An Elastic Combination Forecasting Method for Urban Road Traffic Status

Cheng Wang1*, Xiyu Pang1, Zhonghua Xi2, Guannan Si1
1Department of Information Engineering, Shandong Jiaotong University, Shandong, Jinan, 250357, China
2Shan Dong Eagle Software Technology Co., Ltd. Shandong, Jinan, 25000, China
*Corresponding author e-mail: wangcheng_1001@163.com

Abstract. In intelligent transportation system, urban road traffic flow status prediction plays an important role. Study shows that the traffic flow status has fractal phenomenon in a certain time scale, so using fractal method to excavate the inherent regularity of traffic flow time series can avoid some difficulties of analyzing the traffic flow influencing factors. This paper proposes a new forecasting algorithm of urban road traffic status based on fractal theory, and in the algorithm, the calculation of fractal dimension is based on the structural function method, and the design of the algorithm takes into account the traffic conditions at different time intervals based on the prediction time. The experimental results indicate that the proposed algorithm is capable of dealing with complex nonlinear urban traffic flow forecasting with satisfying accuracy and effectiveness.

1. Introduction
As the road traffic system has the basic characteristics of non-linearity, complexity and uncertainty, combination forecasting model, such as the combination of model algorithm and model-free algorithm, the combination of neural network theory and genetic algorithm, fuzzy theory, wavelet theory, spectral analysis and so on, has more and more extensive research and applications. Fractal theory is a branch of nonlinear scientific theory, used to describe the complexity of chaos phenomena behind the regularity, revealing the relationship between local and the whole, and the key quantitative indicators in the theory - fractal dimension can describe not only the complexity of things quantitatively, but also the change of some characteristics of things. Fractal theory aims at finding out the regularity from the fractal dimension in the seemingly disorderly, high-dimensional and dynamic data in the information world. At present, the concept of fractals and ideas has been abstracted as a methodology [1-4].

The traffic flow and the average speed have fractal phenomenon in a certain time scale, and using fractal method to excavate the inherent regularity of traffic flow time series can avoid some difficulties of analyzing the traffic flow influencing factors. However, if the predicted time interval is short and the fractal characteristic is not obvious, this brings some difficulties to the calculation of fractal dimension. This is the key point in the application of fractal theory to traffic flow forecasting.

2. The Elastic Combination Forecasting Method
This paper presents a flexible combination forecasting method of urban road traffic status and the flow of urban road traffic status combination forecasting method is shown as Fig.1.
First, from the historical database, select m traffic status sequences of the neighbor weeks with the same time points. Then, use the maximum and minimum methods to calculate the weekly similarity of traffic flow and average speed between the current traffic status sequence and the m traffic status sequences of nearest neighbor weeks. If the value of the similarity is within the acceptable range, the traffic status prediction value at the next time is estimated, if the similarity is not within the acceptable range, a prediction algorithm based on the fractal theory is used to predict the traffic flow and the average speed at the next time.

The calculation of fractal dimension is based on the structural function method. The design of the algorithm takes into account the traffic status with different time intervals based on the predicted time, and the prediction value at the next time is restored by using the reciprocal of the distance between the predicted value and the actual value of the time immediately before the predicted time of each traffic status time sequence as the weighting coefficient.

At last, the traffic congestion index of the current road is obtained by the relative evaluation method based on the traffic jam density and the optimum density of the current road section. The final traffic state vector includes the traffic flow at the forecast time, the average vehicle speed at the forecasting moment, the road density of the forecast time, and the traffic congestion index of the forecast time.

![Flow chart of urban road traffic status combination forecasting method](image)

Fig.1 Flow chart of urban road traffic status combination forecasting method
2.1. The Algorithm for Calculating the Weekly Similarity Based on the Maximum and Minimum Method

Maximum and minimum method is one of the similarity coefficient calculation methods, with a small amount of calculation, simple features, and the maximum and minimum methods are suitable for calculating the similarity between similar data sets.

Numerous studies have shown that urban road traffic status have weekly similarity. In this paper, the maximum and minimum methods are used to calculate the weekly similarity between the current traffic status sequence and the traffic status sequence of m nearest neighbor weeks with the same time moments, and m nearest neighbor weeks form the set A.

The equation is as follows:

\[ Rq_j = \frac{\sum_{i=0}^{\alpha-1} \min(q_{j(i-t \Delta t)}, q_{t-i \Delta t})}{\sum_{i=0}^{\alpha-1} \max(q_{j(i-t \Delta t)}, q_{t-i \Delta t})} \]  

(1)

\[ Rv_j = \frac{\sum_{i=0}^{\alpha-1} \min(v_{j(i-t \Delta t)}, v_{t-i \Delta t})}{\sum_{i=0}^{\alpha-1} \max(v_{j(i-t \Delta t)}, v_{t-i \Delta t})} \]  

(2)

In the equation, \(1 \leq j \leq m\), \(0 \leq i \leq \alpha-1\), \(q_{t-i \Delta t}\) represents the traffic flow at time \(t-i \Delta t\) of the current traffic status sequence, \(q_{j(i-t \Delta t)}\) represents the traffic flow at time \(t-i \Delta t\) of the traffic status sequence of the \(j^{th}\) Nearest neighbor week in the set A; \(v_{t-i \Delta t}\) represents the average speed at time \(t-i \Delta t\) of the current traffic status sequence, \(v_{j(i-t \Delta t)}\) represents the average speed at time \(t-i \Delta t\) of the traffic status sequence of the \(j^{th}\) Nearest neighbor week in the set A; \(Rq_j\) represents the similarity of the traffic flow between the current traffic status sequence and the traffic status sequence of the \(j^{th}\) Nearest neighbor week in the set A; \(Rv_j\) represents the similarity of the average speed between the current traffic status sequence and the traffic status sequence of the \(j^{th}\) Nearest neighbor week in the set A, and all the traffic status sequences have the same time moments.

After computing the weekly similarity, the sequence satisfying Equation(3) is selected from the set A, and the sequences form the set B \(B = \{Z_{x_1}, Z_{x_2}, \ldots, Z_{x_h}\}\):

\[ Rq_j \geq S_q \land Rv_j \geq S_v \]  

(3)

In the equation, \(h\) represents the number of sequences in set B, and \(Z_{x \Delta t} = \{z_{x(t-(n-1) \Delta t)}, z_{x(t-(n-2) \Delta t)}, \ldots, z_{x(t-(2 \Delta t)}, z_{x(t-\Delta t)}, z_{x(t)}\}\) represents the traffic status sequence of the \(x^{th}\) nearest neighbor week in the set B, \(1 \leq x \leq h\), \(S_q\) represents the threshold of traffic flow similarity, \(S_v\) represents the threshold of average speed similarity. If there is a sequence satisfying Equation(3), which is, \(h > 0\), then calculate the traffic flow and average speed at the next time according to Equation(4) and Equation(5).
\[ q'_{t+\Delta t} = \sum_{x=1}^{h} \frac{Rq_x}{h} \cdot q_{x(t+\Delta t)} \]  
(4)

\[ v'_{t+\Delta t} = \sum_{x=1}^{h} \frac{Rv_x}{h} \cdot v_{x(t+\Delta t)} \]  
(5)

In the equation, \( Rq_x \) represents the weekly similarity of traffic flow between the current traffic status sequence and the traffic status sequence of the \( x \)th nearest neighbor week in the set B. \( Rv_x \) represents the weekly similarity of average speed between the current traffic status sequence and the traffic status sequence of the \( x \)th nearest neighbor week in the set B.

If there is no sequence that satisfies Equation(3), the traffic status prediction algorithm based on fractal theory is used to forecast the traffic flow and average speed.

2.2. The forecast algorithm based on fractal theory

The flow chart of the forecast algorithm based on fractal theory is shown in Fig.2. First, The forecast algorithm is based on the forecast time \( t + \Delta t \) and select the \( n + 1 \) nearest neighbor traffic time points within the intervals \( \Delta t \) and \( \Delta t \) and \( \Delta t \). These traffic status sequences form the set \( Z = \{Z_{1 \Delta t}, Z_{2 \Delta t}, ..., Z_{\lambda_{\text{max}} \Delta t}\} \). That is to say, all the traffic status sequences include the forecast time \( t + \Delta t \), and the interval of the time points contained in a traffic status sequence is an integer multiple of \( \Delta t \). Likewise, the traffic status at each time point include traffic flow and average speed, and the total number of traffic sequences is \( \lambda_{\text{max}} \), and the set Z is shown as Equation(6):

\[
\begin{align*}
Z_{1 \Delta t} &= \left\{ z_{1(t-\Delta t)}, z_{1(t-2\Delta t)}, ..., z_{1(t-2\Delta t)}, z_{1(t-\Delta t)}, z_{1(t+\Delta t)} \right\} \\
Z_{2 \Delta t} &= \left\{ z_{2(t-2\Delta t)}, z_{2(t-3\Delta t)}, ..., z_{2(t-3\Delta t)}, z_{2(t-\Delta t)}, z_{2(t+\Delta t)} \right\} \\
Z_{\lambda \Delta t} &= \left\{ z_{\lambda(t-\lambda \Delta t)}, z_{\lambda(t-\lambda \Delta t)}, ..., z_{\lambda(t-\lambda \Delta t)}, z_{\lambda(t-\Delta t)}, z_{\lambda(t+\Delta t)} \right\}
\end{align*}
\]  
(6)

In the equation, each traffic status sequence is based on the forecast time point \( t + \Delta t \).

Then, the fractal dimension of the traffic flow and the fractal dimension of the average speed are calculated by using the structural function method as the basis. And the algorithm steps are as follows:

1. For each sequence in set Z, calculate respectively the arithmetic mean of the difference of the traffic status sequence with the time interval is \( \lambda \Delta t \) \( \lambda = (1, 2, 3, ..., \lambda_{\text{max}}) \), using Equation(7) and Equation(8):
\begin{equation}
S_q(\lambda \Delta t) = \frac{\sum_{i=2}^{n} [q_{\lambda(t-(i-1)\lambda \Delta t)} - q_{\lambda(t-\lambda \Delta t)}]^2}{n-1} \tag{7}
\end{equation}

\begin{equation}
S_v(\lambda \Delta t) = \frac{\sum_{i=2}^{n} [v_{\lambda(t-(i-1)\lambda \Delta t)} - v_{\lambda(t-\lambda \Delta t)}]^2}{n-1} \tag{8}
\end{equation}

select the \(n+1\) nearest neighbor traffic time points form a traffic status sequence based on the forecast time, with the different time interval; and these traffic status sequence form a set \(Z\).

Calculate the fractal dimension of the traffic flow and the fractal dimension of the average speed by using the structural function method as the basis.

Calculate the distance between the real traffic flow value and the forecast traffic flow value of the previous time point adjacent to the forecast time of each traffic status sequence in the set \(Z\) as Weighting factor.

calculate the traffic flow forecast value and the average speed forecast value of the forecast time point.

Fig. 2 Flow chart of forecast algorithm based on fractal theory

2). Plot \([\lambda \Delta t, S_q(\lambda \Delta t)]\) onto a coordinate plot, where \(\ln(\lambda \Delta t)\) is the x-axis and \(\ln(S_q(\lambda \Delta t))\) is the y-axis. According to Equation(9), the fractal dimension \(D_q\) and the constant \(C_q\) of the traffic flow are calculated by linear regression of the least squares method.

Similarly, according to Equation (10), the values of the fractal dimension \(D_v\) of the average speed are calculated.

\begin{equation}
S_q(\lambda \Delta t) = C_q ^4 \cdot (\lambda \Delta t) ^{4-2D_q} \tag{9}
\end{equation}

\begin{equation}
S_v(\lambda \Delta t) = C_v ^4 \cdot (\lambda \Delta t) ^{4-2D_v} \tag{10}
\end{equation}

After calculating the fractal dimension, using the fractal dimension \(D_q\), the constant \(C_q\) and Equation(7) and Equation(9), the forecast value \(q_{t-\lambda \Delta t}\) of the traffic flow at the previous time point...
adjacent to the forecast time \( t + \Delta t \) for each traffic status sequence are calculated, and then, calculate the distance \( d_{q\lambda} = (1, 2, 3, 4, ..., \lambda_{\text{max}}) \) between the real traffic flow value and the forecast traffic flow value of the previous time point adjacent to the forecast time \( t + \Delta t \).

Similarly, using the fractal dimension \( D_q \), the constant \( C_q \) and Equation(8) and Equation(10), the forecast value \( q'_{\lambda(t+\Delta t)} \) of the average speed at the previous time point adjacent to the forecast time \( t + \Delta t \) for each traffic status sequence are calculated, and then, calculate the distance \( d_{v\lambda} = (1, 2, 3, 4, ..., \lambda_{\text{max}}) \) between the real average speed value and the forecast average speed value of the previous time point adjacent to the forecast time \( t + \Delta t \).

The fractal dimension \( D_q \) and the constant \( C_q \) are taken into Equation(11) to calculate the traffic flow forecast value \( q'_{\lambda(t+\Delta t)} \) of the forecast time point \( t + \Delta t \) for each traffic status sequence. Similarly, the fractal dimension \( D_v \) and the constant \( C_v \) are taken into Equation(12) to calculate the average speed forecast value \( v'_{\lambda(t+\Delta t)} \) of the forecast time point \( t + \Delta t \) for each traffic status sequence.

\[
q'_{\lambda(t+\Delta t)} = \sqrt{n \cdot C_q \cdot (\lambda \Delta t)^{4-2D_q} - \sum_{i=2}^{n} \left( \left( q_{\lambda(t-(i-1)\Delta t+\Delta t)} - q_{\lambda(t-i\Delta t+\Delta t)} \right) \right)^2 + q_{\lambda(t-i\Delta t+\Delta t)}}
\]

\[
v'_{\lambda(t+\Delta t)} = \sqrt{n \cdot C_v \cdot (\lambda \Delta t)^{4-2D_v} - \sum_{i=2}^{n} \left( \left( v_{\lambda(t-(i-1)\Delta t+\Delta t)} - v_{\lambda(t-i\Delta t+\Delta t)} \right) \right)^2 + v_{\lambda(t-i\Delta t+\Delta t)}}
\]

Finally, with the inverse of the distance as the weighting factor, the forecast value of the traffic flow and the forecast value of the average speed at the forecast time \( t + \Delta t \) are calculated according to Equation(13) and Equation(14):

\[
q'_{\lambda(t+\Delta t)} = \sum_{\lambda=1}^{\lambda_{\text{max}}} \frac{\lambda_{\text{max}}}{} \sum_{i=1}^{1} \frac{1}{d_{q\lambda}}
\]

\[
v'_{\lambda(t+\Delta t)} = \sum_{\lambda=1}^{\lambda_{\text{max}}} \frac{\lambda_{\text{max}}}{} \sum_{i=1}^{1} \frac{1}{d_{v\lambda}}
\]

2.3. The algorithm of traffic congestion index based on relative evaluation

Suppose the traffic congestion density of the current road section is \( K_{\text{stop}} \) \((K_{\text{stop}}^{\text{min}} \leq K_{\text{stop}} \leq K_{\text{stop}}^{\text{max}})\), the traffic congestion density is the density in the case where the current road section is occupied by the vehicle at all, the speed is zero. \( K_{\text{smooth}} \) \((K_{\text{smooth}}^{\text{min}} \leq K_{\text{smooth}} \leq K_{\text{smooth}}^{\text{max}})\) represents the best density of the current road section, then the current road traffic congestion index is calculated as Equation(15):
Where, the vehicle density is calculated by Equation(16):

\[
 B'_{t+\Delta t} = \begin{cases} 
 \left( k'_{t+\Delta t} - K_{\text{min}}^{\text{smooth}} \right) / K_{\text{min}}^{\text{smooth}}, & 0 \leq k'_{t+\Delta t} < K_{\text{min}}^{\text{smooth}} \\
 k'_{t+\Delta t}, & 0, K_{\text{min}}^{\text{smooth}} \leq k'_{t+\Delta t} \leq K_{\text{max}}^{\text{smooth}} \\
 \left( k'_{t+\Delta t} - K_{\text{max}}^{\text{smooth}} \right) / (K_{\text{stop}}^{\text{smooth}} - K_{\text{max}}^{\text{smooth}}), & K_{\text{max}}^{\text{smooth}} < k'_{t+\Delta t} < K_{\text{stop}}^{\text{smooth}} \\
 k_{t+\Delta t}^{\text{smooth}}, & k'_{t+\Delta t} \geq K_{\text{stop}}^{\text{smooth}} 
\end{cases}
\]

(15)

Finally, the traffic status prediction vector of the current road section is obtained as follows:

\[
 R'_{t+\Delta t} = \left[ q'_{t+\Delta t}, v'_{t+\Delta t}, k'_{t+\Delta t}, B'_{t+\Delta t} \right]
\]

It includes: the traffic flow of the prediction time, the average speed of the prediction time, the road density of the prediction time, the traffic congestion index of the prediction time.

3. Conclusion

In this paper, the maximum and minimum methods and the prediction method based on fractal theory are used in urban road traffic flow forecasting. A flexible urban traffic status combination algorithm is proposed. The algorithm has the advantages of large computation flexibility, higher forecast precision, and outputting traffic status data comprehensively.

Fractal theory has long been a research hotpot. Nowadays, fractal theory has been applied in the fields of finance, water quality, mineral, and living standard, but it is still in the initial stage and still needs further exploration in the field of transportation.

Acknowledgments

The research work was supported by Shandong Provincial Natural Science Foundation, A Project of Shandong Province Higher Educational Science and Technology Program under grant NO.J17KA073 and Shandong Province key Research and Development Plan Project under grant NO.GG201710080127 and Shandong Province Housing Urban and Rural Construction Science and Technology Project Plan under grant No.2016-K8-057.

References

[1] Kmianakis Y, Shen WL. Real-time road traffic forecasting using regime-switching space-time models and adaptive LASSO. Applied Stochastic Models in Business and Industry, 2012, 28(4): 297-315.

[2] Abdar, SRN Kalhori, T Sutikno, IMI Subroto, G Arji. Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases. International Journal of Electrical and Computer Engineering (IJECE), 2015, 5(6): 1569-1576.

[3] Hong, WC. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. Neuro computing, 2011, 74(12): 2096-2107.

[4] Fan Na, Zhao Xiang-mo, Dai Ming. Short-term traffic flow prediction model. Journal of Traffic and Transportation Engineering, 2012, 12(4): 114-119.