Traffic state estimation is a key problem with considerable implications in modern traffic management. A simple, general, and complete approach to the design of urban network traffic state and phase estimator has been developed in this paper. A uniform traffic state dynamic estimation method structure is designed which consists of three steps. (1) Floating-car data and radio frequency identification data preprocessing method is proposed to remove the abnormal data and finish the map matching process. (2) Section speed estimation method is proposed based on the degree of confidence. (3) Traffic phase identification method is proposed based on the estimated section speed. A number of simulation and field investigations have been conducted to test the estimator performance. The investigation results indicate that the proposed approach is of high accuracy and smoothness on the section speed estimation and effectively eliminates the influence of abnormal data fluctuations and insufficient data. And the traffic phase identification method can effectively filter out the abnormal distortion of estimated section speed around the threshold value and modify the phase step of traffic status caused by abnormal data.

1. Introduction

Real-time traffic state estimation is a fundamental task for urban traffic management centers and is often a critical element of Intelligent Transportation Systems (ITS). Intelligent Transportation System (ITS) has been regarded as important means to solve traffic jams in the city [1]. Traffic state estimation for an urban network refers to estimating the traffic congestion level of the network at the current time based on available real-time traffic-aware data. For this purpose, various sensors are used to collect traffic information. To restore the real traffic conditions, raw traffic information should be processed by various methods. By using multiple traffic data sources we will make better inferences to obtain the traffic-aware data. And this is a reproduction of real traffic system that could not be achieved from a single source of data. Furthermore, as sources of data are becoming increasingly available, traffic status estimation needs to be applied not only to realize these benefits, but also to use the data more efficiently.

For traffic state estimation, a limited number of research works produced and proposed corresponding estimation algorithms that were almost exclusively based on the seminal methodology of Kalman filtering [2] and its extensions for nonlinear systems [3]. The Kalman filter is an optimal state estimator applied to a dynamic system that involves random noise and includes a limited amount of noisy real-time measurements. Although it was originally derived for linear systems, the Kalman filter can also be extended for its application to nonlinear systems via specific on-line Taylor expansions of the originally nonlinear systems. The Kalman filter obtained in this way is called extended Kalman filter (EKF).

Early applications of traffic state estimation were reported in traffic surveillance systems for short interdetector distances by Gazis and Knapp [4], Knapp [5], Szeto and Gazis [6], Nahi and Trivedi [7], and Nahi [8]. In these and later investigations, for example, Smulders [9] and Bhouri et al. [10], only short freeway sections with a length below 2 km were considered; see a review by Cremer [11]. The applied
models were relatively simple (due to the short section lengths). Later approaches started using more comprehensive dynamic traffic flow models, which opened the way to the consideration of longer freeway stretches (2–4 km); see Cremer [12] and Cremer et al. [13], while more recent investigations elaborated on some technical details based on previously proposed basic ideas; see Kohan and Bortoff [14] and Meier and Wehlan [15]. Other modeling approaches have been used as well as the Kalman filter, such as the neural networks (e.g., [16–20]) and others (e.g., [21, 22]). Karlaftis and Vlahogianni [23] compare statistical methods and neural networks in transportation research, highlighting some of the differences similarities of the two types of data analysis tools. Wang et al. [24, 25] developed a general approach to the real-time estimation of the complete traffic state in freeway stretches based on the extended Kalman filter.

In China, several approaches have been proposed based on the development of social economy and the technology: Jia et al. [26] used the floating-car data in Beijing to identify the traffic state based on average vehicle speed. Dou et al. [27] proposed an algorithm of logistic regression for traffic state probability forecast to obtain the accurate and objective traffic state information to meet the demand of traffic guidance. Wang and Guo [28] present the heterogeneous data fusion methods and road traffic code model in Beijing.

As indicated above, while there has been substantial progress in developing traffic state estimation technologies. Compared with the research on freeway network, fewer researches on the approaches of dynamically estimating urban area network traffic states have been found in the literature. Developing the dynamic estimation of urban road network traffic states is of critical importance in managing urban traffic. The primary objective of this paper is developing a traffic state dynamic estimation method based on degree of confidence for urban road network.

This paper presents an urban road network traffic states estimation approach that combines a section speed estimation procedure based on confidence value and a traffic phase identification method to provide the most reasonable traffic phase identification results. The proposed method was evaluated in part of the roads network in Nanjing, China, and its effectiveness compared with the weighted average algorithm.

This paper is organized as follows. The first part is the introduction of traffic states estimation method. The second part is the section speed estimation model based on confidence value. The third part is traffic phase identification methods based on section speed and performance evaluation. Last part is conclusions.

2. Traffic State Estimation

There are some related terms used in the remainder of this paper; in order to avoid any ambiguity or confusion between them, they are clearly defined here based on existing research [27]. Traffic state: the state in which traffic is at any given time can be described by a number of parameters such as flow, density, and speed; as these are continuous variables, there can be an infinite number of traffic states. Traffic phase: simple traffic flow theory models assume a small (usually 2 or 3) number of traffic phases, which have direct physical interpretations (e.g., congested or uncongested). Traffic-aware data: a number of real-time estimated parameters based on raw traffic data obtained from traffic sensors by big data technology, which is used to describe traffic states.

Figure 1 shows the structure of traffic state estimation method based on traffic-aware data developed in this study.

Input data include floating car data, radio frequency identification data, video data, and network data. Floating car data is generally updated at intervals ranged from 1 second to 10 minutes and is delivered using GPRS sentence. After interpreted and processed, the data includes information about the terminal, time, longitude, latitude, speed, flag, rate, and update time. Electronic license plates with RFID tags are installed on part of vehicles, which enable the road side RFID detectors to get the passing time, encrypted license plate number, and other vehicle properties when they passing by. Both the road side RFID detectors and video detectors are installed at the top of intersection exit. And after being interpreted and processed, the video data enables us to get the section flow volume. Network geometry data are static provided by government. After real-time raw data reading and storage, the raw data will be preprocessed to amend the errors. Degree of confidence calculation which is based on data analysis will be started as soon as raw data is well preprocessed. And historic data storage approach is proposed to support section average speed estimation. Then traffic-aware data is calculated based on the different input data and degree of confidence. Next, degree of confidence based section average speed estimation model is established to calculate traffic phase identification parameters. At last, traffic phase identification method is developed to get the section traffic phase, and final outputs are traffic phase estimated results.

3. Raw Traffic Sensors Data Analysis and Preprocessing

Usually, the raw traffic sensors data is not good enough to be used directly. In this section, raw data analysis results and raw data preprocessing method are proposed.

To remove the abnormal data, floating car data preprocessing is preliminary screening of raw floating car data. Abnormal data mainly includes three categories: abnormal velocity data, wrong direction data, and time duplicate data.

Based on the floating data collected between November 19, 2009, 7:00 to 19:00 in Hangzhou, China, the abnormal data is analyzed. The total number of floating car is 3839. The total sample size is 2045574. Filter results are shown in Table 1.

As the positioning accuracy of GPS data is limited, vehicles cannot be located on the road accurately most of time (as is known to all, vehicles should be on the road). Map matching algorithms integrate positioning data from a Global Positioning System (or a number of other positioning sensors) with a spatial road map with the aim of identifying the road segment on which a user (or a vehicle) is travelling
and the location on that segment. There are many researches about map matching algorithms [29–34] in recent years. Amongst the family of map matching algorithms consisting of geometric, topological, probabilistic, and advanced, topological map matching algorithms are relatively simple, easy, and quick, enabling them to be implemented in real-time. Thus in this research topological map matching algorithms are used to preprocess the floating car data.

Radio frequency identification data is a kind of fixed single point detector data. Points matching algorithm is proposed to filter out records of the same car when passing the detectors installed at upstream and downstream of the section. The algorithm is shown in Figure 2 to preprocess the raw radio frequency identification data in traffic state estimation.

Based on the pair of point record, section travel time is calculated by the passing time. Then it is combined with the section length between this pair of point to obtain travel speed. In this step, abnormal data is removed from the database. Abnormal velocity data is the main abnormal RFID data.

Video data is preprocessed by license plate recognition technology, only to provide the real-time traffic volume of each section in this study because of the low precision of recognition (less than 30%).

Raw traffic sensors data analysis and preprocessing is the first step to estimate traffic state. In next section, section speed estimation method is proposed.

4. Section Speed Estimation Method Based on Degree of Confidence

4.1. Algorithm Outline and Explanation of Parameters. In this study, section speed is used to estimate the traffic phase in each section. Floating car data and RFID data both are collected from part of vehicles in road section. Considering the characteristics of the floating cars and vehicles with RFID tags, weighted average method based on the degree of confidence is used to calculate the section speed. The degree of confidence is obtained by sample size of floating car data and RFID data as well as the data features of floating car data and RFID data.

Concretely speaking, if the sample size of data is enough to estimate the section speed, we adopt weighted average method to calculate the section speed after map-matching. If the sample size of data is not enough, recent data and historic data of the same period fusion method are used. There are some related terms used in the algorithm models of this method; in order to avoid any ambiguity or confusion between them, they are clearly defined here.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Number of abnormal data</th>
<th>Share of total sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal velocity data</td>
<td>19408</td>
<td>1%</td>
</tr>
<tr>
<td>Wrong direction data</td>
<td>141403</td>
<td>7%</td>
</tr>
<tr>
<td>Time duplicate data</td>
<td>19591</td>
<td>1%</td>
</tr>
</tbody>
</table>
Single vehicle weight \( w \): the method assigns different weight for every piece of data (FCD and RFID) by an algorithm proposed in this paper based on speed and sample size of a section.

Degree of confidence \( R \): degree of confidence \( R \) is the sum of sample weights in one section. It is an important parameter in recent data and historic data of the same period fusion method.

Sample size \( n \) and minimum sample size \( n_{\text{min}} \): in order to insure the estimated section speed to reflect the real-time traffic state, sample size in one section should not less than minimum sample size \( n_{\text{min}} \). If the real-time sample size is not enough, we adopt recent data (e.g., 5 minutes ago data) and historic data of the same period to estimate the section speed. Value of \( n_{\text{min}} \) is obtained by simulation experiments [35].

Count variable of small samples \( m \) and confidence attenuation parameter \( m_{\text{max}} \): continuous lack of sample times in section \( i \) is counted as \( m_i \). When \( n < n_{\text{min}}, m_i = m_{i-1} + 1 \) and when \( n > n_{\text{min}}, \) the value of \( m_i \) returns to zero. \( m_{\text{max}} \) is confidence attenuation parameter, which indicates that recent data before \( m_{\text{max}} \) data process cycles is invalid to estimate the section speed.

4.2. Single Vehicle Weight and Degree of Confidence. In this study, single vehicle weight is calculated based on FCD and RFID speed data, sample size, and speed distribution in one section. And the single vehicle weight is used to get the degree of confidence. For each piece of RFID and FCD data, we divide them by the speed: high speed data, medium speed data, and low speed data, and separate weighting factors are assigned to each group (high speed weighting factor: \( C_H \); medium speed weighting factor: \( C_M \), low speed weighting factor: \( C_L \)). At the same time, for each section, count the number of data in each group. \( n_H \) is the number of high speed data; \( n_M \) is the number of medium speed data; \( n_L \) is the number of low speed data. Then we calculate the high speed single vehicle weight as the following formula:

\[
\begin{align*}
\nu_H &= \frac{n_H}{n_H + n_M + n_L} \cdot C_H. \\
\end{align*}
\]  

The same procedure may be easily adapted to obtain the medium speed single vehicle weight and the \( \nu_M \) low speed single vehicle weight \( \nu_L \). Then the high proportion data will be assigned a higher weight. At the same time, adjusting the weighting factors can assign different weight to each group data.

Based on the above method, degree of confidence is calculated as the following formula:

\[
R = \sum_{i=1}^{n} \nu_i. 
\]  

In (2), \( n \) is the sample size of the section.

The next step is to estimate the section speed. Section speed estimation process based on degree of confidence is shown in Figure 3.

If the sample size is enough \( (n > n_{\text{min}}) \), section speed is estimated as

\[
\bar{v} = \frac{\sum v_i \cdot \nu_i}{\sum \nu_i}. 
\]
Mathematical Problems in Engineering

Table 2: Speed distribution by the speed threshold to identify the traffic phase in Nanjing.

<table>
<thead>
<tr>
<th>Time</th>
<th>Amount of data</th>
<th>$v = 0$</th>
<th>$0 &lt; v \leq 15$</th>
<th>$15 &lt; v \leq 30$</th>
<th>$v &gt; 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>AVG (km/h)</td>
<td>Proportion</td>
<td>AVG (km/h)</td>
<td>Proportion</td>
</tr>
<tr>
<td>7:00</td>
<td>9151</td>
<td>65%</td>
<td>6.23</td>
<td>10%</td>
<td>23.20</td>
</tr>
<tr>
<td>7:15</td>
<td>9297</td>
<td>60%</td>
<td>6.52</td>
<td>13%</td>
<td>23.25</td>
</tr>
<tr>
<td>7:30</td>
<td>9403</td>
<td>55%</td>
<td>6.88</td>
<td>15%</td>
<td>22.97</td>
</tr>
<tr>
<td>7:45</td>
<td>9311</td>
<td>52%</td>
<td>6.75</td>
<td>17%</td>
<td>22.94</td>
</tr>
<tr>
<td>8:00</td>
<td>8916</td>
<td>49%</td>
<td>6.92</td>
<td>18%</td>
<td>22.89</td>
</tr>
<tr>
<td>8:15</td>
<td>9010</td>
<td>47%</td>
<td>7.52</td>
<td>19%</td>
<td>22.74</td>
</tr>
<tr>
<td>8:30</td>
<td>8888</td>
<td>46%</td>
<td>7.25</td>
<td>18%</td>
<td>22.91</td>
</tr>
<tr>
<td>8:45</td>
<td>9380</td>
<td>45%</td>
<td>7.48</td>
<td>17%</td>
<td>22.96</td>
</tr>
<tr>
<td>9:00</td>
<td>9240</td>
<td>45%</td>
<td>7.52</td>
<td>17%</td>
<td>22.88</td>
</tr>
</tbody>
</table>

And if the sample size is not enough ($n < n_{\text{min}}$), section speed is estimated based on real-time data, recent data, and historic data as the following formula:

$$\bar{v}_{i} = \frac{\sum v_{i} w_{i} (1 - k) + k \cdot [\bar{v}_{0} R_{0} j + \bar{v}_{s} R_{s} (1 - j)]]}{\sum w_{i} (1 - k) + k \cdot [R_{0} j + R_{s} (1 - j)]}.$$  \hspace{1cm} (4)

In (4),

$$k = \begin{cases} 
1 - \frac{n}{n_{\text{min}}} & n < n_{\text{min}} \\
0 & n \geq n_{\text{min}} 
\end{cases}$$  \hspace{1cm} (5)

$$j = \begin{cases} 
1 - \frac{m}{m_{\text{max}}} & m < m_{\text{max}} \\
0 & m \geq m_{\text{max}} 
\end{cases}$$

$R_{0}$ is degree of confidence of last time period data. $\bar{v}_{0}$ is the estimated section speed in last data process cycle. $R_{s}$ is degree of confidence of historic data in the same period. $\bar{v}_{s}$ is the estimated section speed based on historic data in the same period.

In (4), when $k$ takes value 0 and $j$ takes value 1, it reflects that there is no real-time data and recent data sample size is enough to estimate the section speed, and thus recent data is used to calculate the section speed. When $k$ takes value 1 and $j$ takes value 0, it reflects that there is no real-time data and recent data sample size is not enough to estimate the section speed, and thus historic data is used to calculate the section speed. In (4), when each parameter takes the extreme value, the formula corresponds to a special data status. And when each parameter takes the other value, the formula estimates the section speed by real-time data, recent data, and historic data.

5. Traffic Phase Identification Method Based on Section Speed

In this paper, we analyze the raw FCD and RFID data in March 29, 2011, 7:00 to March 29, 2011, 7:00, Nanjing, China. Table 2 shows some of our results about speed distribution by the speed threshold to identify the traffic phase in Nanjing.

The proportion of each speed group is almost balanced based on present traffic phase threshold. After removing invalid data and zero speed data, the raw speed data that is close to the threshold is analyzed, and the result is shown in Table 3.
Table 3: Near-threshold raw speed data analysis.

<table>
<thead>
<tr>
<th>Time</th>
<th>Amount of valid data</th>
<th>Amount</th>
<th>Proportion</th>
<th>Amount</th>
<th>Proportion</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>7:00</td>
<td>3216</td>
<td>280</td>
<td>8.7%</td>
<td>346</td>
<td>10.8%</td>
<td>19.5%</td>
</tr>
<tr>
<td>7:15</td>
<td>3735</td>
<td>385</td>
<td>10.3%</td>
<td>465</td>
<td>12.4%</td>
<td>22.8%</td>
</tr>
<tr>
<td>7:30</td>
<td>4240</td>
<td>490</td>
<td>11.6%</td>
<td>454</td>
<td>10.7%</td>
<td>22.3%</td>
</tr>
<tr>
<td>7:45</td>
<td>4475</td>
<td>500</td>
<td>11.2%</td>
<td>469</td>
<td>10.5%</td>
<td>21.7%</td>
</tr>
<tr>
<td>8:00</td>
<td>4578</td>
<td>529</td>
<td>11.6%</td>
<td>499</td>
<td>10.9%</td>
<td>22.5%</td>
</tr>
<tr>
<td>8:15</td>
<td>4746</td>
<td>658</td>
<td>13.9%</td>
<td>473</td>
<td>10.0%</td>
<td>23.8%</td>
</tr>
<tr>
<td>8:30</td>
<td>4834</td>
<td>564</td>
<td>11.7%</td>
<td>512</td>
<td>10.6%</td>
<td>22.3%</td>
</tr>
<tr>
<td>8:45</td>
<td>5156</td>
<td>585</td>
<td>11.3%</td>
<td>554</td>
<td>10.7%</td>
<td>22.1%</td>
</tr>
<tr>
<td>9:00</td>
<td>5074</td>
<td>600</td>
<td>11.8%</td>
<td>554</td>
<td>10.9%</td>
<td>22.7%</td>
</tr>
</tbody>
</table>

Table 3 shows that more than 20% of the original data are close to traffic phase threshold. When using these data to identify traffic phase, traffic phase results will vary between two phases frequently. This result is not in accordance with the traffic continuity, and thus traffic phase identification method is proposed to make traffic phase identification results more accurate.

The process of traffic phase identification method is shown in Figure 4.

Speed threshold value and traffic phase are defined as follows.

When section speed is below \( v_1 \), preidentified traffic phase \( S = 0 \) in this section. When section speed is between \( v_1 \) and \( v_2 \), preidentified traffic phase \( S = 2 \) in this section. When section speed is between \( v_2 \) and \( v_3 \), preidentified traffic phase \( S = 3 \) in this section. And when section speed is above \( v_n \), preidentified traffic phase \( S = n \) in this section. In the case described above, we make \( 0 < v_1 < v_2 < v_3 < \cdots < v_n \). In general, \( n \) gets the value 2, which indicates that the section has three traffic phases (free-flow phase, uncongested phase, and congested phase).

Near-threshold speed correcting method and step change correcting method are described as the following formula:

\[
S'_i = \begin{cases} 
S_0, & |\bar{v}_i - \bar{v}_0| < \alpha, |\bar{v}_i - v| < \beta \\
S_0 + 1, & \bar{v}_i - \bar{v}_0 > \gamma, |\bar{v}_i - v| < \beta \\
S_0 - 1, & \bar{v}_0 - \bar{v}_i > \gamma, |\bar{v}_i - v| < \beta \\
S_i', & \text{Other.}
\end{cases}
\]  

(6)

In (6), \( S'_i \) is the estimated traffic phase in recent data process cycle (estimated result). \( S_0 \) is the estimated traffic phase in last data process cycle. And \( S_i \) is the preidentified
traffic phase in recent data process cycle. $\bar{V}_0$ is the section speed in last data process cycle, and $\bar{V}_i$ is the section speed in recent data process cycle. $v$ is the threshold that section speed is close to in recent data process cycle. $\alpha$ is a parameter to judge whether the section speed changing in a data process cycle is tiny. $\beta$ is a parameter to judge whether section speed is close to the traffic phase threshold in recent data process cycle. $\gamma$ is a parameter to judge whether the section speed changes obviously in a data process cycle. 

When $|\bar{V}_i - \bar{V}_0| < \alpha$ and $|\bar{V}_i - v| < \beta$, the section speed changing is tiny and it is close to traffic phase threshold. Because of the traffic continuity, $S'_i = S_0$. 

When $\bar{V}_i - \bar{V}_0 > \gamma$ and $|\bar{V}_i - v| < \beta$, the section speed accelerates significantly and it is close to traffic phase threshold. It shows that the traffic phase changes obviously, so $S'_i = S_0 + 1$. The same procedure may be easily adapted to obtain the estimated traffic phase when $\bar{V}_0 - \bar{V}_i > \gamma$ and $|\bar{V}_i - v| < \beta$.

In other cases, for example, when $\bar{V}_i - \bar{V}_0 > \gamma$ and $|\bar{V}_i - v| > \beta$, the section speed accelerates significantly, and it is not close to traffic phase threshold. The preidentified traffic phase can reflect actual traffic phase, so $S'_i = S$.

In (6), three parameters are calibrated by simulation test in this study (see performance evaluation). And after the traffic phase identification process, traffic phase is obtained finally.

### 6. Performance Evaluation

In this study, FCD and RFID speed data are used to estimate section speed, and data sample size, speed distribution, and the continuity of the traffic flow are considered in the estimation method to improve the accuracy and stability of estimated results. In order to test the effect of section speed estimation method, numerical test is performed by using speed average method and the method in this paper based on simulation data. In this experiment, speed above 30 km/h belongs to high speed data; speed between 15 km/h to 30 km/h belongs to medium speed data; speed below 15 km/h belongs to low speed data. And speed weighting factors are $C_H = 0.4$, $C_M = 0.5$, and $C_L = 0.1$. The result is shown as in Figure 5(a).

As can be seen from the illustration, when the sample size is small, section speed estimated by speed average method has a larger error, and the result has a relatively large fluctuation as time goes on. However the result of method in this study is more relevant to expectation value. This algorithm improves the accuracy and stability of estimated results. In addition, the real-time FCD and RFID data on January 19, 2009, is used to conduct a field test in Nanjing, China. We use the data of November 12, 2008, to November 14, 2008, to establish the historic database. Experimental place is the section between Etiao Rd. and Yixian Bridge on Zhongshan East Rd. The truth value of section speed is obtained by live video capture and analysis. The result is shown in Figure 5(b).

In this experiment, accuracy of the section speed estimation method in this paper is 87.25%, and trend of the result is more relevant to truth value. Error analysis is shown in Table 4.

Traffic phase identification method is also tested by a numerical experiment. According to the MFD in the experiment area section speed above 30 km/h belongs to free-flow phase; section speed between 15 km/h to 30 km/h belongs to uncongested phase; section speed below 15 km/h belongs to congested phase. The traffic phase threshold is $v_i = 15$ km/h, $v_0 = 30$ km/h. And the parameter is $\alpha = 5$ km/h, $\beta = 2.5$ km/h, and $\gamma = 12.5$ km/h. The experiment result is shown in Figure 6(a).

#### Table 4: Error analysis.

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean error (Km/h)</th>
<th>Mean absolute error (Km/h)</th>
<th>Mean percentage error</th>
<th>Mean absolute percentage error</th>
<th>Root absolute error (Km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed average method</td>
<td>-3.40</td>
<td>6.39</td>
<td>-0.10</td>
<td>0.19</td>
<td>15.84</td>
</tr>
<tr>
<td>Method in this paper</td>
<td>1.37</td>
<td>4.21</td>
<td>0.05</td>
<td>0.13</td>
<td>7.58</td>
</tr>
</tbody>
</table>

![Figure 5: The result and verify result of section speed estimation method](image-url)
In this experiment, 600 section speed results have been identified by the method in this paper. 49 results have been corrected (accounts for 8.1%). And 83 traffic phase changes are fixed by this method (accounts for 13.8%). The results show that the traffic phase identification method makes the results more accurate and reasonable.

In addition, a case study based on the estimated section speed data of March 18, 2009, is presented to verify the effectiveness and feasibility of this method. The result is shown in Figure 6(b).

In this case study, 5 traffic phase identification results have been corrected. By comparing the results revised and the original one, it is found that the optimization result is more reasonable.

### 7. Conclusions

The primary objective of this paper is developing a traffic state dynamic estimation method based on degree of confidence for urban road network. A simple, general, and complete approach to the design of urban network traffic state and phase estimator has been developed in this paper. A uniform traffic state dynamic estimation method structure is designed which consists of FCD and RFID preprocessing method to remove the abnormal data and finish the map matching process. Then, section speed estimation method was proposed based on degree of confidence. Traffic phase identification method is proposed based on the estimated section speed. A number of simulation and field investigations have been conducted to test the estimator performance. The investigation results indicate that the proposed approach is of high accuracy and smoothness on the section speed estimation and effectively eliminates the influence of abnormal data fluctuations and insufficient data. And the traffic phase identification method can effectively filter out the abnormal distortion of estimated section speed around the threshold value and modify the phase step of traffic status caused by abnormal data resulting in optimized results of discrimination.

Now, this research is applied in Nanjing, China. A transportation information service system is established to provide a service to the government. Future research includes the development of an efficient calibration method for the parameters in this paper and the enhancement of the simulation model to evaluate the performance of the strategy in large road network.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
Acknowledgments

This research was made possible through funding from the Ministry of Science and Technology of the People's Republic of China under Grant 2012AA12306. The authors wish to acknowledge the vast and helpful assistance from Intelligent Transportation System Research Center Transportation System Engineering & ITS Research Group, School of Transportation Engineering, Tongji University. The Nanjing Traffic Information System was established with the great help of Jitong Traffic Tech Co., Ltd.

References


