2011 International Conference on Computer and Automation Engineering (ICCAE 2011)
IPCSIT vol. 44 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V44.23

Urban Arterial Road Traffic Status Identification Based on Data Fusion

Jiao Zou[1+], Xiaoxu Bei[2] and Jian Wu[3]

[1] Anhui Intelligent Transportation Key Lab. Huangshan Road 628,230088 Hefei

[2]German Aerospace Center (DLR) Institute of Transportation.Rutherfordstr.2,12489 Berlin,

Abstract: The accuracy of traffic status based on FCD (Floating Car Data) is related to taxi coverage rate and sample number. Small sample number leads to less accurate results. If no data is uploaded during some period, it will use historical archive data to infer the results which can cause error. To improve reliability of traffic information and provide accurate information to traffic management department, we carried out the traffic information fusion process which is based on camera and FCD. The two data sources can complement each other. The application scenario of this paper is based on information acquired through camera and FCD on main roads of Hefei city. We adopt decision tree methods to integrate these two data sources and get more reliable traffic status which facilitates the management and control of traffic system.

Keywords: ITS, Camera, FCD, Decision Fusion

1. Introduction

Traffic congestion is an important problem which confronts most big cities. Traffic governors use advanced system to control the whole urban road network, gathering real-time traffic information by detector equipments to regulate control strategy and do guidance to traffic flow which can alleviates traffic congestion to some extent. At present, there are many methods to get the traffic information, and most of them can be grouped into two classes: automatic and manual. The detectors include mobile detector and fixed detector. The data gathered by single type detector is limited and short, which can't reflect the whole road traffic status accurately, so we use the traffic information gathered by various detectors as data sources to do the data fusion, then we can get more exact traffic information to reflect real-time road network running effectively, and provide reliable data to governors and travelers.

Data fusion process can be done in three levels: data level, feature level, decision level. Data level fusion is the lowest level, which fuses original data directly; feature level fusion distills original data character, then analyze and deal with the feature data; decision level is the upmost level, according to practical situation, aiming to make decision by using the result of feature level fusion. This paper, using the traffic information identified based on camera and FCD as data source, fuse the information together by decision level data fusion, then we can get more accurate traffic information of road network.

2. Traffic Status Determination

2.1. Traffic Status Based on Camera

Traffic identification algorithm based on camera uses archive data of last month as training sample, including volume, speed and occupancy. Then these data is standardized, clustered by k-means algorithm, finally we get four clusters (just the four traffic state: free flow, normal flow, retard flow and congestion flow), and according to the distance between the sample and cluster centers, we can assign the sample to the

[+] Jiao Zou Tel.: +0086-05515338221; fax: +0086-05515338248.
E-mail address: xuejiao35@163.com

cluster from which the distance is smallest. If the roads do have clear boundary among the four states, the precision can reach 90%.

2.2. Traffic Status Based on FCD

The underlying principle of FCD method is computing average traveling speed from taxies with GPS. The GPS device can report the time and position at any time point. Therefore, we can get velocity through time difference and position offset for one taxi. Then the average traveling speed of some road is obtained by averaging the velocity among all samples. At last, we set threshold speed for different traffic status according to national standards and empirical experience. The accuracy of FCD depends on many factors, but heavily hinges on the amount of samples. It can get considerable good results when the sample size is greater than 4. However, not all of the roads can be covered by taxis during every time periods. Hence there is a need to include fixed detectors to complement these blind sites.

3. Comparison Analysis

Camera data and FCD data are both stored in Oracle database, but different hosts. We extract both the camera result and FCD result during the same period to ease the comparison analysis. The traffic status is represented by four numbers: 1, 2, 3, 4. Number 1 represents free flow, while number 4 represents congestion flow.

We select four cameras installed on two main roads in Hefei city: Huizhou road (south-north) and Huangshan road (east-west). Meanwhile, the traffic status of corresponding FCD road section is extracted. The camera IDs and corresponding FCD road sections are described in Table. 1.

Table.1. Cameras and road sections

Camera ID	Road Section ID
Camera047	263
Camera048	273
Camera027	834
Camera028	835

3.1. Comparison Analysis (FCD sample size >= 0)

The comparison is carried out between 8:00 am, August 23rd, 2010 and 14:00 pm, August 24th, 2010. The results are shown in Table. 2 and Fig.1. N is FCD sample size.

Table.2. Difference of Traffic State Value of camera and FCD（n>=0）

ROADID	DIFFERENCE：0	DIFFERENCE：1	DIFFERENCE：-1	Else
263	45.1%	23.9%	27.8%	3.2%
273	43.5%	45.8%	0	10.7%
834	35.8%	28.2%	17.1%	8.9%
835	58.9%	7.3%	28.5%	5.3%

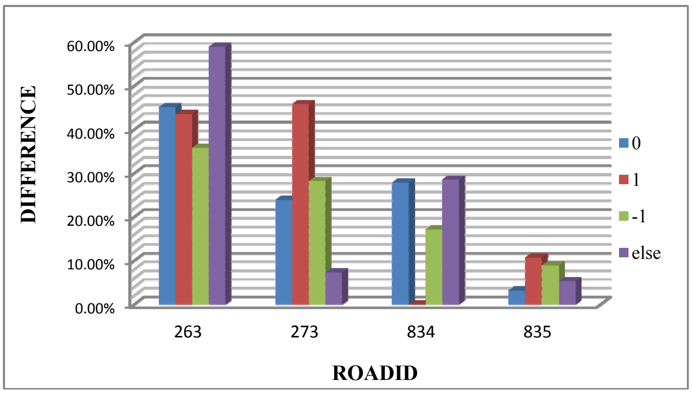

Fig. 1 .Difference of traffic state value between camera and FCD （n>=0）

From Table.2and Fig.1 we can see that there's about 50% chance that the two results coincide with each other.

3.2. Comparison Analysis (FCD sample size > 0)

Since there's no taxi running during night, it's meaningless to compare it with camera results. So we remove the data item of which taxi sample size is 0. Again the results are shown in Table.3 and Fig.2.

Table 3. Difference of traffic state value between camera and FCD （n>0）

ROADID	DIFFERENCE：0	DIFFERENCE：1	DIFFERENCE：-1	Else
263	46%	9%	41%	4%
273	27%	49%	0	24%
834	41%	16%	31%	12%
835	39%	8%	44%	9%

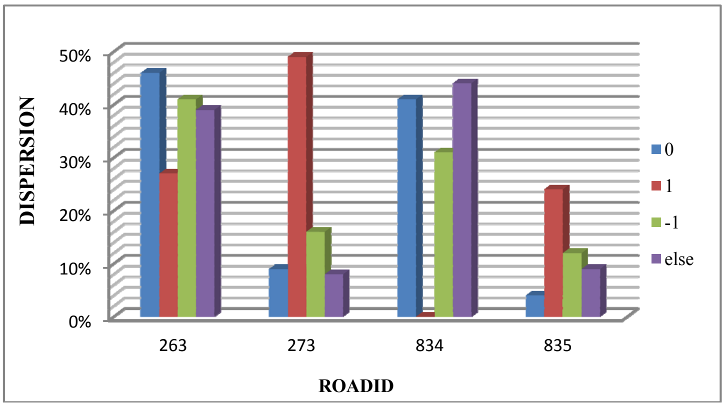

Fig. 2 Difference of traffic state value between camera and FCD（n>0）

We can see from Table. 3 and Fig.2that there's a significant reduction of total consistency. It's reasonable because during night status 1 is dominant for both FCD and camera result.

From the analysis above, we argue that FCD and camera result coincide with each other more or less. It's feasible to merge these two methods.

4. Information Fusion Based on Decision-Tree

Since the raw data of FCD differs vibrantly with that of cameras and it's hard to extract feature data for FCD, we decide to fuse them at decision level using decision tree method. Decision tree is a machine learning approach which tries to learn a bunch of rules from training data sets. Usually the data has the format: $D = (A_1, A_2, \cdots, A_n; C)$; A_i is attribute value, $C \in \{C_1, C_2, \cdots, C_k\}$ is class label.

Each internal node of decision tree represents a test on an attribute, each leaf node indicates final class label.

The most popular methods to build a decision tree is C4.5 algorithm. It starts from root node and performs split test until some criterion is meet. A common criterion used by C4.5 is IG (Information Gain) maximization. Given a probability distribution: $P = (p_1, p_2, \cdots, p_n)$,

the information amount contained in this distribution is

$$I(P) = -(p_1 \times \log_2 p_1 + p_2 \times \log_2 p_2 + \cdots p_n \times \log_2 p_n) .$$

Assume an item set T is grouped into mutually exclusively classes: $(C_1, C_2, \cdots, C_k)$,

then to identify which class does an item belong to, we need information INFO(T)=I(P), where P is distribution of $(C_1, C_2, \cdots, C_k)$,that is : $P = (|C_1|/|T|, |C_2|/|T|, \cdots, |C_k|/|T|)$.

If we split the set T into sets $T_1, T_2, \cdots, T_m$, According to some attribute: $X, X \in \{A_1, A_2, \cdots, A_n\}$, then we have $INFO(X,T) = \sum_{i=1}^{m} |T_i|/|T| \times INFO(T_i)$

Thus we have $IG(X,T) = INFO(T)\text{-}INFO(X,T)$

Infoamtion gain is the difference value two information amount: one is the information needed to identify one element of T, the other is information needed to identify one element of T after knowing the value of some attribute X.

In this paper, we use four attributes: A1, A2, A3, A4 and one class label C. Their meaning is given in Table.4.

Table.4 Attributes and class labels

Attributes	Class lable
A1	Traffic status obtained by FCD
A2	The number of FCD taxi sample
A3	Traffic status obtained by camera
A4	The reliability of camera results
C	Traffic Status after fusion

We get training set by the following steps: pick some time periods and some road, manually observe the traffic status, then compute FCD result and camera result. Decision tree procedure is carried out by J48 implementation in Weka package.

5. Validation

We choose data collected by camera and FCD on road section: 263 and 835 from 6:00 September 6th 2010 to 6:00 September 7th 2010. Then fusion process is carried out to get traffic state. Fig. 3(a) and 3(b) show the results after the fusion process. The red line in Fig. 3 is camera results; blue one is FCD results; green one is merged results. According to our observations, traffic congestion mainly appears at rush hours, which can be seen from Fig. 3. In Fig. 3(a) the merged results overlaps more with camera results than FCD in that taxi coverage rate is low on this road section. However, in Fig. 3(b), the reverse is true.

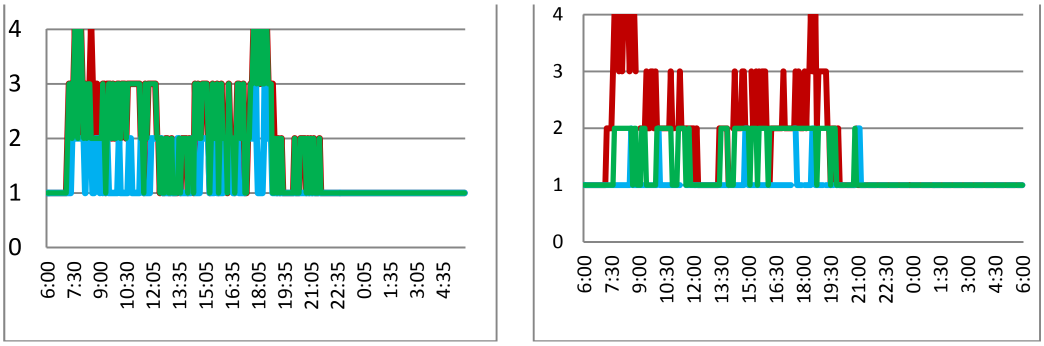

Fig.3. (a) Fusion Results of Camera and FCD(1); (b) Fusion Results of Camera and FCD(2)

It's because the taxi coverage rate is relatively high on this section of road. While congestion seldom occurs on this road, clustering methods may seem less accurate on the determination of traffic status four since it must get four clusters though there may be no congestion flow. FCD classifies the status by absolute threshold speed and may be more applicable in this situation.

In general, camera detectors, or other types of fixed detector installed on urban arterial road can gather fixed spot traffic information, while FCD is moving, and they can cover most of urban roads collectively. In this paper, camera data and FCD data are fused in Hefei city, and precision of traffic status identification can reach 91%. The uncertainty of road traffic status can be reduced, and reliability can be improved. Moreover, it provides real-time dynamic traffic information of the road-networks to governors for traffic signal control and guidance which can evacuate traffic flow, keeping it running smoothly and relax traffic congestion.

6. Acknowledgement

This work is funded by Ministry of Industry and Information Technology of China. Yang and Wu provide FCD and camera data, Zou and Bei perform analysis and fusion process as well as accomplishing manuscript. They all contribute equally to this paper.

7. Reference

[1] Hall, M., E. Frank, et al.. The WEKA data mining software: An update[J]. ACM SIGKDD Explorations Newsletter 11(1).pp: 10-18.2009

[2] Han, J. and M. Kamber. Data mining: concepts and techniques[R], Morgan Kaufmann. 2006

[3] Hawas, Y. E.. A fuzzy-based system for incident detection in urban street networks. [R]Transportation Research Part C: Emerging Techologies 15.pp: 69~95.2007

[4] Jiuh-bijing sheu, S. G. A new methodology for incident detection and characterization on surface streets. [R]transportation research part C6.pp: 315~335.1999.

[5] Pi Xiao Liang，Yang Xiaoguang，Sun Ya. Traffic State Classify Method Research Based on Loop detector [J]. Huadong Highway,pp: 4~6. 2006.1

[6] Xiao Yonglai. Urban Road Traffic State Identification Technology Research Based On SCATS [J]. China Traffic Information Industry. pp: 39~41. 2005.6

[7] Jiang Guiyan. Technologies and Applications of Identification of Road Traffic Condition [M]. Beijing, China Communications Press. 2004.

[8] Jiang Guiyan.，Jiang Longhui，Wang Jiangfeng. Method of Traffic Condition Identification of Urban Free Road [J]. Transportation Engineering Transaction. pp: 87~91. 2006.6