Next Day Load Forecasting Using SVM

Xunming Li, Dengcai Gong, Linfeng Li, and Changyin Sun

College of Electrical Engineering, Hohai University, Nanjing, Jiangsu 210098, China
cysun@ieee.org

Abstract. Based on similar day method and SVM, this paper proposes a new method for next day load forecasting. The new method uses the parameters of several similar days, instead of only selecting one similar day as in similar day method. The parameters of selected similar days are used as inputs to SVM for forecasting the loads of 24 points (one hour per point) of the next day. The method behaves the advantages of both similar day method and SVM method, Corresponding software was developed and used to forecast the next day load in a practical power system and the final forecasting error is low.

1 Introduction

Load forecasting plays an important role in power system planning and operation. Basic operation functions such as unit commitment, economic dispatch, fuel scheduling and unit maintenance can be performed efficiently with an accurate forecast.

A wide variety of methods have been proposed in the last two decades owing to the importance of load forecasting, such as linear regression, exponential smoothing, stochastic process, data mining approach [1]-[4]. However, load forecasting is a difficult task as the load at a given hour is dependent not only on the load at the previous hour but also on the load at the same hour on the previous day, and on the load at the same hour on the day with the same denomination in the previous week. Generally, these methods are based on the relationship between load and factors influencing the load. However, the techniques employed for those models use a large number of complex and nonlinear relationships between the load and factors influencing the load. The traditional prediction methods are difficult to estimate these nonlinear relationships. Therefore, some new forecasting models have been recently introduced as expert systems, artificial neural networks (ANN), and fuzzy systems. Among these different techniques of load forecasting, application of ANN technology for electric load forecasting has received much attention in recently years. The main reason of ANN becoming so popular lies in its ability to learn complex and nonlinear relationships that are difficult to model with conventional techniques. However, there are some disadvantages of ANN method such as network structure is hard to determine and training algorithm has the danger of getting stuck into local minima.

Recently, a novel type of learning machine, called support vector machine (SVM), has been receiving increasing attention in areas from its original application in pattern recognition to the extended application of regression estimation [5]. This is brought about by the remarkable characteristics of SVM, such as good generalization performance, the absence of local minima and sparse representation of solution. One key

J. Wang, X. Liao, and Z. Yi (Eds.): ISNN 2005, LNCS 3498, pp. 634–639, 2005.

characteristic of SVM is that training SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of SVM is always unique and globally optimal, unlike ANN' training which is time-consuming and requires nonlinear optimization with the danger of getting stuck into local minima. Recently, there are also a great deal of researches concentrating on applying regression SVM to short-term electric load forecasting [6],[7],[8]. This paper proposes a new method for next day electric load forecasting, based on similar day method [9] and SVM method. Unlike the similar day method which only uses a similar day with the minimal value of difference evaluation function, here, a series of days are selected as similar day when the value of difference estimation function is less than a given value. The parameters of these selected similar days are used as samples data to train SVM. Finally, the correction should be done with the forecasted value when it violates the general rule. This method is used to forecast the next day load of a practical power system in a week. The experimental result shows that the proposed method performs well on both the forecasting accuracy and the computing speed.

2 SVM for Regression Estimation

Given a set of data points $\{(X_i, y_i)\}_i^N$, ($X_i \in \mathrm{R}^n$, $y_i \in \mathrm{R}$, N is the total number of training sample)randomly and independently generated from an unknown function, SVM approximates the function using the following form:

$$f(X) = <\omega, \varphi(X)> + b \quad . \tag{1}$$

where $\varphi(X)$ represents the high-dimensional feature spaces which is nonlinearly mapped from the input space X. The coefficients ω and b are estimated by minimizing the regularized risk function (2):

$$\text{minimize} \ \frac{1}{2}\|\omega\|^2 + \mathrm{C}\sum_{i=1}^{N}\left|y_i - <\omega, \varphi(X_i)> - b\right|_\varepsilon \quad . \tag{2}$$

$$\left|y_i - <\omega, \varphi(X_i)> - b\right|_\varepsilon = \begin{cases} 0 & \left|y - <\omega, \varphi(X)> - b\right| < \varepsilon \\ \left|y - <\omega, \varphi(X)> - b\right| - \varepsilon & \left|y - <\omega, \varphi(X)> - b\right| \geq \varepsilon \end{cases} \quad . \tag{3}$$

The first term $\|\omega\|^2$ is called the regularized term. Minimizing $\|\omega\|^2$ will make a function as flat as possible, thus playing the role of controlling the function capacity. The second term $\sum_{i=1}^{N}\left|y_i - <\omega, \varphi(X_i)> - b\right|_\varepsilon$ is the empirical error measured by the ε - insensitive loss function(3). This loss function provides the advantage of using sparse data points to represent the designed function (1). C is referred to as the regularized constant. ε is called the tube size. They are both user-prescribed parameters and determined empirically.

To get the estimation of ω and b, (2) is transformed to the primal objective function(4) by introducing the positive slack variables $\xi_i^{(*)}$ ((*) denotes variables with and without *)

$$\text{minimize } \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

subject to

$$\begin{aligned} & y_i - <\omega, \varphi(X_i)> - b \le \varepsilon + \xi_i \\ & <\omega, \varphi(X_i)> + b - y_i \le \varepsilon + \xi_i^* \\ & \xi_i^{(*)} \ge 0 \qquad i = 1, \ldots, N \end{aligned} \quad . \tag{4}$$

Final, by introducing Lagrange multiplier and exploiting the optimality constraints, the decision function (1) has the following explicit form:

$$f(X) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)K(X_i, X) + b \quad . \tag{5}$$

In function(5), $\alpha_i^{(*)}$ are the so-called Lagrange multipliers. They satisfy the equalities $\alpha_i \times \alpha_i^* = 0$, $\alpha_i \ge 0$, and $\alpha_i^* \ge 0$ where i =1.... N ,and they are obtained by maximizing the dual function of (4),which has the following form:

$$W(\alpha_i, \alpha_i^*) = \sum_{i=1}^{N} y_i(\alpha_i - \alpha_i^*) - \varepsilon\sum_{i=1}^{N}(\alpha_i - \alpha_i^*) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(X_i, X_j) \cdot \tag{6}$$

subject to

$$\sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0. \qquad 0 \le \alpha_i, \alpha_i^* \le C. \quad i = 1 \ldots\ldots N$$

$K(X_i, X_j)$ is defined as the kernel function. The value of the kernel is equal to the inner product of two vectors X_i and X_j in the feature space $\varphi(X_i)$ and $\varphi(X_j)$,that is, $K(X_i, X_j)= <\varphi(X_i), \varphi(X_j)>$. The elegance of using the kernel function that one can deal with feature spaces of arbitrary dimensionality without having to compute the map $\varphi(X)$ explicitly. Any function that satisfies Mercer's condition can be used as the kernel function. Common examples of the kernel function are the polynomial kernel $K(X_i, X_j)= (<X_i,X_j>+1)^d$ and the Gaussian kernel $K(X_i, X_j) = \exp(-(1/\sigma^2)(X_i - X_j)^2)$, where d and σ are the kernel parameters.

From the implementation point of view, training SVM is equivalent to solving the linearly constrained quadratic programming problem (6) with the number of variables twice as that of the number of training data points. The sequential minimal optimization (SMO) algorithm extended by Scholkopf and Smola is very effective in training SVM for solving the regression estimation problem. In this paper, an improved SMO algorithm [10] is adopted to train SVM.

3 Method Based on Similar Days and SVM

In short-term load forecasting, loads of two days are generally very close when they have similar weather condition, same day class (workday or weekend) and several other similar factors. From this view, the similar day method has been developed. The advantages of the method are simple, practical and comparatively precise. However, this method has the following disadvantages. First, the selected similar day is not certainly the day which has the closest load to the forecasting day, sometimes the disparity is large. Second, due to the nonlinear characteristic of the relationship between all influencing factors and the load, it is difficult to get a good result and stability by using curve fitting or experiential method to correct the load.

Toward the first problem in similar day method, the load of selected similar day with the least value of difference estimation function is not always close to the load of the forecasting day, But load of some days with comparatively small value of difference estimation function are also possible close to the load of forecasting day. Therefore, in this paper, a series of similar days are selected rather than only the most similar day in similar day method.

Toward the second problem in similar day method, it is hard to fit the load curve using traditional method. This paper applies regression SVM to describe the complex nonlinear relationship between influencing factors and load. To a given forecasting day, it is hard to find out a large number of similar days. SVM is a good learning machine based on SRM principle, and has the capacity to solve the small samples learning problem with a good generalization performance.

The specific steps are presented in following:

First, transforming the non-numerical factors into numerical form. Taking sunlight for example, fine is set as 3, cloudy as 2, rain as 1. The maximal and minimal temperature can apply the actual value.

Second, treating disorder sample. Considering the proportion of the loads in serial time, it is sure appearing disorder data when the change of load violates the general rule. Checking every point according to this principle to pick out all disorder data and correct them.

Third, selecting similar days from four weeks before the forecasting day. When the value of difference estimation function is less than a given value γ, this day will be selected as a similar day, γ is an experiential value.

Fourth, taking parameters of these similar days as training sample to train SVM. Due to the restriction that value of difference estimation function should be less than γ, influencing factors in similar day and forecasting day are close in vector space. It is effective to make use of the generalization capacity of SVM to estimate the nonlinear relationship between load and influencing factors.

Finally, taking the influencing factor vector of forecasting day into the trained SVM, the 24 points load of next day will be forecasted. In the last, the forecasted load also should be corrected. Because load influenced by a series of uncertain factors, it is hard to get a satisfy result absolutely using historical data. Operators should correct the forecasting result from experience. Sometimes, the change of forecasted load violates the regular pattern, then this result is insecure usually, which can be instead by the mean of fore-and-aft loads.

4 Experimental Results

The example data is historical load data in a practical electric network in May and June in 2001. The data includes the data of weather condition and the data of 24 points data in every day. Using these data, 24 load points in a day are forecasted. This paper adopts the improved SMO algorithm to train SVM and RBF kernel is selected as the kernel function.

There are many error evaluation indexes to evaluate the result of daily load forecasting. In this paper, four relative error indexes are selected to evaluate the forecasting result of the proposed method. The four error indexes of forecasting result are represented in table 1.

Table 1. Relative errors of experimental result

Day	E_{MAPE}	E_{MSE}	$E_{\max}$	$E_{\min}$
Monday	3.25	3.92	0.33	2.03
Tuesday	2.85	3.30	1.36	2.08
Wednesday	2.34	2.90	3.13	0.61
Thursday	3.46	3.72	3.02	4.01
Friday	3.08	3.35	1.36	4.80
Saturday	2.99	3.43	4.90	3.29
Sunday	1.90	2.32	1.07	0.60

Table 1 indicates that the maxima of average relative error E_{MAPE} is 3.46 and the minima is 1.90. The forecasting precision of this new method is very high. The mean square root relative error E_{MSE} is bigger than average relative error E_{MAPE} which means that the relative error of forecasting value is large in some points. The forecasting models in these time points should be improved.

5 Conclusions

This paper proposed a new approach for short-term load forecasting based on similar day method and SVM. This new method uses a series of similar days as training data for SVM, which utilizes the information of all these similar days which ensures stabilization and precision of the algorithm. There does not exist the problem of correcting algorithm using SVM to fit the nonlinear relationship between load and influencing factors. Combined with similar day method, new method still keep the characteristic of easy application and effectiveness and the training time of SVM is very short. Experimental result shows that this method is an effective method of high application value for short-term load forecasting.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 50595412, China Postdoctoral Science Foundation, Jiangsu Planned Projects for Postdoctoral Research Funds and a start-up grant from Hohai University of China.

References

1. Papalexopoulos, A.D., Hesterberg, T.C.: A Regression-based Approach to Short-Term Load Forecasting. IEEE Trans. Power Syst, **5** (1990) 1535-1550

2. Haida, T., Muto, S.: Regression Based Peak Load Forecasting Using a Transformation Technique. IEEE Trans. Power Syst., **9** (1994) 1788-1794
3. Rahman, S., Hazim, O.,: A Generalized Knowledge-based Short Term Load-forecasting Technique. IEEE Trans. Power Syst, **8** (1993) 508-514
4. Wu, H., Lu, C.: A Data Mining Approach for Special Modeling in Small Area Load Forecast. IEEE Trans. Power Syst, **17** (2003) 516-521
5. Vapnik, V.N., Golowich, S.E., Smola, A.J.: Support Vector Machine for Function Approximation, Regression Estimation and Signal Procession. Adv. Neural Information Procession Syst, **9** (1996) 281-287
6. Zhao, D.F., Wang, M.: A Support Vector Machine Approach for Short Term Load Forecasting. Proceedings of the CSEE, **17** (2002) 26-30
7. Yang, J.F, Cheng, H.Z.: Application of SVM to Power System Short Term Load Forecasting. Electric Power Automation Equipment, **24** (2004) 30-32
8. Li, Y.C, Fang, T.J., Zhang, G.X.: Wavelet Support Vector Machine for Short-term Load Forecasting. Journal of University of Science and Technology of China, **3** (2003) 726-732.
9. Cheng, S.: A New Approach to Load Forecasting Based on Similar Day. Proceeding of Jiangsu Electrical Engineering Association, **18** (1999) 28-32
10. Zhang, H.R., Han, Z.Z.: An Improved Sequential Minimal Optimization Learning Algorithm for Regression Support Vector Machine. Journal of Software, **14** (2003) 2006-2013.