APPLICATION OF HYBRID MODEL IN DAILY PEAK POWER LOAD PREDICTION PROBLEM

ỨNG DỤNG MÔ HÌNH HỖN HỢP TRONG DỰ BÁO PHỤ TẢI ĐỈNH CỦA BIỂU ĐỒ PHỤ TẢI NGÀY

Nguyen Quan Nhu
Thainguyen University of Technology

Tran Hoai Linh
Hanoi University of Technology

ABSTRACT

The problem of prediction of P_{max} and P_{min} of power daily load is one of fundamental tasks in power system operation and control. There are already a number of different models and solutions including linear and nonlinear ones. In this paper we propose a new hybrid approach including both linear and nonlinear estimators in order to be able to model the complex dependencies between future power daily loads and the data from the past. The Singular Value Decomposition algorithm will be used to estimate the linear relationship between signals where the classic neural network MLP will be used to estimate the nonlinear features. The proposed solution was test on real data taken from Thai Nguyen Power Company.

TÓM TẮT

Bài toán dự báo đỉnh và đáy của biểu đồ phụ tải ngày là một trong những nhiệm vụ quan trọng của công tác điều độ. Đã có rất nhiều mô hình được đề xuất để thực hiện nhiệm vụ sử dụng các mô hình tuyến tính hoặc phi tuyến. Bài báo này đề xuất giải pháp phối hợp cả hai mô hình tuyến tính và phi tuyến để có thể tái tạo được mối phụ thuộc phức tạp phi tuyến giữa đỉnh và đáy của biểu đồ phụ tải ngày hôm sau vào các giá trị trong quá khứ. Thuật toán SVD (Singular Value Decomposition) sẽ được sử dụng để ước lượng thành phần phụ thuộc tuyến tính, còn mạng nơ-rôn kinh điển MLP sẽ được sử dụng để ước lượng thành phần phi tuyến. Các mô hình được thử nghiệm trên bộ số liệu phụ tải ngày của Điện lực Thái Nguyên đã cho các kết quả tốt.

I. INTRODUCTION

The problem of prediction of the highest and lowest peaks of the daily power load has a very high demand in real world applications. There were already a number of different solutions for this problem but until this moment there is no solution which can be seen as a standard one. The main reason for this lack of standard solution is due to the variety of the power load among different regions and the complex dependencies on different factors from a given region. That's why for each region we should design a separate prediction model. Even in the case we use the same mathematical model, its parameters should be tuned adaptively to the data, so the final models are different.

Among proposed models for the peak prediction problems there were linear models as well as nonlinear models [1,2,3]. The most advantage of the linear models is their simplicity, high speed of calculation, easyness to reach the global optimal solutions. For the nonlinear models, they have well performance when we deal with nonlinear problems.

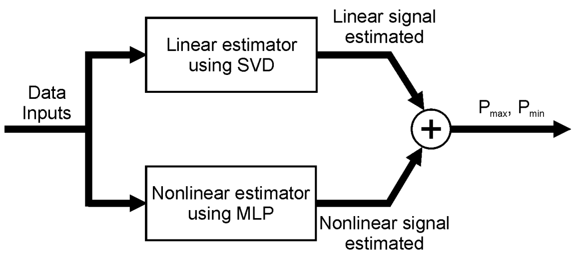

Fig. 1 The general structure of the proposed hybrid model

In this paper we propose a new hybrid model of prediction which consists of both linear and nonlinear estimators. The general structure of this hybrid model is shown on Fig. 1. The linear estimator will be modeled by using the SVD (Singular Value Decomposition) algorithm, where the nonlinear estimator is based on the MLP (MultiLayerPerceptron) neural network.

This hybrid solution allows us to amplify the advantages of the both models in order to get a high accuracy in calculation, when dividing parameters into 2 sets of linear and nonlinear ones and their separated adaptive (or iterative) learning process will have an increased speed (and the effectiveness).

II. HYBRID MODEL OF THE PEAK PREDICTION FOR THE DAILY POWER LOAD

The power load peaks prediction uses the historical data to make the assumption on the future values. Let's assume the value (we'll give the formulas for highest peak, the lowest has the similar form) is estimate by using following form:

$$P_{\max}(d) \approx f_{i=1,2,\ldots,K}\left(P_{\max}(d-i)\right) + \sum_{i=1}^{K} a_i \cdot P_{\max}(d-i) \tag{1}$$

wherethere *f()* is a nonlinear function and a_i are the linear coefficients of the linear estimator.

2.1 The SVD algorithm for linear estimators [4]

The linear optimalization problem can described as the problem of finding vector $\mathbf{x}$ such as $\|\mathbf{A}\cdot\mathbf{x}-\mathbf{b}\| \to \min$ where $\mathbf{A}$ and $\mathbf{b}$ are given matrix and vector respectively (usually we have more equations the unknown, it means $\mathbf{A}$ has more rows than columns) In order to solve this solution, we can decompose the matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ by using SVD algorithm. According to this algorithm, all matrix $\mathbf{A}$ can be decomposed like:

$$\mathbf{A}_{m\times n} = \mathbf{S}_{m\times m} \cdot \mathbf{U}_{m\times n} \cdot \mathbf{V}_{n\times n}^{T} \tag{2}$$

where $\mathbf{S}$ and $\mathbf{V}$ are orthonormal matrices

$$\mathbf{S}\cdot\mathbf{S}^T = \mathbf{S}^T\cdot\mathbf{S} = \mathbf{1}_{m\times m}$$

and $\mathbf{V}\cdot\mathbf{V}^T = \mathbf{V}^T\cdot\mathbf{V} = \mathbf{1}_{n\times n}$

and matrix $\mathbf{U}$ is a diagonal matrix. ($u_{ij} = 0$ when $i \neq j$ and the chain $\{u_{11}, u_{22}, \ldots, u_{nn}\}$ has decreasing values).

Once we have the SVD decomposition of a matrix, it's easy to calculate the so called *pseudo-inverse* matrix of $\mathbf{A}$ (pseudo- means matrix $\mathbf{A}$ can be non-squared matrix)

$$\mathbf{A}^{+} = \mathbf{V}_{n\times n} \cdot \mathbf{U}_{n\times m}^{-1} \cdot \mathbf{S}_{m\times m}^{T} \tag{3}$$

where matrix $\mathbf{U}_{n\times m}^{-1}$ is also a diagonal matrix and its elements have reciprocal values in respecting values of $\mathbf{U}$.

After that, the solution of the problem is:

$$\mathbf{x}_{opt} = \mathbf{A}^{+} \cdot \mathbf{b} \tag{4}$$

From [4] we have this is the global optimal solution of the linear problem posed above.

Applied in the main problem of this paper, the linear relationship between peak value of a day can be estimated from the system of approximations:

$$\begin{array}{lcl}
\sum_{j=1}^{K} a_j \cdot P_{\max}(d-j) & \approx & P_{\max}(d) \\
\sum_{j=1}^{K} a_j \cdot P_{\max}(d-1-j) & \approx & P_{\max}(d-1) \\
\ldots & \ldots & \ldots \\
\sum_{j=1}^{K} a_j \cdot P_{\max}(d-p-j) & \approx & P_{\max}(d-p)
\end{array} \tag{5}$$

Change the above system into the matrix form:

$$\begin{bmatrix} P_{\max}(d-1) & \ldots & P_{\max}(d-K) \\ P_{\max}(d-2) & \ldots & P_{\max}(d-K-1) \\ \ldots & \ldots & \ldots \\ P_{\max}(d-p-1) & \ldots & P_{\max}(d-K-p) \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{bmatrix} = \begin{bmatrix} P_{\max}(d) \\ P_{\max}(d-1) \\ \vdots \\ P_{\max}(d-p) \end{bmatrix} \tag{6}$$

From this equation we will look for a vector $\mathbf{a}=[a_1,a_2,\dots,a_K]^T$ in order to minimize the prediction erros. The fundamental questions are "how many historical data are needed? which data?" In this paper we propose an adaptive way to select the historical data: Starting from a large enough number of *K* (the farest historicla data taken into model) we find the vector **a**, through that we can find out the smallest absolute value of a_i which means the historical day that has smallest influance on the next day value.

Remove that day from the data set, we recalculate the vector $\mathbf{a}=[a_1,a_2,\dots,a_K]^T$ and again we find out and remove the lowest meaning coefficient. The process will be stopped when some criteria are met (we stop the reduction when there are only 10 coefficients left).

2.2 Application of MLP network to estimate the nonlinear dependencies between the power load of different days

After removing the linear dependencies on 10 historical data, the remain $NL(d)=P_{\max}(d)-\sum_{i=1}^{10}a_i\cdot P_{\max}(d-k_i)$ is the nonlinear remain of information. We will use the MLP network to model this relationship. The inputs of the MLP are the historical data $P_{\max}(d-k_i)$ $(i=1,\dots,10)$ where the output is the respecting $NL(d)$. The MLP is a very classical and well exposed nonlinear estimator, which has a 3-layer structure [5] and the parameters can be adaptively trained by backpropagation algorithms [4]. In this paper we will use the 2-nd order Hessian algorithm to train the MLP parameters. The chosen network has 10 inputs, 40 neurons in hidden layer (selected by "trial-and-test" method) and one output.

III. NUMERICAL EXPERIMENTS

The proposed model is tested on the power load of Thai Nguyen province. Data were taken from 9/2000 to 4/2006 (about 290 weeks). The linear estimator was created from $K_{\max}$ = 60 days down to K_{selected} = 10. As the result the best 10 historical days to estimate the day *d* are *d-1, d-3, d-6, d-7, d-14, d-21, d-28, d-35, d-42* and *d-49*.

The coefficients for these 10 past days are: 0,62; 0,032; 0,03; 0,21; 0,24; 0,12; 0,057; 0,042; 0,066; 0,056. Since that, the linear estimator is:

$$\begin{aligned}P_d \approx\ & 0,62P_{d-1}+\ 0,032P_{d-3}+0,030P_{d-6}+\\ &+0,21P_{d-7}+0,24P_{d-14}+0,12P_{d-21}+\\ &+0,057P_{d-28}+0,042P_{d-35}+0,066P_{d-42}\\ &+0,056P_{d-49}\end{aligned} \tag{7}$$

It means that the next $P_{\max}$ depends most strongly on the load of the 3 days of the last week, and only on the same weekday of 7 weeks earlier.

As it can be seen on the figure, the linear prediction error tends to increase with time, which shows that the peak power load tends to have more nonlinear dependencies in present time.

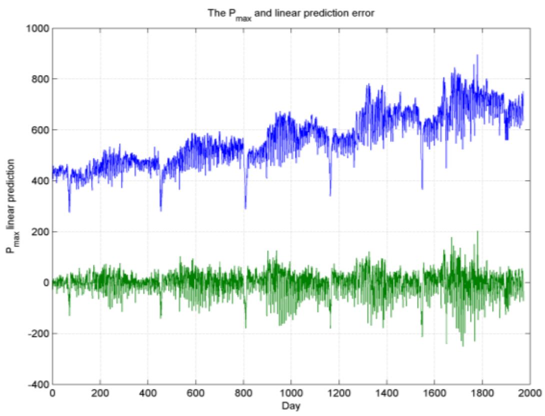

Fig. 2 The P_{max} (upper – blue or darker color) and the linear prediction error (lower – green or lighter color) for the data used in the simulations.

After having the linear estimator indentified, the nonlinear remain part is pushed to the MLP network to train the parameters of the neural networks. With 5 year of daily peak load data, a MLP network was trained, which contains only 1 hidden layer with 40 tansig neuron (neuron with tranfer function $y=f(u)=\tanh(u)$), 10 inputs (according to the 10 best historical days selected before (*d-1, d-3, d-6, d-7, d-14, d-21, d-28, d-35, d-42* and *d-49*) and only one out put (peak load of day *d*). After 50 iterations of the Levenberg – Marquardt learning algorithm [5], the squared error has droped significantly and is equal 8,37/sample. The decreasing process of SSE is presented on Fig. 3.

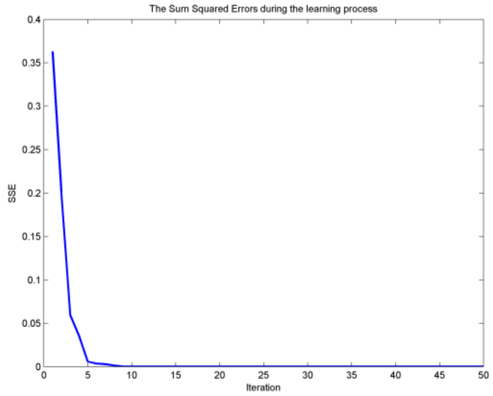

Fig. 3 The Sum Squared Error decreases during the learning process of MLP network

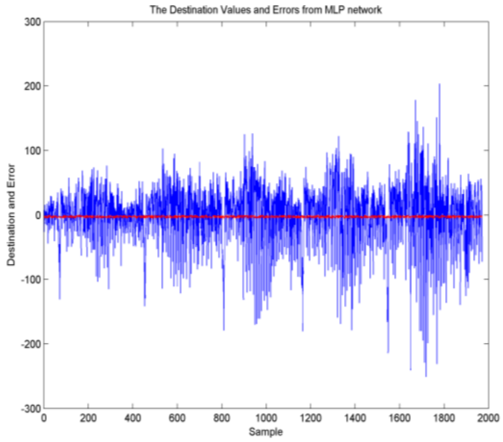

Fig. 4 The destination values and the output errors of the MLP network

The output of the MLP now is shown on the Fig. 4. It can be seen from this that the output of the network strictly follow the destination (required values) and the output error has a much lower level than the required outputs.

IV. CONCLUSIONS

From the numerical experiments, we can show that the idea of splitting a complex problem into linear and nonlinear subproblems work wells for us. It combines both advantages of systems: the simplicity of linear systems and the accuracy of nonlinear models. At the same time, it allows us to split the parameters of the system for learning them independently: the linear parameters are found by SVD algorithm, the nonlinear parameters are trained adaptively by using learning algorithm of the MLP network (for example the famous Levenberg – Marquadrt algorithm).

When applying to the daily peak power load, the results has shown that we can get a very low error of prediction (8MW/day or 0,35% per period). The solution seems to be an effective model for the selected estimation and prediction problems.

REFERENCES

1. *A. Cichocki, R. Unbehauen*; Neural networks for optimization and signal processing, Wiley, N. Y., 1993
2. *Brockwell P. J. and Davis R. A.*; Introduction to Time Series and Forecasting, Springer, 1997.
3. *Tran Ky Phuc, Monique Polit, Nguyen Tien Phong, Le Thi Thanh Ha*; Neural networks and application in short-time power load in Hanoi with the impact of weather data, Today Automation, No. 9, 2006.
4. *S. Osowski, M. Stodolski, P. Bojarczak*; Fast second order learning algorithm for feedforward multilayer neural networks and its applications, Neural Networks, 1995
5. *Hornik K., Stinchcombe M. and White H.*; Multi-layer feedforward networks are universal approximation, Neural Networks 2, pp. 359-366, 1989.
6. H. *Demuth, M. Beale;* Neural Network Toolbox for use with Matlab, The MathWorks, Inc., 1992, Natick

Contact: Tran Hoai Linh - Tel: (+84)912.316.629, Email: linh-3i@mail.hut.edu.vn
Hanoi University of Technology - No. 1, Dai Co Viet road, Hanoi, Vietnam