

Research on Front-End Fusion Processing Technology of Traffic Scenes

Xin Zhou*

China Merchants Chongqing Communications Technology Research & Design Institute Co., Ltd., China

*Corresponding author: Xin Zhou, zhouxin1@cmhk.com

Copyright: © 2022 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the intelligent development of road traffic control and management, higher requirements for the accuracy and effectiveness of traffic data have been put forward. The issue of how to collect and integrate data for traffic scenes has sought importance in this field as various treatment technologies have emerged. A lot of research work have been carried out from the theoretical aspect to engineering application.

Keywords: Traffic scene; Data fusion; Traffic data processing

Online publication: March 4, 2022

1. Overview of front-end fusion processing technology

The acquisition of original traffic data depends on various traffic information acquisition equipment. However, in practical application, drift, deformity, repetition, and even serious errors often occur due to the equipment itself in addition to external interferences and immature existing technologies, which bring great challenges to traffic information mining based on original data. Therefore, original data cleaning and correcting as well as secondary information mining have great practical significance in theoretical research and engineering. Up to now, researchers have carried out a lot of research work in this field.

2. Data preprocessing technology

Data preprocessing is aimed at improving fault data (lost and wrong data), which includes fault data identification, repair, and data filtering.

In data preprocessing research, the theoretical research methods for fault data identification mainly include low-order polynomial sliding regression method, clustering method, association rule method, etc. In practice, threshold method is applied, including time threshold, speed threshold, latitude and longitude threshold, etc. The theoretical research methods for fault data repair mainly include wavelet denoising method, immune algorithm, spectral subtraction, etc. The methods used in practice mainly include the historical trend method, weighted method of historical trend and measured data, adjacent period data repair method, adjacent road section data restoration method, and linear interpolation method.

3. Map matching technology

The GPS system of floating vehicle often causes positioning deviation due to signal propagation error, delay error, reception error, and other factors. Map matching is the technical means to match GPS points with positioning deviation to the two-dimensional road network plane, involving the definition of road network attributes and the division of road section levels. The main algorithms include geometric matching

algorithm, probabilistic matching algorithm, Kalman filtering algorithm, fuzzy theory, artificial neural network, and others. According to different matching methods, map matching algorithms can be divided into point-road section matching, point sequence sub-path matching, global overall matching, and other matching algorithms. The characteristics and descriptions of the map matching algorithms are shown in **Table 1**.

Table 1. Characteristics and descriptions of map matching algorithms

Algorithm		Advantages and disadvantages of the algorithm	References
Point-road section matching	Advantages	The logic is simple, and it has good real-time performance.	Wang M, et al.;
	Disadvantages	The connection between GPS point sequences and the overall trend of trajectory is not considered, and the matching accuracy is low.	Zhang G, et al. [1,2]
Point sequence sub-path matching	Advantages	The relationship between GPS point sequences is considered, and it has good real-time performance.	He Z, et al.; Li Q, et al. [3,4]
	Disadvantages	The overall trend of trajectory is not considered; the efficiency and accuracy of the algorithm are affected by the data sampling interval.	
Global overall matching	Advantages	It has high matching accuracy and is conducive to direct track recovery.	Yu Y, et al.;
	Disadvantages	The efficiency of the algorithm is greatly affected by the data quality; the algorithm is complex, and its real-time performance is poor.	Bouillet E, et al. [5,6]
Other algorithms	Advantages	With the help of advanced technology, real-time performance is better.	H X, et al.; Zou L, et al. [7,8]
	Disadvantages	It has high requirements for hardware and supporting systems, and the universality of the algorithm is poor.	

3.1. Point-road section matching

In the research on application algorithm based on point-road section matching [1], Meiling Wang and other researchers obtained the candidate matching sections of GPS points using the grid method and determined the final matching section according to the distance matching degree, heading matching degree, and accessibility between the GPS positioning points and the candidate sections. In particular, the shortest path algorithm is used to achieve complete topology for the multiple sections between two GPS points caused by large positioning distance.

3.2. Point sequence sub-path matching

In the research on application algorithm based on point sequence sub-path matching [3], several researchers used the point-line matching method to realize the matching of “GPS confidence point” based on error circle analysis. When single point matching is difficult, the “GPS confidence point” is used as the end point of the time window in the longest delay matching tolerance, and the topological connectivity of the two GPS points before and after the window is analyzed to determine the optimal matching sub-path.

3.3. Global overall matching

In the study on application algorithm based on global overall matching [4], Qingquan Li and other researchers found a set of alternative matching sections for all GPS points based on the GPS point error buffer; they connected the set of alternative matching sections to form a set of alternative matching paths according to the traffic constraints of floating vehicles and the geometric connectivity of roads; finally, they selected the most similar path as the final matching path according to the principle of curve similarity, avoiding the setting of factor weight in point based matching method.

3.4. Other algorithms

Massive data real-time matching technology research focuses on the reform of algorithms based on data structure and the application of advanced technology. In a study [5], Yang Yu improved the index structure of the traditional uniform grid map matching algorithm and the delayed matching of historical data; the researcher also designed the matching degree feedback matching algorithm in consideration of the time factor and proposed the processing method of large sample size data in the “cloud environment.”

4. Traffic parameter extraction technology

Accurate traffic parameters are the premise of effective decision-making. The vehicle speed can be used for road traffic state discrimination, traffic event detection, and road traffic management. Travel time can be used for travel guidance, intelligent navigation, and dynamic path planning. The fixed section length also contributes to their roles in characterizing road traffic conditions, conducting road traffic management and control, as well as formulating strategies. From the perspective of application, the research on vehicle speed estimation and travel time estimation includes theoretical analysis and engineering application. The characteristics and descriptions of common theoretical algorithms are shown in **Table 2**.

Table 2. Characteristics of traffic parameter extraction algorithm

Method		Advantages and disadvantages of the algorithm		Applicable object	Expression
Based on knowledge discovery	Neural network	Advantages	There is no need to build a complex mathematical model; high estimation accuracy, and strong adaptability to complex road network and traffic conditions.	Travel time [8,9]	
		Disadvantages	The training samples are large and difficult to ensure convergence; the portability is poor.		
	Evidential reasoning			Average speed [10]	
	Weighted fusion	Advantages	Simple form, low complexity, and good real-time performance.	Average speed [9]	$\bar{v} = w_1 \bar{v}_1 + \dots + w_n \bar{v}_n$
		Disadvantages	The accuracy of the extracted parameters is heavily affected.	Travel time [4-7]	$\bar{T} = w_1 \bar{t}_1 + \dots + w_n \bar{t}_n$
Based on linear system theory	Parameter direct push	Advantages	The derivation logic and form are simple; low complexity and good real-time performance.	Travel time [7,8]	$\bar{T} = t_1 + t_2 + \dots + t_n$ or $\bar{T} = t_n - t_1$
		Disadvantages	The need to deduce the disturbed parking time; the accuracy of the parameters extracted under the complex road network and mixed traffic state is low.		

(Continued on next page)

(Continued from previous page)

Method	Advantages and disadvantages of the algorithm		Applicable object	Expression
Parametric regression	Advantages	Simple form and low complexity.	Average speed [5] Travel time [4]	$\int_{t_1}^{t_2} \bar{v}_x dt = X$ or $\int_{x_1}^{x_2} \frac{1}{v_x} dx = \bar{T}$ $= \sum_{j \in J_i} \int_t^{t+\Delta t_i} V_{i,j}(t) dt$
	Disadvantages	Large amount of pretreatment, weak robustness, and the precision of parameters extracted under the complex road network and traffic conditions is low.		
Based on linear system theory	Advantages	Strong adaptability and robustness to complex road network and traffic state.	Average speed	/
	Disadvantages	High complexity and poor real-time performance.		
Combination extraction	Advantages	Able to make full use of the advantages of each model; the extracted parameters have high accuracy; can be combined freely according to the extracted parameters.	Speed	/
	Disadvantages	High complexity and poor real-time performance.		

4.1. Neural network algorithm

In a study on application algorithm based on neural network [8], Zou and other researchers used BP neural network to fuse floating vehicle and coil data to estimate the travel time. Li and others used BP neural network to dynamically adjust the weight; then, they compared and analyzed two travel time estimation methods based on multi-section detector: speed progressive method and speed average method. The results showed that multi-source data fusion can improve the accuracy of travel time estimates [9].

4.2. Weighted fusion algorithm

In a study on application algorithm based on weighted fusion [10], in reference to low-frequency floating car data, Rahmani and other researchers proposed a path travel time estimation model based on the distance weight of different GPS points in a road section, a path travel time estimation model based on time weight, and a path travel time estimation model based on distance weight and time weight of fusion points, aiming at the difficulty in the estimation of travel time caused by the original floating car data and the complex area of road network.

4.3. Parameter direct push algorithm

In a study on application algorithm based on parameter direct push [11], the researchers first established a road matching model; then, they modified the time node of vehicle operation behavior in the intersection area according to the distribution of front and rear GPS matching points, eliminated the disturbed parking time, and established a single vehicle section travel time estimation model. In view of the phenomenon that the estimation value of adjacent period is unstable in the estimation of average travel time of small sample GPS data and in consideration of the lane flow data collected by induction coil, Fang and other researchers divided the queuing situation at the intersection into unsaturated flow and saturated flow, calculated the

number of dynamic queuing vehicles and the length of queuing area within the cycle, determined the affected area, the delay time within the area, and the normal travel time outside the area, and finally obtained the estimated average travel time of the road section ^[12].

4.4. Parametric regression algorithm

In a study on application algorithm based on parametric regression ^[3], the researchers divided the floating vehicle positioning data into three categories according to different GPS sampling intervals. If the positioning points are continuous in the analysis section, the speed-time integration is used to calculate the effective driving distance of the section; from that, the average speed of the section is deduced. If the positioning points are discontinuous in the sampling time interval, the average speed of the road section is calculated using the vehicle tracking method; if the positioning points are discontinuous in the analysis period, the speed distance integration method is used to calculate the non-zero speed subsection, the travel time of the effective speed part, and then the average speed of the non-zero speed subsection. The parameters of the integral model are solved using the discretization method.

4.5. Combined extraction algorithms

In a study on application algorithm based on the combined model extraction method ^[13], Li and others input the pre-processed average speed estimate obtained based on a variety of detectors into the trained back propagation (BP) neural network model; they used the model to output the basic probability distribution of the average speed estimate interval of multiple detectors and applied the Dempster-Shafer evidence theory (DS theory) to fuse the basic probability assignment of multiple detectors; finally, they obtained the basic probability distribution of the fused average velocity estimation interval and the fall interval of the average velocity estimation interval.

4.6. Overview of engineering application

There are two main methods of engineering application research are as follows: the first is the estimation method based on the information on the instantaneous speed of floating vehicle; the second is the estimation method of floating vehicle timestamp combined with GIS map matching. Details can be seen in **Table 3** and **Table 4**.

Table 3. Road section speed estimation method based on floating vehicle

Road section speed estimation method	Definition	Preliminary treatment	Expression
Average instantaneous velocity method	The average instantaneous speed of all floating vehicles on the road section in a certain period of time represents the average speed of the road section. It has high requirements for sample size.	Combined with road section division, threshold method, adjacent period or road section data repair method, moving average, exponential smoothing method, etc. are adopted to ensure processing accuracy and processing efficiency.	$\bar{v} = \sum v_i/n$
Indirect average speed method	When the length of the road section is known, the average speed of the road section is obtained based on the average travel time for the road section. It is available for large sample interval data, L .		$\bar{V} = L/\bar{T}$

(Continued on next page)

(Continued from previous page)

Road section speed estimation method	Definition	Preliminary treatment	Expression
Weighted speed method	It is a method of weighting the contribution of the estimated average speed of the road section based on the information on the sample speed of each floating car. The adaptability of speed estimation is good.	(See contents in the same column from previous page)	$\bar{v} = w_1 \bar{v}_1 + w_2 \bar{v}_2 + \dots + w_n \bar{v}_n$

Table 4. Travel time estimation method based on floating vehicle

Travel time estimation method	Definition	Preliminary treatment	Expression
Direct method	The travel time of the road section is directly obtained by using the timestamp of the sampling point of the floating car on the road section. It has high requirements for sample frequency.	Combined with road section division, threshold method, adjacent period or road section data repair method, moving average, exponential smoothing method, etc. are adopted to ensure processing accuracy and processing efficiency.	$\bar{T} = t_1 + t_2 + \dots + t_n$ or $\bar{T} = t_n - t_1$
Indirect method	When the length of the road section is known, the travel time is obtained based on the average travel speed. It is available for large sample interval data, L .		

5. Traffic state perception technology

Traffic state perception technology mainly includes traffic state discrimination and dynamic OD estimation.

5.1. Road network traffic state discrimination

Traffic state discrimination can be divided into two categories according to different traffic parameters. The first category is the single factor traffic state discrimination based on road travel time or vehicle speed; the second category is the comprehensive traffic state discrimination considering multiple factors. The main research methods include the California algorithm, McMaster algorithm, exponential smoothing algorithm, standard normal deviation algorithm, improved DS-ANN algorithm, and fuzzy judgment. According to the multidimensional characteristics of urban road network traffic flow, five indicators, including road network traffic congestion index, are set, and the calculation method is given based on real-time travel speed.

5.2. Road network dynamic OD estimation

Road network dynamic OD estimation is the basis for the formulation and evaluation of urban traffic management and control measures. It is also a direct reference for realizing dynamic path planning. There are only a few studies on dynamic OD estimation of road networks using floating vehicles. The idea is to obtain the dynamic OD status of existing floating vehicles and enlarge it to obtain the dynamic OD distribution of the whole road network. Zhang proposed that the OD matrix of floating vehicles arriving at the destination in a certain period of time should be obtained first, followed by the dynamic OD estimation of road network vehicles based on the OD of all vehicles departing during that period of time ^[15].

Disclosure statement

The author declares no conflict of interest.

References

- [1] Wang M, Cheng L, 2012, Research on Floating Vehicle Map Matching Algorithm. *Acta Geodaetica et Cartographica Sinica*, 2012(2): 133-138.
- [2] Zhang G, Li F, 2014, 5th International Conference on Software Engineering and Service Science (ICSESS), June 27-29, 2014: Application of the KNN Algorithm Based on KD Tree in Intelligent Transportation System. IEEE, Beijing, 832-835.
- [3] He Z, She X, Zhuang L, et al., 2013, On-Line Map-Matching Framework for Floating Car Data with Low Sampling Rate in Urban Road Networks. *IET Image Processing*, 7(4): 404-414.
- [4] Li Q, Huang L, 2010, Map Matching Algorithm based on GPS Trajectory Data. *Acta Geodaetica et Cartographica Sinica*, 2010(2): 207-212.
- [5] Yu Y, 2010, Research on Key Technologies of Data Processing of Large Sample Floating Vehicle in Cloud Computing Environment, Wuhan University.
- [6] Bouillet E, Ranganathan A, 2010, Eleventh International Conference on Mobile Data Management, May 23-26, 2010: Scalable, Real-Time Map-Matching Using IBM's System. IEEE, Kansas, 249-257.
- [7] Han X, Liu Y, Yan L, et al., 2011, International Conference on proceedings of the Remote Sensing, Environment and Transportation Engineering (RSETE), June 2011: Parallel Map Matching Algorithm Based on Multi-Core and MPI. Nanjing, 2291-2294.
- [8] Zou L, Xu J, Zhu L, et al., 2007, Travel Time Estimation Model based on Floating Vehicle Movement Detection and Induction Coil Fusion Technology. *Journal of Highway and Transportation Research and Development*, 24(6): 114-117.
- [9] Li R, Chen X, 2014, Research on Road Travel Time Estimation based on Multi-Source Data Fusion. *Journal of Highway and Transportation Research and Development*, 31(2): 99-103.
- [10] Rahmani M, Jenelius E, Koutsopoulos HN, 2013, 16th International IEEE Conference on Intelligent Transportation Systems (ITSC), October 6-9, 2013: Route Travel Time Estimation Using Low-Frequency Floating Car Data. IEEE, Hague, 2292-2297.
- [11] Yu D, Gao X, Yang Z, 2010, Road Travel Time Estimation based on GPS Data and Vehicle Running Characteristics. *Journal of Jilin University (Engineering and Technology Edition)*, 2010(4): 965-970.
- [12] Fang L, Chen S, Zhao F, 2012, Average Travel Time Estimation based on Small Sample Floating Vehicle System. *Computer Simulation*, 2012(9): 367-370.
- [13] Li R, Ma W, 2014, Road Average Speed Fusion Method Based on BP Neural Network and D-S Evidence Theory. *Journal of Traffic and Transportation Engineering*, 2014(5): 111-118.
- [14] Guo JH, Xia JX, Smith BL, 2009, Kalman Filter Approach to Speed Estimation Using Single Loop Detector Measurements Under Congested Condition. *Journal of Transportation Engineering*, 135(12): 927-934.
- [15] Zhang C, Yang X, 2006, The Second China Intelligent Transportation Annual Conference, December 12, 2006: Road Traffic Information Acquisition System Based on Floating Vehicle. Beijing, 108-112.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.