

Parameter Optimization of SVM Based on Improved ACO for Data Classification

Wen Chen^{1,2} and Yixiang Tian¹

¹*School of Management and Economics, University of Electronic Science and Technology of China, Chengdu, 611731 China*

²*Sichuan University of Arts and Science, Dazhou, 635000 China*

Abstract

The parameters of support vector machine have a great influence on the learning ability and generalization ability, so an improved ant colony optimization algorithm is proposed to optimize the parameters of SVM, then an optimized SVM classifier (IMACO-SVM) is proposed for data classification. In the IMACO-SVM, the adaptive adjustment pheromone strategy is used to make relatively uniform pheromone distribution and the improved pheromone updating method is used to submerge the heuristic factor by the residual pheromone information, in order to effectively solve the contradiction between expanding search and finding optimal solution. The selection of parameters of the SVM is regarded as a combination optimization of parameters in order to establish the objective function of combination optimization. The improved ACO algorithm with good robustness and positive feedback characteristics and parallel searching is used to search for the optimal value of objective function. In order to validate the classification effectiveness of the IMACO-SVM algorithm, some experimental data from the UCI machine learning database are selected in this paper. The classification results show that the proposed IMACO-SVM algorithm has higher classification ability and classification accuracy.

Keywords: *Ant colony optimization algorithm; support vector machine; data classification; parameter optimization; kernel function*

1. Introduction

Data classification is an important work in the field of data mining and machine learning by deducing the unknown classes of samples using the learning of the known classes of samples [1]. In recent years, a lot of algorithms have been proposed to directly complete data classification [2], such as decision tree, Bayes, neural network, genetic algorithm, support vector machine, random forest, Adaboost, and so on. These algorithms can effectively solve the discrete attribute data with the low dimension. However, in the real world database, a large number of continuous valued attributes exist. If the data is continuous attributes, then the data are discretized in order to be classified. So it is practical significance to research data classification.

Support vector machine (SVM) is a new learning technique based on the structural risk minimization principle [3]. It likes a neural network that can approximate any complex nonlinear system. It can better solve these classification problems with small sample, nonlinear and high dimensional number, and has good generalization ability. And it well solves the nonlinear, dimension disaster, local optimal solution of neural network. So it has widely been applied in the field of pattern recognition, regression estimation and so on. The SVM is a new machine learning method, which exists some improvements in practical application. The fitting precision and generalization ability of SVM depend on the selection of relevant parameters. The selection results of parameters will directly relate to the

classification effect. Therefore, selection parameters of SVM is a bottleneck problem in the practical application. At present, a lot of optimization methods are proposed to select parameters of SVM, such as experimental method, the grid method, gradient descent algorithm, genetic algorithm, ant colony optimization algorithm, particle swarm optimization algorithm and so on. Rättsch *et al.* [4] exemplified this translation procedure for a new algorithm-one-class leveraging-starting from the one-class support vector machine (1-SVM). Pang *et al.* [5] proposed a method for authenticating an individual's membership in a dynamic group without revealing the individual's identity and without restricting the group size and/or the members of the group. Brudzewski *et al.* [6] proposed the application of support vector machine (SVM) neural approach to the calibration of the electronic nose arrangement for milk recognition. Kalatzis *et al.* [7] proposed an SVM-based computer classification system or discriminating depressive patients from healthy controls. Pang *et al.* [8] proposed a new membership authentication method by face classification using a support vector machine (SVM) classification tree, in which the size of membership group and the members in the membership group can be changed dynamically. Mavroforakis and Theodoridis [9] proposed a geometric approach for the support vector machine (SVM) classification problem to provide an intuitive ground for the understanding and the application of geometric optimization algorithms. Hao *et al.* [10] proposed a novel hierarchical classification method that generalizes support vector machine learning that is based on the results of support vector clustering method, and are structured in a way that mirrors the class hierarchy. Ghoggali and Melgani [11] proposed a novel methodological solution based on semi-supervised multi-temporal classification method. Lin *et al.* [12] proposed an integrated model for recognizing power-quality disturbances (PQD) using a novel wavelet multiclass support vector machine (WMSVM). Yan *et al.* [13] proposed an effective mutual information-based feature selection approach for EMG-based motion classification task. Cho *et al.* [14] proposed a two-stage scheme for the classification of NIR spectra data. Avci and Varol [15] proposed a new methodology based on invariant moments and multi-class support vector machine (MCSVM) for classification of human parasite eggs in microscopic images. Chen *et al.* [16] proposed an adaptive binary tree (ABT) to reduce the test computational complexity of multiclass support vector machine (SVM). Eristi and Demir [17] proposed a new approach for automatic classification of power quality events, which is based on the wavelet transform and support vector machines. Deselaers *et al.* [18] proposed a new technique that employs support vector machines (SVMs) and Gaussian mixture densities (GMDs) to create a generative/discriminative object classification technique using local image features. Pang *et al.* [19] proposed a new concept of a knowledgeable neighborhood and a transductive Support Vector Machine (SVM) classification tree (t-SVMT) for PTL. Zhao *et al.* [20] proposed a weighted maximum margin criterion to optimize the data-dependent kernel, which makes the minority class more clustered in the induced feature space. Fu and Lee [21] proposed a general approach to utilize decision values in SVM to identify indistinguishable documents. Tian *et al.* [22] proposed an improved KPCA/GA-SVM classification model for plant leaf disease recognition. Maji *et al.* [23] proposed a class of nonlinear kernel SVMs admits approximate classifiers with runtime and memory complexity that is independent of the number of support vectors. Soroor and Hossein [24] hybridized the gravitational search algorithm (GSA) with support vector machine (SVM) and make a novel GSA-SVM hybrid system to improve classification accuracy with an appropriate feature subset in binary problems. Liu *et al.* [25] proposed a multi-fault classification model based on the kernel method of support vector machines (SVM) and wavelet frame. Chen *et al.* [26] proposed a novel classification algorithm of hyper-spectral imagery based on

ant colony compositely optimizing support vector machine in spatial and spectral features. Wang *et al.* [27] proposed a novel kNN scheme by incorporating a structurally regularized support vector machines (SVM). Swarnajyoti, and Lorenzo [28] proposed a novel iterative active learning technique based on self-organizing map (SOM) neural network and support vector machine (SVM) classifier. Cheng *et al.* [29] proposed a new incremental learning approach to endow a Takagi-Sugeno-type fuzzy classification model with high generalization ability. Hou *et al.* [30] proposed a novel tensor-based method, *i.e.*, multiple rank multi-linear SVM (MRMLSVM) Maldonado and Montecinos [31] proposed different techniques for classification, such as Support Vector Data Description (SVDD) and two-class SVMs. Martínez *et al.* [32] proposed a new classification methodology based on binary classifiers constructed using support vector machines and applying a one-versus-all approach supported by the use of the directed acyclic graphs. Peng *et al.* [33] combined signal-noise ratio (SNR) and under-sampling technique based on K-means to propose an improved SVM classification method. Li *et al.* [34] proposed a new chaos embedded GSA-SVM (CGSA-SVM) hybrid system based on hybridizing chaotic search and gravitational search algorithm (GSA) with SVM to improve classification accuracy of SVM. Guo and Wang [35] proposed a multi-class classification model that is based on active learning and support vector machines (MC-SVMA), which can be used to address unlabeled data. Xu *et al.* [36] proposed online support vector machine (SVM) classification learning algorithms with uniformly ergodic Markov chain (u.e.M.c.) samples. Yin *et al.* [37] proposed a scene classification approach based on single-layer sparse auto-encoder (SAE) and support vector machine (SVM). García-Gutiérrez *et al.* [38] proposed a novel contextual classifier based on a Support Vector Machine (SVM) and an Evolutionary Majority Voting (SVM-EMV) to develop thematic maps from LiDAR and imagery data. Mrutyunjaya and Ajith [39] proposed a hybrid fuzzy rough with K-nearest neighbor (K-NN)-based classifier (FRNN) to classify the patterns in the reduced datasets, obtained from the fuzzy rough bio-inspired algorithm search.

The proposed algorithms can better solve the values of parameters of support vector machine, but there exists some problems. The adaptive adjustment pheromone strategy is used to make relatively uniform pheromone distribution and an improved pheromone updating method is used to submerge the heuristic factor by the residual pheromone information, in order to effectively solve the contradiction between expanding search and finding optimal solution. Then improved ant colony optimization algorithm is obtained to optimize the parameters of SVM in order to propose an optimized SVM classifier (IMACO-SVM) for solving data classification problem. And some experimental data from UCI machine learning database are selected to test the classification performance.

2. The ACO Algorithm

The ACO algorithm [40] is a metaheuristic inspired by the behavior of real ants in their search for the shortest path to food. The ACO algorithm consists of a number of iterations of solution construction. In the ACO algorithm, it simulates the optimization of ant foraging behavior. The procedure of the ACO algorithm is given in Figure 1.

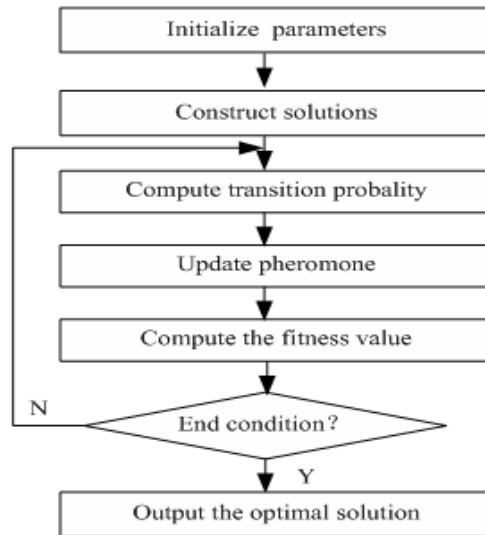


Figure 1. The Flow of ACO Algorithm

The procedure of pheromone update rule is shown as follows:

(1) The transition rule

To visit the next city s with the probability p_k :

$$P_k = \begin{cases} \frac{\tau^\alpha \cdot \eta^\beta}{\sum_{J_r^k} \tau^\alpha \cdot \eta^\beta} & \text{if } s \in J_r^k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In the formula(1), p_k is the transition probability, τ is the intensity of pheromone, η is the length of the path, J_r^k is the set of unvisited cities of the k^{th} ant in the i^{th} population, the parameter α and β are the control parameters.

(2) The pheromone update rule

The local trail updating formula is described:

$$\tau(u) = (1 - \rho)\tau(s) + \sum_{k=1}^m \Delta \tau_k(s) \quad (2)$$

Where ρ ($0 < \rho < 1$) is the pheromone trail evaporating rate. $\Delta \tau_k(s)$ is the amount of pheromone trail added to the edge. It is described:

$$\Delta \tau_k(s) = \begin{cases} \frac{Q}{L_k} & (r, s) \in \pi_k \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

3. An Improved Ant Colony Optimization Algorithm

The ACO algorithm is to combine heuristic algorithm and positive feedback mechanism of pheromone. The random selection strategy is widely applied in the search process. The positive feedback mechanism of pheromone is used to enhance the finding optimal solution by ants. The ACO algorithm uses a fixed pheromone amount for updating the pheromone. This method exists slow convergence phenomenon, resulting in falling into local optimum and causing premature phenomenon. So an

adaptive adjustment pheromone strategy is used to make relatively uniform pheromone distribution, it can effectively solve the contradiction between expanding search and finding optimal solution.

The real variable function $Q(t)$ is selected to replace the constant of pheromone intensity Q in the adjusting pheromone $\Delta \tau_{ij}^k = Q / L_k$.

$$\Delta \tau_{ij}^k(t) = Q(t) / L_k \quad (4)$$

$$Q(t) = \begin{cases} Q_1 & t \leq T_1 \\ Q_2 & T_1 < t \leq T_2 \\ Q_3 & T_2 < t \leq T_3 \end{cases} \quad (5)$$

The real variable function $Q(t)$ is used to replace the constant term Q in order to continue to maintain the exploration and exploitation balance point between the random search of ant and the evocation function of path information under the pheromone evaporation or increasing.

At the same time, an improved pheromone updating method is used to submerge the heuristic factor by the residual pheromone information. The pheromone updating method with elitist strategy of ACO algorithm is described:

$$\tau_{ij}(t+1) = \rho \tau_{ij}(t) + \Delta \tau_{ij} + \Delta \tau_{ij}^* \quad (6)$$

$$\Delta \tau_{ij}^* = \sum_{k=1}^m \Delta \tau_{ij}^k \quad (7)$$

$$\Delta \tau_{ij}^k = \begin{cases} Q / L_k, \text{ passed path } (i, j) \text{ by ant } k \text{ in this iteration} \\ 0, \text{ other} \end{cases} \quad (8)$$

$$\Delta \tau_{ij}^* = \begin{cases} \delta Q / L^*, \text{ the edge } (i, j) \text{ is one of the found optimal solutions} \\ 0, \text{ other} \end{cases} \quad (9)$$

4. Support Vector Machine

Support vector machine (SVM)[3] based on structural risk minimization is one of the popular tools in supervised machine learning methods. It is to map the original nonlinear data into a higher-dimensional feature space. A hyperplane is constructed to maximize the margin of separation between itself and lying nearest points. The hyperplane is used to classify these unknown data.

Give the training sample $S = \{(x_i, y_i) | i = 1, 2, 3, \dots, m\}$, m is the number of samples, the set $\{x_i\} \in R_n$ represents the input vector, $y \in \{-1, 1\}$ represents the corresponding desired output vector, the input data is mapped into the high dimensional feature space by nonlinear mapping function $\phi(\bullet)$. So the classification hyperplane need meet the condition:

$$\begin{cases} \omega^T x_i + b \geq 1, & y_i = 1 \\ \omega^T x_i + b \leq -1, & y_i = -1 \end{cases} \quad (10)$$

Where ω is Omega vector of superplane, b is offset quantity. Then the classification decision function is described as follow:

$$f(x_i) = \text{sgn}(\omega^T x_i + b) \quad (11)$$

The classification model of LS-SVM model is described by the optimization function

$$\min_{\omega, \xi, b} J(\omega, \xi_i) :$$

$$\min_{\omega, \xi, b} J(\omega, \xi_i) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^m \xi_i^2 \quad (12)$$

$$s.t. \quad (13)$$

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

where ξ_i is slack variable, b is offset, ω is support vector, $\xi = (\xi_1, \xi_2, \dots, \xi_m)$, γ is classification parameter in order to balance the fitting 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:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{i=1}^m \xi_i^2 - \sum_{k=1}^m \alpha_k \{ y_k [\omega^T \phi(x_k) + b] - 1 + \xi_k \} \quad (14)$$

The following linear equation is obtained:

$$\begin{bmatrix} 0 & L^T \\ L & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (16)$$

Where $Y = [y_1, y_2, \dots, y_m]^T \in R^m$, $L \in R^m$ is vector of the element m , $y^T = [y_1, y_2, \dots, y_m]$, 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 follow:

$$f(x_i) = \text{sgn} \left(\sum_{i=1}^m \alpha_i y_i K(x, x_i) + b \right) \quad (17)$$

There have linear kernel function, polynomial kernel function, radial basis kernel function (RBF), Sigmoid kernel function and Fourier kernel function and so on. For the data classification, because the RBF has the advantages of simple form, symmetry radial, good smoothness and analyticity, the RBF is selected to be regarded as kernel function of the SVM model. The expression of the RBF is given:

$$K(x, x_i) = \exp[-(x - x_i)^2 / 2\sigma^2] \quad (18)$$

5. Parameter Optimization of SVM Based on Improved ACO

The values of parameters in the SVM have a great influence on the learning ability and generalization ability, so it is a key to determine the values of parameters in the SVM. For RBF kernel function, the tuning parameter C a compromise between the structural risk and sampling error, its value is related to the tolerable error. The larger value of tuning parameter C allows smaller error, the smaller value of tuning parameter C allows larger error. Kernel width σ is related to the inputting space range or width of samples. If the inputting space range is larger, the value of kernel width σ is larger. Conversely, if the inputting space range is smaller, the value of kernel width σ is smaller. The ACO algorithm is a novel simulation evolution algorithm, it belongs to a kind of swarm intelligence algorithm. In order to solve the parameter optimization of the SVM, an improved ACO algorithm is used to optimize the parameters of SVM in order to propose a optimized SVM classifier (IMACO-SVM). The improved ACO algorithm has good robustness and positive feedback characteristics, and parallel searching advantage. The selection of parameters in the SVM is regarded as a combination optimization

of parameters in order to establish the objective function of combination optimization. The improved ACO algorithm is used to search for the optimal value of objective function.

The specific steps of IMACO-SVM are described as follows:

Step 1. Initialize

Set ant size (m), the maximum number of iteration (T). A smaller value of τ_0 is initialized. All τ_d^{ij} are initialized. The first step of all ants is set as $t = 0$.

Step 2. For each ant, the formula (6) to formula (8) are executed.

Step 3. The next visited city is selected according to the formula (1).

Step 4. When each ant completed one path, the local update rule is executed immediately according to the formula (6).

Step 5. Save the optimal solution in each iteration. During each iteration, the ant with most pheromone is saved, then the error value is calculated. Return to Step.2 to continue iteration.

Step 6. Obtain global optimal solution. If the number of iteration meet the maximum number of iteration, the search is completed, so as to obtain the best ant, which is converted into the tuning parameter C and kernel width σ of the SVM.

Step 7. Determine the values of the tuning parameter C and kernel width σ , which to train the training set in order to generate the objective function.

Step 8. Calculate the value of corresponding objective function, and the error is analyzed. Obtain the IMACO-SVM model.

6. Simulation Experiment

6.1. Data Source

In order to verify the effectiveness of the proposed IMACO-SVM algorithm for data classification, some experimental data from the UCI machine learning database are selected in this paper. The data sets are shown in Table1.

Table 1. The Selected Data Sets

Index	Data	Sample number	Attribute	Class
1	Auto	392	7	3
2	Breast	683	9	2
3	Bupa	345	6	2
4	Heart	296	13	2
5	Iris	150	4	3
6	Glass	214	9	7
7	Pima	768	8	2
8	Vehicle	846	18	4
9	Wine	17813	13	3

6.2. Experimental Environment and Parameters

The environments are followed: the Pentium CPU 2.40GHz, 4.0GB RAM, Windows XP system, Matlab2012b. The initial values of parameters could seriously

affect the experiment result, so the most reasonable initial values of parameters are got by testing and modifying them. The obtained initial values of these parameters are: ants $m = 30$, pheromone factor $\alpha = 1.0$, heuristic factor $\beta = 1.0$, evaporation factor $\rho = 0.80$, step length is 0.1, pheromone amount $q = 60$, maximum iteration times $T_{max} = 500$. The tuning parameter C and kernel width σ is coded by binary, $C \in [0, 1000]$, $\sigma \in [0.1, 100]$, the error of optimal solution $\varepsilon = 0.01$. Because the algorithm takes on the volatility, in order to fairly evaluate the accuracy of the algorithms, each data set is executed 30 times. Then the average value is regarded as the comparison results.

6.3. Experimental Results and Analysis

The SVM and PSO-SVM are select to compare with the proposed IMACO-SVM algorithm. The data classification result is shown in Table2.

Table 2. The Data Classification Result

Index	Data	Sample number	Average classification accuracy(%)		
			SVM	PSO-SVM	IMACO-SVM
1	Auto	392	76.79	79.08	81.89
2	Breast	683	94.73	95.17	96.93
3	Bupa	345	63.53	69.28	72.20
4	Heart	296	73.99	78.04	81.77
5	Iris	150	97.33	98.00	98.67
6	Glass	214	41.59	53.74	63.08
7	Pima	768	78.65	84.77	86.98
8	Vehicle	846	77.78	82.98	86.29
9	Wine	178	82.02	90.45	92.70

As can be seen from Table2, the proposed IMACO-SVM algorithm can obtain better classification results for experimental data from the UCI machine learning database than the SVM model and PSO-SVM algorithm. For experimental data of Breast, Iris and Wine, the average classification accuracy respectively is 96.93%, 98.67% and 92.70%. And the PSO-SVM algorithm can obtain better classification results than the SVM for experimental data.

In general, the classification results of proposed IMACO-SVM algorithm are more better and has higher optimization ability and classification accuracy.

7. Conclusion

Support vector machine is a simple method to realize the pattern recognition. It does not require a long training process. The optimal hyperplane is solved according to the initial sample in order to identify other unknown samples. But the values of parameters in the SVM have a great influence on the learning ability and generalization ability, so it is a key to determine the values of parameters in the SVM. The ACO algorithm is simple, versatile, robust and parallel method. It has been successfully applied in quadratic assignment problems, dynamic routing

problems, scheduling problems and so on. The positive feedback mechanism of pheromone is used to enhance the finding optimal solution by ants. So an improved particle swarm optimization algorithm is proposed to optimize parameters of the SVM for obtaining a optimal SVM model(IMACO-SVM). Some experimental data from the UCI machine learning database are selected to validate the classification effectiveness of IMACO-SVM algorithm. The experimental results show that the classification ability of the IMACO-SVM algorithm is better than the SVM and PSO-SVM algorithm under the small sample data. The proposed IMACO-SVM algorithm has higher classification ability and classification accuracy.

Acknowledgements

This research was supported by National Social Science Fund under Grant No. 14BJY174, and The Ministry of education of Humanities and Social Science project under Grant No 12YJA790125.

References

- [1] J. E. Gentle, W. K. Hardle and Y. C. Mori, "Handbook of Computational Statistics: Concepts and Methods", Germany: Springer Press, (2012), pp. 883-926.
- [2] M. Boucekine, A. Loundou and K. Baumstarck, "Using the random forest method to detect a response shift in the quality of life of multiple sclerosis patients: a cohort study", BMC medical research methodology, vol. 13, no. 1, (2013), pp. 1-8.
- [3] V. Cortesc, "Support vector networks. Machine Learning", vol. 20, no. 3, (1995), pp. 273-297.
- [4] G. Rätsch, S. Mika, B. Schölkopf and K. R. Müller, "Constructing boosting algorithms from SVMs: An application to one-class classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, (2002), pp. 1184-1199.
- [5] S. N. Pang, D. J. Kim and S. Y. Bang, "Membership authentication in the dynamic group by face classification using SVM ensemble", Pattern Recognition Letters, vol. 24, no. 1-3, (2003), pp. 215-225.
- [6] K. Brudzewski, S. Osowski and T. Markiewicz, "Classification of milk by means of an electronic nose and SVM neural network", Sensors and Actuators, B: Chemical, vol. 98, no. 2-3, (2004), pp. 291-298.
- [7] I. Kalatzis, N. Piliouras, E. Ventouras, C. C. Papageorgiou, A. D. Rabavilas and D. Cavouras, "Design and implementation of an SVM-based computer classification system for discriminating depressive patients from healthy controls using the P600 component of ERP signals", Computer Methods and Programs in Biomedicine, vol. 75, no. 1, (2004), pp. 11-22.
- [8] S. N. Pang, D. J. Kim and S. Y. Bang, "Face membership authentication using SVM classification tree generated by membership-based LLE data partition", IEEE Transactions on Neural Networks, vol. 16, no. 2, (2005), pp. 436-446.
- [9] M. E. Mavroforakis and S. Theodoridis, "A geometric approach to support vector machine (SVM) classification", IEEE Transactions on Neural Networks, vol. 17, no. 3, (2006), pp. 671-682.
- [10] P. Y. Hao, J. H. Chiang and Y. K. Tu, "Hierarchically SVM classification based on support vector clustering method and its application to document categorization", Expert Systems with Applications, vol. 33, no. 3, (2007), pp. 627-635.
- [11] N. Ghoggali and F. Melgani, "Genetic SVM approach to semi-supervised multi-temporal classification", IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 2, (2008), pp. 212-216.
- [12] W. M. Lin, C. H. Wu, C. H. Lin and F. S. Cheng, "Detection and classification of multiple power-quality disturbances with wavelet multiclass SVM", IEEE Transactions on Power Delivery, vol. 23, no. 4, (2008), pp. 2575-2582.
- [13] Z. G. Yan, Z. H. Wang and X. B. Xie, "The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification", Computer Methods and Programs in Biomedicine, vol. 90, no. 3, (2008), pp. 275-284.
- [14] H. W. Cho, S. H. Baek, E. Youn, M. K. Jeong and A. Taylor, "A two-stage classification procedure for near-infrared spectra based on multi-scale vertical energy wavelet thresholding and SVM-based gradient-recursive feature elimination", Journal of the Operational Research Society, vol. 60, no. 8, (2009), pp. 1107-1115.
- [15] D. Avci and A. Varol, "An expert diagnosis system for classification of human parasite eggs based on multi-class SVM", Expert Systems with Applications, vol. 36, no. 1, (2009), pp. 43-48.
- [16] J. Chen, C. Wang and R. S. Wang, "Adaptive binary tree for fast SVM multiclass classification", Neurocomputing, vol. 72, no. 13-15, (2009), pp. 3370-3375.
- [17] H. Eristi and Y. Demir, "A new algorithm for automatic classification of power quality events based on wavelet transform and SVM", Expert Systems with Applications, vol. 37, no. 6, (2010), pp. 4094-4102.

- [18] T. Deselaers, G. Heigold and H. Ney, "Object classification by fusing SVMs and Gaussian mixtures", *Pattern Recognition*, vol. 43, no. 7, (2010), pp. 2476-2484.
- [19] S. N. Pang, T. Ban, Y. Kadobayashi and N. Kasabov, "Personalized mode transductive spanning SVM classification tree", *Information Sciences*, vol. 181, no. 11, (2011), pp. 2071-2085.
- [20] Z. Y. Zhao, P. Zhong and Y. H. Zhao, "Learning SVM with weighted maximum margin criterion for classification of imbalanced data", *Mathematical and Computer Modeling*, vol. 54, no. 3-4, (2011), pp. 1093-1099.
- [21] J. Fu and S. L. Lee, "A multi-class SVM classification system based on learning methods from indistinguishable Chinese official documents", *Expert Systems with Applications*, vol. 39, no. 3, (2012), pp. 3127-3134.
- [22] J. Tian, Q. X. Hu, X. Y. Ma and M. Y. Han, "An improved KPCA/GA-SVM classification model for plant leaf disease recognition", *Journal of Computational Information Systems*, vol. 8, no. 18, (2012), pp. 7737-7745.
- [23] S. Maji, A. C. Berg and J. Malik, "Efficient classification for additive kernel SVMs", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, (2013), pp. 66-77.
- [24] S. Soroor and N. P. Hossein, "Facing the classification of binary problems with a GSA-SVM hybrid system", *Mathematical and Computer Modeling*, vol. 57, no. 1-2, (2013), pp. 270-278.
- [25] Z. W. Liu, H. R. Cao, X. F. Chen, Z. J. He and Z. J. Shen, "Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings", *Neurocomputing*, vol. 99, no. 1, (2013), pp. 399-410.
- [26] S. J. Chen, Y. H. Hu, L. Shi, L. Wang, S. J. Sun and S. L. Xu, "Classification of hyper-spectral imagery based on ant colony compositely optimizing SVM in spatial and spectral features", *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis*, vol. 33, no. 8, (2013), pp. 2192-2197.
- [27] X. C. Wang, Y. M. Niu, L. W. Tan and S. X. Zhang, "New classification method of combining kNN and structurally regularized SVM", *ICIC Express Letters*, vol. 8, no. 6, (2014), pp. 1689-169.
- [28] P. Swarnajyoti and B. Lorenzo, "A novel SOM-SVM-based active learning technique for remote sensing image classification", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, (2014), pp. 6899-6910.
- [29] W. Y. Cheng and C. F. Juang, "A fuzzy model with online incremental SVM and margin-selective gradient descent learning for classification problems", *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 2, (2014), pp. 324-337.
- [30] C. P. Hou, F. P. Nie, C. S. Zhang, D. Y. Yi and Y. Wu, "Multiple rank multi-linear SVM for matrix data classification", *Pattern Recognition*, vol. 47, no. 1, (2014), pp. 454-469.
- [31] S. Maldonado and C. Montecinos, "Robust classification of imbalanced data using one-class and two-class SVM-based multi-classifiers", *Intelligent Data Analysis*, vol. 18, no. 1, (2014), pp. 95-112.
- [32] J. Martínez, C. Iglesias, J. M. Matías, J. Taboada and M. Araújo, "Solving the slate tile classification problem using a DAGSVM multi-classification algorithm based on SVM binary classifiers with a one-versus-all approach", *Applied Mathematics and Computation*, vol. 230, (2014), pp. 464-472.
- [33] L. Peng, T. T. Bi and L. Yang, "SVM classification for high-dimensional imbalanced data based on SNR and under-sampling", *International Journal of Multimedia and Ubiquitous Engineering*, vol. 10, no. 4, (2015), pp. 105-112.
- [34] C. S. Li, X. L. An and R. H. Li, "A chaos embedded GSA-SVM hybrid system for classification", *Neural Computing and Applications*, vol. 26, no. 3, (2015), pp. 713-721.
- [35] H. S. Guo and W. J. Wang, "An active learning-based SVM multi-class classification model", *Pattern Recognition*, vol. 48, no. 5, (2015), pp. 1577-1597.
- [36] J. Xu, Y. T. Yuan, B. Zou, Z. B. Xu, L. Q. Li and Y. Lu, "The generalization ability of online SVM classification based on Markov sampling", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 3, (2015), pp. 628-639.
- [37] H. P. Yin, X. G. Jiao, Y. Chai and B. Fang, "Scene classification based on single-layer SAE and SVM", *Expert Systems with Applications*, vol. 42, no. 7, (2015), pp. 3368-3380.
- [38] J. G. Gutiérrez, D. M. García, M. Garcia and J. C. Riquelme-Santos, "An evolutionary-weighted majority voting and support vector machines applied to contextual classification of LiDAR and imagery data fusion", *Neurocomputing*, vol. 163, (2015), pp. 17-24.
- [39] P. Mrutyunjaya and A. Ajith, "Hybrid evolutionary algorithms for classification data mining", *Neural Computing and Applications*, vol. 26, no. 3, (2015), pp. 507-523.
- [40] M. Dorigo and M. Gambardella, "Ant colony system: A cooperative learning approach to the traveling salesman problem", *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, (1997), pp. 53-66.

Authors



Wen Chen, He received his MS degree in 2007 from University of Electronic Science and Technology of China. He is currently working toward PhD degree in UESTC. Since 2011, he has been working as a teacher at Sichuan University of Arts and Science, and since 2009, he was a lecturer in school of Mathematics and Finance of Sichuan University of Arts and Science. His research interests are machine learning and financial engineering.



Yixiang Tian, He received the BS degree from Sichuan Normal University (1985), and received the MS degree from Sichuan University (1993), and received PhD degree from Huazhong University of Science and Technology, China. He is currently a professor and doctoral supervisor in School of Management and Economics, University of Electronic Science and Technology of China. His research interests are in the areas of econometrics, financial risk and the data mining.

