Optimizing Support Vector Machine Parameters Using Continuous Ant Colony Optimization

Hiba Basim Alwan
School of Computing
College of Art and Sciences
University Utara Malaysia, 06010 Sintok, Kedah
hiba81basim@yahoo.com

Ku Ruhana Ku-Mahamud
School of Computing
College of Art and Sciences
University Utara Malaysia, 06010 Sintok, Kedah
ruhana@uum.edu.my

***Abstract*-Support Vector Machines are considered to be excellent patterns classification techniques. The process of classifying a pattern with high classification accuracy counts mainly on tuning Support Vector Machine parameters which are the generalization error parameter and the kernel function parameter. Tuning these parameters is a complex process and may be done experimentally through time consuming human experience. To overcome this difficulty, an approach such as Ant Colony Optimization can tune Support Vector Machine parameters. Ant Colony Optimization originally deals with discrete optimization problems. Hence, in applying Ant Colony Optimization for optimizing Support Vector Machine parameters, which are continuous parameters, there is a need to discretize the continuous value into a discrete value. This discretization process results in loss of some information and, hence, affects the classification accuracy and seek time. This study proposes an algorithm to optimize Support Vector Machine parameters using continuous Ant Colony Optimization without the need to discretize continuous values for Support Vector Machine parameters. Seven datasets from UCI were used to evaluate the performance of the proposed hybrid algorithm. The proposed algorithm demonstrates the credibility in terms of classification accuracy when compared to grid search techniques. Experimental results of the proposed algorithm also show promising performance in terms of computational speed.**

***Keywords*-Support Vector Machine; continuous Ant Colony Optimization; parameters optimization**

I. INTRODUCTION

Many decision-making processes are examples of classification difficulty that can be simply transformed into classification difficulty, e.g., prognosis processes, diagnosis processes, and pattern recognition [1]. The majority of recent researches center on enhancing classification accuracy by utilizing statistical approaches [2]. Pattern classification aims to classify input features into predetermined groups consisting of classes of patterns [3]. The Support Vector Machine (SVM) is a present day pattern classification approach. SVM originates from statistical learning approaches that utilize the concept of structural risk minimization [4] and [5]. This concept plans the data into high dimensional domains via a kernel function by using a kernel trick [4] and [6]. Polynomial, Radial Base Function (RBF), and sigmoid kernel function are three examples of kernel functions. RBF is the more popular kernel function because of its capability to manage high dimensional data [7], good performance in major cases [8] and it only needs one parameter, kernel parameter gamma (γ). Two problems in SVM classifier that influence the classification accuracy are: tuning SVM parameters, and selecting an optimal feature subset to be given to the SVM classifier. These problems affect each other [9]. This study focuses on tuning SVM parameters, also known as model selection.

There is no regular methodology that accepts advance approximation of optimal values for SVM parameters. In present classification work, obtaining good values for these parameters is not easy. It requires either an exhaustive search through the space of hyper variables or an optimization approach that searches simply a bounded sub group of the potential values. Currently, almost all SVM research chooses these variables experimentally via searching a bounded number of values and preserving those that supply the lowest amount of mistakes. This approach needs a grid search through the area of variable values and requires identifying the range of executable solution and best sampling step. This is a difficult task because best sampling steps change from kernel to kernel and grid ranges may not be simple to identify without advanced knowledge of the problem. Furthermore, when a hyper parameter exceeds two of the manual prototypes chosen, it may become intractable [10]. Approaches such as trial and error, grid search, cross validation, generalization error estimation and gradient descent, can be used to find optimal parameter values for SVM. Evolutionary approaches such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) may also be utilized [11].

ACO algorithms is applied to tune SVM parameters. These algorithms work through repetitive creation procedures where each procedure directs a dependent heuristic by intelligently mixing various ideas for exploring and exploiting the seek space. The learning fashions are utilized to construct information to efficiently obtain near optimal solutions. Solutions that are built using ACO seek to find the shortest way to the origin of food via pheromones [11]-[13]. ACO algorithms deal with discrete and continuous variables. However, ACO that deals with continuous variables is considered as a modern research field [14]-[17].

Ant Colony Optimization for continuous variables (ACO_R) uses Probability Density Function (PDF) instead of Discrete Probability Distribution, to determine the direction that an ant should follow; Gaussian function, a PDF is one of the most popular as it uses a very simple manner for data sampling. For each built solution, a density function is generated from a set of solutions that the technique preserves at all times. In order to maintain this set, the set is filled with nonsystematic solutions at the beginning. This is similar to initializing pheromone value in a discrete ACO approach. Then, at each loop, the group of created solutions is appended to the set and the equivalent number of worst solutions is deleted from the set. This work is similar to pheromone modification in discrete ACO. The goal is to influence the searching procedure to gain the best solution. Pheromone information is kept in a table when ACO for discrete combinatorial optimization is used. During each loop, when selecting a component to be appended to the current partial solution, an ant utilizes part of the values from that table as a discrete probability distribution. In contrast to the situation of continuous optimization, the selection that the ant makes is not limited to a finite group. Therefore, it is difficult to express the pheromone in the table structure. Instead of using a table, ACO_R uses solution archive to preserve the route for a number of solutions. Solution archive contains values of solution variables and objective functions. These values are then used to dynamically create PDF [16] and [17].

In this study, ACO_R is used to solve the SVM model selection problem. The rest of the paper is organized as follows. Section II reviews several literatures on tuning SVM parameters and Section III describes the proposed algorithm. Section IV presents the experimental results, and concluding remarks and future works are presented in Section V.

II. TUNING SUPPORT VECTOR MACHINE PARAMETER

Imbault & Lebart [18] suggested the use of global minimization approaches, which are GA and SA to solve model selection problems. They measured GA and SA with modified cooling approaches to automatically select the value at each step. Their experiments show that using a global minimization approach guarantees putting them in a good area, thereby preventing very large misclassification ratios. Also in their experiments, they saw that GA tends to be faster, SA needs few variables' setting while GA requires more. The primary disadvantage of these approaches is their calculation time. Frohlich & Zell [19] proposed the use of an online Gaussian Process (GP) from the locations in parameter space that have been visited. From their experiments, they found that online GP can be applied at a cheaper cost. Recent locations in parameter space are sampled based on the predicted enhancement condition. Adankan, Cheriet, & Ayat [20] suggested using a fast enhanced method for tuning SVM parameters based on an approximation of the gradient of the empirical error along with incremental learning, which reduces the resources required both in terms of processing time and of storage space. They tested their method on many benchmark data which produced promising results confirming their approach. The use of GA to optimize C and band width kernel function variable σ of the SVM was suggested by Abbas & Arif [12]. In their study, they proposed seven support vector machines, one for each day of the week, trained on previous data which was then utilized for the predication of daily peak load long range demand. From their results they concluded that their work gave outcomes better than the best paper of the competition. Dong, Xia, Tu, & Xing [21] proposed the cost variable and kernel variable expression as a two level optimization problem, where the values of variables change continuously and thus optimization approaches can be implemented to choose optimal variables. These variables can be calculated through cross-validation. To obtain optimal values, the variables are tested continuously instead of utilizing a discrete approach. Their prototype involves two phases. First, an SVM classifier built on the foundation of training data. Secondly, GA is used to seek optimal values. From their results they concluded that their proposed method often produces better results compared with pre-selected cost methods. Simple pre-selected cost methods work well on some datasets. Zhang [22] suggested using an automatic and successful model selection approach. His work built on evolutionary computation approaches and utilized recollection, accuracy and mistake ratio as optimization goals. The concept of constructing a kernel prototype is used which is then modified to the data group with the help of evolutionary computation approaches. The modification procedure is directed by the feedback information obtained from SVM execution. Both GA and PSO are used as evolutionary computation approaches to resolve optimization difficulty that occurs due to their robustness and global seeking capability. Saini, Aggarwal & Kumar [13] suggested using GA to optimize SVM variables. The regularization parameter C and kernel parameters are dynamically optimized through GA. In their work they used unconnected time strings for each worked trading interval instead of utilizing single time strings to model each day's price profile. From their experiments they concluded that their model supplies better predicting with sensible levels of accuracy and stability. A grid-based ACO technique was introduced by Zhang, Chen, Zhang, & He [23] to select variables C and RBF kernel σ automatically for SVM instead of choosing variables unsystematically through human skill to minimize generalization mistakes and generalization execution which may be enhanced concurrently. Their work provides high accuracy and less calculation time compared with other methods like grid algorithm and cross validation approach. RBF kernel is utilized to enhance the accuracy of SVM. However, one dataset is used to evaluate the performance of the proposed technique. ACO was also used by Fang & Bai [24] to optimize both SVM parameters, C and σ kernel function parameters in continuous fields. Both parameters C and σ are divided into a number of sub intervals. In each sub interval, one point is chosen unsystematically to be the location of artificial ants. Before starting each loop, advance knowledge and heuristic information are modified. In every loop, the transition probability of each ant is predetermined. The ant will move to the next interval if the state transition rule is met, otherwise, the ant will search for optimal variables within local intervals. Their results showed a very promising hybrid SVM model for forecasting share price in terms of accuracy and generalization ability. Lu, Zhou, He, & Liu [27] proposed using PSO for SVM parameter optimization. PSO is very suitable for global optimization. They considered these parameters as particles and PSO is applied to gain optimal values for these parameters. Their work shows that the accuracy and efficiency are enhanced.

III. THE PROPOSED ALGORITHM

This study constructs ACO_R to optimize SVM classifier parameters. An ant's solution is used to represent a combination of the classifier parameters, C and γ, based on the Radial Basis Function (RBF) kernel of the SVM classifier. The classification accuracy of the built SVM classifier is utilized to direct the updating of solution archives. Based on the solution archive, the transition probability is computed to choose a solution path for an ant. In implementing the proposed scheme, this study utilizes the RBF kernel function for SVM classifier because of its capability to manage high dimensional data [7], good performance in major cases [8], and it only needs to use one parameter: kernel parameter gamma (γ) [9]. The overall process to hybridize ACO_R and SVM (ACO_R-SVM) is as depicted in Fig. 1.

The main steps are (1) selecting feature subset (2) initializing solution archive and algorithm parameters, (3) solution construction for C and γ, (4) establishing SVM classifier model, and (5) updating solution archives. In the features subset selection step, F-score is used as a measurement to determine the importance of features. This measurement is used to judge the favoritism capability of a feature. High value of F-score indicates the most favorable feature. The calculation of F-score is as follows [28]:

$$F-Score_i = \frac{\sum_{c-1}^{v}\left(\bar{x}_i^{(c)}-\bar{x}_i\right)^2}{\sum_{c-1}^{v}\left(\frac{1}{N_i^{(c)}-1}\sum_{j=1}^{N_i^{(c)}}\left(x_{i,j}^{(c)}-\bar{x}_i^{(c)}\right)^2\right)}, i = 1,2,\ldots,N_f \quad (1)$$

where v is the number of categories of target variable, N_f is the number of features, $N_i^{(c)}$ is the number of samples of the ith feature with categorical value c, c ∈ {1, 2, …, v}, $\bar{x}_{i,j}^{(c)}$ is the jth training sample for the ith feature with categorical value c, j ∈ {1, 2, …, $N_i^{(c)}$}, $\bar{x}_i$ is the ith feature, and $\bar{x}_i^{(c)}$ is the ith feature with categorical value c.

After computing the F-score for each feature in the dataset, the average F-score is computed and is considered as the threshold for choosing features in the feature subset. Features with F-scores equal to or greater to the threshold are chosen and put in the feature subset and this subset is presented to the SVM.

In the initialization step, for each ant establishing a solution path for parameter C and parameter γ, two solution archives are needed to design the transition probabilities for C and for γ. The range value for C and γ are sampling according to random parameter k which is the size of solutions archives. The weight vector, w is then computed for each sample for C and γ as follows:

$$w_l = \frac{1}{qk\sqrt{2\pi}}e^{-\frac{(l-1)^2}{2q^2k^2}} \quad (2)$$

where k is the size of solution archive, and q is the algorithm's parameter to control diversification of search process. These values are stored in solution archives.

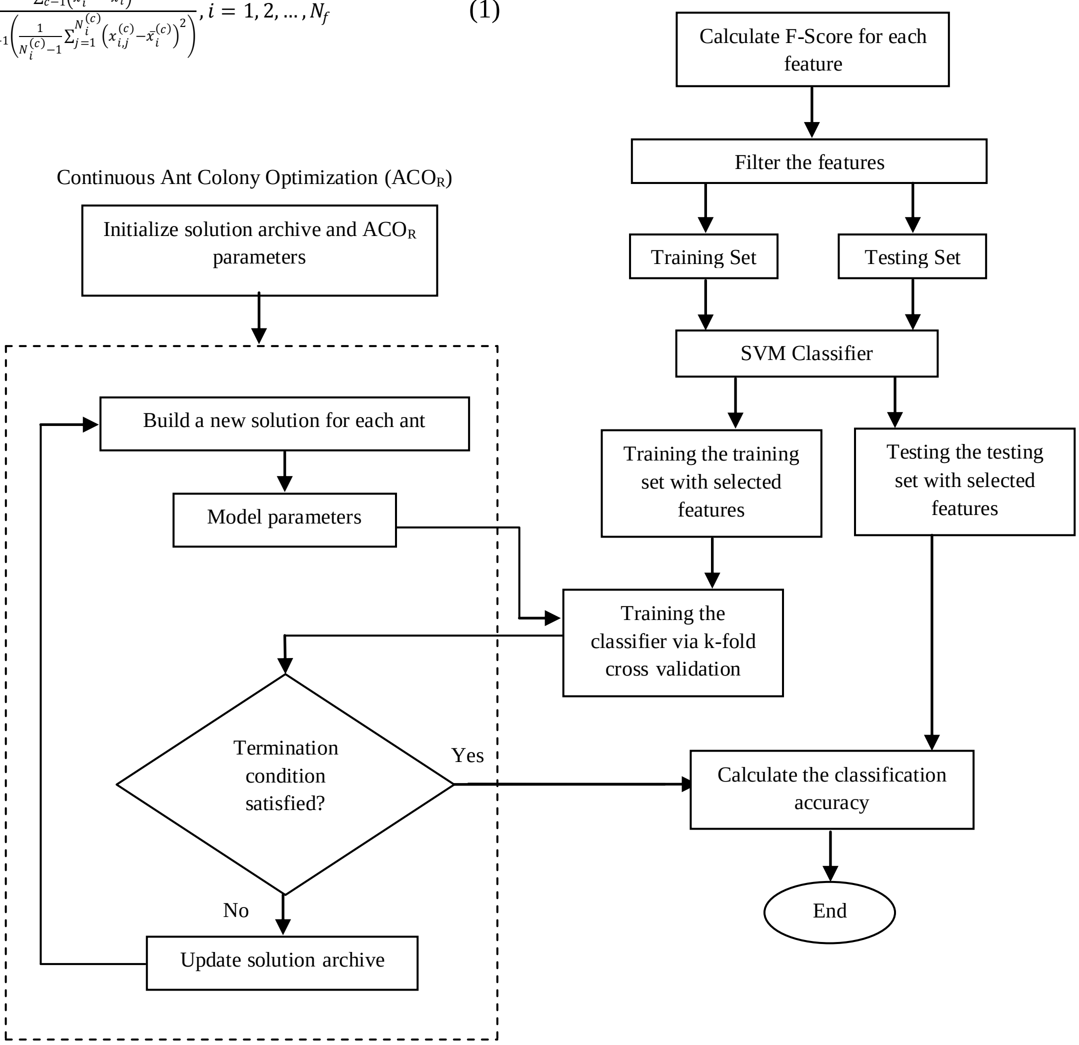

Fig. 1 Hybridize ACO_R and SVM Algorithm

Once this step is completed, the sampling procedure is made through two phases. Phase one involves choosing one of the weight vectors as follows:

$$p_l = \frac{w_l}{\sum_{r=1}^{k} w_r} \quad (3)$$

The Second phase involves sampling selecting w via a random number generator that is able to generate random numbers according to a parameterized normal distribution. This initializing constructs the transition probabilities. Like the solution archives, some important system parameters must be initialized as follows: the number of ants = 2, q = 0.1, and number of runs = 10, C range is ϵ $[2^{-1}, 2^{12}]$ and $\gamma \epsilon$ $[2^{-12}, 2^{2}]$.

The third step relates to solution construction where each ant builds its own solution. This solution is a combination of C and γ. In order to construct the solution, two transition probabilities with various solutions archives are needed. These transitions are computed according to Eq. 2 and Eq. 3.

A classifier model is constructed in step four. Solution is generated by each ant and is evaluated based on the classification accuracy obtained by the SVM model utilizing k-fold Cross Validation (CV) with the training set. In k-fold CV, the training data group is portioned into k subgroups, and the holdout approach is repeated k times. One of the k sub-groups is utilized as the test set and the remaining k-1 sub-groups are combined to construct the training group. The average mistakes along with all the k trails are calculated. CV accuracy is calculated as follows:

$$\mathrm{CV}_{\mathrm{accuracy}} = \frac{\sum_i \mathrm{test\_accuracy}}{\mathrm{k}}, i = 1, 2, \ldots, k \quad (4)$$

Test_accuracy evaluates the percentage of samples that are classified in the correct way to determine k-folds and is computed as follows:

$$\mathrm{Test\,Accuracy} = \frac{\mathrm{no.of\ correctly\ \ predicted\ \ data}}{\mathrm{total\ \ testing\ \ data}} * 100\% \quad (5)$$

The benefits of using CV are (1) each of the test groups are independent and (2) the dependent outcomes can be enhanced [28].

The final step is related to updating solution archives. This modification is completed by appending the newly generated group solutions that gave the best classification accuracy to solution archive and then deleting the exact number of worst solutions. This ensures the size of solution archive does not change. This procedure guarantees that only good solutions are stored in the archive, and it will efficiently influence the ants in the seek process.

IV. EXPERIMENTAL RESULTS

Seven datasets were used in evaluating the proposed $\mathrm{ACO_R}$-SVM algorithm. The datasets are Australian, Pima-Indian Diabetes, Heart, Ionosphere, German, Sonar, Splice datasets, available from UCI Repository of Machine Learning Databases [29]. The summary of these datasets is presented in Table I.

TABLE I SUMMARIZATION OF UCI'S DATASETS REPOSITORY

Dataset	No. of Instances	No. of Features
Australian	690	14
Pima-Indian Diabetes	760	8
Heart	270	13
Ionosphere	351	34
German	1000	24
Sonar	208	60
Splice	1000	60

All input variables were scaled during the data pre-processing phase to avoid features with higher numerical ranges from dominating those in lower numerical ranges and to minimize complexity of computation. The following formula was used to linearly scale each feature to [0, 1] range:

$$\bar{x} = \frac{x - min_i}{max_i - min_i} \quad (6)$$

where x is the original value, $\bar{x}$ is the scaled value, and max_i and min_i are the maximum and minimum values of feature i, respectively [28].

Each dataset was randomly re-arranged and divided into ten approximately equal sized subsets, one subset is a testing set and the remaining are training sets and repeated ten times. The performance of the proposed $\mathrm{ACO_R}$-SVM was compared with the grid search approach [28] and [30] which was considered as the basic approach to optimize SVM parameters.

C programming language was used to implement $\mathrm{ACO_R}$-SVM. Experiments were performed on an Intel(R) Core (TM) 2 Duo CPU T5750, running at 2.00 $\mathrm{GH_Z}$ with 4.00 GB RAM and 32-bit operating system.

Table II shows the optimal values for C and γ that were produced by the proposed algorithm and these values were used to produce the classification accuracy depicted in Table III. The average number of selected features and time to classify pattern of the proposed $\mathrm{ACO_R}$-SVM algorithm together with grid search results [28] and [30]. The proposed approach classifies patterns with higher accuracy compared to grid search for all seven datasets. The average percentage increased in accuracy for all datasets is approximately 7.85. This is because the integration of $\mathrm{ACO_R}$ with SVM, $\mathrm{ACO_R}$ as an optimization approach improves SVM classification accuracy through optimizing its parameters which are the regularization parameter C and gamma (γ) of RBF kernel function.

TABLE II OPTIMAL VALUE FOR C AND γ

Dataset	C	γ
Australian	473.39	0.63
Pima-Indian Diabetes	2464.50	2.42
Heart	372.50	0.36
Ionosphere	633.44	0.60
German	109.50	0.11
Sonar	291.27	0.23
Splice	244.28	0.19

For each iteration, $\mathrm{ACO_R}$ generates SVM parameters' values and introduces it to SVM and SVM uses these values to classify patterns. The proposed algorithm stops if the classification accuracy or maximum number of iteration satisfies user specification, otherwise, $\mathrm{ACO_R}$ searches for other optimal values for SVM parameters to work with.

Table IV shows the best features chosen by filter F-score technique to generate features subsets to be introduced to SVM. All features displayed in this table are important based on their threshold values. The reason for using filter F-score technique to select features subset was because RBF would fail for large numbers of features [31]. Table IV shows that the biggest reduction in number of features is 71% for the Australian dataset while the smallest feature reduction is 47% for the Ionosphere dataset.

TABLE III EXPEREMENT RESULTS OF THE PROPSED ACO_R-SVM AND GRID SEARCH

Dataset	Number of Features	ACO_R-SVM			Grid Search
		Classification accuracy (%)	Average number of selected features	Time (sec.)	Classification accuracy (%)
Australian	14	96.14	3.2	511	84.7
Pima-Indian Diabetes	8	87.79	2.3	162	77.3
Heart	13	89.99	5.9	276	83.7
Ionosphere	34	89.87	8.6	343	89.44
German	24	94.00	5.6	3718	76
Sonar	60	90.41	20.2	176	87
Splice	60	96.22	6.7	792	91.31

TABLE IV FREQUENSIES OF SELECTED FEATURES

Data																					
Australian																					
Feature#	5	7	8	9																	
Frequencies	10	2	10	10																	
Pima-Indian Diabetes																					
Feature#	2	5	6	8																	
Frequencies	10	1	3	9																	
Heart																					
Feature#	3	8	9	10	12	13															
Frequencies	7	10	10	10	10	10															
Ionosphere																					
Feature#	3	4	5	6	7	8	9	12	14	16	21	22	23	24	25	29	31	33			
Frequencies	10	4	10	3	10	5	8	5	4	3	2	1	1	1	1	7	7	4			
German																					
Feature#	1	2	3	5	6	7	12														
Frequencies	10	10	10	9	9	2	6														
Sonar																					
Feature#	1	2	3	4	5	8	9	10	11	12	13	20	21	22	31	36	43	44	45	46	47
Frequencies	10	9	7	9	7	3	10	10	10	10	8	1	2	2	1	3	6	8	10	10	10
Feature#	48	49	50	51	52	54	58														
Frequencies	9	10	5	10	10	7	4														
Splice																					
Feature#	15	16	18	19	20	21	22	23	26	29	30	31	32	33	34	45	49	51	54	58	60
Frequencies	1	5	5	1	5	2	5	2	2	5	5	1	4	5	5	4	1	5	1	1	1

V. CONCLUSIONS AND FUTURE WORKS

This study investigated a hybrid ACO_R and SVM technique to obtain optimal model parameters. Experimental results on seven public UCI datasets showed promising performance in terms of test accuracy and training time. Possible extensions can focus on the area where ACO_R-SVM can simultaneously optimize both SVM parameters and features subset using mixed-variable ACO ($ACO_{R\text{-}MV}$). Incremental Continuous ACO ($IACO_R$) may also be a good alternative for optimizing the classifier parameter values. Other kernel parameters besides RBF, application to other SVM variants and multiclass data are considered possible future work in this area.

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Higher Education Malaysia for funding this study under Fundamental Research Grant Scheme, S/O code 12377 and RIMC, Universiti Utara Malaysia, Kedah for the administration of this study.

REFERENCES

[1] H. Orkcu and H. Bal, “Comparing performance of back propagation and genetic algorithms in the data classification,” *Expert Systems with Applications*, vol. 38, no. 4, pp. 3703-3709, Apr. 2011.

[2] V. Tseng and C. Lee, “Effective temporal data classification by integrating sequential pattern mining and probabilistic Induction,” *Expert Systems with Applications*, vol. 36, no. 5, pp. 9524-9532, Jul. 2009.

[3] R. Sivagaminathan and S. Ramakrishnan, “A hybrid approach for feature subset selection using neural networks and ant colony optimization,” *Expert Systems with Applications*, vol. 33, no. 1, pp. 49-60, Jul. 2007.

[4] W. Liu and D. Zhang, “Feature subset selection based on improved discrete particle swarm and support vector machine algorithm,” *Information Engineering and Computer Science*, Wuhan, China, 2009, pp. 586-589.

[5] H. Zhang and H. Mao, “Feature selection for the stored-grain insects based on PSO and SVM,” *Knowledge Discovery and Data Mining*, Moscow, Russia, 2009, pp. 586-589.

[6] Y. Ye, L. Chen, D. Wang, T. Li, Q. Jiang, and M. Zhao, “SBMDS: an interpretable string based malware detection system using SVM ensemble with bagging,” *Computer Virology*, vol. 5, no. 4, pp. 283-293, Nov. 2009.

[7] S. Moustakidis and J. Theocharis, “SVM-FuzCoC: A novel SVM-based feature selection method using a fuzzy complementary criterion,” *Pattern Recognition*, 43, no. 11, pp. 3712-3729, Nov. 2010.

[8] H. Zhang, M. Xiang, C. Ma, Q. Huang, W. Li, W. Xie, Y. Wei, and Yang, S., “Three-class classification models of LogS and LogP derived by using GA-CG-SVM approach,” *Molecular Diversity*, vol. 13, no. 2, pp. 261-268, May 2009.

[9] C. Huang and C. Wang, “A GA-based feature selection and parameters optimization for support vector machines,” *Expert Systems with Applications*, vol. 31, no. 2, pp. 231-240, Aug. 2006.

[10] N. Ayat, M. Cheriet, and C. Suen, C., “Automatic model selection for the optimization of SVM kernels,” *Pattern Recognition*, vol. 38, no. 10, pp. 1733-1745, Oct. 2005
[11] X. Zhang, X. Chen, and Z. He, “An ACO-based algorithm for parameter optimization of support vector machines,” *Expert Systems with Applications*, vol. 37, no. 9, pp. 6618-6628, Sep. 2010.
[12] S. Abbas and M. Arif, “Electric load forecasting using support vector machines optimized by genetic algorithm,” *INMIC*, Islamabad, India, 2006, pp. 395-399.
[13] L. Saini, S. Aggarwal, and A. Kumar, “Parameter optimization using genetic algorithm for support vector machine-based price-forecasting model in national electricity market,” *IET Generation, Transmission & Distribution*, vol. 4, no. 1, pp. 36-49, 2010.
[14] K. Socha and C. Blum, ”Ant colony optimization,” in Metaheuristic procedures for training neutral networks, vol. 36, E. Alba and R. Martí, Eds. US : Springer, 2006, pp. 153-180.
[15] C. Blum, “Ant colony optimization: introduction and recent trends,” *Physics of Life Reviews*, vol. 2, no. 4, pp. 353-373, Dec. 2005.
[16] K. Socha and M. Dorigo, “Ant colony optimization for continuous domain,” *European Journal of Operational Research*, vol. 185, no. 3, pp. 1155-1173, Mar. 2008.
[17] K. Socha, “Ant colony optimization for continuous and mixed-variables domain,” Ph.D. dissertation, Universite’ Libre de Bruxelles, 2008.
[18] F. Imbault and K. Lebart, “A stochastic optimization approach for parameter tuning of support vector machines,” *Proc. of the IEEE International Conference on Pattern Recognition (ICPR’04), 2004,* pp. 597-600, Washington, USA.
[19] H. Frohlich and A. Zell A., “Efficient parameter selection for support vector machines in classification and regression via model-based optimization,” *Proc. of the IEEE International Joint Conference on Neural Network (IJCNN’05), 2005,* pp. 1431-1436 , Montreal, Canada, pp. 597-60.
[20] M. Adankan, M., Cheriet, and N. Ayat, “Optimizing resources in model selection for support vector machines,” *Proc. of the IEEE International Joint Conference (IJCNN ‘05), 2005,* pp. 925-930, Montreal, Canada.
[21] Y. Dong, Z. Xia, M. Tu, and G. Xing, “An optimization method for selecting parameters in support vector machines,” *IEEE International Conference on Machine Learning and Applications (ICMLA’06) 2007*, pp. 1-6, OH, USA.
[22] Y. Zhang, “Evolutionary computation based automatic SVM model selection,” *IEEE International Conference on Natural Computation (ICNC’08) 2008*, pp. 66-70, Jinan, China.
[23] X. Zhang, X. Chen, Z. Zhang, and Z. He, “A grid–based ACO algorithm for parameters optimization in support vector machines,” *IEEE International conference on Granular Computing (GrC 2008), 2008*, pp. 805-808, Hangzhou, China.
[24] X. Fang and T. Bai, “Share price prediction using wavelet transform and ant colony algorithm for parameters optimization in SVM,” Congr. of the IEEE International Congress on Intelligent Systems (GCIS ’09), 2009, pp. 288-292, Xiamen, China.
[25] Y. Dong, Z. Xia, M. Tu, and G. Xing, “An optimization method for selecting parameters in support vector machines,” *IEEE International Conference on Machine Learning and Applications (ICMLA 2007), 2007*, pp. 1-6. Cincinnati, OH.
[26] H. Frohlich and A. Zell, “Efficient parameter selection for support vector machines in classification and regression via model-based optimization,” *Proc. of the IEEE International Joint Conference on Neural Networks (IJCNN ’05), 2005*, pp. 1431-1436, Montreal, Canada.
[27] N. Lu, J. Zhou, Y. He, and Y. Liu, “Particle swarm optimization for support vector machine model,” *IEEE International Conference on Intelligent Computation Technology and Automation (ICICTA 2009), 2009*, pp. 283-286, Changsha, China.
[28] C. Huang, “ACO-based hybrid classification system with feature subset selection and model parameters optimization,” *Neurocomputing*, vol. 73, no. 1-3, pp. 438-448, Dec. 2009.
[29] *UCI Repository of machine learning databases, Department of Information and Computer Science, University of California, Irvine, CA, <http://www.ics.uci.edu/mlearn/MLRepository>, 2012.*
[30] S. Ding and S. Li, “Clonal Selection Algorithm for Feature Selection and Parameters Optimization of Support Vector Machines,” *Symp. of the IEEE on knowledge acquisition and modeling (KAM ’09)*, 2009, Wuhan, China, pp. 17-20.
[31] C. Huang and J. Dun, “A Distributed PSO-SVM hybrid system with feature selection and parameter optimization”. *Applied Soft Computing*, vol. 8, no. 4, pp. 1381-1391, Sep. 2008.