitness values calculated by the proposed NBPSOSEE and these comparison algorithms show an increasing trend with the increase in the number of iterations. Moreover, the number of selected features by NBPSOSEE is decreasing. LBPSOSEE and NBPSOSEE get higher fitness values and select fewer features than the other seven algorithms. The fitness values calculated by LBPSOSEE are 2.21%, 4.17%, 6.94%, 4.41%, 5.60%, 4.41%, 1.09% higher than those of BWOA, BABC, BGA, BGWO, BPSO, CBPSO and HI-BQPSO, respectively. Moreover, the number of features selected by LBPSOSEE is 122, 46, 106, 51, 80, 105, and 59 less than those of the seven algorithms. From these results, it can be seen that LBPSOSEE has better ability to avoid falling into local optimal solutions than these seven comparison algorithms. After 58 iterations, the fitness value of NBPSOSEE is 0.939, and the selected optimal feature subset has 213 features. However, the fitness value calculated by NBPSOSEE is 1.62% higher than that of LBPSOSEE, and the number of features selected is 69% less than that of LBPSOSEE. This proves that the global search ability of NBPSOSEE is stronger than LBPSOSEE, and the performance of NBPSOSEE is significantly better than other feature selection algorithms.

Figure 4 shows the test results of the improved NBPSOSEE and other eight comparison algorithms for ECR risk in August. The search capabilities of LBPSOSEE and NBPSOSEE are significantly higher than other algorithms. The optimal fitness values obtained by LBPSOSEE and NBPSOSEE are increasing, and the selected features are decreasing. The optimal fitness value calculated by LBPSOSEE is 0.957, which is 2.46%, 3.12%, 1.27%, 1.70%, 7.71%, 2.79%, and 0.95% higher in BWOA, BABC, BGA, BGWO, BPSO, CBPSO, and HI-BQPSO, respectively. LBPSOSEE selects 172 features, and the number of selected features is reduced by 169, 144, 208, 137, 175, 184, and 111, respectively, compared with seven algorithms. However, the optimal feature subset selected by NBPSOSEE contains 102 features, and the optimal fitness value calculated is 0.977.

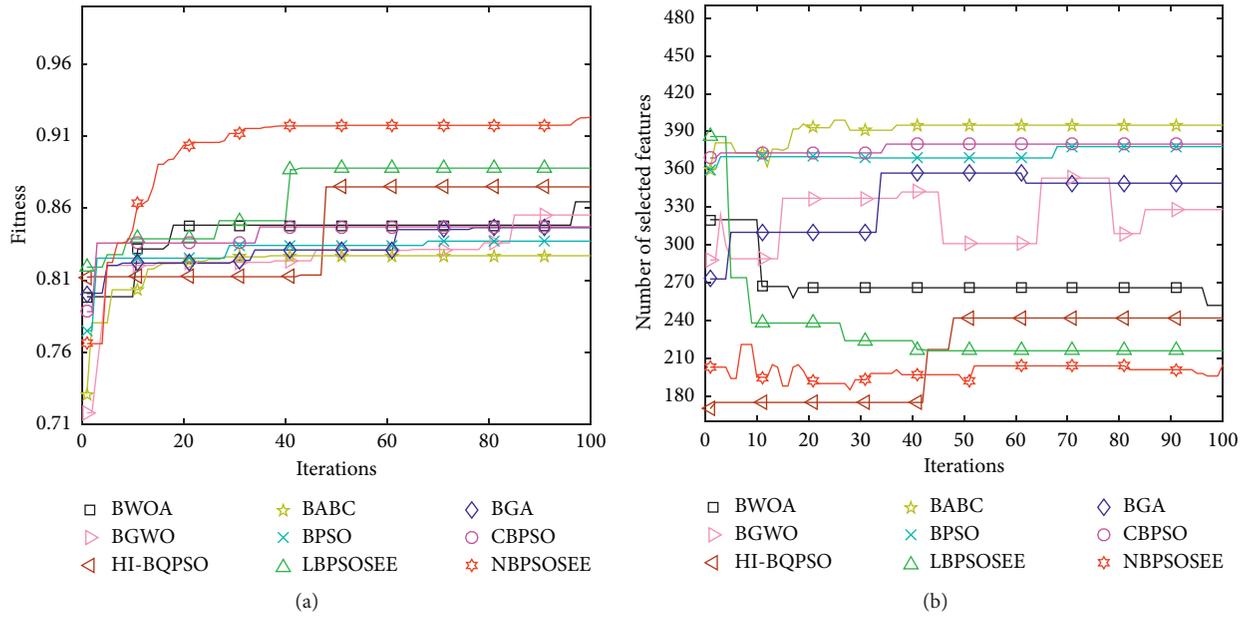


FIGURE 2: Experimental result based on NBPSOSEE in June 2018.

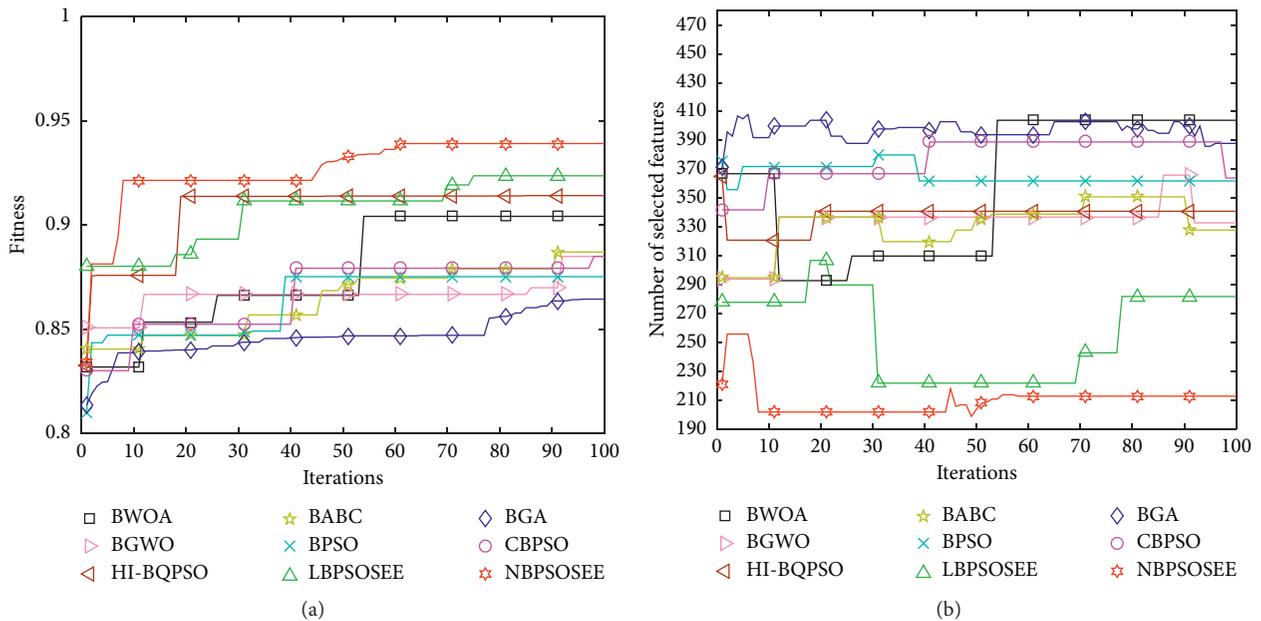


FIGURE 3: Experimental result based on NBPSOSEE in July 2018.

NBPSOSEE deletes 70 more redundant or unrelated features than LBPSOSEE, and the optimal fitness value calculated is 2.09% higher than LBPSOSEE. The experimental results show that NBPSOSEE obtains the highest ability to jump out of the local optimal solution and gets the highest fitness value under the condition of selecting the least features. In conclusion, NBPSOSEE has a higher convergence speed and global search ability than BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, and LBPSOSEE.

As shown from Figures 2–4, the performance of NBPSOSEE is significantly better than BWOA, BABC, BGA,

BGWO, BPSO, CBPSO, HI-BQPSO, and LBPSOSEE. The proposed NBPSOSEE selects the least number of features to obtain the highest fitness value for the test of ECR risk in June, July, and August 2018. This verifies that the NBPSOSEE convergence speed and global search ability are higher than other algorithms and also demonstrates the effectiveness and stability of the proposed NBPSOSEE.

6.4. *Experimental Results of NBPSOSEE-SBS.* Although the results obtained by NBPSOSEE have been improved, the feature subset selected by NBPSOSEE still contains many

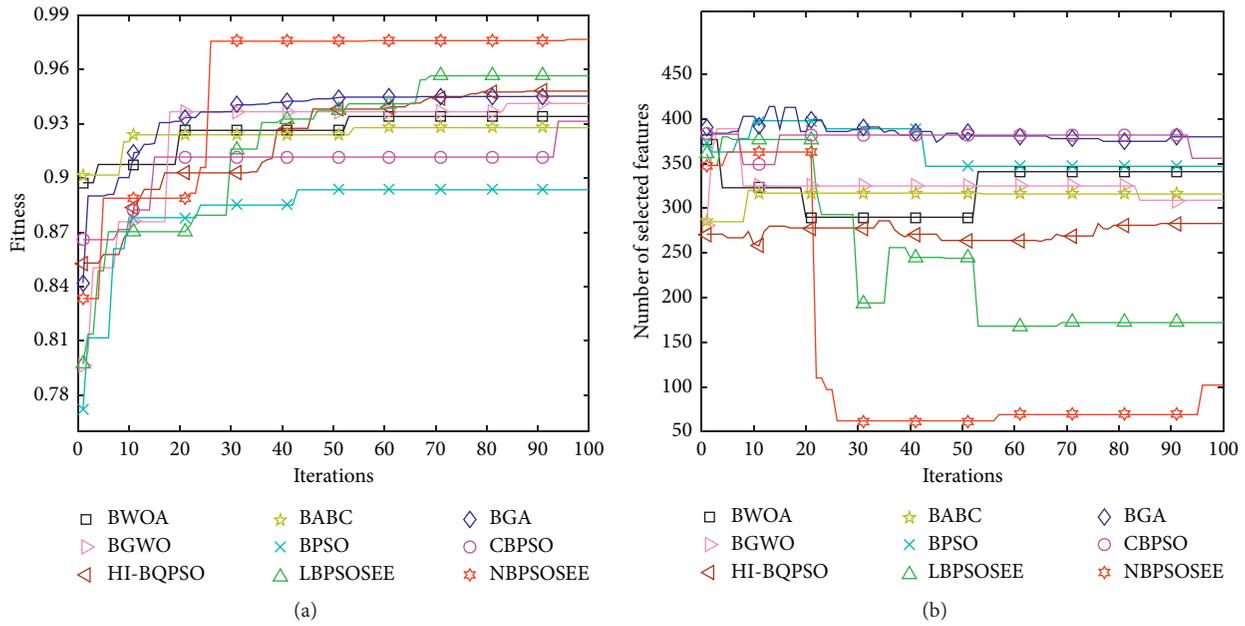


FIGURE 4: Experimental result based on NBPSOSEE in August 2018.

redundant features. Therefore, a hybrid feature selection algorithm is proposed, namely, NBPSOSEE-SBS, to select optimal feature subset in this paper, which can not only effectively reduce the number of features but also improve the accuracy. Figures 5–7 show the test results calculated by BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, LBPSOSEE-SBS, and the proposed NBPSOSEE-SBS algorithms for the risk prediction of ECR in June, July, and August 2018. The x -axis represents the number of iterations, where the maximum number of iterations represents the number of features selected by the NBPSOSEE.

Figure 5 shows the test results of the improved NBPSOSEE-SBS for ECR risk in June 2018. With the constantly deleting redundant or irrelevant features, the fitness values calculated by NBPSOSEE-SBS and other eight feature selection algorithms show an increasing trend. The optimal fitness value calculated by BWOA-SBS is 3.01% higher than that of BWOA, and the number of deleted redundant features is 9 more than BWOA. The best fitness value calculated by BABC-SBS is 0.24% higher than that of BABC, and the selected features are reduced by 6 compared to BABC. The optimal fitness value obtained by BGWO-SBS is increased by 1.29% compared with BGWO, and the selected features are 21 fewer than that of BGWO. The optimal fitness value obtained by BPSO-SBS is 1.9% more than BPSO, and the number of selected features is reduced by 12 compared with BPSO. The optimal fitness value obtained by CBPSO-SBS is 0.7% higher than CBPSO, and 9 more redundant features are deleted than CBPSO. The best fitness value of HI-BQPSO-SBS calculation is 3.54% higher than that of HI-BQPSO, and the deleting redundant features are 35 more than HI-BQPSO. The best fitness value of LBPSOSEE-SBS is 0.79% higher than that of LBPSOSEE, and selecting redundant features are 14 fewer than LBPSOSEE. In addition, the

optimal fitness value obtained by NBPSOSEE-SBS is 0.54% more than NBPSOSEE, and the number of selecting features is reduced by 11 compared with NBPSOSEE. The experimental results show that NBPSOSEE-SBS obtains the highest fitness value and removes the most redundant or unrelated features.

Figure 6 shows the test results based on NBPSOSEE-SBS and other feature selection algorithms for ECR risk in July. The best fitness values obtained by BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, LBPSOSEE-SBS, and the proposed NBPSOSEE-SBS are significantly superior to BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, LBPSOSEE, and NBPSOSEE, respectively, and the more redundant features are removed. NBPSOSEE-SBS obtains the highest fitness value, which increased by 3.68%, 3.68%, 8.79%, 5.73, 1.18%, 1.95%, 2.84%, and 0.64% compared to BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, and LBPSOSEE-SBS, respectively. NBPSOSEE-SBS selects the least number of features, which, respectively, reduced by 184, 88, 171, 121, 134, 158, 114, and 60 redundant features compared with the eight comparison algorithms. This shows that the proposed NBPSOSEE-SBS has better performance than other comparative algorithms in balancing local search and global search.

Figure 7 shows the test results of the improved NBPSOSEE-SBS for ECR risk in August. The results obtained by NBPSOSEE-SBS are significantly better than other algorithms. NBPSOSEE-SBS deletes the most redundant or unrelated features and obtains the highest fitness value. The optimal fitness values obtained by NBPSOSEE-SBS increased by 4.45%, 5.79%, 4.34%, 4.23%, 6.71%, 5.45%, 3.79%, and 2.60% compared with the other eight algorithms, respectively. The number of features selected by NBPSOSEE-SBS is reduced by 256, 236, 299, 227, 259, 280, 198, and 85,

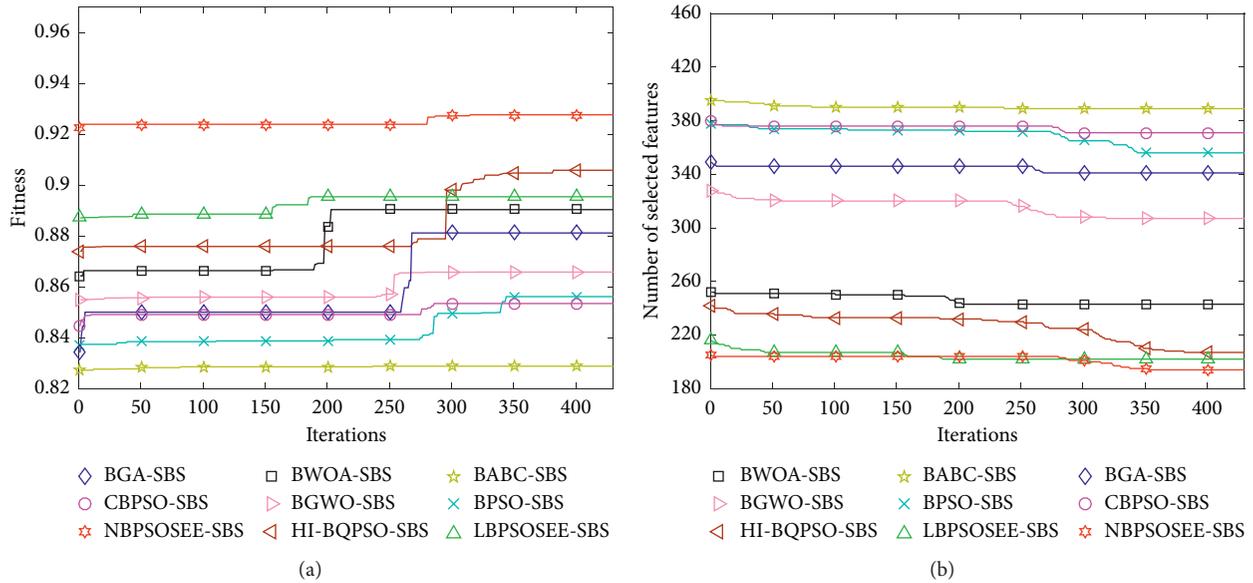


FIGURE 5: Experimental result based on NBPSOSEE-SBS in June 2018.

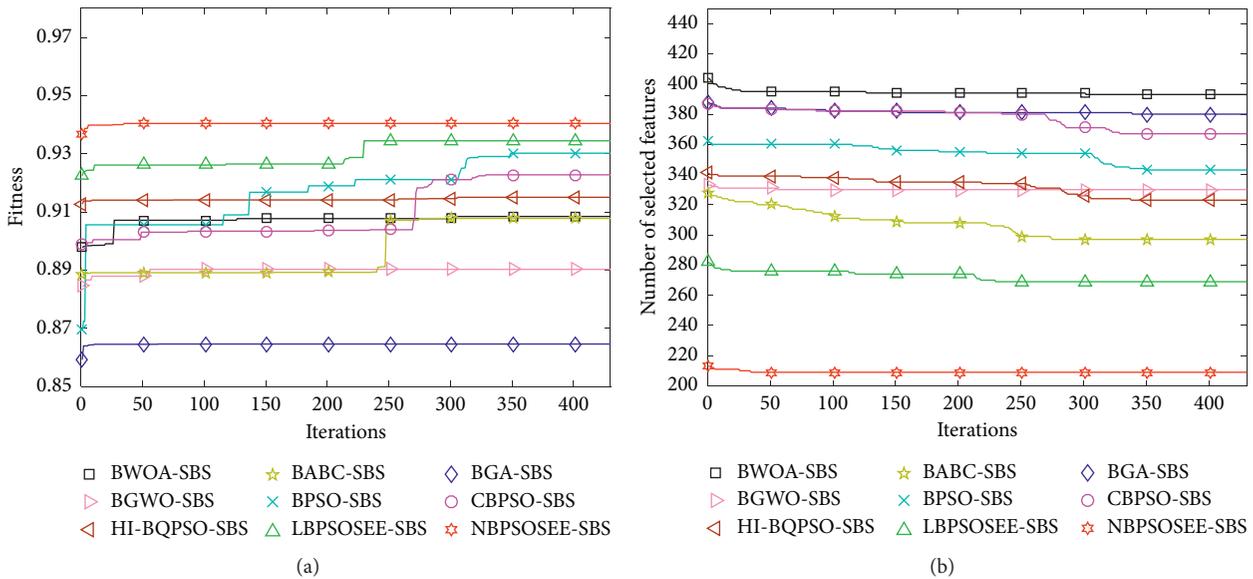


FIGURE 6: Experimental result based on NBPSOSEE-SBS in July 2018.

respectively, over these eight algorithms. The experimental results show that the proposed NBPSOSEE-SBS has the best performance.

The test results are shown in Figures 5–7. BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, LBPSOSEE-SBS, and the proposed NBPSOSEE-SBS are significantly better than BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, LBPSOSEE, and NBPSOSEE. However, the proposed NBPSOSEE-SBS gets the best results, which not only deletes the most of redundant or unrelated features but also calculates the optimal fitness value. The experimental results show that the proposed NBPSOSEE-SBS improves the convergence speed of

particles and also enhances the ability to avoid falling into premature.

7. Discussion

Tables 5–7 show the comparative experiments of the sixteen traditional algorithms, a new feature selection algorithm, and the proposed NBPSOSEE-SBS for the risk prediction of ECR in June, July, and August 2018. According to the number of selected features, accuracy, precision, recall, and F1-measure, the convergence speed and search performance of the proposed algorithm can be evaluated verifying its effectiveness and superiority. In this table, d^* , Acc, Pre, Rec,

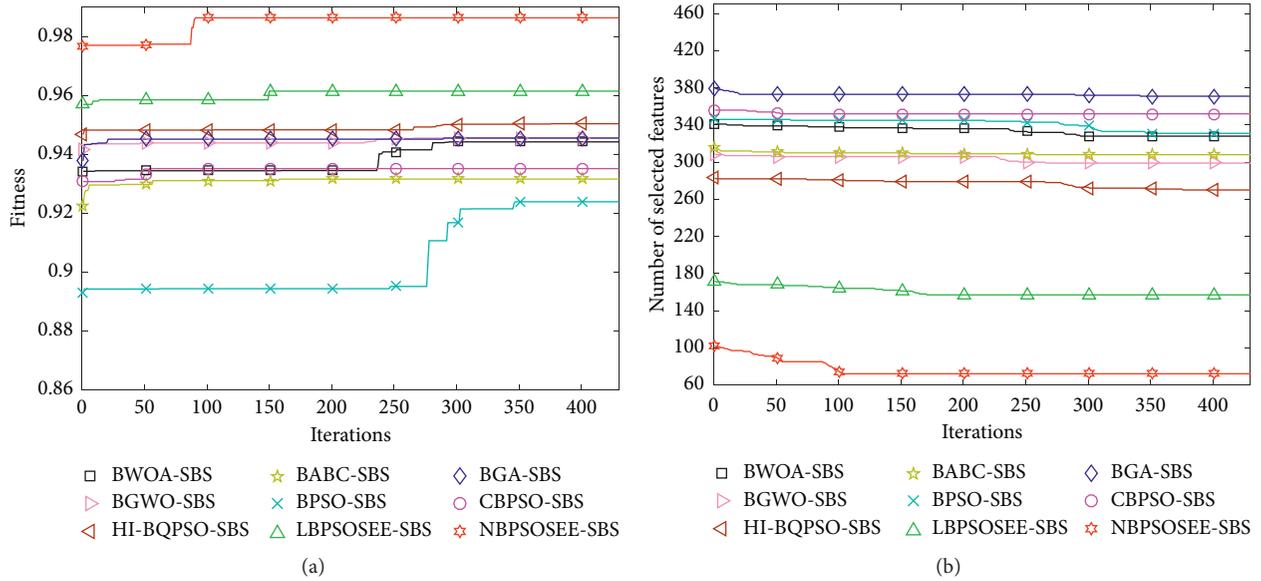


FIGURE 7: Experimental result based on NBPSOSEE-SBS in August 2018.

TABLE 5: Results of NBPSOSEE-SBS and contrast algorithms for ECR risk in June 2018.

Algorithms	d^*	High risk				Medium risk				Low risk			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
Without selected features	748	1.0	1.0	1.0	1.0	0.736	0.824	0.843	0.833	0.779	0.210	0.742	0.327
BWOA	252	1.0	1.0	1.0	1.0	0.859	0.996	0.861	0.924	0.928	0.526	0.921	0.669
BABC	395	1.0	1.0	1.0	1.0	0.761	0.901	0.818	0.858	0.912	0.549	0.720	0.623
BGA	349	1.0	1.0	1.0	1.0	0.828	0.982	0.824	0.896	0.924	0.499	0.905	0.643
BGWO	328	1.0	1.0	1.0	1.0	0.838	0.996	0.839	0.911	0.925	0.514	0.900	0.655
BPSO	378	1.0	1.0	1.0	1.0	0.855	0.986	0.858	0.917	0.914	0.426	0.978	0.594
CBPSO	380	1.0	1.0	1.0	1.0	0.832	0.989	0.837	0.907	0.924	0.450	0.869	0.634
HI-BQPSO	242	1.0	1.0	1.0	1.0	0.803	1.0	0.802	0.890	0.945	0.618	0.903	0.734
LBPSOSEE	216	1.0	1.0	1.0	1.0	0.789	0.975	0.806	0.882	0.961	0.759	0.804	0.781
NBPSOSEE	205	1.0	1.0	1.0	1.0	0.771	1.0	0.771	0.870	0.981	0.886	0.912	0.899
BWOA-SBS	243	1.0	1.0	1.0	1.0	0.778	1.0	0.778	0.875	0.957	0.680	0.960	0.796
BABC-SBS	389	1.0	1.0	1.0	1.0	0.764	0.905	0.819	0.860	0.913	0.558	0.714	0.626
BGA-SBS	341	1.0	1.0	1.0	1.0	0.748	1.0	0.748	0.856	0.955	0.686	0.925	0.788
BGWO-SBS	307	1.0	1.0	1.0	1.0	0.844	1.0	0.843	0.915	0.930	0.539	0.930	0.683
BPSO-SBS	356	1.0	1.0	1.0	1.0	0.886	0.991	0.889	0.938	0.922	0.465	0.982	0.631
CBPSO-SBS	371	1.0	1.0	1.0	1.0	0.845	0.989	0.851	0.915	0.927	0.512	0.875	0.646
HI-BQPSO-SBS	207	1.0	1.0	1.0	1.0	0.797	1.0	0.797	0.887	0.965	0.748	0.933	0.831
LBPSOSEE-SBS	202	1.0	1.0	1.0	1.0	0.804	0.974	0.821	0.891	0.963	0.780	0.811	0.795
NBPSOSEE-SBS	194	1.0	1.0	1.0	1.0	0.782	1.0	0.782	0.878	0.982	0.897	0.914	0.905

and F1 represent the selected number of features, accuracy, precision, recall, and F1-measure, respectively.

Table 5 shows the test results of the improved NBPSOSEE-SBS and comparison algorithms for ECR risk in June 2018. For high-risk arrears users, the accuracy, precision, recall, and F1-measure calculated by these comparison algorithms are all 100%. This is because the number of high-risk arrears is small, and the characteristics of reflecting arrears behavior are obvious. Among the arrears' users of the medium risk level, HI-BQPSO, NBPSOSEE, BWOA-SBS, BGA-SBS, BGWO-SBS, HI-BQPSO-SBS, and the proposed NBPSOSEE-SBS achieve the highest precision, reaching 100%. The proposed NBPSOSEE-SBS obtains the highest accuracy, precision, and F1-measure for the arrears' users of the low risk level. The

accuracy, precision, and F1-measure obtained by NBPSOSEE-SBS are 0.1%, 1.24%, and 0.67% higher than NBPSOSEE, respectively. Moreover, the proposed NBPSOSEE-SBS removes the most redundant or unrelated features and calculates the highest average value of F1-measure.

Table 6 shows the test results of the improved NBPSOSEE-SBS and comparison algorithms for ECR risk in July 2018. The results calculated using all feature subsets are the worst, but the results of the proposed NBPSOSEE-SBS are significantly better than other algorithms. The accuracy and F1-measure of NBPSOSEE-SBS are higher than other algorithms except for CBPSO in the arrears' users of the medium risk level. Furthermore, NBPSOSEE-SBS achieves a precision of 100%. In addition, the accuracy and precision

TABLE 6: Results of NBPSOSEE-SBS and contrast algorithms for ECR risk in July 2018.

Algorithms	d^*	High risk				Medium risk				Low risk			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
Without selected features	748	1.0	1.0	1.0	1.0	0.830	0.913	0.891	0.902	0.839	0.253	0.826	0.387
BWOA	404	1.0	1.0	1.0	1.0	0.850	1.0	0.850	0.919	0.970	0.732	0.867	0.794
BABC	328	1.0	1.0	1.0	1.0	0.849	1.0	0.849	0.918	0.957	0.628	0.908	0.743
BGA	388	1.0	1.0	1.0	1.0	0.826	0.968	0.845	0.902	0.934	0.661	0.724	0.691
BGWO	333	1.0	1.0	1.0	1.0	0.860	1.0	0.859	0.924	0.958	0.685	0.782	0.730
BPSO	362	1.0	1.0	1.0	1.0	0.843	1.0	0.843	0.915	0.962	0.573	0.936	0.711
CBPSO	387	1.0	1.0	1.0	1.0	0.921	1.0	0.924	0.961	0.946	0.553	0.932	0.694
HI-BQPSO	341	1.0	1.0	1.0	1.0	0.844	1.0	0.844	0.916	0.970	0.770	0.893	0.827
LBPSOSEE	282	1.0	1.0	1.0	1.0	0.826	1.0	0.826	0.905	0.978	0.837	0.898	0.866
NBPSOSEE	213	1.0	1.0	1.0	1.0	0.868	1.0	0.868	0.929	0.982	0.910	0.867	0.888
BWOA-SBS	393	1.0	1.0	1.0	1.0	0.850	1.0	0.850	0.919	0.972	0.753	0.867	0.806
BABC-SBS	297	1.0	1.0	1.0	1.0	0.818	1.0	0.818	0.900	0.970	0.754	0.907	0.823
BGA-SBS	380	1.0	1.0	1.0	1.0	0.828	0.969	0.846	0.903	0.935	0.662	0.727	0.693
BGWO-SBS	330	1.0	1.0	1.0	1.0	0.870	1.0	0.870	0.930	0.959	0.703	0.782	0.740
BPSO-SBS	343	1.0	1.0	1.0	1.0	0.790	1.0	0.790	0.883	0.980	0.888	0.928	0.908
CBPSO-SBS	367	1.0	1.0	1.0	1.0	0.821	1.0	0.821	0.902	0.980	0.862	0.871	0.866
HI-BQPSO-SBS	323	1.0	1.0	1.0	1.0	0.851	1.0	0.851	0.919	0.970	0.761	0.903	0.826
LBPSOSEE-SBS	269	1.0	1.0	1.0	1.0	0.786	1.0	0.786	0.880	0.987	0.944	0.904	0.924
NBPSOSEE-SBS	209	1.0	1.0	1.0	1.0	0.871	1.0	0.871	0.931	0.983	0.913	0.869	0.891

TABLE 7: Results of NBPSOSEE-SBS and contrast algorithms for ECR risk in August 2018.

Algorithms	d^*	High risk				Medium risk				Low risk			
		Acc	Pre	Rec	F1	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
Without selected features	748	1.0	1.0	1.0	1.0	0.770	0.821	0.896	0.857	0.936	0.607	0.699	0.650
BWOA	341	1.0	1.0	1.0	1.0	0.816	0.998	0.816	0.899	0.982	0.895	0.913	0.904
BABC	316	1.0	1.0	1.0	1.0	0.892	0.999	0.893	0.943	0.970	0.799	0.887	0.841
BGA	380	1.0	1.0	1.0	1.0	0.842	1.0	0.842	0.914	0.983	0.955	0.891	0.922
BGWO	309	1.0	1.0	1.0	1.0	0.816	1.0	0.816	0.899	0.986	0.933	0.918	0.925
BPSO	347	1.0	1.0	1.0	1.0	0.887	0.975	0.905	0.939	0.950	0.625	0.912	0.741
CBPSO	356	1.0	1.0	1.0	1.0	0.819	0.997	0.821	0.901	0.979	0.887	0.900	0.893
HI-BQPSO	283	1.0	1.0	1.0	1.0	0.863	1.0	0.863	0.927	0.983	0.939	0.898	0.918
LBPSOSEE	172	1.0	1.0	1.0	1.0	0.961	0.998	0.964	0.980	0.978	0.801	1.0	0.890
NBPSOSEE	102	1.0	1.0	1.0	1.0	0.877	1.0	0.877	0.934	0.999	0.996	0.996	0.996
BWOA-SBS	328	1.0	1.0	1.0	1.0	0.863	1.0	0.863	0.926	0.981	0.895	0.918	0.906
BABC-SBS	308	1.0	1.0	1.0	1.0	0.893	0.999	0.894	0.944	0.972	0.815	0.890	0.851
BGA-SBS	371	1.0	1.0	1.0	1.0	0.840	1.0	0.840	0.913	0.983	0.958	0.892	0.923
BGWO-SBS	299	1.0	1.0	1.0	1.0	0.836	1.0	0.836	0.911	0.985	0.950	0.903	0.926
BPSO-SBS	331	1.0	1.0	1.0	1.0	0.937	1.0	0.937	0.968	0.960	0.696	0.951	0.804
CBPSO-SBS	352	1.0	1.0	1.0	1.0	0.827	0.998	0.827	0.905	0.980	0.901	0.899	0.901
HI-BQPSO-SBS	270	1.0	1.0	1.0	1.0	0.882	1.0	0.882	0.937	0.982	0.938	0.890	0.913
LBPSOSEE-SBS	157	1.0	1.0	1.0	1.0	0.980	0.999	0.981	0.990	0.979	0.809	1.0	0.894
NBPSOSEE-SBS	72	1.0	1.0	1.0	1.0	0.926	0.999	0.927	0.962	0.999	1.0	0.994	0.997

calculated by the proposed NBPSOSEE-SBS are close to that of LBPSOSEE-SBS among the arrears' users of the low risk level. On the whole, the proposed algorithm selects the lowest number of features, less than one-third of the original features, and the calculated average accuracy and F1-measure are the highest.

Table 7 shows the test results of the improved NBPSOSEE-SBS and comparison algorithms for ECR risk in August 2018. Among the arrears' users of the low risk level, the proposed NBPSOSEE-SBS achieves the highest accuracy, precision, and F1-measure. The accuracy, precision, and F1-measure of NBPSOSEE-SBS increased by 6.73%, 64.74%, and 53.38%, respectively, compared with the results

calculated using all features. Furthermore, NBPSOSEE-SBS selects the least number of features in the optimal feature subset obtained. The experimental results show that the proposed NBPSOSEE-SBS deletes irrelevant or redundant features more than 90.4% of the original features, getting the highest average value of F1-measure.

The execution time of algorithm is also an important indicator for evaluating the performance of algorithm. The long running time of algorithm means that the complexity of algorithm is high, and the short running time indicates that the complexity of algorithm is low. In this paper, the running time of the proposed NBPSOSEE, NBPSOSEE-SBS, and other algorithms is shown in Figures 8 and 9. In order to

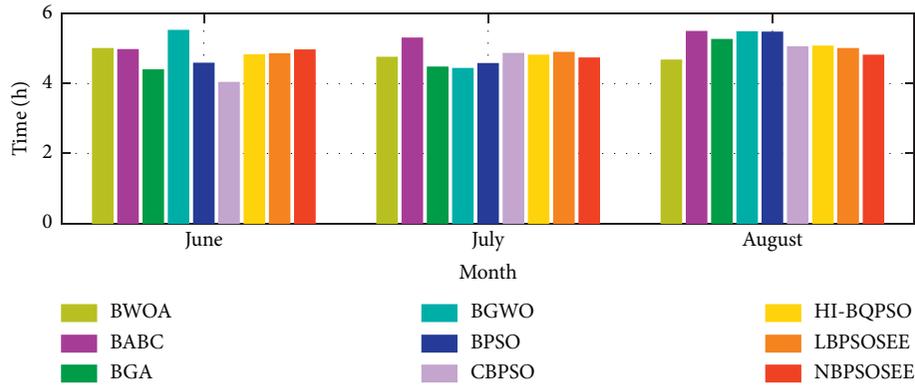


FIGURE 8: Execution time of NBPSOSEE and comparison algorithms.

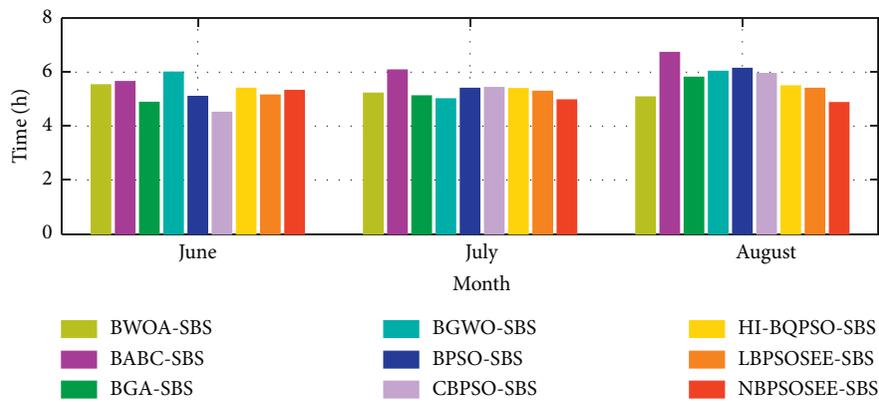


FIGURE 9: Execution time of NBPSOSEE-SBS and other algorithms.

remove more redundant features and improve the classification results, these hybrid feature selection algorithms with SBS take slightly more execution time than those without SBS. The execution time of NBPSOSEE-SBS is not significantly different from other algorithms. The execution speed of NBPSOSEE-SBS is faster than BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, and LBPSOSEE-SBS for the testing in July and August. In summary, the proposed NBPSOSEE-SBS outperforms other algorithms. NBPSOSEE-SBS can effectively reduce a large of redundant features and stably improve the prediction results while keeping low execution time.

8. Conclusion

To solve the problem of accurately predict the risk of ECR for power customers, a hybrid nonlinear inertia weight binary particle swarm optimization with shrinking encircling and exploration mechanism (NBPSOSEE) is proposed for solving feature selection tasks. In addition, an improved feature selection approach that is based on the NBPSOSEE and SBS is proposed, namely, NBPSOSEE-SBS, to select the optimal feature subset. The experimental results prove that the proposed NBPSOSEE-SBS can steadily remove more redundant or irrelevant features and obtain better prediction

results of ECR for power customers in the case of low execution time, compared with one state-of-the-art optimization algorithm, and seven well-known wrapper-based feature selection approaches for the risk prediction of ECR for power customers.

Data Availability

The experimental data contain the user’s privacy. Therefore, in order to protect the security of users, we cannot upload the data set.

Conflicts of Interest

There are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (no. 61672470), Science and Technology Project of Henan Province (no. 202102210351), Doctoral Program of Zhengzhou University of Light Industry (no. 2017BSJJ046), and Key Research Projects of Henan Higher Education Institutions (no. 20A120011).

Supplementary Materials

The experimental results file directory contains three sub-directories: 201806, 201807, and 201808. There are 18 CSV files in the subdirectories 201806, 201807, and 201808, respectively. There are 18 comparison algorithms in the manuscript, which are BWOA, BABC, BGA, BGWO, BPSO, CBPSO, HI-BQPSO, LBPSOSEE, NBPSOSEE, BWOA-SBS, BABC-SBS, BGA-SBS, BGWO-SBS, BPSO-SBS, CBPSO-SBS, HI-BQPSO-SBS, LBPSOSEE-SBS, and NBPSOSEE-SBS. These CSV files in the 201806, 201807, and 201808 represent the experimental results calculated by these comparison algorithms for the risk prediction of electric charge recovery in June, July, and August 2018. The file BWOA_results.csv represents the calculation result of the algorithm BWOA. The file BABC_results.csv represents the calculation result of the algorithm BABC. The file BGA_results.csv represents the calculation result of the algorithm BGA. The file BGWO_results.csv represents the calculation result of the algorithm BGWO. The file BPSO_results.csv represents the calculation result of the algorithm BPSO. The file CBPSO_results.csv represents the calculation result of the algorithm CBPSO. The file HPSO_SSM_results.csv represents the calculation result of the algorithm HI-BQPSO. The file LBPSOSEE_results.csv represents the calculation result of the algorithm LBPSOSEE. The file NBPSOSEE_results.csv represents the calculation result of the algorithm NBPSOSEE. The file BWOA_SBS_results.csv represents the calculation result of the algorithm BWOA-SBS. The file BABC_SBS_results.csv represents the calculation result of the algorithm BABC-SBS. The file BGA_SBS_results.csv represents the calculation result of the algorithm BGA-SBS. The file BGWO_SBS_results.csv represents the calculation result of the algorithm BGWO-SBS. The file BPSO_SBS_results.csv represents the calculation result of the algorithm BPSO-SBS. The file CBPSO_SBS_results.csv represents the calculation result of the algorithm CBPSO-SBS. The file HPSO_SSM_SBS_results.csv represents the calculation result of the algorithm HI-BQPSO-SBS. The file LBPSOSEE_SBS_results.csv represents the calculation result of the algorithm LBPSOSEE-SBS. The file NBPSOSEE_SBS_results.csv represents the calculation result of the algorithm NBPSOSEE-SBS. In the headline of these CSV files, features_num means selected number of features. Accuracy_high, Loss_high, Precision_high, Recall_high, and F1_high represent the accuracy, loss, precision, recall, and F1-measure of predicting the high-risk level of customer electric charge recovery, respectively. Accuracy_med, Loss_med, Precision_med, Recall_med, and F1_med represent the accuracy, loss, precision, recall, and F1-measure of predicting the medium-risk level of customer electric charge recovery, respectively. And Accuracy_low, Loss_low, Precision_low, Recall_low, and F1_low represent the accuracy, loss, precision, recall, and F1-measure of predicting the low-risk level of customer electric charge recovery, respectively. (*Supplementary Materials*)

References

- [1] X. Zhang, M. R. Nakhai, G. Zheng, S. Lambotharan, and J. A. Chambers, "Distributed foresighted energy management in smart-grid-powered cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4064–4068, 2019.
- [2] H. Zhou, Y. Wang, W. Wang, T. Li, and H. Yang, "Predication of default probability of clients' electricity charges arrears based on logistic regression model," *Power System Technology*, vol. 31, no. 17, pp. 85–88, 2007.
- [3] H. Zhou, Y. Wang, W. Wang, T. Li, and H. Yang, "Prediction of default probability of clients' electricity charge arrears," in *Proceedings of the IEEE International Conference on Service Operations and Logistics*, pp. 2582–2586, Beijing, China, October 2008.
- [4] H. Zhou, Y. Wang, W. Wang, T. Li, and H. Yang, "Arrears forewarning model for power clients in electricity market," *Proceedings of the Chinese Society for Electrical*, vol. 28, no. 22, pp. 107–112, 2008.
- [5] J. M. Wang and Y. Q. Wen, "Application of data mining in arrear risks prediction of power customer," in *Proceedings of the International Symposium on Knowledge Acquisition and Modeling*, pp. 206–210, Wuhan, China, December 2008.
- [6] Y. Huang and Y. Yan, "Research of evaluating credit-risk in power enterprise based on SVM and VIKOR method," in *Proceedings of the International Conference on Industrial Engineering and Engineering Management*, pp. 1596–1599, Singapore, December 2008.
- [7] Z. Yang and K. Hong, "Credit evaluation for mobile customers using artificial immune algorithms," in *Proceedings of the 7th World Congress on Intelligent Control and Automation*, pp. 7021–7025, Chongqing, China, June 2008.
- [8] S. Zhang, S. Yang, and B. Wang, "A trust evaluation model of electricity customers based on cloud model," in *Proceedings of the International Conference on Computer Application and System Modeling (ICCASM)*, pp. 644–649, Taiyuan, China, October 2010.
- [9] H. T. Xia, Y. Y. Shi, and W. J. Ling, "Research and practice on risk prediction model of electricity tariff recovery," *Electric Power Information and Communication Technology*, vol. 14, no. 8, pp. 58–63, 2016.
- [10] R. G. Lopes, R. N. Carvalho, M. Ladeira, and R. S. Carvalho, "Predicting recovery of credit operations on a Brazilian Bank," in *Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 780–784, Anaheim, CA, USA, December 2016.
- [11] Z. Xiao, X. Feng, L. Dang et al., "Research and application of subtilized customer clustering algorithm in power marketing," in *Proceedings of the Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1–6, Beijing, China, November 2017.
- [12] W. Guo, W. Hong, W. Li, and K. Guo, "Design and implementation of electric charge arrears prediction system," in *Proceedings of the 12th Web Information System and Application Conference (WISA)*, pp. 309–313, Jinan, China, September 2015.
- [13] X. Zhang, T. Wang, J. Hu, C. Li, Y. Xu, and Z. Zhang, "Study on tariff risk early warning of electric power users based on PSO-SVM algorithm," in *Proceedings of the International Conference on Big Data and Artificial Intelligence (BDAl)*, pp. 92–96, Beijing, China, June 2018.

- [14] W. Fu, D. Zhang, Y. Fu, J. Li, and Y. Xie, "Arrears prediction for electricity customer through Wgan-Gp," in *Proceedings of the 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 1667–1670, Chengdu, China, December 2017.
- [15] S. Adams and P. A. Beling, "A survey of feature selection methods for Gaussian mixture models and hidden Markov models," *Artificial Intelligence Review*, vol. 52, no. 3, pp. 1739–1779, 2019.
- [16] A. Melo and H. Paulheim, "Local and global feature selection for multilabel classification with binary relevance," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 33–60, 2019.
- [17] C. Wang, Y. Lin, and J. Liu, "Feature selection for multi-label learning with missing labels," *Applied Intelligence*, vol. 49, no. 8, pp. 3027–3042, 2019.
- [18] J. Liu, Y. Lin, M. Lin, S. Wu, and J. Zhang, "Feature selection based on quality of information," *Neurocomputing*, vol. 225, pp. 11–22, 2017.
- [19] M. Hammami, S. Bechikh, C.-C. Hung, and L. Ben Said, "A multi-objective hybrid filter-wrapper evolutionary approach for feature selection," *Memetic Computing*, vol. 11, no. 2, pp. 193–208, 2019.
- [20] G. S. Krishnan and S. Sowmya Kamath, "A novel GA-ELM model for patient-specific mortality prediction over large-scale lab event data," *Applied Soft Computing*, vol. 80, pp. 525–533, 2019.
- [21] M. Lu, "Embedded feature selection accounting for unknown data heterogeneity," *Expert Systems with Applications*, vol. 119, pp. 350–361, 2019.
- [22] Y. Li and Z. Yang, "Application of EOS-ELM with binary Jaya-based feature selection to real-time transient stability assessment using PMU data," *IEEE Access*, vol. 5, pp. 23092–23101, 2017.
- [23] Z. Yang, Y. Guo, Q. Niu, H. Ma, Y. Zhou, and L. Zhang, "A novel binary Jaya optimization for economic/emission unit commitment," in *Proceedings of the 2018 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–6, Rio de Janeiro, Brazil, July 2018.
- [24] M. Aslan, M. Gunduz, and M. S. Kiran, "JayaX: Jaya algorithm with xor operator for binary optimization," *Applied Soft Computing*, vol. 82, pp. 105576–105593, 2019.
- [25] M. Sharawi, H. M. Zawbaa, and E. Emary, "Feature selection approach based on whale optimization algorithm," in *Proceedings of the Ninth International Conference on Advanced Computational Intelligence (ICACI)*, pp. 163–168, Doha, Qatar, February 2017.
- [26] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," *Applied Soft Computing*, vol. 62, pp. 441–453, 2018.
- [27] E. H. Houssein, A. G. Hussien, A. E. Hassanien, S. Bhattacharyya, and M. Amin, "S-shaped binary whale optimization algorithm for feature selection," in *Proceedings of the First International Symposium on Signal and Image Processing (ISSIP)*, pp. 166–172, Cairo, Egypt, November 2017.
- [28] M. A. Tawhid and A. M. Ibrahim, "Feature selection based on rough set approach, wrapper approach, and binary whale optimization algorithm," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 3, pp. 573–602, 2020.
- [29] H. Rao, X. Shi, A. K. Rodrigue et al., "Feature selection based on artificial bee colony and gradient boosting decision tree," *Applied Soft Computing*, vol. 74, pp. 634–642, 2019.
- [30] G. Feng and A. P. Engelbrecht, "Binary artificial bee colony optimization," in *Proceedings of the 2011 IEEE Symposium on Swarm Intelligence (SWIS)*, pp. 170–177, Paris, France, April 2011.
- [31] S. Wang and R. Dong, "Feature selection with improved binary artificial bee colony algorithm for microarray data," *International Journal of Computational Science and Engineering*, vol. 19, no. 3, pp. 387–399, 2019.
- [32] S. Oreski and G. Oreski, "Genetic algorithm-based heuristic for feature selection in credit risk assessment," *Expert Systems with Applications*, vol. 41, no. 4, pp. 2052–2064, 2014.
- [33] H. Chen, W. Jiang, C. Li, and R. Li, "A heuristic feature selection approach for text categorization by using chaos optimization and genetic algorithm," *Mathematical Problems in Engineering*, vol. 2013, Article ID 524017, 6 pages, 2013.
- [34] A. K. Shukla, P. Singh, and M. Vardhan, "A new hybrid feature subset selection framework based on binary genetic algorithm and information theory," *International Journal of Computational Intelligence and Applications*, vol. 18, no. 3, pp. 1950020:1–1950020:22, 2019.
- [35] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371–381, 2016.
- [36] Q. Al-Tashi, H. Rais, and S. Jadid, "Feature selection method based on grey wolf optimization for coronary artery disease classification," in *Proceedings of the International Conference of Reliable Information and Communication Technology*, pp. 257–266, Kuala Lumpur, Malaysia, July 2018.
- [37] A. Adeli and A. Broumandnia, "Image steganalysis using improved particle swarm optimization based feature selection," *Applied Intelligence*, vol. 48, no. 6, pp. 1609–1622, 2018.
- [38] Y. Zhang, H.-G. Li, Q. Wang, and C. Peng, "A filter-based bare-bone particle swarm optimization algorithm for unsupervised feature selection," *Applied Intelligence*, vol. 49, no. 8, pp. 2889–2898, 2019.
- [39] A. P. Engelbrecht, J. Grobler, and J. Langeveld, "Set based particle swarm optimization for the feature selection problem," *Engineering Applications of Artificial Intelligence*, vol. 85, pp. 324–336, 2019.
- [40] E. García-Gonzalo and J. L. Fernández-Martínez, "A brief historical review of particle swarm optimization (PSO)," *Journal of Bioinformatics and Intelligent Control*, vol. 1, no. 1, pp. 3–16, 2012.
- [41] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95—International Conference on Neural Networks*, pp. 1942–1948, Perth, Australia, November 1995.
- [42] D. Das, D. K. Pratihar, G. G. Roy, and A. R. Pal, "Phenomenological model-based study on electron beam welding process, and input-output modeling using neural networks trained by back-propagation algorithm, genetic algorithms, particle swarm optimization algorithm and bat algorithm," *Applied Intelligence*, vol. 48, no. 9, pp. 2698–2718, 2018.
- [43] K. R. Harrison, B. M. Ombuki-Berman, and A. P. Engelbrecht, "A parameter-free particle swarm optimization algorithm using performance classifiers," *Information Sciences*, vol. 503, pp. 381–400, 2019.
- [44] S. Barak and T. Mokfi, "Evaluation and selection of clustering methods using a hybrid group MCDM," *Expert Systems with Applications*, vol. 138, pp. 112817–112835, 2019.
- [45] X. Li, X. Wu, S. Xu, S. Qing, and P.-C. Chang, "A novel complex network community detection approach using discrete particle swarm optimization with particle diversity and mutation," *Applied Soft Computing*, vol. 81, pp. 105476–105489, 2019.

- [46] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, pp. 4104–4108, Orlando, FL, USA, October 1997.
- [47] J. C. Bansal and K. Deep, "A modified binary particle swarm optimization for knapsack problems," *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 11042–11061, 2012.
- [48] G. Lin, J. Guan, Z. Li, and H. Feng, "A hybrid binary particle swarm optimization with tabu search for the set-union knapsack problem," *Expert Systems with Applications*, vol. 135, pp. 201–211, 2019.
- [49] A. Ahmadi, Y. Alinejad-Beromi, and M. Moradi, "Optimal PMU placement for power system observability using binary particle swarm optimization and considering measurement redundancy," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7263–7269, 2011.
- [50] Q. Zou and S. Chen, "Enhancing resilience of interdependent traffic-electric power system," *Reliability Engineering & System Safety*, vol. 191, pp. 106557–106574, 2019.
- [51] S. M. Vieira, L. F. Mendonça, G. J. Farinha, and J. M. C. Sousa, "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients," *Applied Soft Computing*, vol. 13, no. 8, pp. 3494–3504, 2013.
- [52] Y. Zhang, S. Wang, P. Phillips, and G. Ji, "Binary PSO with mutation operator for feature selection using decision tree applied to spam detection," *Knowledge-Based Systems*, vol. 64, no. 1, pp. 22–31, 2014.
- [53] B. Tran, B. Xue, and M. Zhang, "A new representation in PSO for discretization-based feature selection," *IEEE Transactions on Cybernetics*, vol. 48, no. 6, pp. 1733–1746, 2017.
- [54] C. S. Yang, L. Y. Chuang, C. H. Ke, and C.-H. Yang, "Boolean binary particle swarm optimization for feature selection," in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 2093–2098, Hong Kong, China, June 2008.
- [55] A. Unler and A. Murat, "A discrete particle swarm optimization method for feature selection in binary classification problems," *European Journal of Operational Research*, vol. 206, no. 3, pp. 528–539, 2010.
- [56] Y. Li, J. Xiao, Y. Chen, and L. Jiao, "Evolving deep convolutional neural networks by quantum behaved particle swarm optimization with binary encoding for image classification," *Neurocomputing*, vol. 362, pp. 156–165, 2019.
- [57] L. Y. Chuang, J. C. Li, and C. H. Yang, "Chaotic binary particle swarm optimization for feature selection using logistic map," in *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, pp. 131–136, Hong Kong, China, March 2008.
- [58] L.-Y. Chuang, C.-H. Yang, and J.-C. Li, "Chaotic maps based on binary particle swarm optimization for feature selection," *Applied Soft Computing*, vol. 11, no. 1, pp. 239–248, 2011.
- [59] L.-Y. Chuang, C.-S. Yang, K.-C. Wu, and C.-H. Yang, "Gene selection and classification using Taguchi chaotic binary particle swarm optimization," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13367–13377, 2011.
- [60] J. Liu, Y. Mei, and X. Li, "An analysis of the inertia weight parameter for binary particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 666–681, 2016.
- [61] Q. Wu, Z. Ma, J. Fan, G. Xu, and Y. Shen, "A feature selection method based on hybrid improved binary quantum particle swarm optimization," *IEEE Access*, vol. 7, pp. 80588–80601, 2019.
- [62] M. S. Uzer, O. Inan, and N. Yilmaz, "A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA," *Neural Computing and Applications*, vol. 23, no. 3-4, pp. 719–728, 2013.
- [63] R. C. Eberhart, Y. Shi, and J. Kennedy, *Swarm Intelligence*, Morgan Kaufmann, San Mateo, CA, USA, 2001.
- [64] M. Clerc and J. Kennedy, "The particle swarm—explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.
- [65] B. H. Menze, B. M. Kelm, R. Masuch et al., "A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data," *BMC Bioinformatics*, vol. 10, no. 1, p. 213, 2009.
- [66] S. H. Walker and D. B. Duncan, "Estimation of the probability of an event as a function of several independent variables," *Biometrika*, vol. 54, no. 1-2, pp. 167–179, 1967.
- [67] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215–232, 1958.
- [68] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.