

An ELM-based traffic flow prediction method adapted to different data types

Wang Xingchao^{1, a}, Hu Jianming^{2, b}, Zhang Yi³ and Wang Zhenyu⁴

¹ Department of Control Science and Engineering, Tongji University, Shanghai, 201804, China

^{2, 3, 4} Department of Automation, Tsinghua University, Beijing, 100084, China

^aemail: 1352805@tongji.edu.cn, ^bemail: hujm@mail.tsinghua.edu.cn

Keywords: ELM; non-uniform data; traffic flow prediction

Abstract. Traffic flow prediction has been the crucial aspect in the Intelligent Transport Systems (ITS) and it will help ITS to induce the traffic flow rationally. In this paper, a prediction model based on ELM has been discussed through testing this model with actual data. This model takes advantage of the speed and ability of handling nonlinear system of the ELM algorithm. By analyzing the actual data in Haining County in 2016, the feasibility and advantages of ELM prediction model have been shown when compared with other algorithms. Analyzing the non-uniform and lost data, this model shows the performance that the model can be used in actual application. These characters can remedy the shortcomings which other models have.

Introduction

With the development of society, the need of Intelligent Transport Systems (ITS) has been stronger in case of the traffic jam. Moreover, in the small-medium cities this problem has been more serious because of the increasing vehicles and narrow road. Short term travel flow prediction based on the time series takes an important role in the ITS to ease the road. Traditional prediction models include: Kalman filter, auto-regressive integrated moving average model (ARIMA), neural networks, support vector regression (SVR). Besides, neural networks has been another major tool to predict the travel flow.

However, these current prediction models are imperfect to predict. Although neural networks can predict accurately, but the time used by the processing is much longer than other models. The ARIMA is fast but not accurate. Therefore the research wants to take advance of the speed and accuracy of Extreme Learning Machine (ELM) in order to predict the travel flow better.

What's more, most forecasting models cannot handling the non-uniform sampling data and missing data which may be attained in the actual society. If the prediction models always use the method based on uniform sampling data to handling these different type data, the prediction results may be not accurate. Thus, the research also test this prediction model based on the different datasets.

For further accuracy and rapidity travel flow prediction, this forecasting model based on ELM has been used to predict the travel flow in this paper. This experiment used ARIMA and wavelet neural network (WNN) to be the comparison to test the performances of this model. Besides, the research also uses the non-uniform and missing data to prove that this model can satisfy the need of actual application.

Extreme Learning Machine

The ELM algorithm (proposed by Huang) is a fast learning method. This method is a kind of Single-Layer Feedforward Neural Networks (SLFNs). The activation function in the hidden layer can be infinitely differentiable so the weight of input can be randomly settled. Moreover, the least squares solutions can be obtained by minimizing quadratic loss function. As a result, the iteration is not required when the network parameters are under randomly determining, which benefits the reducing of the amount of calculation and the time of forecasting. The network structure of ELM

shows in the following figure.

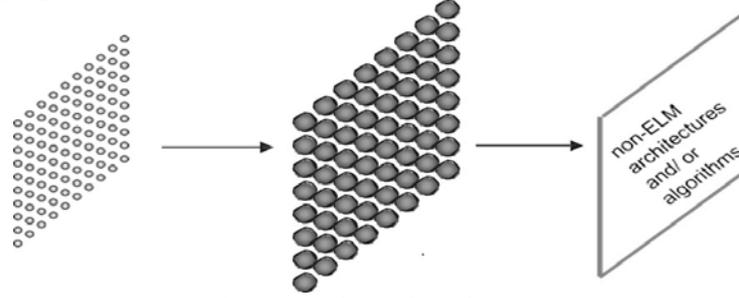


Fig. 1. ELM algorithm diagram

Theorem2.1. For N arbitrary distinct samples (x_i, t_i) , $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbb{R}^n$ is the input vector (n is the number of input node), and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in \mathbb{R}^m$ is the output vector. $G(x)$ is the activation function. The model can be described as the following function.

$$\sum_{i=1}^m \beta_i g_i(x_j) = \sum_{i=1}^m \beta_i g(\omega_i \times x_j + b_j) = t_j, \quad (1)$$

$j=1, 2, \dots, N$,

Where, $\omega_i = [\omega_{i1}, \omega_{i2}, \dots, \omega_{in}]^T$ is the weight vector to connect the i^{th} hidden node and the input nodes, b_i is the threshold of the i^{th} hidden node. $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is the weight vector, which connects the i^{th} hidden node and the output nodes. The N equations can be simplified as the following equation.

$$H\beta = T, \quad (2)$$

$$H = \begin{bmatrix} g(w_1 \times x_1 + b_1) & \cdots & g(w_m \times x_1 + b_m) \\ \vdots & \ddots & \vdots \\ g(w_1 \times x_N + b_1) & \cdots & g(w_m \times x_N + b_m) \end{bmatrix}, \quad (3)$$

$$\beta = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]_{k \times m}^T, \quad (4)$$

$$T = [t_1, t_2, \dots, t_m]_{k \times N}^T, \quad (5)$$

Where, H is the hidden layer output matrix; β is the weight vector of output layer; T is the desired output and k is the output node number. The unique smallest norm least squares solution of the equation (1) is:

$$\hat{\beta} = H^+ T, \quad (6)$$

H^+ is the Moore-Penrose generalized inverse of output matrix H .

Steps of ELM algorithm:

- (1) Randomly assign input weights and hidden layer threshold;
- (2) Calculate the hidden layer output vector H ;
- (3) Calculate the output weights vector $\beta = H^+ T$.

Data type

In this section, three types data which have been used in this paper will be introduced. Most travel prediction models base on the data sampled uniform, but actual data are sampled with some non-uniform even lost. Thus, this paper uses 3 kinds of data to test the model: uniform sampling data; non-uniform sampling data; missing data. Each data type has the unique parameter which will influence the dataset. However, most test data are uniform sampled and the non-uniform sampling data or missing data cannot be obtained. Thus, in this paper the research adopts the randomly merge method to obtain the non-uniform data and the randomly lost method to attain the missing data. When handling different data, this research uses different ways to predict.

Table 1 Data type

Data	Parameter	Signal
Uniform sampling data	Merge number	Merge_Num
Non-uniform sampling data	Max merge-number	Merge_Max_Num
Missing data	Missing rate	Lost_Rate

Prediction algorithm based on ELM

To predict based on different type data, the algorithm differs when handling the input dataset. For non-uniform sampling data the method is the same as the uniform data but the sequences number is fewer. Handling missing data, this research uses the abandon forecasting method which means that the algorithm will abandon the days' data when meeting the lost data.

The structure of this prediction algorithm shows in the flowing diagram

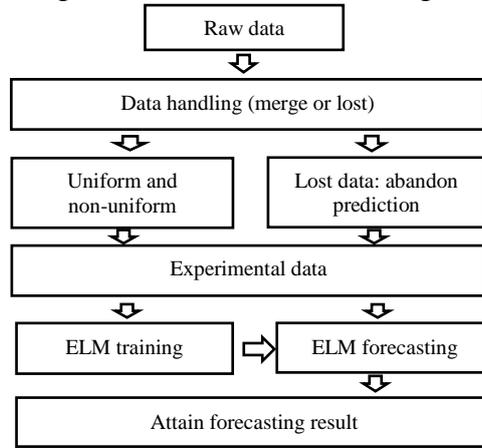


Fig. 2. Steps of prediction

When the research wants to forecast the $j+1^{\text{th}}$ sequence data, we can get the j^{th} traffic flow data. It is that there are N days' data and one day has T sequences. Therefore, the data vector is $N \times T$ and the i^{th} observation point's j^{th} traffic flow data is x_{ij} .

The training input of ELM is:

$$\begin{bmatrix} x_{N-p1+1, j+2-p2} & \cdots & x_{N-p1+1, j+1} \\ \cdots & & \cdots \\ x_{N-1, j+2-p2} & \cdots & x_{N-1, j+1} \end{bmatrix}_{p1 \times p2}, \quad (7)$$

The training output is:

$$\begin{bmatrix} x_{N-p1+1, j+1} \\ \cdots \\ x_{N-1, j+1} \end{bmatrix}_{p1 \times 1}, \quad (8)$$

The testing input is:

$$\begin{bmatrix} x_{N, j+2-p2} & \cdots & x_{N, j} \end{bmatrix}_{1 \times p2-1}, \quad (9)$$

Raw Data

The original data are used in this paper is collected in Haining County, Jiaxing City, Zhejiang Province, China by microwave in 2016. The raw data uniform sampled have been detected all days and the detection interval is 1 minute. The dataset covers 7 observation points and 20 days with one day has 1440 sequences. In this case, the first 19 days data are training data and the 20th day data are testing data. In the following figure, the locations of observation points are showed.



Fig. 3. Distribution of observation points

When the research plot the data and several properties can be found: little travel flow volume; strong fluctuation; no obvious crest. So the research merges and cuts out the raw data to find out the experimental data whose detection interval is 5 minutes.

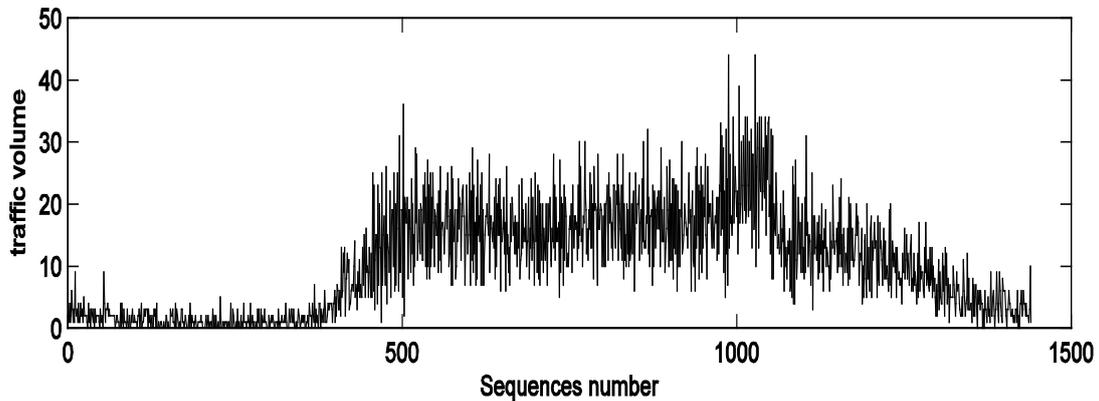


Fig. 4. Raw data

The forecasting process is programmed in Matlab 2014a environment running in an Intel Core i5-3230M, 2.6GHZ CPU. Running the program with the history data and testing data will get the forecasting performance.

Prediction results

Uniform data.

In this section, the experiment choose the 5th observation point to be the test point. Based on the uniform sampling data and comparison with ARIMA and WNN, the performances of this model will be shown in table and figures.

Table 2 Time and accuracy comparison

Model	Time	Error
ELM	0.375s	7.66%
ARIMA	0.570s	13.23%
WNN	34.63s	10.43%

In the table 2, the speed can be found faster and more accurate than other two models. The forecasting are showed in the following figures. The first one is the ELM forecasting results.

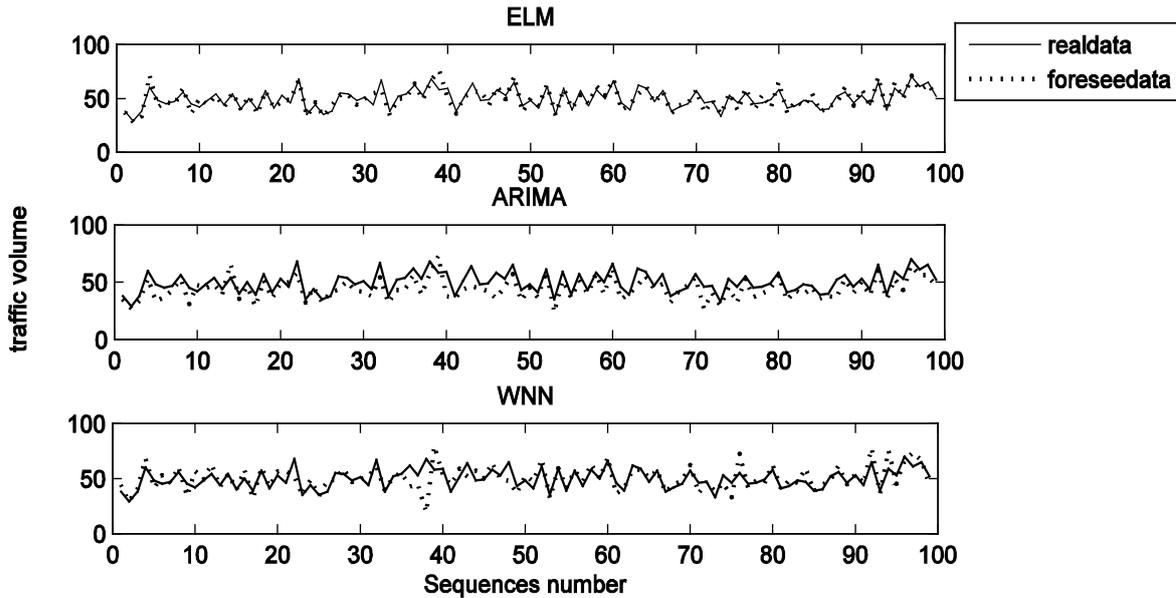


Fig. 5. Prediction comparison figure

Non-uniform data.

In this paper, the max merge number is 2 which means the adjacent data have 50% probability to combine. By using this method, we can get the traffic volume which is non-uniform sampled to do experiment. And the prediction results show in the following picture:

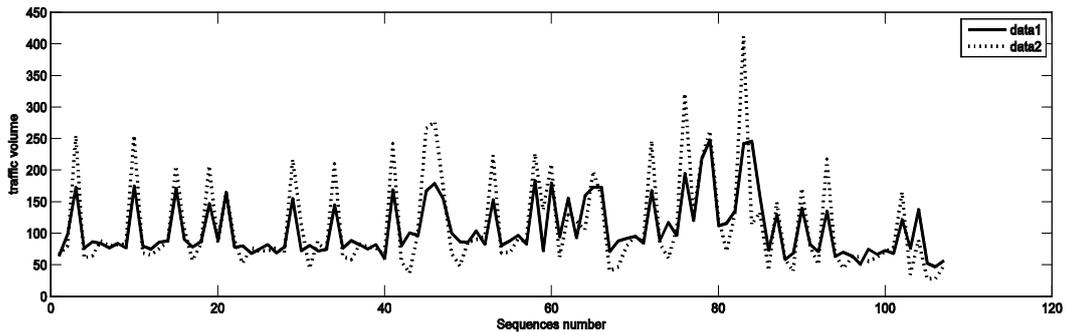


Fig. 6. Non-uniform sampling data prediction

From the figure 5, the result shows that the error rate of this method is 18.75%. Although this result is not good when compare with the uniform result. But if we analyze the data, the extremely strong fluctuation is the important influence to predict. And in the actual environment the random non-uniform sampling data hardly exist. Therefore, this method can also apply to predict based on non-uniform data.

Missing data.

This section will test whether this model can predict based on the data with lost data. The 5th observation point's data are used in this test. By setting the lost rate to get the sequences with missing data. This research sets the lost rate is 10%., which means every datum has 10% probability to lose. Following picture shows the prediction result.

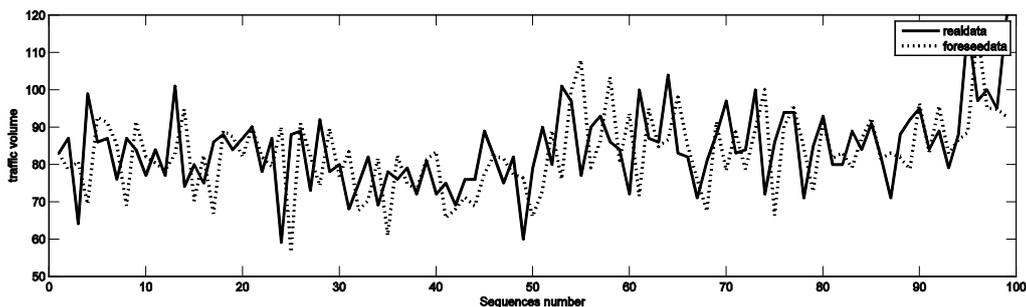


Fig. 7. Missing sampling prediction

In this picture, the prediction tendency can follow the real tendency well. By calculating, the error rate is 11.74%. In case of the using of abandon prediction method, the strong fluctuation cannot influence the performances of forecasting. Moreover, this model can predict accurately in this environment that some data lost. Therefore, this method will help the ITS to predict traffic flow when the actual data miss.

Conclusion

This paper concludes prediction methods which have been used and offer the obvious shortcomings in these methods. Besides, this paper discusses the time series prediction method based on ELM. By handling raw data, the training dataset and testing input will be determined. And because of that the weight of hidden layer can be randomly determined and the hidden layer is single, the ELM algorithm is faster than the normal neural networks. Moreover, the ability to adapt to the nonlinear system make the model can predict better. By analyzing the actual data in Haining County in 2016, the performances can be showed in tables and figures when it compared with other algorithms. Moreover, when testing the model with non-uniform and missing data, this method can satisfy the need of the actual application. Meanwhile, this model can provide better support for traffic guidance.

Acknowledgement

This work is supported by the national science and technology support program (2014BAG03B01), National Natural Science Foundation China 61673232, 61273238, and Tsinghua University Program (20131089307).

References

- [1] Guang-Bin Huang, Qin-Yu Zhu, Chee-kheong Siew, Extreme learning machine: Theory and applications, *Neurocomputing* 70 (2006) 489-501.
- [2] Guang-Bin Huang, What are Extreme Learning Machines? Filling the Gap Between Frank Rosenblatt's Dream and John von Neumann's Puzzles, 2015.04.09, *Cogn Compute* (2015) 7:263-278.
- [3] Lun Li¹, Dong Wang¹, Zhu Xiao¹, and Xiaohong Li, Urban Road Travel Time Prediction based on ELM, 5th International Conference on Advanced Materials and Computer Science (ICAMCS 2016)(2016).
- [4] Oh S, Byon Y J, Jang K, et al., "Short-term Travel-time Prediction on Highway: A Review of the Data-driven Approach", *Transport Reviews A Transnational Transdisciplinary Journal*, vol.35, no.1, pp.4-32 (2015).
- [5] Tianfan Ye, A pruning algorithm with L1/2 regularizer for extreme learning machine, *Journal of Zhejiang University Science C(Computers & Electronics)*(2014).
- [6] Jianding Xiao, Active set strategy of optimized extreme learning machine, *Chinese Science Bulletin* (2014).
- [7] Baojian Li, Monthly discharge forecasting using wavelet neural networks with extreme learning machine, *Science China Technological Sciences* (2014).