Research on Parameters Optimization of SVM Based on Improved Fruit Fly Optimization Algorithm

Qiantu Zhang, Liqing Fang, Leilei Ma, and Yulong Zhao

Abstract—The performance of the support vector machine (SVM) is determined to a great extent by the parameter selection. In order to improve the learning and generalization ability of SVM, in this paper, an improved fruit fly optimization algorithm (IFOA) was proposed to optimize kernel parameter and penalty factor of SVM. In IFOA, the fruit fly group is dynamically divided into advanced subgroup and drawback subgroup according to its own evolutionary level. A global search is made for the drawback subgroup under the guidance of the best individual and a finely local search is made for the advanced subgroup in which the fruit flies do Levy flight around the best individual. Two subgroups exchange information by updating the overall optimum and recombining the subgroups. Getting rid of local optimum and improve search ability are ensured by making those changes in basic FOA. The performance of the IFOA and classification accuracy of optimized SVM based on IFOA are respectively examined through several typical benchmark functions and classical data sets from UCI benchmark. The experiment results show that the performance of the new algorithm is obviously more successful than FOA and it is also an effective SVM parameter optimization method which has better performance than some other methods.

Index Terms—Support vector machine, fruit fly optimization algorithm, parameter optimization, Levy flight.

I. INTRODUCTION

In pattern recognition field, support vector machine as a new machine learning method was widely used due to its strong classification capability and good generalization even though the samples are few. 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 methods have been carried out on SVM parameters optimization. The commonly used are grid search and cross-validation. But these methods have several drawbacks, for example, both the grid search and cross-validation require long and complicated calculations. In recent years, some intelligent evolutionary algorithm, such as genetic algorithm (GA) [1], particle swarm optimization (PSO) [2], ant colony optimization algorithm (ACO) [3], artificial fish swarm algorithm (AFSA) [4] and artificial bee colony optimization (ABC) [5] have also been used to

optimize the SVM parameters for their good global search abilities.

The Fruit Fly Optimization Algorithm (FOA), which was introduced by Pan [6], 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. Due to its merits, the FOA has been applied successfully to solve some academic and engineering optimization problems, including financial distress [6], general regression neural network optimization [7], and fuzzy-PID controller parameters optimization [8]. However, the basic FOA and some modified version can easily fall into the local optimum in global optimization computations, and the stability of the algorithm must be further promoted [9], [10].

To strengthen the search ability of FOA and overcome the problem of being trapped into the local optimum, an improved fruit fly optimization algorithm (IFOA) was proposed. 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 is guaranteed, and the global and local search ability is balanced. Furthermore, the fruit fly group can also escape from local optimum thanks to the occasional long jumps of Levy flight. Finally, IFOA was used to optimize the kernel function parameters and the penalty factor of SVM. The performance of the optimized SVM is tested by selecting several typical data sets from UCI benchmark, successful results have been obtained.

II. PRINCIPLE OF SUPPORT VECTOR MACHINE AND ITS PARAMETERS

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 mapping function φ . Suppose there is a training set $T = \{(x_1, y_1), \dots, (x_l, y_l)\}_i^l, x_i \in \mathbb{R}^n, y_i \in \{1, -1\}$, 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:

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

Manuscript received March 13, 2015; revised August 25, 2015.

Qiantu Zhang, Liqing Fang, and Yulong Zhao are with the First Department, Mechanical Engineering College, PR China (e-mail: qiantuz@sina.com).

Leilei Ma is with the Basic Courses Department, Mechanical Engineering College, PR China.

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:

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

$$(2)$$

To nonlinear classification problems, by introducing the kernel function $K(x, x_i) = (\varphi(x_i), \varphi(y_i))$ which maps the sample space into a high dimension feature space, then the formula (2) can be rewritten as follows:

$$\min \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x \cdot x_i) - \sum_{i=1}^{l} \alpha_i$$

$$\text{s.t.} \quad \sum_{i=1}^{l} \alpha_i y_i = 0, \ 0 \le \alpha_i \le C, \ i = 1, \cdots, l$$

$$(3)$$

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

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_{i} y_{i} K(x, x_{i}) + b\right)$$
(4)

where α_i is the optimal solution, *b* is the classification thresholds, and *x* is the sample which needs to be identified.

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:

$$K(x, x_i) = \exp\left(-g \left\|x - x_i\right\|^2\right)$$
(5)

where g denotes the width of RBF kernel function. From (3), (4) and (5), 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.

III. THE DESIGN OF IFOA

A. Basic Principle of FOA

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 [6]-[8]:

Step1. Randomly initialize fruit fly group location:

Init
$$X _ axis$$
; Init $Y _ axis$

Step2. Give the random direction and distance for food search:

$$X_i = X _ axis + RandomValue$$

 $Y_i = Y _ axis + RandomValue$

Step3. 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:

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

Step4. 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:

$$Smell_i = Function(S_i)$$

Step5. Find out the fruit fly with maximal smell concentration (finding the maximal value or the minimum value) among the fruit fly group:

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

Step6. Keep the best smell concentration value and coordinate (X, Y). The fruit fly group will use vision to fly towards that location:

Step7. 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.

B. The Proposed Algorithm 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 [11], [12].

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 (6):

$$\begin{cases} Dist_{i_best} = \sqrt{(X_i - X_b)^2 - (Y_i - Y_b)^2} \\ Dist_{i_worst} = \sqrt{(X_i - X_w)^2 - (Y_i - Y_w)^2} \end{cases}$$
(6)

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 (7). 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 (8).

$$\begin{cases} X'_{i} = X_{b} + RandomValue \\ Y'_{i} = Y_{b} + RandomValue \end{cases}$$
(7)
$$\begin{cases} X'_{i} = X_{i} + \alpha (X_{i} - X_{b}) \oplus L(\lambda) \\ Y'_{i} = Y_{i} + \alpha (Y_{i} - Y_{b}) \oplus L(\lambda) \end{cases}$$
(8)

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 subgroups update its location by using (7) and (8) 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 can avoid being attracted by the local optimum when making local search.

In Ref. (8), the $L(\lambda)$ is calculated by using Mantegna's algorithm [13]:

$$s = \mu / \left| v \right|^{1/\beta} \tag{9}$$

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

$$\mu \sim \left(0, \sigma_{\mu}^{2}\right), \mathbf{v} \sim \left(0, \sigma_{\nu}^{2}\right)$$
(10)

with

$$\sigma_{\mu} = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \sigma_{\nu} = 1$$
(11)

where Γ is standard Gamma function.

C. IFOA Algorithm Testing and Results Analysis

In order to test the performance of IFOA, two benchmark functions are selected to test it.

Test function 1: Rosenbrock (n=30)

$$f_{\text{Ro}}(x) = \sum_{i=1}^{n-1} \left(100 \left(x_{i+1} - x_i^2 \right)^2 + \left(x_i - 1 \right)^2 \right), \quad -100 < x_i < 100$$

The global minimum value of this function is 0 within the scope of its definition domain and the optimal point is $(1,1,\dots,1)$.

Test function 2: Ackley (n=30)

$$f_{Ac}(x) = -20 \exp\left(-0.2\sqrt{\frac{1}{30}\sum_{i=1}^{n}x_{i}^{2}}\right) + 20 + e$$
$$-\exp\left(\frac{1}{30}\sum_{i=1}^{n}\cos(2\pi x_{i})\right), -100 < x_{i} < 100$$

The global minimum value of this function is 0 within the scope of its definition domain and the optimal point is $(0,0,\dots,0)$.

In the experiments, IFOA was compared with FOA. Table I shows the results of 30 independently runs with 2000 iterations of the two functions about the maximum value (Max), minimum value (Min), mean value (Mean) and standard deviation (Std). Convergence process contrast curve of the two algorithm are given in Fig. 1 and 2.

TABLE I: COMPARISON OF OPTIMIZATION RESULTS AMONG
IFOA AND FOA

Function	Index	IFOA	FOA
	Max	27.9717	28.8394
f_	Min	26.4716	28.3812
JRo	Mean	27.1292	28.7450
	Std	0.1128	0.2692
	Max	1.8200E-05	0.0965
C	Min	4.4700E-09	0.0002
JAc	Mean	9.7100E-09	0.0414
	Std	3.6958E-09	0.0385

From Fig. 1 and 2, the convergence curve shows clearly that FOA has a problems of being trapped in local optimum and has early convergence while IFOA escapes local optimum and continues to improve solutions. From Table I we can see that results of IFOA are closer to the theoretical optimal value compared with the FOA and the standard deviation shows that IFOA is also more robust than FOA. The experiment results above indicate that IFOA is superior to FOA in terms of search quality. IFOA may find the global optimum quickly with the least iterations and better convergence results.



IV. 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. As IFOA can not only make the search avoid trapped in local optimum, it can also help to search the optimum quickly. Therefore, the method is employed to determine the parameters of SVM. Here the fruit fly is composed of the parameter C and g. Fig. 3 presents the process of optimizing the SVM parameters with IFOA, which is described as follows:



Fig. 3. The process of optimizing the SVM parameters with IFOA.

Step 1: Divided the data into two parts with one part as the training sample and the other part as the testing sample

respectively.

Step 2. 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 *a*. Set iteration variable: t = 0 and perform the training process from step 3 to 7.

Step 3: Set iteration variable: t = t + 1.

Step 4: Evaluate the quality of every fruit fly by using fitness function which must be designed before searching for the optimal values of the SVM parameters. The fitness function is based on the classification accuracy of a SVM classifier, which is as follows:

$$fitness = \frac{y_t}{y_t + y_f} \tag{12}$$

where y_t and y_f denote the number of true and false classifications, respectively.

Step 5: Updating the best fruit fly and global best fruit fly according to the fitness value. Then divided the group into two subgroups according to (6) and update the location of the fruit flies in each subgroups according to (7) and (8).

Step 6: Go to step7 if the stopping criteria is satisfied. Otherwise, go to step3 to continue the operation.

Step 7: End the training procedure and now the parameter C and g gotten is the final model parameters.

V. EXPERIMENT RESULTS AND ANALYSIS

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 data sets. Instruction of the three data sets can be found in Table II. German is a binary class problem while the other two data sets are multi-class problem and the samples of each class in Glass are different while it is the same in Segment. Therefore, the three data sets can evaluate the performance from the different aspects. All samples of each data sets are divided into two sets: the training set and the testing set, in which the training set is used to calculate the fitness function and train SVM. The testing set is used to test the accuracy of the classification algorithm.

TABLE II: INSTRUCTION OF UCI DATA SETS

Name	Train	Test	Input	Class
Glass	107	107	9	6
Segment	700	1610	18	7
German	200	800	24	2

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 are fixed with the values given in literature [15], that is 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 a = 0.5.

82

In Fig. 4, we have plotted the search performance of all the four methods. From the fitness performance that is shown in Fig. 4, 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. It indicates that IFOA is superior to FOA, GA and PSO in SVM parameters optimization.



Fig. 4. Fitness performance for the proposed IFOA, FOA, GA and PSO algorithm of the three data sets.

In the optimized SVM, the parameters are optimized by adopting the four methods respectively. The adjusted parameters with the maximal classification accuracy of training set are selected as the most appropriate parameters. Then, the optimal parameters are utilized to train SVM model. At last, all the testing set of the three data sets is adopted to verify the superior classification performance of the proposed method by comparing with FOA-SVM, GA-SVM and PSO-SVM. The comparison results of classification results among them are shown in Table III, IV and V. From the three tables, we can see that the values of *C* and *g* obtained by each method are different. The classification accuracy of IFOA-SVM is better than that of the FOA-SVM, the GA-SVM and the PSO-SVM. Especially for the unbalanced multi-class data set Glass, the classification accuracy has made a considerable improvement, and it has great significance from the point of the view of the engineering application. 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 methods. The above analysis indicates that the proposed method are more suitable in classification problems.

TABLE III: EXPERIMENT RESULTS OF GLASS						
Method	Accuracy/%	Cost time/s	С	g		
IFOA-SVM	67.4418	10.5046	9.6151	10.7151		
FOA-SVM	64.8598	10.4452	80.1397	8.0140		
GA-SVM	62.0155	24.7808	75.1109	14.6309		
PSO-SVM	64.3410	15.0544	10.0138	10.1400		
TABLE IV: EXPERIMENT RESULTS OF SEGMENT						
Method	Accuracy/%	Cost time/s	С	g		
IFOA-SVM	82.6087	141.9624	21.0546	0.0875		
FOA-SVM	82.5466	138.6546	3.8242	0.3824		
GA-SVM	82.1739	373.9882	4.5046	0.1737		
PSO-SVM	82.4442	285.1161	8.9041	0.1534		
TABLE V: EXPERIMENT RESULTS OF GERMAN						
Method	Accuracy/%	Cost time/s	С	g		
IFOA-SVM	72.8250	33.0549	3.0731	0.1128		
FOA-SVM	72.6250	30.9033	3.3764	0.1013		
GA-SVM	71.8750	79.3651	1.7015	0.0547		
PSO-SVM	72.5000	56.9700	3.1763	0.1020		

VI. CONCLUSION

The choice of the parameters directly affects the performance of SVM. The best classification and generalization ability of SVM may be obtained when the appropriate parameters are selected. In this paper, an improved FOA algorithm which can avoid being trapped into local optimum, speed up the convergence rate and reduce the total iterations is used to optimize the parameters of SVM. The proposed IFOA-SVM method is applied to classify some benchmark data sets from UCI benchmark. Compared with some other methods, the results show that the proposed method is both effective and computational efficient in searching for the optimal parameters for SVM. Of course, an in-depth study on SVM is still needed in the future.

ACKNOWLEDGMENT

The authors would like to thanks UCI benchmark for the data sets provided.

REFERENCES

- [1] C. L. Huang and C. J. Wang, "A GA-based feature selection and parameters optimization for support vector machine," *Expert Systems with Application*, vol. 31, pp. 231-240, August 2006.
- [2] S. W. Lin, K. C. Ying, S. C. Chen, and Z. J. Lee, "Particle swarm optimization for parameter determination and feature selection of

support vector machine," *Expert Systems with Application*, vol. 35, pp. 1817-1824, November 2008.

- [3] X. L. Zhang, X. F. Chen, and Z. J. He, "An ACO-based algorithm for parameter optimization of support vector machine," *Expert Systems* with Application, vol. 37, pp. 6618-6628, September 2010.
- [4] M. H. Sulaiman, M. W. Mustafa, H. Shareef, and S. N. Abd. Khalid, "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*, vol. 37, pp. 67-77, May 2012.
- [5] J. Bai, L. H. Yang, and X. Y. Zhang, "Parameter optimization and application of support vector machine based on parallel artificial fish swarm algorithm," *Journal of Software*, vol. 8, pp. 673-679, March 2013.
- [6] W. T. Pan, "A new fruit fly optimization algorithm: taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, pp. 69-74, February 2012.
- [7] H. Z. Li, S. Guo, C. J. Li, and J. Q. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network woth fruit fly optimization algorithm," *Knowledge-Based Systems*, vol. 37, pp. 378-387, January 2013.
- [8] W. Sheng and Y. Bao, "Fruit fly optimization algorithm based fractional order fuzzy-PID controller for electronic throttle," *Nonlinear Dynamics*, vol. 73, pp. 611-619, February 2013.
- [9] L. Wang, X. L. Zheng, and S. Y. Wang, "A novel binary fruit fly optimization algorithm for solving the multidimensional knapsack problem," *Knowledge-Based Systems*, vol. 48, pp. 17-23, August 2013.
- [10] X. W. Wu and Q. Li, "Research of optimizing performance of fruit fly optimization algorithm and five kinds of intelligent algorithm," *Fire Control & Command Control*, vol. 38, pp. 17-20, April 2013.
- [11] M. Yahya and M. P. Saka, "Construction site layout planning suing multi-objective artificial bee colony algorithm with Levy flights," *Automation in Construction*, vol. 38, pp. 14-29, March 2014.

- [12] E. Dogan, "Solving design optimization problems via hunting search algorithm with Levy flights," *Structural Engineering and Mechanics*, vol. 52, pp. 351-368, 2014.
- [13] R. N. Mantegna, "Fast, accurate algorithm for numerical simulation of Levy stable stochastic process," *Physical Review E*, vol. 49, pp. 451-458, 1992.
- [14] A. Frank and A. Asuncion. UCI machine learning repository. [Online]. Available: http://archive.ics.uci.edu/ml/datasets.html
- [15] F. Ardjani, K. Sadouni, and M. Benyettou, "Optimization of SVM multiclass by particle swarm (PSO-SVM)," in *Proc. the 2nd International Workshop on Database Technology and Application* (DBTA'10), November 2010.



Qiantu Zhang was born in 1991. He is a postgraduate at the First Department, Mechanical Engineering College, PR 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 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.