ABSTRACT. Traffic congestion visualization is an important part in traffic information service. However, the real-time data is difficult to obtain and its analysis method is not accurate, so the reliability of congestion state visualization is low. This paper proposes a visualization analysis algorithm of traffic congestion based on Floating Car Data (FCD), which utilizes the FCD to estimate and display dynamic traffic state on the electronic map. Firstly, an improved map matching method is put forward to match rapidly the FCD with road sections, which includes two steps of coarse and precise matching. Then, the traffic speed is estimated and classified to display different traffic states. Eventually, multi-group experiments have been conducted based on more than 8000 taxies in Xian. The experimental results show that FCD can be matched accurately with the selected road sections which accuracy can reach up to 96%, and the estimated traffic real-time state can achieve 94% in terms of reliability. So this visualization analysis algorithm can display accurately road traffic state in real time.

1. Introduction. Traffic congestion in urban road is a very serious problem that most cities are facing with it, which has great influences on people’s life, property and air quality. A convenient and accurate real-time traffic information display method is very necessary for efficient traffic.

To monitor traffic conditions or collect traffic volume data, several traditional methods such as remote sensing are widely used. The existing collection methods of road traffic flow mainly adopt the fixed-point modes including induction loop, microwave, radar and video techniques, which have some drawbacks such as limited coverage region, high cost, and difficult to maintenance [4]. FCD (Floating Car Data) has the advantages of easy maintenance, low cost, comprehensive information, so FCD has received significant attention in the past decades [10]. Guo used the FCD to get the real time traffic flow [6]. Yuan designed a scheme to detect the traffic congestion and dissemination based on FCD [16]. However, they usually focused on the estimation of link travel time, and link travel speed, and seldom paid attention to the visualization of traffic congestion.
Table 1. THE FORMAT OF ORIGINAL DATA.

| Number  | Time         | Longitude  | Latitude  | Speed | Direction |
|---------|--------------|------------|-----------|-------|-----------|
| S A **** | 2015/3/14 00:13:31 | 108.995444 | 34.262183 | 76    | 5         |

Map matching algorithms integrate positioning data with spatial road network data to identify the correct link on which vehicle is moving and determine the location of a vehicle on a link. Inaccurate and low-frequency positioning data has two challenges in the study of Map-matching. Bierlaire developed a post-processing map-matching algorithm for GPS-enabled smartphones only for driving [1]. Position, speed and heading data from GPS are used to assign a probability to each candidate path based on horizontal accuracy of GPS data and the road network. Miwa [11] and Lou [9] developed a similar algorithm specifically for vehicles with lower positioning data frequency and only position coordinates and timestamp. Sparse positioning data, with gaps up to 1.5, 3 and 5 min with latitude, longitude and timestamp, were used. Chen proposed FCD map-matching algorithm based on local path searching [4]. The information of the previous matched GPS point is utilized to reduce the search space significantly. Chen et al. proposed a map-matching algorithm for large-scale low-frequency floating car data [3]. They used a multi-criteria dynamic programming technique to minimize the number of candidate routes for each GPS point. Lou et al. [9] considered two important points in their post-processing algorithm: (a) true paths tend to be straight, rather than roundabout and (b) true paths tend to follow posted speed limits on roads. Yang et al.[15] proposed to project GPS points within 20 m from intersection nodes onto the intersection node itself, postponing resolution of the problem to the next point. Quddus [12] developed an algorithm which was used together with the outputs of an extended Kalman filter formulation for the integration of GPS, dead reckoning data, and a spatial digital database of the road network, to provide continuous, accurate and reliable vehicle location on a given road segment.

This paper proposes a dynamic traffic visualization algorithm based on FCD. The proposed algorithm is composed with map matching, velocity estimation and data visualization. It has fast speed, high reliability, and more conducive to meet the people’s daily travel traffic demand. The remaining part of this paper is organized as follows: the improved map matching algorithm based on a weighted topology and trajectory shape is introduced in the next section; the section 3 puts forward a compensation method of the road speed estimating; the section 4 describes a traffic congestion data visualization method; some experiments using real data are conducted and analyses are shown in section 5. Finally, the conclusions are summarized in the last section.

2. Map matching.

2.1. Data pre-processing. In this paper, the FCD data is acquired from GPS of 8000 Xi’an taxis. The data sampling interval is 30 seconds. The original format of received data includes six kinds of information, which are respectively the plate number, time, longitude, latitude, speed and direction, as shown in Table 1.

FCD data cannot be directly used for the actual ITS, which reasons are shown as follows:
GPS coordinate errors: the prevailing error, 3-15 m, exists in most of the commercial GPS, which causes the vehicles’ GPS coordinates and road position deviation. The attribute information is insufficient: GPS data has a latitude and longitude, but we need to associate them and specific network number in the dynamic navigation.

Because of equipment quality and transmission, there are errors in the original data. If the map matching method is directly used for them, it has a great influence for efficiency of the algorithm, so the study cleans original data by preprocessing, which main idea is to eliminate data that speed is less than 0 km/h or greater than 120 km/h, latitude and longitude are out of range, and driving direction is abnormal.

2.2. Map matching algorithm.

2.2.1. Coarse matching. There are four weights in this map matching algorithm: heading, proximity, link connectivity and turn restriction\[14\]. We can get the heading weight from the vehicle movement direction and link direction [5]. It is calculated by (1).

\[ W_h = H_w f(\theta) \] (1)

Where, we define the \( f(\theta) \) as \( f(\theta) = |\cos(\theta)| \) [13].

The weight for proximity is based on the perpendicular distance (\( D \)) from the GPS point to the link. Generally speaking, the error in a GPS measurement can be reasonably described as a normal distribution \( N(\mu, \sigma) \), so the \( f(D) \) is define as

\[ f(D) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(d_j-u)^2}{2\sigma^2}} \] [7], where \( d_j \) is the distance between the correct point to the candidate point. We can use a zero-mean normal distribution with a standard deviation of 20 m based on empirical evaluation.

\[ W_p = D_w f(D) \] (2)

When the vehicle is at a junction, the link connectivity and turn restriction will be considered necessarily [14]. They can be defined by

\[ W_c = C_w X \] (3)

\[ W_t = T_w Y \] (4)

Where, \( Y = \{1, -1\} \), \( X = \{1, -1\} \).

Setting the \( TWS \) as total weight score, it is the sum of four weights for a link at a junction, as given in (5).

\[ TWS = H_w f(\theta) + D_w f(D) + C_w X + T_w Y \] (5)

Where, \( H_w \), \( D_w \), \( C_w \) and \( T_w \) are the weight coefficients of heading, proximity, link connectivity and turn restriction respectively. These coefficients represent the relative importance of different factors in calculating \( TWS \).

According to this method above, the candidate point set of a FCD can be obtained. If the difference between the maximum probability and the second of candidate points is less than a threshold, the first two candidate points are retained; on the contrary, we only get the one with maximum probability.
2.2.2. Precise matching. The possibility described in Fig. 1 may be produced after the coarse matching above, which four points are matched to the wrong road section, namely the 5th, 6th, 7th and 8th points. In order to further improve accuracy and reliability of the matching results, this paper makes full use of the result of the coarse matching, and determines the unique position estimated point according to the overall moving trends within a certain period of time. The method connects the FCD to obtain a trajectory, and then it is matched with the line produced by the estimated point on the real road to make sure the movement trajectory unique. That is, the position estimated point is the unique existence.

Assuming that vehicle location data \( D_1, D_2, \cdots, D_k \) are obtained by GPS in \( t = 1, 2, \cdots, k \) and \( k \) is less than the update time of road conditions, the corresponding candidate points \( E_1, E_2, \cdots, E_k \) in road network are obtained by the coarse matching algorithm. Each candidate point of the FCD may not be unique, so it can be optimized as follows:

Step1. All candidate sections corresponding to the set \( E = \{ E_1, E_2, \cdots, E_k \} \) are found to build a road network structure, as shown in Fig. 1. In the figure, the arrow refers to the direction of a passable road; the blue circle is original location data collected by floating car; the red circle represents candidate point after matching road section, in which the 5th, 6th, 7th and 8th points have two red points; the green line is the candidate road section; the yellow line is the line \( L \) to be matched that is got by connecting \( D_i, i = 1, 2, \cdots, 9 \) in turn.

![Figure 1. The road network structure.](image)

Step2. On the basis of assuming the correct initial and final positions, according to the candidate points, all candidate routes can be determined, such as \( L_1=1-2-3-5-7 \), \( L_2=1-2-4-6-7 \). If the candidate route is only one, it is considered directly to be the best matching route.

Step3. \( L_1, L_2 \) and \( L \) are discretized by using the interval \( d \), and Freeman chain (8 directions) is applied to code those discrete points, respectively obtaining feature vectors \( C_1, C_2 \) and \( C_L \).

Step4. The feature vector \( C_L \) is matched to \( C_1 \) and \( C_2 \) in turn, to find the best matching route \( L_{\max} = L_1 \), so it is considered to the movement track of vehicle
in this time. The similarity measurement standard in this paper is the Hausdorff distance (Max-min Distance), which means the dissimilar degree between two point sets. Setting $C_L = \{c_{L1}, c_{L2}, \ldots, c_{LP}\}$ and $C_k = \{c_{k1}, c_{k2}, \ldots, c_{kq}\}$, the Hausdorff distance $H$ between $C_L$ and $C_k$ is defined by

$$H(C_L, C_k) = \max[h(C_L, C_k), h(C_k, C_L)]$$ (6)

Where,

$$h(C_L, C_k) = \max_{c_L \in C_L} \min_{c_k \in C_k} \|c_L - c_k\|$$

is the directed distance from $C_L$ to $C_k$; \(h(C_k, C_L) = \max_{c_k \in C_k} \min_{c_L \in C_L} \|c_k - c_L\|\) is the directed distance from $C_k$ to $C_L$; \(\|\cdot\|\) is defined as the distance norm in points sets of $C_L$ and $C_k$.

**Step 5.** Candidate points without the best matching route $L_{max}$ are eliminated to obtain the accurate and unique result.

Through the above method, error candidate points can be found and eliminated, to get more accurate result. It is of great significance for real-time traffic state estimation.

3. **Traffic speed estimation.** Assume the GPS data sequence covering a section of the road as $P_0 \rightarrow P_k$, the positioning time sequence as $t_0 \rightarrow t_k$, the instantaneous speed as $v_0 \rightarrow v_k$, the distance can be calculated by (7) during the time $t_0 \rightarrow t_k$ [15][16].

$$d = \int_{t_0}^{t_k} v dt \approx v_0 \left( \frac{t_k-t_0}{2} \right) + \sum_{i=1}^{k-1} \left( \frac{t_{i+1}-t_i}{2} \right) + v_k \left( \frac{t_k-t_{k-1}}{2} \right)$$ (7)

Where, $i = 1, 2, \ldots, k-1, k$.

With the fixed sampling interval GPS data, equation (7) can be simplified as:

$$d = t_c \left( \frac{v_0}{2} + \sum_{j=1}^{k-1} v_j + \frac{v_k}{2} \right)$$ (8)

The vehicle speed in the road is

$$v_1 = \frac{d}{t_k-t_0}$$ (9)

The current road speed is estimated with the instantaneous velocity and average velocity, which error can be reduced effectively. So $V$ is calculated by

$$V = b_1 v_1 + b_2 v_2$$ (10)

Where, $b_1$ and $b_2$ are correlation coefficients of instantaneous velocity and average velocity respectively. We set $b_1=0.6$, $b_2=0.4$ according to the experimental data.

4. **Visualization analysis.** In order to make the visualization result more hierarchical and distinguishable, we need to conduct pseudo color processing on result. In practical application, HIS transform, rainbow coding and hot metal coding are the main methods. We prefer to use the HIS transform which is easy to realize and has better visual effect.

4.1. **HIS transform.** The HIS space is a frequently-used perceptual color space. The symbol $H$ represents the hue, which is used to distinguish different color features. The symbol of $I$ represents the intensity of image. The saturation is represented by the symbol $S$, which reflects the concentration of color [2].

Firstly, the value of intensity $I$, hue $H$, saturation $S$ can be gotten through brightness transform, hue transform and saturation transform. Then the value of
RGB is obtained by HIS transform to generate a new image. The procedure of processing is shown in Fig. 2.

![Figure 2. The procedure of HIS transform.](image)

The transform relationship between the HIS space and RGB space is expressed in (11)-(13).

\[
\begin{bmatrix}
  R \\
  G \\
  B
\end{bmatrix} =
\begin{bmatrix}
  1 & -0.204124 & 0.612372 \\
  1 & -0.204124 & -0.612372 \\
  1 & 0.408248 & 0
\end{bmatrix}
\begin{bmatrix}
  I \\
  P_1 \\
  P_2
\end{bmatrix}
\]

(11)

\[
P_1 = S \times \cos(H)
\]

(12)

\[
P_2 = S \times \sin(H)
\]

(13)

4.2. Visualization procedure. The Fig. 3 shows the various results got by different pseudo color processing methods [7][19]. The results of rainbow coding and hot metal coding have too many colors, which are not needed. But the result of HIS transform has enough kinds of color, and all of them can be used for congestion visualization. Hence, in order to avoid the loss or redundancy of information, HIS transform is used to process pseudo color, and five kinds of color are chosen in our paper.

Visualization procedure (seeing Fig. 4) is described specifically as follows:

**Step1.** Divide the city map into 1024*1024 images.

**Step2.** Initialize speed value of every area.

**Step3.** According to the update frequency (5 min), data during this period can be acquired.

**Step4.** Based on this divided result, we estimate speed of every sub-area using section 3.

**Step5.** The speed value of every sub-area can be converted as gray value with [0, 255].

As our research object is city roads, we set that the up line of speed is 80 km/h and the low line is 0 km/h. Equation (14) shows the conversion process.

\[
g = \frac{v}{v_{up} - v_{low}} \times (g_{up} - g_{low})
\]

(14)
Where, $g$, $g_{up}$ and $g_{low}$ represent respectively the gray value, the up line and the low line of the gray value. $v$, $v_{up}$ and $v_{low}$ are respectively the road speed, the up line and the low line of the road speed.

**Step6.** Using the HIS transform, we can get the color image by combining the RGB result of every sub-area.

**Step7.** Add the traffic congestion layer in the map to display for users.

**Step8.** Update the map every five minutes.

![Figure 3. Results of different methods.](image)

![Figure 4. Visualization procedure.](image)

5. **Experimental results and analysis.** The program of this algorithm uses mainly the PostGIS, PostgreSQL database and OpenLayers library. Java is applied to realize the map operating and algorithm. In order to verify the accuracy of the proposed algorithm, the experiment is conducted by the GPS data of 8000 taxis from March 1, 2015 to April 1, 2015, and the map of Xian. The comparison of the
Table 2. The Comparison of Map Matching Performance.

| Testing algorithm                     | Matching efficiency | Error matching rate | Correct rate |
|---------------------------------------|---------------------|---------------------|--------------|
| Traditional distance map matching     | General             | 21.60%              | 73.80%       |
| The algorithm in this paper           | High                | 0.60%               | 96.00%       |

Table 3. The Relationship of States and Speed.

| States                  | Speed (km/h) |
|-------------------------|--------------|
| Enormous congestion     | 0–5          |
| Large congestion        | 5–10         |
| Congestion              | 10–20        |
| Slow moving traffic     | 20–40        |
| Smooth traffic          | >40          |

Traditional distance map matching and the study is shown in the Fig. 5. It shows that the matching efficiency and accuracy are improved in the paper. After a large number of comparative experiments, some detailed parameters are summarized in Table 2.

![Figure 5. The comparison between distance map matching and our method.](image)

After map matching, based on the current road speed (section 3), we use five colors to express the traffic congestion of urban road by HIS transform, which are shown in Fig. 6. The relationship of speed and them is listed in Table 3, which is divided by our experience.
Based on the platform and the algorithm above, we achieve congestion visualization, and the result of March 12, 2015 PM 7 is shown in Fig. 7. Through the actual investigation in a large number of sections, we find its reliability is high, and the accuracy can be more than 94%.

6. Conclusions. Traffic congestion visualization analysis algorithm based on FCD is proposed in the paper, which can analyze accurately real-time road traffic condition to display. It has better anti-noise performance for the GPS data, and the reliability and real-time performance are superior to the traditional algorithm; The FCD can be accurately matched by an improved topology map matching algorithm, which accuracy rate is up to 96%, and is obviously higher than the traditional distance map matching algorithm; The current road speed is estimated by instantaneous velocity and average velocity, to describe congestion states. Then, after HIS transform, they are visualized in the map by five kinds of colors. So this visualization method is more clear; With the high reliability and fault tolerance, the algorithm proposed in this paper can meet the basic requirements of people’s daily travel.

REFERENCES

[1] M. Bierlaire, J. Chen and J. Newman, A probabilistic map matching method for smartphone GPS data, Transportation Research Part C Emerging Technologies, 26 (2013), 78–98.
[2] M. Y. Cao and D. Y. Yu, Pseudocolor coding of gray image based on perceptual color space, Optical Technology, 28 (2002), 338–367.
[3] B. Y. Chen, H. Yuan and Q. Lam, Map-matching algorithm for large-scale low-frequency floating car data, International Journal of Geographical Information Science, 28 (2014), 22–38.
[4] F. Chen, M. Shen and Y. Tang, Local path searching based map matching algorithm for floating car data, Procedia Environmental Sciences, 10 (2011), 576–582.
[5] J. S. Greenfeld, Matching GPS observations to locations on a digital map, in Proceedings of Annual Meeting of the Transportation Research Board Washington D.C., 2002.
[6] J. Guo, W. Huang and B. M. Williams, Real time traffic flow outlier detection using short-term traffic conditional variance prediction, Transportation Research Part C Emerging Technologies, 50 (2015), 160–172.
Figure 7. Results of congestion visualization.

[7] J. Hu, X. Peng and Z. Xu, Study of gray image pseudo-color processing algorithms, in Proc. SPIE 8415 International Symposium on Advanced Optical Manufacturing and Testing Technologies: Large Mirrors and Telescopes, 8415, International Society for Optics and Photonics, 2012, 842519.

[8] G. Y. Jiang, A. D. Chang and Q. Li, Estimation models for average speed of traffic flow based on GPS data of taxi, Journal of Southwest Jiaotong University, 46 (2011), 638–644.

[9] Y. Lou, C. Zhang and Y. Zheng, et al., Map-matching for low-sampling-rate GPS trajectories, ACM SIGSPATIAL GIS 2009, 2009, 352–361.

[10] S. Messelodi, M. C. Modena and M. Zanin, Intelligent extended floating car data collection, Expert Systems with Applications, 36 (2009), 4213–4227.

[11] T. Miwa, D. Kiuchi and T. Yamamoto, Development of map matching algorithm for low frequency probe data, Transportation Research Part C Emerging Technologies, 22 (2012), 132–145.

[12] M. A. Quddus, W. Y. Ochieng and L. Zhao, A general map matching algorithm for transport telematics applications, GPS Solutions, 7 (2003), 157–167.

[13] N. R. Velaga and A. L. Bristow, Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems, Transportation Research Part C Emerging Technologies, 17 (2009), 672–683.

[14] H. Yang, S. Cheng and H. Jiang, An enhanced weight-based topological map matching algorithm for intricate urban road network, Procedia Social and Behavioral Sciences, 96 (2013), 1670–1678.

[15] J. S. Yang, S. P. Kang and K. S. Chon, The map matching algorithm of GPS data with relatively long polling time intervals, Journal of the Eastern Asia Society for Transportation Studies, 6 (2005), 2561–2573.
[16] Q. Yuan, Z. Liu and J. Li, A traffic congestion detection and information dissemination scheme for urban expressways using vehicular networks, *Transportation Research Part C Emerging Technologies*, 47 (2014), 114–127.

[17] Y. C. Zhang, X. Q. Zuo and L. T. Zhang, Traffic congestion detection based on GPS floating-car data, *Procedia Engineering*, 15 (2011), 5541–5546.

[18] Y. Zheng and M. A. Cheng, Weight-based shortest-path aided map matching algorithm for low-frequency positioning data, Transportation Research Board Meeting, 2011.

[19] S. Zhong, X. Y. Jiang and Z. Wei, Pseudo-Color Coding with phase modulated image density, in *International Conference on Micro/Nano Optical Engineering (ICOME)*, 2011.

Received May 2015; revised September 2015.

E-mail address: jingmzhou@163.com
E-mail address: xmzhao@chd.edu.cn
E-mail address: promotercx@126.com
E-mail address: 38301187@qq.com