Research Journal of Applied Sciences, Engineering and Technology 5(4): 1271-1277, 2013
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: June 28, 2012 Accepted: August 17, 2012 Published: February 01, 2013

The Method of Traffic State Identification and Travel Time Prediction on Urban Expressway

[1, 2]Yanguo Huang, [1]Lunhui Xu and [1]Qiang Luo
[1]Department of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong 510640, China
[2]Department of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

Abstract: This study presented a method of real-time traffic condition identification based on the fuzzy c-means clustering and travel time prediction on urban expressway. The traffic flow characteristics of expressway were analyzed and the traffic flow states were divided into four classes. Then the fuzzy c-means clustering technique was used to classify the sampled historical data and the clustering center of different traffic condition was gotten. In the test module, the real-time traffic data were used to identify which state the traffic data belong to. Based on the analysis, a travel time prediction model of urban expressway was given by using fuzzy regression. According to the collecting real-time traffic dada and the result of traffic state identification, a method of predicting travel time was introduced. Finally, an urban expressway in Guanzhou was as an example and the result of traffic state identification was same with the results of actual measurement data and questionnaire survey through drivers. The result of travel time prediction show that the predicted results had better fitting degree and precision and the feasibility of this method was verified. It can provide the basis for urban expressway traffic control and traffic induction.

Keywords: Fuzzy C-means clustering, fuzzy regression, traffic flow state, travel time prediction, urban expressway

INTRODUCTION

The traffic state dynamic estimation and travel time dynamic prediction of urban expressway are important parts of Intelligent Transportation System (ITS). Reliable transportation state information is a necessary condition in traffic control and management system. According to the estimation of traffic condition, a proper effective measure of traffic control and induction can be implemented, which makes traffic flow smooth and avoids traffic jam. At present, the traffic status evaluation methods include classical algorithms (Coifman, 2003), such as California algorithm, standard deviation algorithm, McMaster Machine (SVM) method, etc., Jiang (2004) and Li *et al*. (2009). In these methods, the road traffic congestion state judgment usually depended on the basic traffic data such as flow, speed and occupancy or density. These data are calculated and converted to specific characterization quantity. If one or more index of the characterization quantity beyond the threshold value that is determined beforehand during a few test cycles, the traffic condition may be crowded (Zhuang *et al*., 2006; Luo *et al*., 2010). The existing research results were usually mainly on the testing and discrimination for traffic abnormal events, but could not judgment the real-time traffic state of urban expressway.

On the other hand, it is more practical significance for traveler if the travel time of sections can be forecasted than the road traffic state can be known and the travel time prediction of expressway caused many scholars' attention. There are many travel time prediction methods (Xu *et al*., 2008), such as historical trends method, multiple linear regression method, time series analysis, Kalman filter and neural network and random queuing theory model. For these methods and models, each has its advantages and disadvantages. On one hand, the influence factors of traffic condition are more complicated, so it is difficult to establish precise model and evaluation index. On the other hand, each traffic condition has certain similarity, which makes traffic state division fuzziness. According to the traffic characteristics of urban expressway, a new traffic state identification method based on Fuzzy C-Means clustering (FCM) is presented. The model is built and the algorithm realization of road traffic condition real-time identification is given in this study. In addition to this, travel time prediction model of expressway is constructed by using fuzzy regression to forecast the travel time of sections.

Corresponding Author: Yanguo Huang, Department of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China

EXPRESSWAY TRAFFIC FLOW CHARCTERISTIC ANALYSIS

Traffic flow characteristics: Three parameters are usually used to describe the traffic flow characteristic, which are volume of traffic, speed and traffic density, the relationships among the 3 parameters can be expressed as follows:

$$Q = v.k$$

where,
Q = Traffic volume (pcu/h)
V = Traffic interval velocity (km/h)
k = Traffic density (pcu/km)

This formal is most frequently used in the traffic engineering and management and the flow-density curve is called the traffic fundamental diagram.

Urban expressway is different from the general urban road and highway, it has the characteristics of car special, high small. These determine the specific needs for model and method to analyze it. The traffic data of an speed and capacity and the import and export spacing is expressway in Guangzhou were taken, which included flow, occupancy, speed and density and so on. After filtering and statistics analysis these data, we got flow-density scatter plot chart as shown in Fig. 1.

The flow-density scatter plot chart from unblocked to crowded under different traffic flow on the expressway is shown in Fig. 1a, b and c is the flow-density scatter plot chart from crowded to unblock of traffic. According to the scatter plot distribution, the fitting method was used to analyze relationship of density and flow and we found that it was approximate the curve "λ" form as shown in Fig. 2. There is a semi-stable state area from non-crowded area to crowded area. B.S. Kerner thought that the semi-stable state area led to the hysteresis phenomenon in his phase change theory, which makes the traffic flow phase transition from smooth traffic to crowded area happen at high density k_1 and the traffic flow phase transition from crowded area to smooth traffic happen at lower density k_2.

Traffic state division and description: Road traffic state refers to the real-time traffic flow condition of one

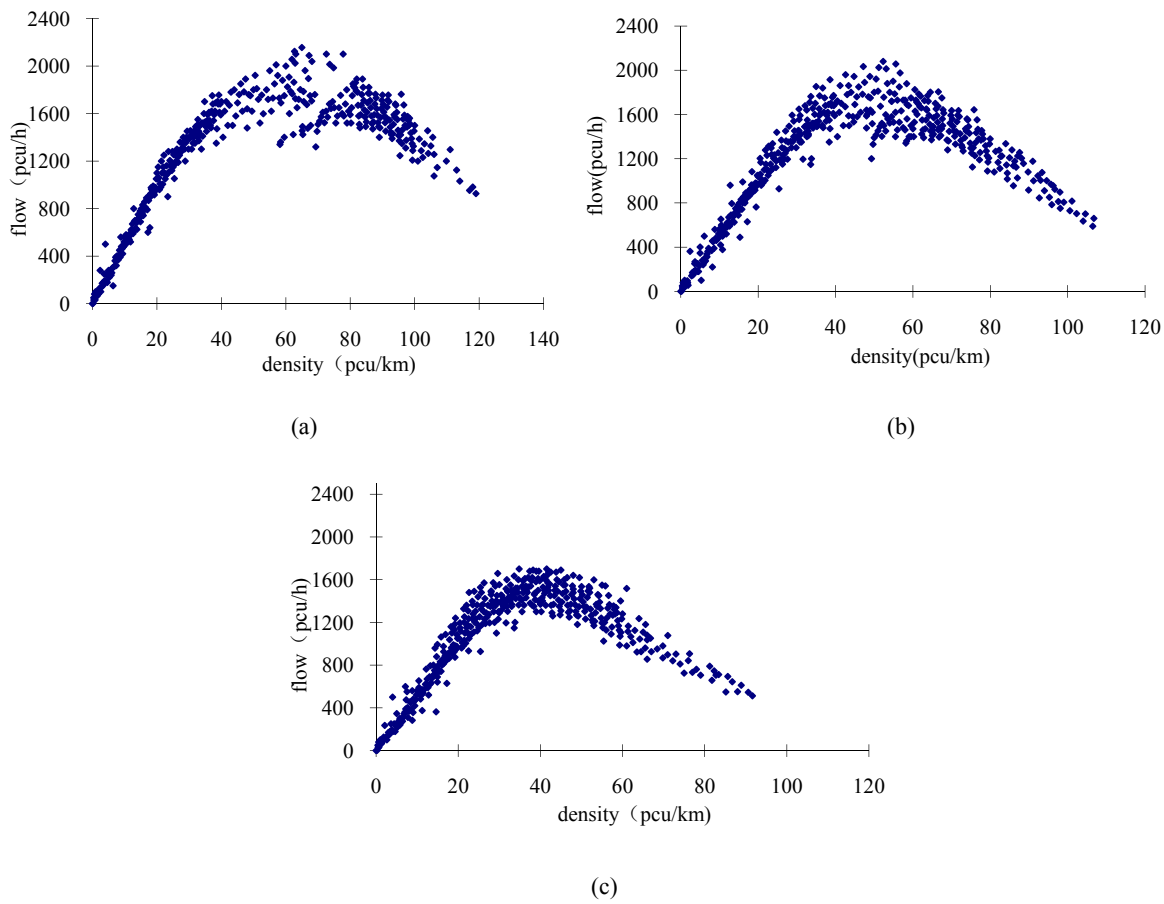

Fig. 1: Flow-density scatter plot chart of different traffic flow (a) flow-density scatter plot chart from unblocked to crowded under small traffic flow, (b) flow-density scatter plot chart from unblocked to crowded under large traffic flow, (c) flow-density scatter plot chart from crowded to unblock of traffic

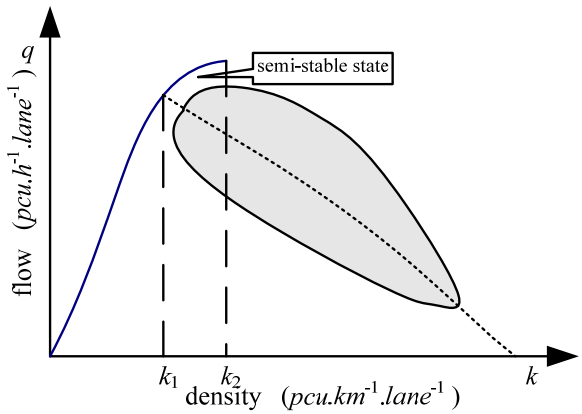

Fig. 2: "λ" shape of relationship of density and flow

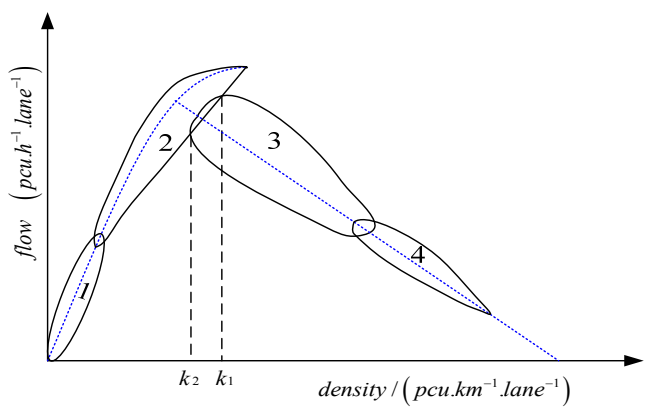

Fig. 3: State classification of traffic flow

road, but the road congestion definition is a vague concept and it is difficult to use specific data to define. Now present, there is still no unified regulation. According to the above analysis for the traffic flow characteristics, the traffic condition of one road can be divided into different feature states. Through analysis and comparison for large amounts of traffic data of different sections of expressway, the traffic condition is divided into 4 features states as shown in Fig. 3.

The first state is freely driving state. Under this condition, flow and density are very low and speed is high. The vehicle almost cannot be suffered from the influence of the vehicle ahead or behind of it and driver has good free driving degrees.

The 2th state is steady flow state. Speed and flow are very high, but density is medium. Along with the increase of density, flow is on the increase and even can be increased to the traffic capacity. Under this traffic condition, the road can get to be fully used and driver can drive in larger freedom.

The 3th state is crowded flow state. Along with the increase of flow, speed falls sharply. In this traffic condition, the traffic flow volatility is bigger and average speed is about 35 km/h.

The 4th state is serious crowded state. The density is very high and traffic jam often happen. The average speed is under 20 km/h and the whole the road traffic condition is under the state of vehicle-following synchronization.

TRAFFIC STATE IDENTIFICATION MODEL BASED ON FCM CLUSTERING

FCM clustering principle: Fuzzy C-Means (FCM) clustering is a typical fuzzy clustering method, which was proposed by Bezdek (1981). In this method, every one data point belongs to a kind of class with some degree and a membership is used to describe the relation degree that one data point belongs to one class.

A sample data set is $X = \{X_1, X_2,... X_n\}$, The sample data are divided into C classes and $1<C<n$. For any $X_i \in X$, X_i is characteristic vector, $X_i = \{x_{i1}, x_{i2},... x_{ip}\}$, x_{ik} is the first k characteristic attributes of the first i sample data. The target function with membership functions defined is as follows Luo *et al*. (2010):

$$\min\{J(U,V)\} = \min\left\{\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m (d_{ik})^2\right\} \quad (1)$$

$$\text{s.t.} \begin{cases} 0 \le u_{ik} \le 1 \\ \sum_{i=1}^{c} u_{ik} = 1 \end{cases} \quad (2)$$

$$(d_{ik})^2 = \| x_k - v_i \| = (x_k - v_i)^{\mathrm{T}} A (x_k - v_i) \quad (3)$$

$$V = [v_1, v_2, \ldots, v_c] \quad (4)$$

where,

u_{ij} (X_i)	=	The membership function that the first i sample belong to the first j class
m	=	Fuzzy weighted exponent
d_{ij}	=	The Euclidean distance of the sample data to clustering center
U	=	Fuzzy classification matrix
V	=	Clustering center
V_c	=	The center for each clustering class
J (U, V)	=	The weighted value of the distance of all samples to clustering center

Algorithm realization of FCM clustering: In this study, the three parameters (flow, speed and occupancy) are selected as input variables and that is $x_i = (x_{i1}, x_{i2}, x_{i3})$ and x_i is the first i sample that include flow, speed and occupancy, three characteristic attributes. Because the values of three parameters are different, the sampled data should be standardized and the extreme standardization formula is used to normalize the sample parameters and compressed to interval [0, 1]. By using the Lagrange multiplier method, the parameters calculation is as follows:

$$F = \sum_{i=1}^{c}(u_{ij})^m (d_{ij})^2 + \lambda\left(\sum_{i=1}^{c} u_{ij} - 1\right) \quad (5)$$

The optimized condition is that:

$$\begin{cases} \dfrac{\partial F}{\partial \lambda} = \left(\sum_{i=1}^{c} u_{ij} - 1\right) = 0 \\ \dfrac{\partial F}{\partial u_{ij}} = \left[m(u_{ij})^{m-1}(d_{ij})^2 - \lambda\right] \end{cases} \quad (6)$$

The solution of the equation is as follows:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\dfrac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \quad (7)$$

$$v_i = \frac{\sum_{k=1}^{n}(u_{ik})^m x_k}{\sum_{k=1}^{n}(u_{ik})^m} \quad (8)$$

Travel time prediction based on fuzzy regression:

Road travel time prediction method analysis: Travel time prediction is one of the core research content of ITS, There are many travel time prediction methods, such as historical trends method, multiple linear regression method, time series analysis, Kalman filter and neural network and random queuing theory model and so on. Fuzzy regression is a new kind of regression prediction method and it has the characteristics of recognition of complex nonlinear system. In this study, travel time prediction model of expressway is built by fuzzy regression and how to calibrate the model parameter is introduced. According to the real-time traffic flow data and the result of the traffic state identified through FCM and the traffic condition that the road can provide, the dynamic travel time between any 2 points of the road can be forecasted.

Fuzzy linear regression model: Fuzzy Linear Regression (FLR) is similar with classic linear regression analysis, assuming that variables y and its related factors $x_1, x_2, \ldots, x_N$ is linear related and the fuzzy linear regression model is described as fellow:

$$\hat{y} = A_0 x_0 + A_1 x_1 + \cdots + A_N x_N = AX \quad (9)$$

while,

$\hat{y}_i$ = fuzzy output

$X = (x_0, x_1, \ldots, x_N)$ is an independent variable vector

$A = (A_0, A_1, \ldots, A_N)$ is fuzzy coefficient vector

Regression analysis problem is to use the known M group observation data, i.e., $y_i, x_{i1}, x_{i2}, \ldots, x_{ij}$ (i = 1, 2, ⋯, M) to estimate the regression coefficients A_j. Regression coefficient A_j is fuzzy number, which causes deviation between the fitted values $\hat{y}_i$ and observation values y_i. Usually A_j is taken triangular fuzzy function A_j (α_j, c_j) and its membership function is as follows:

$$\mu_A(x) = \begin{cases} 1 - \dfrac{|x-a|}{c} & , \ |x-a| \le c \\ 0 & , \text{ others} \end{cases} \quad (10)$$

The membership functions of output variable y is as following shown:

$$\mu_{\tilde{Y}}(y_i)=\begin{cases}1-\dfrac{\left|y_i-\sum_j a_j x_{ij}\right|}{\sum_j |x_{ij}|c_j}, & \left|y_i-\sum_{j=0}^{N} a_j x_{ij}\right|\le\sum_{j=0}^{N}|x_{ij}|c_j \\ 0 & \text{, others}\end{cases} \quad (11)$$

The model is to determine the parameters $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $c_1, c_1, \ldots, c_n$.

Model parameters: In order to get the best fitting function (as shown in Eq. (9)) from the existing M group observation data, the 2 requirements must be met in the FLR analysis:

- The sum of fuzzy range of regression coefficients must be minimum, which make precision is the biggest.
- According to certain confidence levels H, the results must be able to reflect all the observation data y_i.

It is that:

$$\mu_y(y_i)\ge H \quad (12)$$

Then the above parameters can be transformed into the problem of solving linear programming problem as shown in fellows [15]:

$$\min \sum_{j=0}^{N}\left(c_j\sum_{i=1}^{M}|x_{ij}|\right) \quad (13)$$

$$\text{s.t.}\begin{cases}\sum_{j=0}^{N} a_j x_{ij}+(1-H)\sum_{j=0}^{N}|x_{ij}|c_j\ge y_i \\ \sum_{j=0}^{N} a_j x_{ij}-(1-H)\sum_{j=0}^{N}|x_{ij}|c_j\le y_i\end{cases} \quad (14)$$

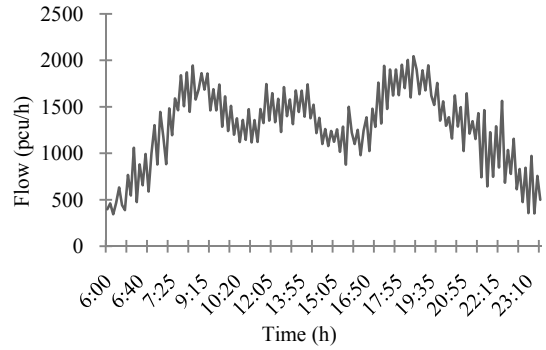

Fig. 4: Traffic flow

while,
H = given in advance
x_{i0} = 1
i = 1, 2, …, M

CASE ANALYSIS

The simulation of traffic environment: This study selects one expressway in Guangzhou as an example, which is long about 3.2 km, 3 lanes. The continuous seven days traffic data were collected in May 2010, 10 to 16 (from Monday to Sunday). The traffic data were sampled every 5 min from 6:30 AM to 20:30. The data that were filtered and processed were shown in Fig. 4.

Traffic state identification based on FCM: The implementation of the method was completed with software METLAB 7.0 and the realization steps were fellows:

Step 1: Normalize the collected sample data
Step 2: Initial membership matrix $U^{(0)}=(u^{(0)}_{ij})$
Step 3: Use the formula 5, 6 and 7 to calculate the new membership matrix U
Step 4: Use the formula 8 to calculate the clustering center feature vector V

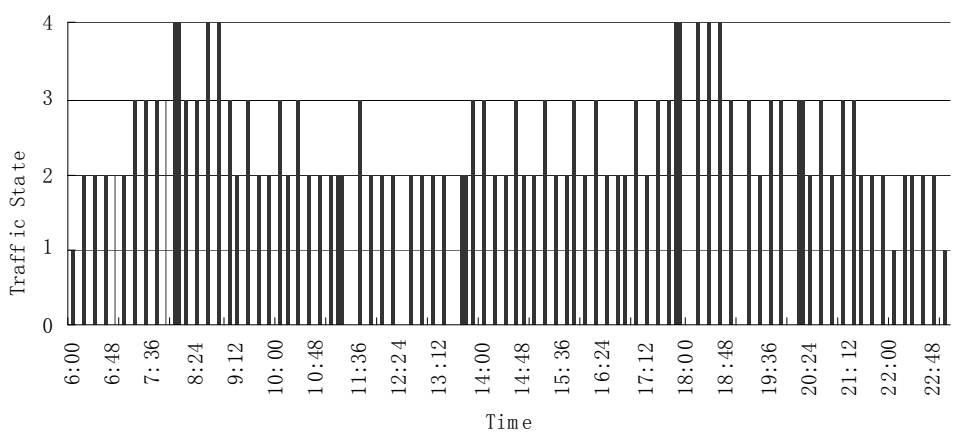

Fig. 5: The identification result of traffic condition

Table 1: Results of traffic state identification

Sample time	Flow q (pcu/5 min)	Speed v (km/h)	Occupancy (%)	Traffic state
9:00	312	38.5	38.01	3
9:05	296	41.2	30.24	3
9:10	265	44.8	25.14	3
9:20	245	46.3	20.36	2
9:30	219	45.5	19.25	2
9:40	158	52.4	16.35	2
9:50	121	58.8	14.48	2
10:00	109	60.1	11.38	1
10:10	103	63.5	9.210	1
10:20	81	65.2	8.210	1
10:30	84	66.7	8.350	2
10:40	72	69.8	7.430	1

Step 5: Repeat steps 3 and 4 until the objective function for the optimal solution meet the following conditions, that is $\| v_i^{(k+1)} - v_i^{(k)} \| < \varepsilon$, then stop and find clustering center *Vc*

Step 6: Predict the travel time of urban expressway by fuzzy regression method. The program included 2 modules: training module and test module. Firstly, selected 100 group sample dada and fuzzy C-means clustering method was used to classify the sample data and got the clustering center of different states of this section:

$$V_c = \{v_1, v_2, v_3, v_4\} = \begin{bmatrix} 530 & 1370 & 1550 & 1680 \\ 68 & 57 & 33 & 18 \\ 8 & 16 & 38 & 60 \end{bmatrix}$$

Then, in the test module, the traffic test data of working days and holiday were used to identify which state the traffic data belong to. The data in May 2010, 10 (Monday) as a test data, the result of state identification was shown in Fig. 5. Through the actual traffic data in the test, the questionnaire survey of feeling of the traffic state for drivers was done and the results were same with the judgment of FCM algorithm

TRAVEL TIME FORECASTING RESULTS AND ANALYSIS

In order to verify the validity of the model, the data in May 2010, 13 (Wednesday) were selected, which were sampled every 5 min from 9:00 to 11:30, as shown in Table 1.

While x_1 is flow, x_2 is occupancy and y is the observation travel time. There are 30 group data and the first 20 group data of them are used to analyze the FLR model and the rest 10 sets of the data are used to detect verification. The general fuzzy regression model is as follows:

$$y = A_0 + A_1 x_1 + A_2 x_2 \tag{15}$$

While regression coefficient $A_j(\alpha_j, c_j)$ was taken triangle function, $c_j \geq 0$, (j = 0, 1, 2). The confidence levels H was that H>0.5. We could get parameters A_j through solving the Eq. (13) and (14) and the FLR model was built. Then the travel time of this section could be forecasted with the model and the real-time traffic data. Then use the following method to evaluate the prediction values and the observation values:

$$\mu_y(y_i) \geq H \tag{16}$$

$$\Omega_1 = \frac{\left| y_i - \sum_j a_j x_{ij} \right|}{y_i} \tag{17}$$

$$\Omega_2 = \frac{\sum_j |x_{ij}| c_j}{y_i} \tag{18}$$

while,

Ω_1 : The relative deviation of fitting center value to observation yi

Ω_2 : The ratio of fuzzy range to observation yi

uij : The membership degree of the observations yi to model

Generally, if every u_{ij} is larger than 0.5 and the 2 ratio values Ω_1 and Ω_2 are within 30%, the model would be thought a better fitting model. The predication values and evaluation index were shown in Table 2. From the results, we could find that there were few data whose uij value was less than 0.5 and the value of Ω_1 and Ω_2 were all less than 0.3, so the fitting model was better, it can be used to forecast the travel time of urban expressway.

Table 2: Results of travel time fuzzy prediction

			Evaluation index		
Sample time	Travel time (min)	Traffic state	u_{ij}	Ω_1	Ω_2
10:45	165	1	0.158	0.152	0.141
10:50	176	1	0.075	0.126	0.089
10:55	185	1	0.602	0.264	0.168
11:00	190	2	0.341	0.147	0.205
11:05	196	2	0.216	0.155	0.113
11:10	215	2	0.133	0.203	0.152
11:15	253	2	0.075	0.146	0.109
11:20	298	2	0.194	0.181	0.174
11:25	381	3	0.217	0.162	0.265
11:30	398	3	0.186	0.117	0.187

CONCLUSION

The traffic state identification and travel time prediction are an important content in road traffic control system and guidance system. If the road traffic condition can be identified and the travel time can be forecasted beforehand, the drivers can select rational entrance on the expressway. In this study, the traffic characteristics of expressway were analyzed and traffic condition was divided into 4 traffic states. With flow, speed, occupancy as a characteristic attributes, the built model by fuzzy c-means clustering could be used to identify the traffic state of urban expressway and the traffic clustering center of different state be gotten through the algorithm. Then the constructed travel time prediction model with FLR was used to forecast the travel time of sections. Through the case analysis, the results verified the feasibility of the method.

ACKNOWLEDGMENT

This study is supported by National Natural Science Foundation of China (No. 60664001), Natural Science Foundation of Jiangxi Province (No. 2010GQS0076), Science and Technology Plan of Jiangxi Province Education Department (No. GJJ10481).

REFERENCES

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.

Coifman, B., 2003. Identifying the onset of congestion rapidly with existing traffic detectors. Transport. Res., 37: 277-291.

Jiang, G.Y., 2004. Technologies and Applications of the Identification of Road Traffic Conditions. China Communications Press, Beijing.

Li, Q.Q., D. Gao and B. Yang, 2009. Urban road traffic status classification based on fuzzy support vector machines. J. Jilin Univ., Eng. Technol. Edn., 39: 131-134.

Luo, X.Q., K.M. Chen and T.F. Zhang, 2010. Congestion incident and degree identification of urban expressway. J. Chang Univ. Nat. Sci. Edn., 30: 71-75.

Xu, T.D., L.J. Sun and Y. Hao, 2008. Real-time traffic state estimation and travel time prediction on urban expressway. J. Tongji Univ., Nat. Sci., 36(10): 1355-1361.

Zhuang, B., X. Yang and K. Li, 2006. Criterion and detection algorithm for road traffic congestion incidents. China J. Highway Transport, 19: 82-86.