



Support vector machine algorithm for artificial intelligence optimization

Xian Tan^{1,2} · Fasheng Yu¹ · Xifeng Zhao²

Received: 29 January 2018 / Revised: 6 March 2018 / Accepted: 10 March 2018 / Published online: 14 March 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

To improve the short-term power prediction accuracy, a comparative analysis of the PSO and SVM algorithm was carried out. Then, the two were combined and, the penalty factor and kernel function parameters in SVM model were optimized by the improved PSO algorithm. The SVM algorithm with optimized parameters and model were applied to predict and control and form PSO-SVM algorithm. Finally, the short-term power load was modelled and predicted based on PSO-SVM algorithm and it was compared with the conventional SVM algorithm. The results showed that the relative error of the average absolute value of PSO-SVM method was 1.62%, while the relative relative error of the average absolute value of conventional SVM using particle swarm optimization algorithm was 3.52%. It can be seen that the error adopting the new algorithm is reduced by 1.9%. It shows that the precision of the improved power load forecasting model is greatly improved.

Keywords Particle swarm optimization (PSO) · Support vector machines · Artificial intelligence · Power load

1 Introduction

With the gradual progress of science and technology and the development of society, the modern industrial process presents many complex characteristics such as multi-variable coupling, strong non-linearity and uncertain model, which seriously restrict the production and rapid development of industry. How to accurately predict and control the complex industrial production process has become a hot issue of today's scholars. The so-called artificial intelligence optimization algorithm, in essence, is a summary and improvement of some natural phenomena in daily life or the law of human activity. First of all, it needs to determine how to transform the practical problems into corresponding mathematical theory model that we know. Then, the analysis of and solution to the problems are actually transformed into analysis of and solution to the

mathematical model and the formula. The main research directions include two aspects: the first is a kind of improvement and innovation to the existing algorithm, and another is to use different characteristics of different algorithms to combine some algorithms. In fact, there are many examples for this, which are not described in detail. In order to obtain the optimum solution, it drives people to conduct continuous research and exploration for many years, and the optimization problem becomes the emerging popular practical science [1].

Support vector machine is another new artificial intelligence algorithm. Its proposed time is the same time as that of particle swarm optimization algorithm, which is raised by the famous scholar Vapnik. This algorithm is also different from other algorithms. Especially in dealing with the small samples and very complicated nonlinear model problems, it has great advantages. In addition, because its essence is the mapping and conversion of a dimension, for the high latitude pattern recognition problem, it also has special advantages and at the same time used in many other algorithms [2]. Compared with the traditional neural network based on empirical risk minimization, it has fast learning speed, global optimization and generalization

✉ Fasheng Yu
yufasheng@mail.ccnu.edu.cn

¹ Communication & Journalism College, Central China Normal University, Wuhan, China

² College of Literature and Communication, Hubei University for Nationalities, Enshi, China

ability. Its learning result is more obvious than other pattern recognition and regression prediction methods.

2 PSO and SVM algorithm

2.1 Improved PSO algorithm

Particle swarm optimization (PSO) algorithm is an artificial intelligence algorithm summed up based on the action laws of birds and fish in the biological clock. The process of the algorithm is that when fish are foraging, they have both group action and other foraging processes. This group and individual behaviour is destined for some fish to find better food ways, while some fish can't find better food sources. When there is a good food way, it will continue to pass this way. Its path information is constantly transmitted and optimized, and ultimately finds the best way. This food way is the global optimal solution of this problem [3]. The way of particle swarm finding the optimal solution through collaboration among individuals is compared with genetic algorithm adhering to the "survival of the fittest". It takes particles with no quality and no volume as individuals, the behaviour rules are simple and complex characteristics of each particle can solve the complex optimization problems.

Assuming that in a fixed period of time, every particle is a possible solution. At the same time, the feasible solution is likely to be selected as the model global optimal solution for many. Different feasible solutions, namely the times for corresponding particles selected as the global optimal solution, are also different. Therefore, in selecting the global optimal solution each time, the particles with the most global optimal solution should be considered with priority [4].

In the model, for the number of the particles in the model selected as the optimal solution, a corresponding tabu list is established. At the same time, for the table, a specific vector is constructed. The vector dimension is determined by the number of particle algorithm, each dimension in vector represents the times of a particle, namely the corresponding feasible solution chosen as the global optimal solution of the model. The table has a reference standard value, which is selected as 100 in this article. The penalty function is designed as a recursive function to maintain the historical value of penalty value, and as a function of the tabu list, taking the frequency components as independent variables, so as to achieve the principle of global extremum selection. When a penalty function value is larger on a particle, it means that it is chosen to do the extreme value for many times. Next time, we need to continue to reduce its chance to be selected, and maintain the penalty function to achieve the possibility of larger function value [5]. If a penalty function selected on a

particle is small or 0 at a time, it means that the times of it selected to do the global extreme value is less or not selected. Then, next time, it is necessary to maintain or increase the chance to be selected, so that the penalty function value has a smaller function value.

The step process for improving the PSO algorithm is shown in Fig. 1.

2.2 SVM model

The idea of support vector machines (SVM) is established based on the optimal classification surface on the linear separable basis. The optimal classification point set diagram is shown in Fig. 2.

In Fig. 2, the optimal classification surface is relative to the multidimensional space. In two-dimensional space, the optimal classification surface is the best classification line. It means that, in the premise of ensuring empirical risk minimization, it can accurately separate the two types of data samples and make the interval the maximum, that is,

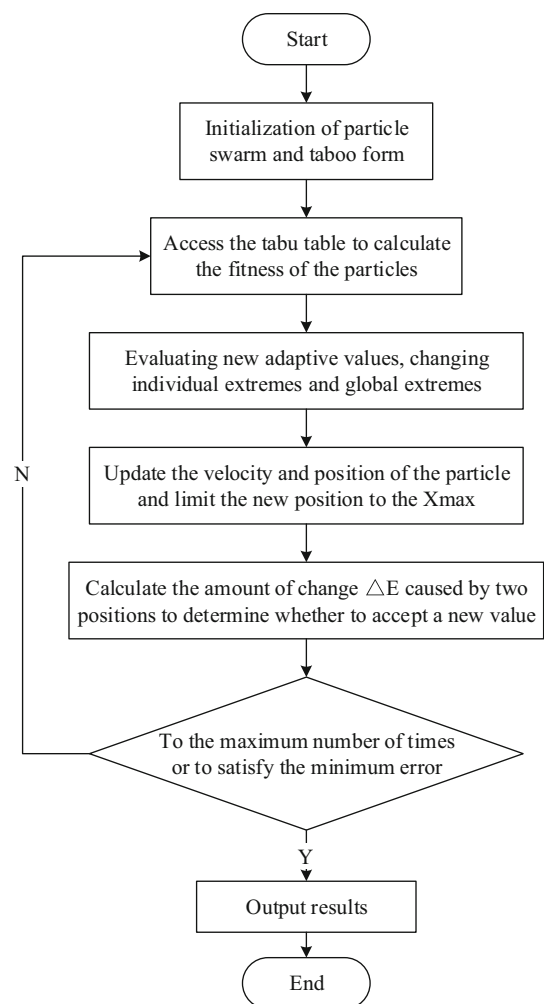


Fig. 1 The flow chart of PSO model improvement

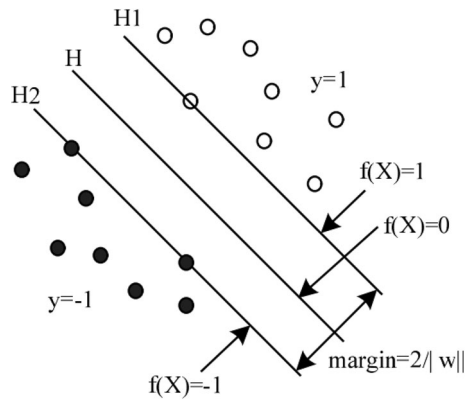


Fig. 2 The best classification hyper-plane

the classification interval is the maximum [6]. H is a classification line. H_1 and H_2 represent the smallest line that separates different categories, respectively, and the distance between them is different. They balance each other, and at the same time, they are also balanced with classification. The classification interval is the distance between two kinds of classification lines. In the above figure, the two types of data samples are solid points and hollow points.

2.3 PSO optimized SVM regression estimation model

From the principle of support vector machine algorithm, it can be seen two parameters of support vector machine have great influence on the diagnosis effect of the model. A parameter is the penalty factor C , which controls the penalty degree of samples beyond the error, and another is the kernel function parameter σ , which represents the width of the radial basis function. The two parameters determine the generalization ability of SVM. As a result, to get the best diagnostic results, it needs to obtain the optimal parameters [7]. The traditional methods of selecting parameters include empirical method, experimental comparison method, crossover validation method and grid search method, all of which are time-consuming and inefficient. In order to solve this problem, this chapter combines the advantages of particle swarm optimization algorithm and support vector machine algorithm, and uses particle swarm optimization algorithm to optimize the parameters of SVM regression estimation model, so as to further improve its learning ability and convergence speed.

The prediction accuracy of SVM is closely related to the value of the penalty parameter C , RBF kernel function σ , and the value of the insensitive loss parameter ϵ . Therefore, the optimal combination of support vector machine parameters is the key to obtain better predictive performance.

Particle swarm optimization (PSO) is a kind of swarm intelligence optimization algorithm. PSO then produces a group of particles, which are the initial model feasible solutions. Then, it continues to search and iterate through the particles, constantly modifies the feasible solutions, and finally finds the global optimal solution [8]. When the location is updated each time, every particle is updated by tracking personal extremum p_{best} and global minimum g_{best} . The personal extremum p_{best} indicates the optimal solution of the particle itself; the global extremum g_{best} refers to the optimal solution found by the entire population, as shown in Fig. 3.

The PSO-SVM intelligent algorithm combines PSO algorithm model and SVM regression model. Its essence is to optimize parameters in SVM by particle swarm optimization algorithm, and the specific calculation steps are as follows:

First of all, the initial value of the particle group is set. T_{max} is the maximum number of iterations, w is the inertia weight, and c_1 and c_2 are the acceleration factors. Assuming that a particle iteration is at the current time of t , then the initial particle swarm is $X(t)$, x_1, x_2, \dots, x_s represents every particle in the particle swarm, and they are all generated randomly. In the meanwhile, corresponding to the particle population, corresponding initial speed is set, each particle is v_1, v_2 and v_s , and the speed of the group is $V(t)$.

Secondly, it is the evaluation of the whole particle group. It is assumed that the fitness function is as follows:

$$-F = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{1}$$

For each particle in a particle population, the closer the distance from the target is, the greater the value of the fitness is. Among them, \hat{y}_i and y_i represent the actual and target output results of SVM model, respectively.

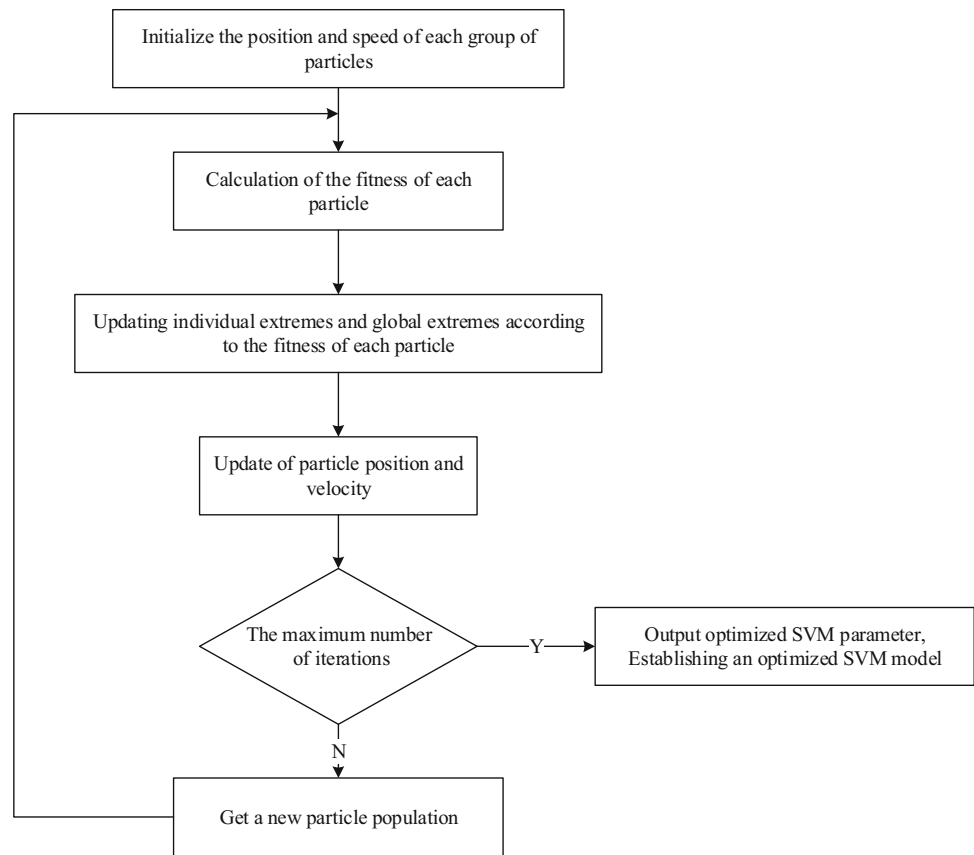
Thirdly, the position of each particle and the velocity value are updated by two extremes, and the specific updating equation is as follows:

$$v_{id}(t+1) = wv_{id}(t) + c_1r_1(p_{id}(t) - x_{id}(t)) + c_2r_2(p_{gd}(t) - x_{id}(t)) \tag{2}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{3}$$

w is called inertia weight; $d = 1, 2, \dots, n$; t suggests the current evolution iteration; $i = 1, 2, \dots, S$ and S represents the group scale, used for controlling the random numbers r_1 and r_2 of effect for previous velocity on the current velocity; the interval is $[0,1]$; c_1 and c_2 are two positive constants; acceleration factor x_{id} indicates the particle position; v_{id} refers to the particle velocity; p_{id} denotes the individual best solution p_{best} and p_{gd} means the global best solution g_{best} ;

Fig. 3 The PSO based on SVM parameter optimization process



Fourth, the iteration termination condition is detected. There are two main conditions to terminate the iteration: the first one is the number of iterations to reach the maximum number of iterations set in the first step and the second is not reached the maximum number of iterations, but has found the accuracy less than the set error value and at this time, iteration can be stopped [9]. If none of them is reached, then return to the second step to continue the iteration;

Fifth, after the iteration, the parameters of the support vector machine are extracted, and the PSO-SVM control prediction model is established;

Sixth, the obtained model is used to train the sample data, and the sequence minimum method is improved to solve the model;

Seventh, the parameters obtained in the sixth step are substituted into (4) to determine the final function equation of the model. The new sample data is predicted with this equation.

But the larger inertia weight w has better global search ability. Because the inertia weight will decrease correspondingly with the increasing number of iterations, when initializing, we must carefully select the inertia weight and it should not be too small. In this way, the PSO model algorithm can improve the global or local search and final

convergence. Suppose the number of iterations is the law of linear change, where the maximum and the minimum are 0.9 and 0.4, then we can get:

$$w(t) = 0.9 - \frac{t}{MaxNumber} \times 0.5 \quad (4)$$

The MaxNumber is specified as the maximum number of iterations. Another important factor is the acceleration factors c_1 and c_2 and the selection will also affect the convergence ability of particles. If c_1 value is great, it will result, in the search process, a large part of the particle trapped in range of the local circulation, ultimately unable to find the optimal solution in the global scope. Similarly, if c_2 value is relatively large, then the process of the whole particle swarm optimization will quickly converge to a local extremum, so the biggest drawback is that the random factor cannot be balanced, so after the comprehensive analysis above, $c_1 = c_2 = 2$.

3 Short-term power load prediction based on PSO-SVM

Short-term load prediction is a traditional research problem in the field of power system, which is of great significance to the planning and scheduling of power systems. With the

development of modern power grid, the theory and method of forecasting in the electric power market have been paid more and more attention to by relevant departments and researchers. In recent years, SVM has been applied to short-term power load prediction and also achieved good results, but there are outstanding problems to be solved for the method in practical applications, which is how to include balance parameter C , insensitive parameter ε for setting to directly affect the key parameters of the algorithm performance [10]. In the studies having applied this method, crossover validation is usually used to get all kinds of key parameters, but this method is blind and time-consuming [11]. In this paper, the particle swarm optimization (PSO) is proposed to optimize the parameters of SVM and to predict the short-term load.

3.1 Prediction steps

In order to reduce the scale of the solution, we can set up load prediction for every point on a working day. Because the load rule of general working days and holidays is quite different, for the prediction accuracy, we must establish a sub sample model in turn. The steps of short-term power load prediction based on PSO-SVM are as follows:

First of all, the exception samples are processed. Considering the proportion of previous and back load, if there is an obvious load value beyond the load range, then the data is an anomaly. In order to eliminate such abnormal data, we can take an estimated value, which is within normal load range, and then check and calibrate the load data one by one;

Secondly, the influence factors of non-numerical value are numerically processed. For example, in the numerical process of illumination, it can be set up: the number of sunny days is 5, and the numerical value of rain is 1, of which cloudy and overcast days take 4 and 3, respectively. Therefore, the non-numerical factors are numerically processed;

Thirdly, the sample of similar daily type load is selected and divided into two parts, one is training sample, and the other is the test sample. We choose the known days as the first 4 weeks of the whole sample, and give a constant r . We take the evaluation function less than the constant as a similar load day type. r value is an empirical value, which is selected according to people's experience;

Fourth, if the directly collected data is used for direct prediction and calculation, it may be saturated. In order to avoid this problem, we conduct normalization on load data so that the input load data is between $[0, 1]$, and at t moment, the load data uses the following normalization formula:

$$\hat{L} = \frac{L_t - L_{\min}}{1.5L_{\max} - L_{\min}} (t = 1, 2, \dots, 24) \quad (5)$$

L_{\max} and L_{\min} are the maximum and the minimum training sample collective load, respectively. The data of temperature and light are also treated with normalization.

Fifth, the influence factors of all similar day types are taken as input values of PSO-SVM model, and the load values of t of each similar day are taken as output samples. When the parameter vector (C , ε and σ) is selected and determined, we use the optimization algorithm to search the global optimal value, and assign the value to the SVM prediction model. The SVM is tested and trained with the corresponding sample data;

Sixth, to predict the load at a certain time in the future, we need to use the SVM that has been trained;

Finally, the predictive value is corrected. Because there are many uncertain factors in the real environment, the future load demand based on historical data is also need to be corrected manually, but this process is more complicated. In addition, there will be a great difference between the prediction result and the original load change rule. For such data problem, we need to replace all the data with the average value of the prediction value before and after the data.

3.2 Prediction results

In this paper, our specific application is to carry out 24-points short-term load forecast for the grid of a city in Northeast. We collect 24-points load data and various meteorological factors including average daily temperature, peak temperature, and the illumination condition in the city within 1 year and set up the model of load prediction [12].

Matlab software is used for simulation and the simulation program is compiled. Here, the overall size of the particle swarm is selected as 20, three-dimensional solution space corresponds to C , ε and σ , the initial weight is 0.9 and the maximum number of iterations of the program is determined to 200 times. The acceleration factor is $c_1 = c_2 = 2$; C is in the range of $[0.1, 150]$, ε is ranged in $[0, 0.8]$ and σ is ranged in $[0.1, 10]$. The maximum velocity vector corresponding to (C , ε , σ) is (0.5, 0.1, 0.1). Figure 4 is the prediction result of the city on March 17, 2016. The conventional SVM method and the PSO-SVM method are used to predict. Table 1 is the relative error list of the two methods.

From the table, we can see that the relative error of the absolute value of PSO-SVM method is 1.62%, while the relative error of the absolute value of conventional SVM using particle swarm optimization is 3.52%. It can be seen that the error of the new algorithm is reduced by 1.9%.

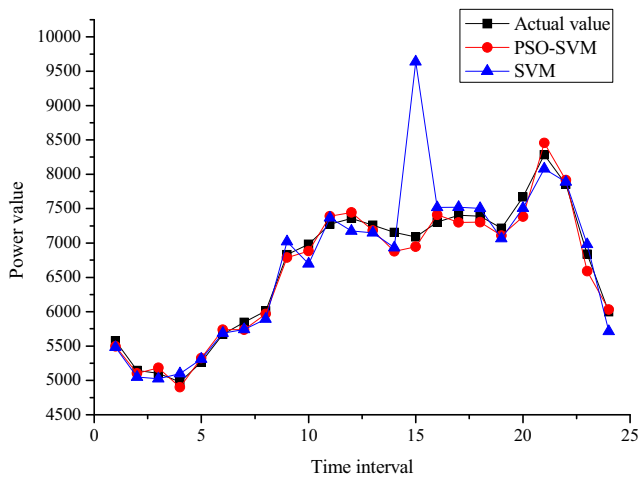


Fig. 4 Short-term power load forecasting of SVM and PSO-SVM

Table 1 SVM and PSO-SVM short-term power load forecasting error comparison table

Time interval	PSO-SVM error (%)	SVM error (%)
1	1.43	1.68
2	0.87	1.89
3	1.62	1.5
4	1.62	2.42
5	1.06	0.82
6	1.21	0.29
7	1.84	1.92
8	0.74	1.97
9	0.65	2.82
10	1.51	4.21
11	1.62	1.32
12	1.21	2.51
13	1.16	1.54
14	3.75	3.22
15	2	35.97
16	1.43	2.94
17	1.47	1.57
18	1.21	1.54
19	1.51	2.1
20	3.68	2.22
21	2.13	2.48
22	0.73	0.43
23	3.75	2.17
24	0.62	4.83
The mean value	1.62	3.52

When the system is predicted, it can be found that the training time is within 2 s. It can be concluded that the support vector machine optimized by particle swarm optimization algorithm effectively improves the accuracy of

short-term load forecasting, and the speed of operation has been significantly improved.

4 Conclusion

The problem of particle swarm optimization (PSO) and support vector machines (SVM) is mainly studied. On the basis of load forecasting for traditional SVM method, a short-term load forecasting method based on SVM optimized by PSO is proposed. This method effectively overcomes the blind selection of system parameters in SVM. Using a test day consumption data and feature data as the experimental data, and the maximum temperature, minimum temperature, wind, humidity and weather conditions as the feature data of input and electricity for output data, PSO-SVM experimental samples are constructed. The experimental results showed that the power consumption prediction based on particle swarm optimization algorithm and support vector machine had a high prediction accuracy, which was 1.9 percentage points lower than the traditional algorithm, very suitable for electricity consumption prediction. Compared with the SVM method, which has been widely used at present, the precision of short-term load forecasting is improved, and the speed of operation has been greatly improved.

Acknowledgements The authors acknowledge the National Natural Science Foundation of China (Grant: 111578109), the National Natural Science Foundation of China (Grant: 11111121005).

References

- Xiao, Y., Kang, N., Hong, Y., et al.: Misalignment fault diagnosis of DFWT based on IEMD energy entropy and PSO-SVM. *Entropy* **19**(1), 6–21 (2017)
- Du, J., Liu, Y., Yu, Y., et al.: A prediction of precipitation data based on support vector machine and particle swarm optimization (PSO-SVM) algorithms. *Algorithms* **10**(2), 57–68 (2017)
- Long, Y., Du, Z.J., Wang, W.D., et al.: PSO-SVM-based online locomotion mode identification for rehabilitation robotic exoskeletons. *Sensors* **16**(9), 1408 (2016)
- Nieto, P.J.G., García-Gonzalo, E., Sánchez, A.B., et al.: Air quality modeling using the PSO-SVM-based approach, MLP neural network, and M5 model tree in the metropolitan area of Oviedo (Northern Spain). *Environ. Model. Assess.* **4**, 1–19 (2017)
- Kapoor, N., Ohri, J.: Fuzzified PSO-SVM controller for motion control of robotic manipulator. *Int. J. Ind. Syst. Eng.* **24**(3), 361 (2016)
- Zhou, C., Yin, K., Cao, Y., et al.: Application of time series analysis and PSO-SVM model in predicting the Bazimen landslide in the Three Gorges Reservoir, China. *Eng. Geol.* **204**, 108–120 (2016)
- Hayat, M., Tahir, M.: PSO fuzzy SVM-TMH: identification of transmembrane helix segments using ensemble feature space by incorporated fuzzy support vector machine. *Mol. Biosyst.* **11**(8), 2255–2262 (2015)

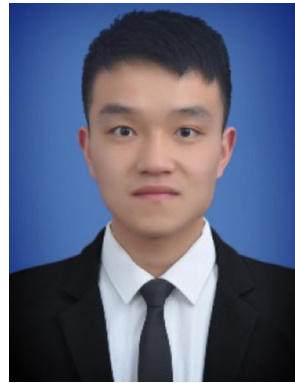
8. Hameed, S.S., Hassan, R., Muhammad, F.F.: Selection and classification of gene expression in autism disorder: use of a combination of statistical filters and a GBPSO-SVM algorithm. *PLoS ONE* **12**(11), e0187371 (2017)
9. Guo, X., Guo, X., Su, J.: Improved support vector machine short-term power load forecast model based on particle swarm optimization parameters. *J. Appl. Sci.* **13**(9), 1467–1472 (2013)
10. Selakov, A., Mellon, S., Bekut, D.: Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank. *Appl. Soft Comput.* **16**(3), 80–88 (2014)
11. Li, H.Z., Guo, S., Li, C.J., et al.: A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl.-Based Syst.* **37**(2), 378–387 (2013)
12. Lee, W.J., Hong, J.: A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. *Int. J. Electr. Power Energy Syst.* **64**, 1057–1062 (2015)



Xian Tan born in July 1974, Ph.D. candidate in Central China Normal University and Senior engineer of Hubei University for Nationalities. His main research interest include Big data, Internet search and mining and database technology.



Fasheng Yu born in Mar 1969, who is Associate-Dean, Professor and doctoral supervisor of Journalism and Communication of Central China Normal University. His main research interest include Database of the emergency, online public opinion, Com-connection and et al.



Xifeng Zhao born in Nov 1993, master candidate in Literature and Media College of Hubei University for Nationalities, His main research interest include big data and public emergency.