First Vehicle Arrival Time Prediction at Signalized Intersection Based on Wavelet-Elman Neural Network

Guchang Ao¹, Huiling Zhang²*, Linyu Yang¹, Wuting Jiang¹

ABSTRACT
There is a close relationship between the vehicle arrival time at signalized intersection and the vehicle delay at intersection entrance. When the red light of the signalized intersection is on, the first vehicle stopping at the stop line of the entrance is taken as the study object, and the time difference in the red light’s turning on and the first vehicle's arriving at the stop line is defined as the first vehicle arrival time at signalized intersection. There is great randomness in the first vehicle arrival time series at signalized intersection. Firstly, the wavelet transform (WT) method is adopted to decompose the non-stationary original arrival time series into low-frequency signal and high-frequency signal. Then the dynamics and fast feedback of Elman neural network are used to predict different signals respectively. Finally, a final predicted result of the first vehicle arrival time is obtained when they are subject to linear superposition. The result shows that the error of the first vehicle arrival time prediction at signalized intersection based on wavelet-Elman neural network is small, and the predicted value is highly consistent with the actual value, which can provide reliable data source for delay parameter extraction and signal timing optimization at signalized intersection.

Key Words: Traffic Engineering, Signalized Intersection, First Vehicle Arrival Time, Wavelet Transform (WT), Elman Neural Network, Short-Time Prediction

Introduction
The vehicle arrival-departure information of intersection is the basic supporting data of intersection delay acquisition, and the delay reflects the driver’s obstruction and expense at intersection, which can provide evaluation basis for facility setting and timing improvement at signalized intersection. However, due to the discontinuity of artificial observation and the limitation of detection technology, the delay value at signalized intersection hasn’t been obtained directly and in real time so far (Zhang et al., 2017). Therefore, researchers at home and abroad have carried out a series of studies on the acquisition of delay parameters at signalized intersection. The previous studies pay more attention to the acquisition method of delay the deduction of delay models, including Webster classical delay model and the improved models based on Webster model (Webster, 1958), such as Miller model (Miller, 1963), Robertson model (Hurdle, 1984), Akcelik model (Akçelik and Rouphail, 1994), and HCM 2010 model (Darma et al., 2005). In addition, delay acquisition methods get some attention (Cai et al., 2014; Mousa, 2002). The delay at signalized intersection is related to the characteristics of vehicle's arriving at signalized intersection (Pan et al., 2009; Li et al., 2019).

Corresponding author: Huiling Zhang
Address: ¹School of Traffic & Transportation, Chongqing Jiaotong University, Chongqing, 400074, China; ²Chongqing Key Laboratory of Traffic & Transportation, Chongqing Jiaotong University, Chongqing, 400074, China
e-mail: 326991773@qq.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 6 March 2018; Accepted: 25 April 2018
In recent years, more and more people have begun to explore the delay distribution of intersection (Gao et al., 2016; Chen et al., 2016), but they mainly focus on the statistical law of vehicle arrival. However, there are few studies on vehicle arrival time prediction, and the traffic studies on short-time prediction mainly focuses on traffic volume, travel time and traffic state (Yao et al., 2010, Cetin and Comert, 2006, Yang et al., 2012).

The traffic flow at signalized intersection has strong randomness that becomes stronger and stronger with the shortening of time scale, which makes it difficult to analyze the characteristics of traffic flow and conduct short-time prediction at signalized intersection (Weng et al., 2006). In order to improve the intelligent management level of urban signalized intersections, video-units have been set up at many intersections, providing an opportunity to extract arrival-departure information of the first vehicles at signalized intersection. Based on the first vehicle arrival time series at signalized intersection and aiming at the non-stationarity and nonlinearity of the time series, this study constructs an arrival time prediction model combining with wavelet analysis and Elman neural network in order to improve the accuracy of real-time prediction.

**Characteristic Analysis of First Vehicle Arrival at Signalized Intersection**

For signalized intersections with cameras, it is known from previous studies that the arrival and departure time of the first vehicles can be obtained after the vehicles enter the entrance (Zhang et al., 2010). Considering the close relationship between the parking and starting time of the first vehicle and the delay of the entrance, this study takes the first vehicle stopping at the parking line as the study object and takes the time of red light signal’s turning on of intersection as the time of signal cycle. The first vehicle arrival time of signal intersection is defined as the time difference in the red light’s turning on and the first vehicle’s entering the entrance and arriving at the stop line.

The six-kilometer intersection in Nanan District of Chongqing is a T-shaped intersection of arterial road (Xuefu Avenue, north-south direction) and branch road (Ertang Road, east-west direction). The signal timing adopts a 3-phase timing control scheme with a period of 90 seconds. The distance between the south entrance road stop line and the upstream intersection is 600 meters. The red-light duration and yellow light duration of road straight traffic flow phase at the south entrance are 30 seconds and 3 seconds respectively. 260 first vehicle arrival time sample data at the south entrance are obtained through the continuous video data of Ertang Road intersection and the manual processing in the later stage. The maximum arrival time is 30 seconds, the minimum arrival time is 0 seconds, and the average arrival time is 14.2 seconds. The saturation range of the entrance corresponding to 260 sample data is 0.43~0.80.

The first vehicle arrival time is drawn into the time series shown in Figure 1. It can be seen from the figure that the first vehicle arrival time series with the unit of period has large nonlinearity, randomness and non-stationarity.
Wavelet-Elman Neural Network Prediction Method of First Vehicle Arrival Time

The neural network model is featured with identifying nonlinear complex systems and is often used for short-term prediction of traffic flow (Hodge et al., 2014). However, the neural network easily falls into local minima and causes oscillation effects. Thus, the predicted effect of non-stationary time series will have large random fluctuations (Ji et al., 2016). Wavelet analysis has the capability of multi-scale decomposition for non-stationary series, which can reduce the non-stationarity of series and realize the prediction of non-stationary series (Lu and Wang, 2016). The prediction model of wavelet-neural network by combining wavelet analysis with neural network not only has more freedom than wavelet decomposition, but also has stronger fault-tolerant ability and better function approximation ability than general neural network.

**Prediction process**

In order to predict the first vehicle arrival time at signalized intersection based on wavelet-Elman neural network, the arrival time series at signalized intersection shall be decomposed by WT to obtain the low-frequency signal and high-frequency signal under the influence of different factors. Then all the signals are predicted separately by the Elman neural network with fast training characteristic. Finally, the predicted results of different signals are superimposed to obtain the final predicted result. The specific prediction process is shown in Figure 2, and the main steps are as follows:

1. **Step 1**: Count the sample data, and compose the sample data into a signal $S$ of length $M$.
2. **Step 2**: Decompose and reconstruct signal $S$ by multi-resolution $M$ scale wavelet to obtain one low-frequency signal and $M$ high-frequency interference signals.
3. **Step 3**: Input the low-frequency signal and high-frequency signal after decomposition and reconstruction into an Elman neural network for training and predicting;
4. **Step 4**: Linearly superpose $M+1$ prediction time series, namely, the predicted result of the first arrival time;
5. **Step 5**: Carry out error test on the prediction data to verify the validity of the prediction method.

![Figure 2. Prediction process of wavelet-Elman neural network](image)

**WT**

The purpose of WT is to decompose the original time series into two parts by low-pass filter and high-pass filter. One part is low-frequency coefficient which expresses signal characteristics by numerical value. The other part is high-frequency coefficient that expresses micro nuance of signals by numerical value. For discrete time series, Mallat algorithm is one of the most commonly used WT algorithms. Firstly, WT is performed on large scale signals (original time series) to obtain $s_1$ and $d_1$. Then the low-frequency part $s_1$ of $s_1$ and $d_1$ is selected to perform WT on the 1/2 scale of the original scale (the original wavelet function is shortened by 1/2) to obtain higher-frequency $s_2$ and $d_2$ to another set decomposition scale. The Mallat algorithm is (Mallat, 1999):

$$\begin{align*}
    s_{m+1} &= HS_m, \quad m = 0, 1, 2, \ldots, M \\
    d_{m+1} &= GS_m
\end{align*}$$

Where, H and G are respectively a low-pass filter and a high-pass filter, $M$ is a decomposition scale, and $S$ is an original signal (composed of time series $S_1$, $S_2$, $\ldots$, $S_M$). The wavelet decomposition process automatically produces a pyramid-shaped successive
decomposition result, as shown in Figure 3. The original time series can be decomposed into a high-frequency coefficient vector \( d_1, d_2, \ldots, d_M \) and a low-frequency coefficient vector \( S_{\hat{M}} \) through formula (1).

![Signal decomposition flow chart](image)

**Figure 3.** Signal decomposition flow chart

When Mallat algorithm is adopted for WT, the samples in the time series will be reduced by half after decomposition for each time, that is, the number of sample will decrease as the observation scale increases. Thus, the prediction of the time series is obviously disadvantageous. Therefore, the series decomposed by the Mallat algorithm needs to be reconstructed and the specific algorithm is:

\[
S_m = H^* s_{m+1} + G^* d_{m+1}, \quad m = \hat{M} - 1, \hat{M} - 2, \ldots, 0 \quad (2)
\]

Where, \( H^* \) and \( G^* \) is the dual operator of \( H \) and \( G \) respectively. The flow of signal reconstruction is shown in Figure 4. Reconstruction can increase the number of points of signal. \( D_1, D_2, \ldots, D_{\hat{M}} \) and \( S_{\hat{M}} \) are reconstructed separately to get \( d_1, d_2, \ldots, d_M \) and \( S_{\hat{M}} \) that are the same as the number of points of the original time series, and have \( S = S_{\hat{M}} + D_1, D_2, \ldots, D_{\hat{M}} \).

**Prediction of Elman neural work**

Elman neural network is a typical dynamic neural network. Based on the basic structure of BP artificial neural network, it makes the system has the ability to adapt to the time-varying characteristics by storing the internal states and making it have the function of mapping dynamic characteristics (Elman, 1990). Compared with BP neural network, Elman neural network has an additional receiving layer in addition to input layer, hidden layer and output layer. The receiver layer is mainly used for the feedback connection in or between the layers so that it can express the delay between the input and the output in time, which needs to be described by the dynamic equation (Abdalla, 2002; Gonçalves, 2016). It is this feedback that makes Elman neural network has memory function and has stronger dynamic behavior and computing ability compared with feedforward neural network.

The structure of the Elman neural network is shown in Figure 5, where \( Y, X, S, X_c \) respectively represents \( L \)-dimensional output node vector, \( N \)-dimensional node unit vector in the hidden layer, \( M \)-dimensional input vector and \( N \)-dimensional feedback state vector; and \( w_3, w_2, w_1 \) represents the connection weights of the hidden layer to the output layer, the input layer to the hidden layer and the receiving layer to the hidden layer respectively.

The input signal \( S \) will forward propagation to the hidden layer node. After the action function, the output signal \( X \) of the hidden layer node connects to the input of the hidden layer through the delay and storage of the receiving layer, and then propagates to the output node \( Y \) to obtain the output result. The input information is processed layer by layer from the input layer to the hidden layer and to the receiving layer, and transmitted to the output layer. If the output layer does not obtain the desired output, the error signal will be returned along the original connection channel. The weight of neuron is each layer is modified to minimize the error signal.

![Signal reconstruction flow chart](image)

**Figure 4.** Signal reconstruction flow chart
The existence of the receiving layer makes the feedforward connection part carry out the connection weight correction, while the recursive part is fixed, that is, the learning correction cannot be carried out. It can be seen from Figure 5 that the output of the receiving layer at time point \( k \) is \( \alpha \) times of that of the hidden layer at the time point \( k-1 \), namely,

\[
x_{c,j}(k) = \alpha \cdot x_{c,j}(k-1) + x_i(k-1), \quad l = 1, 2, \cdots, N
\]

(3)

Where, \( x_{c,l}(k) \) represents the output of the \( j \)-th receiving layer unit at the time point \( k \); \( x_i(k) \) represents the output of \( l \)-th hidden layer unit at the time point \( k \); \( k \) is an iterative time step; \( \alpha \) is a self-connected feedback gain factor. When \( \alpha \) is fixed to zero, the network is a standard Elman network; when \( \alpha \) is not zero, it is an improved Elman network.

\[
f(x) = \frac{1}{1+e^{-x}}
\]

(7)

\( g(x) \) is the transfer function of the output neuron, which is a linear combination of the output of the hidden layer:

\[
y(k) = w_3x(k)
\]

(8)

There are \( \tilde{M}+1 \) time series for the low-frequency signal \( S_{\tilde{M}} \) and high-frequency signal \( D_1, D_2, \cdots, D_{\tilde{M}} \) obtained by WT. The prediction model of Elman neural network is established respectively for the series and \( \tilde{M}+1 \) predicted results are obtained by network training, and correction of weight and the closed value. Finally, the predicted \( \tilde{M}+1 \) time series are linearly accumulated to get the predicted result of the first vehicle arrival time at the signalized intersection.

**Error test**

For the error test of the first vehicle arrival time prediction, the average absolute error \( AAE \) and root mean square error \( RMSE \) are selected. If the actual value is \( \hat{S} \) and the predicted value is \( \hat{S} \), the calculation formula of \( AAE \) is:

\[
AAE = \frac{1}{N} \sum_{i=1}^{N} |s_i - \hat{s}_i|
\]

(9)

The calculation formula of \( RMSE \) is:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (s_i - \hat{s}_i)^2}
\]

(10)

**Case Analysis of First Vehicle Arrival Time Prediction**

**WT**

The decomposition scale is assumed to be \( \tilde{M}=31 \) and WT is performed by the original time series shown in Figure 1 to obtain a low-frequency part \( S_3 \) and a high-frequency part \( D_1, D_2, D_3 \) of the original signal, as shown in Figure 6.
Data prediction
For the four signal series shown in Figure 6, the first several invalid data are eliminated and the first 160 data are taken as training samples, and the last 93 data are taken as test samples to respectively establish Elman neural network prediction model. The comparison of the predicted value and the actual value is as shown in figure 7.

Analysis of result
For the predicted comparison data shown in Figure 7, an error test is performed according to formulas (9) and (10) to obtain AAE=3.13 and RMSE=3.95. The absolute error between the predicted value and the actual value is shown in Figure 8, from which it can be seen that there is a high consistency between the predicted value and the actual value. The prediction error is small, which can reflect the rule of vehicle’s arriving at the intersection.

Figure 6. WT of original time difference series

Figure 7. Comparison of predicted value and actual value of arrival time series

Figure 8. AAE diagram of error test
Conclusions
The real-time accurate prediction of vehicle arrival parameters at signalized intersection is beneficial to the on-line optimal control of intersection. Aiming at the merits and demerits of wavelet analysis and Elman neural network, this study constructs a prediction model of first vehicle arrival time series at signalized intersection based on wavelet-Elman neural network. The application example and the result show that the prediction error of the model is very small, which can accurately describe the vehicle arrival characteristic at signalized intersection. This method is of great significance for traffic flow modeling and intelligent control at signalized intersection.

Acknowledgements
The research is supported by the National Natural Science Foundation of China (51508061), the Open Project of Chongqing Key Laboratory of Traffic & Transportation, Chongqing Jiaotong University, China (2016CQJY005) and the Open Project of National Engineering and Research Center for Highways in Mountain Area, China (gsgzj-2013-06).

References
Gao ZH, Qu ZW, Hu HY. Study on vehicle delay based on the vehicle arriving distribution at entrance lanes of intersection. Procedia Engineering 2016; 137: 599-608.
Hurdle VF. Signalized intersection delay models—a primer for the uninitiated. Transportation Research Record 1984; 971: 96-105.