Journal of Network Intelligence
Taiwan Ubiquitous Information

Volume **8**, Number **1**, February **2023**

GPS Map Matching Algorithm Based on Improved D-S Evidence Theory in Complex Road Sections

Zi-Ming Teng
College of Communications Engineering, Jilin University, China.
Nanhu Campus, Jilin University, No. 5372, Nanhu Road, Changchun City, Jilin Province
1020299792@qq.com

Ze-Nan Huangfu
School of Electrical Engineering, Northeast Electric Power University, China
No. 169, Changchun Road, Chuanying District, Jilin City, Jilin Province
1251536891@qq.com

Jin-Liang Gu*
School of Electrical Engineering, Northeast Electric Power University, China
No. 169, Changchun Road, Chuanying District, Jilin City, Jilin Province
1963631427@qq.com

Xing-Xing Wang
School of Electrical Engineering, Northeast Electric Power University, China
No. 169, Changchun Road, Chuanying District, Jilin City, Jilin Province
1471385830@qq.com

Zhuo Yang
Criminal Investigation Police University of China
No.83 Tawan Street, Huanggu District, Shenyang, Liaoning Province
7718944@qq.com

Yu Zhang
East China Electric Power Design Institute CO LTD of China Power Engineer Consulting Group
Shanghai Putuo District Guchuan Road Zhongjun Tianyue
228226327@qq.com

*Corresponding author: Jin-Liang Gu

Received September 17, 2022, revised November 21, 2022, accepted January 2, 2023.

Abstract. *When a vehicle is driving at the intersection, mismatching will occur when traditional D-S evidence theory algorithm is matched. This paper improves the matching algorithm to solve this problem. The improved algorithm eliminates the abnormal location points and fills the vacancy. Candidate roads are identified using a simplified error ellipse to reduce matching time. Combining directional probability function, distance probability function and confidence function to improve the probability formula of candidate segments. The experimental results show that the improved algorithm has a matching accuracy of about 97%. Compared with existing map matching algorithms, the accuracy can be improved by about 4%, and the single point matching time can be reduced by about 1 ms.*
Keywords: GPS, D-S evidence theory, map matching, distance threshold, matching time.

1. **Introduction.** With the rapid development of the economy and the rise of population in large cities, urban transportation has become developed and roads have become complex [1]. In the actual vehicle navigation, due to GPS system error, signal transmission error, accidental error of data, and accuracy problems of electronic maps, there is a deviation between GPS track point and actual road point [2]. This will make the driver deviate from the correct driving route, drive incorrectly, and more serious may cause traffic accidents. Therefore, a fast, effective and accurate method is especially important to improve navigation performance [3]. At present, there are two main solutions [4]. One is to improve the accuracy of GPS track points and electronic maps, but this method is more complex and costly, which makes it difficult to update electronic maps in time. Another method is to perform a series of processing on GPS track points by map matching, especially to correct the track point data which deviates from the correct road to reposition it on the correct road. This method is easy to achieve, low cost and has become a research hot point [5, 6, 7].

The map matching technology is widely used to provide location services-related help in areas such as vehicle navigation, traffic flow analysis, Geo-social network analysis, and so on [8, 9, 10]. At present, the proposed algorithms are mainly for most ordinary sections [11]. The algorithms mainly include probability matching algorithm [12], geometric matching algorithm [13], topology constraint matching algorithm [14, 15], weight matching algorithm [16, 17]. These algorithms have their own advantages and disadvantages. Some algorithms are simple in principle and easy to implement, but their matching accuracy is not high. Others are complex in principle and difficult to implement, but have high matching accuracy. Crossing road segment [18] is the most complex situation in the road network, with high road density and complex connection. Current map matching algorithms generally treat it as a normal road, which may cause mismatching, mismatch, low matching accuracy, etc. Therefore, the matching of intersecting road segments is the difficult point of map matching [19]. In Figure 1, a car is traveling on Section A with a destination of Z and is preparing to enter a continuous intersection. The vehicle's true track is Section A to Section B. However, due to mismatching, the track point is mismatched to Section C. To reach the destination, it needs to turn right into Section D. For a real track, turning right is equivalent to entering Section E. As we get farther and farther from our destination, the final driving route is very different from the correct one, so improving the map matching accuracy of the intersection is very important for us to get to our destination quickly and accurately [20]. In order to solve the problem of intersecting road segments, this paper presents an improved D-S evidence theory algorithm under the environment of intersecting road segments. It mainly uses directional evidence to improve the probability formula of candidate road segments to make up for the low accuracy of intersecting road segments in traditional D-S evidence theory.

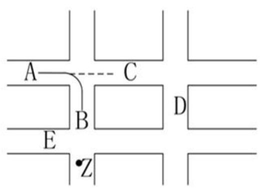

Figure 1. Mismatching diagram

2. The Implementation of Map Matching.

2.1. **Date pre-processing.** Before map matching, the positioning data must be preprocessed [21]. Data pre-processing consists of two parts: exceptional data is first eliminated and [22] interpolated to complete missing data, then grid index is generated.

2.1.1. *Eliminate abnormal data and complete missing data.* There are tunnels, viaducts and other complex situations on urban roads, which may affect the transmission of GPS signals and make the data obtained by receivers abnormal. In order to reduce the influence of abnormal data on matching performance, it is necessary to process GPS data, eliminate abnormal data and use the interpolation method to complete the next missing data according to two consecutive vehicle historical registration points. The information contained in the location point includes latitude and longitude, vehicle speed, time stamp Angle between the direction of the vehicle and due north, etc. The elimination principle is shown in Figure 2. The interpolation principle is shown in Figure 3.

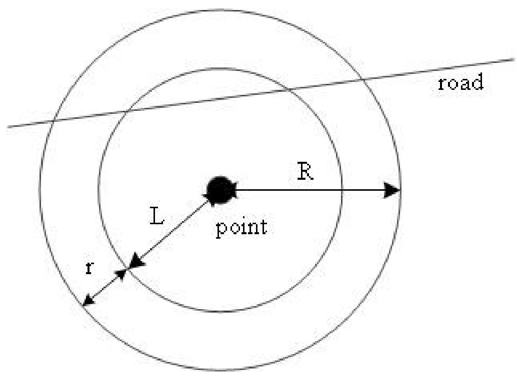

Figure 2. Elimination principle

This paper sets a distance threshold R to remove locations points beyond the threshold range. The distance threshold is determined by two parts, which are the maximum offset of the vehicle and the positioning error.

First, the maximum speed $V_{\max}$ of the vehicle is calculated from Equation (1).

$$V_{\max} = V_{t\max} + V_s \tag{1}$$

$V_{t\max}$ is the max highway speed limit and V_s is the instantaneous speed error.

The maximum vehicle offset L is:

$$L = V_{\max} + \Delta t \tag{2}$$

Δt is the sampling interval. r is the error circle radius of positioning data under a certain confidence level, from which the distance threshold can be obtained.

$$R = L + r \tag{3}$$

If the distance between the location point and the road section exceeds the distance threshold, the location point will be removed, otherwise the location point will be retained.

Due to the high sampling frequency of vehicles in the intersection, any two adjacent sampling points can be regarded as a straight road. The coordinates of the positioning points after interpolation are shown in Equation (4) (5). (x_i, y_i) represents the coordinates of the current positioning points to be interpolated, and (x_{i-1}, y_{i-1}) represents the coordinates of the previous historical positioning points of the vehicle. v_{i-1} is the driving speed of the vehicle at the previous historical positioning point, ψ_{i-1} and ψ_{i-2} respectively

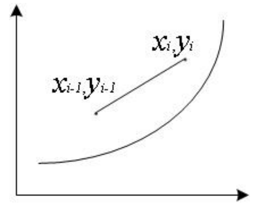

Figure 3. Schematic of interpolation

represent the angle between the driving direction and the positive and north direction of the vehicle at the first two positioning points.

$$x_i = x_{i-1} + v_{i-1} \times \Delta t \times \sin\left[(\psi_{i-1} - \psi_{i-2}) + \psi_{i-1}\right] \tag{4}$$

$$y_i = y_{i-1} + v_{i-1} \times \Delta t \times \cos\left[(\psi_{i-1} - \psi_{i-2}) + \psi_{i-1}\right] \tag{5}$$

2.1.2. *Generating grid index.* In the process of map matching, it is necessary to obtain the candidate road sections of GPS positioning points. The initial way is each location point for the query the whole road network in the electronic map every time a positioning point is obtained. This method is more accurate, but it will take a long time, which will affect the matching performance. Therefore, this paper introduces grid index [23], so as to reduce the query time and quickly obtain the candidate road sections. The basic idea of grid indexing is to divide the whole electronic map road network into several grids and calculate the sections contained or intersected in each grid in advance. When querying, first determine the grid where the object is located, and then quickly query the candidate road sections contained in the grid, so as to greatly improve the query speed.

2.2. **Candidate set acquisition.** After pre-processing the positioning data, the approximate area of the actual road is determined according to the positioning data information. At present, the common method is to determine an error ellipse according to the probability, which contains the vehicle position.

The derivation formula of error ellipse is as follows:

$$a = \sigma_0 \sqrt{\frac{1}{2\left(\sigma_X^2 + \sigma_Y^2\right)} + \sqrt{\sigma_X^2 - \sigma_Y^2 + 4\sigma_{XY}^2}} \tag{6}$$

$$b = \sigma_0 \sqrt{\frac{1}{2\left(\sigma_X^2 - \sigma_Y^2\right)} + \sqrt{\sigma_X^2 - \sigma_Y^2 + 4\sigma_{XY}^2}} \tag{7}$$

$$\varphi = \frac{\pi}{2} - \frac{1}{2}\arctan\left(\frac{2\sigma_{XY}}{\sigma_X^2 - \sigma_Y^2}\right) \tag{8}$$

a and b is the long and short half axes of the ellipse, σ_X and σ_Y is the standard deviation of longitude and latitude of the location point, σ_{XY} is the covariance, φ is the angle between the long half axis of the ellipse and due north direction, and σ_0 is the adjustable factor. As can be seen from the above ellipse formula, the calculation of relevant parameters is relatively complex, so this paper simplifies it to reduce the calculation cost. The simplified ellipse center is the current positioning point $P_i\left(x_i, y_i\right)$.

The calculation of relevant parameters is as follows:

$$a' = R \times \Delta t \tag{9}$$

$$c' = \sqrt{\left(x_i - x_{i-1}\right)^2 + \left(y_i - y_{i-1}\right)^2} \tag{10}$$

$2a'$ and $2c'$ is the major axis and focal length of the simplified ellipse respectively. In the simplified ellipse, the grid of the trajectory points is determined based on the grid index, and the links included in or tangent to the error circle in the 3×3 grid centered on the grid are queried as candidate links. Let all candidate road sets extracted from the error circle is $D = \{S_1, S_2, \ldots, S_i, \ldots, S_k \mid i = 1, 2, \ldots, k\}$, The process is as shown in Figure 4. Point O is the locating point, and K, L, M, N are the road sections in the 3×3 grid with the locating point as the center. In these sections, find the sections contained in the error circle or tangent to the error circle, and finally get the candidate sections of point O as M and N.

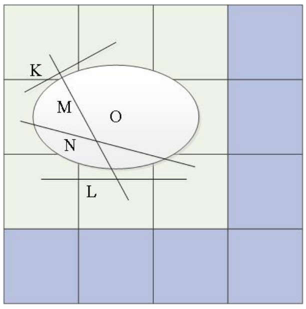

FIGURE 4. The diagram of obtaining the candidate road section

2.3. **Application of algorithm in intersection.**

2.3.1. *Probability function* $\mathbf{M}_{\mathbf{1}(s_i)}$ *of distance.* Distance is the most important criterion in the map matching algorithm. The real position of a certain positioning point may be on the road which closest to it, but sometimes there will be wrong judgment. According to the distance, different probabilities $\mathbf{M}_{\mathbf{1}(s_i)}$ are assigned to all candidate links according to the distance, where $i = 1, 2, 3, \ldots, k$. And the sum of the probabilities of all the candidate link distances of the location points is 1.

The distance evidence function α_i is constructed as follows (11):

$$\alpha_i = \frac{a'}{d_i} \tag{11}$$

d_i is the shortest distance from the GPS positioning point to the candidate road section s_i.

The probability distribution function of the distance is constructed as follows (12).

$$m_{1(s_i)} = \frac{\alpha_i}{\sum_{j=1}^{k} \alpha_j} \tag{12}$$

As shown in Figure 5, when calculating the distance d_i, according to whether the projection of the GPS anchor point is on the road section, this paper analyses the two situations. Assuming the coordinates of the positioning point is (x, y) and the coordinates of the road endpoints are $N(x_1, y_1)$, $M(x_2, y_2)$.

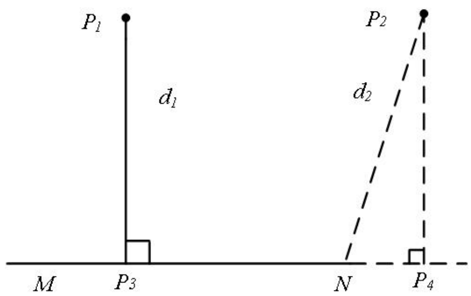

Figure 5. Calculate the shortest distance

If the projection point of locating point P_1 on road MN is P_3 and P_3 is on the road, d_1 is the shortest distance of segment P_1P_3.

$$d_1 = \frac{|(y_2 - y_1)\,x + (x_1 - x_2)\,y + (y_1x_2 - y_2x_1)|}{\sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2}} \tag{13}$$

The projection point of location point P_2 on the road segment MN is P_4, but P_4 is not on the road segment, and on its extension line. At this time, the nearest road point to P_2 is N, and the shortest distance d_2 is the length of segment P_2N.

$$d_2 = \sqrt{(y_1 - y)^2 + (x_1 - x)^2} \tag{14}$$

2.3.2. *Probability function* $\mathbf{M}_{\mathbf{2}(s_i)}$ *of direction.* Direction is another criterion to determine the matching road section when there are many candidate roads around the vehicle. When considering the direction, it is more complicated to calculate the angle between the direction of vehicle and the direction of the road to which it belongs. In this paper, the angle between the direction of vehicle driving and the direction of the north and the direction of the road to the north is obtained. Then, the angle between the direction of vehicle driving and the direction of the road belongs to can be obtained by making a difference between the two angles. According to the angle between the two, different probability $\mathbf{M}_{\mathbf{2}(\boldsymbol{s_i})}$ is assigned to all the candidate links of the location point, where $i = 1, 2, 3, \dots, k$ The sum of the probabilities of all candidate link directions of the same location point is 1.

Directional evidence function β_i is constructed as:

$$\beta_i = \frac{\pi}{\theta_i} \tag{15}$$

θ_i is the angle between the direction of the vehicle and the direction of the road.

The probability distribution function of the direction is defined as:

$$m_{2(s_i)} = \frac{\beta_i}{\sum_{j=1}^{k} \beta_j} \tag{16}$$

As shown in Figure 6, select the straight line section on road S_i and select two points (denoted as P and Q) on it. Take P as the origin, set Q coordinates as (x_3, y_3), and calculate the angle γ_i between road S_i and north direction through Equation (17).

$$\gamma_i = \begin{cases} \frac{\pi}{2} - \arctan\left(\frac{y_3}{x_3}\right), & x_3 > 0 \\ \frac{3\pi}{2} - \arctan\left(\frac{y_3}{x_3}\right), & x_3 < 0 \end{cases} \tag{17}$$

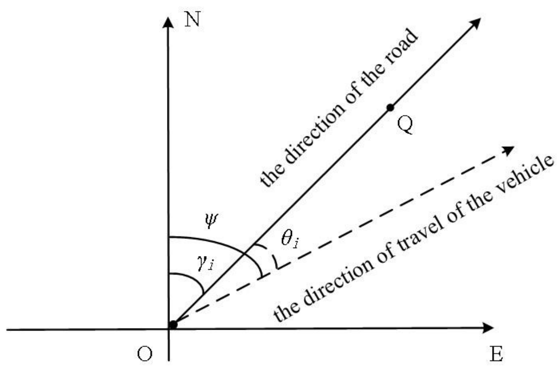

FIGURE 6. Muti-angle diagram of road direction and vehicle direction

If $x_3 = 0$, $y_3 < 0$, then $\gamma_i = \pi$; if $x_3 = 0$, $y_3 > 0$, then $\gamma_i = 0$.

As for the angle between the vehicle and the north, many positioning software can obtain it directly.

The difference between the direction of the vehicle and the direction of the road is Equation (18).

$$\theta_i = \begin{cases} |\psi - \gamma_i|, & |\psi - \gamma_i| < 180° \\ 360° - |\psi - \gamma_i| & |\psi - \gamma_i| > 180° \end{cases} \tag{18}$$

2.4. **Probability calculation of candidate road sections.** The research object of this paper is the intersecting road. It is possible that the vehicle location point is between the roads. Because the direction angles between the roads are quite different, the direction information is more reliable than the distance information. Location points are matched based on direction information. The matching result is the largest in I road segment and the probability of the other roads is small. The more reliable the matching result is in I road segment.

The definition of credibility is as follows (19).

$$\varepsilon = e^{-\left(1-m_{2(s_i)}\right)} \tag{19}$$

After figuring out the probability distribution function $m_{1(s_i)}$ of the distance, the probability distribution function $m_{2(s_i)}$ of the direction and the credibility, the probability calculation of the candidate road sections was carried out. The formula of the probability calculation of the candidate road sections is as follows (20). The F and G points in the formula are candidate road segments for GPS points. The candidate road segment corresponding to the maximum probability of the true GPS point is selected as the matching road segment of the GPS point, that is, the vehicle is considered on the road segment at the current moment.

$$m_{(s_i)} = \sum\nolimits_{s_i \subseteq (F \cup G)} m_{1(F)} \times m_{2(G)} \times \varepsilon \tag{20}$$

2.4.1. *Algorithm steps and process.*

Step 1: Pre-processing positioning data. Set thresholds to remove outliers, interpolate missing data, and index the grid to reduce matching time.
Step 2: Get the candidate road set according to the simplified error circle.
Step 3: Calculate the basic probability function of distance and direction.
Step 4: Define credibility based on directional evidence.
Step 5: Using improved candidate segment probability formula to calculate all candidate segment probabilities for track points.
Step 6: Select the segment with the maximum probability as the matching segment

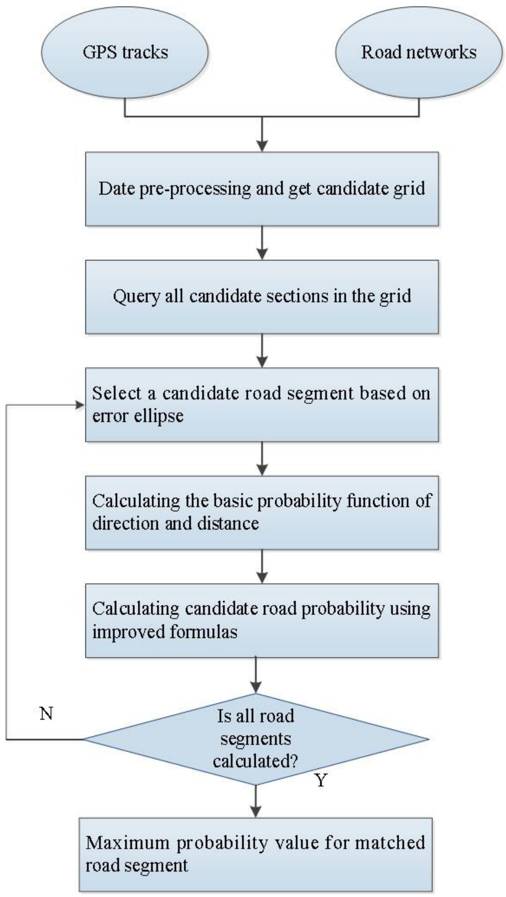

FIGURE 7. Process of the improved D-S evidence theory map matching algorithm

3. Algorithm Demonstration and Simulatinhe.

3.1. Comparison of simulation results. Figure 8 is the simulation chart of the matching accuracy of four different map matching algorithms: Direct projection algorithm, Traditional D-S evidence theory algorithm, Curve fitting matching algorithm and Text matching algorithm. The accuracy of the algorithm is measured by the ratio of the number of correctly matched locations to the total number of locations obtained. The matching accuracy of text matching algorithm in parallel sections is lower compared to other map matching algorithm, but it has a very high matching accuracy at intersecting road. The matching accuracy is about 97%, compared with other methods in accuracy increased by 4% This is because the algorithm in this paper is aimed at intersection roads, focus on using the directional information, so the algorithm is best used at cross roads.

Figure 9 is four simulation results of single point matching time for four algorithms when the number of candidate road segments is 2, 3, 4 and 5. The dark blue curve is the direct projection algorithm, the light blue curve is the traditional D-S evidence theory algorithm, the green curve is the curve fitting algorithm, and the yellow curve is the text matching algorithm. The matching time is the average time from getting the location point to determining the candidate road segment. Compared with the four pictures, the single point matching time of text matching algorithm increases with the increase of the candidate road segments in the candidate area, but the increase is the least. When there are two candidate road segments in the candidate area, the single point matching time of text matching algorithm is about 4.2 ms. When there are five candidate roads in the candidate area, the single point matching time of text matching algorithm is only about 5.4 Ms. Among the four algorithms, the single point matching time of the paper improved algorithm is lower than that of other algorithms regardless of the number of candidate roads in the candidate area. Overall, the matching time of the text improved algorithm can be improved by at least 1 ms, and the real-time performance is good.

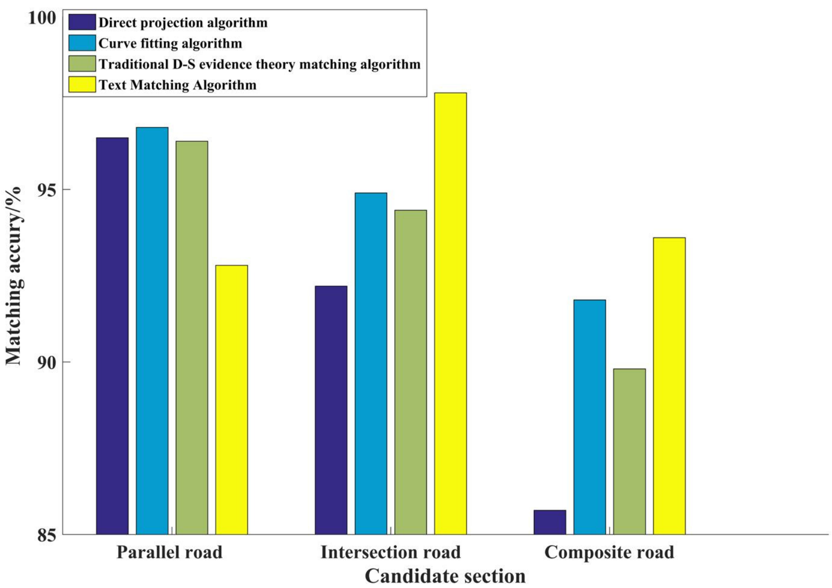

FIGURE 8. Comparison of matching accuracy

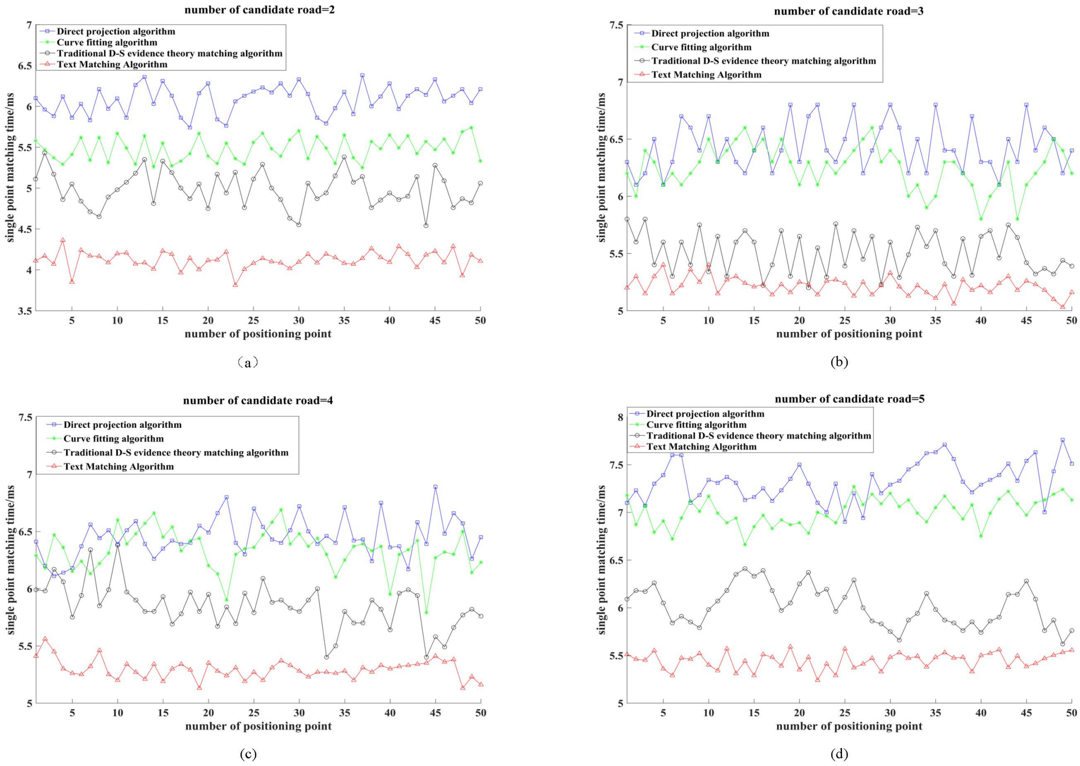

Figure 9. Single point matching time for different number of roads: (a) Single point matching time for two candidate roads (b) Single point matching time for three candidate roads (c) Single point matching time for four candidate roads (d) Single point matching time for five candidate roads

4. **Conclusion.** In this paper, an improved D-S evidence theory map matching algorithm is proposed under the crossing road environment. By setting a distance threshold to eliminate outliers and fill in missing points, the error ellipse is simplified and the matching time is reduced. Combining directional probability function, distance probability function and confidence function to improve the probability formula of candidate segments. The results show that the algorithm improves the matching accuracy and shortens the single point matching time of the locations.

REFERENCES

[1] S.-M. Zhang, X. Su, X.-H. Jiang, M.-L. Chen, and T.-Y. Wu, "A traffic prediction method of bicycle-sharing based on long and short term memory network," *Journal of Network Intelligence*, vol. 4, no. 2, pp. 17-29, 2019.

[2] G. Wu, Y. Qiu, and G. Wang, "Map matching algorithm based on hidden Markov model and genetic algorithm," *Journal of Northeastern University (Natural Science)*, vol. 38, no. 4, pp. 472-475, 2017.

[3] C.-M. Chen, B. Xiang, Y. Liu, and K.-H. Wang, "A secure authentication protocol for Internet of vehicles," *IEEE Access*, vol. 7, pp. 12047-12057, 2019.

[4] J. Yan, Z. Hu, and X. Yu, "Research on coordinate transformation based on CORS and static relative positioning technology," *Bulletin of Surveying and Mapping*, vol. 12, pp. 54-56, 2015.

[5] J.-C. Fan, C. Fu, K. Stewart, and L. Zhang, "Using big GPS trajectory data analytics for vehicle miles traveled estimation," *Transportation Research Part C*, vol. 103, pp. 298-307, 2019.

[6] Y.-L. Hsueh, H.-C. Chen, " Map matching for low-sampling-rate GPS trajectories by exploring real-time moving directions," *Information Sciences*, vol. 34, no. 1, pp. 433-434, 2018.

[7] M. Čelan, M. Lep, "Bus-arrival time prediction using bus network data model and time periods," *Future Generation Computer Systems*, vol. 110, pp. 364-371, 2020.

[8] T.-Y. Wu, X.-L. Guo, L. Yang, Q. Meng, and C.-M. Chen, "A lightweight authenticated key agreement protocol using fog nodes in Social Internet of vehicles," *Mobile Information Systems*, vol. 2021, 3277113, 2021.

[9] Z. Li, Q.-K. Miao, S.-A. Chaudhry, and C.-M. Chen, "A provably secure and lightweight mutual authentication protocol in fog-enabled social Internet of vehicles," *International Journal of Distributed Sensor Networks*, vol. 18 (6), 2022.

[10] T.-Y. Wu, Z.-Y. Lee, L. Yang, and C.-M. Chen, A Provably Secure Authentication and Key Exchange Protocol in Vehicular Ad Hoc Networks, Security and Communication Networks, vol. 2021, 9944460, 2021.

[11] Z.-J. Teng, Z.-Q. Z, and Y.-C. He, "Fast map-matching algorithm for highway based on initial position," *Journal of Chinese Inertial Technology*, vol. 254, pp. 461-465, 2017.

[12] H. Guo, M. Ren, and P. Wang, "Indoor map matching algorithm based on conditional random field model," *Journal of Instrumentation*, vol. 39, no. 5, pp. 107-114, 2018.

[13] F. Chen, T. Zou, and L. Wang, "Improved projection-based map matching algorithm," *CEA*, vol. 47, no. 31, pp. 242-244+248, 2011.

[14] L. Yuan, D. Li, and H. Song, "A map-matching algorithm with low-frequency floating car data based on matching path," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 146-152, 2018.

[15] C. Blazquez, J. Ries, P.-A. Miranda, and R-J. León, "An instance-specific parameter tuning approach using fuzzy logic for a post-processing topological map-matching algorithm," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 4, pp. 87-97, 2018.

[16] M.-J. He, L.-J. Zheng, W. Cao, J. Huang, X. Liu, and W.-N. Liu, "An enhanced weight-based real-time map matching algorithm for complex urban networks," *Physica A: Statistical Mechanics and its Applications*, vol. 534, pp. 122318, 2019.

[17] W.-X. Teng, Y.-H. Wang, "Real-time map matching: A new algorithm integrating spatio-temporal proximity and improved weighted circle," *Open Geosciences*, vol. 11, no. 1, pp. 288-297, 2019.

[18] L. Tang, Z. Kan, and F. Huang, "Intersection transit time detection using low-frequency spatio-temporal GPS trajectories," *Wuhan University Journal (Information Science)*, vol. 41, no. 1, pp. 136-142, 2016.

[19] H. Qi, X.-Q. Di, and J.-Q. Li, "Map-matching algorithm based on the junction decisiondomain and the hidden Markov model," *PloS one*, vol. 14, no. 5, pp. 0216476, 2019.

[20] T.-Y. Wu, Z.-Y. Lee, L. Yang, J.-N. Luo, and R.-L. Tso, "Provably Secure Authentication Key Exchange Scheme Using Fog Nodes in Vehicular Ad-Hoc Networks," *The Journal of Supercomputing*, vol. 77, pp. 6992-7020, 2021.

[21] V. Kumar, R. Kumar, S. Jangirala, S. Kumari, S. Kumar, and C.-M. Chen, "An Enhanced RFID-Based Authentication Protocol using PUF for Vehicular Cloud Computing," *Security and Communication Networks*, vol. 2022, 8998339, 2022.

[22] X.-M. Wang, S.-Y. Shen, D. Bezzina, J.-R. Sayer, H.-X. Liu, and Y.-H. Feng, "Data infrastructure for connected vehicle applications," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 5, pp. 85-96, 2020.

[23] M. Liu, M. Li, X.-Y. Xu, and S.-J. Mao, "An improved map matching algorithm based on HMM model," *Journal of Peking University (Natural Science)*, vol. 54, no. 6, pp. 1235-1241, 2018.