Fault Diagnosis Strategy of Polymerization Kettle Equipment Based on Support Vector Machine and Cuckoo Search Algorithm

Shuo Wen, Jie-Sheng Wang, and Jie Gao

Abstract—**For the requirements of the real-time fault diagnosis and optimized monitoring of the polymerization kettle, a large key device of polyvinyl chloride resin (PVC) production process, a fault diagnosis strategy of polymerization kettle based on support vector machine (SVM) is proposed. Firstly, a mapping between polymerization process datum and the fault mode is established by analyzing the production technology of PVC polymerization process. Then the cuckoo search (CS) algorithm is used to optimize the penalty factor and kernel function parameters of SVM. After optimizing the SVM parameters by CS algorithm, the fault pattern classification of polymerization kettle equipment is to realize the nonlinear mapping from symptom set to fault set according to the given symptom set. Finally, the fault diagnosis simulation experiments are conducted by combining with the industrial on-site historical datum of polymerization kettle and the results show that the CS - SVM fault diagnosis strategy is effective.**

Index Terms—**polymerization kettle equipment, fault diagnosis, support vector machine, cuckoo search algorithm**

I. Introduction

POLYVINYL chloride (PVC) is one of the five largest thermoplastic synthetic resins, and its production is second only to the polyethylene (PE) and polypropylene (PP). PVC is a kind of general colophon, which is good in quality and is widely used. It has good mechanical properties, ant chemical properties and it is corrosion-resistant and difficult to burn [1]. With vinyl chloride monomer (VCM) as a raw material, the suspension method to produce polyvinyl chloride (PVC) resin is a kind of typical batch chemical production process. PVC polymerization process is a complex control system with muti-variable, uncertain, nonlinear, and strong coupling. Polymerization kettle is the key equipment of the PVC production process, where vinyl chlorides go on the polymerization reaction to generate polyvinyl chloride [1]. Whether the polymerization kettle can run steadily is directly related to the working conditions of the PVC production device. On the other hand, the motor, reducer, and machine seal are key equipment to ensure that the polymerization kettle device runs normally. Once they failed to work, the serious losses will be brought to the PVC polymerizing process. Therefore, the earlier diagnosis of the fault type and location of polymerization kettle can avoid the huge economic losses which are caused by the parking of polymerization kettle, which has the important practical significance to improve the product quality and reduce the production costs [2].

Support vector machines (SVM) is a new pattern recognition method developed based on the statistical learning theory, which is also a reflection in reality of the structural risk minimization thought of the statistical learning theory [3-4]. The basic idea of SVM is that the input space is transformed into a higher dimensional space by the nonlinear mapping to obtain the optimal separating hyper-plane in the new space. It exhibits many unique advantages in solving the problems of small samples, nonlinear and high dimensional pattern recognition [5-6]. For effectively improving the efficiency and accuracy of automotive systems fault diagnosis, a motor fault diagnosis method based on SVM optimized by the particle swarm optimization (PSO) algorithm is put forward [7]. The PSO algorithm is used to optimize the SVM with the multilevel binary tree structure in order to realize the fault classification. In order to realize the intelligent fault diagnosis of motor bearing, the artificial bee colony (ABC) algorithm is used to optimize the SVM parameters [8]. The artificial fish swarm algorithm (AFSA) is used to optimize the SVM parameters. The simulation experiments results on the hydro-logical forecast field show that the training speed of the proposed method is better than the standard SVM [9].

In this paper, a fault diagnosis strategy of the polymerization kettle based on SVM optimized by the improved CS algorithm is proposed. The simulation results show the effectiveness of the proposed fault

Manuscript received June 29, 2017; revised July 30, 2017. This work was supported by the Program for Research Special Foundation of University of Science and Technology of Liaoning (Grant No. 2015TD04) and the Opening Project of National Financial Security and System Equipment Engineering Research Center (Grant No. USTLKFGJ201502 and USTLKEC201401).

Shuo Wen obtained the bachelor degree in the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail: 249195878@qq.com).

Jie-Sheng Wang is with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China; National Financial Security and System Equipment Engineering Research Center, University of Science and Technology Liaoning. (phone: 86-0412-2538246; fax: 86-0412-2538244; e-mail: wang_jiesheng@126.com).

Jie Gao is a postgraduate student in the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail: 540948998@qq.com).

diagnosis strategy. The paper is organized as follows. In section 2, the technique flowchart of the PVC polymerization process is introduced. The support vector machine is presented in section 3. In section 5, the SVM fault diagnosis method of polymerization kettle optimized by improved CS algorithm is introduced. The simulation experiments and results analysis are introduced in details in section 5. Finally, the conclusion illustrates last part.

II. Polyvinyl Chloride (PVC) Polymerization Process

A. Technique Flowchart

Four methods (suspension polymerization, emulsion polymerization, bulk polymerization and liquor polymerization) are usually used in the PVC polymerization process of. Among them, the suspension polymerization is one of the most widely used methods. The typical technique process of PVC polymerization kettle is shown in Figure 1 [2]. In PVC polymerization process, various raw materials and additives are added to the reaction kettle, which are full evenly dispersed under the mixing action. Then the suitable amounts of the initiators are added to the kettle and start to react. The cooling water is constantly poured into the jacket and baffle of reaction kettle to removing the reaction heat. The reaction will be terminated and the final products are obtained when the conversion ratio of the vinyl chloride (VCM) reaches a certain value and a proper pressure drop appears. Finally, after the reaction completed and VCM contained in slurry separated by the stripping technique, the remaining slurry is fed into the drying process for dewatering and drying.

B. Information Table of Fault Diagnosis System

The proposed fault diagnosis system it applied a certain 70 M^3 polymerization kettle from a large chemical company with the measured data. The main parameters of the polymerization kettle fault diagnosis are shown in Table 1. It can be seen form Table 1 that the main parameters of polymeric kettle include stirring speed (r/min), stirring electric current (A), polymeric kettle pressure (MPa), polymeric kettle temperature (℃), mechanical seal pressure (MPa) and mechanical seal temperature (℃). These six parameters of polymerization kettle are respectively noted as a, b, c, d, e and *f*. The motor fault, shaft seal fault, the damage of sealing components and the running smoothly are respectively represented by 1, 2, 3 and 4. Large amounts of on-spot data are collected from the PVC polymerization kettle as input samples and testing samples of the neural network fault diagnosis system. The history working data of polymeric kettle are shown in Table 2.

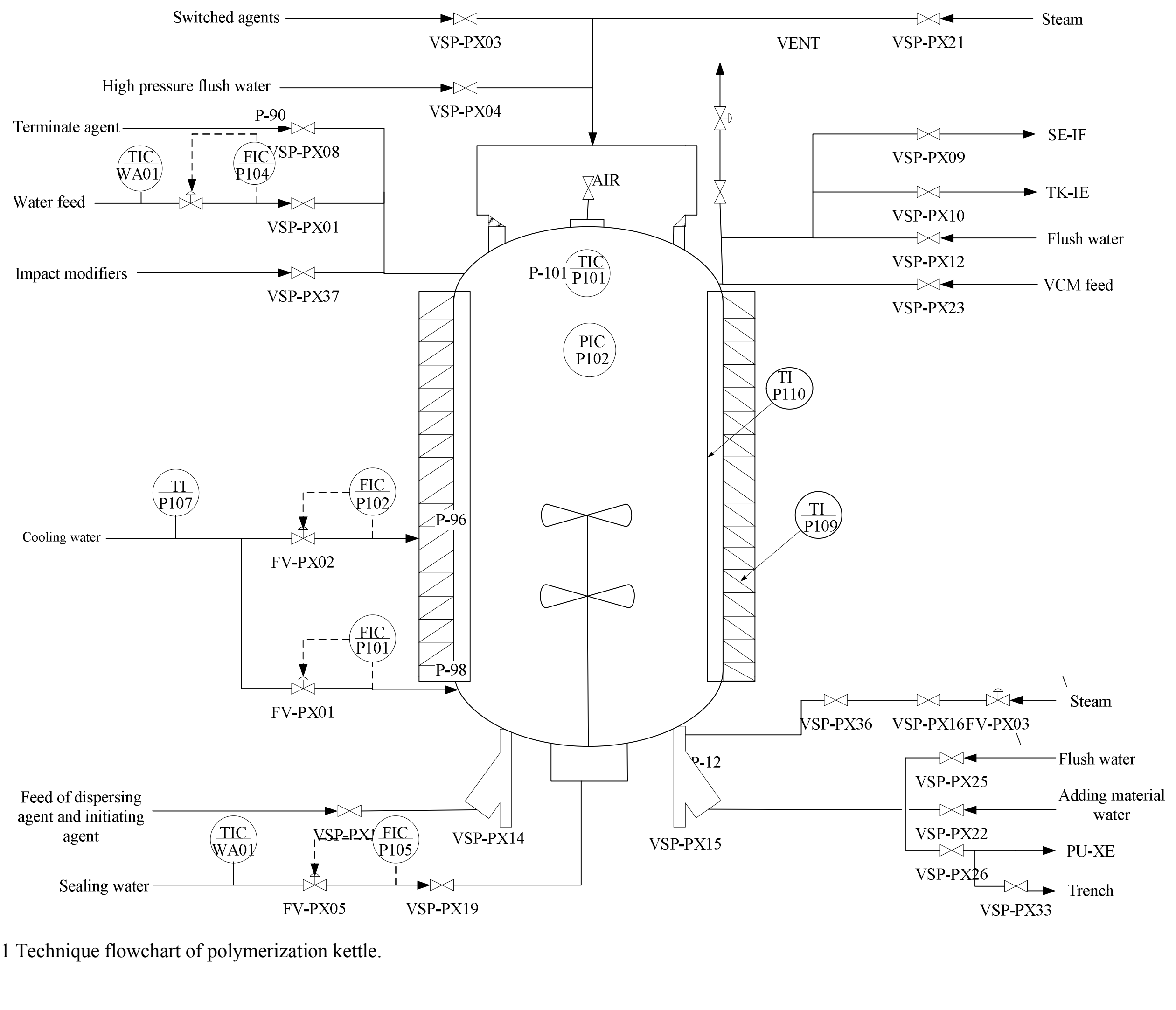

Fig. 1 Technique flowchart of polymerization kettle.

TAB. 1 MAIN PARAMETERS OF POLYMERIZATION KETTLE

Parameters	Symbol	Parameters range	Fault type
Stirring speed	r/min	45r/min~98 r/min	Motor fault
Stirring electric current	A	120A~175A	Motor fault
Polymerization kettle pressure	MPa	0.8MPa~1.3MPa	Shaft seal fault
Polymerization kettle temperature	℃	50℃~69℃	Shaft seal fault
Mechanical seal pressure	MPa	0.8MPa~1.5MPa	Damage of sealing components
Mechanical seal temperature	℃	55℃~80℃	Damage of sealing components

TAB. 2 HISTORICAL DATA OF POLYMERIC KETTLE

Sample	Historical datum of polymerizer						Diagnosis type
	a	*b*	*c*	*d*	*e*	*f*	
1	56.71	128	1.90	54.96	0.98	58.84	2
2	56.67	160	0.82	55.16	0.98	60.75	4
3	113	128	0.86	54.51	0.97	58.75	1
4	56.67	136	0.80	54	2.6	58.75	3
5	56.75	192	0.82	55.13	0.99	59.36	1
6	56.77	137	0.82	79	1	59.52	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
380	56.48	146	0.83	55.49	1	109	3

III. SUPPORT VECTOR MACHINE

Support vector machine (SVM) is a new kind of machine learning method based on statistical learning theory (SLT) [10]. SVM obtains the optimal promotion ability for unknown sample classification by constructing the optimal hyper-plane, which transforms the hyper-plane optimization into a dual problem of solving the convex quadratic programming (CQP) problem and obtain a global unique optimal solution [11-12].

SVM is a supervised learning method. Suppose the linear separable samples are (x^1, y^1), (x^2, y^2), …, (x^n, y^n), $x \in R^d$. $y \in \{+1,-1\}$ is the category label. The objective function, the constraint conditions and the optimal classification function of SVM are respectively described as follows.

$$Q(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j (x_i\cdot x_j) \quad (1)$$

$$\sum_{i=1}^{n}\alpha_i y_i=0, \alpha_i\ge 0, i=1,2,\cdots,n \quad (2)$$

$$\begin{aligned} f(x)&=\operatorname{sgn}\{(w^*\cdot x)+b^*\} \\ &=\operatorname{sgn}\left\{\sum_i \alpha^*{}_i y_i (x_i\cdot x)+b^*\right\} \end{aligned} \quad (3)$$

where $Q(\alpha)$ is the objective function, $f(x)$ is the optimal classification function, α_i^* is the Lagrange multiplier for the *i*th support vector, b^* is the classification threshold calculated by any support vector and sgn $(\cdot)$ is the symbolic function.

For the nonlinear (linear inseparable) problem, the input space is mapped to a high-dimensional feature space through the proper nonlinear transformation. In this space the optimal function is constructed to make the defined risk function minimal. Based on the Mercer theorem, the mapping ϕ and the kernel function $K(x_i,x_j)$ exists to make $K(x_i,x_j)=\varphi(x_i)\cdot\varphi(x_j)$. Then the optimal hyper-plane in the high-dimensional space Z is solved. In other words, the optimal classification function is determined by selecting the proper kernel function to maximize the objective function parameters.

If it is still inseparable in high dimensional space, the standard C-support vector machine (C-SVC) with nonlinear soft interval having the slack variables and penalty factor is adopted to realize the correct classifies. This classifier allowed a certain classification error (soft interval) exists. In order to enhance the promotion capacity, the concept of the soft edge optimal hyper-plane is introduced. That is to say a non-negative variables $\xi_i \ge 0$ is introduced, so the constraint condition is relaxed as follows.

$$y_i(w \cdot x_i + b) - 1 + \xi_i \geq 0 \quad (4)$$

Minimizing $\sum_{i=1}^{n}\xi_i$ can minimize the wrong classification of samples, so the optimization problem is described as follows.

$$\begin{aligned} &\min(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i) \\ &s.t\ \ y_i(w \cdot x_i + b) - 1 + \xi_i \geq 0\ (i = 1,2,...,n) \\ &\qquad \xi_i \geq 0 \end{aligned} \quad (5)$$

where $\xi_i \geq 0$ is the slack variable, C is the penalty factor for a comprehensive consideration of the minimum wrong classification of samples and the largest classification interval. According to the Lagrange method, the original problem is transformed into the following dual problem to be solved.

$$\begin{aligned} &\max \sum_i \alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j K(x_i, x_j) \\ &s.t \qquad \sum_i \alpha_i y_i = 0 \\ &\qquad\qquad 0 \leq \alpha_i \leq C \end{aligned} \quad (6)$$

The accordance discriminator function described in Eq. (3) can be changed as follows.

$$f(x) = \mathrm{sgn}(\sum_i \alpha_i y_i K(x_i, x) - b) \quad (7)$$

Now the commonly used inner product kernel functions includes the polynomial function (PF), the radial basic function (RBF) and the sigmoid function, etc. The different kernel function is in corresponding to the different support vector. In this paper, the RBF is adopted shown as Eq. (8).

$$K(x_i, x) = \exp\{-\frac{|x_i - x|^2}{\gamma^2}\} \quad (8)$$

where, x is the input vector, x is the support vector (i = 1, ⋯ , s), s is the number of support vectors and γ is the width of kernel function.

Vapnik found that the kernel parameter γ and error penalty factor C are the key factors affecting the SVM performance. The kernel function parameters γ mainly affects the distribution complexity of the sample data in high dimensional feature space. While the effect of error penalty factor C is to regulate the proportion of the confident range of machine learning and experience risk in the certain feature space. The SVM model based on the swarm intelligence optimization algorithms (such as genetic algorithm, particle swarm algorithm, etc) is widely used in the fault diagnosis systems [13-14. In this paper, the SVM model parameters (γ and C) are optimized by the improved CS algorithm.

IV. SVM Fault Diagnosis Method of Polymerization Kettle Optimized by Improved CS Algorithm

A. Basic Principle of Cuckoo Search Algorithm

Cuckoo search (CS) algorithm is proposed based on the magical natural phenomenon and an artificial processing [15-16], which is mainly based on cuckoo's nest - parasitic propagation mechanism and the Levy flights search principle. This simulation algorithm is one of the latest available heuristic algorithms inspired by nature. In nature, the cuckoo seeks bird's nest location in a random way or a way that similar to the random way. Firstly, in order to imitate the mode that the cuckoo seeks the optimal nest, three ideal states are supposed [17-18].

(1) The cuckoo selects bird's nest location randomly and only lays one egg each time.

(2) The best bird's nest location is chosen from a group of random selected bird's nests and reproduced to the next generation.

(3) The size n of the bird's nests is fixed. Suppose the probability of a nest master finding eggs are not its owns is $p_a = [0,1]$.

Based on the above conditions, the procedure of CS algorithm steps are described as follows.

Step 1: Initialize settings. N bird's nest location $X_0 = (x_1^0, x_2^0, \cdots, x_N^0)$ are generated randomly, which are put into fitness function to conduct the select operator. The optimal bird's nest location can be chosen, which will be reproduced to the next generation.

Step 2: Search the bird's nest locations. By calculating and updating the locations by using Eq. (9), the nest location of the next generation begins to be searched to get a new group of bird's nest locations, and then they are brought into the fitness function. The more optimal nest location is selected to enter the next step by comparing it with the bird's nest location of previous generation.

$$x_i^{(t+1)} = x_i^t + \alpha \oplus levy(\lambda)\,,\quad i = 1,2,\cdots,n \quad (9)$$

Step 3: Bird's nest location selection. The probability of the host discovering alien eggs $p_a = 0.25$ is compared with the random number $r \in [0,1]$, which complies with the uniform distribution. If $r > p_a$, x_i^{t+1} is be changed randomly, otherwise, it remains fixed. The fitness of the bird's nest location having been changed is calculated, then it is compared with the optimal bird's nest of previous generation. The optimal bird's nest location $X_t = (x_1^t, x_2^t, \cdots, x_N^t)$ is recorded. Finally, the optimal nest location pb_t^* is selected.

Step 4: Accuracy or iterative judgment. Judge whether $f(pb_t^*)$ achieves the target accuracy or the terminating conditions of iteration is satisfied. If it meets the requirements, pb_t^* is the global optimal solution gb, otherwise pb_t^* is kept to the next generation, return to

Step (2) and continue the iterating and updating process.

According to the above four steps, the CS algorithm not only uses the Levy flight search method (global search) but also introduce the elite reserved strategy (local search), which makes the CS algorithm owns the global and local search abilities. The purpose of the step (3) fed into CS algorithm is to increase the diversity of solutions, which will prevent the algorithm from falling into the local optimum and make it achieve the global optimal.

Levy flight in the CS algorithm has the character of strong randomness. Broadly speaking, the Levy flight is a random walk, whose step size complies with the Levy distribution, while whose direction of walking is the uniform distribution. The step length vector of CS algorithm is determined by Mantegna rule with the characteristics of Levy distribution. In Mantegna rule, the step size s is designed as follows.

$$s = \frac{u}{|v|^{1/\beta}} \tag{10}$$

where u and v comply with the normal distribution, which can be expressed as follows.

$$u \sim N(0, \sigma_u^2), v \sim N(0, \sigma_v^2) \tag{11}$$

$$\sigma_v = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta} \tag{12}$$

$$\sigma_u = 1 \tag{13}$$

In this paper, the extraction method of direction complies with the uniform distribution. The search model of the CS algorithm is the Levy flight. For instance, the i th cuckoo in the t th generation produces the solution of the next generation x_i^{t+1}, which is described as follows.

$$x_i^{t+1} = x_i^t + \alpha \oplus Levy(\lambda) \tag{14}$$

where $\oplus$ is the point to point multiplication. The step size $Levy(\lambda)$ complies with the random walk of Levy distribution, which can be expressed as follows.

$$Levy \sim u = t^{-\lambda}, \quad (1 < \lambda \leq 3) \tag{15}$$

The Mantegna rule is used to calculate the specific step size. In Eq. (14), α is the step size controlled variable, which is mainly used to control the direction and size of the step.

$$\alpha = O(L/10) \tag{16}$$

where L is the size of the search space for optimization problem.

Thus some new solutions are generated gradually to achieve the optimal solution through the Levy flight walking around the optimal solution, which can speed up local search. On the contrary, some part new solutions are far from the current optimal solution, which is randomly generated by deviating from the remote location. The main purpose of these solutions is to ensure that the system does not fall into the local optimal solution.

The parameters of CS algorithm include the step length control variable α, the scale of bird's nest group n and the detection probability p_a. The size of the search space has influence on the value α. For instance, if the search space is relatively small and the search step length is very big, it is easy to ignore the space region that contains a relatively optimal solution. At this time, the step length control variable α will be used. The specific extraction method of α is related to the dimensionality of the search space. If the search space is determined, the value of step length control α is fixed. In principle, the bigger the bird's nest group scale, the faster the search speed. A large numbers of simulation experiments prove that when the value of bird's nest group is $n = 15 \sim 40$ and $p_a = 0.25$, most of the optimization problems can be solved effectively. Once the nest group scale n is fixed, the detection probability p_a is an important parameter for balancing the local and the global search and the elitist selection. Therefore, the CS algorithm has many good characteristics, such as the little parameters, the strong global optimization ability.

Markov chain definition: a random process $\{X_n, n \in T\}$, $T = \{0,1,2,...\}$, $I = \{0,1,2,...\}$. If $\forall n \in T$, $\forall i_0, i_1, ..., i_n, i_{n+1} \in I$, then

$$\begin{aligned} &P(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, ..., X_n = i_n) \\ &= P(X_{n+1} = i_{n+1} \mid X_n = i_n) \end{aligned} \tag{17}$$

Then $\{X_n, n \in T\}$ is called as a Markov chain. It can be seen from the above definition:

$$\begin{aligned} &P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) \\ &= P(X_n = i_n \mid X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}) \\ &\quad \times P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}) \\ &= P(X_n = i_n \mid X_{n-1} = i_{n-1}) \\ &\quad \times P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}) \\ &\quad \cdots\cdots \\ &= P(X_n = i_n \mid X_{n-1} = i_{n-1}) \times P(X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}) \\ &\quad \times .... \times P(X_1 = i_1 \mid X_0 = i_0) \times P(X_0 = i_0) \end{aligned} \tag{18}$$

It can be seen that once the initial distribution of the Markov chain $P(X_0 = i_0)$ is given, its statistical properties will be entirely determined by the conditional probability $P(X_n = i_n \mid X_{n-1} = i_{n-1})$.

B. Improved Cuckoo Search Algorithm

Though the CS algorithm has stronger global search ability, if the detection probability p_a is fixed, both the better solution and the worse solution will be replaced with the probability p_a. If p_a is relatively small, it will affect the convergence speed of current worse solution. If p_a is relatively big, it is difficult for the better solutions converge to the optimal solution [19-20]. Aiming at the above shortcomings, a dynamic method is adopted to adjust the detection probability p_a and make p_a change according to the current states of the overall solutions. In addition, in order to find the optimal solution effectively in the search space and improve the accuracy, three improved CS algorithm are proposed to balance the global convergence and local convergence.

1) Dynamic Detection Probability p_a

In order to adapt to the searching characteristic of CS algorithm, all the values that have been searched are compared with the optimal value in the beginning, which requires the step of bird's nest location is very big. With the number of iteration increasing, a growing number of bird's nests are becoming more and more close to the optimal value. At this time, the change rate of bird's nest location is more and more tiny. So the dynamic detection probability shown in Eq. (19) is adopted.

$$p_a = \frac{p_{a.\max} - p_{a.\min}}{p_{IterNum} - 1} \times (p_{IterNum} - 1) \tag{19}$$

where $p_{a.\max}$ is the maximum detection probability, $p_{a.\min}$ is the minimum detection probability, $p_{IterNum}$ is the current iteration number and $p_{Iter\max}$ is the maximum iteration number.

2) Improvement of the CS Algorithm

In nature, most of creatures not only use Levy flight but also combine other flight patterns. For example, a shark is accompanied by a Levy flight with the brown characteristic in its foraging process, which can improve the searching efficiency and lively of this algorithm. In this algorithm, the "cognitive part", "social part" and "ring topology part" are introduced in Levy flight, which can be expressed as follows.

$$a = u(\lambda) + R_1(N_{1b} - x_i^{(t)}) + R_2(N_{gb} - x_i^{(t)}) + R_3\left[\frac{\sum_{i=1}^{r} N_r(i)}{r} - x_i^{(t)}\right] \tag{20}$$

$$x_i^{(t+1)} = x_t^{(t)} + R_4 a \tag{21}$$

where N_{1b} is the individual best position of each bird's nest; N_{gb} is the global best location of all the bird's nest; $N_r(i)$ is the i th bird's nest position for a ring topology; R_1, R_2, R_3 and R_4 respectively complies with the uniform distribution. If the number of bird's nest group is $n = 5$, the corresponding bird's nest number is (N_1, N_2, N_3, N_4, N_5). If the bird's nest number with the ring topology is $r = 3$, the corresponding ring topology respectively are (N_1, N_2, N_3), (N_2, N_3, N_4), (N_3, N_4, N_5), (N_4, N_5, N_1) and (N_5, N_1, N_2).

3) Improvement of Levy Flight Method in CS Algorithm

In the standard CS algorithm, the most simple method to calculate Levy flight is described as follows.

$$s = \frac{u}{|v|^{1/\beta}} \tag{22}$$

where β is a parameter to control the random process, s is a random variable and u complies with the $N(0, \sigma_u^2)$ distribution, in which:

$$\sigma_u = \left[\frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{2^{(\beta-1)/2}\Gamma((1+\beta)/2)\beta}\right]^{-1/\beta} \tag{23}$$

where v compiles with the $N(0,1)$ distribution and $\Gamma(\bullet)$ is the Euler function.

For the bigger random variables, this method has an excellent levy flight characteristic. But for the smaller random variables, this characteristic is not obvious. Therefore, a method put forward by Leccardi M is adopted to improve the Levy calculation method in CS algorithm, which is described as follows.

$$w = \left[(K(a)-1)\exp\left[\frac{-|s|}{C(a)}\right]+1\right] \bullet s \tag{24}$$

where a is the exponent parameter, w is the independent random variable and $C(a)$ is the pending optimal value.

$$K(a) = \frac{a\Gamma((a+1)/2a)}{\Gamma(1/a)}\left[\frac{a\Gamma((a+1)/2)}{\Gamma(1+a)\sin(\pi a/2)}\right]^{1/a} \tag{25}$$

The relationship between $K(a)$ and $C(a)$ is described as follows.

$$\begin{aligned} &\frac{1}{\sigma_u}\int_0^{+\infty} q^{1/a} \exp\left[-\left[\frac{q^2}{2} + \frac{q^2 C(a)^2}{2\sigma_u^2}\right]\right] dq \\ &= \int_0^{+\infty} \cos\left[\left[\frac{K(a)-1}{e}+1\right]C(a)\right] \exp(-q^a) dq \end{aligned} \tag{26}$$

where q is the integral variable. Then the Levy distribution is described as follows.

$$z = c^{1/a}\frac{1}{n^{1/a}}\sum_{i=1}^{n} w_i \tag{27}$$

where c is the scaling factor and w_i is a component of w.

Based on the above discussion, the flowchart of the improved cuckoo algorithm (ICSA) is shown as Figure 2.

C. SVM Optimized by Improved CS Algorithm

The kernel function used in SVM is the RBF kernel function, which has two very important parameters (the kernel function parameter γ and the penalty factor C). γ has a great influence on the RBF kernel function and the kernel function has a great influence on SVM, which mainly affects the complexity of sample data distributing in subspace. The penalty factor C mainly determines the confidence interval range of data subspace when adjusting learning machine. The flowchart of the improved CS algorithm based SVM is shown in Figure 3. The algorithm procedure is described as follows.

Step 1: Initialize the parameters of CS algorithm, generate the bird's nests randomly and set up the eggs' detection probability.

Step 2: Initialize the cuckoo species and the position of each cuckoo laying eggs. Select the optimal bird's nest location by the calculated fitness and reserve it to the next generation.

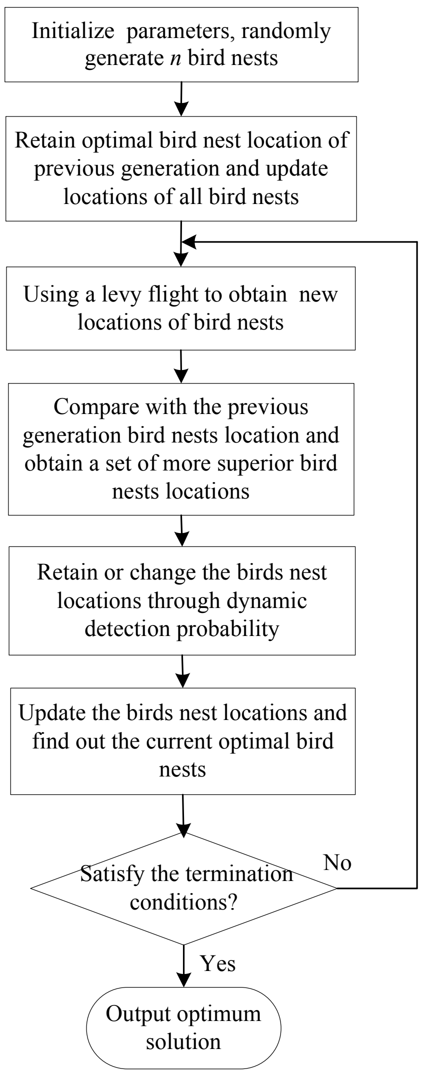

Fig. 2 Flowchart of the improved cuckoo search algorithm.

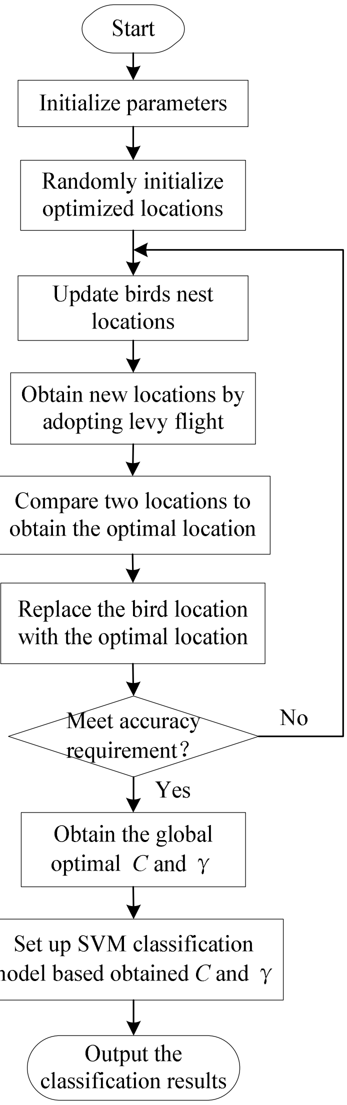

Fig. 3 Flowchart of SVM optimized by improved CS algorithm.

Step 3: Train SVM for each iteration and calculate the error precision. When the requited precision or the iterations is reached, the optimal parameters of C and γ are obtained to be used to realize the optimized SVM.

Step 4: Otherwise, when it does not reach the required iterations or precision, return to the previous step, reselect the bird's nest locations and update the nest location.

Step 5: Repeat from the Step 2 to the Step 4 until the termination conditions are met (reach the maximum iteration number or meet the precision requirement).

Step 6: Finally, the optimized the penalty factor C and the kernel function parameter γ is used in the SVM classifier.

V. SIMULATION EXPERIMENTS AND RESULTS ANALYSIS

The parameters of CS algorithm are initialized as follows. The initial population of CS algorithm is 40, the initial positions are generated randomly. The probability of finding eggs is 0.25, the precision requirement is 0.00001, the initial penalty factor C is 1 and the parameter γ is 0.1. In the fault diagnosis system, 380 groups of sample data are divided into 8 parts, in which the 7 sample sets have 47 samples and the last sample set has 51 samples. The 8-fold cross-validation is adopted in this paper, that is to say one sample set is selected randomly as the testing sample set and the other 7 sample sets are selected as the training samples. Eight times training simulation experiments are carried out. The diagnosis results are shown in Figure 4 and Table 3.

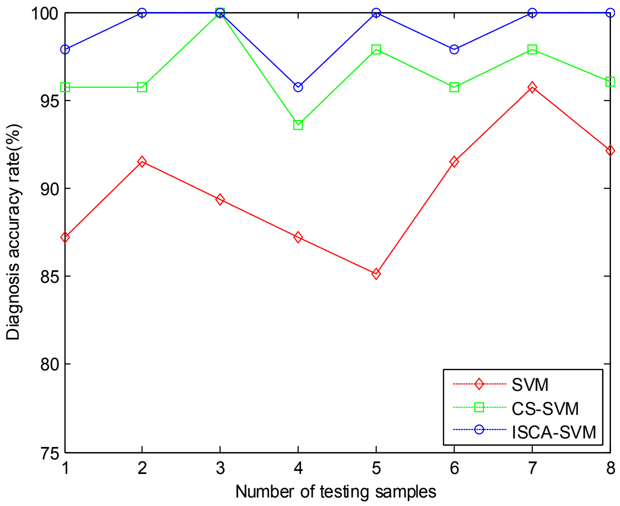

Fig. 4 Comparison of diagnosis accuracy rate.

TAB. 3 FAULT DIAGNOSIS RESULTS

Sample	Method	C	γ	Accuracy ratio (%)
	SVM	21.56	0.76	87.23
1	CS-SVM	34.45	0.90	95.74
	ICSA-SVM	70.50	0.81	97.87
	SVM	52.84	0.75	91.49
2	CS-SVM	63.86	0.59	95.74
	ICSA-SVM	164.42	1.54	100
	SVM	54.12	1.98	89.36
3	CS-SVM	75.98	1.24	100
	ICSA-SVM	86.45	0.57	100
	SVM	34.75	0.76	87.23
4	CS-SVM	37.45	1.92	93.62
	ICSA-SVM	153.42	0.31	95.74
	SVM	86.41	1.87	85.12
5	CS-SVM	95.86	1.64	97.87
	ICSA-SVM	129.26	0.98	100
	SVM	29.45	1.75	91.49
6	CS-SVM	33.56	1.45	95.74
	ICSA-SVM	55.87	0.85	97.87
	SVM	54.86	1.86	95.74
7	CS-SVM	68.45	1.68	97.87
	ICSA-SVM	99.78	1.35	100
	SVM	36.45	1.15	92.17
8	CS-SVM	44.75	1.32	96.07
	ICSA-SVM	117.41	1.54	100

The average of accuracy rate under three diagnosis methods are 89.98%, 96.58% and 98.94% respectively. Then 300 groups samples are selected randomly as the training samples and the rest 80 groups' samples are treated as the test samples.

The simulation experiments results are shown in Figure 5 - Figure 7. In Figure 5, among 80 groups test samples, 9 samples are diagnosed falsely and the diagnosis rate is 88.75%. In the similar way, four samples are diagnosed falsely and diagnosis rate is 95% in Figure 6. In Figure 7, only one sample is diagnosed falsely and diagnosis rate is 98.75%. Through the validation of two methods, the support vector machine (SVM) optimized by the improved cuckoo algorithm (ICSA) has a better diagnosis accuracy. The 10-fold cross-validation adopted in this paper is more convincing than only a single experiment.

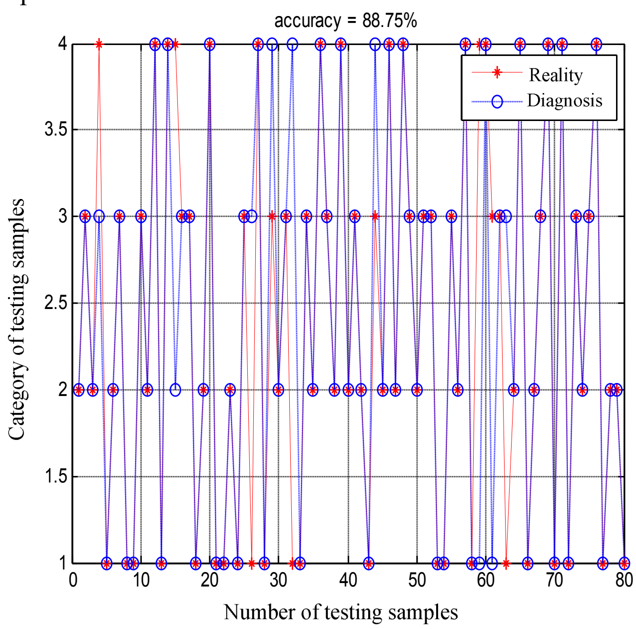

Fig. 5 Comparison of diagnosis results based on SVM.

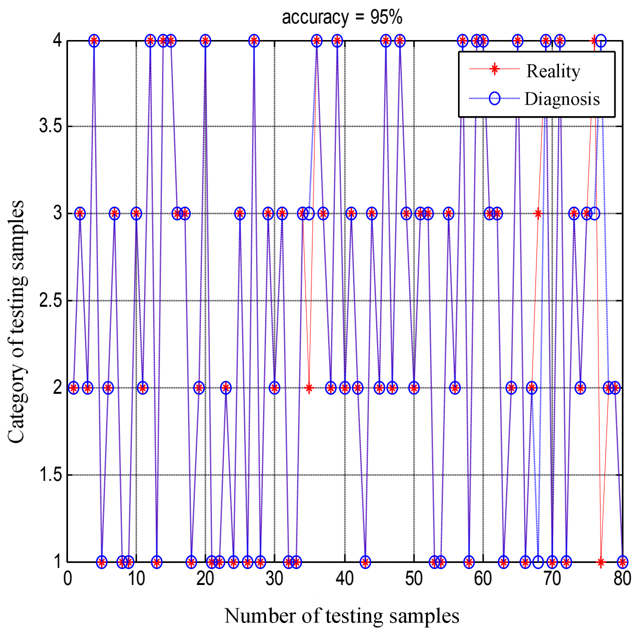

Fig. 6 Comparison of diagnosis results based on SVM optimized by CS algorithm.

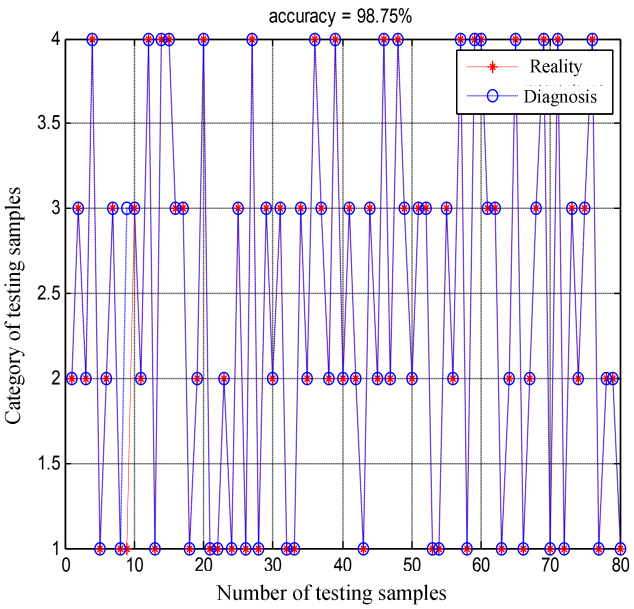

Fig. 7 Comparison of diagnosis results based on SVM optimized by improved CS algorithm.

VI. Conclusion

A fault diagnosis strategy of polymerization kettle based on SVM is put forward in this paper. Fault data are classified through SVM optimized by CS algorithm. Combining with the industrial site historical datum of polymerization kettle, the simulation experiments of fault diagnosis are conducted. The results show that the SVM fault diagnosis strategy proposed is effective.

References

[1] S. Zhou, G. Ji, Z. Yang, and W. Chen, "Hybrid intelligent control scheme of a polymerization kettle for ACR production," *Knowledge-Based Systems*, vol. 24, no.7, pp. 1037-1047, 2011.

[2] S. Z. Gao, J. S. Wang, and N. Zhao, "Fault diagnosis method of polymerization kettle equipment based on rough sets and BP neural network," *Mathematical Problems in Engineering*, vol. 2013, no.2, pp. 1-8, 2013.,

[3] C. P. Shen, J. W. Lin, F. S. Lin, et al, "GA-SVM modeling of multiclass seizure detector in epilepsy analysis system using cloud computing," *Soft Computing*, vol. 21, no.8, pp. 1-11, 2017.

[4] S. Ding, Z. Zhu, and X. Zhang, "An overview on semi-supervised support vector machine," *Neural Computing & Applications*, vol. 28, pp. 1-10, 2017.

[5] I. Aydin, M. Karakose, and E. Akin, "A multi-objective artificial immune algorithm for parameter optimization in support vector machine," *Applied Soft Computing*, vol. 11, no.1, pp. 120-129, 2011.

[6] D. Niu, Y. Wang, and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization," *Expert Systems with Applications*, vol. 37, no.3, pp. 2531-2539, 2010.

[7] Z. T. Yu, "Automotive fault diagnosis based on SVM and particle swarm algorithm," *Application Research on Computers*, vol. 29, no.2, pp. 572-574, 2012.

[8] L. Liu, T. Y. Wang, "Support vector machine optimization based on artificial bee colony algorithm," *Journal of Tianjin University*, vol. 44, no.9, pp. 803-808, 2011.

[9] Y. P. Zhang, X. H. Lu, and M. Jin, "Hydrological forecasting system model based on AFSVM," *Application Research on Computers*, vol. 27, no.8, pp. 2902-2906, 2010.

[10] X. Zhang, J. Zhou, J. Guo, et al, "Vibrant fault diagnosis for hydroelectric generator units with a new combination of rough sets and support vector machine," *Expert Systems with Applications*, vol. 39, no.3, pp. 2621-2628, 2012.

[11] N. Saravanan, V. N. S. Siddabattuni, K. I. Ramachandran, "Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM)," *Applied Soft Computing*, vol. 10, no.1, pp. 344-360, 2010.

[12] S. Fei, and X. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Systems with Applications*, vol. 36, no.8, pp. 11352-11357, 2009.

[13] J. Huang, X. Hu, and F. Yang, "Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker," *Measurement*, vol. 44, no.6, pp. 1018-1027, 2011.

[14] C. Zhao, X. Sun, S. Sun, and T. Jiang, "Fault diagnosis of sensor by chaos particle swarm optimization algorithm and support vector machine," *Expert Systems with Applications*, vol. 38, no.8, pp. 9908-9912, 2011.

[15] J. S. Wang, and S. X. Li, "PID decoupling controller design for electroslag remelting process using cuckoo search algorithm with self-tuning dynamic searching mechanism," *Engineering Letters*, vol. 25, no. 2, pp. 125-133, 2017.

[16] Q. Liao, S. Zhou, H. Shi, and W. Shi, "Parameter estimation of nonlinear systems by dynamic cuckoo search," . *Neural Computation*, vol. 29, no.4, pp. 1103-1123, 2017.

[17] K. Chandrasekaran, and S. P. Simon, "Multi-objective scheduling problem: hybrid approach using fuzzy assisted cuckoo search algorithm," *Swarm and Evolutionary Computation*, vol. 5, pp. 1-16, 2012.

[18] G. Kanagaraj, S. G. Ponnambalam, and N. Jawahar, "hybrid cuckoo search and genetic algorithm for reliability–redundancy allocation problems," *Computers & Industrial Engineering*, vol. 66, no.4, pp. 1115-1124, 2013.

[19] E. Valian, S. Mohanna, and S. Tavakoli, "Improved cuckoo search algorithm for feedforward neural network training," *International Journal of Artificial Intelligence & Applications*, vol. 2, no.3, pp. 36-43, 2011.

[20] S. X. Li, and J. S. Wang, "Improved cuckoo search algorithm with novel searching mechanism for solving unconstrained function optimization problem," *IAENG International Journal of Computer Science*, vol. 44, no.1, pp. 8-12, 2017.

Shuo Wen obtained the bachelor degree in the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail: 249195878@qq.com).

Jie-Sheng Wang is with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China; National Financial Security and System Equipment Engineering Research Center, University of Science and Technology Liaoning. (phone: 86-0412-2538246; fax: 86-0412-2538244; e-mail: wang_jiesheng@126.com).

Jie Gao is a postgraduate student in the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, 114051, PR China (e-mail: 540948998@qq.com).