A GPS-BASED ENHANCED APPROACH FOR TRAFFIC INCIDENT AND
CONGESTION DETECTION

Qinglu Ma
Associate Professor
Chongqing Jiaotong University
No. 66 Xuefu Road, Nanan Distr.
Chongqing China 400074
Tel: +86 18602303294; mql@cqu.edu.cn

Kara Kockelman
E.P. School Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu
Tel: 512-471-0210;

ABSTRACT
Evaluating traffic networks is crucial for administration of roadway systems, to better address
congestion, safety, and air quality issues around the globe, however, challenges in implementation
abound, including huge investment, big data mazes and obtaining timely and accurate feedback.
Recently developed low-cost Global Positioning System (GPS) data loggers are a promising tool
for traffic monitoring, thanks to their low cost, ready availability on smartphones, and ability to
simultaneously track many travelers and vehicles, relative to expensive, built-in traffic GPS. GPS
data from many travelers provides real-time estimates of traffic conditions and can improve active
traffic management using various Big Data analytical techniques. This paper demonstrates a
method to couple GPS data with real-time road conditions to estimate Relative Roadway Velocities
in responses to the traffic evaluation of the improved estimation model of Average Roadway
Velocities. Through analyzing real-time traffic surveillance software with high data coupling and
concurrent processing, a new data coupling analysis method for real-time traffic evaluation is
proposed. Experimental results show efficient coupling of all available GPS data with road
condition could improve the accuracy of the estimation of traffic state. This proposed method could
potentially increase the estimation accuracy from more than ten meters to one meter. Over 98% of
GPS data can be successfully matched to service routes when the low-cost GPS devices were used
to detect real-time traffic conditions. The results of traffic network evaluation could well serve as
a driving assistant for connected and autonomous vehicles (C/AVs) and other traffic operations.

Keywords: Traffic Conditions, Global Positioning Systems (GPS), Roadway Levels of Service,
Relative Velocity, Data Processing

INTRODUCTION
As an essential technology in Intelligent Transportation Systems (ITS), the real-time, accurate
traffic state estimation serves as a foundation of most ITS applications and attracts the
extraordinary attention of an increasing number of researchers (Elkafoury, 2015). Presently, one
of the most effective methods of traffic state estimation is to collect and analyze traffic data using
GPS-equipped vehicles (e.g., taxi, bus and shared vehicle) because of their mobility, wide coverage and easy data access. Many technology companies (such as Google and Baidu) have even integrated real-time traffic information with their mapping services. However, due to the constraint in size, power, and cost of GPS receivers, the investigation of efficient estimation algorithms which satisfy the basic accuracy requirement for low-cost GPS receivers meet new challenges. (Islam and Kim, 2014)

Generally, GPS data can provide a series of track-points at certain time intervals (Khan and Kockelman 2012). Each GPS track-point is influenced by various external factors, like weather and road conditions, which affect data accuracy. Errors in longitude and latitude coordinates have been found to lie between 10 to 15 meters in 95% of low-cost GPS readings (Zandbergen 2011).

For urban road networks, researchers have utilized various traffic state estimation methods. (Jia et al. 2011) estimated the urban traffic state by considering resident travel characteristics and road network capacity. (Lovisari 2016) proposed a kernel-based density/flow estimation method based on probe-vehicle data. (Kong et al. 2013) applied the curve-fitting method into the fused estimation of urban traffic states based on the research results of (Shi et al. 2008). In addition, (Ndoye et al. 2011) proposed a recursive multiscale correlation-averaging algorithm to combine the measures from GPS probe vehicles to produce higher quality traffic state data. (Zhao et al. 2011) improved Chen et al.’s (2007) method and verified the improved method through real traffic data.

To sum up, the existing methods are usually intended to estimate the state of traffic’s mean velocity, and for a large-scale urban road network, such as the center region of Beijing (around 10000 links), they need to account for the traffic capacity of different road grades to increase estimation accuracy. Therefore, a relative velocity mode for traffic state estimation is developed based on the above research and road conditions.

METHODOLOGY

This study describes a systematic approach to evaluating site-specific traffic conditions across networks, by relying on low-cost, real-time, and spatially detailed GPS sources. As traffic becomes less stable (shifting from relatively free-flow conditions to critical congestion levels), the approach identifies anomalous traffic conditions by appraising the likelihood of various traffic attributes. Warning information can be disseminated via roadside equipment (e.g., VMS) or in-vehicle systems (like car radios). Figure 1 shows the approach’s operation algorithm.
**Point Matching**

Traditional map matching methods locate points and map coordinates between checkpoints along designated routes (Chen, 2014). In this paper, a matching module can transfer trip data (message points) into the electronic map, whereas GPS may be prone to disturbances and inaccuracies. Positioning errors could be significantly reduced from more than ten metres to one metre by the above algorithm, which has contributed to solving the problem of map matching with low-cost GPS data. Both direction and distance to the nearest road and direction are used for more accurate map matching. The Pointwise Convergence approach considers a nearby area, which is centered at each GPS point, where all road segments in the area are regarded as candidates for matching. Then, the projection distance is calculated, along with the crossing angle between the road and any vehicle’s travel direction. Due to errors embedded in both GPS positioning and map data, GPS points may not appear on roads in the network map. So road segment buffers are created to catch individual deviations. To address some uncertain points, such as those at intersections, a vehicle-tracking method based on the heuristic search algorithm is employed here. As a result, the final match is determined based on shortest projection distance while taking a reasonable intersection angle with the vehicle’s heading direction.

FIGURE 1 Approach framework for traffic network evaluation.
Figure 2 presents the results of this map matching method. The black points grouped onto the road segments’ buffers represent their vehicles’ corresponding positions. The figure suggests a systematic location bias in these GPS data, which is common in satellite data and is a reason engineers often use inertial movement units and other devices to provide devices with redundancies.

Relative Velocity Estimation Methods

The typical traffic estimation method involves estimating the average speed of vehicles or travelers’ GPS units on every road link. This paper analyzes on the existing estimation model, and advances the relative velocity model for improving the current mean-speed model of the traffic state estimation. The derivation and realization of improved model are described in the following subsections.

Firstly, consider an object moving along path $P$ and traveling the distance $L_i$ during the time interval $\Delta t_i = t_{i+1} - t_i$, then, the actual distance $L_i$ of the object runs along the path $p$ is defined as,

$$L_i = \int_{t_i}^{t_{i+1}} v(i, t) dt \approx v(0, t) \left( \frac{t_i - t_0}{2} \right) + \sum_{i=1}^{n-1} v(i, t) \left( \frac{t_{i+1} - t_i}{2} \right) + v(p, t) \left( \frac{t_n - t_{n-1}}{2} \right)$$

where $t_i, (i=0,1,2,\ldots,n)$ and $v(i, t)$ are a series of times and speeds, which are collected by running the vehicle on the road segment $L_i$.

In the case of real data, the average travel time is computed by equation 2,

$$t_s = \int_{x_i}^{x_{i+1}} \frac{1}{v(x, t)} dx \approx \frac{1}{v(x, 0)} \left( \frac{x_i - x_0}{2} \right) + \sum_{j=1}^{n-1} \frac{1}{v(x, j)} \left( \frac{x_{j+1} - x_j}{2} \right) + \frac{1}{v(x, n)} \left( \frac{x_n - x_{n-1}}{2} \right)$$

where $x_j (j=0,1,2,\ldots,n)$ and $v(x, t)$ are series in the length and spot speeds, which are collected by running the vehicle on the road segment $L_i$.

In order to simplify the calculation process and make the method more useful in practice, each discrete instantaneous speed sample $v_i^x$ contains measurement error $v_i^e$. Hence true speed $v_i$ is $v_i = v_i^x - v_i^e$. For a discrete series of speed samples $v_i^x$ acquired at $N$ uniform time intervals over the time $T$, the integral can be approximated in equation 2 by a sum, where the expression for the unknown average speed becomes,

$$\bar{v} = \frac{1}{N} \sum_{i=1}^{N} v_i^x - \frac{1}{N} \sum_{i=1}^{N} v_i^e = \bar{v}_s - \bar{v}_e$$

where $\bar{v}_s$ is the exact average of all measured GPS samples and $\bar{v}_e$ is the measurement error.
Secondly, speed time series are smoothed in order to reduce noise in speed measurements, which is typical for traffic in urban networks. Li (2009) proposed an algorithm that estimates this speed based on historical bus travel speed along the route segment and the current travel speed of the bus derived from GPS data,

$$\hat{v}(x, t) = \frac{\sum_{i=1}^{N-1} v_a(x, t) + v_r(x, t)}{2N}$$

(4)

where $\hat{v}(x, t)$ is the predicted speed of route segment $x$ at $t$ time (to the immediate downstream route segments), $v_a(x, t)$ is the historical average speed of route segment $x$ ($x = 1, 2, \ldots, N - 1$), $v_r(x, t)$ is the bus’ current speed (as obtained from the GPS data), and $N$ is the number of route segments before reaching the station of interest.

In this algorithm, the predicted speed $\hat{v}$ will depend primarily on its historical average speed along the route $v_{ai}$ ($i = 1, \ldots, n - 1$) rather than its current speed $v_r$. In real operating conditions, however, the current speed of the bus is usually a more important factor influencing how fast the bus will travel over the distance to the station of interest. Furthermore, this method would predict non-zero speeds even when $v_r = 0$. This model could therefore perform poorly in applications where the bus speed changes frequently, the GPS speed is not sufficiently accurate or the historical average speed of the route does not adequately reflect the current speed of the traffic.

Sun et al. (2007) presented an improved method based on equation 3, in which the average speed can be updated dynamically.

$$v(x, t + \tau) = \frac{L_{xe}v(x, t) + L_{xb}v(x, t - \tau)}{L_{xe} + L_{xb}}$$

(5)

where $L_{xe}$ is the distance from the current bus location to the end of the route segment $x$, and $L_{xb}$ is the distance from the beginning of the route segment $x$ to the current bus location.

Finally, it is more reasonable to develop the estimation model based on the average speeds of route segments rather than the historical average velocity. This leads to a superior result, as this model only estimates the speed to the end of route segment instead of the speed to downstream bus stations, when the bus is far from the station.

To take into account the uncertain nature of traffic conditions, this paper proposes a new scheme for predicting travel time to a downstream stop. The proposed method first estimates the road factor to the immediate downstream route segments, denoted by $i$. The adjustment factors’ estimate is updated dynamically as follows:

$$v(x, t + \tau) = pv(x, t) + (1 - p)v(x, t - \tau) - v(x, t)\tan(w_x)$$

(6)

where $p$ is the adjustment factor used ($0 \leq p \leq 1$), $w_x$ is the road surface grade at $x$, and $v(x, t)\tan(w_x)$ is relative velocity of road grade. This equation aims to simulate traffic status: the steeper it is, the slower the relative velocity.

Then, based on Eqns. 3, 4 and 5, let $A = pL_{xe}, B = (1 - p)L_{ef},$ then, average speed estimation is updated dynamically as follows:
\[
\begin{align*}
  v(x, t + \tau) &= \frac{A v(x, t) + B v(x, t - \tau)}{A + B} - v(x, t) \tan(w_x) \\
  A &= p L_{se} \\
  B &= (1 - p) L_{sf}
\end{align*}
\] (7)

Thus, if \( p=0.5 \), Equation 7 simplifies to become Equation 5, as shared by Sun et al. (2007).

Similarly, when \( p=1 \), \( v(x, t + \tau) \) it is only related with current data, whereas when \( p=0 \), \( v(x, t + \tau) \) it is only related with historical data. Generally, \( v(x, t + \tau) \) will depend on the fusion of current and historical data.

If a road network is regarded as a connected graph \( G \), then there are any two road segments (i.e., segment \( i \) and \( j \)) where the traffic service level should be the same, even if it differs from the road’s average speeds. In order to approximate the actual traffic situation more effectively, this paper introduces an evaluation index to improve the accuracy of the estimation. The evaluation index provides a means to estimate whether or not congestion is over or under its free flow traffic speed. The traffic condition is evaluated in equation 8.

\[
\frac{v(t, i)}{v_f^i} \geq \frac{v(t, j)}{v_f^j} = r(t)
\] (8)

where \( v(t, i) \) and \( v(t, j) \) are the average speeds of the road segment \( i \) and \( j \) in time \( t \) respectively, \( v_f^i \) and \( v_f^j \) are the speed of the free flow traffic state of the road segment \( i \) and \( j \) respectively.

As expected, this equation is better at distinguishing the traffic status when the road’s average speeds are same. That is to say, the traffic condition of road segment \( i \) is better than that of \( j \) when \( r(t, i) \geq r(t, j) \) even if the road’s average speeds are same. \( r(t, i) \) and \( r(t, j) \) are the evaluation indices of traffic condition of the road segment \( i \) and \( j \) respectively. Similarly, the evaluation index for traffic conditions across the whole road network is computed:

\[
R(i) \approx \frac{\sum_{i=1}^{n} v(t, i)}{n} \left/ \frac{\sum_{j=1}^{n} v_f^j}{n} \right\} \approx \frac{\sum_{i=1}^{n} v(t, i)}{\sum_{j=1}^{n} v_f^j}
\] (9)

where \( n \) is the number of road segments in the network.

23

Quantification of Traffic Condition

A quantitative and discrete evaluating model is proposed to estimate the traffic condition by utilizing the data from low-cost GPS receivers. The model appears most effective in collecting traffic information from the heavy traffic in CBD area, especially during peak hours. At the end of each time interval, the module produces a score that can be connected to the possible presence of traffic slow down. The score can be easily mapped into an indicator of the real traffic situation, where different traffic data correspond to different score ranges, resulting in four levels of traffic conditions:

I. Free-flow

Qinglu Ma, Kara Kockelman
In general, the evaluation index is more than 0.67; there are few vehicles, typically night traffic. The color green is used to map in a road section that has a good weather and road condition, relatively. In this case, it can freely travel without any restrictions.

II. Light traffic (normal condition)

The evaluation index ranges from 0.34 to 0.67 and the color yellow is used to map a road segment. The traffic is flowing smoothly here.

III. Heavy traffic

The evaluation index is in the range 0.17 to 0.34 and the color orange is used to map a road segment. Traffic moves slowly here and congestions may occur.

IV. Stationary traffic

The evaluation index is less than 0.17 and the color red is used to map a road segment. In this case, severe congestions happen in the network and traffic is almost at a standstill.

Traffic condition could be automatically determined with a comprehensive quantitative evaluation. In this study, the influence of the evaluation period on the real-time character and availability of the model was analyzed and tested with a case study.

Road Section Coloring

Considering the computation speed, the discrete evaluation method is presented to approximate the sample set on each road link. And then, according to the results of discrete evaluation, the different traffic service levels are displayed in real-time on the E-map represented by corresponding colors based on congestion level. The results of region coloring are shown in Figure 3.

FIGURE 3 Tangential path areas and region coloring.

CASE STUDY

Data Processing

The case study used GPS data from Chongqing’s Municipal Committee of Communications. The data were collected from approximately 2000 vehicles traveling via the urban test road network (as shown in Fig.4) every day, each with a unique vehicle ID. The GPS recording frequency is 0.1Hz, i.e., one data point per 10 seconds, which satisfies the purpose of the case study. The data application platform receives GPS data along the traveled routes by matching the related trajectories to the road/street network in order to estimate average speeds, then computes the evaluation index and disseminates them through the traveler information systems. The evaluation
index is presented in 4 categories (as mentioned above: free-flow, light traffic, heavy traffic, stationary traffic), updated every three minutes throughout the day.

Route linearization provides a traffic network model that is required in the later steps of map matching and distance calculation. In a situation where a large transit bus drives through a location multiple times, the road link number needs to be used in order to determine the exact status of the bus. Figure 4 shows the result of route-linearization.

![Sectionalized road and flyover](image)

**FIGURE 4. Sectionalized straight road layer of map**

In order to approximate the real traffic situation, real-time road conditions were surveyed in field. The road used in performance testing is in the same location as the above test field, paved with asphalt. Road condition is represented in Table 1.

**TABLE 1 Data Measured by the Vehicles on the Test Road Network**

<table>
<thead>
<tr>
<th>Road segment</th>
<th>Grade</th>
<th>Number of Lane</th>
<th>Estimated Velocity (km/h)</th>
<th>Adjustable speed (km/h)</th>
<th>Free flow speed (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jianbei branch junction</td>
<td>1</td>
<td>4</td>
<td>15</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>Jianxin west junction</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Jianxin south junction</td>
<td>2</td>
<td>4</td>
<td>18</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>Jianxin east junction</td>
<td>2</td>
<td>4</td>
<td>15</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Jianbei branch junction</td>
<td>0</td>
<td>4</td>
<td>20,15</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>Farmer’s market(traffic signals)</td>
<td>0</td>
<td>4</td>
<td>10,25</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>Three-points intersection</td>
<td>-2</td>
<td>4</td>
<td>8,15</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>Jianxin north junction (traffic signals)</td>
<td>-2</td>
<td>4</td>
<td>10</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>Loop road north junction</td>
<td>2</td>
<td>4</td>
<td>15</td>
<td>15</td>
<td>40</td>
</tr>
</tbody>
</table>
Map matching can be a problem of statistical estimation. In this experiment, a sequence of GPS points in the buffer (a rectangular region and its centerline along the road segment in the digital map) are explored.

Table 2 shows how important the buffer is. The road segments L1-L18 were recorded. The objective is to fit GPS points to a buffer, but the buffer is constrained to the network. In this case, the buffer size can be adjusted to match the actual road condition.

<table>
<thead>
<tr>
<th>Road segment</th>
<th>Section Label</th>
<th>Buffer diameter (/m)</th>
<th>Diameter Revise (m)</th>
<th>Road width (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hongshi East Road</td>
<td>1</td>
<td>26.55</td>
<td>27</td>
<td>11.55</td>
</tr>
<tr>
<td>Hongshi West Road 2</td>
<td>2</td>
<td>26.05</td>
<td>26.05</td>
<td>11.05</td>
</tr>
<tr>
<td>Huatang Road</td>
<td>3</td>
<td>30</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Jianbei Branch Junction</td>
<td>4</td>
<td>29.93</td>
<td>30</td>
<td>14.93</td>
</tr>
<tr>
<td>Guanyinqiao west Road</td>
<td>5</td>
<td>32.32</td>
<td>32.32</td>
<td>17.32</td>
</tr>
<tr>
<td>Jianxin West Road</td>
<td>6</td>
<td>32.06</td>
<td>32.06</td>
<td>17.06</td>
</tr>
<tr>
<td>Guanyinqiao East Road</td>
<td>7</td>
<td>37.33</td>
<td>37.33</td>
<td>22.33</td>
</tr>
<tr>
<td>Yu’ao Bridge</td>
<td>8</td>
<td>31.68</td>
<td>32</td>
<td>16.68</td>
</tr>
<tr>
<td>Beichengtian South Street</td>
<td>9</td>
<td>29.73</td>
<td>30</td>
<td>14.73</td>
</tr>
<tr>
<td>Beichengtian North Street</td>
<td>10</td>
<td>24.75</td>
<td>25</td>
<td>9.75</td>
</tr>
<tr>
<td>Huaxin Street</td>
<td>11</td>
<td>42.62</td>
<td>43</td>
<td>27.62</td>
</tr>
<tr>
<td>Hongshi Road</td>
<td>16</td>
<td>47.23</td>
<td>47.23</td>
<td>32.23</td>
</tr>
<tr>
<td>Hongjin avenue</td>
<td>17</td>
<td>63.34</td>
<td>63.34</td>
<td>48.34</td>
</tr>
<tr>
<td>Guanyinqiao East Road</td>
<td>18</td>
<td>34.62</td>
<td>35</td>
<td>19.62</td>
</tr>
</tbody>
</table>

GPS data cannot be directly used by the traditional traffic flow models, due to the strong discreteness and randomness, as they only provide individual measures, such as vehicle ID, position coordinates, time, velocity, moving direction, instead of explicit consequences of traffic conditions. These measures’ results can be used to estimate the spatiotemporal mean speed, which is defined as the average speed of vehicles on a specific road link during a period of time. If the road link is a two-way structure, the heading directions must be identified before the estimation.

**Identification Results**

The road’s average and free-flow speeds are used in estimating the evaluation index to identify the traffic condition. Figure 5 shows the identification results of traffic congestion by evaluation index for a specific road segment.
The performance of the proposed approach was measured in the urban traffic state surveillance system. Figure 6 illustrates final outputs of the system in these experiments. The color of the road sections are shown in the legend, with: road links in red, orange, yellow, green or black representing the following traffic states: “Free-flow,” “Light traffic,” “Heavy traffic,” “Stationary traffic” or “Lack of data,” respectively.

To test this study’s implementation ability, the trend of traffic condition is recorded in three special road segments within a day in aforementioned experiments. Simultaneously, traffic surveillance video was used to serve the real traffic conditions among the interested road segments. Traffic
conditions were determined according to vehicle density and speeds shown in the video. Figure 7 shows the results comparing the estimated value and the true value of traffic conditions, which are calculated per 5 mins. It is evident that the estimation of the road’s average speeds is efficient in identifying traffic conditions.

**FIGURE 7 Comparison of traffic conditions as estimated versus actual human evaluation, over time.** By comparison of these estimated results with real-time video stream data, both evaluated methods can fit for most real traffic state evaluation, but it is also shown from the results that the difference of the road conditions has a great effect on the evaluation. The video showed that the roadworks are in progress in a time series from 110 to 180, although all speed has slowed down, the overall traffic here remains relatively smooth. Therefore the estimation of relative speed is closer to practice by comparison with the estimation of average speed.

Lastly, this research evaluated the performance of the estimation method of relative speed. Two segments of the tested road were selected to estimate the traffic state at different times, and the results are summarized in Table 3.

| Table 3 Performance of estimation method based on the relative velocity |
|--------------------------------------------------|----------------|----------------|----------------|----------------|
| Time series (/5mins) | 6:00-10:00 | 10:00-16:00 | 16:00-20:00 | 20:00-00:00 |
| Number of samples     | 2,362     | 1,836       | 2,226        | 1,032         |
| Maximal error(s)      | 372       | 220         | 304          | 340           |
| Minimal error(s)      | 0         | 0           | 0            | 0             |
| MAPE(%)               | 11.97     | 12.15       | 12.03        | 12.10         |

The estimated method was perfect in detecting traffic state and had a success rate of over 98% in matching the service route. The small proportion of route matching errors is probably due to the errors in GPS location data and inaccuracy of the transit network model. These matching errors can be reduced by using higher quality GPS devices and more accurate digital street maps. In addition, after comparing and contrasting the results of experiments and the results of artificial judgments, the experimental results indicate that the estimation value aligns closely with the actual traffic status despite divergences at specific points. These differences could be due to the fact that only shorter periods were considered, for every day, in which case the results would be more precise.
CONCLUSIONS

In this paper a relative velocity estimation model is proposed by synthesizing the average relative velocity estimation model and road condition. This improved approach makes road-difference-estimation more accuracy, subjectivity less and easy to realize intelligent evaluation. The results show that the estimated model presented could not only increase the accuracy of estimation, but also reduce the cost of operation and ensure economic benefit. Further work may include the assessment of the estimation error by comparing the obtained mean speed results to state of the art methods.

Based on the above, the improved model provides a more appropriate way for traffic state estimation, and the relative travel speeds are found to closely track real traffic conditions. The results of the case study presented could well serve as guidelines for the design and analysis of traffic circles under mixed traffic conditions of developing cities, and finally offers a useful tool for the policy makers.

ACKNOWLEDGMENTS

This research is supported by the China Scholarship Council (CSC) and the China Postdoctoral Science Foundation (2016M592645). The authors would like to thank the Chongqing Municipal Transportation Information Centre for providing real traffic data used in this paper, Dr. Chris Claudel for his constructive comments, and Scott Schauer-West for editing and administrative support.

REFERENCES


Qinglu Ma, Kara Kockelman


