

## A Novel Data Classification Method and its Application in IRIS Flower Shape

Chong Wu<sup>1</sup>, Chonglu Zhong<sup>2</sup> and Yanlei Yin<sup>\*1</sup>

<sup>1</sup>*Shan Dong Institute of Pomology, Taian 271000 China*

<sup>2</sup>*Research institute of tropical forestry Chinese academy of forestry, Guangzhou 510520 China*

### Abstract

*IRIS flower data is a class of multi variable data set, which is widely applied in data classification. This paper aims at the parameter optimization problem of least squares support vector machine (LS-SVM) in data classification, an improved particle swarm optimization (IMPSO) algorithm is introduced into the LS-SVM model for improving the learning performance and generalization ability of LS-SVM model. A new data classification method based on IMPSO algorithm and LS-SVM (IMPSO-LS-SVM) model is proposed. First, the numbers of current iteration and population are added into the control strategy of adaptive adjustment inertia weight in order to improve the performance of inertia weight of PSO algorithm. Then the IMPSO algorithm is used to search the optimal combination values of the parameters of kernel function for obtaining the IMPSO-LS-SVM. Finally, the training samples are used to comprehensively train the IMPSO-LS-SVM, and the best large-scale data classification model is constructed. The IRIS flower data is used to validate the effectiveness of the IMPSO-LS-SVM model. The result indicates that the IMPSO algorithm can effectively search the optimal combination values of the parameters, and the proposed data classification model has better generalization performance, faster training speed and higher classification precision.*

**Keywords:** *particle swarm optimization, least squares support vector machine, classification method, parameter optimization, IRIS flower data*

### 1. Introduction

Data classification is one of the most important branches of data mining. It is to select the classified training set from the data in order to build a classification model based on data mining technology in the training set, which is used to carry on classifying for the unclassified data [1-2]. In data analysis, data classification methods have become more and more important. With the continuous search of the data analysis, more and more data classification algorithms are proposed. They are widely applied in a variety of different fields [3]. The classification technology is mainly divided into two steps: training and classification. In the training process, the algorithm selects the feature from training set with class label, and then trained classifier. In the classification process, the classifier selects the feature from the sample without class label, and then classifies and determines the category. So data classification is a crucial step to construct a data classifier. A classifier is constructed in order to carry on the accurate classification for some unknown type of data.

At present, the main data classification methods are divided into the supervised learning and semi supervised learning. However, most mainstream classification methods are decision tree algorithm, BP neural network algorithm, K nearest neighbor algorithm, support vector machine (SVM) algorithm, and so on [4-7]. The SVM algorithm is widely used and also the research hotspot in recent years. It is a new machine learning method based on the statistical learning theory, VC dimension

theory and structural risk minimization principle. The basic idea of SVM method is: the optimal hyperplane is defined, the searching optimal linear hyperplane algorithm is boiled to solve an optimization convex problem. On the basis of Mercer kernel expansion theorem, the sample space is mapped to a high dimensional or infinite dimensional feature space by using the nonlinear mapping in order to apply linear machine learning methods in the feature space for solving nonlinear classification and regression problems in the sample space.

The LS-SVM takes on excellent learning ability and generalization ability [8]. However, it exists some weaknesses and limitations. At present, the mainly existed problem is how to select kernel function and optimize the parameters. For some data sets, the different kernel functions have similar influence on the classification problem. But for the other data sets, the different kernel functions have different influence on the classification results. There is no uniform solution to the different data sets for selecting the kernel function. It can only rely on the experience and validation. In addition, there has no reference standard for selecting the related parameters of the kernel function. In allusion to the selecting parameters of LS-SVM, many researchers have deeply studied and explored from the different views. They proposed some methods for selecting parameters of LS-SVM. Guo, *et al.*, [9] proposed three step search method for selecting the parameters of least squares support vector machines. Theoretical analysis indicates that this method is effective and superior for selecting the optimized parameters of least squares support vector machines. Chen, *et al.*, [10] used evolutionary algorithm to optimize several parameters of LS-SVM. This method is used in the soft sensor modeling for temperature measurement of Texaco gasifier. Luo, *et al.*, [11] proposed quantum-inspired evolutionary optimization to adjust the parameters of least squares support vector machines (LS-SVM). Classification and function estimation are studied using LS-SVM with wavelet kernel and Gaussian kernel. Wang, *et al.*, [12] proposed a new fuzzy least squares-support vector machines classifier based on chaos genetic algorithm (FLS-SVMBCGA). The method constructs a fuzzy membership set and the chaos genetic algorithm to optimize the parameters. Zhu, *et al.*, [13] proposed a hybrid QPSO algorithm for LS-SVM parameter selection to improve the learning performance and generalization ability of LS-SVM model. The Powell algorithm was used to obtain the initial position, local optimal position and global optimal solution. The proposed hybrid QPSO method combines with the global search ability of QPSO algorithm with the local search ability of Powell algorithm, to improve the solving speed and the accuracy of the solution. Wei, *et al.*, [14] proposed an adaptive penalty learning procedure called evolution strategies (ES) based adaptive Lp least squares support vector machine (ES-based Lp LS-SVM) to address the above issue. Glauber, *et al.*, [15] proposed to combine of LS-SVM and a new chaotic differential evolution optimization approach based on Ikeda map (CDEK) to the identification of a thermal process. Yang and Zhao [16] proposed a multi-class classification model based on improved least squares support vector machine (LSSVM). In the model, the non-sensitive loss function is replaced by quadratic loss function, and the inequality constraints are replaced by equality constraints. Zhang, *et al.*, [17] proposed a novel classification system to distinguish among elderly subjects with Alzheimer's disease (AD), mild cognitive impairment (MCI), and normal controls (NC). The kernel support vector machine decision tree (kSVM-DT) is constructed, and the error penalty parameter C and kernel parameter  $\sigma$  were determined by Particle Swarm Optimization (PSO). Wang, *et al.*, [18] proposed a preclassification and later regression based method optimized by improved particle swarm optimization (IPSO) for sales growth rate forecasting. Support vector machine (SVM) is used as a classification model. The nonlinear relationship in sales growth rate forecasting is efficiently represented by SVM, while IPSO is optimizing the training parameters of SVM. IPSO addresses issues of traditional PSO, such as relapsing into local optimum, slow convergence speed, and low convergence precision in the later evolution. Xue, *et al.*,

[19] proposed a novel hyperspectral image classification approach named as HA-PSO-SVM based on integrating the harmonic analysis (HA), particle swarm optimization (PSO), and support vector machine (SVM). In the combined method, HA is first proposed to transform the pixels from spectral domain into frequency domain expressed by amplitude, phase and residual, yielding more functional and discriminative features for classification purpose. Then, PSO is adapted to optimize the penalty parameter  $C$  and the kernel parameter  $\gamma$  for SVM, which leads to improved classification performance. Finally, the extracted features are classified with the optimized model. Nieto, *et al.*, [20] proposed a cyanotoxin diagnostic model by using support vector machines (SVMs) in combination with the particle swarm optimization (PSO) technique from cyanobacterial concentrations determined experimentally in the Trasona reservoir (recreational reservoir used as a high performance training center of canoeing in the Northern Spain). The agreement of the PSO-SVM-based model with experimental data confirmed its good performance. Finally, conclusions of this innovative research work are exposed.

In allusion to the existing shortcomings of the LS-SVM and improved LS-SVM for classification, a new large-scale data classification method based on improved PSO algorithm and LS-SVM(IMPSO-LS-SVM) model is proposed to improve the learning performance and generalization ability of the LS-SVM model. The effectiveness of the proposed IMPSO-LS-SVM model is verified by using data set.

## 2. PSO Algorithm and LS-SVM Model

### 2.1. PSO Algorithm

The PSO algorithm is a population-based search algorithm based on the simulation of the social behavior of birds within a flock. In PSO algorithm, individuals, referred to as particles, are “flown” through hyper dimensional search space. The particles’ positions within the search space are changed based on the social-psychological tendency of individuals in order to delete the success of other individuals. The changing of one particle within the swarm is influenced by the experience, or knowledge. The consequence of modeling for this social behavior is that the search is processed in order to return toward previously successful regions in the search space. Namely, the velocity( $v$ ) and position( $x$ ) of each particle will be changed by the particle best value ( $pB$ ) and global best value ( $gB$ ). The velocity and position updating of the particle is shown by the followed expression:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1r_1(pB_{ij}(t) - x_{ij}(t)) + c_2r_2(gB_{ij}(t) - x_{ij}(t)) \quad (1)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (2)$$

Where  $v_{ij}(t+1)$ , velocities of particle  $i$  at iterations  $j$ ,  $x_{ij}(t+1)$ , positions of particle  $i^{th}$  at iterations  $j^{th}$ .  $w$  is inertia weight to be employed to control the impact of the previous history of velocities. Accordingly, the parameter  $w$  regulates the trade-off between the global and local exploration abilities of the swarm. A large inertia weight facilitates the global exploration, while a small one tends to facilitate the local exploration. A suitable value of the inertia weight  $w$  usually provides the balance between the global and local exploration abilities and consequently results in a reduction of the number of iterations required to locate the optimum solution.  $t$  denotes the iteration number,  $c_1$  is the cognition learning factor,  $c_2$  is the social learning factor,  $r_1$  and  $r_2$  are random numbers uniformly distributed in  $[0, 1]$  which denote remembrance ability for study. Thus, the particle flies through potential solutions towards  $pB$  and  $gB$  in a navigated way while still exploring new areas by the stochastic mechanism to escape from

local optima. Generally, the value of each component in  $v$  can be clamped to the range  $[-V_{\max}, V_{\max}]$  to control excessive roaming of particles outside the search space. The PSO is terminated with a maximal number of generations or the best particle position of the entire swarm cannot be improved further after a sufficiently large number of generations. The PSO has shown its robustness and efficacy in solving function value optimization problems in real number spaces.

## 2.2. LS-SVM Model

Support vector machine (SVM) [1] introduced by Vapnik, is one of the popular tools for a supervised machine learning method based on structural risk minimization. The basic characteristic of the SVM model is to map the original nonlinear data into a higher-dimensional feature space where a hyperplane is constructed to bisect two classes of data and maximize the margin of separation between itself and those points lying nearest to the support vectors. The LS-SVM model is to use the least square linear system as the loss function, and the inequality constraints are revised as the equality constraints in the LS-SVM model.

The given the training sample is  $S = \{(x_i, y_i) | i = 1, 2, 3, \dots, m\}$ ,  $m$  is the number of samples, the set  $\{x_i\} \in R_n$  represents the input vector,  $y \in \{-1, 1\}$  indicates the corresponding desired output vector, the input data is mapped into the high dimensional feature space by using nonlinear mapping function  $\phi(\bullet)$ . Then the existed optimal classification hyperplane must meet the following conditions:

$$\begin{cases} \omega^T x_i + b \geq 1, & y_i = 1 \\ \omega^T x_i + b \leq -1, & y_i = -1 \end{cases} \quad (3)$$

where  $\omega$  is Omega vector of superplane,  $b$  is offset quantity. Then the classification decision function is described as follow:

$$f(x_i) = \text{sgn}(\omega^T x_i + b) \quad (4)$$

The classification model of the LS-SVM is described by the optimization function  $\min_{\omega, \xi, b} J(\omega, \xi_i)$ :

$$\min_{\omega, \xi, b} J(\omega, \xi_i) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^m \xi_i^2 \quad (5)$$

$$s.t. \quad y_i [\omega^T \phi(x_i) + b] = 1 - \xi_i, i = 1, 2, 3, \dots, m \quad (6)$$

where  $\xi_i$  is slack variable,  $b$  is offset,  $\omega$  is support vector,  $\xi = (\xi_1, \xi_2, \dots, \xi_m)$ ,  $\gamma$  is classification parameter to balance the fitness error and model complexity.

The optimization problem transforms into its dual space. Lagrange function is introduced to solve it. The corresponding optimization problem of the LS-SVM model with Lagrange function is shown:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^m \xi_i^2 - \sum_{k=1}^m \alpha_k \{y_k [\omega^T \phi(x_k) + b] - 1 + \xi_k\} \quad (7)$$

where  $\alpha_i$  is the Lagrange multiplier, and  $\alpha_i \geq 0 (i = 1, 2, 3, \dots, m)$ . The optimal conditions are described as follow:

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \Rightarrow \omega = \sum_{i=1}^m \alpha_i y_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^m \alpha_i y_i = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \Rightarrow \alpha_i = \gamma \xi_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow y_i (\omega^T \phi(x_i) + b) - 1 + \xi_i = 0_i \end{cases} \quad (8)$$

The following linear equation is obtained:

$$\begin{bmatrix} 0 & L^T \\ L & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (9)$$

where  $Y = [y_1, y_2, \dots, y_m]^T \in R^m$ ,  $L \in R^m$  is vector of the element  $m$ , then  $Y^T = [y_1, y_2, \dots, y_m]$ ,  $I$  is unit matrix,  $I_m = [1, 1, \dots, 1]^T$ ,  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_m]^T$ ,  $\Omega = [\Omega_{ij}]_{m \times m}$ ,  $\Omega_{ij} = y_i y_j K(x_i, x_j)$ . Then the classification decision function is described as follows:

$$f(x_i) = \text{sgn} \left( \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b \right) \quad (10)$$

### 3. IMPSO-LS-SVM Classification Model

#### 3.1. Improve Inertia Weight

The inertia weight( $w$ ) is one of the most important parameters in the PSO algorithm. The appropriate weight can effectively balance the convergence speed and accuracy of this algorithm. In order to improve the global and local search capability of PSO algorithm, on the basis of the control strategy thought of adaptive adjustment  $w$ , the numbers of current iteration and population are introduced to further improve the inertia weight( $w$ ) in order to obtain improved PSO(IMPSO) algorithm when the PSO algorithm updates the iteration. In the early stage, the larger inertia weight( $w$ ) is used, the PSO algorithm has strong global search ability. In the later stage, the smaller inertia weight( $w$ ) is used, the PSO algorithm enhances the local search ability. In many experiments, when the inertia weight( $w$ ) changes according to the following expression, the running time, average convergence iteration and segmentation accuracy of the PSO algorithm are improved.

$$w = w_{\min} + \frac{(w_{\max} - w_{\min}) \times (G_{\max} \times P_{size} - q \times p)}{G_{\max} \times P_{size}} \quad (11)$$

where  $w_{\min}$  is the minimum value of the inertia weight,  $w_{\max}$  is the maximum value of the inertia weight.  $G_{\max}$  is the maximum number of iterations,  $P_{size}$  is the number of populations.  $q$  is the number of the current iteration, and  $p$  is the number of the current updated population.

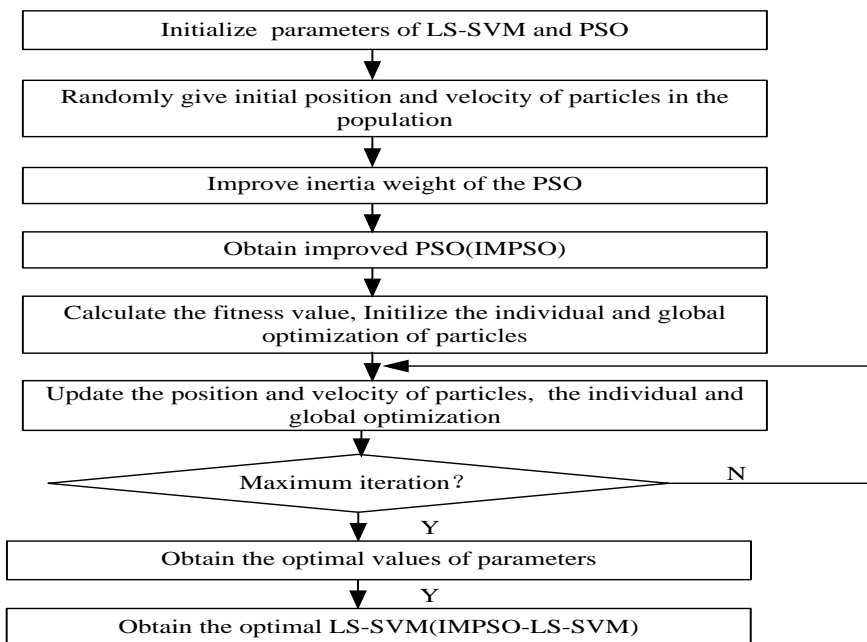
#### 3.2. LS-SVM Classification Model based on Improved PSO (IMPSO-LS-SVM) Algorithm

Because the different kernel function is selected to form LS-SVM model with different performance, it is extremely important to select the kernel function to form a good performance of the LS-SVM model. At present, there are linear kernel function,

polynomial kernel function, radial basis kernel function (RBF), Sigmoid kernel function, Fourier kernel function, etc. But there doesn't exist one kernel function, which is suitable for all applications. The LS-SVM is used to classify in this paper. Because the RBF takes on good learning ability, simple form, symmetry radial, good smoothness and analyticity, the RBF is widely applied to classify the data. So the RBF is an ideal kernel function to form the LS-SVM model for solving the data classification. The expression of RBF is describing by follows:

$$K(x, x_i) = \exp[ -(x - x_i)^2 / 2\sigma^2 ] \quad (12)$$

The kernel function width  $\sigma$  and penalty factor  $C$  is two important parameters in the LS-SVM model. They have considerable influence on the classification performance of LS-SVM. The penalty factor  $C$  is to control the punishment degree of complexity and error classifying samples of LS-SVM model. If the value of penalty factor  $C$  is too large, the classification accuracy is higher in the training phase and very low in testing phase. If the value of penalty factor  $C$  is too small, it is difficult to obtain satisfactory classification accuracy. The kernel function width  $\sigma$  is to reflect the distributional complexity of data sample in the high dimensional feature space, and determine the complexity of linear classification. If the input sample space is larger, the value of kernel function width  $\sigma$  is larger. If the value of kernel function width  $\sigma$  is smaller, the relationship among support vectors is more relaxed, learning machine is relatively complex, and the generalization ability of LS-SVM cannot be guaranteed. If the value of kernel function width  $\sigma$  is larger, the influence among support vectors is too strong, classification model is difficult to achieve sufficient accuracy. The influence of classification accuracy of kernel function width  $\sigma$  is larger than the penalty factor  $C$ . So it is particularly important for selecting suitable values of the penalty factor  $C$  and kernel function width  $\sigma$  in the training LS-SVM. In this paper, the IMPSO-LS-SVM classification model is proposed. The proposed model uses the IMPSO algorithm to automatically select the appropriate values of parameters of LS-SVM model, in order to obtain the optimal performance of LS-SVM classifier. The optimization process of LS-SVM classifier based on the IMPSO algorithm is shown Figure 2.



**Figure 2. The Optimization Process of LS-SVM Classifier based on the IMPSO Algorithm**

#### 4. Experimental Results and Analysis

In order to verify the effectiveness of the proposed IMPSO-LS-SVM classification method, IRIS flower data set is selected in this paper. The experiment works on Intel(R) Core i5-4200U, 2.30GHz, 4G RAM, Windows 8 and Matlab 2010. IRIS is a classic pattern recognition problem. There have three kinds of flower types. They are Setosa, Versicolor and Virginica. The type of each flower is constructed the sample of four dimensions according to the length and width of the petal and calyx. In this paper, three kinds of flower types generate 150 samples, the front 35 samples are regarded as the training sample set and the final 35 samples are regarded as the testing sample set. The different initial samples are selected to be regarded as the training samples for train LS-SVM. In this experiment, the particle sizes are 30, the maximum number of iterations are 500, the learning factor  $c_1 = c_2 = 1.6$ , the initial inertia weight  $w = 0.2$ , the penalty parameter  $C = 100$ , the radial basis kernel width is  $\sigma = [0,10]$ , the regularization parameter is  $\gamma = [1,1000]$ . The RBF kernel function is selected as the kernel function.

The proposed IMPSO-LS-SVM classification method is used to classify IRIS data set. Experiments are performed 20 times, and the average values are regarded the final classification results. The classification results are shown in Table 1.

**Table 1. The Classification Results**

Method	Classification	Samples	Correct number	Accuracy(%)	Average accuracy(%)
PSO-LS-SVM	Setosa	35	35	100	93.3
	Versicolor	35	32	91.4	
	Virginica	35	31	88.6	
APSO-LS-SVM	Setosa	35	35	100	97.1
	Versicolor	35	34	97.1	
	Virginica	35	33	94.3	
IMPSON-LS-SVM	Setosa	35	35	100	98.1
	Versicolor	35	34	97.1	
	Virginica	35	34	97.1	

As can be seen from the Table 2, the classification accuracy of PSO-LS-SVM, APSO-LS-SVM and IMPSON-LS-SVM algorithms are 100% for Setosa under the same number of training samples. For Versicolor, the classification accuracy of PSO-LS-SVM algorithm is 97.1%, the classification accuracy of APSO-LS-SVM algorithm is 97.1%, and the classification accuracy of IMPSON-LS-SVM algorithm is 97.1%. For Virginica, the classification accuracy of PSO-LS-SVM algorithm is 88.6%, the classification accuracy of APSO-LS-SVM algorithm is 94.3%, and the classification accuracy of IMPSON-LS-SVM algorithm is 97.1%. So the classification accuracy of the IMPSON-LS-SVM algorithm is better than the APSO-LS-SVM and IMPSON-LS-SVM algorithms. The results show that the IMPSON-LS-SVM algorithm takes on the strong generalization ability.

## 5. Conclusion

Data classification is an important concept in the data mining. And support vector machine (SVM) is based on statistical learning theory, and can effectively solve the classification and regression estimation with small sample and nonlinear problem. In allusion to the existing shortcomings of the LS-SVM, an improved particle swarm optimization algorithm is proposed to optimize and select parameters of the LS-SVM with radial basis kernel function for obtaining IMPSO-LS-SVM algorithm. IRIS data is used to test the effectiveness of the IMPSO-LS-SVM classification algorithm by computer simulation. The simulation results show that the classification ability of the IMPSO-LS-SVM is better than the PSO-LS-SVM algorithm and APSO-LS-SVM algorithm under the small sample data. The IMPSO-LS-SVM classification algorithm takes on the strong generalization ability, better generalization performance, faster training speed and higher classification precision.

## References

- [1] V. Nedic, S. Cvetanovic, and D. Despotovic, "Data mining with various optimization methods", *Expert Systems with Applications*, vol. 41, no. 8, (2014), pp. 3993-3999.
- [2] J. X. Dong, A. Krzyzak and Y. Suen, "Fast SVM training algorithm with decomposition on very large data sets", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, (2005), pp. 603-618.
- [3] V. Vapnik, "Statistical Learning Theory", New York: John Wiley & Sons, (1998), pp. 253-256.
- [4] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization", *Proceeding of the IEEE International Conference on Neural Networks*, Piscataway, NJ, (1995), pp. 1942-1948.
- [5] G. Acampora, M. Gaeta and V. Loia, "Combining multi-agent paradigm and memetic computing for personalized and adaptive learning experiences", *Computational Intelligence*, vol. 27, no. 2, (2011), pp. 141-165.
- [6] I. Ellabib, P. Calamai and O. Basir, "Exchange strategies for multiple ant colony system", *Information Sciences*, vol. 177, no. 5, (2007), pp. 1248-1264.
- [7] G. Shen and Y. Q. Zhang, "A new evolutionary algorithm using shadow price guided operators", *Applied Soft Computing*, vol. 11, no. 2, (2011), pp. 1983-1992.
- [8] J. A. K. Suykens, T. Van Gestel and J. De Brabanter, "Least squares support vector machines", Singapore: World Scientific, (2002).
- [9] H. Guo, H. P. Liu and L. Wang, "Method for selecting parameters of least squares support vector machines and application", *Journal of System Simulation*, vol. 18, no. 7, (2006), pp. 2033-2036+2051.
- [10] S. Chen, J. N. Zhu, J. Pan and H. B. Shi, "Parameters optimization of LS-SVM and its application", *Journal of East China University of Science and Technology*, vol. 34, no. 2, (2008), pp. 278-282.
- [11] Z. Y. Luo, P. Wang, Y. G. Li, W. F. Zhang, W. Tang and M. Xiang, "Quantum-inspired evolutionary tuning of SVM parameters", *Progress in Natural Science*, vol. 18, no. 4, (2008), pp. 475-480.
- [12] H. Wang, F. Q. Deng and Z. M. Chen, "Fuzzy LS-SVM classifier based on chaos genetic algorithm and its application", *Journal of South China University of Technology (Natural Science)*, vol. 39, no. 5, (2011), pp. 49-54.
- [13] H. Q. Zhu, C. H. Yang, J. Wang and W. H. Gui, "Hybrid-QPSO-based parameters optimization of LS-SVM and its application", *Journal of Central South University (Science and Technology)*, vol. 42, no. 4, (2011), pp. 1000-1004.
- [14] L. W. Wei, Z. Y. Chen and J. P. Li, "Evolution strategies based adaptive Lp LS-SVM", *Information Sciences*, vol. 181, no. 14, (2011), pp. 3000-3016.
- [15] S. D. S. Glauber, L. G. J. Luvizotto, V. C. Mariani and D. S. Coelho, "Least squares support vector machines with tuning based on chaotic differential evolution approach applied to the identification of a thermal process", *Expert Systems with Applications*, vol. 39, no. 5, (2012), pp. 4805-4812.
- [16] K. H. Yang and L. L. Zhao, "A new optimizing parameter approach of LSSVM multiclass classification model", *Neural Computing and Applications*, vol. 21, no. 5, (2012), pp. 945-955.
- [17] Z. Q. Zhang, L. Zhen, N. Y. Deng and J. Y. Tan, "Sparse least square twin support vector machine with adaptive norm", *Applied Intelligence*, vol. 41, no. 4, (2014), pp. 1097-1107.
- [18] X. B. Wang, J. H. Wen, S. Alam, X. Gao, Z. Jiang and J. Zeng, "Sales growth rate forecasting using improved PSO and SVM", *Mathematical Problems in Engineering*, (2014).
- [19] Z. H. Xue, P. J. Du and H. J. Su, "Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, (2014), pp. 2131-2146.
- [20] P. J. Nieto, J. R. A. Fernández, V. M. Suárez, C. D. Muñiz, E. García-Gonzalo and R. M. Bayón, "A hybrid PSO optimized SVM-based method for predicting of the cyanotoxin content from experimental



cyanobacteria concentrations in the Trasona reservoir: A case study in Northern Spain”, Applied Mathematics and Computation, vol. 260, (2015), pp. 170-187.

## Authors



**Chong Wu**, Assitant researcher, received the Doctor degree in forest tree genetics and breeding from research institute of tropical forestry Chinese academy of forestry in 2013, Guangzhou, China. The main research directions: Forest tree germplasm resources and genetic diversity, Data classification.



**Chonglu Zhong**, Researcher, Ph.D. supervisor, received the Doctor degree from Chinese academy of forestry in 2000, Beijing, China. The main research directions: Tropical forest tree genetic breeding.

**Yanlei Yin**, Associate researcher, received the Master degree from Shandong Agricultural University in 2008, Taian, China. The main research directions: Collection and evaluation of forest plant germplasm.

