
Research on parameters optimisation of SVM based on swarm intelligence

Shifei Ding*

School of Computer Science and Technology,
China University of Mining and Technology,
Xuzhou, 221116, China

and

Key Laboratory of Intelligent Information Processing,
Institute of Computing Technology,
Chinese Academy of Science,
Beijing 100190, China

E-mail: dingsf@cumt.edu.cn

*Corresponding author

Huajuan Huang

School of Computer Science and Technology,
China University of Mining and Technology,
Xuzhou, 221116, China

and

Guangxi Key Laboratory of Hybrid Computation and
IC Design Analysis,

Nanning, 530006, China

E-mail: hhj-025@163.com

Abstract: Support vector machine (SVM) is a new machine learning method based on statistical learning theory, which has become a hot research topic in the field of machine learning because of its excellent performance. However, the performance of SVM is very sensitive to its parameters. At present, swarm intelligence is the most common method to optimise the parameters of SVM. In this paper, the research on parameters optimisation of SVM based on swarm intelligence algorithms is reviewed. Firstly, we briefly introduce the theoretical basis of SVM. Secondly, we describe the latest progress of parameters optimisation of SVM based on swarm intelligence in recent years. Finally, we point out the research and development prospects of this kind of method.

Keywords: support vector machine; SVM; parameter optimisation; swarm intelligence; machine learning.

Reference to this paper should be made as follows: Ding, S. and Huang, H. (2014) 'Research on parameters optimisation of SVM based on swarm intelligence', *Int. J. Collaborative Intelligence*, Vol. 1, No. 1, pp.4–17.

Biographical notes: Shifei Ding received his BSc and MSc from Qufu Normal University in 1987 and 1998 respectively. He received his PhD from Shandong University of Science and Technology in 2004. He received postdoctoral degree from Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences in 2006. And now, he

works in China University of Mining and Technology as a Professor and PhD Supervisor. His research interests include intelligent information processing, pattern recognition, machine learning, data mining, and granular computing, etc. He has published three books, and more than 150 research papers in journals and international conferences.

Huajuan Huang received her BSc and MSc in Applied Computer Technology from Guangxi University for Nationalities, Guangxi, China, in 2006 and 2009 respectively. Since 2011, she has been a PhD candidate in Applied Computer Technology from the China University of Mining and Technology, Xuzhou, China. Her current research interests include data mining, pattern recognition and computational intelligence.

1 Introduction

Support vector machine (SVM), introduced by Vapnik and co-workers, is an excellent kernel-based tool for solving classification and regression problems (Vapnik, 1995; Ding and Qi, 2012; Huang et al., 2012, 2013a). Unlike other machine learning methods, such as artificial neural network (Xu et al., 2012; Ding et al., 2012), the mathematical theory of SVM is based on the statistical learning theory, which has many advantages in solving small samples, non-linear and high dimensional pattern recognition problems. Therefore, SVM can have better generalisation ability. Within a few years after its introduction, SVM has played excellent performance on many real-world predictive data mining applications such as text categorisation (Pan et al., 2013), time series prediction (Chen and Zhi, 2012), pattern recognition (Morales et al., 2013) and image processing (Wu, 2012), etc.

One of the main challenges for SVM is the choice of parameters. As we know, the learning performance and generalisation ability of SVM is very dependent on its parameters. If the choice is not reasonable, it will be very difficult to approach superiorly. At present, the grid search method is the commonly used parameter selection method for SVM. However, the search time of this method is too long, especially in dealing with the large dataset. In order to solve this problem, many parameter selection methods of SVM have been proposed. Among them, swarm intelligence is the most common approach.

As a new evolutionary computation technology, swarm intelligence has been a research hotspot (Lin et al., 2013; Nezhad et al., 2013; Verwaeren et al., 2013). The basic idea of swarm intelligence is to imitate the solving problem ways of gregarious biological entities by simulating their behaviours such as foraging, resisting enemies and nesting. With the development of swarm intelligence, at present, it has many kinds of algorithms, such as genetic algorithm (GA), particle swarm optimisation (PSO) algorithm, ant colony optimisation (ACO), artificial fish swarm algorithm (AFSA), differential evolution (DE), artificial glowworm algorithm (AGA), invasive weed optimisation (IWO) algorithm and shuffled frog leaping algorithm (SFLA), etc. (Drezner and Misevicius, 2013; Wong and Ngan, 2013; Bai et al., 2013; Zhou and Huang, 2012; Brest et al., 2013; Zhou and Zhou, 2013; Ahmadi and Mojallali, 2012; Vatani et al., 2013). So far, these swarm intelligence algorithms have been successfully applied in many fields, which prove that swarm intelligence is a kind of new method which can effectively solve most global optimisation problems. This is because swarm intelligence is easy to implement. Furthermore, its

output only needs the values of objective function, without gradient information. In particular, swarm intelligence has the characteristics of parallel and distributed, which provides technical guarantee for processing in large database data.

In view of the good optimisation performance of swarm intelligence, how to use it to optimise the parameters of SVM has become a hot topic (Huang and Ding, 2013; Huang et al., 2013b). In this paper, we mainly describe the latest progress of parameters selection methods of SVM based on swarm intelligence. Finally, we point out the research and development prospects of this kind of method.

This paper is organised as follows: In Section 2, the basic theory of SVM is briefly introduced. In Section 3, we describe the latest progress of parameters selection methods of SVM based on swarm intelligence. In Section 4, we provide summary and prospects of this kind of method.

2 Basic theory of SVM

2.1 Theoretical model of SVM

SVM is a new data mining method based on statistical learning theory and its target is to find the optimal separating hyperplane which can meet the classification requirements. The mathematical model of SVM is shown as follows.

Consider the classification problem with the training set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$, where $x_i \in R^n$ are inputs and $y_i \in \{+1, -1\}$ are the corresponding outputs.

Linear SVM searches for a separating hyperplane

$$f(x) = w^T x + b = 0, \quad (1)$$

where $w \in R^n$ and $b \in R$. To measure the empirical risk, the soft margin loss function

$\sum_{i=1}^l \max(0, 1 - y_i (w^T x_i + b))$ is used. By introducing the regularisation term $1/2 \|w\|^2$ and the slack variable $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_l)$, the primal problem of SVM can be expressed as

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \zeta_i \\ \text{s.t.} \quad & y_i (w^T x_i + b) \geq 1 - \zeta_i, \zeta_i \geq 0, i = 1, 2, \dots, l \end{aligned} \quad (2)$$

where $C > 0$ is a penalty parameter.

In order to solve this constraint optimisation problem, the Lagrangian function is introduced:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i ((w^T x_i) + b) - 1) \quad (3)$$

where $\alpha_i > 0$ are the Lagrange multiplier.

By considering the Karush-Kuhn-Tucker (KKT) conditions for the Lagrangian function (3), we can obtain the corresponding dual QPP:

$$\begin{aligned}
 \max \quad & Q(\alpha) = \sum_{j=1}^l \alpha_j - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \\
 \text{s.t.} \quad & \sum_{j=1}^l \alpha_j y_j = 0, j = 1, 2, \dots, l \\
 & \alpha_j \geq 0, j = 1, 2, \dots, l
 \end{aligned} \tag{4}$$

After obtaining the optimal solutions $\alpha^* = (\alpha_1^*, \dots, \alpha_l^*)^T$, we can get the optimal classification function as follows:

$$f(x) = \text{sgn} \left\{ (w^* \cdot x) + b^* \right\} = \text{sgn} \left\{ \left(\sum_{j=1}^l \alpha_j^* y_j (x_j \cdot x_i) \right) + b^* \right\}, x \in R^n \tag{5}$$

For the non-linear case, we need mapping the data samples from the original input space R^n into potentially higher dimensional feature space denoted by χ where linear SVM algorithm can be used. So we can consider the non-linear map $\varphi: R^n \rightarrow \chi$. In general, under the Mercer theorem, it is possible to use some kernel $K(u^T, v)$ to represent the inner product in χ , i.e., $K(u^T, v) = \varphi(u)^T \varphi(v)$. Based on the above idea, the SVM formulation for the non-linear case is of the form

$$\begin{aligned}
 \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\
 \text{s.t.} \quad & y_i (w^T K(x_i, x) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, l
 \end{aligned} \tag{6}$$

By using KKT optimality conditions for constrained optimisation problems, we obtain the Wolfe dual optimisation problem of (6) as

$$\begin{aligned}
 \max \quad & Q(\alpha) = \sum_{j=1}^l \alpha_j - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \\
 \text{s.t.} \quad & \sum_{j=1}^l \alpha_j y_j = 0, j = 1, 2, \dots, l \\
 & \alpha_j \geq 0, j = 1, 2, \dots, l,
 \end{aligned} \tag{7}$$

After SVM have been trained, it can be used to predict of a new coming data by using the following relation

$$f(x) = \text{sgn} \left\{ (w^* \cdot K(x, x_i)) + b^* \right\} = \text{sgn} \left\{ \left(\sum_{j=1}^l \alpha_j^* y_j K(x_j \cdot x_i) \right) + b^* \right\}, x \in R^n \tag{8}$$

2.2 Analysis the parameters of SVM

2.2.1 The penalty parameter

The role of penalty parameter C is to adjust the ratio between the confidence range with the experience risk in the defining feature, so that the generalisation ability of SVM can

achieve the best state. The value of C smaller expresses the punishment on empirical error smaller. Do it this way, the complexity of SVM is smaller, but its fault tolerant ability is worse. The value of C is greater, the data fitting degree is higher, but its generalisation capacity will be reduced. From the above analysis, we can know that the parameter selection is very important for SVM.

2.2.2 The kernel parameter

By introducing the kernel function, SVM can achieve the linear classification in the dimensional feature for the non-linear problems. Therefore, kernel function takes an important role in the non-linear SVM. At present, the most commonly used kernel functions in SVM are as follows:

- The linear kernel function:

$$K(x, x_i) = x, x_i \quad (9)$$

- The polynomial kernel function:

$$K(x, x_i) = (\gamma(x, x_i) + 1)^d, \gamma > 0 \quad (10)$$

- The gauss kernel function:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (11)$$

- The sigmoid kernel function:

$$K(x, x_i) = \tanh(p_1(x, x_i) + p_2) \quad (12)$$

Similar to the penalty parameter, the kernel parameter also has a significant effect on the performance of SVM.

3 Parameters optimisation of SVM based on swarm intelligence

As swarm intelligence algorithms, GA and PSO algorithm are the earliest proposed. Furthermore, the mathematical theory of two algorithms is the most perfect between the swarm intelligence algorithms. Therefore, at present, the researchers mainly use GA and PSO to optimise the parameters of SVM.

3.1 Parameters optimisation of SVM based on GA

3.1.1 The basic theory of GA

GA is an adaptive method, which is considered to be the most typical swarm intelligence algorithm. In GA, a candidate solution for a specific problem is called an individual and consists of a linear list of genes. Each individual represents a point in the search space, and hence a possible solution to the problem. Each individual is decided by an evaluating mechanism to obtain its fitness value. According to this fitness value and undergoing genetic operators, a new population is generated iteratively with each successive

population referred to as a generation. Then GA employs mutation, crossover and selection as the primary operators to manipulate the genetic composition of a population. Reproduction is a process by which the most highly rated individuals in the current generation are reproduced in the new generation. When GA reached the stop condition, please stop and output the best value.

3.1.2 Research progress on parameters optimisation of SVM based on GA

For linear SVM, there is only one parameter, i.e., the penalty parameter, to be optimised. For the non-linear case, the parameters of SVM contain the penalty parameter and the kernel parameters. In order to get the reasonable parameter values, GA takes a sample of possible solutions and employs mutation, crossover and selection as the basic operators for optimisation. At present, researchers have completed many papers about this issue. We will discuss in detail the progress on parameters optimisation of SVM based on GA as follows.

In 2004, Huang et al. (2004) developed a new improved SVM algorithm, called QGA-SVM. In their paper, the quasi genetic algorithm (QGA) strategy was integrated in the SVM learning procedure to further optimise and accelerate the training procedure. The experiments on some datasets proved its advantages, especially for the multi-class pattern classification with unbalanced classes. In 2008, in order to solve the parameter selection problems, Zhou et al. (2008) proposed a new culture genetic algorithm (CGA) to optimise the parameters of SVM. Through embedding GA into the cultural algorithm framework, this CGA algorithm constructed the population space and the knowledge space based on GA. The two spaces evolved independently, at the same time, the population space continuously transferred the evolving knowledge to the knowledge space, and then the knowledge space to achieve global optimisation. Experimental results showed that CGA-SVM performed good prediction accuracy and generalisation, implying that the hybrid of CGA with traditional SVM can serve as a promising alternative for predicting share price. In 2011, Ma et al. (2011) established the mapping between the natural frequency of a rolling bearing rotor and the various parameters, which reduced the rotor structure for the study similar to the natural frequency of the calculation of the workload greatly. Based on this idea, they proposed a model called GA-SVM. Using the model to identify the natural frequency of bearing rotor under different parameters, experiments showed that projections are good agreement with the experimental data. According to the principles of a GA and SVM, Yao et al. (2011) developed a GA-SVM programme and applied it to human cytochrome P450s (CYP450s). The final predictive model had satisfactory performance, with the prediction accuracy of 61% and cross-validation accuracy of 73%. The results indicated that the GA-SVM programme was a powerful tool in optimising mutation predictive models of nsSNPs of human CYP450s. In 2011, Zhu et al. (2011b) used GA-SVM to construct a reliable computational model for the classification of agonists and antagonists of 5-HT1A receptor. This model was successfully developed to effectively distinguish agonists and antagonists among the ligands of the 5-HT1A receptor. To our knowledge, this was the first effort for the classification of 5-HT1A receptor agonists and antagonists based on a diverse dataset. In the same year, based on empirical mode decomposition (EMD) method and SVM, a new method for the fault diagnosis of high voltage circuit breaker (CB) was proposed in Huang et al. (2011). The purpose of this study was to develop a genetic algorithm-based support vector machine (GA-SVM) model that can determine the

optimal parameters of SVM with the highest accuracy and generalisation ability. Finally, the experimental results indicated that the classification accuracy of this GA-SVM approach was more superior than that of the artificial neural network and the SVM which had constant and manually extracted parameters. As we know, county Industry-Population-Knowledge matching degree analysis and prediction played an important role in region economic development and improved the transformation of national economic growth pattern. Jing et al. (2012) proved that GA-SVM was an effective method for county industry-population-knowledge matching degree classification and prediction. In 2012, Yang et al. (2012) used GA-SVM to deal with the estimation of amplification effect of mining-induced blast vibration on surrounding structures. The results showed that this method was effective. In the same year, Liu and Fu (2012) used EGA, which was based on elite survival strategy, to optimise the parameters of SVM. Iris dataset and one hundred pieces of news reports in Chinese news were chosen to compare EGA-SVM, GA-SVM and traditional SVM. The results of numerical experiments showed that EGA-SVM could improve classification performance effectively than the other algorithms. This text classification algorithm could be extended easily to apply to literatures in the field of electrical engineering.

In the existing literatures, we find that GA is not only used to optimise SVM, but also used to optimise least square support vector machine (LSSVM). In 2010, Liang et al. (2010) proposed an algorithm of multiple information fusion based on GA-LSSVM. The results showed that the model and algorithm had certain superiority in measuring precision and are easy to be promoted. In order to improve the performance of GA-LSSVM, Wang et al. (2011a) proposed a new fuzzy LSSVM classifier based on chaos GA, term as FLS-SVMWBCGA. Experimental results showed that FLS-SVMWBCGA was effective in improving the prediction accuracy of the classification problems with noises or outliers. In the same year, a self-adaptive binary GA was introduced to optimise the hyper-parameters of LSSVM, and the individual fitness values in GA were determined by cross-validation in Chen (2011). Then they used the proposed algorithm to build an hourly water consumption forecasting model. Case study showed that the proposed model had higher computing speed and better estimating performance than the traditional LSSVM. In 2012, Mustafa and Sulaiman (2012) presented a new method for reactive power tracing in a pool-based power system by introducing the hybrid GA and least squares support vector machine (GA-LSSVM). The idea was to use GA to obtain the optimal values of regularisation parameter, gamma, and kernel radial basis function (RBF) parameter, and adopt a supervised learning approach to train the LSSVM model. The 25-bus equivalent system of southern Malaysia was used to illustrate the effectiveness of the proposed GA-LSSVM model compared to PSM and artificial neural network.

3.2 Parameters optimisation of SVM based on PSO

3.2.1 The basic theory of PSO

PSO is in principle a much simpler algorithm. It operates on the principle that each solution can be represented as a particle denoted by x_i in a swarm. A population of particles is randomly generated initially. Then a swarm of particles moves through the problem space, with the moving velocity of each particle represented by a velocity vector v_i . Like GA, PSO must also have a fitness evaluation function that takes the particle's

position and assigns to it a fitness value. Each particle keeps track of its own best position, which is associated with the best fitness it has achieved so far in a vector p_i . Furthermore, the best position among all the particles obtained so far in the population is kept track of as p_g . In addition to this global version, another local version of PSO keeps track of the best position among all the topological neighbours of a particle.

At each time step t , by using the individual best position, $p_i(t)$, and global best position, $p_g(t)$, a new velocity for particle i is updated by

$$v_i(t+1) = v_i(t) + c_1\phi_1(p_i(t) - x_i(t)) + c_2\phi_2(p_g(t) - x_i(t)) \quad (13)$$

where c_1 and c_2 are positive constants, ϕ_1 and ϕ_2 are uniformly distributed random numbers in $[0, 1]$. The term v_i is limited to the range $\pm v_{\max}$. If the velocity violates this limit, it is set at its proper limit. Changing velocity this way enables the particle i to search around its individual best position, p_i , and global best position, p_g . Based on the updated velocities, each particle changes its position according to the following:

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (14)$$

Based on (13) and (14), the population of particles tends to cluster together with each particle moving in a random direction.

3.2.2 Research progress on parameters optimisation of SVM with PSO

With self-organising, adaptive, self-learning and the nature of parallelism, etc., PSO has been widely used in parameter estimation. So far, there are many papers about using PSO or its improved algorithm to optimise the parameters of SVM. We will track the latest progress of SVM optimised by PSO as follows. In Wang and Wang (2012), in order to overcome the difficulty in selecting parameters of SVM when modelling the PT fuel system fault diagnosis, SVM optimised by PSO algorithm was proposed. The result of experiment confirmed the validity of this method through comparison of the BP-NN, SVM. Literature (Ren and Zhou, 2011) used PSO-SVM model to forecast the traffic safety. The experimental results showed that traffic safety forecasting by PSO-SVM was better than that by BP neural network. At the same time, in literature (Jin et al., 2011), the authors used PSO-SVM to solve the face recognition problem. The results indicated that PSO-SVM had higher face recognition accuracy than normal SVM, BPNN. Therefore, PSO-SVM was well chosen in face recognition. In 2011, Mao et al. (2011) proposed a P2P protocol identification algorithm based on PSO-SVM. Experimental results of the P2P and non-P2P applications showed that this algorithm can identify the HTTP applications 100% and for a variety of P2P applications, the accuracy of classification can also reach to 95%. In order to further improve the performance of PSO-SVM, in Li and Zhao (2012), the authors firstly proposed chaotic PSO (CPSO), and then used CPSO to optimise SVM. Finally, CPSO-SVM model was used to predict the convergence deformation of the Xiakeng tunnel in China. The results indicated that the proposed method can describe the relationship of deformation time series well and was proved to be more efficient. In view of the perfect performance of CPSO-SVM, Zhao and Yin (2010) applied CPSO-SVM model to predict the ultimate bearing capacity of shallow foundations. Results indicated that the proposed methods can appropriately describe the relationship between ultimate bearing capacity and its affective factors, and make good

predictions. In literature (Yang et al., 2009), a chaotic adaptive particle swarm optimisation (CAPSO) method was applied to select parameters of SVM and genetic characteristics of the subset of choices (GA_FSS) was applied to reduce large dimensions and improve greatly the accuracy of classification. The experimental results on heart disease diagnosis problem showed that CAPSO-SVM classifier algorithm was effective and correct. Li et al. (2009) proposed an improved CPSO-SVM method and applied it to solve the face recognition problem. Experimental results on face database showed that the presented SVM method optimised by CPSO can achieve higher recognition performance. Wang et al. (2011b) presented a method that used an interval adaptive PSO to optimise the parameters of SVM. Then they applied this method to the intrusion detection systems. The experimental results showed that this method improved the classification accuracy by 9.7%, and the response time was shortened by 40.6%~56.5%.

In the spirit of GA-LSSVM, at present, many researchers have used PSO or its improved algorithms to optimise LSSVM. In order to enhance fault diagnosis precision for electric equipment, the LSSVM algorithm based on PSO was proposed in Zhang et al. (2009). The experiments showed that the PSO-LSSVM algorithm had better fault diagnosis ability than LSSVM. In order to improve the measurement accuracy of oil holdup of oil-water two phase flow, a modelling methodology of oil holdup based on PSO-LSSVM was proposed in Zhang and Zhang (2010). The experimental results indicated that the PSO-LSSVM model performed better than the other model in term of measurement accuracy and calculating speed. The average measurement errors were 0.93% in the range of 4% to 60% oil holdup. To deal with the difficulty in parameter adjustment and the low precision of the traditional heat-conduction model, the authors in Li et al. (2011) built a prediction model for the steel plate temperature, based on LSSVM which was optimised by the adaptive chaotic PSO. The experimental results indicated that the proposed prediction model had higher prediction accuracy than the tradition one. Aiming at the parameter optimisation problem in LSSVM, a hybrid QPSO algorithm for LSSVM parameter selection was proposed in Zhu et al. (2011a) to improve the learning performance and generalisation ability of LSSVM model. Then the production data from a purification process of zinc hydrometallurgy was used to test the model precision. The test results showed that the proposed model had better generalisation performance and higher precision.

3.2.3 Research progress on parameters optimisation of SVM with other swarm intelligence algorithms

In recent years, some new swarm intelligence algorithms with perfect performance are presented by researchers. However, the current research on these algorithms optimising the parameters of SVM is very scarce. In this section, we will introduce the latest progress of these algorithms optimising the parameters of SVM.

For the ACO, some researchers have used it to optimise the parameters of SVM. In 2010, Zhang et al. (2010a) adopted ACO algorithm to develop a novel ACO-SVM model to solve the parameters optimisation of SVM problem. The proposed algorithm was applied on some real world benchmark datasets to validate the feasibility and efficiency, which showed that the new ACO-SVM model can yield promising results. In the same year, Zhao et al. (2010b) presented a fault diagnosis method based on SVM with parameter optimisation by ant colony algorithm (ACO) to attain a desirable fault diagnosis result, which was performed on the locomotive roller bearings to validate its

feasibility and efficiency. The experiment found that the proposed algorithm of ant colony optimisation with support vector machine (ACO-SVM) can help one to obtain a good fault diagnosis result, which confirmed the advantage of the proposed ACO-SVM approach. In view of the good performance of ACO-SVM, in 2012, Cui and Huang (2012) used ACO-SVM to select the Chinese text feature. The experimental results showed that the proposed method was feasible and lead to a considerable increase of classification accuracy.

For DE algorithm, there also are some researches. In 2009, Jun and Jian (2009) employed a DE SVM model that hybridised the DE and SVMs to improve the classification accuracy for rainstorm forecasting. This optimisation mechanism combined the DE to optimise the SVM parameter setting. Based on the European Centre for Medium-Range Weather Forecasts (ECMWF), Japan and T213 precipitation data from 2003 to 2006, using DE-SVM, the 24 hour's storm models for five sub-regions in Hubei province were created, which have been used in the real-time running work from May to July in 2007. The results have shown the forecasting ability and reference value of the SVM method. In literature (Li and Cai, 2008), the authors applied DE to optimise support vector regression (SVR). Based on this idea, they proposed the DE-SVR model. Experimental results on several real-world datasets demonstrated that, comparing with the GA-based SVR and the grid search methods, the DE-SVR can search the optimal parameters much more rapidly with less training time to build the SVR model, and had the comparable prediction accuracy as grid search, even better than GA-based SVR.

For AFSA, in 2010, Zhang et al. (2010c) adopted AFSA to optimise SVM. The results showed that this proposed algorithm was fast than SVM, and provided a fast technical method for hydrological forecasting. Based on the perfect performance of AFSA-SVM, it was used in the fault prediction of condenser in naval vessel propulsion plant (Liu and Jiang, 2008). The experimental results showed that the proposed method can select the best fault features in shorter time and improve the performance of SVM.

For SFLA, in 2011, by introducing SFLA to solve the random-choice problem of the parameters of SVM, Song et al. (2011) SFLA-SVM algorithm. In order to verify the optimised effect, based on the training exempling energy demand from 1979 to 1999, the improved SVM was used to forecast the total energy demand of China in 2000 to 2009. The results showed that compared to the normal PSO-SVM, the time of the proposed algorithm increased by 51 s, and the precision of the proposed method increased by 2.34%. In order to improve the performance of SFLA-SVM, Zhang et al. (2011) used the improved SFLA to optimised SVM. Finally, they used the proposed model to solve the emotion recognition of practical speech problem. The test result indicated that the improvement mechanisms brought an outstanding improvement in the classification ability and provided a new method and idea for speech emotion recognition.

4 Summary and outlook

As we know, for SVM, its performance depends largely on its parameter values. In order to find a perfect parameter optimisation method, many researchers have done a lot of research on this issue. Swarm intelligence is a kind of good performance optimisation algorithm. At present, some researchers have used some swarm intelligence algorithms to optimise the parameters of SVM, whose results indicate that this kind of methods can

reach perfect performance. In this paper, we firstly introduce the basis theory of SVM and analyse its parameter selection problem. Then we introduce GA-SVM, PSO-SVM and SVM optimised by some new swarm intelligence algorithms. Meanwhile, we also describe the application of these models. From several models described above, we can see that, at present, only GA-SVM and PSO-SVM are more mature, while other models that SVM optimised by some new swarm intelligence algorithms are still in the initial stage of the study. Therefore, some new swarm intelligence algorithms with good optimisation ability such as AGA, IWO algorithm and so on could be used to optimise the parameters of SVM. Furthermore, those swarm intelligence algorithms also should be used to optimise the parameters of the improved SVM. However, at present, the literatures in this area are very scarce. So SVM optimised by swarm intelligence needs further development and refinement.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 61379101), the National Key Basic Research Program of China (No. 2013CB329502), and the Opening Foundation of Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis.

References

- Ahmadi, M. and Mojallali, H. (2012) 'Chaotic invasive weed optimization algorithm with application to parameter estimation of chaotic systems', *Chaos Solitons & Fractals*, Vol. 45, No. 9, pp.1108–1120.
- Bai, J., Yang, G.K. and Chen, Y.W. (2013) 'A model induced max-min ant colony optimization for asymmetric traveling salesman problem', *Applied Soft Computing*, Vol. 13, No. 3, pp.1365–1375.
- Brest, J., Korosec, P. and Silc, J. (2013) 'Differential evolution and differential ant-stigmergy on dynamic optimization problems', *International Journal of Systems Science*, Vol. 44, No. 4, pp.663–679.
- Chen, L. (2011) 'Genetic least squares support vector machine approach to hourly water consumption prediction', *Journal of Zhejiang University (Engineering Science)*, Vol. 45, No. 6, pp.1100–1103.
- Chen, Z. and Zhi, Z. (2012) 'Distributed customer behavior prediction using multiplex data: a collaborative MK-SVM approach', *Knowledge-Based Systems*, Vol. 35, pp.111–119.
- Cui, J.X. and Huang, X.X. (2012) 'Application of ACO-SVM in Chinese text feature selection', *Applied Mechanics and Materials*, Vols. 155–156, pp.770–775.
- Ding, S. and Qi, B. (2012) 'Research of granular support vector machine', *Artificial Intelligence Review*, Vol. 38, No. 1, pp.1–7.
- Ding, S., Xu, L., Su, C. and Jin, F. (2012) 'An optimizing method of RBF neural network based on genetic algorithm', *Neural Computing and Applications*, Vol. 21, No. 2, pp.333–336.
- Drezner, Z. and Misevicius, A. (2013) 'Enhancing the performance of hybrid genetic algorithms by differential improvement', *Computers & Operations Research*, Vol. 40, No. 4, pp.1038–1046.
- Huang, H. and Ding, S. (2013) 'A PSO-SVM method for time series prediction based on fuzzy information granulation', *INFORMATION: An International Interdisciplinary Journal*, Vol. 16, No. 6B, pp.4165–4174.

- Huang, H., Ding, S. and Shi, Z. (2013a) 'Primal least squares twin support vector regression', *Journal of Zhejiang University – SCIENCE C*, Vol. 14, No. 9, pp.722–732.
- Huang, H., Ding, S. and Wu, F. (2013b) 'Invasive weed optimization algorithm for optimizing the parameters of mixed kernel twin support vector machines', *Journal of Computers*, Vol. 8, No. 8, pp.2077–2084.
- Huang, H., Ding, S., Jin, F., Yu, J. and Han, Y. (2012) 'A novel granular support vector machine based on mixed kernel function', *International Journal of Digital Content Technology and its Applications*, Vol. 6, No. 20, pp.484–492.
- Huang, J., Hu, X.G. and Yang, F. (2011) 'Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker', *Measurement*, Vol. 44, No. 6, pp.1018–1027.
- Huang, J.T., Ma, L.H. and Qian, J.X. (2004) 'A new improved support vector machine: QGA-SVM', *ICCC2004: Proceedings of the 16th International Conference on Computer Communication*, Vol. 2, No. 1, pp.1749–1753.
- Jin, W., Zhang, J.Q. and Zhang, X.A. (2011) 'Face recognition method based on support vector machine and particle swarm optimization', *Expert Systems with Applications*, Vol. 38, No. 4, pp.4390–4393.
- Jing, Z., Duo, H. and Wei, Z. (2012) 'Matching degree prediction of region industry-Population-Knowledge system based on GA-SVM: county-level evidence from Guangzhou urban agglomeration in China', *Information – An International Interdisciplinary Journal*, Vol. 15, No. 11, pp.4797–4802.
- Jun, S. and Jian, L. (2009) 'A combination of differential evolution and support vector machine for rainstorm forecast', *2009 Third International Symposium on Intelligent Information Technology Application*, pp.392–395.
- Li, J., Wang, J. and Yang, L. (2011) 'Optimized prediction algorithm with adaptive chaos particle swarm optimization-least squares support vector machine for steel plate temperature prediction in heat treatment furnace', *Control Theory & Applications*, Vol. 28, No. 12, pp.1825–1830.
- Li, J.W. and Cai, Z.H. (2008) 'A novel automatic parameters optimization approach based on differential evolution for support vector regression', *3rd International Conference on Intelligence Computation and Applications*.
- Li, M., Sun, X.F. and Wang, Y. (2009) 'Application of improved CPSO-SVM approach in face recognition', *2009 International Conference on Artificial Intelligence and Computational Intelligence*, pp.119–123.
- Li, S.J. and Zhao, H.B. (2012) 'Deformation prediction of tunnel surrounding rock mass using CPSO-SVM model', *Journal of Central South University*, Vol. 19, No. 11, pp.3311–3319.
- Liang, Y., Wu, P. and Liu, Y. (2010) 'An algorithm of multiple information fusion based on GA-LSSVM', *Information And Electronic Engineering*, Vol. 8, No. 6, pp.697–701.
- Lin, R., Sir, M.Y. and Pasupathy, K.S. (2013) 'Multi-objective simulation optimization using data envelopment analysis and genetic algorithm: specific application to determining optimal resource levels in surgical services', *Omega-International Journal of Management Science*, Vol. 41, No. 5, pp.881–892.
- Liu, S. and Jiang, N. (2008) 'Research on parallel optimization simulation of fault prediction based on SVM', *Journal of System Simulation*, Vol. 20, No. 13, pp.3445–3448.
- Liu, X.Y. and Fu, H. (2012) 'A hybrid algorithm for text classification problem', *Przegląd Elektrotechniczny*, Vol. 88, No. 1, pp.8–11.
- Ma, T.R. and Qin, F.H. (2011) 'Identification of natural frequency of bearing rotor based on GA-SVM', *Manufacturing Science and Materials Engineering*, Vol. 2, No. 1, pp.27–33.
- Mao, L., Chen, X.S. and Wu, Z.G. (2011) 'A P2P protocol identification algorithm based on PSO-SVM', *Key Engineering Materials*, Vols. 467–469, pp.1528–1534.

- Moraes, R., Valiati, J.F., Gaviao, N. and Wilson, P. (2013) 'Document-level sentiment classification: an empirical comparison between SVM and ANN', *Expert Systems with Application*, Vol. 40, No. 2, pp.621–633.
- Mustafa, M.W. and Sulaiman, M.H. (2012) 'Reactive power tracing in pool-based power system utilising the hybrid genetic algorithm and least squares support vector machine', *IET Generation Transmission & Distribution*, Vol. 6, No. 2, pp.133–141.
- Nezhad, A.M., Shandiz, R.A. and Jahromi, A.E. (2013) 'A particle swarm-BFGS algorithm for nonlinear programming problems', *Computers & Operations Research*, Vol. 40, No. 4, pp.963–972.
- Pan, H., Zhu, Y. and Xia, L. (2013) 'Efficient and accurate face detection using heterogeneous feature descriptors and feature selection', *Computer Vision and Image Understanding*, Vol. 117, No. 1, pp.12–28.
- Ren, G. and Zhou, Z.P. (2011) 'Traffic safety forecasting method by particle swarm optimization and support vector machine', *Expert Systems with Applications*, Vol. 38, No. 8, pp.10420–10424.
- Song, X.H., Yang, S.D. and Liu, D. (2011) 'Improved support vector machine forecasting model by shuffled frog leaping algorithm and its application', *Journal of Central South University (Science and Technology)*, Vol. 42, No. 9, pp.2737–3740.
- Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.
- Vatani, A., Mehrpooya, M. and Pakraves, H. (2013) 'Modification of an industrial ethane recovery plant using mixed integer optimization and shuffled frog leaping algorithm', *Arabian Journal for Science and Engineering*, Vol. 38, No. 2, pp.439–455.
- Verwaeren, J., Scheerlinck, K. and De Baets, B. (2013) 'Countering the negative search bias of ant colony optimization in subset selection problems', *Computers & Operations Research*, Vol. 40, No. 4, pp.931–942.
- Wang, D. and Wang, X.Q. (2012) 'Fault diagnosis of PT fuel system based on particle swarm optimization-support vector machine', *Applied Mechanics and Materials*, Vols. 121–126, pp.2809–2813.
- Wang, H., Deng, F. and Chen, Z. (2011a) 'Fuzzy LS-SVM classifier based on chaos genetic algorithm and its application', *Journal of South China University of Technology (Natural Science Edition)*, Vol. 39, No. 5, pp.49–54.
- Wang, J., Jiang, N. and Zhang, Y. (2011b) 'SVM algorithm based on interval adaptive PSO and its application', *Journal of Zhengzhou University (Engineering Science)*, Vol. 32, No. 1, pp.75–79.
- Wong, T.C. and Ngan, S.C. (2013) 'A comparison of hybrid genetic algorithm and hybrid particle swarm optimization to minimize makespan for assembly job shop', *Applied Soft Computing*, Vol. 13, No. 3, pp.1391–1399.
- Wu, J. (2012) 'Efficient HIK SVM learning for image classification', *IEEE Transactions on Image Processing*, Vol. 21, No. 10, pp.4442–4453.
- Xu, X., Ding, S., Shi, Z. and Zhu, H. (2012) 'A novel optimizing method for RBF neural network based on rough set and AP clustering algorithm', *Journal of Zhejiang University – SCIENCE C*, Vol. 13, No. 2, pp.131–138.
- Yang, C.X., Zhang, X.W. and Zhang, F.P. (2012) 'Estimation of amplification effect of mining-induced blast vibration on surrounding structures using a hybrid GA-SVM', *Disaster Advances*, Vol. 5, No. 4, pp.1537–1543.
- Yang, H.Y., Luo, Q.F. and Zhou, Y.Q. (2009) 'Using chaotic adaptive PSO-SVM for heart disease diagnosis', *2nd IEEE International Conference on Computer Science and Information Technology*.
- Yao, Y., Zhang, T. and Xiong, Y. (2011) 'Mutation probability of cytochrome P450 based on a genetic algorithm and support vector machine', *Biotechnology Journal*, Vol. 6, No. 11, pp.1367–1376.

- Zhang, C. and Zhang, T. (2010) 'Soft measurement method for oil holdup of two phase flow based on least squares support vector machine and particle swarm optimization', *Proceedings of the CSEE*, Vol. 30, No. 2, pp.86–91.
- Zhang, S.T., Zhang, K. and Jiang, J. (2009) 'Study of fault diagnosis method for three-phase high power factor rectifier based on PSO-LSSVM algorithm', *2009 International Conference on Applied Superconductivity and Electromagnetic Devices*, pp.221–224.
- Zhang, X.D., Hu, F. and Zhao, L. (2011) 'Recognition of practical speech emotion based on improved shuffled frog leaping algorithm and support vector machine', *Signal Processing*, Vol. 27, No. 5, pp.678–689.
- Zhang, X.L., Chen, X.F. and He, Z.J. (2010a) 'An ACO-based algorithm for parameter optimization of support vector machines', *Expert Systems with Applications*, Vol. 37, No. 9, pp.6618–6628.
- Zhang, X.L., Chen, X.F. and He, Z.J. (2010b) 'Fault diagnosis based on support vector machines with parameter optimization by an ant colony algorithm', *Proceedings of the Institution of Mechanical Engineers Part C – Journal of Mechanical Engineering Science*, Vol. 224, No. 1, pp.217–229.
- Zhang, Y., Lu, X.H. and Jin, M. (2010c) 'Hydrological forecasting system model based on AFSVM', *Application Research of Computer*, Vol. 27, No. 8, pp.2902–2905.
- Zhao, H.B. and Yin, S.D. (2010) 'A CPSO-SVM model for ultimate bearing capacity determination', *Marine Georesources & Geotechnology*, Vol. 28, No. 1, pp.64–57.
- Zhou, J.G., Bai, T. and Suo, C. (2008) 'The SVM optimized by culture genetic algorithm and its application in forecasting share price', *2008 IEEE International Conference on Granular Computing*, Vol. 1, No. 2, pp.838–843.
- Zhou, Y. and Huang, X. (2012) 'Hybrid optimization algorithm based on mean particle swarm and artificial fish swarm', *Information – An International Interdisciplinary Journal*, Vol. 15, No. 2, pp.763–777.
- Zhou, Y. and Zhou, G. (2013) 'A hybrid glowworm swarm optimization algorithm for constrained engineering design problems', *Applied Mathematics & Information Sciences*, Vol. 7, No. 1, pp.379–388.
- Zhu, H.Q., Yang, C.H. and Wang, J. (2011a) 'Hybrid-QPSO-based parameters optimization of LS-SVM and its application', *Journal of Central South University (Science and Technology)*, Vol. 42, No. 4, pp.1000–1004.
- Zhu, X.L., Cai, H.Y. and Xu, Z.J. (2011b) 'Classification of 5-HT1A receptor agonists and antagonists using GA-SVM method', *ACTA Pharmacologica Sinica*, Vol. 32, No. 11, pp.1424–1430.