Traffic State Detection Based on Equipped Vehicle Data
Yun-feng SHI, Li-cai YANG* and Shen-xue HAO
Department of Control Science and Engineering, Shandong University, Jinan, China
*Corresponding author

Keywords: Traffic state, Smart-phone, KHDD, Smart vehicle, Traffic congestion.

Abstract. To detect traffic state using Smart-phone equipped vehicles (for short, equipped vehicles) data is popular in recent years. Calculating the average speed of all equipped vehicles simply is no longer reasonable without considering random distribution of equipped vehicles. In this paper, a new method of detecting traffic state based on equipped vehicle data was presented. Inspired by information entropy theory, the key homogeneous distribution degree (KHDD) can be obtained, which contributes to correct the detection error caused by the random distribution of equipped vehicles. In simulations, the results show the validity of the proposed method to detect traffic state under the different KHDD.

Introduction
Tom Tom, which is the traffic navigation service provider of Netherlands, released urban traffic congestion level and the rank of congested cities all over the world on April 4, 2015 [1]. The report indicated the commuters living in Istanbul could waste 125 hours every year in traveling due to traffic congestion. Moreover, Istanbul was appointed the most congested city in 2014. Urban traffic congestion is becoming a tough problem, which restricts the development of society and economy seriously. Thus, congestion detection is viewed as a primary consideration of intelligent transportation systems (ITS) to vehicle navigation, traffic control, and traffic management. Also, it is one of the main application of ITS to help drivers to avoid congested roads through delivering traffic information [2].

New technologies, such as GPS-enabled smart-phone and smart-phone signal triangulation, to register vehicle trajectories have opened an opportunity to collect valuable real-time information on traffic conditions [3]. The sources can provide the related vehicle data, such as position, velocity, acceleration, and so on [4]. Based on the above analysis, this paper presents a method to detect traffic congestion considering equipped vehicles’ random distribution under different penetration rates. The basic data of traveling vehicles in the area of interest need to be gathered and analyzed to evaluate traffic congestion [5]. Xu et al. [6] collected the basic trajectory data of buses to calculate the travel time to identify the congestion state. Some cities like Jinan, China established dedicated lanes for buses, thus the buses cannot really represent traffic state of the relevant road segment. In [7], traffic congestion is classified into two types: Recurrent Congestion (RC) and Non-Recurrent Congestion (NRC). RC is caused by excess travel demand, inadequate traffic capacity or poor signal control [8]; NRC is caused by unexpected events like traffic accidents or vehicle breakdowns. Reference [7] proposed a NRC detection methodology through calculating Link Journey Time estimates (LJTs) to support the accurate detection of NRCs on large urban road networks. Hiribarren et al. [9] collected the trajectory data of equipped vehicle running on the road of interest with the help of smart-phone to evaluate traffic congestion. The authors confirmed equipped vehicles play important roles in detecting congestion, even for very low penetration rates. Because the core of the paper is to study the penetration rates of vehicles, the detection results probably lost the accuracy without considering the random distribution of equipped vehicles on road. We should note that equipped vehicles to detection traffic congestion are indeed a good method at present.
Proposed Method

Speed Entropy

Rudolf Julius Emanuel Clausius, German physicist who initially presented a definition of entropy in 1850. The definition describes the relationship between energy entropy and the degree of uniformity, i.e., the entropy reaches the maximum when the energy in a system distributes uniformly. In information theory, the entropy is a mathematical expectation, which can be represented by calculating the summation after each variable multiplies its corresponding probability. Also, it indicates the occurrence probability for some particular information.

In ITS, the equipped vehicle periodically transmits its own speed to other equipped vehicles or roadside infrastructures. Due to the speed gap to the different position along a road, whether the detected result is accurate which is subjected to the distribution of all equipped vehicles. If we assume that only three equipped vehicles on lane A or lane B in Figure 1, simply calculating all equipped vehicles’ average speed on lane A using general method (Eq. 7) is not logical compared with lane B where equipped vehicles show homogeneous distribution nearly. The result is the worst when all equipped vehicles distribute intensively like the situation of lane A shown in Figure 1. Inspired by the entropy theory, this paper aims to solve the key homogeneous distribution degree (KHDD) through establishing entropy function to acquire an accurate road average speed, and the ‘entropy’ is called Speed Entropy.

Inspired by the entropy theory, this paper aims to solve the key homogeneous distribution degree (KHDD) through establishing entropy function to acquire an accurate road average speed, and the ‘entropy’ is called Speed Entropy.

Figure 1. Random distribution of equipped vehicles on road. It is easy to know that the three equipped vehicles on lane A can not represent the actual traffic condition compared with the ones on lane B.

The equipped vehicle running on road continually changes its position, and correspondingly the behavior naturally induces a kind of distribution correlation with respect to other vehicles. \( k^m \) is the homogeneous distribution degree (HDD) which is a discrete variable, \( k^m \in [0, 1] \). \( k^m = 0 \) when all equipped vehicles distribute uniformly, in contrast, \( k^m = 1 \) means all equipped vehicles distribute intensively. \( \omega^m \) indicates the change rate of distribution for \( m \)th equipped vehicles on road which obeys normal distribution \( W \sim N(\mu, \sigma^2) \). Probability density function and distribution function of \( \omega^m \) is as Eq. 1 follows, respectively.

\[
f(\omega^m) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\omega^m - \mu)^2}{2\sigma^2}}, (-\infty \leq \omega^m \leq +\infty) \tag{1}
\]

Although \( \omega^m \) belongs to \((-\infty, +\infty)\), the probability of concentrated and homogeneous distribution should be extremely low considering the real traffic conditions, i.e., \( P(\alpha) \rightarrow 0, P(\beta) \rightarrow 0 \). \( \alpha, \beta \) is the real number, which presents the change rate for concentrated and homogeneous distribution, respectively. Alternatively, all equipped vehicles mainly show random distribution if \( \omega^m \in (\alpha, \beta) \). For arbitrary domain \((\omega^m_{i-1}, \omega^m_i)\), having

\[
P\{\omega^m_{i-1} < W \leq \omega^m_i\} = \Phi(\frac{\omega^m_i - \mu}{\sigma}) - \Phi(\frac{\omega^m_{i-1} - \mu}{\sigma}) \tag{2}
\]

We adopt standard normal distribution. Equation (2) can be converted into as

\[
P\{\omega^m_{i-1} < W \leq \omega^m_i\} = \Phi(\omega^m_{i-1}) - \Phi(\omega^m_i) \tag{3}
\]

Different \( k^m \) has the respective corresponding probability intervals. \( \omega^m \) is divided into \( n \) groups of probability intervals based on the number of \( k^m \). The corresponding probability intervals can be
constructed based on the probability theory. Based on Boltzmann's H-theorem, KHDD can be solved for the finite \( k^m \), as Eq. 4 follows

\[
H(k^m) = \sum_i k^m_i p(k^m_i)
\]  

(4)

**Calculation of KHDD Based on Random Distribution**

The aim of this subsection is to design a valid \( \omega \) for catering the KHDD. Equipped vehicles divide the road into multiple detected areas. The current traffic situation of each area is almost different. For example, the three equipped vehicles show the homogeneous distribution regardless of how other normal vehicles distribute (one of the directions in section ① of Figure 2). Each equipped vehicles is in charge of recording and transmitting the current speed to other equipped vehicles or road infrastructure. Likewise, the vehicles on road ② also can transmit the speed, however, the accuracy needs to be deliberated before through the KHDD correcting. In addition, the middle distance between the two adjacent equipped vehicles is viewed as the boundary for dividing the detected areas clearly.

![Figure 2](image)

Figure 2. Two different type of distribution of equipped vehicles (the little black rectangle stands for an equipped vehicle and dashed line stands for the actual detected length, and the length of each area is \( l (l') \) in section ②).

Figure 2 shows that the change rate \( \omega_i \) can be designed depending on \( \Delta l \) for the random distribution as (5) follows

\[
\omega_i^1 = \frac{l + \Delta l_1}{l}, \quad \omega_i^2 = \frac{l - \Delta l_1 + \Delta l_2}{l}, \quad \omega_i^3 = \frac{l - \Delta l_2}{l}.
\]  

(5)

where \( \Delta l \) is the length difference between the homogeneous area \( l \) and actual area \( l_i \). \( \omega_i \) changes continually with the vehicles moving. Eq. 5 is the calculation of \( \omega_i \) one time for three equipped vehicles. In addition, \( \alpha=0, \beta=2 \) can be inferred based on Eq. 5. Discrete variable \( k^m \) has the expressions as Eq. 6 follows

\[
k_i^1 = \left| \frac{\Delta l_1}{l} \right|, \quad k_i^2 = \left| \frac{\Delta l_2 - \Delta l_1}{l} \right|, \quad k_i^3 = \left| \frac{\Delta l_3}{l} \right|
\]  

(6)

**Calculation of Road Average Speed**

This study proposes a method using KHDD to improve the accuracy of road average speed. The primary consideration for congestion detection is to calculate the average speed. The general method is as below

\[
\bar{v} = \frac{\sum_{i=1}^{m} \bar{v}_i}{m}
\]  

where \( \bar{v}_m \) is the average speed of equipped vehicle \( m \) at given time. We can easily infer the result loses its accuracy after analyzing Figure 1. If we take Figure 2 (road ②) as an example (three equipped vehicles) and modify (11) using KHDD, then (11) can be transformed as

\[
\bar{v} = \frac{\bar{v}_1[1-H(k^1)] + \bar{v}_2[1-H(k^2)] + \bar{v}_3[1-H(k^3)]}{3}
\]  

(8)
Considering vehicles running regularly, the positive or negative sign before $H(k^m)$ depends upon the length of the detection area. The sign is positive if the length is more than $l$, otherwise is negative. $H(k^m)$ can be obtained according to Eq. 4. In addition, $\bar{v}$, $\bar{r}$ are time mean speed (TMS).

**Equipped Vehicle Trajectory Data**

In this section, the field data are collected using smartphone to validate our proposed method based on the previous assumptions. Every smartphone installs a kind of traveling software called Speed Tracker. This software can record real trajectory data conveniently and transmit it to traffic control center with the help of communication function of smartphone, such as travelling distance, travelling time, average speed, current speed, and so on. To perform the experiment successfully, each of the drivers carries an interphone to communicate each other in real time. The segment of Wenhuaxi Road in Jinan, China, is considered as the testing road after careful consideration. Car A, B, C randomly drives into the appointed area in turn. The cars labeled in Figure 3 take charge of recording their respective trajectory data. It’s time to record the data when the last car (car C) starts to enter the testing road, until the first car (car A) bears off the segment. The three testing cars equipped with smartphones are used to collect trajectory data. Thus, the complete work is done to record as the detailed trajectory data as possible at given time.

![Figure 3. A section of Wenhuaxi Road of Jinan, China.](image)

In addition, $\Delta l_1$, $\Delta l_2$ can be calculated based on the collected 108 sets of typical data. Correspondingly, $\theta_i$ is easy to figure out, and the KHDD can be solved through solving Eq. 4. Table 1 is the KHDD of the selected time segments during two days. The value on the first day is lower than that of on the second day.

<table>
<thead>
<tr>
<th>Time Segments</th>
<th>8:00-9:05</th>
<th>9:25-10:05</th>
<th>10:00-10:35</th>
<th>11:20-11:35</th>
<th>12:00-12:05</th>
<th>8:00-9:05</th>
<th>9:25-10:05</th>
<th>10:00-10:35</th>
<th>11:20-11:35</th>
<th>12:00-12:05</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(k^1)$</td>
<td>0.58</td>
<td>0.71</td>
<td>0.83</td>
<td>0.65</td>
<td>0.55</td>
<td>0.27</td>
<td>0.37</td>
<td>0.25</td>
<td>0.19</td>
<td>0.23</td>
</tr>
<tr>
<td>$H(k^2)$</td>
<td>0.64</td>
<td>0.59</td>
<td>0.53</td>
<td>0.71</td>
<td>0.61</td>
<td>0.33</td>
<td>0.28</td>
<td>0.34</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>$H(k^3)$</td>
<td>0.62</td>
<td>0.67</td>
<td>0.47</td>
<td>0.52</td>
<td>0.76</td>
<td>0.31</td>
<td>0.30</td>
<td>0.18</td>
<td>0.35</td>
<td>0.28</td>
</tr>
</tbody>
</table>

**Discussion**

In this section, the main work is to demonstrate the validity of our proposed method compared with the general method. Under 5% penetration rates, the detection results at three same time segments during two successive days can be calculated using two kinds of methods. Traffic state of Wenhuaxi Road is detected for several given typical time segments during two successive days based on the trajectory data. The value on the first day is lower than that of on the second day. Based on [10], traffic congestion degree of the specified road is divided and contrasted with or without considering the random distribution of equipped vehicles, as shown in Figure 4. Also, the detection results and comparisons between our proposed method and general method are shown. In Figure 4a, the results appear three differences (marked by dashed oval) when the KHDD is lower. However, there has only one difference in Figure 4b when the KHDD is higher. The results with our method are basically identical in Figure 4, which strongly demonstrates the validity of our proposed method,
even for a low penetration rate (5% for 12:00–12:05 am). The comparisons indicate our proposed method is reasonable, and also congestion detection needs to consider the random distribution of equipped vehicles.

Figure 4. Comparison of Traffic state or congestion level on detected result of Wenhuaxi Road. a: lower KHDD with 5% equipped vehicles on the first day. b: higher KHDD with 5% equipped vehicles on the second day.

Summary

The aim of this study is to provide a valid method for detecting traffic congestion on urban roads. This method considers the random distribution of equipped vehicles that influences the detection results. Inspired by information entropy, the concept of Speed Entropy is presented for further calculating the KHDD. This contributes to correct the detection error caused by the random distribution of equipped vehicles. From Figure 4, the comparison results show three differences when the KHDD is lower. In contrast, only one difference appears when the degree is higher. This indicates the validity of our proposed method to detect traffic congestion.

Acknowledgement

This research was financially supported by the Shandong Provincial Natural Science Foundation (ZR2014FM022).

References


