

ROAD TRAFFIC FLOW FOREWARNING AND CONTROL MODEL WITH THE SLOPE OF THE CHANGE RATE

Xiao Qiang, He Ruichun, Yu Jianning, Zhang Wei, Ma Changxi

Preliminary communication

Real-time, accurate and efficiency short term traffic flow prediction is one of the key technologies to realize traffic flow guidance and traffic control, which has been widely concerned in the domain of ITS (Intelligent Transport System) during recent years. Through the study of the existing traffic flow prediction model, road traffic flow control model with the slope of the change rate is proposed. The model can find out abnormal point from the traffic flow time series by the use of the slope change rate, and it can analyse this trend of traffic flow changes for control purposes of traffic flow. The achieved results indicate that the algorithm is suitable for road traffic flow peak control problem and could be effective for road traffic flow control.

Keywords: *traffic flow; slope change rate; time series; abnormal data*

Model za reguliranje toka cestovnog prometa padom promjene brzine

Prethodno priopćenje

Zadnjih je godina točno i učinkovito kratkoročno predviđanje toka prometa u realnom vremenu jedna od ključnih tehnologija u ostvarenju upravljanja i reguliranja tokom cestovnog prometa iz ITS područja (Intelligent Transport System). Analizirajući postojeći model predviđanja toka prometa, predlaže se model za reguliranje cestovnog toka prometa. Model može pronaći nenormalnu točku analizom vremenskih serija toka prometa primjenom pada promjene brzine (slope change rate), i može analizirati taj trend promjena toka prometa u svrhu reguliranja toka prometa. Rezultati pokazuju da je algoritam pogodan za problem reguliranja vršnog cestovnog opterećenja prometa, a može biti učinkovit u reguliranju cestovnog prometa.

Ključne riječi: *tok prometa; pad promjene brzine; vremenske serije; nenormalni podaci*

1 Introduction

Traffic jams in urban areas have attracted widespread attention in a global context, with the development of socio-economic and transport industry [1, 2]. Particularly in China, car ownership has reached millions of vehicles [3], but the Chinese road construction now falls behind development of the automotive industry. Traffic congestion has become a widespread social phenomenon in Beijing, Shanghai, Guangzhou and other cities, and naturally it has become people's focus.

To some extent, traffic congestion has become an important problem in economic development and social stability, however, limited land and restricted economic and other factors fail to achieve relatively satisfactory road construction mileage [4, 5]. Therefore, the use of modern information technology and data analysis techniques to improve transportation efficiency and the traffic capacity of the road network is necessary and it has become the way to solve the road congestion in terms of the existing conditions.

At present, the research on traffic flow prediction methods is mainly focused on the traditional methods based on statistical theory, the method based on neural network, the method based on nonlinear theory and the method based on combined algorithm [6]. Traditional theoretical predictions include: the historical average model, regression analysis, time series models, Kalman filter model and Markov etc. [7]. However, these methods model is relatively simple, and the prediction accuracy is poor, although it can solve the problem of changing the traffic flow at different times, it is a poor deal for unexpected traffic flow and other special circumstances. Artificial neural network technology was born in the 1940s, and it was used for long-term traffic flow prediction by Chin in 1992, Dougherly and Clark, who

respectively used neural network technology for short-term traffic flow forecasting in 1993 and 1994 [8]. The emergence of neural network algorithm got rid of the trouble to establish a precise mathematical model for the study of work, which opened up a new way of thinking. However, the accurate prediction of neural networks is based on a large number of historical training or unknown data, the amount of data train and the length of training time, which determines the accuracy of the model prediction. Therefore, if it is predicted to no historical data, there will be a big predictive error. Nonlinear prediction mainly in chaos theory [9, 10], dissipative structure theory, self-organization theory and other theories as theoretical foundation, uses the concept of chaotic attractor, fractal concept phase space reconstruction methods to build predictive models. However, this model includes many parameters, and is not easy to determine. Thus, it makes nonlinear algorithms in prediction accurate so that it is not easy to control and reduce the usefulness of the nonlinear algorithm. Combined algorithm is based on the advantages of various algorithms [11, 12], which takes complementary advantages principle and combines the advantages of different algorithms, and makes up for the lack of a single algorithm, leading the wide application of the type of prediction algorithm. However, this embodiment is relatively complex to structure of the algorithm, and the algorithm is more difficult to fuse in practical application, which means there are still many problems.

Using traffic data to predict traffic patterns, as a research direction, is studied by a lot of scholars. The given traffic prediction models become inaccurate under sensor malfunction or noise-contaminated data for a certain period of time in part of a transportation network. This problem frequently occurs in transportation networks

[13–15]. Many studies address the issue of traffic prediction with partially missing traffic data. For instance, van Lint et al. [23] presented a neural network for travel time prediction under missing traffic data. Sun et al. [16, 17] introduced a Bayesian method to forecast traffic flows where a certain period of historical data is missing for some links of the transportation network. The missing portion of historical traffic data is approximated by using a Gaussian mixture model. Moreover, other statistical and probabilistic methods are used to address the missing traffic data problems [18–22]. Therefore, the problem of short-term traffic flow prediction based on completely unavailable traffic data in some links due to lack of sensors in the network is an open challenging problem.

In this paper, traffic flow prediction model with abnormal data mining algorithm has been proposed in current research methods. Based on the linear slope, model algorithm makes use of the slope calculation principle and the technology of data mining. To find out the abnormal data in traffic flow, it uses these abnormal points as traffic forewarning basis, helps the traffic management departments to make reasonable decision, guides the car pass smoothly, and avoids congestion phenomenon.

The rest of this paper is organized as follows: Section 2 describes the overall structure of early forewarning model based on data mining technology. Section 3 presents abnormal data mining algorithms based on the slope calculation principles, as well as early forewarning processing approach. Section 4 introduces algorithm simulation and prediction processing simulation by the use of real data. Section 5 summarizes the paper and outlines the application of the model algorithm.

2 Mode descriptions

In order to realize the traffic flow prediction, predict the road traffic conditions, make reasonable traffic guidance and avoid the occurrence of traffic jams, the early forewarning control model of traffic is presented. Realization diagram of the model is as follows in Fig. 1.

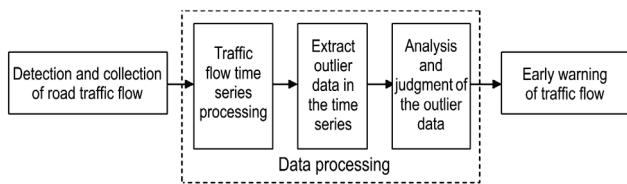


Figure 1 Processing model of traffic flow based on abnormal data mining

2.1 Detection and collection of road traffic flow

The forewarning effect to sources of traffic data in traffic flow forewarning will be determined. Therefore in this paper, which uses traffic flow detector acquisition technology based on microwave detector RTMS (Remote Traffic Microwave Sensor) [7], it can achieve traffic flow data collection.

As can be seen from Fig. 2, this road is a crossroads. Detector is placed in the four corners of the road, two detectors separated by 150 meters on each road. Within the prescribed time, the collection data are detected by the

two detectors on each road. The collected data according to $S_i(d_i, t_i)$ for storage, where S_i is the acquisition time sequence group in i roads and i time, d_i denotes the road occupancy rate, t_i is collection point in time.

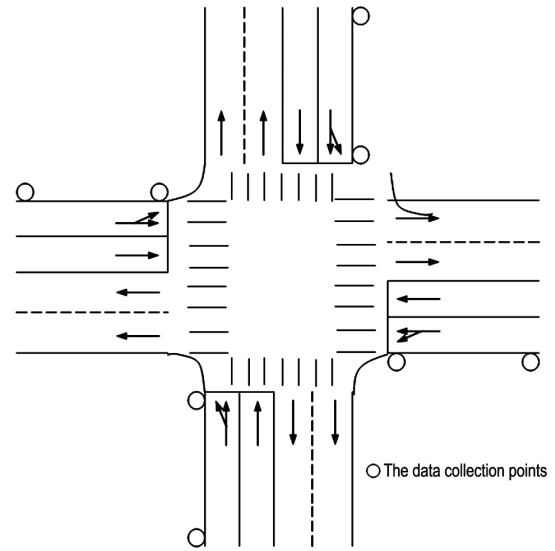


Figure 2 Schematic of traffic flow data collection

2.2 Traffic flow data processing

Time series of traffic flow will be calculated by use of the slope of the principle for abnormal data mining. With the abnormal points from the abnormal data analysed, it can determine the road traffic situation which will be changed. From the analysis of results, this change will be sent to confident control department. Traffic management departments will use the relevant processing mode, such as traffic electronic screen guidance, traffic broadcast, and traffic grooming. These ways will be used to change the route and prevent traffic jams (Fig. 3).

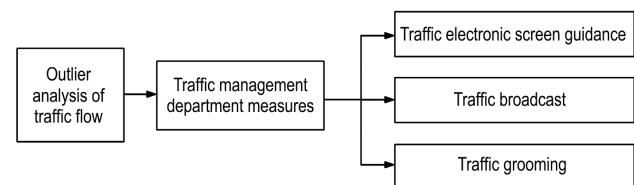


Figure 3 Traffic forewarning handling diagram

3 Traffic flow data processing and alert processing

In order to find the traffic flow abnormal data, road traffic flow control model with the slope of the change rate is proposed in this paper. The algorithm included collecting the road traffic flow data and determined the data point in the time axis and computed the slope of the line passing through each pair of adjacent data points.

Definitions 3.1: Any three consecutive points are s_{i-1} and s_i and s_{i+1} in traffic flow time series, k_i is the slope the straight line $s_{i-1}s_i$, k_{i+1} is the slope the straight line $s_i s_{i+1}$, these two straight lines directions are different, if $k_i \cdot k_{i+1} < 0$ These two straight lines directions are the same, if $k_i \cdot k_{i+1} > 0$ [24].

Corollary 3.1: Given the slope change threshold of two straight lines with m_1 , if the two straight lines directions are different. There is no abnormal data, if

$|k_{i+1}-k_i|<m_1$. There is abnormal data, if $|k_{i+1}-k_i|>m_1$, and s_i is considered abnormal point.

Corollary 3.2: Given the slope change threshold of two straight lines with m_2 , if the two straight lines directions are the same. There is no abnormal data, if $|k_{i+1}|<m_2$. There is abnormal data, if $|k_{i+1}|>m_2$, and s_i is considered abnormal point.

Definition 3.1 is proved:

Given three consecutive points with s_{i-1} and s_i and s_{i+1} in traffic flow series times, their axes are (x_{i-1}, y_{i-1}) and (x_i, y_i) and (x_{i+1}, y_{i+1}) in the coordinate axes, k_i is the slope the straight line $s_{i-1}s_i$, k_{i+1} is the slope the straight line $s_i s_{i+1}$. The following four cases will be discussed, because the two straight lines which can be constructed by three consecutive points have different direction.

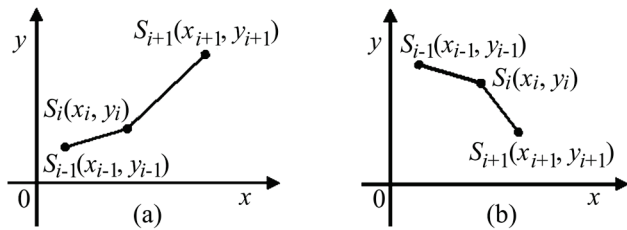


Figure 4 The two straight lines have the same direction

Case 1: If $x_{i-1}<x_i$, $y_{i-1}<y_i$ and $x_i<x_{i+1}$, $y_i<y_{i+1}$, the establishment of the formula (1):

$$K_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} > 0, K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} > 0. \quad (1)$$

According to the principle of slope, $K_i>0$ and $K_{i+1}>0$, as can be seen from Fig. 4(a). Two lines have the same directions, and their directions are upward and rightward.

Case 2: If $x_{i-1}<x_i$, $y_{i-1}>y_i$ and $x_i<x_{i+1}$, $y_i>y_{i+1}$, the establishment of the formula (2):

$$K_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} < 0, K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} < 0. \quad (2)$$

According to the principle of slope, $K_i<0$ and $K_{i+1}<0$, as can be seen from Fig. 4(b). Two lines have the same directions, and their directions are up down and rightward.

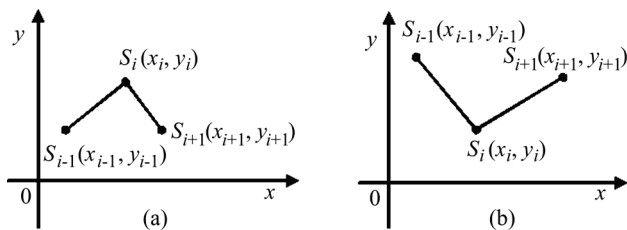


Figure 5 The two straight lines have different direction

Case 3: If $x_{i-1}<x_i$, $y_{i-1}<y_i$ and $x_i<x_{i+1}$, $y_i>y_{i+1}$, the establishment of the formula (3):

$$K_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} > 0, K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} < 0. \quad (3)$$

According to the principle of slope, $K_i>0$ and $K_{i+1}<0$, as can be seen from Fig. 5(a), lines $S_{i-1}S_i$ direction is upward and rightward, lines $S_i S_{i+1}$ direction is up down and rightward.

Case 4: If $x_{i-1}<x_i$, $y_{i-1}>y_i$ and $x_i<x_{i+1}$, $y_i<y_{i+1}$, the establishment of the formula (4):

$$K_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}} > 0, K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} < 0. \quad (4)$$

According to the principle of slope, $K_i<0$ and $K_{i+1}>0$, as can be seen from Fig. 5(b), lines $S_{i-1}S_i$ direction is up down and rightward, lines $S_i S_{i+1}$ direction is upward and rightward.

Corollary 3.1 is proved:

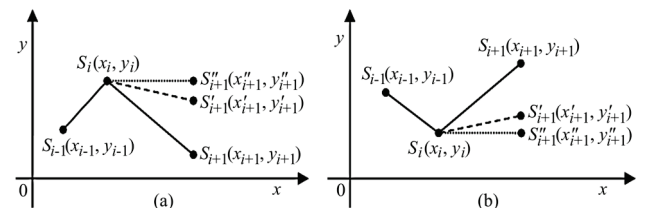


Figure 6 Abnormal point between different direction of the line

(1) Assuming the two lines have different direction, and meet the following requirements $x_{i-1} < x_i$, $y_{i-1} < y_i$ and $x_i < x_{i+1}$, $y_i > y_{i+1}$. Let S'_{i+1} and S''_{i+1} be two points in coordinate axis, their coordinates are (x'_{i+1}, y'_{i+1}) and (x''_{i+1}, y''_{i+1}) , and meet the following requirements $x_{i+1} = x'_{i+1} = x''_{i+1} > x_i$, $y_{i+1} > y'_{i+1} > y''_{i+1} > y_i$. As can be seen from Fig. 6(a), the line slope of $S_i S'_{i+1}$ and $S_i S''_{i+1}$ and $S_i S''_{i+1}$ can be given:

$$K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, K'_{i+1} = \frac{y'_{i+1} - y_i}{x'_{i+1} - x_i}, K''_{i+1} = \frac{y''_{i+1} - y_i}{x''_{i+1} - x_i}. \quad (5)$$

According to the case 3 and the slope theory, there are the following relations

$$|K_{i+1} - K_i| > |K'_{i+1} - K_i| > |K''_{i+1} - K_i|. \quad (6)$$

As can be seen from formula (6), the slope change rate of $S_i S_{i+1}$ and $S_{i-1} S_i$ is larger than the other slope change rate. Setting the threshold is m_1 in the slope change rate, if the slope change rate is bigger than m_1 , there is abnormal data, and s_i is considered abnormal point.

(2) Assuming the two lines have different direction, and meeting the following requirements $x_{i-1} < x_i$, $y_{i-1} > y_i$ and $x_i < x_{i+1}$, $y_i < y_{i+1}$. Let S'_{i+1} and S''_{i+1} be two points in coordinate axis, their coordinates are (x'_{i+1}, y'_{i+1}) and (x''_{i+1}, y''_{i+1}) , and meet the following requirements $x_{i+1} = x'_{i+1} = x''_{i+1} > x_i$, $y_{i+1} > y'_{i+1} > y''_{i+1} > y_i$. As can be seen from Fig. 6(b), the line slope of $S_i S'_{i+1}$ and $S_i S''_{i+1}$ and $S_i S''_{i+1}$ can be given:

$$K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, K'_{i+1} = \frac{y'_{i+1} - y_i}{x'_{i+1} - x_i}, K''_{i+1} = \frac{y''_{i+1} - y_i}{x''_{i+1} - x_i}. \quad (7)$$

According to the case 4 and the slope theory, there are the following relations

$$|K_{i+1} - K_i| > |K'_{i+1} - K_i| > |K''_{i+1} - K_i|. \tag{8}$$

As can be seen from formula (8), the slope change rate of $S_i S_{i+1}$ and $S_{i-1} S_i$ is larger than the other slope change rate. Setting the threshold is m_1 in the slope change rate, if the slope change rate is bigger than m_1 , there is abnormal data, and s_i is considered abnormal point.

Corollary 3.2 is proved:

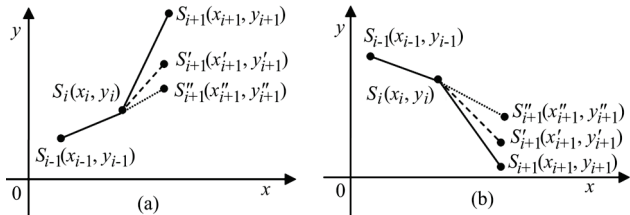


Figure 7 Abnormal point between the different direction of the line

(1) The two lines have the same direction, and meet the following requirements $x_{i-1} < x_i$, $y_{i-1} < y_i$ and $x_i < x_{i+1}$, $y_i < y_{i+1}$. Let S'_{i+1} and S''_{i+1} are two points in coordinate axis, their coordinates are (x'_{i+1}, y'_{i+1}) and (x''_{i+1}, y''_{i+1}) , and meet the following requirements $x_{i+1} = x'_{i+1} = x''_{i+1} > x_i$, $y_{i+1} > y'_{i+1} > y''_{i+1} > y_i$. As can be seen from Fig. 7(a), the line slope of $S_i S_{i+1}$ and $S_i S'_{i+1}$ and $S_i S''_{i+1}$ can be given:

$$K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, K'_{i+1} = \frac{y'_{i+1} - y_i}{x'_{i+1} - x_i}, K''_{i+1} = \frac{y''_{i+1} - y_i}{x''_{i+1} - x_i}. \tag{9}$$

According to the case 3 and the slope theory, there are the following relations

$$|K_{i+1}| > |K'_{i+1}| > |K''_{i+1}|. \tag{10}$$

As can be seen from formula (10), the slope change rate of $S_i S_{i+1}$ and $S_{i-1} S_i$ is larger than the other slope change rate. Setting the threshold is m_2 in the slope change rate, if the slope change rate is bigger than m_2 , there is abnormal data, and s_i is considered abnormal point.

(2) Assuming the two lines have the same direction, and meeting the following requirements $x_{i-1} < x_i$, $y_{i-1} > y_i$ and $x_i < x_{i+1}$, $y_i > y_{i+1}$. Let S'_{i+1} and S''_{i+1} be two points in coordinate axis, their coordinates are (x'_{i+1}, y'_{i+1}) and (x''_{i+1}, y''_{i+1}) , and meet the following requirements $x_{i+1} = x'_{i+1} = x''_{i+1} > x_i$, $y_{i+1} < y'_{i+1} < y''_{i+1} < y_i$. As can be seen from Fig. 7(b), the line slope of $S_i S_{i+1}$ and $S_i S'_{i+1}$ and $S_i S''_{i+1}$ can be given:

$$K_{i+1} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, K'_{i+1} = \frac{y'_{i+1} - y_i}{x'_{i+1} - x_i}, K''_{i+1} = \frac{y''_{i+1} - y_i}{x''_{i+1} - x_i}. \tag{11}$$

According to the case 4 and the slope theory, there are the following relations

$$|K_{i+1}| > |K'_{i+1}| > |K''_{i+1}|. \tag{12}$$

As can be seen from formula (12), the slope change rate of $S_i S_{i+1}$ and $S_{i-1} S_i$ is larger than the other slope change rate. Setting the threshold to m_2 in the slope change rate, if the slope change rate is bigger than m_2 , there is abnormal data, and s_i is considered abnormal point.

3.2 Algorithm and program of traffic flow abnormal data mining

According to the above principle, the traffic flow sets $S = \{S_1, S_2, \dots, S_n\}$ and the slope sets $K = \{K_1, K_2, \dots, K_n\}$ can be calculated and decided as follows. Where $S_i = (d_i, t_i)$, $S_2 = (d_2, t_2)$, $S_n = (d_n, t_n)$, S_n is data point, d_n is traffic flow value, t_n is a collection point in time, K_n is a slope of linear which can be constructed by two continuous points.

(1) If $k_i > 0$, $k_{i+1} > 0$ and $k_i \cdot k_{i+1} > 0$, $|k_{i+1} - k_i| > m_1$, there is abnormal data, and s_i is considered as an abnormal point, this abnormal point s_i means traffic flows continue to show signs of improvement.

(2) If $k_i < 0$, $k_{i+1} < 0$ and $k_i \cdot k_{i+1} > 0$, $|k_{i+1} - k_i| > m_1$, there is abnormal data, and s_i is considered as an abnormal point, this abnormal point s_i means traffic flow continues to show signs of decrease.

(3) If $k_i > 0$, $k_{i+1} < 0$ and $k_i \cdot k_{i+1} < 0$, $|k_{i+1}| > m_2$, there is an abnormal data, and s_i is considered as an abnormal point, this abnormal point s_i means the traffic lights have changed from large to small.

(4) If $k_i < 0$, $k_{i+1} > 0$ and $k_i \cdot k_{i+1} < 0$, $|k_{i+1}| > m_2$, there is an abnormal data, and s_i is considered as an abnormal point, this abnormal point s_i means the traffic lights have changed from small to large.

In order to achieve the exception processing of traffic flow and set array SN_1 as an abnormal storage point of traffic flow which include these continuous points to rise or change from small to large, set array SN_2 as an abnormal storage points of traffic flow and include these points of continuous to falling or from large to small.

3.3 Outlier identification and early forewarning

Array SN_1 and SN_2 is analysed, to determine the current road conditions. According to the result of processing, the traffic administration can decide with the abnormal points processing results to realize traffic flow data processing and prove experiment simulation testing results, there is an assuming condition that these measures can be effectively reduced by 30 % in large traffic flow, these measures include light time changes and the traffic police guidance and electronic screen guidance and traffic radio guidance.

4 Experimental emulation

In order to demonstrate the effectiveness of the algorithm model, this model used traffic data from the university of Minnesota Duluth (<http://www.d.jumn.edu/tdrl/index.htm>). Their traffic data research laboratory depends on the traffic management center of Minnesota, and by more than 5000 toroidal coil detectors to collect all the year round, and the data stored in the "traffic." format. This article downloaded data from <http://www.d.jumn.edu/tdrl/traffic/>, and we used the software extracted data for analysis, and obtained the traffic flow data. We adopted data on October 1, 2013, on September 20, on September 4, stimulated by the MATLAB software. The computer

used was Pentium (R) Dual - core CPU E5300 2.6 GHZ, 1.98 GB of memory. The experiment was stimulated by R2007b MATLAB software.

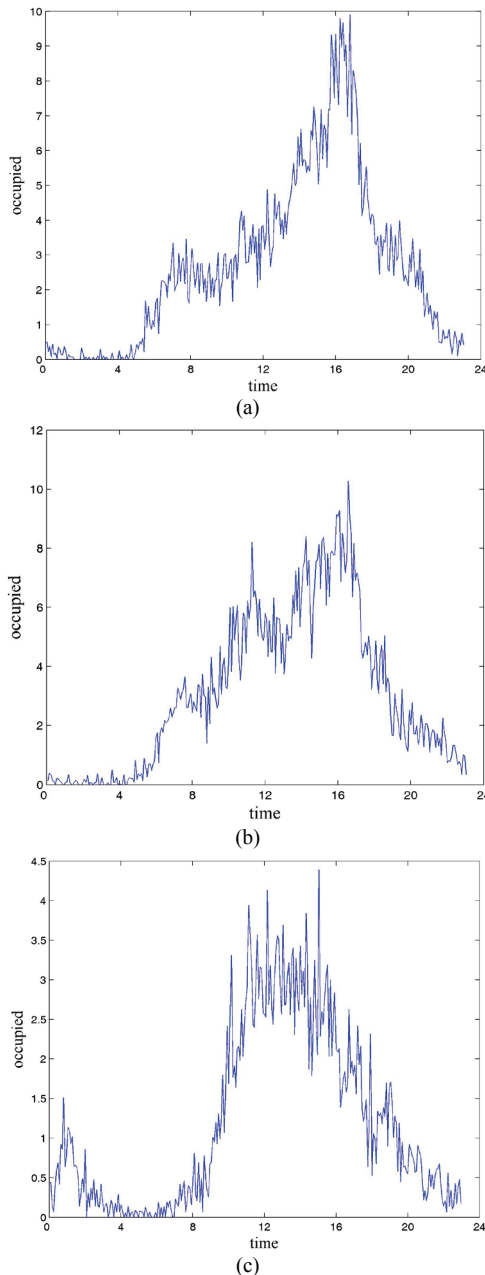


Figure 8 Traffic flow: a) Traffic flow on October 1, b) Traffic flow on September 20, c) Traffic flow on September 4.

As can be seen from Fig. 8, on October 1st, September 20, September 4, road traffic flow was small at 0:00-7:00 PM and 18:00-24:00 PM, and it shows that the capacity of road is good. At 11:00-16:00, traffic reached the peak, road traffic flow was larger. In this period of time, the road traffic flow is high, the capacity of road is poor, and it is easy to cause the road congestion.

To test and verify the feasibility of the algorithm, traffic flow time series are processed by using the slope gradient algorithm principle of time series and different slope gradient thresholds. As can be seen from Fig. 9(a) and Fig. 9(b), the early forewarning and control effect of traffic flow time series depends on the threshold; if the threshold is large, the forewarning and control effect is

better at rush hour. If the threshold is small, the forewarning and controlling effect is worse at rush hour. But if the threshold is set too high, traffic flow control effect will decline, the reason is abnormal data points of the traffic flow time series were difficult to find by using the principle of the slope change rate, and unable to realize the early forewarning and controlling of traffic flow.

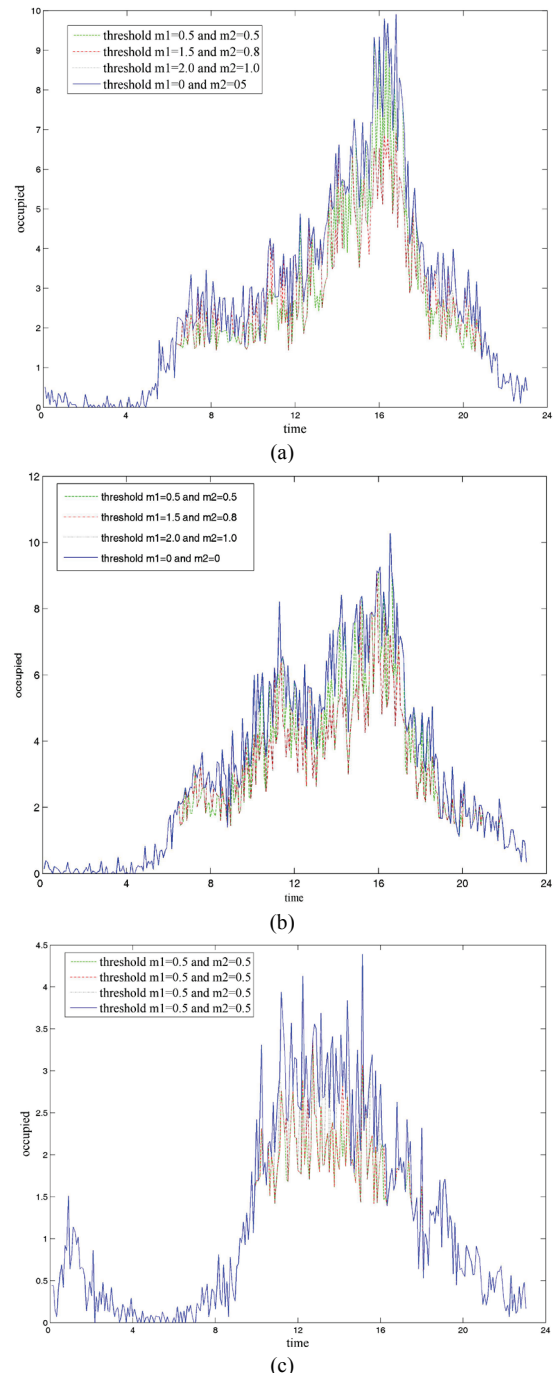


Figure 9 Different slope gradient thresholds of road traffic stream processing: a) Traffic flow process on October 1st, b) Traffic flow process September 20, c) Traffic flow process September 4.

As you can see from Fig. 9(c), for the low traffic flow on 1 October peak, forewarning threshold set at a small traffic flow control effect is better than the threshold value set. This algorithm is suitable for the peak forewarning and control of traffic flow in the threshold setting is opposite bigger, to the low peak forewarning

and control of traffic flow in the threshold value set is relatively small. As can be seen from Fig. 9(c), for traffic flow peak is low on 1 October, forewarning threshold is set small to control the effect of traffic flow better than the threshold value is set large. This means this algorithm is suitable for the large traffic flow peak, so if the early traffic flow peak is large, the threshold should be set large, the forewarning and control effect of traffic flows is better. If the early traffic flow peak is small, the threshold should be set small, the forewarning and control effect of traffic flow is better.

In short, the advantages of this algorithm do not rely on historical data and the size of the data, these data can be real-time collection, real-time processing and real-time analysis, this algorithm reduces the complexity of road traffic flow control.

5 Conclusions

Because time series data are massive and complex, storage and computation cost a high price, and if the data mining directly on the time series these may affect the accuracy and reliability of the algorithm. This paper draws on the basic theory of analytic geometry in slope, and combines the slope of the line segment with the characteristics of time series, the linear slope change of adjacent data points on the time axis is used to analyse the abnormal points in the traffic flow, so take the control measures to avoid road congestion.

According to the experimental test, we get some conclusions:

(1) Traffic reaches the peak, threshold is set large better than threshold is set small - it can effectively reduce the traffic flow and avoid road congestion. However, if the threshold setting is too large, the control effect of traffic flow will be decreased.

(2) Traffic reaches a low peak, threshold is set small better than threshold is set large - it can effectively reduce the traffic flow and avoid road congestion. However, if the threshold setting is too small, traffic flow control is difficult to achieve.

The algorithm is feasible in traffic flow abnormal data mining, but it can be seen from this paper, algorithm of the key is set the threshold. The threshold value determines the accuracy of the prediction and control, it will be discussed in the subsequent paper.

Acknowledgements

The work is supported by the National Natural Science Foundation of China (Grant no. 6106402 and 61364026 and 51408288) and young scientific research Foundation of LAN Zhou Jiao Tong University (Grant no. 2011044).

6 References

- [1] Chen, D. W.; Zheng, C. Q.; Zhang, C. B. An Algorithm for Judging Abnormal Data of Expressway Traffic Flow and Its Validation. // *China Safety Science Journal*. 16, 7(2006), pp. 122-127
- [2] Xu, S.; Lu, H. P.; Chu, W. J. A Low-Carbon-Based Bilevel Optimization Model for Public Transit Network. // *Mathematical Problems in Engineering*. 2013, (2013), pp. 1-7.
- [3] Lei, Shao-Mei; Jia, Xu-Jie; Yu, Zai-Yang. Short-term Traffic Flow Forecast Based on the Gaussian Kernel. // *Journal of MUC (Natural Sciences Edition)*. 22, (2013), pp. 82-87.
- [4] Mei, Hong-Biao; Wang, Jian; Zhang, Hui-Zhe. Research of Fractal Forecasting Algorithm of Traffic Flow on Urban Expressway. // *Journal of Highway and Transportation Research and Development*. 26, (2009), pp. 105-110.
- [5] Peng, Zheng Hong; Song, Bin. The Applications of Data Mining Method in Urban Traffic Flow Prediction. // *Journal of Computational Information Systems*. 5, (2007), pp. 1957-1962.
- [6] Zhang, Hui-Zhe; Wang, Jian. Application of data mining on short-term traffic flow forecasting model. // *Computer Integrated Manufacturing Systems*. 14(2008), pp. 690-695.
- [7] Yu, Yan-Hua; Song, Mei-Na; Zhang Wen-Ting; Song Jun-De. A dynamic computation approach to deter mining the Threshold in Network Anomaly Detection. // *Journal of Beijing University of Posts and Telecommunications*. 34, (2011), pp. 45-49.
- [8] Min, W.; Wynter, L. Real-time road traffic prediction with spatio-temporal correlations. // *Transportation Research Part C*. 19, (2011), pp. 606-616. DOI: 10.1016/j.trc.2010.10.002
- [9] Wang, Chao; Sun, Wei-Hua; He, Yuan-Lie. The Application of Gray Prediction Model in Highway Traffic Prediction. // *The Application of Gray Prediction Model in Highway Traffic Prediction*. 29, (2012), pp. 32-34.
- [10] Xue, Jieni; Shi, Zhongke. Short-Time Traffic Flow Prediction Based on Chaos Time Series Theory. // *Journal of Transportation Systems Engineering and Information Technology*. 8, (2008), pp. 68-72. DOI: 10.1016/S1570-6672(08)60040-9
- [11] Yang, Fan, Yan, Zhong-Zhen. Study on Hybrid Neural Network Data Mining Model for Traffic Flow Prediction. // *Journal of Chinese Computer Systems*. 33, (2012), pp. 1978-1981.
- [12] Hu, Junhui; Yang, Li; Kong, Lingjiang; Yang, Yongxu. Urban mixed traffic flow considering the influence by origin-destination of public transportation. // *Journal of Transportation Systems Engineering and Information Technology*. 11, (2011), pp. 102-107. DOI: 10.1016/S1570-6672(10)60107-9
- [13] Smith, B.; Scherer, W.; Conklin, J. Exploring imputation techniques for missing data in transportation management systems. // *Transp. Res. Rec.* 1836, (2013), pp. 132-142. DOI: 10.3141/1836-17
- [14] Sun, S.; Zhang, C.; Yu, G.; Lu, N.; Xiao, F. Bayesian network methods for traffic flow forecasting with incomplete data. // in *Proc. Mach. Learn., ECML*. 2004, pp. 419-428.
- [15] Li, L.; Li, Y.; Li, Z. Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. // *Transp. Res. Part C Emerging Technol.* 34, (2013), pp. 108-120. DOI: 10.1016/j.trc.2013.05.008
- [16] Haworth, J.; Cheng, T. Non-parametric regression for space-time forecasting under missing data. // *Comput. Environ. Urban Syst.* 36, 6(2012), pp. 538-550. DOI: 10.1016/j.compenurbsys.2012.08.005
- [17] Tan., H. A tensor-based method for missing traffic data completion. // *Transp. Res. Part C Emerging Technol.* 28, (2013) pp. 15-27. DOI: 10.1016/j.trc.2012.12.007
- [18] Khosravi, A. Prediction intervals to account for uncertainties in travel time prediction. // *IEEE Trans. Intell. Transp. Syst.* 12, 2(2011), pp. 537-547. DOI: 10.1109/TITS.2011.2106209
- [19] Cascetta, E. Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. //

- Transp. Res. 18B, 4/5(1984), pp. 289-299. DOI: 10.1016/0191-2615(84)90012-2
- [20] Codina, E.; Barceló, J. Adjustment of O-D trip matrices from observed volumes: An algorithmic approach based on conjugate directions. // Eur. J. Oper. Res. 155, (2004), pp. 535-557. DOI: 10.1016/j.ejor.2003.08.004
- [21] Nie, Y.; Zhang, H. M.; Recker, W. W. Inferring origin destination trip matrices with a decoupled GLS path flow estimator. // Transp. Res. Part B Methodol. 39, (2005), pp. 497-518. DOI: 10.1016/j.trb.2004.07.002
- [22] Maher, M. J.; Zhang, X.; van Vliet, D. A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows. // Transp. Res. Part B Methodol. 35, (2001), pp. 23-40. DOI: 10.1016/S0191-2615(00)00017-5
- [23] van Lint, J. W. C.; Hoogendoorn, S. P.; van Zuylen, H. J. Accurate freeway travel time prediction with state-space neural networks under missing data. // Transp. Res. Part C Emerg. Technol. 3, 5/6(2005), pp. 347-369. DOI: 10.1016/j.trc.2005.03.001
- [24] Zhan, Yanyan; Xu, Rong Cong; Chen, Xiaoyun. Time Series Piecewise Linear Representation Based on Slope Extract Edge Point. // Computer science. 33, (2006), pp. 139-142.

Authors' addresses***Xiao Qiang, PhD***

School of Traffic and Transportation,
Lanzhou Jiao Tong University
Lanzhou 730070, China
School of Economics and Management,
Lanzhou Jiao Tong University,
Lanzhou Gansu 730070, China
Email: lzjt_xq@126.com

He Ruichun, Prof. PhD

Corresponding author
School of Traffic and Transportation,
Lanzhou Jiao Tong University
Lanzhou 730070, China
Email: herc@mail.lzjtu.cn

Yu Jianning, Prof. PhD

School of Traffic and Transportation,
Lanzhou Jiao Tong University
Lanzhou 730070, China
Email: lzjt_yjn@126.com

Zhang Wei, PhD

School of Traffic and Transportation,
Lanzhou Jiao Tong University
Lanzhou 730070, China
Email: zhangwei@mail.lzjtu.cn

Ma Changxi, PhD

School of Traffic and Transportation,
Lanzhou Jiao Tong University
Lanzhou 730070, China
Email: Mcx9815@sina.com