A New Support Vector Machine Model Based on Improved Imperialist Competitive Algorithm for Fault Diagnosis of Oil-immersed Transformers

Yiyi Zhang*, Hua Wei*, Ruijin Liao**, Youyuan Wang[†], Lijun Yang** and Chunyu Yan***

Abstract – Support vector machine (SVM) is introduced as an effective fault diagnosis technique based on dissolved gases analysis (DGA) for oil-immersed transformers with maximum generalization ability; however, the applicability of the SVM is highly affected due to the difficulty of selecting the SVM parameters appropriately. Therefore, a novel approach combing SVM with improved imperialist competitive algorithm (IICA) for fault diagnosis of oil-immersed transformers was proposed in the paper. The improved ICA, which is proved to be an effective optimization approach, is employed to optimize the parameters of SVM. Cross validation and normalizations were applied in the training processes of SVM and the trained SVM model with the optimized parameters was established for fault diagnosis of oil-immersed transformers. Three classification benchmark sets were studied based on particle swarm optimization SVM (PSOSVM) and IICASVM with four multiple classification schemes to select the best scheme for transformer fault diagnosis. The results show that the proposed model can obtain higher diagnosis accuracy than other methods. The comparisons confirm that the proposed model is an effective approach for classification problems.

Keywords: Power transformer, Fault diagnosis, Dissolved gases analysis, Support vector machine, Improved imperialistic competitive algorithm, Cross validation, Classification

1. Introduction

An oil-immersed power transformer is the core equipment for electric energy conversion in the electric power system. A fault in a transformer may result in not only substantial repair cost but also power interruptions to thousands of customers, therefore it is essential to assess its working condition online and detect the potential fault as soon as possible [1].

In the past few years, some monitoring methods, including windings displacement [2] and hot spot temperature [3], were applied to detect faults of oil-immersed power transformers, however DGA is still a more convenient and effective online monitoring method comparing to the above methods [4]**.** If an oil-immersed transformer is subjected to electrical or thermal stress, some gases, emitting from its oil-paper insulation system, mainly including H_2, CH_4, C_2H_6, C_2H_4 and C_2H_2, can be regarded as a symbol of a potential fault [5]. Many approaches such as IEEE standard [6], International Electro technical Commission (IEC)

Method 60599 [7], and Duval's Triangle [8], have been applied to detect faults of oil-immersed transformers based on the ratios of the gases for many years. The main drawback of these approaches is that the observed gas ratios do not match all the fault types [9, 10]. To improve the drawback and to dig the potential fault law from the gases content, some new fault diagnosis methods, including fuzzy logic [11], Dempster Shafer theory [12], grey clustering [13] and rough set [14], self-adaptive RBF neural network [15], auto-associative neural networks and mean shift [16], evolutionary wavelet neural network based on genetic algorithm [17], and support vector machine (SVM) based methods [18-20], have been proposed in the recent years. In these methods, SVM is an effective method and is wildly applied in classification problems (fault diagnosis of transformers based on DGA is often considered as a classification problem) [21-22] because its advantages of small sample learning, global optimization, and structural risk minimization [23-24]. In general, the performance of the support vector machine (SVM) have a significant impact factors, including the kernel function, the punishment coefficient C, and kernel parameter γ, which are uncertain and should be optimized by artificial intelligence (AI) approaches. Some AI methods such as clonal selection algorithms [18], PSO [19], GA [20], were used to optimize the two parameters in the classification problems and receive a good performance.

ICA, firstly put forward by Esmaeil Atashpaz Gargari

and Caro Lucas in 2007 [25], is an effective AI optimization algorithm for optimizing the SVM parameters because this algorithm has good global convergence solutions and efficient local search ability. Therefore, in this paper, a new support vector machine model based on improved imperialist competitive algorithm was proposed to find out whether the approach can detect faults of transformers accurately and effectively.

The following is the main content of this paper. The first section is a short introduction. Section 2 briefly describes the imperialist competition algorithm. In section 3, a new SVM based on the improved ICA is proposed to settle the problems of parameters optimization. In section 4, the multiclass SVM model based on the improved ICA is explored for fault diagnosis of power transformers. Results are discussed and conclusions are drawn in section 5 and 6, respectively.

2. Improved Imperialist Competitive Algorithm Based on Differential Evolution

The imperialist competition algorithm (ICA) is a new intelligence optimization algorithm inspired by the competition of the Empire. In 2007, Esmaeil Atashpaz Gargari and Lucas firstly put forward this algorithm, which offered a new way to solve continuous-optimization problems effectively [26-27].

The first step of the ICA is to initialize the population. The second step of this algorithm is to search the scope efficiently through several specific steps. At last, this algorithm ends to the optimal solution or near optimal solution. The initial population is made up of the countries, and the most powerful countries are called imperialists, the other countries are called the colonies. Imperialists and the colonies together constitute the empire. In ICA, an array is used to represent the initial countries, and the array dimension N is the number of countries N, defined as $[p_1, p_2, …,p_n]$. The following formula (1) is used to compute the cost of the ith country s_i:

$$s_i = F(country_i) = F(p_{i1}, p_{i2}, ..., p_{iN}) \qquad (1)$$

Some powerful countries (the countries with minimum cost), N_{imp} countries, are chosen to be the imperialists from the N initial countries. In the meanwhile, the remaining countries are colonies that belong to the imperialists. The goal of ICA processes is to obtain the most powerful country with minimum cost.

In traditional ICA, the information interaction between empires are shown in the imperialist competitive step, however imperialist competition just returns the weakest colony to the strongest empire, which will weaken information interaction degree between empires. Therefore, the improved imperialist competitive algorithm (IICA) based on a differential evolution arithmetic operator was

used in the paper- the following equations are added between the assimilation step and competitive step [26]:

1. Every colony has a differential mutation with a possibility of MR calculated by Eq. (2)

$$G = Cl_{r3} + S(Cl_{r1} - Cl_{r2}) \qquad (2)$$

where Cl_{r1}, Cl_{r2}, Cl_{r3} are three random colonies; $S \in [0,2]$ is the scaling factor.
2. Every Dimension Has Differential Crossover Calculated by Eq. (3)

$$B_i = \begin{cases} G_i, if \ rand < RN \\ Cl_i, otherwise \end{cases} \qquad (3)$$

where $RN \in [0,1]$ is a random number in $[0,1]$.
3. Greed strategy is used when the power of the new colony B is bigger than the original colony; when $f(B) < f(Cl)$, the position changes.

Fig. 1 is the flow chart of IICA for parameters optimization, the specific steps described as follows [26]:

Step 1: Initialize parameters of ICA, and set up fitness function model.

Step 2: By moving the colonies, imperialist and its colonies constitute an empire.

Step 3: Use a differential evolution arithmetic operator and calculate the total cost of the empires.

Step 4: Implement the competition of the imperialists.

Step 5: If the termination condition is satisfied, that is, the selected condition is satisfied, and the algorithm is terminated and the optimal parameters are selected. Loop to Step 2 if the stopping condition is not satisfied. The ending condition is shown as follows:

1. All empires and colonies are controlled by the unique empire. In that case, the position of the united empire is the best solution of the optimal problem.
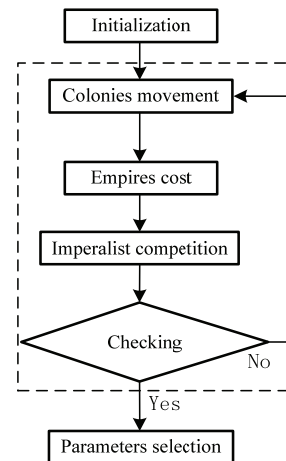2. A maximum generations are reached.

Fig. 1. Flow chart of parameters optimization

3. Classification using Multiclass SVM Based on IICA

3.1 Establishment of the nonlinear-multiple classify-cation SVM

In a multiple classification problem, how to establish the optimal objective function is the most important process because it largely determines the classification accuracy. To form the optimal objective function in a multiple classification problem, the traditional SVM (linear-two classification) should first be replaced by a nonlinear-multiple classification SVM. Two steps should be done as follows.

3.1.1 Transform the linear SVM to the nonlinear SVM

Supposing that $\{(x_1, y_1), \ldots, (x_i, y_i), \ldots, (x_l, y_l)\}$ is a training set (including input data $x_i \in \mathbf{R}^n$ and output $y_i \in \mathbf{R}$) and a class label $y_i \in \{-1, +1\}$ is determined by input x_i. The data in the training set can be considered as a separable set only when the training set is bound to satisfy the formula (4) [19].

$$\begin{cases} \omega^T \varphi(x_i) + b \leq -1, & \text{if } y_i = -1 \\ \omega^T \varphi(x_i) + b \geq +1, & \text{if } y_i = +1 \end{cases} \qquad (4)$$

The vector ω with the bias b can be defined as the optimal hyper plane. In the above formula, the mapping $\varphi(x_i)$ is defined as a nonlinear mapping. Therefore, $\varphi(x_i)$ and the ω vector are both infinite dimension.

In order to solve the following quadratic programming in (5) with the constraints in (6), the concept of SVM is introduced.

$$\min \Phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{l} \xi_i \qquad (5)$$

$$\begin{cases} y_i[\omega^T \varphi(x_i) + b] \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, 2, \ldots, l \end{cases} \qquad (6)$$

In the above formula, ξ_i is the slack variable and C is the penalty parameter. Therefore, formula (7) shows the Lagrange function simulation.

$$\begin{aligned} L(\omega, b, \xi, \alpha, \beta) &= \Phi(\omega, \xi) \\ &- \sum_{i=1}^{l} \alpha_i \{y_i[\omega^T \varphi(x_i) + b] - 1 + \xi_i\} - \sum_{i=1}^{l} \beta_i \xi_i \end{aligned} \qquad (7)$$

In the formula, the Lagrange multiplier is α_i and β_i. The Lagrangian function is transformed to the following quadratic programming (QP) problem (8) with the constraints in (9).

$$\max \Psi(\alpha) = -\frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) + \sum_{i=1}^{l} \alpha_i \qquad (8)$$

$$\sum_{i=1}^{l} \alpha_i y_i = 0, \alpha_i \in [0, C], i = 1, \ldots, l \qquad (9)$$

In the formula (8), $K(x_i, x_j) = \varphi(x_i)\varphi(x_j)$ represents a kernel function.

At last, the formula (10) represents the nonlinear two-class classifier.

$$y(x) = \text{sign}\left[\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b \right] \qquad (10)$$

3.1.2 Transform the two-class SVM to the multiple-class SVM

Traditional SVM is a dimidiate classification method for two-class classification problems. In practical application, it often needs to extend the two classification support vector machine to multi classification support vector machine for solving multiple-class classification problems. The OAA (one-against-all), ECOC (error correcting output codes), MOC (minimal output coding), OAO (one against one) scheme are good methods [28] to transform and extend, so the methods are compared and with each other in the next section. Take OAO for example, using OAO scheme to solve K-class problems needs to construct $K (K - 1)/2$ discriminant function, in other words, this scheme requires $N = K(K-1)/2$ times training by using two-class classifiers. If the two-class SVM classifier is used to solve a three-class (A_1, A_2, A_3) classification problem, two-class SVM is trained to obtain the classification results between A_1 and "not A_1" for the first time. For the second time, two-class SVM is trained to obtain the classification results between A_2 and "not A_2". For the third time, two-class SVM is trained to obtain the classification results between "not A_1 and not A_2" (class A_3) and "not A_3". Therefore, the data is divided into three classes by three times.

3.2 Establishment of the optimal objective function

After obtaining the model of the nonlinear-multiple classification SVM, the selection of kernel function of SVM is important for establishing the optimal objective function. There are four kernel functions in the present study, which are linear, polynomial, sigmoid colon and radial basis function (RBF). The RBF is a real valued function that depends only on the specific value of the distance. The RBF kernel function, shown in Eq. (11), is effective and less parameters-setting needed [29, 30], so it is used in the paper.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \ \gamma > 0 \qquad (11)$$

In formula (11), γ is a free parameter in the RBF kernel function and it is inversely proportional to the width of RBF.

The performance of classification accuracy is the key of classification problems, therefore, the optimized objective function can be defined as the average classification accuracy during m-fold cross validation in (12). In the equation, the value of the target function equals to the negative of the classification accuracy. To apply the IICA for parameters selection, the country cost function, shown in (1), is set as the optimized objective function. Two important parameters of objective function (C and γ) need to be set up in advance. The smaller the objective function value (target value) is, the better the classification accuracy of the classification. In a word, the best value of the target function is the same as the cost of the best country in IICA. When all the countries are united to the unique empire, the cost of the unique empire will be the lowest value in IICA, and the target function will get its minimum value at the same time. Assuming that the multiple classification SVM is an uncertain function (UF) in a black box and C and γ are the unknown parameters of the function. If the lowest value of the target function is obtained, the best parameters C and γ can be calculated by the inverse function of the function (UF).

$$Target = -\frac{1}{m}\sum_{i=1}^{m}\left(\frac{p_T^i}{p^i}\times 100\%\right) = UF(c, \gamma)$$
$$= s_i = F(country_i) = F(p_{i1}, p_{i2}) \tag{12}$$

In the formula (12), m is the number of folds in the cross validation; p^i is the number of validated subsets; in this subset, p^i_T is the number of correctly classified samples; UF is the uncertain function.

According to the IICA steps in section 2, the optimal parameters are obtained. The steps can be summarized as follows [24]:

Step 1: Initialize parameters of IICA, and set up fitness function model.

Step 2: By moving the colonies, imperialist and its colonies constitute an empire.

Step 3: Exchange positions between the imperialist and the colony.

Step 4: Count the total cost of the empires.

Step 5: Implement the IICA algorithm.

Step 6: If the termination condition is satisfied, the algorithm is completed and the optimal parameters are selected.

3.3 Processes of classification based on IICASVM

After the establishment of the optimal objective function, IICA is applied to select the best parameters of the nonlinear-multiple class SVM (C andγ).The processes can be shown in Fig. 2 and summarized as following steps:

Step 1: Collect data of the characteristic quantities and the classification results to form the attributes set and the classes set, respectively. Combine the attributes set and the
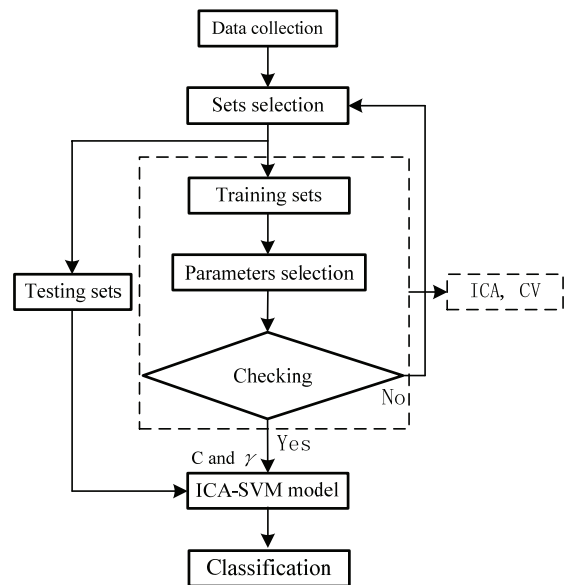

Fig. 2. Processes of classification based on IICASVM

classes set to be the samples set.

Step 2: Select a large portion of the attributes set and the classes set to be the training set. Set the remaining portion of the attributes set and the classes set to be the testing set and the verifying set, respectively.

Step 3: Use the training set and the nonlinear-multiple classification SVM with undetermined parameters to form the optimal objective function whose best solution is the best parameters of the nonlinear-multiple classification SVM with the best classification accuracy. The optimal objective function is formed by the average classification accuracy during m-fold cross validation which is better than traditional optimal objective function.

Step 4: Apply IICA to find out the best solution of the optimal objective function to determine the best parameters of SVM for classification. If the parameters meet the criterion of the parameters checking, the parameters are reserved to be the best parameters of SVM; otherwise, go to the step 3.

Step 5: Use the training set and the best parameters to train the nonlinear-multiple classification SVM to be a trained classifier.

Step 6: Apply the trained classifier to determine the classification results of the testing set (testing samples). Compare the classification results obtained based on the proposed approach and the verifying set (actual samples classification results) to analyze the effectiveness of the proposed approach.

3.4 Classification and verification

The nonlinear-multiple class SVM with the determined parameters is trained by the training set to form a trained classifier. The trained classifier is applied to determine the classification results of the testing set (testing samples).

Classification accuracy is the ratios between testing set classification results and the classification results of actual samples.

4. Benchmark Cases Analysis for Selecting Multiple Classification Schemes

In order to test the general applicability of the IICASVM model for classification problems, three benchmark data sets are analyzed. These sets include the iris data set, wine data set, and glass identification data set [31]. Table 1 shows some of the basic information of the three data sets. Since the PSOSVM and IICASVM were the two effective approaches for classification problems, three benchmark data sets were classified and compared by the two approaches with four different multiple classification approaches including OAA (one-against-all), ECOC (error correcting output codes), MOC (minimal output coding), OAO (one against one) multiple classification schemes.

The number of countries (IICA) and the particle number (PSO) are set to be 10; the candidate parameters C and γ are set and they vary in the range of two fixed ranges, $[10^{-1}\ 10^2]$ and $[10^{-2}\ 10^2]$. Before the training, all input data are normalized within the [0, 1] range. Table 2 shows the classification results based on benchmark data sets. As shown in Table 2, the two approaches with four multiple classification schemesobtained 100.00% classification accuracy in the testing phases on the iris data. However, the training performance of IICASVM is better than the PSOSVM in the iris data (MOC: 98.3333% VS 97.5000%, ECOC: 98.3333% VS 97.5000%, OAO: 100.00% VS 98.3333%, OAA: 100.00% VS 97.5000%). The two approaches with four multiple classification schemesobtained 100% classification accuracy in the training phases on the wine data; however, the testing performance of IICASVM with ECOC scheme is better than the PSOSVM with ECOC scheme in iris data (100% VS 96.6667%). Besides, the IICASVM had a better training and testing performance than the PSOSVM on the glass identification data (MOC: 92.5287% VS 92.5287%, ECOC: 89.6552% VS 87.3563%, OAO: 91.80% VS 89.6552%, OAA: 89.6552% VS 89.6552%), (MOC: 72.5000% VS 72.5000%, ECOC: 72.5000% VS 67.5000%, OAO: 73.9100% VS 72.5000%, OAA: 67.5000% VS 67.5000%). Compared with 47.96% to 73.81% testing accuracy by applied other approaches in glass identification

Table 1. Benchmark classification data based on UCI repository

Data sets	Class	Attributes	Instances	Sample number	
				Training	Testing
Iris	3	4	150	120	30
Wine	3	13	178	148	30
Glass identification	6	9	214	174	40

Table 2. Experimental results using benchmark data

Data sets		Parameters		Classification accuracies		Gener-ation
		C	γ	Training	Testing	
Iris (IICASVM)	MOC	15.1245	3.3551	98.3333%	100.00%	100
	ECOC	66.7089	1.8437	98.3333%	100.00%	100
	OAO	46.5246	4.8349	100.00%	100.00%	100
	OAA	50.9111	4.2415	100.00%	100.00%	100
Iris (PSOSVM)	MOC	18.1919	14.6835	97.5000%	100.00%	100
	ECOC	42.6324	14.6900	97.5000%	100.00%	100
	OAO	99.9568	22.1372	98.3333%	100.00%	100
	OAA	36.5871	18.9373	97.5000%	100.00%	100
Wine (IICASVM)	MOC	77.7161	2.5249	100.00%	100.00%	100
	ECOC	80.2676	1.2745	100.00%	100.00%	100
	OAO	85.6377	1.0338	100.00%	100.00%	100
	OAA	27.4499	2.7922	100.00%	100.00%	100
Wine (PSOSVM)	MOC	96.1237	93.0120	100.00%	100.00%	100
	ECOC	40.0040	32.3873	100.00%	96.6667%	100
	OAO	16.3059	92.4436	100.00%	100.00%	100
	OAA	62.1652	43.6327	100.00%	100.00%	100
Glass identification (IICASVM)	MOC	30.5801	8.0412	92.5287%	72.5000%	100
	ECOC	85.3111	14.8039	89.6552%	72.5000%	100
	OAO	25.9828	2.0285	91.80%	73.9100%	100
	OAA	1.6252	29.0383	89.6552%	67.5000%	100
Glass identification (PSOSVM)	MOC	88.0702	9.9412	92.5287%	72.5000%	100
	ECOC	30.7324	7.1844	87.3563%	67.5000%	100
	OAO	19.3800	10.3313	89.6552%	72.5000%	100
	OAA	42.9012	6.7332	89.6552%	67.5000%	100

[13], the accuracy of the glass identification using the IICASVM reach to 73.913%, which demonstrates the effectiveness of the proposed method. Experimental results of IICASVM and PSOSVM on benchmark data are shown in Fig. 3. The figures with no grid show the average cost and minimum cost of the objective function versus iteration by IICA. The figures with grid show the average value and minimum value of the objective function versus iteration by PSO. According to Table 2 and Fig. 3, IICA has a much better performance in processing generation than PSO.

5. Fault Diagnosis of Power Transformers Based on IICASVM

Dissolved gas data of transformers is the key symbol for fault diagnosis of transformers. Fault diagnosis of oil-immersed transformers based on DGA is often considered as a multiple class problem [18, 19], therefore the proposed IICASVM algorithm described in section 3 was explored for fault diagnosis of transformers in the paper.

This paper collects DGA data from several electric power companies to build fault types of models. The data contains 270 samples, which matched with five diagnosis types, i.e. low-energy discharge, high-energy discharge, thermal fault of low and medium temperature, thermal fault of high temperature and normal condition.
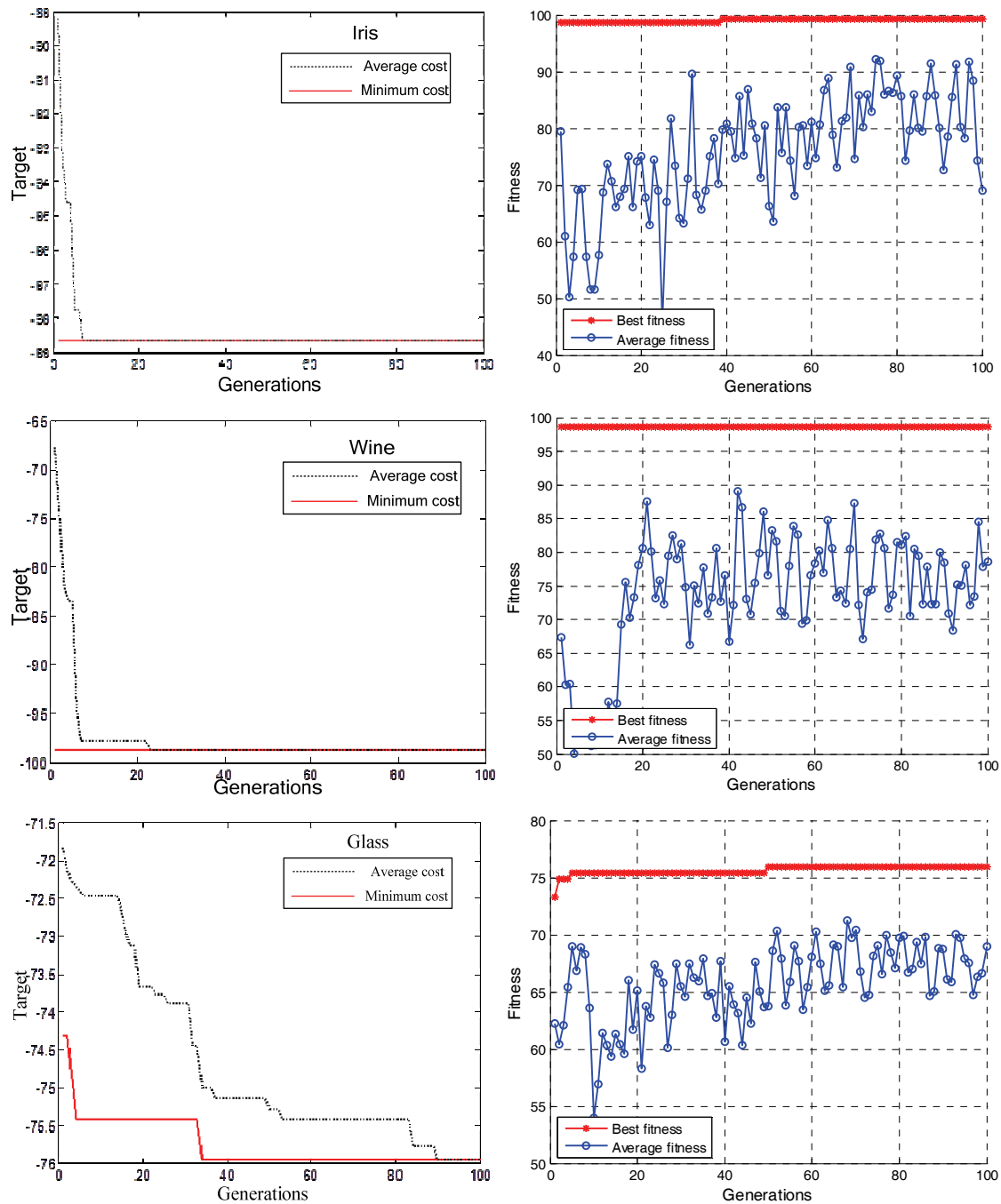
Fig. 3. Experimental results of IICAOAOSVM and PSOOAOSVM on benchmark data

The 270 samples were randomly divided into the training sample set containing 243 samples and the test sample set containing 27 samples. The input of the classifier is the six commonly gases used to detect transformer faults, i.e. H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, and total hydrocarbon content (THC). All input data were normalized in the range of [0 1] before training to improve the generalization performance of SVM [30].

The target function shown in (12) was defined to be the optimized objective function. Cross validation (10-fold, which means m=10) was applied in the function. The

target of IICA is to realize the process of the optimization of the parameters. When the candidate parameters C and γ are set, they vary in the range of two fixed ranges, $[10^{-1} \ 10^2]$ and $[10^{-2} \ 10^2]$, respectively. The parameters of the IICA algorithm used in this paper are as follows: the number of countries and the number of initial imperialists were fixed to 20 and 6, respectively; the dimension of the optimized function was set to 2; the maximum number of generations was 100; revolution rate were set to 0.3; assimilation coefficient equaled to 2 and assimilation angle coefficient equaled to 0.5.

Table 3. Diagnosis results of the five approaches

NO.	DGA content （μL/L）						Results					
	H_2	CH_4	C_2H_2	C_2H_4	C_2H_6	THC	IEC	RBFNN	SVM	PSO OAOSVM	IICA OAOSVM	Actual fault
1	11.9	12.4	1.0	13.6	5.9	32.9	LM-T	LE-D	LE-D	LE-D	LE-D	LE-D
2	145	68.4	578.2	151.2	1.4	799.2	None*	HE-D	HE-D	HE-D	HE-D	HE-D
3	5.1	9.5	1.3	47.9	5.9	64.6	HT	HT	HT	LM-T	LM-T	LM-T
4	30	62	3.4	460	60	585.4	HT	HE-D	HE-D	HE-D	HT	HT
5	63.0	20.1	93.5	49.0	18.6	181.2	LE-D	HE-D	HE-D	HE-D	HE-D	HE-D
6	12.1	3.3	0.2	2.2	9.3	15	Normal	Normal	Normal	Normal	Normal	Normal

* 'None' means no matching code based on IEC
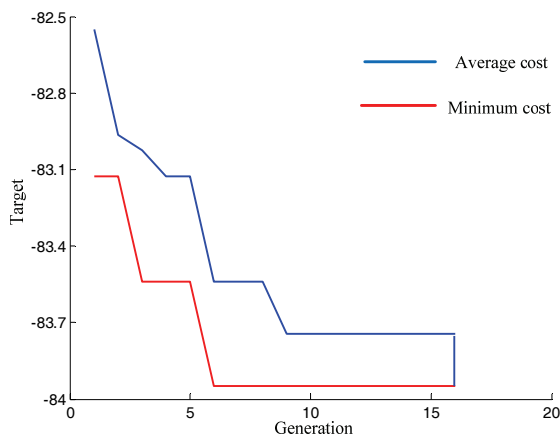
Fig. 4. Average cost and minimum cost of all imperialists versus iteration by IICA

Fig. 5. Classification accuracy of testing samples

The average cost and the minimum cost of all imperialists in each generation are shown in Fig. 4. At the beginning, the minimum cost of all imperialists is much smaller than the average cost of all imperialists. However, following by the increasing generation, the difference between them is increasingly small. At the 16th generation, the minimum cost of all imperialists is equal to the average cost of all imperialists. At the first generation, the number of imperialists is 6, however the calculation ends and the number changes to1 at the 16th generation. The cost of the imperialist is the minimum value of the target function. The average cost and minimum cost of all imperialists versus iteration by IICA can be shown in Fig. 3. The value of the target function was -83.95. The elapsed time of fault diagnosis is about 17.46 seconds. At last, the optimal parameters of the multi class support vector machine classifier based on IICA algorithm are $C = 62.3245$, $\gamma = 0.0187$.

According to the selected two parameters, the OAO-nonlinear multiple classification SVM classifier is trained and applied for fault diagnosis of transformers. According to the actual condition, the collected samples of 243 transformers are divided into five classes. This condition include low-energy discharge (LE-D, class 1), high-energy discharge (HE-D, class 2), thermal fault of low and medium temperature (LM-T, class 3), thermal fault of high temperature (HT, class 4), and normal condition (Normal,
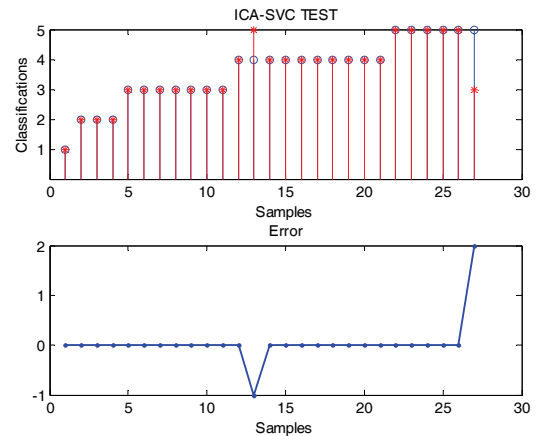
class 5). After cross validation (10-fold) and nominalization, the classification accuracy of training samples hit 90.535% (220/243). Besides, Fig. 5 shows the testing results of 27 samples for fault diagnosis. Apparently, only the 13th sample and the 27th sample were classified incorrectly and the classification accuracy of test samples hit 92.59% (25/27). The possible reason of the error may be that the two cases are bad data with measuring errors, or they are near the classifying surfaces' border between normal condition and HT, LM-T and normal condition, respectively.

Five fault diagnosis methods are applied to compare separately. Those methods include IEC criteria [6], RBFNN, SVM, PSOSVM [19] and IICASVM. Since OAO scheme is proofed to be the best multiple classification scheme in section 4, the OAO scheme is applied in PSOSVM and IICASVM. The work was done by Matlab in a laptop (Inter Core (TM) i3-2330M CPU, 2.20 GHz). The first training of RBFNN was performed in 30 experiments. Then the best networks and spread of the radial basis kernel can be chosen. RBFNN was searched from 1 to 10^6. When the candidate parameters C and γ are set, they vary in the range of two fixed ranges, $[10^{-1} \ 10^2]$ and $[10^{-2} \ 10^2]$, respectively. The population sizes were set to be 20. The maximum number of generations equal to 100. Before the training, all input data are normalized within the [0 1] range. By using the PSOSVM [19], the optimal parameters C and γ were 53.2457 and 0.0142 respectively.

Table 4. Test and training accuracies using five diagnosis methods

Diagnosis approaches	C	σ^2	Classification accuracies	
			Training	Testing
IEC criteria	/	/	78.6%(191/243)	74.07%(20/27)
RBFNN	/	/	100%(243/243)	77.78% (21/27)
SVM	50.5555	2.4865	76.95%(187/243)	85.18%(23/27)
PSOOAOSVM	56.2548	2.6541	90.54%(220/243)	88.89%(24/27)
IICAOAOSVM	43.8516	3.6517	94.24%(257/243)	92.59%(25/27)

Six samples of the diagnosis results using the five approaches are shown in Table 3. The training accuracies by the five transformer fault diagnosis methods and the diagnosis results in testing accuracies (i.e., the ratio of the total number of test data correctly classified to the total number of test data) are shown in Table 4. According to Table 3 and Table 4, classification accuracy of the RBFNN, the IEC criteria, and the SVM with random parameters is below 85.18%. The IICASVM and the PSOSVM algorithms with the OAO scheme have the better diagnosis accuracy, and the IICASVM hits 92.59% and has the best classification accuracy in the testing phase. It can be concluded that the proposed approach is an effective tool for fault diagnosis of transformers.

6. Conclusion

This paper proposes a new nonlinear-multiple classification SVM model based on improved imperialist competitive algorithm for fault diagnosis of oil-immersed transformers. Conclusions can be summarized as follows:

1) The IICASVM model and the PSOSVM model were tested for solving classification problems, including iris, wine, and glass classification problem by using three benchmark data sets of UIC. By using the two classification approaches with the four multiple classification schemes, the classification results show that PSOSVM and IICASVM are two effective approaches for solving classification problem and the OAO performs better performance than other schemes in the three benchmark classification problems.
2) Five classification approaches were studied for fault diagnosis of transformers. The classification accuracy of the five approaches, including IEC criteria, RBFNN, SVM, PSOOAOSVM, the nonlinear-multiple classification IICAOAOSVM, reaches 74.07%, 77.78%, 85.18%, 88.89%, 92.59%, respectively. The comparing results demonstrate that the proposed IICASVM approach has better performance than the other four approaches.

In order to offer more useful information and obtain higher accuracy for transformer fault analysis, it is necessary to find out and update (or delete) the bad data of the training samples by AI approaches. Thus, a subsequent work needs to be supplemented in the future study.

References

[1] X. Liu, R. J. Liao, Y. Lv, "Study on Influences and Elimination of Test Temperature on PDC Characteristic Spectroscopy of Oil-Paper Insulation System," *Journal of Electrical Engineering & Technology*, vol. 10, no. 3, pp. 1107-1113, 2015.
[2] C. H. Shon, S. H. Yi, H. J. Lee, "Study on the Transfer Functions for Detecting Windings Displacement of Power Transformers with Impulse Method," *Journal of Electrical Engineering & Technology*, vol. 7, no. 6, pp. 876-883, 2012.
[3] D. J. Kweon, K. S. Koo, J. W. Woo, J. S. Kwak, "A Study on the Hot Spot Temperature in 154kV Power Transformers," *Journal of Electrical Engineering & Technology*, vol. 7, no. 3, pp. 312-319, 2012.
[4] S. Khan, M. D. Equbal, T. Islam, "A Comprehensive Comparative Study of DGA Based Transformer Fault Diagnosis Using Fuzzy Logic and ANFIS Models," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 1, pp. 590-596, 2015.
[5] R. E. James, Q. Su, "Condition assessment of high voltage insulation in power system equipment," *The Institution of Engineering and Technology*, London, 2008.
[6] "IEEE Std C57.104," *IEEE guide for the interpretation of gases generated in oil-immersed transformers*, 2008.
[7] "IEC Std 60599," *Guide to the interpretation of dissolved and free gases analysis*, 1999.
[8] M. Duval, A. Depabla, "Interpretation of oil in gas analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electrical Insulation Magazine*, vol. 17, no. 2, pp. 31-41, 2001.
[9] "CIGRE working group 09 of study committee 12," *Life time evaluation of transformer*, no. 150, pp. 39-51, 1993.
[10] M. Schou, A. Amdisen, S. E. Jensen, T. Olsen, "A Review of Dissolved Gas Analysis Measurement and Interpretation Techniques," *IEEE Electrical Insulation Magazine*, vol. 30, no. 3, pp. 39-49, 2014.
[11] J. Aghaei, A. Gholami, H. A. Shayanfar, "Dissolved gas analysis of transformers using fuzzy logic approach," *European Transactions on Electrical Power*, vol. 20, no. 5, pp. 630-638, 2010.
[12] A. C. M. D. Silva, A. R. G. Castro, V. Miranda, "Transformer failure diagnosis by means of fuzzy rules extracted from Kohonen Self-Organizing Map,"

International Journal of Electrical Power & Energy Systems, vol. 43, no. 1, pp. 1034-1042, 2012.

[13] C. H. Lin, C. H. Wu, P. Z. Huang, "Grey clustering analysis for incipient fault diagnosis in oil-immersed transformers," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1371-1379, 2009.

[14] Z. Y. Wang, C. X. Guo, Q. Y. Jiang, Y. Cao, "A fault diagnosis method for transformer integrating rough set with fuzzy rules," *Transactions of the Institute of Measurement and Control*, vol. 28, no. 3, pp. 243-251, 2006.

[15] K. Meng, Z. Y. Dong, D. H. Wang, K. P. Wong, "A self-adaptive RBF neural network classifier for transformer fault analysis," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1350-1360, 2010.

[16] V. Miranda, C. Garcez, R. Adriana, S. Lima, "Diagnosing Faults in Power Transformers With Autoassociative Neural Networks and Mean Shift," *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1350-1357, 2012.

[17] C. Pan, W. Chen, Y. Yun, "Fault diagnostic method of power transformers based on hybrid genetic algorithm evolving wavelet neural network," *Iet Electric Power Applications*, vol. 2, no. 1, pp. 71-76, 2008.

[18] H. Y. Wu, C. Y. Hsu, T. F. Lee, "Improved SVM and ANN in incipient fault diagnosis of power transformers using clonal selection algorithms," *International Journal of Innovative Computing Information and Control*, vol. 5, no. 7, pp. 1959-1974, 2009.

[19] R. J. Liao, H. B. Zheng, S. Grzyborwski, "A multiclass SVM-based classifier for transformer fault diagnosis using a particle swarm optimizer with time-varying acceleration coefficients," *International Transactions on Electrical Energy Systems*, vol. 23, no. 2, pp. 181-190, 2013.

[20] S. W. Fei, X. B. 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.

[21] M. S. David, C. Dictino, M. C. Jesus, "Identification of a surface marine vessel using LS-SVM," *Journal of Applied Mathematics*, vol. 2013, pp. 1-11, 2013.

[22] F. Mei, J. Mei, J. Zheng, Y. Wang, "Development and Application of Distributed Multilayer On-line Monitoring System for High Voltage Vacuum Circuit Breaker," *Journal of Electrical Engineering & Technology*, vol. 8, no. 4, pp. 813-823, 2013.

[23] V. N. Vapnik, "The nature of statistical learning theory," *New York: Springer-Verlag*, 1998.

[24] Y. Bazi, F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3374-3385, 2006.

[25] E. A. Gargari, C. Lucas, "Imperialist Competitive Algorithm: An Algorithm for Optimization Inspired by Imperialistic Competition," *IEEE Congress on*

Evolutionary Computation, vol. 21, no. 1, pp. 4661-4667, 2007.

[26] W. Sun, Y. Liang, "Least-Squares Support Vector Machine Based on Improved Imperialist Competitive Algorithm in a Short-Term Load Forecasting Model," *Journal of Energy Engineering*, vol. 141, no. 4, pp. 1-8, 2014.

[27] H. M. Moghimi, B. Vahidi, "A solution to the unit commitment problem using imperialistic competition algorithm," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 117-124, 2012.

[28] C. W. Hsu, C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 1, pp. 415-425, 2002.

[29] F. Melgani, Y. Bazi, "Classification of electro-cardiogram signals with support vector machines and particle swarm optimization," *IEEE Transactions on Information Technology in Biomedicine*, vol. 5, pp. 667-677, 2002.

[30] C. C, Chang, C. J. Lin, "LIBSVM: A library for support vector machines," 2001.

[31] A. Frank, A. suncion. "UCI machine learning repository", *http://archive.ics.uci.edu/m*l, 2012.

Yiyi Zhang was born in Guangxi, China in 1986. He received his Bachelor and Ph.D. degrees in electrical engineering in 2008 and 2014, respectively from Guangxi University, Nanning, China and Chongqing University, Chongqing, China. In 2014, He joined Guangxi University where he is a lecturer of Guangxi key Laboratory of Power System Optimization and Energy Technology. He is now a member and young professionals of IEEE. His current research interests include intelligent diagnosis and maintenance decision making of power transformers. Dr. Yiyi Zhang is author and co-author of over 30 papers published in journals and conferences.

Hua Wei was born in Guangxi, PR China. He received the M.S. and PhD. degrees in power engineering from Guangxi University, Guangxi, PR China and Hiroshima University, Hiroshima, Japan, respectively. Now, he is a Professor of the department of Electrical Engineering, Guangxi University. His research interest is power system operation and planning, particularly in the application of optimization methods to power systems. Prof. Wei is a member of CSEE and IEEE.

Ruijin Liao was born in Suining, Sichuan, China, on July 28, 1963. He received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2003. He became Professor of Chongqing University in 1999. He has been Dean of College of Electrical Engineering at Chongqing University since 2002. His current research includes the online detection of insulation condition and insulation fault diagnosis for high-voltage (HV) apparatus and high-voltage testing technique. He is an author of more than 100 publications and inventions.

Youyuan Wang received the B.Sc. degree from Southwest Normal University in 1994, and obtained the M.Sc. and Ph.D. degrees both from Chongqing University in 2003 and 2008, respecttively. Now he is a professor in Chongqing University. His major research interests are online insulation condition monitoring and fault diagnosis of high voltage equipment.

Lijun Yang was born in Sichuan, China in 1980. She received her M.S. and Ph.D. degrees in electrical engineering from Chongqing University, China in 2004 and 2009. Her major research interests include online detection of insulation condition of electrical devices, partial discharges, and insulation fault diagnosis for high voltage equipment.

Chunyu Yan was born in 1960. He is a professor level senior engineer. His main research interests include condition based maintenance and fault diagnosis techniques.