International Journal of Hybrid Information Technology
Vol.9, No.1 (2016), pp. 387-396
http://dx.doi.org/10.14257/ijhit.2016.9.1.33

Bacterial Foraging Optimization Based on LS-SVM for BTP Forecasting

[1]SONG Qiang, [2]GUO Xiao-bo and [3]LI hua

[1]Mechanical engineering department of Anyang institute of technology songqiang01@126.com
[2]School of computer and Information Engineering of Anyang institute of technology, songqiang08@126.com
[3]The Transport Department of Anyang iron and Steel Group Company, E-mail: 13523323305@126.com

Abstract

Because of the nonlinear characteristics of the BTP in sintering process, the BTP forecasting is difficult to realize. The LS-SVM was employed in this study for forecasting. However, Because SVM is using the two programming support vector, computing and solving two quadratic programming will involve matrix of order m, when the M number is large storage and computing the matrix will consume a large amount of computer memory and calculation time. The traditional training methods based on searching technique are not effective and fast. Therefore, bacterial foraging optimization (BFO) was adopted to optimize the LS-SVM. BFO is a novel and powerful global search technique, It is found that Bacteria Foraging Algorithm (BFO) is capable of improving the speed of convergence as well as the precision in the desired result. Simulation results clearly illustrate that the proposed approach is very efficient and could easily be extended for other global optimization problems. It can conclude that BFO is effective and rapid for the cluster analysis problem.

***Keywords**: Bacteria Foraging Algorithm (BFO), LS-SVM, Optimization, BTP*

1. Introduction

The sintering process is an important step in stove smelting. It not only can agglomerate power materials, but also has pretreatment function of burning method for raw materials so that the target of high productivity, high industry, but also from nonferrous industry in stove smelting. The study of sintering process has attracted interest, not only from iron and steel The Burning-Through-Point (BTP) is an important parameter in sintering process. In the iron and steel enterprises, sintered ore is the main raw material for blast furnace. In the sintering process, stability and quality of sinter is a decision factor to production efficiency of blast furnace. The sintering process is a preprocess for blast-furnace materials. The quality of sinter is very important for smooth operation and high productivity of the blast furnace since it improves the permeability and reducibility of the burden material. Burning-Through-Point for judging the quality of sinter important indicator is a measure of good and bad sintered minerals important parameters [1, 2].

The BTP is a very important parameter in the sintering process. The sinter quality can be improved, and the energy consumption can be reduced, if one can accurately predict the BTP value. That means the accurate prediction of BTP can bring significant economic benefits and is of important practical significance. Traditional methods using features such as exhaust gas temperature, negative pressure and exhaust gas composition to predict the BTP [3].

ISSN: 1738-9968 IJHIT

Research on BTP has attracted wide attention for many years. Several major methods and techniques have been proposed and developed. These methods can roughly be categorized into two types: One is classical methods, such as time series models , regression models, and ARX models [3]. The other is artificial intelligence methods, such as expert system, neural network, fuzzy logic .Among these methods, the neural network (NN) has received much attention in last few decades. The advantage of NN lies in the fact that they do not require any complex mathematical formulations or quantitative correlation between inputs and outputs. However, the NN faced with two shortcomings, viz., it converges too slowly and the optimal weights/biases are difficult to set. Although genetic algorithm can solve above flaws in some measure, but the results are not satisfactory. In this study, a novel bacterial foraging optimization (BFO) was proposed. The bacterial foraging optimization (BFO) is a new global-search technique and it has achieved widespread success in solving practical optimization problem in different fields.

2. Sintering Process and the Analysis of BTP

2.1 . The Analysis of BTP

The BTP is characterized by the maximum of the exhaust gas temperature, which is measured at six locations towards the end of the strand. It is primarily controlled by variable strand speed. The temperature is measured at a certain point. The method employed here is not to explicitly control the temperature maximum, but to yield the expected BTP by keeping the exhaust temperature distribution on a pre-defined curve. The BTP cannot be tested on-line, and the judgment based on the observation data by operators is usually inaccurate (Fan [1999]). Although many successful attempts have been made to model the sintering process (Wang [2002]), it is very difficult to obtain some important parameters in these models, which indicate the physical properties of the sinter material. Therefore, Soft-sensing method is adopted to solve this problem.

The model will be built up based on the BTP related variables in real-time. The value of BTP can be calculated according to the temperature curve of waste gas in wind-boxes.

It is generally believed that the temperature of waste gas in wind-boxes is highest when the sinter-mix bed is just burned through. Thermocouples are put along five wind-boxes at discharge end. Quadratic curve is fitted according to three points including the highest temperature. The general expression of quadratic curve is:

$$T_i = aX_i^2 + bX_i + c \quad , \quad (i=1,2,\ldots..,22) \qquad (1)$$

where T is temperature of waste gas in wind-boxes, x is number of wind-boxes, A,B,C are coefficients. With the three know values, ((x1,T1),(x2 ,T2),(x3,T3)), and the highest temperature, using the Eq.1 , the following expression can be obtained:

$$X_{\max} = -\frac{b}{2a}$$

$$T_{\max} = aX_{\max}^2 + bX_{\max} + c \qquad (2)$$

Burning Through Point (BTP) is the position when sinter process is finished and is expressed by corresponding wind-box position when sinter bed is burned through. Accurate control of Burning Through Point (BTP) can not only make the sintering process stable but also use the sintering area effectively. The BTP affects the sintering output and quality. It is generally believed that BTP should be at the second wind-box counting from backward. If BTP is ahead of this point, effective grate area of sintering machine will not be fully utilized. On the contrary, if BTP lags behind the point, sinter bed cannot be burned through, this will result in the increase of re-sinter and decrease of sinter rate. BTP may be affected by many factors. Row material parameters, operation parameters and process condition parameters may have a great influence on BTP direct or indirect. Moreover, BTP cannot be measured direct. Up till now, there are no instruments that can be used to measure BTP online. In the paper, the online prediction of BTP based on combined BFO-SVMis analyzed. Mathematical method will be used to solve the problem of online measurement of BTP [3].

In the practical sintering process, BTP is affected by a number of factors, such as the ignition temperature, the material thickness, mixture of water, trolley speed, lime dosage, fuel consumption, exhaust gas temperature, the alkalinity.

3. The Bacteria Foraging Optimization Algorithm

The survival of species in any natural evolutionary process depends upon their fitness criteria, which relies upon their food searching and motile behavior. The law of evolution supports those species who have better food searching ability and either eliminates or reshapes those with poor search ability. The genes of those species who are stronger gets propagated in the evolution chain since they posses ability to reproduce even better species in future generations. So a clear understanding and modeling of foraging behavior in any of the evolutionary species, leads to its application in any non-linear system optimization algorithm. The foraging strategy of Escherichia coli bacteria present in human intestine can be explained by four processes, namely chemotaxis, swarming, reproduction, elimination–dispersal [8, 9].

During foraging of the real bacteria, locomotion is achieved by a set of tensile flagella. Flagella help an E.coli bacterium to tumble or swim, which are two basic operations performed by a bacterium at the time of foraging. When they rotate the flagella in the clockwise direction, each flagellum pulls on the cell. That results in the moving of flagella independently and finally the bacterium tumbles with lesser number of tumbling whereas in a harmful place it tumbles frequently to find a nutrient gradient. Moving the flagella in the counterclockwise direction helps the bacterium to swim at a very fast rate. In the above-mentioned algorithm the bacteria undergoes chemotaxis, where they like to move towards a nutrient gradient and avoid noxious environment. Generally the bacteria move for a longer distance in a friendly environment. Figure 1 depicts how clockwise and counter clockwise movement of a bacterium take place in a nutrient solution.

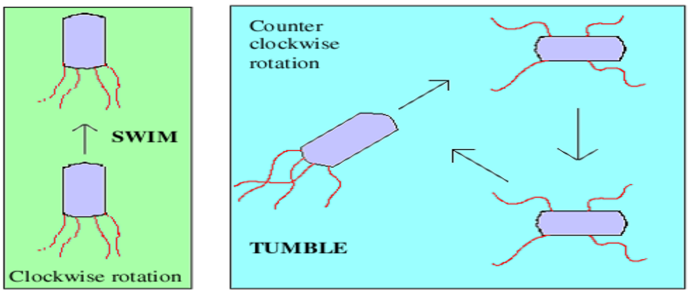

Figure 1. Swim and Tumble of a Bacterium

When they get food in sufficient, they are increased in length and in presence of suitable temperature they break in the middle to from an exact replica of itself. This phenomenon inspired Passino to introduce an event of reproduction in BFOA. Due to the occurrence of sudden environmental changes or attack, the chemotactic progress may be destroyed and a group of bacteria may move to some other places or some other may be introduced in the swarm of concern. This constitutes the event of elimination-dispersal in the real bacterial population, where all the bacteria in a region are killed or a group is dispersed into a new part of the environment [11-14].

3.1.1. Chemotaxis

In the original BFO, a unit walk with random direction represents a “tumble” and a unit. Represents walk with the same direction in the last step indicates a “run.” Suppose $\theta^i(j,k,l)$ the bacterium at jth chemotactic, kth reproductive, and lth elimination-dispersal step.

$C(i)$ is the chemotactic step size during each run or tumble .Then in each computational chemotactic step, the movement of the ith bacterium can be represented as:

$$\theta^i(j+1,k,l) =$$

$$\theta^i(j,k,l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta'(i)}} \tag{3}$$

Where Δi is the direction vector of the jth chemotactic step. When the bacterial movement is run, Δi is the same with the last chemotactic step; otherwise, Δi is a random vector whose elements lie in[-1,1]. With the activity of run or tumble taken at each step of the chemotaxis process, a step fitness, denoted as J(i,j,k,l) will be evaluated [15-18].

Swarming: It is always desired that the bacterium that has searched the optimum path of food should try to attract other bacteria so that they reach the desired place more rapidly. Swarming makes the bacteria congregate into groups and hence move as concentric patterns of groups with high bacterial density. Mathematically, swarming can be represented by

$$J_{CC}(\theta) = \sum_{i=1}^{S} J_{CC}(\theta) = \sum_{i=1}^{S}\left\{-d_{attract}\exp\left(-\omega_{attract}\sum_{m-1}^{p}\left(\theta_m-\theta_m^i\right)^2\right)\right\} +$$

$$\sum_{i=1}^{S}\left\{-h_{repellant}\exp\left(-\omega_{repellant}\sum_{m-1}^{p}\left(\theta_m-\theta_m^i\right)^2\right)\right\} \tag{4}$$

Where $J_{CC}(\vartheta, P(j,k,l)$ is the cost function value to be added to the actual cost function to be minimized to present a time varying cost function. “S” is the total number of bacteria. “p” is the number of parameters to be optimized that are present in each bacterium. $d_{attract}$ $h_{repellant}$、 $\omega_{attract}$、 $\omega_{repellant}$ are different cofficients that are to be chosen judiciously.

3.1.2. Reproduction

The health status of each bacterium is calculated as the sum of the step fitness during its life, that is $\sum_{j=1}^{N_c} J(i,j,k,l),$, where N is the maximum step in a chemotaxis process. All bacteria, are sorted in reverse order according to health status. In the reproduction step, only the first half of population survives and a surviving bacterium splits into two identical ones, which are then placed in the same locations. Thus, the population of bacteria keeps constant.

3.1.3. Elimination and Dispersal

The chemotaxis provides a basis for local search, and the reproduction process speeds up the convergence which has been simulated by the classical BFO. While to a large extent, only chemotaxis and reproduction are not enough for global optima searching. Since bacteria may get stuck around the initial positions or local optima, it is possible for the diversity of BFO to change either gradually or suddenly to eliminate the accidents of being trapped into the local optima. In BFO, the dispersion event happens after a certain number of reproduction to be killed processes. Then some bacteria are chosen, according to a preset probability Ped and moved to another position within the environment.

The flow chart of the iterative algorithm is shown in the following Figure 2.

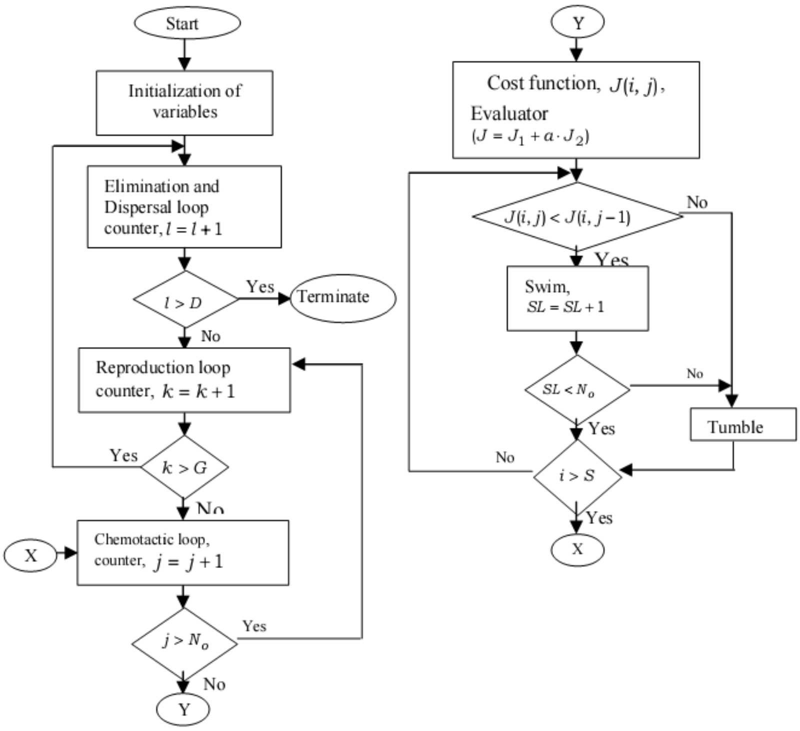

Figure 2. Flowchart of Bacterial Foraging Algorithm

4. Prediction Model based on Bayesian Framework and LS-SVM

4.1 LS-SVM Algorithm

The LS-SVM, evolved from the SVM, changes the inequality constraint of a SVM into all equality constraint and forces the sum of squared error (SSE) loss function to become an experience loss function of the training set. Then the problem has become one of solving linear programming problems. This call be specifically described as follows：

Given a training set $\{x_t, y_t\}_{t=1}^{N}$,with $x_t \in R^n$, $\mathrm{y}_t \in \mathrm{R}$, $x_t \in R^n$ is input vector of the first t samples , $\mathrm{y}_t \in \mathrm{R}$ is the desired output value of the first t corresponds to samples, N is the number of samples data, the problem of linear regression is to find a linear function y(x) that models the data. In feature space SVM models take the form:

$$y(x) = \omega^T \varphi(x) + b \tag{5}$$

Where the nonlinear function mapping $\varphi(\cdot): R^n \to R^{n_h}$ maps the high-dimensional space into the feature space.

Having comprehensively considered the complexity of function and fitting error, we can express the regression problem as the constrained optimization problem according to the structural risk minimization principle：

$$\min J(w,e)=\frac{1}{2}w^{T}w+\gamma\frac{1}{2}\sum_{t=1}^{N}e_t^2 \quad (6)$$

subject to the restrictive conditions,

$$y(x)=w^{T}\varphi(x_t)+b+e_t\,,\quad \text{for } t = 1,\ldots, N$$

Where γ is margin parameter, and e_i is the slack variable for x_i.

In order to solve the above optimization problems, by changing the constrained problem into an unconstrained problem and introducing the Lagrange multipliers, we obtain the objective function：

$$L(w,b,e,\alpha)=J(w,e)-\sum_{t=1}^{N}\alpha_t\{w^{T}\varphi(x_t)) \quad (7)$$

Where α_t is Lagrange multipliers. According to the optimal solution of Karush-Kuhn-Tucker (KKT) conditions, take the partial derivatives of (5) with respect to w, b and e, respectively, and 1et them be zero, we obtain the optimal conditions as follows:

$$\begin{cases}\dfrac{\partial L}{\partial w}=0\rightarrow w=\sum_{t=1}^{N}\alpha_t\varphi(x_t)\\ \dfrac{\partial L}{\partial b}=0\rightarrow \sum_{t=1}^{N}\alpha_t=0\\ \dfrac{\partial L}{\partial e_t}=0\rightarrow \alpha_t=\gamma e_t\\ \dfrac{\partial L}{\partial \alpha_t}=0\rightarrow w^{T}\varphi_t+b+e_t-y_t=0\end{cases}$$

After elimination of e_t and w , the equation can be expressed as a linear function group:

$$\begin{bmatrix}0 & \mathbf{1}^T\\ \mathbf{1} & \varphi(x_t)^T\varphi(x_t)+\mathbf{D}\end{bmatrix}\begin{bmatrix}b\\ \boldsymbol{\alpha}\end{bmatrix}=\begin{bmatrix}0\\ \mathbf{y}\end{bmatrix}$$

where $\mathbf{y}=[y_1,\cdots,y_N]$, $1=[1,\cdots,1]$, $\boldsymbol{\alpha}=[\alpha_1,\ \ldots,\ \alpha_N]$, $\mathbf{D}=diag[\gamma_1,\cdots,\gamma_N]$,

Select γ > 0, and guarantee matrix $\boldsymbol{\varphi}=\begin{bmatrix}0 & \mathbf{1}^T\\ \mathbf{1} & \varphi(x_t)^T\varphi(x_t)+\mathbf{D}\end{bmatrix}$ $\begin{bmatrix}b\\ \boldsymbol{\alpha}\end{bmatrix}=\boldsymbol{\varphi}^{-1}\begin{bmatrix}0\\ \mathbf{y}\end{bmatrix}$

Finally, the LS-SVM regression model can be expressed as

$$y(x)=\sum_{t=1}^{N}\alpha_t\exp\{-\|x-x_t\|_2^2/2\sigma^2\}+b$$

Where σ is a positive real constant. Note that in the case of RBF kernel function, one has only two additional turning parameters σ and γ, which is less than standard SVM.

There exist two parameters (γ and σ) selected by the user in advance. γ is viewed as a regularization parameter, which controls the trade-off between complexity Of the machine and the number of non. separable points The kernel parameter denotes the width of Gauss function The difference mainly attributes to the different effects of parameters to LS-SVM regression namely γ only influences the training step in the whole process while σ affects both the training step and the predicting one.

The role of the kernel function parameters σ affect the complexity of the distribution of sample data in high-dimensional feature space. The change in the parameters of the kernel function is actually implied changing mapping function, thereby changing the number of dimensions of the sample space. The dimension of the sample space determines linear classification surface VC dimension can be constructed in this space, and also determines the the linear classifier surface to achieve the minimum empirical error. Example RBF kernel (kernel function parameter is the greater, the wider the range of the output response of the samples, the smaller the optimal classification surface structure risk, but experience increased risk; smaller the output response of the sample, the narrower the interval, the optimal hyper plane experience risk will get smaller, the structural risk will increase, and likely to cause over-fitting, to reduce the performance of the SVM. choose the δ2 need to trade-off between these two. The role of γ is to adjust the proportion of learning machines confidence range of experience and risk data subspace determined to have the best ability to promote the learning machine. The optimal γ are different according to the different data subspace. Data determined subspace, small values of γ indicates that the punishment of the empirical error is small, the degree of complexity of the learning machine small experience greater risk value; vice versa. When γ exceeds a certain value, the degree of complexity of SVM reached the maximum value allowed by the sample space. Almost no experience in risk and the ability to promote change. The each data subspace least there is an optimal γ SVM promote the ability to achieve the best. Not yet a unified way to decide the best values of the parameters of SVM, the choice is a trial and error method to get satisfactory results through continuous experiments. Not yet a unified approach to determine the the SVM parameters best take Value, the general selection method is trial and error method, which through continuous experiments to be satisfied with the Results.

Obviously it plays an important role to select appropriate parameters in the whole process of prediction using LS-SVM. Motivated by pursuing smaller error we have discussed the methods of selecting parameters There have existed several effective optimization approaches for LS-SVM, including the method based on genetic algorithm (GA) and multilayer adaptive parameters optimization(MAPO) approach, etc. However, they have some disadvantages such as costing longer training time easily destroying the searched results for crossover operator existing in GA, and easily trapping into those parameters with local minimization using MAPO.

This LS-SVM regression leads to solving a set of linear equations, which is for many application in different areas. Especially, the solution by solving a linear system is instead of quadratic programming. It can decrease the model algorithm complexity and shorten computing time greatly. The LS-SVM algorithm software package is run in MATLAB 7.0.1 software.

5. BFO Optimize SVM Parameters

5.1 Prediction Step BFO-SVM Prediction Steps are as Follows:

a) Analysis of the experimental data, selected treated prediction indicators have important influence on the relationship between parameters, pretreatment of the sample set is formed and the sample set; b) choose the of SVM types and nuclear function, to determine the operating parameters required by the model to determine the operating parameters of the BFO or PSO algorithm, the SVM regression model; c) call the optimal parameters of the PSO-SVM the algorithm search SVM regression model; d) the optimal parameters of re-training the SVM regression machine regression model was established for solving the regression equation; e) promote the proficiency test with the test sample set.

5.2 Experiments and Discussions

In order to evaluate the accuracy of the model, we use RMSE to evaluate, Mean-Square-error(RMSE)and coincidence probability P are taken to evaluate the performance of the LS-SVM regression The first criterion(RMSE)is calculated

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i' - x_i)^2} \qquad (8)$$

where x_i' and x_i stand for the predicted and desired values respectively, and N denotes the total number for prediction.

The second one which shows the coincidence degree between the predicted and desired values is evaluated as

$$p = \frac{N_c}{N} \qquad (9)$$

where N_c represents the number of points whose errors are less than certain value .

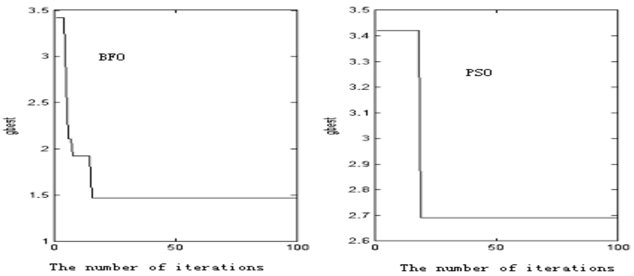

Figure 3. Iterations Result

To compare algorithms fairly, we tested the optimization ability of BFO algorithm. Original BFO, PSO and Genetic Algorithm (GA) were employed for comparison. It is clear from Figure 3 that BFO obtained the best values .BFO is best on the rest one. PSO performed worst.

In order to search the global optimum more efficiently planned further work is to improve the mutation operation BFO and investigate how to choose the evolutions steps k and k to deal with different issues. BFO can be applied to not only LS-SVM but also other modified versions of SVM. Its development will stimulate the application of SVM to more general fields.

The data set as follows:

Number of bacterium (s):20; Number of chemotatic steps (n_c):20; Swimming length (ns):3; Number of reproduction steps (n_{re}):4; Number of elimination and dispersal events:5; Depth of attractant ($d_{attract)}$;0.12;Width of attractant ($d_{attract}$):0.25; Height of repellent ($h_{repellant}$) :0.1; Width of repellent ($r_{epellant}$) :10; Probability of elimination-dispersal events (p_{ed}):0.02.

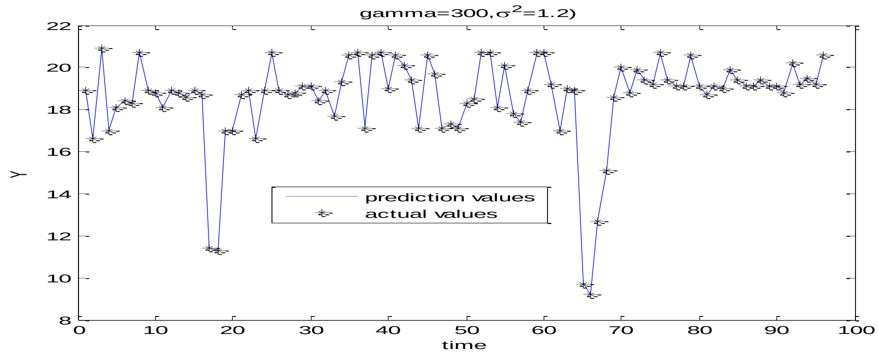

Figure 4. Prediction Diagram of the BTP based on BFO and LS-SVM

Figure 4 the comparison of forecast load with the actual. The results are found to be quite accurate. From Table 2 it can be seen that three models present quite satisfactory forecasting results. By comparing the Mean relative error, Mean absolute error and Mean square error of BFO-SVM is smaller than that of BFO and LS-SVM. Moreover, BFO-SVM has higher precise prediction than BFO and LS-SVM.

6. Conclusion

This paper presented an approach to optimize the parameters of LS-SVM regression. The effectiveness of the proposed was verified and was also compared against results produced by BFO. From the undertaken experiments, it is shown that the approach of using BFO in optimizing parameters of the Least Squares Support Vector Machines (LS-SVM) is beneficial. This can be seen in the obtained results of Mean Squared Error and prediction accuracy. The hybridization of BFO-LSSVM can be considered as a promising technique that is suitable for predicting nonvolatile data. There are ways to improve our proposed algorithms. Further research efforts should focus on the tuning of the user-defined parameters for ABFO algorithms based on extensive evaluation on many benchmark functions and real-world problems.

Acknowledgements

In writing this paper, I have benefited from the presence of my teachers and my classmates. They generously helped me collect materials I needed and made many invaluable suggestions. I hereby extend my grateful thanks to them for their kind help, without which the paper would not have been what it is. Particularly, I am deeply indebted to Professor YANG Cheng-zhi, my supervisor, who guided me throughout my writing of this thesis. He carefully read the whole draft and offered painstaking and precious criticism. His standards of academic excellence have made my revision and exciting and gratifying experience. I also wish to sincerely thank my The National Natural Science Foundation of Henan province talent training joint fund project (U1204613), whose brilliant ideas and perceptive observations have proved immensely constructive.

References

[1] F. Xiao-Hui and W. Hai-Dong, “Mathematical model and Artificial Intelligence of sintering process”, Central South University Press, vol. 47, no. 1, **(2013)**, pp. 064 - 072.
[2] W. Yiwen, G. Weihua and W. Yalin, “Integrated model for predicting burning through point of sintering process based on optimal combination algorithm”, The Chinese Journal of Nonferrous Metal, vol. 12, no. 1, **(2002)**, pp. 191–195.
[3] Z. Xiao-Long, Forecasting method and application of BTP based on neural network”, Central South University, **(2006)**.

[4] C. Wushan, "A building of the genetic-neural network for sinter's burning through point", Sintering and Pelletizin, vol. 29, no. 5, **(2004)**, pp. 18–22.
[5] X. Qi-Liang, W. Min, X. Jie, *et al.*, "Prediction method of BTP Based on multiple fuzzy linear regression", Control Engineering, vol. 14, no. 6, **(2012)**, pp. 603-605.
[6] W. Xiao-Feng and F. Min-Rui, "Fuzzy control applied to burning through point based on support vector machines prediction model", Journal of Zhejiang University (Engineering Science), vol. 41 no. 10, **(2007)**, pp. 1721-1725.
[7] L. Yu-Chang, G. Wei-Hua and Z. Jie-Min, "Study on intelligent control of BTP based on soft-sensor technology", Sintering and Pelletizing, vol. 27, no. 2, **(2003)**, pp. 27-30.
[8] Z. Xue-Gong, "On the statistical learning theory and support vector machines", Automation Journal, vol. 26, no. 1, **(2006)**, pp. 1-6.
[9] L. Guo-Zheng, W. Meng and Z. Hua-Jun, "Support vector machine Introduction", Beijing: Electronics Industry Publish press, **(2004)**.
[10] M. H. Li and J. Wang, "The Research for Soft Measuring Technique of Sintering Burning Through Point", Industrial Electronics and Applications, 2006 1ST IEEE Conference on, **(2006)** May 24-26, pp. 1-4.
[11] H. C. Yuan, F. L. Xiong and X. Y. Huai, "A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy", Computers and Electronics in Agriculture, vol. 40, no. 1-3, **(2003)**, pp. 57-64.

Author

SONG Qiang, male, He was born in Xintai city of Shandong province. He is a Associate professor of Anyang Institute of Technology. He received his master's degree from Kunming University of science and technology in 2006.His research interests are in the areas of intelligence control theory and rough set theory. GUO Xiao-bo, male, was born in 1957, he is a Associate professor of Anyang Institute of Technology. He received his doctor's degree from DongNan University. His research direction is SVM and rough set of intelligent decision-making research.

.

