Parameters Optimization of SVM Based on Improved FOA and Its Application in Fault Diagnosis

Qiantu Zhang^{1*}, Liqing Fang¹, Sicai Su², Yan Lv¹

¹ First Department, Mechanical Engineering College, Shijiazhuang, Hebei Province, P. R. China. ² Fourth Department, Mechanical Engineering College, Shijiazhuang, Hebei Province, P. R. China.

* Corresponding author. Tel.: +8618632119821; email: qiantuz@sina.com Manuscript submitted July 7, 2015; accepted September 8, 2015. doi: 10.17706/jsw.10.11.1301-1309

Abstract: In most cases, fault diagnosis is essentially a pattern recognition problem and support vector machine (SVM) provides a new solution for the diagnosis problem of systems in which the fault samples are few. However, the parameters selection in SVM has significant influence on the diagnosis performance. In this paper, improved fruit fly optimization algorithm (IFOA), which is basically the standard fruit fly optimization algorithm (FOA) combined with Levy flight search strategy, is proposed to determine the SVM parameters. Some benchmark datasets are used to evaluate the proposed algorithm. Furthermore, the proposed method is used to diagnose the faults of hydraulic pump. Experiments and engineering application show that the proposed method outperforms standard FOA, genetic algorithm (GA) and particle swarm optimization (PSO) methods.

Key words: SVM, fruit fly optimization algorithm, Levy flight, fault diagnosis.

1. Introduction

With the development of artificial intelligence, many methods have been applied to fault diagnosis, such as expert systems [1], fuzzy logic [2], neural network [3], and successful results have been obtain. However, in practical engineering application, due to the complex system with multi-factor, strong coupling and intensive nonlinearity, it is difficult to obtain much useful fault information which are necessary for the above methods.

Support vector machine as a new machine learning method was widely used in fault diagnosis due to its strong classification capability and good generalization even though the fault samples are few [4]-[6]. But there is an existing fact that the kernel function parameters and the penalty factor of SVM affect its classification performance seriously and the parameters are difficult to select due to the lack of corresponding theoretical basis. In order to find the appropriate parameters of SVM, many intelligent evolutionary algorithm, such as genetic algorithm (GA)[7], particle swarm optimization (PSO)[8], ant colony optimization algorithm (ACO)[9], and artificial bee colony optimization (ABC)[10], have been carried out on SVM parameters optimization, and improved the performance of SVM to some extent. However, due to the defects of algorithm itself, those methods are still time consuming and does not performed very well at many situations.

The fruit fly optimization algorithm (FOA), which was introduced by Pan[11], is a novel global optimization computational method which was inspired by the foraging behavior of fruit flies. The FOA has few parameters to adjust, and it is easy to understand and more suitable for computer programming

processing. In this paper, aiming at the problems that the FOA can easily fall into the local optimum in global optimization computations and the stability of the algorithm do not performs well, an improved FOA (IFOA) was proposed and was used to optimize SVM parameters. In IFOA, by dividing the fruit fly group into two subgroups, introducing Levy flight search strategy, and changing the location updating way of each fruit fly, not only the diversity of the fruit fly group guaranteed, but also the global and local search ability balanced, at the same time, the fruit fly group can also escape from local optimum thanks to the occasional long jumps of Levy flight. The experiment results on fault data of hydraulic pump indicate that the proposed can reduce the training time while increasing diagnosis accuracy, and has certain superiority when compared with the method of SVM with FOA, GA, PSO separately.

2. Principle of Support Vector Machine

The main purpose of SVM is to construct optimal separating hyper-plane and maps the training samples from the input space into a higher dimensional feature space via a kernel function. Suppose there is a training set $T = \{(x_1, y_1), \dots, (x_l, y_l)\}_l^l, x_i \in \mathbb{R}^n, y_i \in \{1, -1\}$, where x_i is the input vector and y_i is the label of the x_i , and l is the number of the input vectors and n is the number of input dimension. Structure hyper-plane is $\omega \cdot x + b = 0$, where ω is a weight vector and b is a scalar. The goal of maximizing the margin width is equivalent to the following optimization problem:

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \varepsilon_i \\ s.t. \quad y_i \left[(\omega \cdot x_i) + b \right] \ge 1 - \varepsilon_i \quad \varepsilon_i \ge 0, i = 1, 2, \cdots, l \end{cases}$$
(1)

where *C* is the penalty parameter and C > 0, ε_i are positive slack variables which are necessary to allow misclassification. According to the Lagrangian principle, the above optimization problem can be transformed into its dual form as follows:

$$\begin{cases} \min \ \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_{i} \alpha_{j} y_{i} y_{j} K(x \cdot x_{i}) - \sum_{i=1}^{l} \alpha_{i} \\ st. \ \sum_{i=1}^{l} \alpha_{i} y_{i} = 0, \ 0 \le \alpha_{i} \le C, \ i = 1, \cdots, l \end{cases}$$
(2)

where $K(x, x_i) = (\varphi(x_i), \varphi(y_i))$ is the kernel function, which maps the sample space into a high dimension feature space.

After solving the above problem, the corresponding decision function can be shown as:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_{i} y_{i} K(x, x_{i}) + b\right)$$
(3)

The typical kernel functions are polynomial kernel function, radial basis kernel function and sigmoid kernel function. In the above kernel functions, the radial basis kernel function (RBF) is widely used due to its universal adaptability and good performance. The form of RBF is as follow:

$$K(x, x_i) = \exp(-g ||x - x_i||^2)$$
(4)

where g denotes the width of RBF kernel function. From (2), (3) and (4), we can see that the parameters that need to be optimized in RBF-SVM is penalty parameter C and kernel parameter g. The penalty parameter C is used to regulate the proportion between the confidence range and the empirical risk in

determinate feature space to maximize the generalization ability of SVM. The kernel parameter *g* mainly affects the distribution complexity of samples in the high-dimension feature space. The performance of SVM is highly depends on the two parameters and it is of great importance to use optimization algorithm which can give useful guidance on its selection.

3. Parameters Optimization of SVM with IFOA

3.1. FOA Algorithm

The fruit fly optimization algorithm proposed by Pan in 2011 is a new global optimization algorithm inspired by the food search behaviors of fruit fly. The olfactory organ of the fruit flies can collect scents floating in the air, and then, after it gets close to the food location, it can also use its sensitive vision to find the food and fly to that direction. The main steps of FOA are as follow[11]:

Step 1. Randomly initialize fruit fly group location: Init *X*_*axis*, Init *Y*_*axis*

Step 2. Give the random direction and distance for food search:

$$\begin{cases} X_i = X _ axis + RandomValue \\ Y_i = Y _ axis + RandomValue \end{cases}$$
(5)

Step 3. As the food location cannot be known, the distance to the origin is thus estimated first (*Dist*), then the smell concentration judgment value (*S*) is calculated, and this value is the reciprocal of distance:

$$\begin{cases} Dist_i = \sqrt{X_i^2 + Y_i^2} \\ S_i = 1/Dist_i \end{cases}$$
(6)

Step 4. Substitute smell concentration judgment value (*S*) into smell concentration judgment function (or called Fitness function) so as to find the smell concentration (*Smell*_{*i*}) of the individual location of the fruit fly:

$$Smell_i = Function(S_i)$$
⁽⁷⁾

Step 5. Find out the fruit fly with maximal smell concentration (finding the maximal value or the minimum value) among the fruit fly group:

$$[bestSmell, bestIndex] = \max(Smell)$$
(8)

Step 6. Keep the best smell concentration value and coordinate (*X*,*Y*). The fruit fly group will use vision to fly towards that location:

$$\begin{cases}
Smellbest = bestSmell \\
X_axis = X (bestIndex) \\
Y_axis = Y (bestIndex)
\end{cases}$$
(9)

Step 7. Repeat the implementation of steps $2\sim5$, then judge if the smell concentration is superior to the previous iterative smell concentration, if so, implement step 6.

3.2. Design of IFOA

From the basic steps of the FOA, we can see that the fruit fly group only study the best individual in the whole iteration process, when the best individual are discovered, all the fruit flies will fly to that location. This kind of location updating way not only decreases diversity of the fruit fly group, but also makes algorithm easily trapped in local optimum and affects its convergence speed and precision when that

individual is not the best.

Sociology experience tells us that the global optimum always exists around the local optimum and the evolutionary rate of the group depends on its worse individuals to a greater extent instead of the better individuals. On the other hand, it is reported that many animals and insects, such as fruit fly and honey bee, will take a Levy flight search strategy in which exploratory short jumps alternate with occasional long jumps when they are searching food. The exploratory short jumps can ensure they have a finely local search in their landscape, and the occasionally long jumps can help them get into another area to make more extensive search. Considering the merits of Levy flight, such strategy has been applied to optimization and optimal search, and preliminary results have shown its promising capability[12][13].

According to the above analysis, an improved FOA was proposed. The IFOA is based on the basic FOA. In the whole evaluation process, the distance (D_{i_best} and D_{i_worst}) of each fruit fly individual *i* to the best individual and worst individual in contemporary group are calculated by using (10):

$$\begin{cases} Dist_{i_best} = \sqrt{(X_i - X_b)^2 - (Y_i - Y_b)^2} \\ Dist_{i_worst} = \sqrt{(X_i - X_w)^2 - (Y_i - Y_w)^2} \end{cases}$$
(10)

where (X_i, Y_i) are the location of fruit fly individual i, (X_b, Y_b) and (X_w, Y_w) are the location of the best individual and worst individual in the contemporary group. If $D_{i_best} > D_{i_worst}$, divide the individual i into drawback subgroup, else, divide the individual i into advanced subgroup (The fruit fly individual and individual number in two subgroups are exchanging in iteration process). As for drawback subgroup, a global search is made with the guidance of the best individual, fruit flies updating location according to (11). As for advanced subgroup, the Levy flight strategy was introduced. A local search is made around the best individual, fruit flies updating location according to (12).

$$\begin{cases} X'_{i} = X_{b} + RandomValue \\ Y'_{i} = Y_{b} + RandomValue \end{cases}$$
(11)

$$\begin{cases} X_i' = X_i + \alpha (X_i - X_b) \oplus L(\lambda) \\ Y_i' = Y_i + \alpha (Y_i - Y_b) \oplus L(\lambda) \end{cases}$$
(12)

where (X'_i, Y'_i) are the new location of each fruit flies, $L(\lambda)$ is the random search pathway of Levy flight and $\alpha > 0$ is the step size which should be related to the scale of the particular problem under study. The symbol \oplus means entrywise multiplications.

Fruit flies in each subgroup update its location by using (11) and (12) respectively, displaced the location updating way that all the fruit flies fly to the best location. The new location updating way not only prevent loss of diversity, but also balanced the global search ability and local search ability. Meanwhile, since the size and direction of $L(\lambda)$ are highly random, its occasional long jumps can make the fruit flies jump from one area to another, which can ensure the fruit flies avoid being attracted by the local optimum when making local search.

In (12), the $L(\lambda)$ is calculated by using Mantegna's algorithm[14]:

$$s = \mu / |v|^{1/\beta} \tag{13}$$

where *s* is the random search pathway $L(\lambda)$, the scope of β is $0 < \beta < 2$, μ and *v* are drawn from normal

distributions. That is

$$\mu \sim \mathcal{N}(0, \sigma_{\mu}^{2}), \nu \sim \mathcal{N}(0, \sigma_{\nu}^{2})$$
(14)

with

$$\sigma_{\mu} = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \sigma_{\nu} = 1$$
(15)

where Γ is standard Gamma function.

3.3. Parameters optimization of SVM with IFOA

SVM classification model constructed by radial basis kernel function only has two parameters need to be optimized, which are penalty parameter C and the kernel function parameters g. In this paper, the IFOA method is employed to determine the parameters of SVM. Here the fruit fly is composed of the parameter C and g, and the process of optimizing the SVM parameters with IFOA can been described as follows:

Step1. Randomly generate the initial location of the each fruit fly which determines the scope of SVM parameter vector array (C, g). Set the number of fruit flies, the maximum iteration number and step size α . Set iteration variable: t = 0 and perform the training process from Step2~Step6.

Step2. Set iteration variable: t = t + 1.

Step3. Evaluate the quality of every fruit fly by using fitness function which was calculated with five-fold cross validation method.

Step4. Updating the best fruit fly and global best fruit fly according to the fitness value. Then divided the group into two subgroups according to (10) and update the location of the fruit flies in each subgroups according to (11) and (12).

Step5. Go to Step6 if the stopping criteria is satisfied. Otherwise, go to Step2 to continue the operation. Step6. End the training procedure and now the parameter C and g gotten is the final model parameters.

3.4. Verification of IFOA-SVM Algorithm

To evaluate the performance of the proposed IFOA-SVM method, we have used three common benchmark data sets from UCI benchmark[14], the Glass, Segment and German datasets. Instruction of the three datasets can be found in Table 1.

Table 1. Instruction of the UCI Datasets						
Name	Attribution	Class	Training set	Testing set		
German	24	2	200	800		
Segment	18	7	700	1610		
Glass	9	6	107	107		

In the experiment, the proposed IFOA method and other three classical optimization algorithm, i.e., traditional FOA, GA and PSO are used to optimize SVM and classify these data sets. To make a fair comparison, the same parameter value was used, in addition to some new parameters. The population size for the four methods is 20, and maximum evaluation generation is set to 100. For the PSO, the parameters W = 0.75, $c_1 = c_2 = 1.5$. For the GA, the mutation probability $p_m = 0.01$ and the crossover probability $p_c = 0.7$. In IFOA, the step size $\alpha = 0.5$.

In Fig. 1, we have plotted the fitness curve of the three data sets with four methods. From the fitness curve that shown in Fig. 1, it can be clearly seen that FOA, GA and PSO may either fall into a local optimum

easily, or have a rather slow evolution speed, while IFOA is built with good global search ability and fast convergence. The test results of the three data sets are shown in Table 2. It can be well seen that the IFOA-SVM has higher recognition accuracy than that of the FOA-SVM, the GA-SVM and the PSO-SVM. It indicates that IFOA is superior to FOA, GA and PSO in SVM parameters optimization.



Fig. 1. Fitness curve for three datasets with the four methods.

Table 2. Testing Results of the Three Datasets (%)								
Datasets	IFOA-SVM	FOA-SVM	GA-SVM	PSO-SVM				
German	72.8800	72.6520	71.8750	72.5000				
Segment	82.6087	82.5466	82.1739	82.4442				
Glass	67.4418	64.8598	62.0155	64.3410				

4. Application in Fault Diagnosis of Hydraulic Pump

Hydraulic pump are important and frequently used in hydraulic system, which are found widely used in industrial application. Any defects occur to one of its operation may lend to serious damage to the whole hydraulic system. Therefore, it is essential to develop a reliable condition monitoring and fault diagnosis method to prevent it from malfunction.

The use of vibration signals is quite common in the field of condition monitoring and diagnostics of hydraulic pump damage detection. It is possible to obtain diagnosis information from the vibration signals by using some signal processing techniques. Then, feature extraction and selection are undertaken. At last, the pattern recognition methods are used to diagnose the faults. In this section, the proposed method is applied in fault diagnosis of hydraulic pump.



Fig. 2. Time domain waveform of four fault type.

4.1. Data Collection

In this paper, the test object is a swash-plate axial piston hydraulic pump in the type of SY-10MCY14-1EL, which includes seven pistons and the rated speed is 1500rpm. A piezoelectric accelerometer which attached to the end cap was used to collected vibration signal. Four different fault type: normal (N), one loose slipper (F1), two loose slipper (F2), slipper wear (F3) were obtained. The sampling frequency is 50kHz. 100 samples of each fault type are acquired and each sample contain 2048 data points. Therefore, 400 samples in total are collected for further analysis. Fig. 2 demonstrates the time domain waveforms of four fault type. From Fig. 2 it may be possible differ from different fault type, but it is unreliable to make decision only according to the time domain waveforms and further analysis is needed.

4.2. Feature Extraction

The autoregressive (AR) model [15] is an effective approach to extract the fault feature of vibration signal but it can only be applied to stationary signals, while the fault vibration signals of hydraulic pump are non-stationary. The local characteristic-scale decomposition (LCD) [16] method is a new signal analysis method which can decompose the non-stationary vibration signal into several intrinsic scale components (ISC) which are stationary. In this paper, the input feature of SVM is obtained by a novel hybrid method based on LCD and AR model. The four kinds of hydraulic pump fault signals are original decomposed into several ISCs, then the AR model of the first three ISCs of each vibration signal which contain main fault information are established and the order of the AR model is determined as eight by AIC criterion. The eight AR parameters and the variance of the AR model of each ISC are regard as the feature vectors. Therefore, dimension of the feature vector is 27 and the size of feature matrix is 400 × 27.

4.3. Fault Diagnosis

In this study, to verify the performance of IFOA-SVM in fault diagnosis, a fair comparison is made with FOA-SVM, GA-SVM, and PSO-SVM. All 400 data samples are divided into two data sets: the training set and the testing set, in which the training set including 240 samples are used to calculate the fitness and train the diagnosis model, and the testing set including 160 samples are used to examine the classification accuracy of each model.



Fig. 3. Fitness optimization search curve.

Fig. 4. Visual expression of the classification results.

In the optimized SVM, the parameters C and g are optimized by IFOA, FOA, GA and PSO respectively, and the adjusted parameters with maximal cross validation accuracy are selected as the most appropriate parameters. Then, the optimal parameters are utilized to train SVM model. The search span of parameter C and g is [0, 1000]. The population size for the four methods is 20, and maximum evaluation generation is set to 100. The remaining parameters needed in the four methods is the same as in Section 3.4.

1307

Method	Values of <i>C</i>	Values of g	Classification Accuracy/%	Training Time/s
IFOA-SVM	198.4376	0.0496	96.2500	42.3219
FOA-SVM	43.9734	0.6596	92.5000	40.3507
GA-SVM	187.6431	2.1419	90.6250	88.3441
PSO-SVM	210.4944	0.7708	93.1250	107.2874

Table 3. Diagnostic Results of Different Methods

The fitness curve of the four methods of the training set is shown in Fig. 3. From it we can see that the highest five-fold cross validation is 98.75%, which was achieved by IFOA-SVM with only 16 iterations. Faster convergence and higher classification accuracy can be obtained when compared with other three methods. Table 3 given the diagnostic results by different methods and Fig. 4 shows the visual expression. As can be seen from Table 3 and Fig. 4, the values of *C* and *g* obtained by each method are different. The classification accuracy of the proposed IFOA-SVM is 96.25%, which has a 3.125% improvement than PSO-SVM which has highest accuracy in the other three methods. Furthermore, although the computation time of the proposed method is a bit more than that of the FOA-SVM, but it is much less than that of the GA-SVM and the PSO-SVM method. The above analysis indicates that the proposed method are more suitable in fault diagnosis of hydraulic pump.

5. Conclusion

In this paper, a fault diagnosis method using SVM with improved FOA was proposed. The improved FOA is basically the standard FOA combined with Levy flight search strategy, and was used to optimize the SVM parameters. Then the IFOA-SVM was applied to fault diagnosis of hydraulic pump. The experiment results of hydraulic pump fault diagnosis indicated that the IFOA is feasible to optimize the SVM parameters and IFOA-SVM reduces the training time when compared with GA-SVM and PSO-SVM method while improving the fault diagnosis accuracy when compared with FOA-SVM, GA-SVM and PSO-SVM method.

Acknowledgment

This work is supported by the National Nature Science Foundation of China (No.51275524). The authors would like to thanks UCI benchmark for the datasets provided.

References

- [1] Liu, S. C., & Liu, S. Y. (2003). An efficient expert system for machine fault diagnosis. *International Journal of Advanced Manufacture Technology*, *21*, 691–698.
- [2] Tan, D. M., & Qu, W. L. (2011). Finite element model updating of structure based on lifting wavelet packet transform and fuzzy pattern recongnition. *Journal of Vibration and Shock, 30(8)*, 194-198.
- [3] Liu, X. J., & Liu, X. M. (2011). Modified PSO-based artificial neural network for power electronic devices fault diagnosis modeling. *Energy Procedia*, *13*, 748-752.
- [4] Chao, J. H. (2013). Fault diagnosis for electrical control system by support vector machine and chaotic particle swarm optimization algorithm. *Journal of Computational Information System*, 9(12), 4931-4938.
- [5] Rondi, Y., Dan, P., Feng, H., & Min, H (2014). Fault diagnosis of engine by support vector machine and improved particle swarm optimization algorithm. *Journal of Information & Computational Science*, 11(13), 4827-4835.
- [6] Jing Bai, Lihong Yang & Xueying Zhang (2013). Parameter optimization and application of support vector machine based on parallel artificial fish swarm algorithm. *Journal of Software*, *8*(*3*), 673-679.
- [7] Huang, C. L., & Wang, C. J. (2006). A GA-based feature selection and parameters optimization for support vector machine. *Expert Systems with Application*, *31*, 231-240.

- [8] Lin, S. W., Ying, K. C., Chen, S. C., & Lee, Z. J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machine. *Expert Systems with Application*, 35, 1817-1824.
- [9] Zhang, X. L., Chen, X. F., & He, Z. J. (2010). An ACO-based algorithm for parameter optimization of support vector machine. *Expert Systems with Application*, *37*, 6618-6628.
- [10] Sulaiman, M. H., Mustafa, M. W., Shareef, H., & Khalid, S. N. A. (2012). An application of artificial bee colony algorithm with least squares support vector machine for real and reactive power tracing in deregulated power system. *International Journal of Electrical Power & Energy Systems*, 37, 67-77.
- [11] Pan, W. T. (2012). A new fruit fly optimization algorithm: Taking the financial distress model as an example. *Knowledge-Based Systems*, *26*, 69-74.
- [12] Yahya, M., & Saka, M. P. (2014). Construction site layout planning suing multi-objective artificial bee colony algorithm with Levy flights. *Automation in Construction*, *38*, 14-29.
- [13] Dogan, E. (2014). Solving design optimization problems via hunting search algorithm with levy flights. *Structural Engineering and Mechanics*, *52*, 351-368.
- [14] Mantegna, R. N. (1992). Fast, accurate algorithm for numerical simulation of levy stable stochastic process. *Physical Review E*, *49*, 451-458.
- [15] Frank, A., & Asuncion, A. UCI machine learning repository. Retrieved, from http://archive.ics.uci. edu/ml/datasets.html
- [16] Cheng, J. S., Yang, Y., & Yang, Y. (2012). Local characteristic-scale decomposition method and its application to gear fault diagnosis. *Journal of Mechanical Engineering*, *48*(9), 64-71.
- [17] Cheng, J. S., Yu, D. J., & Yang, Y. (2006). A fault diagnosis approach for roller bearings based on EMD method and AR model. *Mechanical Systems and Signal Processing*, *20(2)*, 350-362.



Qiantu Zhang was born in Chongqing, P. R. China in 1991. He is a postgraduate student at the First Department, Mechanical Engineering College, P. R. China. He received his bachelor degree from Mechanical Engineering College in 2013. His research interests include intelligence fault diagnosis, signal process and support vector machine.



Liqing Fang was born in Shijiazhuang, Hebei province, P. R. China in 1969. He is a professor at the First Department, Mechanical Engineering College, PR China. He received his Ph.D from Beijing Institute of Technology in 2005. His currently research include intelligence fault diagnosis and test technology. He is the author or coauthor of more than 50 scientific papers.