Identification for road traffic state

A new method based on driver evaluation

Lao Yunteng[a], Yang Xiaoguang[b]

[a]Room 15-503, No. 4800 Cao'an Road, Tongji University, 201804, Shanghai, P.R. China; Tel or Fax +86-21-69588994; email: laoyunteng@163.com

[b]School of Transportation Engineering, Tongji University, 2000092, Shanghai, P.R. China; Tel or Fax : +86-21-65988372; email: yangxg@mail.tongji.edu.cn

Abstract

It is difficult to evaluate or optimize a road traffic state identification method, since traffic state is a fuzzy variable based on sense. And the accuracy of previous traffic state identification methods is still not well analyzed. This paper puts forward a driver-based evaluation approach which can solve the fuzziness in the determination of traffic state. Based on the evaluation, an optimization model is presented, and the parameters in identification algorithms can be amended. Thus, a new identification conception for traffic state is presented. This research suggests that driver-based evaluation should be needed a necessity for the improvement of the traffic management system.

Keywords: Traffic state identification; Driver-based evaluation; Optimization method

1. Introduction

The identification of road traffic state (free, crowd and jam) is carried out on manual program at first. And this method is still effective in the situation where it is lack of any detection system. However, this method has become discommodious with the development of Intelligent Transportation Systems, and it has its limitation since people could not observe it all the time.

With the growing concern about traffic congestion, the characters of traffic congestion are studied (Daganzo, 1995; B.S. Kerner and H. Rehborn, 1996; Cassidy, 1999) and many identification algorithms of traffic state have been put forward for the traffic guidance system. Either the operational characteristics, e.g., speed, delay, travel time, density, etc., or the volume characteristics, e.g., operating traffic volume, volume to capacity ratio, traffic volume per lane, etc., has been used in threshold models to describe the road traffic state (Mateen 1990; Ghiria 1991; Bhargab, 1999; Robert, 2006); Also, some new methods are raised with the development of Applied Math (Narayanan, 2003; Jiang, 2004). However, very few research has been focused on that whether the traffic state identified by these algorithms are accorded with the true situation, or whether the thresholds in the model are suitable, or how the time window of data collection and the thresholds in traffic state identification algorithms are optimized.

Recently, a driver-based evaluation method is presented by Lao et al. (2006), and the detail application of this method is still not carried out in the early research. Moreover, the road traffic state determined by the evaluation could not meet the requirement of traffic guidance system, since the system needs to know the information timely and we are not able to organize the drivers to evaluate all the time. Newly, we have applied the driver-based evaluation method to the experimental unit project in Out Loop Line of Shanghai and developed it, such as optimizing the identification algorithm and choosing a better algorithm. A new conception to identify a certain traffic state on the urban road was developed based on the evaluation method.

2. Identification Conceptualization

One of the aims of the traffic guidance system is to improve the accuracy of the traffic information which illuminates the coincidence between the detected traffic state and the actual state. However, the identification of the actual traffic state is various among different areas and different drivers, since the traffic state is subjective information strongly depended on sense. Moreover, the traffic state modeled by the detected data, such as speed, volume, occupation, etc., could not be used as actual state, since different models may lead to different results. It is hard to determine which result is better accorded with the actual traffic state and the drivers' sense.

It is more reasonable to define the identification state of all drivers in the system as the actual state, since the drivers are the users of the information service. Thus, a driver-based evaluation method is put forward in our research (Lao, 2006). And the traffic state identified by the drivers is defined as optimization benchmark for the traffic state identification algorithm. With the optimization benchmark, the optimization function can be provided, and the effect of different parameters in algorithm, such as different time window of data collection or different thresholds, can be analyzed. According to this principle, a new conception of traffic state identification method can be carried out in the following procedure:

(1). Quantifying the benchmark traffic state by applying a driver-based evaluation method.
(2). Optimizing the relative parameters in identification algorithms based on the benchmark state.
(3). Choosing a better identification algorithm for road traffic state information based on the optimization result.
(4). Identifying the road traffic state by applying the new optimized model.

3. Identification procedure

3.1 Driver-based Evaluation method

3.1.1 Determining Evaluation Sample Size

The more drivers the evaluation invites, the better the result would be. However, a large sample will cost much. Thus, it is necessary to determine a rational evaluation sample size. We assume that the evaluation of drivers as a normal distribution; therefore, the sample n can be obtained from the theory of statistics as follow:

$$n = \frac{Z_{a/2}{}^{2}\sigma^{2}}{\Delta^{2}} \quad (1)$$

Where $Z_{a/2}$ is known as the critical value, the positive Z value that is at the vertical boundary for the area of a/2 in the right tail of the standard normal distribution; σ is the population standard deviation; the margin of error Δ is the maximum difference between the observed sample mean and the true value of the population mean.

Mark the free, crowd, jam as 1, 2, 3, respectively. The standard deviation is maximal when the traffic state at the critical condition between free and crowd or between crowd and jam, since different drivers may have a large diverge in this situation. In this case, σ =0.5. We assume that the margin of error Δ =0.3. Thus:

$$n = \frac{25}{9} Z_{a/2}{}^{2} \quad (2)$$

The critical value $Z_{a/2}$, corresponding confidence degree H and sample n, which can be obtained from formula (2), are shown in the table 1. The sample can be determined by the requirement of precision. It can be found that a small sample of drivers can make a high precision, since the evaluation has only three ranks.

Table 1 Critical value $Z_{a/2}$ and confidence degree H for different sample n

n	7	8	9	10	11
$Z_{a/2}$	1.59	1.70	1.80	1.90	1.99
H	0.89	0.91	0.93	0.94	0.96

3.1.2 Quantifying Benchmark Traffic State

At first, we calculate the average evaluation result, and then determine the corresponding perception traffic state by using the adjacency principle. The average result A_i of traffic state is:

$$A_i = \frac{1}{n}\sum_{j=1}^{n} A_{ij} \qquad (3)$$

Where i is the evaluation sequence; A_{ij} is the evaluation result of driver j at sequence i; n is the evaluation sample size. Using the adjacency principle in pattern recognition, we can get perception traffic state D_i at sequence i:

$$D_i = \text{int}(A_i + 0.5) \qquad (4)$$

Where int(A) is a function that gets a biggest integer not more than A. The perception traffic state evaluated by the drivers can be used as a benchmark state.

3.2 Optimization Method

3.2.1 Accuracy function

According to the evaluation principle, the accuracy of detection algorithms can be obtained by comparing the detection state and the benchmark state. Thus, the false identified rate f_j for benchmark state j could be calculated as follow:

$$f_j = \frac{\sum_{k=1}^{N_j} \left| j - d_{ji} \right|}{2N_j} \times 100\% \qquad j=1, 2, 3 \qquad (5)$$

Where d_{ji} is the detection traffic state at sequence i for benchmark state j, d_{jk} = 1, 2, 3; N_j is the number of all sequence for benchmark state j. And the accuracy F of detection algorithm is:

$$F = 1 - \sum_{i=1}^{3} a_j f_j \qquad j=1, 2, 3 \qquad (6)$$

Where a_j is the weight of benchmark state j; in general, it can be a_1=a_2=a_3=1/3. In function (6), the higher the F is, the more accuracy the detection algorithm would be. The value of F is determined by the type of algorithm and the parameters in this algorithm.

3.2.2 Parameter Optimization Model

For certain detection method, if the complexity of an algorithm do not affect the detection time so much, the aim of optimization should be to improve the accuracy of detection. Thus, the optimization function could be:

$$\max F(x) = 1 - \sum_{i=1}^{3} a_j f(x)_j \qquad \text{j=1, 2, 3} \qquad (7)$$

Where the optimization parameter x can be a vector including number of unit time, speed, occupation, etc. The unit time of collection data can constrain the change of time window T. This constraint condition can be stated as:

$$T = nt \qquad \text{n=1, 2, 3 …} \qquad (8)$$

Where t is unit time of collection data; n is the number of unit time t in a time window T. We mark that T_1, T_2 are the identification thresholds for free, crowd and jam (T_1、T_2 are single value for single threshold models, while vector for multi-threshold models, this paper only focuses on the single threshold models). Then the optimization parameter vector x= (n, T_1, T_2). In some threshold models, such as the speed threshold model, the traffic condition would become better, if the value of the parameter increases. In this case, the thresholds should be satisfied:

$$b_1 < T_2 < T_1 < b_2 \qquad (9)$$

Where b_1, b_2 are the lower limit and the upper limit of threshold, respectively. Thus, the value of d_{ji} in formula (5) is:

$$d_{jk} = \begin{cases} 1 & h_i > T_1 \\ 2 & T_2 < h_i \le T_1 \\ 3 & h_i \le T_2 \end{cases} \qquad (10)$$

Where h_i is the corresponding detected data. The optimization parameters can be calculated by (5)-(10). In some threshold models, such as the occupation, volume and delay threshold model, if the value of the parameter increases, the traffic condition would become worse. In this case, the thresholds should be satisfied:

$$b_1 < T_1 < T_2 < b_2 \qquad (11)$$

Thus, the value of d_{ji} in formula (5) is:

$$d_{jk} = \begin{cases} 1 & h_i \le T_1 \\ 2 & T_1 < h_i \le T_2 \\ 3 & h_i > T_2 \end{cases} \qquad (12)$$

The optimization parameters can be calculated by (5)-(8) and (11)-(12).

3.2.3 Solution method for the model

In most threshold algorithm, the range of the parameters is not so large, and the needed optimization parameters are not so many, the optimization result can be obtained by searching directly. Some heuristic algorithms (Jorge, 1999) can be used to solve the optimization model when the parameters are in a large group. Iterative Optimization is used in this research.

3.3 Choosing better detection algorithm

A reasonable detection algorithm should be considered in three aspects (Lao, 2006). First of all, the accuracy of a certain detection algorithm is the most important factors for the guidance system, since inaccurate traffic information may heavily reduce the whole efficiency of road traffic system. In addition, the timeliness of a detection algorithm should not limit the requirement of traffic management and the economic factor should also be considered for choosing a better algorithm. The combination evaluation function E of these three factors could be defined as follow, for more detail is discussed in (Lao, 2006).

$$E = F \times (\alpha \times F_T) \times (\beta \times F_C) \qquad (13)$$

Where F, F_T, F_C are accuracy, time efficiency and economical efficiency, respectively; α、β are the weights of time efficiency and economical efficiency, respectively. Generally, it can be $\alpha = \beta = 1$, and their value can be changed depend on the importance of these three factors. We can choice a better traffic state identification algorithm base on the value of E.

For a certain road traffic system, the benchmark traffic state in certain time can be calculated by formal (3)-(4); and comparing with the identification traffic state, the accuracy of a certain identification method can be obtained by formal (5)-(6). The parameters in the identification methods can be optimized based on formal (5)-(12). And then using the optimization result, we can choose a better traffic state identification method for a certain road traffic system via formal (13).

4 Case Study

The driver-based evaluation method in this research has been applied to the experimental unit project in Out Loop Line, which is an expressway in Shanghai with the limited speed of 80km/h. There are two kinds of lanes along the road, one is cars' lane, while the other is trucks' lane. The unit time of data collection is 20s (t=20s). The data in this research is from the loop detectors and evaluation results of drivers.

There are ten drivers taking part in the evaluation, all of them are familiar with the traffic condition of Out Loop Line of Shanghai, which can contribute to the precision of our evaluation. Moreover, our research team made a spot questionnaire survey in experimental unit project in Out Loop Line in July, 2006 for two weeks. Then, we make up a series of indices of traffic state evaluation by analyzing 788 questionnaires of survey (see Table 2).

These indices act as guiding conclusions provided to the drivers, on purpose of enabling the drivers to be more familiar with the traffic condition of the road. Combining with their experience, the drivers taking part in the evaluation will evaluate the traffic state they observe.

Table 2 Indices of traffic state for Out Loop Line

Traffic state	Free	Crowd	Jam
Distance between vehicles (m)	>20	10～20	<10
Speed (km/h)	>45	20～45	<20
Stabilization of driving	Good	Bad	Depend on the front vehicle
Disturbance among vehicles	Seldom	Sometimes	Severe

5 Result and discussion

Three commonly used thresholds in different time window are optimized in our research. And two types of lanes (cars' lane and trucks' lane) are analyzed. The optimization results are showed in Fig.2-12 and the main results are listed in Table 3, the Figure on the left is for the cars' lane with three lanes' data, while the right one is for the trucks', where data in lane1 comes from the combination data of lane2 and lane3, data in lane4 comes from the combination data in lane5 and lane6.

5.1 Speed threshold optimization

Fig.1-3 is the optimized result for speed threshold model based on (5)-(10). Fig.1 is the optimized accuracy F in different time window of data collection, where n is the number of unit time of data collection in a time window; Fig.2 is the optimized speed threshold in different time window, where T1, T2 are speed thresholds for free and crowd, crowd and jam, respectively; Using the optimization result in Fig.2, we put the optimized T1 and T2 into (5)-(10) and optimize the time window again, see Fig.3; then the T1 and T2 can be optimized again based on Fig.3.

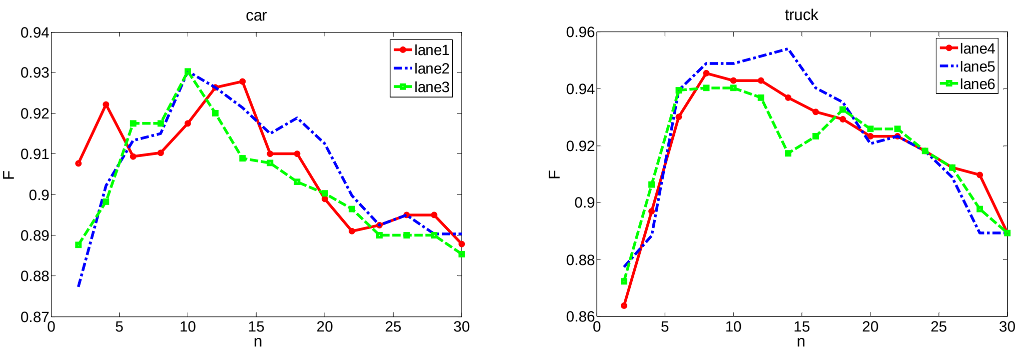

Fig.1 Optimization accuracy in different time window for speed threshold model.

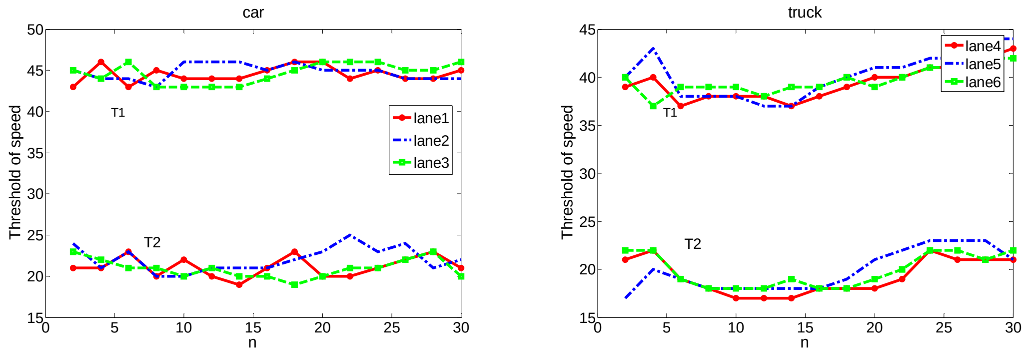

Fig.2 Optimization thresholds in different time window for speed threshold model.

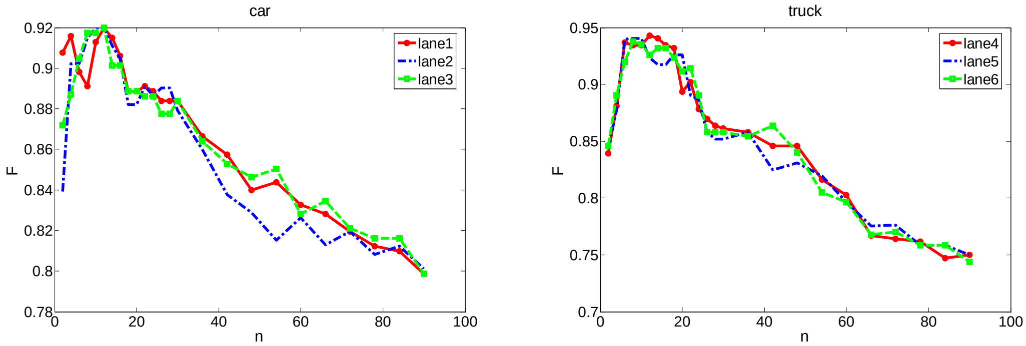

Fig.3 Accuracy of traffic state in different time window with fixed thresholds for speed threshold model. For cars' T1=44km/h, T2=21 km/h. For trucks' T1=38 km/h, T2=18 km/h.

5.2 Occupation threshold optimization

Fig.4-6 is the optimized result for occupation threshold model based on (5)-(8), (11)-(12). Fig.4 is the optimized accuracy F in different time window of data collection; Fig.5 is the optimized occupation threshold in different time window, where T1,T2 are occupation thresholds for free and crowd, crowd and jam, respectively; Using the optimization result in Fig.5, we put the optimized T1 and T2 into (5)-(8), (11)-(12) and optimize the time window again, see Fig.6; then the T1 and T2 can be optimized again based on Fig.6.

5.3 Volume threshold optimization

Fig.7 is the optimized result for volume threshold model based on (5)-(8), (11)-(12). It can found that the accuracy is very low when using the volume alone as detection parameter and the optimized result is not in regular form. Therefore, it is not commended to use volume alone as threshold to detect the traffic state, since low volume can occur in both free flow and congested states.

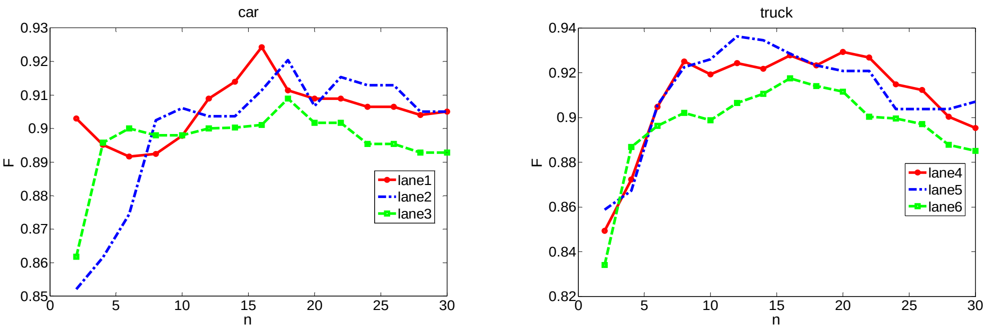

Fig.4 Optimization accuracy in different time window for occupation threshold model.

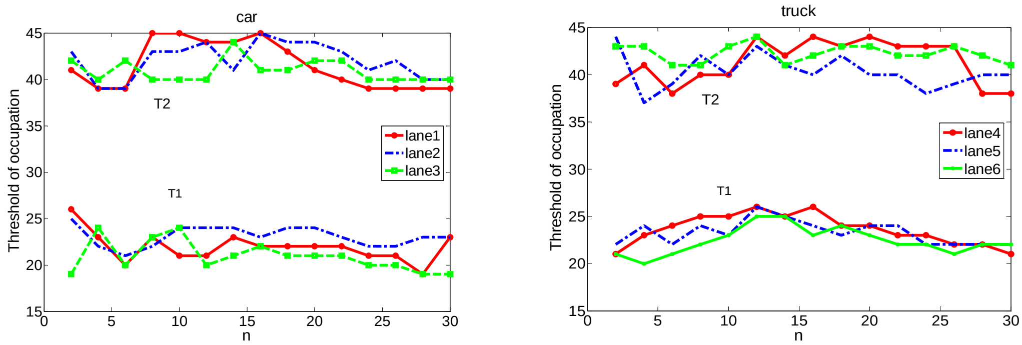

Fig.5 Optimization threshold in different time window for occupation threshold model.

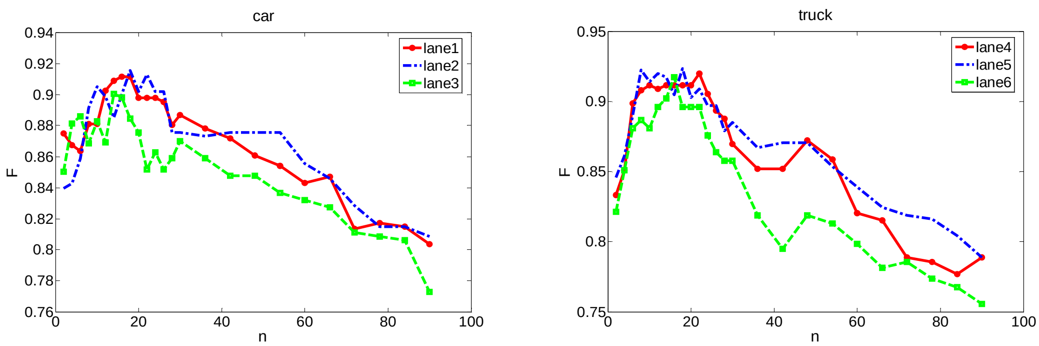

Fig.6 Accuracy of traffic state in different time window with fixed thresholds for occupation threshold model. For cars' T1=22%, T2=44%. For trucks' T1=24%, T2=42%.

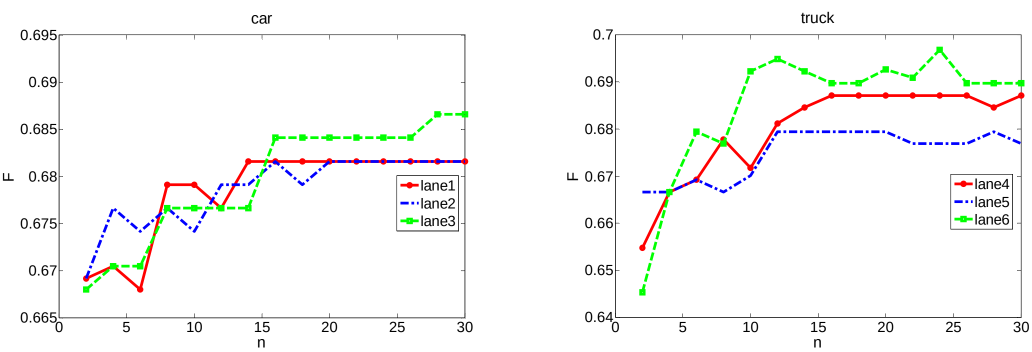

Fig.7 Optimization accuracy in different time window for volume threshold model.

5.4 Discussion

The main results are listed in Table 3. Some further results can be got from the optimization as follow, and these can be useful for the improvement of traffic management system.

(1) The optimized result of speed threshold model and occupation threshold model is much better than that of volume threshold model.

(2) The optimized result of time window is related to the algorithm, it is not necessary to collect data in every 20s. The best data collection cycle is 4 minutes for speed threshold model, and 6 minutes for occupation threshold model. Neither the optimized result nor the collection cycle is good in volume threshold model.

(3) It could be optimized to a good result when data collection time window change from 3 minutes to 6 minutes in speed or occupation threshold model. It may be a little different in different model, and 4 minutes can be recommended in these models. Of course, other traffic factors, such as traffic incident detection and travel time prediction, as well as the users' requirement for traffic guidance system may affect the width of time window. It needs further study on other factors for the determining of time window.

(4) There is a best value for threshold in identification algorithm for certain traffic condition.

(5) The optimizing thresholds are stable for a certain threshold model when the time window changes in a certain range.

(6) The optimization result will be a little fluctuant, since some stochastic factors may affect the detector data. However, this will not affect the optimizated result in total.

Table 3 Optimization result for three threshold models

	speed		occupation		volume	
	car	truck	car	truck	car	truck
F	0.93	0.94	0.92	0.93	0.68	0.69
n	12	10	18	16	——	——
T(s)=n×t	240	200	360	320	——	——
T1	44km/h	38 km/h	22%	24%	720veh/h	540 veh/h
T2	21 km/h	18 km/h	44%	42%	1980 veh/h	1260 veh/h

The time efficiency F_T in these three threshold models is equal if they use the same detection system, the only difference among them is the processing time of certain algorithm, and this difference is so small comparing with the whole detection time in this situation. Also the economical efficiency F_C is equal, since distance between two detectors is the same in these threshold models. However, if the distance is different, the value of F_C should be calculated respectively. It needs further research that whether there is an optimization distance between two detectors.

The only contributing factor among these three models is the accuracy. In this case, we can evaluation these detection models only depend on the value of F. Thus, we can choose the speed threshold model as our identification method, and the correlative parameters can be seen in Table 3.

Some identification methods have not been optimized in this research, and the time efficiency and economical efficiency factors may become important in these methods, which still need further study.

6 Conclusion

This paper presents a new conception for road traffic state identification. To identify more accurately, we put forward a driver-based evaluation method, which distinctly solves the fuzziness in determination of traffic state and provides a suitable benchmark for optimization. Based on the evaluation, the optimization function is provided, and the correlative parameters in identification algorithm can be optimized. In this research, the time window of data collection and three commonly used thresholds are optimized. And then using the optimization result, we choose the speed threshold model as our identification method. For cars' lane, the optimization time window T=240s, speed threshold T1=44km/h and T2=21km/h; while for trucks' lane, T=200s, speed threshold T1=38km/h and T2=18km/h.

Our evaluation method has been applied to the experimental unit project in Out Loop Line of Shanghai. And the development of this method has been carried out. However, further researches are still needed. First of all, we assume that the drivers' evaluation as a normal distribution while it may not in this case, since the drivers' perceptions of traffic congestion are related to their purpose of traveling. Whether more drivers are necessary for other distributions needs further research. Moreover, the time efficiency and economical efficiency factors may become important in other identification methods, and some other parameters such as distance between two detectors are not analyzed in this research. Finally, other traffic factors, such as traffic incident detection and travel time prediction, as well as the users' requirement for traffic guidance system may affect the width of time window.

The road traffic state identification method presented in this paper provides a method for the improvement of the traffic management system. We suggest that the traffic guidance system needs to use the driver-based evaluation method to optimize the identification methods they used, since the identification for the traffic information in traffic system should not only be determined by the detected data, but also be affected by their users. By doing this, the traffic state identification can be more reasonable.

Acknowledgements

The research was supported by Key Program of National Science Foundation of P.R. China: Urban Transportation Network Optimization and Management (Grant No. 70631002) and ours thanks are to Highway Management Department of Shanghai for the help in the preparation of the data.

References

Bhargab Maitra, P. K. Sikdar, S. L. Dhingra. (1999) 'Modeling congestion on urban roads and assessing level of service', Journal of Transportation Engineering, 508-514

Cassidy, M.J. and Bertini, R.L, 'Observations at a Freeway Bottleneck', International Symposium of Traffic and Transportation Theory, Jerusalem, Israel, 1999.

Daganzo, C.F. (1995) 'The nature of freeway gridlock and how to prevent it', Institute of Transportation Studies, Research Report UCB-ITS-RR-95-1; abridged in Transportation and Traffic Theory, pp 629-646, J.B. Lesort, editor, Pergamon-Elsevier, New York, N.Y.

Ghiria, N. (1991) ''Evaluation and pricing of congestion in a local area network of roads.'' Proj. Rep., Civ. Engrg. Dept., Indian Institute of Technology, Bombay.

Jiang Guiyan , Wang Jiangfeng, Zhang Xiaodong, Gang Longhui. (2003) The Study on the Application of Fuzzy Clustering Analysis in the Dynamic Identification of Road Traffic State, The IEEE 6' International Conference on Intelligent Transportation Systems-Shanghai, CHINA-October 12-15

Jiang Guiyan , Wang Jiangfeng, Zhang Xiaodong,. (2004) Research on algorithm of ACI for urban expressway by using ANN, The 8th International Conference on Application of Advanced Technologies in Transportation Engineering.

Jiang Guiyan. (2004) Technologies and applications of the ldentification of road traffic conditions, China Communications Press.

Jorge Nocedal, Stephen J. Wright. (1999) Numerical optimization. Springer Press, Berlin.

Kerner, B.S and Rehborn, H. (1996) Experimental features and characteristics of traffic jams Physical Review E 1297-1300

Kerner, B.S. and Rehborn, H. (1996) Experimental properties of complexety in traffic flow Physical Review E 53 5

Lao Yunteng, Yang Xiaoguang Yun Meiping et al. (2006) evaluation of traffic state detection method, Computer and communications, vol. 24, pp. 74-77

Narayanan R., Udayakumar R. (2003) Quantification of congestion using Fuzzy Logic and Network Analysis using GIS, Transportation. Map India

Robert L. Bertini. (2006) You Are the Traffic Jam: An Examination of Congestion Measures, Submitted for presentation at the 85th Annual Meeting of the Transportation Research Board January 2006, Washington, D.C.