TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7379~7385
e-ISSN: 2087-278X

Wavelet Neural Network-based Short-Term Passenger Flow Forecasting on Urban Rail Transit

Xiaojie Zhang*, Baohua Mao, Yongliang Wang, Jia Feng, Minggao Li
MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University (BJTU), The 8th Building, No. 3 Shang Yuan Cun, Hai Dian District Beijing, China, 100044 Ph. /Fax: +8610-5168 2264
*Corresponding author, e-mail: zhangxiaojie@bjtu.edu.cn

Abstract

Accurate forecasting of short-term passenger flow has been one of the most important issues in urban rail transit planning and operation. Considering the shortcomings of traditional forecasting methods, and in order to improve forecasting accuracy of passenger flow, this paper presents a wavelet neural network (WNN) for short-term passenger flow forecasting. One real urban rail transit station with large and significantly changed passenger flow is chosen to be the example. The proposed method and BP neural network have been compared with the results, which show that the WNN model has more advantages. The WNN model features higher learning speed and drastically less convergence time, showing that it is meaningful in practical application. Furthermore, the calculated relative errors using BP neural network are nearly in the range [-0.4, 0.3] and the calculated relative errors are in the range [-0.25, 0.1] using the WNN, which demonstrate the superior accuracy of using this approach.

Keywords: *urban rail transit, short-term passenger flow forecasting, wavelet neural network*

1. Introduction

Urban rail transit short-term passenger flow forecasting is one of the essential elements in transportation systems which can be used to fine-tune travel behaviors, reduce passenger congestion, and enhance service quality of transportation systems. The forecasting results of short-term passenger flow can be applied to support transportation system management such as operation planning, and station passenger crowd regulation planning [1].

In recent years, a number of techniques are used in short-term passenger flow forecasting. Kalman filtering model is simple in calculation and fast in speed; however, it fails to reflect uncertainty and nonlinearity in traffic flow process and is unable to handle the rapid variation and complicated process changes under lying of traffic flow [2-3]. Support vector regression (SVR) has been successfully used to predict traffic parameters such as hourly flow, and travel time. However, when the new traffic data become available in every couple of minutes or seconds, the traditional SVR method is not a practical option because it requires complete model training whenever a new data point is added [4]. Genetic algorithm is able to reserve a few best fitted members of the whole population for the next generation in the operation process, however, after some generations genetic algorithm may lead to a premature convergence to a local optimum in the searching the suitable parameters of a model [5-6]. Simulated annealing (SA) is a stochastic based general search tool that mimics the annealing process of material physics: however, it costs more computation time [7]. The classical representative is artificial neural network (ANN) model due to its superior performance to approximate any degree of complexity and without prior knowledge of problem solving [8]. ANN model is based on a model of emulating the processes of the human neurological system to determine the numbers of vehicle and temporal characteristics from the historical traffic flow patterns, especially for nonlinear and dynamic evolutions [9].

Wavelet analysis is better method which is applied to the non-stationary signal analysis [10-11]. So this paper combines the wavelet transform and BP neural network, and presents the wavelet neural network. It describes the wavelet neural network that is developed to predict the urban rail transit short-term passenger flow. The feasibility of this method is demonstrated through the basis of time series of passenger flow. The results reported in this

Received May 3, 2013; Revised July 4, 2013; Accepted August 23, 2013

paper clearly show the advantages of implementing this approach for urban rail transit short-term passenger flow prediction.

This paper is organized as follows: Section 2 presents the wavelet neural network models; Section 3 introduces the data processing; Section 4 illustrates the forecasting performance and makes some discussions; Section 5 gives the conclusions.

2. Wavelet Neural Network

2.1. Wavelet Transform

Wavelet is a type of transformation that retains both time and frequency information of the signal [12]. In wavelet transform, all basic function $\psi_{a,b}(x)$ can be derived from a mother wavelet $\psi(x)$ through the following dilation and translation processes:

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \qquad a\,,\, b \in R \text{ and } a \neq 0 \tag{1}$$

Where *a* and *b* are the dilation and translation parameters, respectively.

Given a time-varying signal $f(t)$, then, the Continuous Wavelet Transform is defined as follows:

$$CWT\,(f(t):a,b) = \int f(t)\psi^{*}_{a,b}(t)dt \tag{2}$$

Where "*" denotes the complex conjugation. When $a = 2^{j}$, $b = k2^{j}$, $j,k \in Z$ (Z is the set of integers), it can be written as:

$$\psi_{j,k} = 2^{-j/2}\psi(2^{-j}t-k) \tag{3}$$

The fast algorithm of DWT (Discrete Wavelet Transform) can be written as following [13]:

$$\begin{cases} A^{d}_{2^{j+1}}f = \sum\limits_{k=-\infty}^{+\infty} h(k-2n)A^{d}_{2^{j}}f \\ D_{2^{j+1}}f = \sum\limits_{k=-\infty}^{+\infty} g(k-2n)A^{d}_{2^{j}}f \end{cases} \tag{4}$$

Where $g(n) = (1-n)^{1-n}h(1-n)$, $n \in Z$. In (4), $g(n)$ and $h(n)$ are the high-pass and low-pass filters, respectively, f is the discrete signal. A signal or function $f(t)$ decomposed by wavelet transform is expressed finitely as follows:

$$f(t) = f_0(t) + \sum_{j=0}^{N-1} w_j(t) = \sum_{k} C^{0}_{k}\psi_{0k} + \sum_{j=0}^{N-1}\sum_{k} d^{j}_{k}\psi_{jk} \tag{5}$$

Where $f_0(t)$ represents the lowest frequency component, w_j represents different frequency component and *N* represents decomposition level; d^{k}_{j} is the wavelet coefficient at scale *j*.

2.2. Wavelet Neural Network

The wavelet neural network (WNN) consists of three layers: input layer, hidden layer and output layer. Unlike a traditional back-propagation neural network that applies activation functions to both hidden and output layers, WNN employs mother wavelets (or wavelet functions) in the hidden layer only [14]. The connections between input units and hidden units, and between hidden units and output units are called weights w_{ti} and W_t, respectively. The network structure is shown in Figure 1.

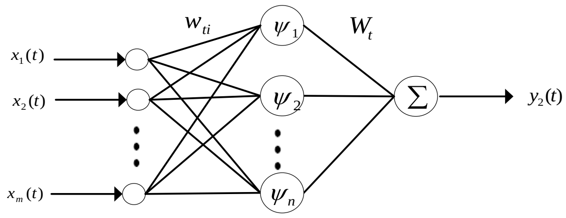

Figure 1. Wavelet Neural Network Structural Diagram

In this WNN, the training procedure is described as follows:

Initializing the dilation parameter a_t, translation parameter b_t and node connection weights w_{ti}, W_t to some random values. All those random values are limited in the interval (0, 1).

Input data $X_n(i)$ and the corresponding output values V_n^T, where i varies from 1 to S, representing the number of the input nodes, n represents the nth data sample of training set, and T represents the target output state.

The output value of the sample V_n is calculated with the following formula:

$$V_n = \sum_{t=1}^{T} W_t \psi\left(\frac{\sum_{i=1}^{S} w_{ti} x_n(i) - b_t}{a_t}\right) \tag{6}$$

Where ψ is considered a mother wavelet, such as the Morlet wavelet filter, and is represented by:

$$\psi(t) = \cos(\omega_0 t)\exp(-0.5t^2) \tag{7}$$

To reduce the error, W_t, w_{ti}, a_t, b_t are adjusted using ΔW_t, Δw_{ti}, Δa_t, Δb_t. In the WNN, the gradient descend algorithm is employed, through the following equations:

$$\Delta W_t(j+1) = -\eta \frac{\partial E}{\partial W_t(j)} + \alpha \Delta W_t(j) \tag{8}$$

$$\Delta w_{ti}(j+1) = -\eta \frac{\partial E}{\partial w_t(j)} + \alpha \Delta w_t(j) \tag{9}$$

$$\Delta a_t(j+1) = -\eta \frac{\partial E}{\partial a_t(j)} + \alpha \Delta a_t(j) \tag{10}$$

$$\Delta b_t(j+1) = -\eta \frac{\partial E}{\partial b_t(j)} + \alpha \Delta b_t(j) \tag{11}$$

Where the error function E is taken as:

$$E = \frac{1}{2}\sum_{n=1}^{N}(V_n^T - V_n)^2 \tag{12}$$

And, N standing for the data number of training set, η , α being the learning rate and the momentum term, respectively.

The process continues until E satisfies the given error criteria, and the whole training of the WNN is completed [15].

3. Data Collection and Pre-processing

The entering passenger flow dataset of one urban rail transit station is collected to investigate the viability of the proposed WNN approach for forecasting the short-term passenger flow. The dataset was collected during 6:00 AM to 9:05 AM, on five working days, 2012, and the sampling period was 2 minutes. One day's passenger flow data are shown in Figure 2. In order to examine whether WNN gives better results or not, the dataset is divided into two parts to be used for training and testing. One part on the first four days are used as the training sample, for determining the WNN parameters; the other part on the fifth day are used as the testing sample, for validating the performance of the trained model.

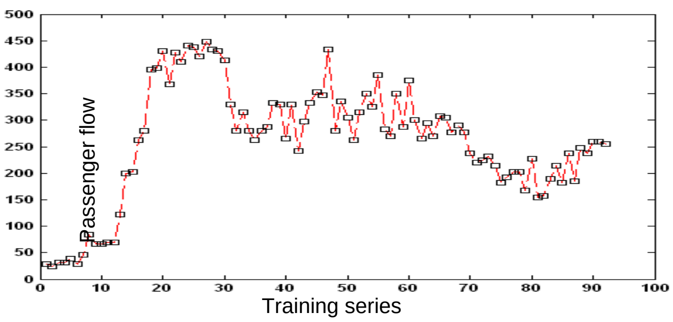

Figure 2. One day's Passenger Flow

In order to reduce the influence of the prediction performance due to the different dimensions of sample data, the sample data are normalized according to the following formula:

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{13}$$

Where x , is the data before normalization, $\bar{x}$ is the data after normalization, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the raw data.

After training the WNN and obtaining the forecasting output, the forecasting model should renormalize the output data as the following:

$$y = u * (x_{\max} - x_{\min}) + x_{\min} \tag{14}$$

Where y is the output of the network, u is the normalized output.

4 Forecasting Performance and Discussion

4.1. Structure of WNN

The wavelet neural network used for predicting short-term passenger flow consists of three layers. It is developed using the 4 neurons as input layer. The output layer has one neuron that predicts short-term passenger flow by the model. The number of neuron in the hidden layer is unknown and needs to be optimized. In addition to the number of neurons in the hidden layer, the WNN parameters consist of the learning rate, the momentum and the number of iterations should also be optimized. In this paper, the number of neurons in the hidden layer and other parameters, except the number of iterations, are optimized simultaneously.

Deciding the number of neurons in the hidden layers is a very important part of deciding your overall neural network architecture. There are many rule-of-thumb methods for determining the correct number of neurons to use in the hidden layers, such as the following:

$$m = \sqrt{n+l} + \alpha \tag{15}$$

$$m = \log_2 n \tag{16}$$

$$m = \sqrt{nl} \tag{17}$$

Where m , is the number of neurons in the hidden layer, n is the number of neurons in the input layer, l is the number of neurons in the output layer, α is the constant between 1-10.

In this paper, the formula (15) is selected, after many times of experiments, the hidden layer adopts 6 neurons with faster speed and better learning effect, so the structure of WNN is 4-6-1. The learning rate of all those neural network is determined as 0.02, the momentum term as 0.6. The training procedure is described in section 2.2.

4.2. Forecasting Results

After the urban rail transit short-term passenger flow forecasting model is established, the training dataset is used for training the BP neural network and WNN by MATLAB. (The BP neural network has the same parameters with the WNN.) Recently, some researchers have tried to develop the BP neural network approach for the city traffic flow forecasting [16]. In order to evaluate the forecasting accuracy and stability, this study compares BP neural network with WNN.

The forecasting results indicate that the proposed WNN model is feasible and effectivefor the short-term passenger flow forecasting. The results are analyzied as follow:

(1) The WNN neural network model can get higher learning speed and less convergence time that are used for short-term passenger flow prediction. For the BP neural network, a very satisfactory result is obtained after about 239 training epochs, but the WNN is about 106 training epochs. The training time of BP neural network is 3.925ms, and the training time of WNN is 1.256ms.

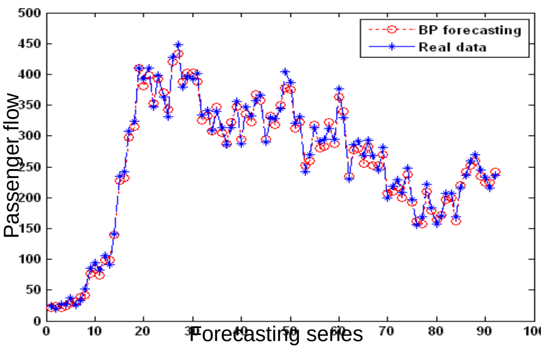

Figure 3. Forecasting Curves of BP Neural Network

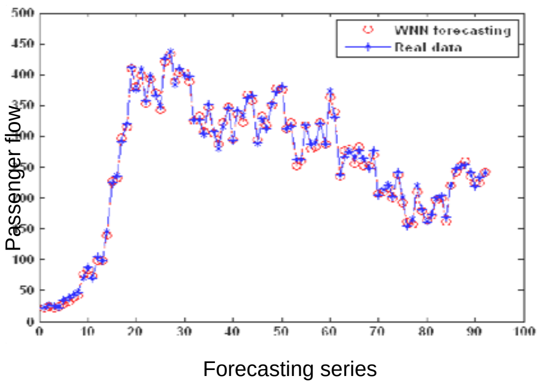

Figure 4. Forecasting Curves of WNN

(2) The improved prediction results of the WNN algorithm correspond to the real passenger flow better. Figure 3, Figure 4 illustrate that the prediction value curve of WNN fits the real value curve better than the BP neural network does. To better appreciate the performance of the WNN, the relative error between the actual and the forecasting flow series is considered as appropriate and used here. From the Figure 5, we can see that the relative errors of BP neural network are nearly in the range [-0.4, 0.3], and that of WNN are in the range [-0.25, 0.1].

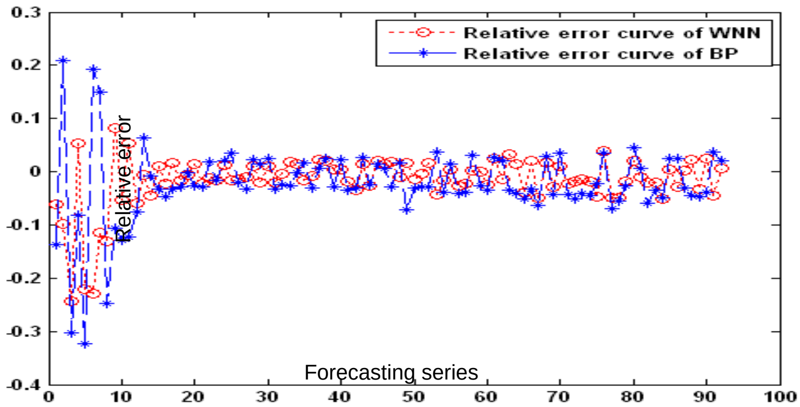

Figure 5. The Relative Error Curve of BP Neural Network and WNN

The BP algorithm has several drawbacks, for example, the performance of the network learning is strictly dependent on the shape of the error surface, values of the initial connection weights, and the convergence to the global optimum is not guaranteed. The WNN is a novel approach towards the learning function. It combines the wavelet theory and feed-forward neural networks, and utilizes wavelets as the basis function to construct a network. Wavelet function is a local function and influences the networks' output only in some local ranges. The wavelet neural network shows surprising effectiveness in solving the conventional problems of poor convergence or even divergence encountered in other kinds of neural networks.

5. Conclusion

The traditional prediction models have some weaknesses; therefore this paper established a wavelet neural network-based short-term passenger flow forecasting model for urban rail transit station, combining the theory of wavelet transform with the BP neural network.

Through analyzing in the paper, the following conclusions may be inferred:

(1) The WNN model features a higher learning speed, reduced convergence time, and being appropriate to practical application, so this method has fair prospects of application for the short-term passenger flow forecasting.

(2) From the relative error curve of BP neural network and WNN, we can see that the testing performance of the WNN is found to be better than BP neural network in accuracy and robustness.

Therefore, the favorable results obtained in this work reveal that the proposed model is a valid alternative for the short-term passenger flow forecasting. In addition, even the proposed WNN model is one of the hybrid forecasting models; some other advanced optimization algorithms can be applied for the WNN model to improve the accuracy of the neural network, and this would be valuable future work.

Acknowledgement

The research described in this paper was substantially supported by a grant from National Basic Research Program of China (2012CB725406) and the National Natural Science Foundation of China (Project No. 71131001; Project No. 70971010).

References

[1] Yu Wei, Muchen Chen. *Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks*. Transportation Research Part C. 2012; 21:148-162.
[2] Y Wang, M Papageorgiou, A Messmer. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transport. Res*. 2007; 41(2): 167–181.
[3] ZHANG Chun-hui, SONG Rui, SUN Yang. Kalman Filter-Based Short-Term Passenger Flow Forecasting on Bus stop. *Journal of Transportation Systems Engineering and Information Technology*. 2011; 11(4): 154-159.
[4] Manoel Castro Neto, YoungSeon Jeong. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications*. 2009; 36: 6164–6173.
[5] WeiChiang Hong, Yucheng Dong. Hybrid evolutionary algorithms in a SVR traffic flow forecasting model. *Applied Mathematics and Computation*. 2011; 217: 6733–6747.
[6] Yan Xuesong, Wu Qinghua. An improved genetic algorithm and its application. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(5): 1081-1086.
[7] S Kirkpatrick, CD Gelatt, MP Vecchi. Optimization by simulated annealing. Science. 1983; 220: 671–680.
[8] Mardiyono, Suryanita. Intelligent monitoring system on prediction of building damage index using neural-network. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(1): 155-164.
[9] MingWei Li, WeiChiang Hong. Urban traffic flow forecasting using Gauss–SVR with cat mapping, cloud model and PSO hybrid algorithm. *Neurocomputing.* 2013; 99: 230–240.
[10] Qin Ling Yang Jun. Application of Wavelet Packets Analysis to the Fault Diagnosis of Rotating Machinery. *Electronic Engineer*. 2006; 32(1): 48-50
[11] Jianwei Yang. A Method of Bearing Fault Feature Extraction Based on Improved Wavelet Packet and Hilbert Analysis. *International Journal of Digital Content Technology and its Applications*. 2010; 4(4):127-139.
[12] I Daubechies. Ten Lectures on Wavelet, Philadelphia, PA, SIAM Press. 1992: 78-80.
[13] Pang Peilin, Ding Guangbin. *Wavelet-based Diagnostic Model for Rotating Machinery Subject to Vibration Monitoring.* Proceedings of the 27th Chinese Control Conference, Yantai. 2008: 303–306.
[14] Jongsoo Lee, Kwang Ho Shin. A conservative method of wavelet neural network based meta-modeling in constrained approximate optimization. *Computers and Structures*. 2011; 89: 109–126.
[15] M Amina, VS Kodogiannis. Identification of the Listeria monocytogenes survival curves in UHT whole milk utilising local linear wavelet neural networks. *Expert Systems with Applications.* 2012; 39: 1435–1450.
[16] Wu Tiefeng. Study of Improved BP Neural Network on Forecasting City Traffic Flow. Transportation Science& Technology. 2010; 242(5): 92–94.