The Selforganizing Neural Network Approach to Load Forecasting in the Power System

S. Osowski, K. Siwek
Institute of the Theory of Electrical Engineering and Electrical Measurements
Warsaw University of Technology, Warsaw, POLAND
email: sto@iem.pw.edu.pl

Abstract

The paper presents the neural network approach to the prediction of 24-hour load consumption in the power system. Two different structures are studied: the conventional Kohonen network and the extended form of selforganizing network taking into account the activity of both winner and neighbouring neurons (so called fuzzy selforganizing network). These methods have been combined with the multilayer perceptron approach to the prediction of mean and variance of the load. Thanks to such solution the average accuracy of the power consumption prediction for the power system may be greatly improved. The method based on selforganization is universal, flexible and easy for use in any power system.

Introduction

The improvement of the system economy is the motivation for the daily forecasting of the load of the electric power system in the country. A lot of methods have been developed for this purpose [?, ?, ?, ?, ?, ?]. The most known are the methods based on the statistics and linear regression approaches. Recently a great interest is in the application of neural networks. Multilayer perceptron, recurrent networks as well as selforganizing neural networks are the most important representatives of this group. Recently [?] we have proposed the method based on the Kohonen network used for 24 hour load prediction. This paper is the continuation and generalization of this research.

First we will propose the improved method of the prediction of the mean and variance of the daily load. It is based on the application of the multilayer perceptron and supervised learning. Thanks to this the accuracy of prediction may be significantly improved. Secondly we propose more advanced profile prediction techniques. Instead of simple Kohonen network representing the output signal by the winner only we propose here the extension of it, taking into account the activity of not only the winner but also the neighbouring neurons. The results of the numerical experiments, concerning prediction of the load for the Polish Power System will be given and discussed in the paper. They show the significant improvement of the proposed method in comparison to the previously used.

General description of method

The selforganizing neural network is a single layer neural network, working in a competitive mode. The detailed treatment of this structure as well as the variety of the learning algorithms can be found in many textbooks and papers [?, ?] and will not be mentioned here.

The idea of applying selforganizing networks for load forecasting has been presented, among others in [?, ?]. To make the prediction independent on the general trend, changing from year to year, the input data (the real consumption of power at different hours of the day) is transformed to the so called profiles $p(d,t)$ [?] by cutting out the mean value and dividing the result by the variance of the data for this day. The profiles are defined in the following way

$$p(d,t) = \frac{P(d,t) - P_m(d)}{\sigma(d)} \tag{1}$$

where $P(d,t)$ is the real load of dth day at tth hour, $P_m(d)$ is the mean value of the load of dth day and $\sigma(d)$ is the variance of the power consumption of this particular day. The days of the same type belonging to the same seasons of the year have similar profile characteristics and form clusters, grouping the similar data. Each cluster is represented by one neuron acting in the competitive mode.

The learning data for the selforganizing network is composed of a set of profiles of the past years. The network is trained using one of the learning algorithms (neural gas, Winner Takes All with conscience mechanism, classical Kohonen algorithm [?]) and as a result of learning each neuron of the net represents the data closest to its weight vector in the chosen metric space. This is so called classical Kohonen neural approach. In more advanced representation of the data we take into account not only the activity of winner but also of the neighbouring neurons. In such case the data is represented by the weights of the winner and some limited number of neurons closest to the winner, while each representation is weighted by the so called membership function, like it is in the fuzzy systems.

In either case after learning, the weights of the neurons are frozen and the results of analysis memorized. The system is ready for the retrieval mode. If we now want to make prediction for 24-hour load of the particular day of the year we simply take the reversed form of equation (??). Substituting $p(d,t)$ by its estimation $\hat{p}(d,t)$, delivered by the selforganizing neural network, we get

$$P(d,t) = \sigma(d)\hat{p}(d,t) + P_m(d) \tag{2}$$

However two problems remain to be solved now. One is the accurate prediction of the profiles on the basis of activities of neurons of the neural network and the second – good estimation of the mean and variance for the particular day.

Prediction of the daily mean and variance

The correct prediction of the mean and variance for the particular day is an important factor in accurate prediction of the load. Our previous solutions [?] were based on either the mean of these values for the same types of day in the past or on the linear regression model. In linear model of prediction we take into account the appropriate, values from the last 4 years. Let us denote by $P_m(d,y)$ the mean load of dth day of the yth year. Then the linear prediction of it is given in the form

$$\hat{P}_m(d,y) = ay + b \tag{3}$$

The coefficients a and b are obtained from the solution of quadratic programming problem, for which the minimized cost function is defined as

$$E = \sum_{k=i}^{i+3} \left[(ak+b) - P_m(d,y)\right] w_k \tag{4}$$

In this formulation w_k is the weighting coefficient adjusted for each year, where we have taken 4 last years in definition of E. Solving this quadratic problem we get

$$b = \frac{\sum_{k=i}^{i+3} k w_k \frac{\sum_{k=i}^{i+3} k P_m(d,y) w_k}{\sum_{k=i}^{i+3} k^2 w_k} - \sum_{k=i}^{i+3} P_m(d,y) w_k}{\sum_{k=i}^{i+3} k w_k \frac{\sum_{k=i}^{i+3} k w_k}{\sum_{k=i}^{i+3} k^2 w_k} - \sum_{k=i}^{i+3} w_k}$$

$$a = \frac{\sum_{k=i}^{i+3} k P_m(d,y) w_k - b \sum_{k=i}^{i+3} k w_k}{\sum_{k=i}^{i+3} k^2 w_k}$$

Different simulations have been made to find the best values for weighting coefficients w_k. As a result of these experiments we have assumed $w_1 = 1$ for the data of actual year, $w_2 = 0.9$, $w_3 = 0.7$, and $w_4 = 0.5$ for the previous years, respectively. Similar estimation is repeated for the variance σ.

Although the linear fit provides better accuracy of mean and variance prediction it still suffers the limited accuracy. In this paper we propose the nonlinear estimation based on the application of the multilayer perceptron for the prediction of the mean and variance of the day. On the basis of observation of the appropriate values in the data set, we noticed the correlation of present value with the appropriate values from the closest past. In our model we take the input vector to the network containing 9 nodes, representing the mean (variance) of the days of the previous years, actual season of the year and type of the day. The type of the day may be coded in one node (0 - holidays, 1 - workdays). The code for the season takes two nodes. We have assumed: 11 for winter, 01 for spring, 00 for summer and 10 for autumn. The destination is associated with the predicted value of the mean (variance) for dth day. The particular structure of the proposed feedforward neural network for prediction of mean or variance is shown in Fig. 1. The number of hidden neurons has been chosen on the ground of good generalization

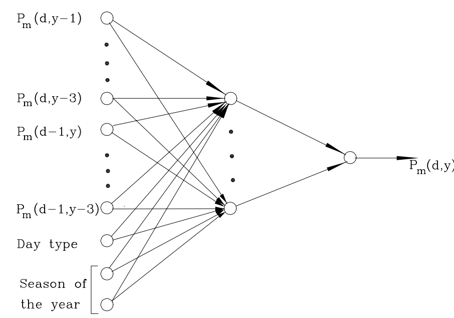

Figure 1: *The structure of the MLP network used for prediction of the mean and variance*

ability requirements of the network. Many numerical experiments have shown that the best number of hidden neurons for this case is equal 5 and this number was applied in prediction.

The network has been trained and tested on the available data of the Polish Power System from the years 1990 - 1995. The data of 1995 has been used only for testing. Fig. 2 presents the results in the form of MAPE (Mean Absolute Percentage Error) for prediction of mean and variance for these years. Fig. 2a is the distribution of MAPE errors for mean and Fig. 2b for the variance. Comparing these results with the classical statistical approach we notice the great improvement, especially for the mean values, where the MAPE error has been reduced below 2.5%.

Prediction of the profiles

In the classical approach presented in the past we estimated the profiles averaging the winner vectors for this particular day from the past, i.e.,

$$\hat{\mathbf{p}}(d) = \frac{\sum_{i=1}^{n} k_{di} \mathbf{w}_i}{\sum_{i=1}^{n} k_{di}} \tag{5}$$

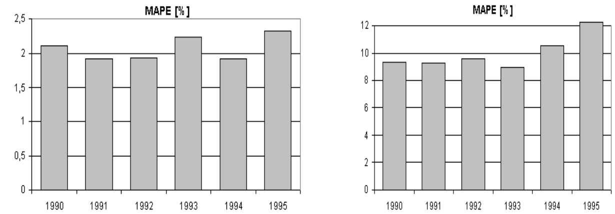

Figure 2: *The MAPE errors of prediction of the mean (left) and variance (right) of the power for Polish Power System within the years 1990 - 1995*

where k_{di} is the number of appearances of ith neuron among the winners in the past for this particular day. In this paper we propose the application of fuzzy selforganizing approach. It is also associated with the one-layer selforganizing network, however this time in prediction process we take into account not only the winner but also the activity of some loosers, closest to the winner. The learning phase is performed in the same way as it was done in the first case (neural gas algorithm). However as a result of learning we memorize not only the winner but also some limited number q (in practical implementation we assumed $q = 5$) of neurons closest to the winner. At the same time we keep also their relative activities, that represent the membership values.

If the activities of the winner and the neighbouring neurons are denoted by u_w and u_i respectively, we introduce their relative activities, that are defined by

$$y_i = e^{-\alpha(u_w - u_i)^2} \tag{6}$$

This value for the winner is $y_w = 1$ and for all other neurons $0 \leq y_i < 1$. The coefficient α is the decaying parameter of this transformation. On the basis of these distributed activities of neurons we can define the memebership value of ith neuron at the presentation of the input vector $\mathbf{x}$ in the form

$$\mu_i = \frac{y_i}{\sum_{i=1}^{q} y_i} \tag{7}$$

The phase of prediction is performed in the similar way as it was done in the first solution, however this time we take into account not only the winners but also their neighbours and their relative acitivities, described by μ_i. As a result the prediction of the profile vector for any dth day is given in the form

$$\hat{\mathbf{p}}(d) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{q} \mu_{di}^{(j)} \mathbf{w}_i^{(j)}}{\sum_{i=1}^{n} \sum_{j=1}^{q} \mu_{di}^{(j)}} \tag{8}$$

The parameter $\mu_{di}^{(j)}$ denotes the membership value of jth neuron taking part in prediction of the load for dth day and i is the notation of the particular day of the previous years taking part in the prediction process. The number of past days influencing the prediction is denoted here by n. The relations (??) - (??) resemble the fuzzy relationships, thus the method is called fuzzy s[illegible]istra[illegible]a-

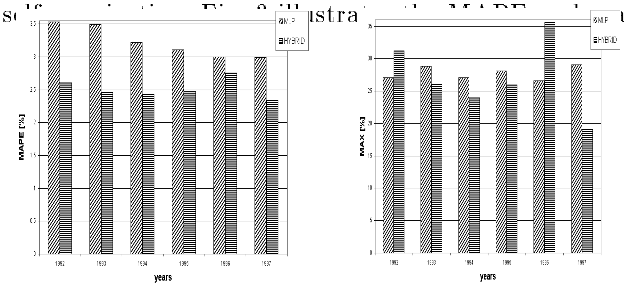

Figure 3: *The MAPE (left) and maximum (right) errors of profile prediction of the power for Polish Power System within the years 1990 - 1995 (CSO - Classical and FSO - Fuzzy SelfOrganization)*

ximum errors of the profile prediction for Polish Power System within the years 1990 − 1995. As it is seen the maximum MAPE error of prediction of profiles using FSO (year 1995) was still below 2.5% and minimum is 1.8%. It is much better than the best results obtained using the classical method of prediction based on the Kohonen network.

Results of numerical experiments

The numerical experiments have been carried out on the data of the Polish Power System and used the real data for years 1986 - 1995. The data have been first converted to the profiles and these profiles have been used in training of the network. The Kohonen network of 100 neurons has been used in experiments. The weights of each neuron or the activity of neurons (in the case of fuzzy selforganizing network) represent the profile vector corresponding to the energy consumption of the particular day of the year.

Table 1: *The MAPE errors of final prognosis*

Date	CSO	FSO
1990	2.64%	2.44%
1991	2.38%	1.84%
1992	2.42%	1.94%
1993	2.73%	2.26%
1994	2.48%	2.14%
1995	2.86%	2.45%

On the basis of predicted profiles the final prognosis for 24 hours is made by applying equation (??). The important point is to estimate the mean value $P_m(d)$ of the power of the particular dth day and the variance $\sigma(d)$ for the same day. These have been done by using the multilayer perceptron approach, presented in the previous section.

Table 1 presents the obtained results in the form of MAPE (Mean Absolute Percentage Error) and table 2 the maximum errors for the years 1990 - 1995. The data corresponding to 1990 - 1994 have been used also in learning phase while 1995 was used only in testing mode. The first column of data denotes the year of prediction, second column - the results of classical selforganizing network (CSO) applying rule (??) for prediction of profiles and the third column - the fuzzy selforganizing network (FSO) based on the equation (??).

Table 2: *The maximum errors of prognosis*

Date	CSO	FSO
1990	23.49%	21.98%
1991	21.03%	19.62%
1992	22.07%	20.48%
1993	20.79%	19.11%
1994	24.03%	20.13%
1995	20.87%	19.87%

As it is seen, in all cases both MAPE and maximum errors have been reduced by applying more sofisticated methods of selforganization. The relative improvement of accuracy is in the range of 7-15%. It is important that the accuracy of prediction is satisfactory from the practice point of view. Also the maximum errors of prediction are acceptable.

Very interesting is the distribution of 24-hour prediction of power for different seasons of the year. Fig. 1 presents chosen curves of the load prediction for four

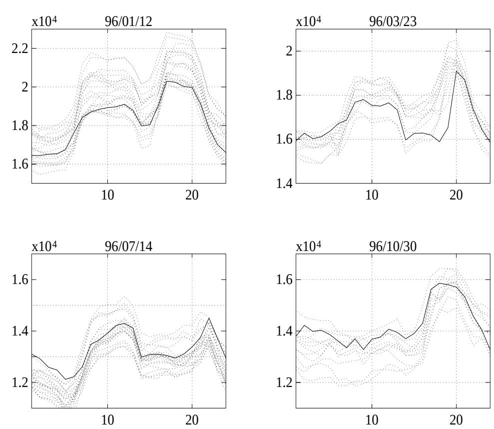

Figure 4: *The load curves of the days forming the base for prediction of the load for 4 different days of the year*

different days belonging to different seasons of the year, taking part in prognoses. The solid lines represent the final prediction and the dot lines show the variability of the loads corresponding to the days taking part in final prediction. The prediction (solid line) is the result of the weighted average of the dotted lines and the actual trend. As it is seen from the figures the prediction of the data follows the real load changes and the discrepancy between both curves is acceptable, especially for workdays.

Conclusions

The paper has presented selforganizing neural network approach to the short term load forecasting in the power system. The obtained average accuracy is above the level of 97% and exceed both classical method and multilayer perceptron approaches. The method based on the so called profiles is universal, flexible and insensitive to the global changes following from the development of the economy of the country.

It can be used for prediction of the energy consumption at any date ahead (even one year ahead). The accuracy of prediction is dependent on the availability of the past data. The network can be retrained any time with the additional actually obtained data base, allowing in this way to refresh the weights and adapt the system to the new conditions of the power system performance.

References

[1] S. Haykin, Neural networks, a comprehensive foundation, Macmillan, 1994, N. Y.

[2] Osowski S., Neural network - algorithmic approach, WNT, (in Polish), Warsaw, 1996

[3] Hagan M. T., Behr S. M., The time series approach to short term load forecasting, IEEE Trans. PWRS-2, 1987, pp. 785-791

[4] Brace M. C., Schmidt J., Hadlin M., Comparison of forecasting accuracy of n.n. with other techniques, Proc. Int. Forum on Appl. of N. N. to power systems, Seattle, 1991

[5] K. Lee, T. Choi, C. Ku, J. Park, Neural network architectures for short term load forecasting, IEEE ICNN, Orlando, 1994

[6] H. Mori, T. Ogasawara, A recurrent n. n. approach to short term load forecasting in electric power systems, IEEE ICNN, Portland, 1993

[7] T. Onoda, Next day peak load forecasting using ANN with modified backpropagation learning algorithm, IEEE ICNN, Orlando, 1994

[8] A. Germond, N. Macabrey, T. Baumann, Application of artificial neural networks to load forecasting, 1992 NATO Conference, Brussels, 1992

[9] S. Osowski, K. Siwek, Neural networks for load forecasting in the power system based on selforganization, Proc. of KKTOiUE, 1998

[10] M. Cottrell, B. Girard, Y. Girard, C. Muller, P. Rousset, Daily electrical power curve: classification and forecasting using a Kohonen map, IWANN, Malaga, 1995, pp. 1107-1113

[11] T. Kohonen, Self-organizing maps, Springer Verlag, 1995, Berlin