4. Marinakis Y, Marinaki M, Dounias G. Honey bees mating optimization algorithm for the Euclidean traveling salesman problem [J], Information Sciences, 2011, 181(20): 4684-4698.
5. SALIM M, VAKIL-BAGHMISHEH M T. Discrete bee algorithms and their application in multivariable function optimization [J], Artificial Intelligence Review, 2011, 35(1): 73-84.
6. Shuanfeng Zhao, Guanghua Xua, Tangfei Tao, Lin Liang. Real-coded chaotic quantum-inspired genetic algorithm for training of fuzzy neural networks. Computers and Mathematics with Applications, 2009, 57: 2009-2015
7. XU C, DUAN H. Artificial bee colony (ABC) optimized edge potential function (EPF) approach to target recognition for low-altitude aircraft[J], Pattern Recognition, 2011, 31: 1759-1772.
8. SINGH A. An artificial bee colony (ABC) Algorithm for the leaf-constrained minimum spanning tree problem [J], Applied Soft Computing, 2009, 9: 625-631.
9. Yang X S, Sadat Hosseini S S, Gandomi A H. Firefly algorithm for solving non-convex economic dispatch problems with valve loading effect [J], Applied Soft Computing, 2012, 12(3): 1180-1186.
10. Pan Q K, Fatih Tasgetiren M, Suganthan P N, et al. A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem [J], Information Sciences, 2011, 181(12): 2455-2468.
11. Parpinelli R S, Lopes H S. New inspirations in swarm intelligence: a survey [J], International Journal of Bio-Inspired Computation, 2011, 3(1): 1-16.
12. Wang L, Zhou G, Xu Y, et al. An effective artificial bee colony algorithm for the flexible job-shop scheduling problem [J], The International Journal of Advanced Manufacturing Technology, 2012, 60(1-4): 303-315.
13. QUANKE P, TASGETIREN M F, SUGenetic algorithm NTHAN P N, CHUA T J. A discrete artificial bee colony algorithm for the Lot-streaming flow shop scheduling problem [J], Information Sciences, 2011, 181(12): 2455-2468.
14. SUNDAR S, SINGH A, ROSSI A. An artificial bee colony algorithm for the 0–1 multidimensional knapsack problem [J], Communications in Computer and Information Science, 2010, 94(1): 141-151.
15. OMKAR S N, SENTHILNATH J, KHANDELWAL R, NARAYANA NAIK G, GOPALAKRISHNAN S. Artificial bee colony (ABC) for multi-objective design optimization of composite structures [J], Applied Soft Computing, 2011, 11(1): 489-499.

Research on Power Load Forecasting Models Based on Simulated Annealing Support Vector Machine (SA-SVM) Algorithm Mathematical

Liu Qi[1]*, Huang Zhenzhen[2]

School of Mechanical and Electrical Engineering, Zhoukou Normal University, Zhoukou, Henan, China
School of Political Science and Law, Zhoukou Normal University, Zhoukou, Henan, China

Abstract
A Load forecasting belongs to stratagem forecasting, which is an important research content of electrical system planning and running and a premise for reliable supplying and economic running. With the development of the state electrical system, the electric network management modernizes day by day and load forecasting is one of the most important tasks of the modern electrical system operation research, which arouses increasing more and more interests from researchers and is a major foundation for the research of electrical system planning and electrical system economic operation and automatic dispatch. Therefore, finding an appropriate load forecasting method to improve the accuracy of precision has important application value.But non-mathematical model based on neural network provide a new way for solving the problem of mathematical model.Therefore, this thesis proposes a method which combined simulated annealing (SA) algorithm and support vector machine (SVM), SA algorithm is used to optimize the parameters of SVM and acquire simulated annealing support vector machine model, denotedas SA-SVM model.
Keywords: SUPPORT VECTOR MACHINES (SVM), ELECTRICITY LOAD FORECASTING, SIMULATED ANNEALING ALGORITHMS (SA), ADAPTIVE SIMULATED ANNEALING ALGORITHM.

1. Introduction

In today's increasingly complex development of electrical system, with the commercialization and marketization of electric electrical system, the accuracy of power load forecasting of electrical system safe and economic operation and national economic development is of great significance. Various traditional load forecasting technology has more and more difficult to satisfy the load forecasting precision of the power sector is higher and higher requirements, so a new algorithm was used for power load forecasting, improving the accuracy of load forecasting, which has the very vital significance. Studies have shown that no accurate power load forecasting will greatly increase the operation cost of electrical system [1]. As early as 1985, a study in the UK showed that [2], when the forecasting load error of electrical system increases at one percent, the losses will reach ten million pounds a year., therefore, the power load forecasting is accurate or not directly is linked with the benefit of the electrical system, power load forecasting has being paid more and more attention to.

2. Loda Forecasting of Power System

2.1. The Content of the Power System Load Forecasting

The level of power load forecasting work has become the one of the significant signs either the management of electric power enterprise is towards modernization. Load forecasting of electrical system operation scheduling is a very important content, which is an important module of the energy management system (EMS), and is the base to ensure safe and economical operation of electrical system and realize the scientific management and scheduling of power grid. Electrical system is composed of power grid, power users. Its mission is to provide the users uninterrupted economical, reliable electricity which conforms to the quality standard, to meet the needs of all kinds of load, power for social development. Due to the particularity of the production and use of electricity, means that the electricity is hard to be stored in a large amount, and the demands of all kinds of users of electricity change all the time, which requires the power output of the system should be ready to follow the changeable dynamic balance of system load, namely system is to maximize the equipment capacity, to keep the system stable and efficient operation, to meet the needs of users. Otherwise, it will affect the quality of power supply, and even endanger the security and stability of the system. Therefore, the electrical system load forecasting technology has developed, which is the premise and foundation for all this to go smoothly.

Load forecasting of electrical system refers to setting off according to the known demand situation of the electrical system, under the guidance of correct theory, and on the basis of mastering a large number of detailed historical data and analysis through the investigation and study, with the use of reliable methods and means, finally estimating and forecasting on the development trend of power load in the future within a certain period of time. Power system load forecasting can be divided into annual forecast, monthly forecast and daily prediction according to different deadlines, they can also be divided into long-term, medium-term, short-term and ultra-short-term load forecasting from the classification of according to big aspects [3].There is no clear line between long-term and medium-term. In general, long-term forecast can be as long as 30 years, the time of medium-term forecast is usually 5 to 6 years. The meaning of the long-term and medium-term prediction lies in: the installation of the new generator set (including the size of the installed capacity, form, place and time) and the

expansion and renovation of grid, all depends on the load forecast of the future after several years. Short-term forecasting refers a few months, weeks, days, hours or less time. The meaning of the short-term prediction is to help determine the fuel supply plan, forecast the operating power plant output request in order to reach the purpose of prior estimate output changes of generator, and arrange the stop and start of each unit in this net rev economically reasonably.. Super short-term prediction level generally refers to hours or minutes, which can make on-line control of power grid, to realize reasonable scheduling of the generating capacity, can satisfy the demands of a given operation, and minimize the cost at the same time. Classification according to the forecast indicators, the contents of the medium and long-term load forecasting can be divided into two parts as power prediction (such as electricity society electric quantity, network power supply and power supply of each business and industry, etc.) and power prediction (such as maximum differential power, minimum power, peak valley, load rate, etc.).

2.2. The Basic Principles of Power System Load Forecasting

Power load forecasting is the activity to predict and judge the trend of the future development and the status according to the change of power load development rules. So its research object is uncertain, random events. And forecasting of power load is to predict the development trend and the condition that will happen of the load, the following introduces the basic principles of load forecasting, which are used to guide work load forecasting.

(1)The knowable principle

That is to predict development rules of the object, its future development trend and status can be known by people, this is the basic foundation of people's prediction activities.

(2)Possibility principle

Since development and change of things are carried on under the action of internal cause and external cause, the effect of change of internal cause and external cause are different, so it is much more possible to develop and change the things. Forecast for specific indicators are various scheme predictions often according to the development and changes of various possibilities.

(3)Continuity principle

It is also called the inertia principle, continuity principle refers that the development of the predicted object is a continuum process; its future development is the continuation of the process. The principle holds that in the process of development and change some original features of things will be kept down, lasted down. There is also inertia existing during the development of power system load changes; the inertia is the main basis for load forecasting.

(4)Similarity theory

In many cases, a thing as the object of forecasting, its present development process and situation may be similar to these of a certain stage in the past, so you can predict future development process and status of objects according to the known development process and conditions. At present, the analogy methods of prediction technology are based on this principle of prediction methods.

(5)Feedback principle

Feedback is to make use of output returning to the input, and then adjust the output. People found in predicting activity practices that when predicting results and the actual value after a period of practice are different, the gap can be used to carry on feedback adjustment for long-term forecast in order to improve the accuracy of the prediction. When feedback adjustment is carried on, it is essential the combination of prediction theoretical value and practical value, first the gap between predicted and actual values and the cause for it should be analyzed. Then according the found reason, change input data and parameters appropriately, to carry on feedback adjustment, and further improve the prediction quality.

(6)Systemic principle

Systemic principle holds that predicting object is a complete system, itself has internal system, and external system formed because of the contact with the outside world things. The future development of the predicted object is a whole dynamic development of the system, and the dynamic development of the whole system is closely related to the interaction and influence between the factors and its each component. Systemic principle emphasizes the best of overall; only the whole system best prediction is the forecast of high quality, which can provide the best prediction schemes for decision makers [4].

2.3. The Procedure of Power Load Forecasting

To forecast the power load should have a basic program, which is to consider how forecasting works. Only the entire program of the load forecasting work is clear can a good prediction job be done.

(1)Determine the prediction purpose and forecast plan.

The purpose of load forecasting should be clear and specific, which should be closely linked with power industry actual need, and a load forecasting work plan should be written accordingly.

(2)Survey data and choose data

Multi-aspect survey to collect data should be taken, including the internal and external data of power enterprise, data of the national economy of relevant departments, and published and unpublished data, and then pick out the useful part from several data. The standards of selecting materials: one of them is directly pertinence, second is reliability, and third is the latest. The selection should in accordance with the three points at first, and then in-depth study should be taken, considering whether it is need to collect other information after that.

(3)Data compilation

Data compilation purpose is to ensure a complete and accurate data, so as to lay a foundation for improving the prediction precision.

(4)Establish prediction model

Load forecasting model is varied, in order to apply to the data of different structures. Therefore, there is problem for selecting the appropriate forecasting model as to a concrete prediction instance. Correct selection of forecasting model is a key step in the load forecasting.

(5)Write forecast report and bring into service

According to the analysis and judgment of the finally determined forecast results, the report of load forecasting can be written. Because the prediction results are always multivariate, detail descriptions should be given on the forecast condition, assumption and limiting factors etc. which gain the results in the repot. Data information, report analysis, mathematical model, prediction result and essential charts,

(6)Management of load prediction

Submit the load prediction report to competent authorities, but it only shows this prediction is over, which can't be considered as the end of whole prediction work. Then tested should be taken according to the change of subjective and objective conditions and information used by the prediction, the predicted value should be amended when it is necessary.

3. Support Vector Machine (SVM)

Support Vector machine (SVM) as a new method of nonlinear time series analysis and forecasting is proposed by Cortes and Vapnik in 1995 [5], it is based on the statistical theory, and it is still at the stage of development at present. Compared with the traditional methods of artificial neural network it has the advantage that support vector has a certain insensitivity and does not depend on the choice of the model on different methods. Through narrow the upper bound of the generalization error, minimize the structural risk of model, the SVM improves the generalization ability of the models [5].This makes the SVM can be used as the ideal predictor of time

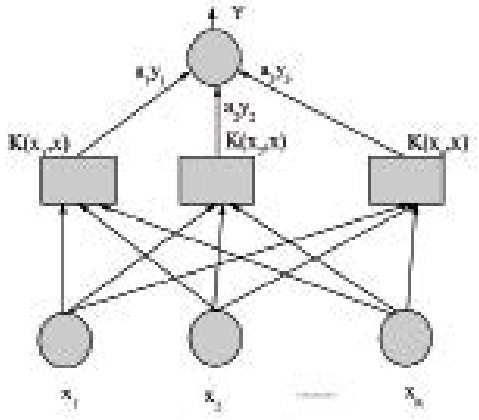

Figure 1. Structure Diagram of SVM

series analysis. SVM is a kind of new machine learning method that studies under the condition of finite samples based on statistical learning theory of VC dimension theory and on the base of the structural risk minimization principle, this method uses kernel function to map sample to high-dimensional feature space and structure the optimal linear classification hyper plane in this space, to achieve maximum generalization ability [6].

Assume that $(x_i, y_i)_{i=1,2,\ldots,n}$, $x_i \in R^d$, $y_i \in \{-1,+1\}$as sample training set, of which $x_i \in R^d$ express dimension feature vector of d dimensional, $y_i \in \{-1,+1\}$ express the type of feature vector x_i belongs to, n is the sample number. In the case of nonlinear, nonlinear transform $\phi(\cdot)$ will be used to transform the original space of sample set, two classification problem of sample space is expressed as:

$$y_i\left(\omega \cdot \varphi\left(x_i\right)+b\right)-1 \geq 0 \quad i=1,2,\ldots,n \qquad (1)$$

In the formula: the ω is the weight vector, b is the deviation.

In addition, considering there are some sample types classification hyper plane that cannot be divided right. Therefore, the introduction of nonnegative relaxation factor ε_i to specified maximum classification interval and minimum classification error sample. Regularization constant C'' decides the punishment degree of the wrong classification of sample error [7]. Therefore, the problem of the optimal separating hyper plane is transformed into:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\varphi\left(x_i\right) \cdot \varphi\left(x_j\right)\right)$$

$$s.t. \sum_{i=1}^{n} \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, i = 1,2,\ldots,n \qquad (2)$$

In type (2), there are only a small fraction of α_i is not zero, the corresponding training sample is support vector, and then optimal decision function is:

$$f(x)=\mathrm{sgn}\left(\sum_{i=1}^{n}\alpha_i y_i K(x\cdot x_i)+b\right) \quad (3)$$

In the formula, K(x·xi)=(x) (xi) is kernel function, and sgn is symbolic function, the results are +1 or -1, n is the number of support vector, this paper select radial basis kernel function. Regularization parameter and kernel function parameters' C sigma selection directly affects the size of the classification accuracy of SVM, so the reasonable selection of parameters is very important for the SVM theory. But the existing methods of selecting SVM's parameters are tedious, time-consuming and the result is not good. Therefore, this article puts forward to use the simulated annealing (SA) algorithm for SVM parameters.

Due to the problem set is a convex set, global optimal solution can be got. In addition, due to some Lagrange multipliers can be zero, not all of the training points are meaningful for the final result. If the training points are removed in advance, the same conclusion can be obtained. Training set with a nonzero Lagrange multiplier is called support vector which plays a crucial role for the conclusion. The fewer number of support vector, the more reasonable conclusion, less calculation amount and more ideal result will be got.

SVM is known as the first choice to solve the problem of nonlinear because of the uniqueness of its solution, the sparseness and good generalization of solution. The traditional artificial intelligence technology is based on empirical risk minimization principle, which is easy to cause over-fitting and under-fitting, and the SVM is based on structural risk minimization principle, it is not only controlling the training error, but at the same time controlling the complexity and generalization risk of model. So compared with the traditional method, the SVM can get better performance, this article through improved simulated annealing algorithm to optimize the parameters of SVM, then uses well-optimized SVM to realize the prediction of grid resources.

4. Simulated Annealing Algorithm

4.1. Introduction of Simulated Annealing (SA) AlgorithmMain text

Simulated Annealing (SA) algorithm is put forward by Metropolis etc., Kirkpatrick etc. used it in combinatorial optimization in 1983 [8-9].The SA algorithm is a kind of random optimization algorithm based on the Monte Carlo iterative solution strategy, its starting point is based on the between similarity physical annealing process of the solid matter and general combinatorial optimization problem. Simulated annealing algorithm under a certain initial temperature, accompanied by falling of the temperature parameters, combined with the kick characteristics of probability in the solution space of random to find the global optimal solution of the objective function, namely the local optimal solution which can jump out of the probabilistic and ultimately tend to the global optimal. The general form of SA is: starting from the selected initial state, in the use of controlling a series of "Markova chain" produced when parameter T declines, use a new generated state and acceptance criteria, to repeat including "produce new state - compute cost function value of the target - determine whether can accept the new state – accept (or abandon) new state" these four steps process, constantly iterate on the current state and target optimal solution is obtained.

Generally speaking, the simulated annealing algorithm includes two loops: an inner loop and an outer loop. Inner loop is perturbing many times under the same temperature to produce different states, and receives a new state according to the Metropolis probability criterion, so the inner loop is controlling through the frequency of state disturbance. Outer loop includes the SA of temperature drop (control through annealing rate r), the condition of algorithm and iteration times I incremental and the stop of algorithm, so the external loop is basically controlled by the number of iterations.

4.2. Advantages of Simulated Annealing (SA) Algorithm

Compared with the previous optimization algorithm, simulated annealing algorithm has the characteristics that easy to descript, flexible to be used, widely used, high efficiency of running and less restricted by the initial conditions, and it is especially suitable for parallel computing. So it not only has the very high practical value, but also the important theoretical significance of parallel computing.

Through the analysis of the experiment of simulated annealing algorithm performance, it can be concluded that simulated annealing algorithm has the following advantages:

(1)Efficiency: as there are no mandatory requirements of simulated annealing algorithm in the selection of initial solution and no, and in the case of the initial solution of random, more ideal optimization solution can be gotten at the end of the algorithm, so we can reduce a lot of work used in the choice of initial solution. Due to the use of shorter time, simulated annealing algorithm can obtain the ideal optimization solution; it is also the more genius point than other traditional optimization algorithm.

(2)Robustness: for different practical optimization problems, the influences on the time efficiency and

the algorithm performance of simulated annealing algorithm are not very big.

(3)The flexibility and versatility: for the solving of most optimization problems, simulated annealing algorithm can be used to achieve the goal of the optimization results. Just as a result of the above features that simulated annealing algorithm has, this paper uses simulated annealing algorithm to optimize the parameters of SVM, to take advantage of the optimal SVM to forecast the grid resources and procedures.

4.3. Parameter Selection of Simulated Annealing Algorithm

Simulated annealing algorithm is a common kind of optimization algorithm based on the Monte Carlo iterative solving portfolio strategy; it is based on fixed material in physics, and finds out the similarities between and on the premise of annealing process and general combinatorial optimization. The general application form of simulated annealing algorithm is that: starting from the selected initial state, in the use of control the decreasing of parameters in a series of "Markova chain", use the generated new state and acceptance criteria to repeat including "producing new status, calculating cost function value of target, determine whether to accept the new rules, accept(or leave) new state" the four-step process, continuously iterate on the current status, to obtain the optimal solution of goal [10]. Main ideas of simulated annealing algorithm are as follows:

(1)Initialization, T represents the initial temperature, and S represents the state of initial solution, and iterations L displayed by different T values are different;

(2)As to k = 1,...L, complete steps (3) and (6);

(3)Promote the produce of new explanation of S';

(4)Complete the incremental calculation Δt '= C(S')C(S), C(S) here represents the evaluation function;

(5)If Δt<0, S 'can be accepted to represent the current new explanation, otherwise probability exp(t/T) will be used to complete the new explanation of S';

(6)If the prerequisite conditions are met, then the output value is the optimal solution, the program is completed. Termination condition is to use different continuous value as a new explanation commonly.

(7)T constantly decreases and T ≥ 0, then turn into the step (2).

5. Empirical Research on Power Load Forecasting

5.1. Measure of Power Load Forecasting Model Based on Generalized Regression Neural Network

As to generalize regression neural network model, different parameters are adapted to test, and then compare the MAPE values obtained model. The MAPE values of generalized regression neural network under different parameters is as shown in Figure 5-3. We can see from the figure that when the value of σ is greater than 0.41, MAPE values of the model will begin to increase, therefore, in the load forecasting, σ= 0.41 will be used. When σ= 0.41 generalized regression neural network model of prediction results as shown in Table 1:

Table 1. Prediction Results of Generalized Regression Neural Network Model

Year	GRNN
1999	3216449278
2000	3494288704
2001	3267446989
2002	3598033042
2003	3473339319
2004	3525524067

5.2. Measure of Power Load Forecasting Model Based on Autoregressive Smooth Movement Model

As to ARIMA model, the parameter selections are mainly p, d, q, here we use SPSS to analysis the electric power load data from 1949 to 1998.Observing the related figure of events sequence: ACF drops rapidly; Second, PACF lag sharply decreases at first, and all PACF on the lag 1 are not significant statistically, thus, we determine that the time sequence is smooth, no need to proceed a difference, therefore, d =0, q =1, p =2. After the parameters is determined, estimation model for residual analysis, results show that the estimation model of Figure ACF and PACF are not significant in statistics.

When the d=0, q=1, p=2 the predicted results obtained from smooth movement model are as shown in Table 2:

Table 2. Prediction Results of ARIMA Model

YEARS	ARIMA
1999	3060211746
2000	3221928908
2001	3075910345
2002	3276127707
2003	3283799093
2004	3304514028

5.3. Comparison and Analysis of Prediction Result

This paper uses MAPE, MAD and NRMSE to carry on the comparison of model prediction accuracy of. MAPE, MAD and NRMSE values of each model is presented in Table 3.

Table 3. Comparison of Prediction Results of Models

	ARIMA	GRNN	SVMSA
MAPE	10.42%	5.22%	1.82%
MAD	12.781	5.687	2.378
NRMSE	0.108232	0.043893	0.026648

From the table we can see that in the same forecast period, there are minimum values of MAPE, MAD and NRMSE of simulated annealing support vector machine (SVM) model built in this paper. From the above numerical comparison of prediction precision test indicators, you can see that the accuracy test rating of the simulated annealing method of support vector machine (SVM) constructed by this article is significantly better than that of autoregressive moving average model and generalized regression neural network model.

These two tests show that the predictive results of the simulated annealing algorithm of support vector machine (SVM) constructed by this paper is significantly better than the prediction results of the other two models.

6. Coclusion

Power system load forecasting is of important significance for electric power industry and even the country's national economy. But there are still many problems exist need to be solved of support vector machine (SVM) theory in practice. Especially in the case of the parameter selection of support vector machines (SVM). Support vector machine (SVM) of parameter selection has a great influence on the performance of support vector machine (SVM). In this paper, by introducing a simulated annealing algorithm for parameter selection, simulated annealing support vector machine (SVM) model was constructed. In this paper the simulated annealing algorithm of support vector machine (SVM) is applied to implement the study of power load forecasting is far from over, so to achieve practical application also requires a lot of research and development work.

Acknowledgements

The research of this paper has been sponsored by Henan Provincial Research Foundation for Science and Technological Breakthroughs, China (Grant No. 132102210179; 142102210597), Natural Science Foundation of Henan Educational Committee, China (Grant No.2011B510021).

References

1. Gross G, Galiana FD. Short term load forecasting. Proc IEEE 1987,(75):1558–73.
2. Bunn DW, Farmer ED. Comparative models for electrical load forecasting. New York: John Wiley; 1985.
3. Niu Dongxiao, Cao Shuhua, Zhao Lei, Zhang Wenwen. Forecast technology and its application on power load. Beijing: China Electric Power Press, 1998.
4. Liu Chenhui. Theory and method of the load forecast on electric power system. Harbin: Harbin Institute of Technology, 1987
5. Lee L T, Tao D F, Tsao C. An adaptive scheme for predicting the usage of grid resources[J]. computers and Electrical Engineering archive. 2007, vol. 33, no. 1 , p.p.1-11.
6. Li Lingjun, Zhang Zhousuo, He Zhengjia. Study on the application of SVM on mechanical fault diagnosis. Computer Engineering and Application, 2002, vol.38, no.19, pp.19-21.
7. Deng Naiyang, Tian Yingjie, New method of data mining-SVM, Beijing: Science Press, 2004
8. N. Metropolis, A. Rosenbluth, and M. Rosenbluth,et al. Equtation of state calculations by fast computing machines. Journal of Chemical Physics, vol. 1953, no. 21, pp. 1087-1092.
9. S. Kirkpatrick, Jr. CD. Gelat, and MP. Vecchi. Optimization by simulated annealing. Science, vol. 1983, no. 220, pp. 671- 680.
10. Zou Xiaojin, Research on transformer condition monitoring based on simulated annealing algorithm [J]. Application of Power technology,2012, vol.14, no.3 p.p. 45-49.