Journal of Highway and Transportation Research and Development
2024, 18(4): 42−50

Research Article
ISSN 2095−6215
DOI: 10.26599/HTRD.2024.9480031

Vehicle spatial and time trajectory filling based on dynamic road network

Ruojian Li[1*], Hongsheng Qi[2**]

1. *Polytechnic Institute & Institute of Intelligent Transportation Systems, Zhejiang University, Hangzhou, Zhejiang 310058, China*;
2. *College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Abstract: Complete vehicle trajectory data is essential for urban traffic flow modeling studies. This study proposes a framework for filling vehicle trajectories in spatial and time for automatic vehicle identification (AVI) data. Based on the particle filter, the dynamic correction factor is innovatively used to improve algorithm accuracy. After four resamplings, such as traffic situation index and traffic event factor, spatial trajectory filling is completed. The Copula function fills the time trajectory by analyzing the correlation between upstream and downstream paths. Finally, the experiment was conducted in Xiaoshan District, Hangzhou, China. The results show that for spatial trajectory filling, the average accuracy exceeds 97% with 75% camera coverage. In time trajectory filling, the time trajectory filling error is reduced by 35% compared to the Hellinga algorithm.

Keywords: traffic engineering; particle filtering; Copula functions; dynamic traffic conditions; automatic vehicle identification

1 Introduction

Automated vehicles play an essential role in improving road safety, increasing traffic efficiency, and promoting the development of intelligent mobility. On urban roads, vehicles using autonomous driving technology and traditional vehicles driven by humans will run together for a long time, resulting in high costs for traffic flow data acquisition[1]. At present, automatic vehicle identification systems are commonly installed at intersections of large cities in China. Compared with other traffic collection sources, AVI data has the advantages of easy access, rich data format, and long collection duration. However, due to equipment constraints and other factors, there are incomplete, duplicate, and unavailable data that defy the laws of physics. The results of filling in the spatial and time trajectories of vehicles can provide data support for studies such as the modeling of mixed traffic flows at this stage.

This study proposes a framework for filling vehicle time and spatial trajectory data, which uses a dynamic particle filtering algorithm to accomplish vehicle spatial trajectory filling. Existing studies, with limited capacity for urban traffic data collection and analysis, sample the importance of particles through static road network structural features, such as gravity flow models and journey time consistency, without considering dynamic road network operations-event data for resampling[2]. The average spatial filling accuracy exceeds 97%, which was significantly improved compared with the static particle filter model.

This study uses the Copula function to fill vehicle track time. Existing studies perform vehicle trajectory time filling for floating vehicle data and do not consider the correlation between upstream and downstream traffic on the path[3]. In the study, it is found that there are many phenomena in the upstream and downstream traffic flow of urban road networks. Modeling the correlation between upstream and downstream traffic flow can effectively improve the accuracy of vehicle time trajectory filling. This study uses dynamic road network operation

Received: 25 April 2024; **Revised:** 9 August 2024; **Accepted:** 25 August 2024

* *E-mail address*: liruojian@zju.edu.cn

** *Corresponding author*: qihongsheng@zju.edu.cn

characteristics to optimize vehicle congestion probability and parking probability calculation on urban roads. The correlation between upstream and downstream traffic flows on urban roads is also described using Copula functions to complete the vehicle time trajectory filling. The filling method using Copula functions has been tested to reduce errors by over 35% compared to the Hellinga algorithm.

The content of this paper is as follows: Chapter 2 provides an overview of existing research on vehicle trajectory filling. Chapter 3 presents the components of the trajectory space and time-filling framework. Chapter 4 uses accurate AVI data from Xiaoshan District, Hangzhou, China, and introduces the evaluation details and results. Chapter 5 summarises the conclusions drawn from this study and suggests future research directions.

2 Literature review

Researchers in different countries have carried out several studies related to the filling of vehicle trajectories. Feng et al.[2] used particle filter theory to estimate vehicle trajectories in combination with five trajectory correction factors (i.e., path consistency, AVI measurability criterion, travel time consistency, gravity flow model, and path-link flow matching model). Tang et al.[4] added a degree of randomness to the resampling process of correction factors containing subjective factors in the particle filter to simulate the subjective behavior of vehicles. Qi et al.[5] used an improved K-Shortest Path (KSP) algorithm to generate candidate trajectories based on spatiotemporal prism theory and proposed decision metrics such as travel time consistency and path preference as multiple attributes in the self-encoder model to fill incomplete vehicle paths. Chen et al.[6] built an extended variational solution network and used Kalman filtering to characterize the stochastic property of queueing boundary curves to fill the vehicle trajectories at the intersection. Xiao et al.[7] propose a bidirectional RNN-based algorithm for reconstructing personal car trajectories by introducing neural arithmetic logic units (NALUs) into the trajectory filling model to enhance the extrapolation capability of neural networks and ensure the accuracy of trajectory prediction. Zhang et al.[8] use hybrid attention consisting of multi-headed attention and dynamic convolution to learn the contextual structure of intersections in vehicle trajectories to fill incomplete vehicle trajectories. The decision features used in the above studies are based on static roads network structures such as lane width and number, ignoring the dynamic operation of the road network and the dynamic characteristics of vehicles operating on the road network. This paper proposes dynamic and static correction factors by describing the static and dynamic road network operation characteristics. The refined description of the road network traffic operation state effectively improves the accuracy of spatial trajectory filling.

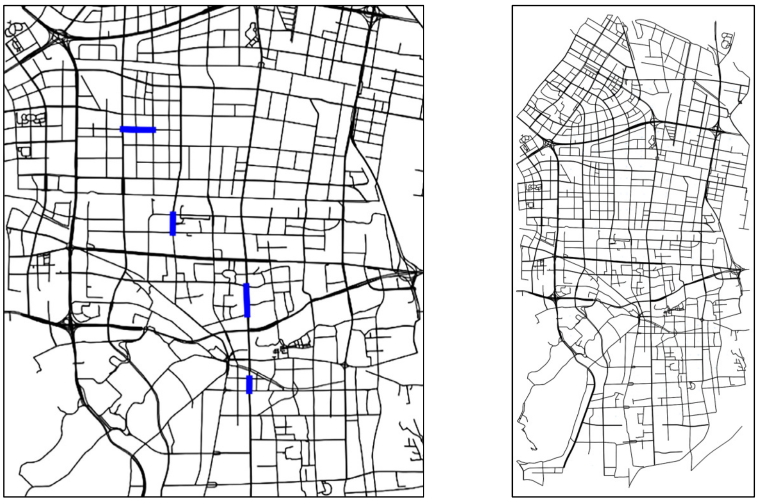
Chen et al.[9] used a Copula function to model the road segment travel time distribution based on spatial correlation. The Copula function method has advantages over convolutional and empirical distribution fitting methods that do not capture the correlation of road segments. Hellinga et al.[3] define a likelihood function based on the stopping probability and congestion of vehicles on each road section, and obtain the most reasonable travel time for the road section by solving for the maximum value of the function. Tang et al.[10] use a third-order tensor to model the travel times of different road sections under different traffic conditions during specific periods. The missing entries in both tensors were estimated by a contextual tensor factorization method to obtain the travel times for any road section in the urban road network. Most of the above studies are data-driven travel time estimation methods, which require simulation of the entire regional road network to obtain data on actual traffic conditions, such as signal timing. In practice, signal timing data is challenging to obtain, making the above method challenging to apply to large-scale urban road networks. The study found that the vehicle time trajectory filling error could be effectively reduced by modelling the upstream and downstream traffic travel time distribution in the absence of traffic signal timing data. This study proposes using Copula functions to model the correlation between upstream and downstream traffic flows, further improving the accuracy of filling vehicle time trajectories.

3 Methodology

3.1 Description of the problem

The framework for filling the spatial and temporal trajectory of the vehicle is briefly described below, as shown in Figure 1(a). The blue bolded line segment in

Figure 1 shows the vehicle trajectory information recorded in the AVI data, including the license plate number, the number of the intersection, and the time of passage through the corresponding intersection. AVI has difficulty capturing complete vehicle trajectories due to factors such as intersection camera coverage and license plate recognition rates. As the network size increases, the trajectory-filling algorithm becomes more complex. The main work in this paper is to complete the filling of the spatial and time trajectories of the vehicles. The complete trajectory data can be used as a basis for subsequent studies.



(a) Example of vehicle trajectory (b) Xiaoshan Road network, China

Fig. 1 Vehicle trajectory filling area

3.2 General framework

This study aims to fill in the AVI data. The framework for filling vehicles' time and spatial trajectories is outlined below. The framework for filling trajectories consists of two parts. The first part completes the vehicle space trajectory filling. The particle set is first initialized using the depth-first algorithm. The actual state-space probabilities of these particles are then updated using two dynamic correction factors and two static correction factors. These correction factors are the traffic pattern index, the traffic event situation, the path size, and the path consistency factor. After four resamplings, the trajectory with the highest particle aggregation number results from filling the vehicle spatial trajectory.

The second part completes the vehicle time trajectory filling. Stopping probabilities and congestion levels on road sections are modeled based on dynamic road network operating characteristics. Each road section's most probable travel time is obtained by solving for the maximum likelihood function. The correlation between upstream and downstream traffic is then modeled using the Copula function to complete the vehicle time trajectory filling.

3.3 Module 1: Particle filtering module

In particle filtering theory, sets of particles are usually built to represent particles with different state probabilities. The distribution is represented by extracting particles with random states from the posterior probabilities. The random variables in the non-parametric particle filtering algorithm do not have to obey a Gaussian distribution. Particle filtering has a solid ability to model parameters in non-linear systems[2].

3.3.1 Initializes the set of particles

This study obtains a sufficient number of paths using the depth-first algorithm as the initial particle set. The prior probability of the particles is then set to $1/I$, where I is the number of particles in the particle set.

3.3.2 Importance sampling

This study assumes that the importance of initial particles obeys uniform distribution. The sampling equation of the importance of path consistency is shown in Equation (1).

$$W_1^{(i)} = W_0^{(i)} \times \frac{P\left[y_{t:t+\Delta t} \mid \hat{x}_1^{(i)}\right] P\left[\hat{x}_1^{(i)} \mid x_0^{(i)}\right]}{q\left[\hat{x}_1^{(i)} \mid x_0^{(i)}, y_{t:t+\Delta t}\right]}, \tag{1}$$

where the explanation of the variables is shown in *List of symbols and abbreviations*.

$$N_1^{(i)} = \frac{W_1^{(i)}}{\sum_{i=1}^{N} W_1^{(i)}} \times N \tag{2}$$

After the path consistency resampling, two dynamic correction factor resamples and one static correction factor resample are completed. The trajectory is filled with the maximum number of particles aggregated.

3.3.3 Dynamic correction factor

Different from using static road network characteristics as resampling indicators, this study proposes to resample the importance of particles from dynamic road network operation characteristics to obtain trajectory-filling results that match the actual road network operation. This study proposes two dynamic correction factors: the traffic situation index and the traffic event factor.

(1) Traffic Situation Index. Within a local road network, traffic flow on a congested section is more significant than on a clear section. The more likely the section is to be chosen by drivers, the more likely it is to be an actual trajectory. There are four traffic dynamics levels: unimpeded, slowdown, congestion, and heavy congestion. The traffic pattern is described in terms of T_{TS} and is defined in Equation (3).

$$I_{\mathrm{TS}} = \frac{\sum_{j\in C_4} \delta_j S_j}{S_{\text{heavy congestion}}}. \tag{3}$$

When the I_{TS} value of a road section is higher, the higher the traffic flow on the road section, the more vehicles will drive over that road section, and the more likely the trajectory will be an actual trajectory. A logit model describes the traffic dynamics in Equation (4).

$$P_{\mathrm{si}}^{i} = \frac{1}{1+\mathrm{e}^{-(\beta_0+\beta_1 I_{\mathrm{TS}})}}. \tag{4}$$

(2)Traffic Event Factor. Urban roads are often subject to traffic control for reasons such as construction work and the holding of significant events. Alternate trajectories should not pass through traffic-controlled roads. Historical road event information includes road name, event type, event start time, and expected event end time. The probabilistic update process follows an all-or-nothing allocation. The probability update formula is given in Equation (5).

$$\begin{cases} P_{\mathrm{in}}^{i} = 0, \text{ when } N_{\mathrm{pt}} \in N_{\mathrm{in}}, \\ P_{\mathrm{in}}^{i} = 1, \text{ when } N_{\mathrm{pt}} \notin N_{\mathrm{in}}. \end{cases} \tag{5}$$

3.3.4 Static correction factor

This study completes static importance sampling regarding path size and consistency factors.

Path Consistency Factor. Partial trajectories captured by AVI data exist in the set of possible path trajectories, so path consistency analysis aims to match partial trajectories to complete trajectories. The probability update formula is as follows.

$$\begin{cases} P_{N_{\mathrm{ct}}}^{i} = 1, \text{ when } N_{\mathrm{pt}} \in N_{\mathrm{ct}}, \\ P_{N_{\mathrm{ct}}}^{i} = 0, \text{ when } N_{\mathrm{pt}} \notin N_{\mathrm{ct}}. \end{cases} \tag{6}$$

Path Size The index path size was proposed by Tang et al.[11] in the path size logit model. P_{ps}^{i} was used to describe the path size of possible trajectory i, as defined in Equation (8).

$$P_{\mathrm{ps}}^{i} = \sum_{k\in\Gamma_i} \left(\frac{1}{L_i}\right) \frac{1}{\sum_{j\in C_n} \delta_{kj}}, \tag{7}$$

$$P_{\mathrm{ps}}^{i} = \begin{cases} \exp\left(\frac{1}{2} \times \frac{\min_{j\in M}\left(S_{\mathrm{ps}}^{i}\right) - S_{\mathrm{ps}}^{i}}{\max_{j\in M}\left(S_{\mathrm{ps}}^{i}\right) - \min_{j\in M}\left(S_{\mathrm{ps}}^{i}\right)}\right), \max_{j\in M}\left(S_{\mathrm{ps}}^{i}\right) \neq \min_{j\in M}\left(S_{\mathrm{ps}}^{i}\right), \\ 1, \quad \max_{j\in M}\left(S_{\mathrm{ps}}^{i}\right) = \min_{j\in M}\left(S_{\mathrm{ps}}^{i}\right). \end{cases} \tag{8}$$

3.4 Module 2: copula function module

This section describes the work involved in filling in vehicle time trajectories. Firstly, AVI data describes the method of assigning travel times between intersections.

The path travel time can be divided into three components, namely the free-flow travel time $\tau_{\mathrm{f}}\left(l_{k,i,j}\right)$, the stopping time $\tau_{\mathrm{s}}\left(l_{k,i,j}\right)$ and the congestion time $\tau_{\mathrm{c}}\left(l_{k,i,j}\right)$. The calculation of the path travel time between two consecutive data collection points is shown in Equation (9).

$$t_{k,i+1} - t_{k,i} = \sum_{j=0}^{|J(ki)|} \left\{\tau_{\mathrm{f}}\left(l_{k,i,j}\right) + \tau_{\mathrm{s}}\left(l_{k,i,j}\right) + \tau_{\mathrm{c}}\left(l_{k,i,j}\right)\right\}. \tag{9}$$

Considering that the AVI data collection point is the beginning or end of the road section, this study reduces the power of the calculation method for congestion time $\tau_{\mathrm{c}}\left(l_{k,i,j}\right)$ and stopping time $\tau_{\mathrm{s}}\left(l_{k,i,j}\right)$. The traffic pattern index I_{TS} mentioned above is also used to effectively improve the accuracy and operational efficiency of the algorithm.

The calculation of $\tau_{\mathrm{c}}\left(l_{k,i,j}\right)$, $H_{\mathrm{s}}\left(l_{k,i,j}, w\right)$, and $Q_{\mathrm{s}}(k,i)$ is carried out using the traffic situation index I_{TS} so that the calculation results are in line with the operational characteristics of the dynamic traffic road network. The formula for calculating the free-flow travel time τ_{f} is shown in Equation (10).

$$\tau_{\mathrm{f}}=\frac{|l(n_a,n_b)|}{S_{\mathrm{f}}(n_a,n_b)}. \tag{10}$$

The congestion time $\tau_{\mathrm{c}}\left(l_{k,i,j}\right)$ on road segment j is calculated as shown in Equation (11).

$$\tau_{\mathrm{c}}\left(l_{k,i,j}\right)=\int_0^{I_{\mathrm{TS}}}\delta_{k,i,j}\tau_{\mathrm{c}}\frac{\sum_{j=0}^{|J(k,i)|}P_{\mathrm{w}}(k,i,w)P_{\mathrm{s}}(l_{k,i,j},w)}{Q_{\mathrm{s}}(k,i)}\mathrm{d}w, \tag{11}$$

where w is the congestion level indicator, as defined in Equation (12).

$$w=\frac{\sum_{j=0}^{|J(k,i)|}\{\tau_{\mathrm{c}}(l_{k,i,j})\}}{\sum_{j=0}^{|J(k,i)|}\{\tau_{\mathrm{c}}\left(l_{k,i,j}\right)+\tau_{\mathrm{f}}(l_{k,i,j})\}}. \tag{12}$$

The stopping time $\tau_{\mathrm{s}}\left(l_{k,i,j}\right)$ on road segment j is calculated as shown in Equation (13).

$$\tau_{\mathrm{s}}\left(l_{k,i,j}\right)=\tau_{\mathrm{s}}\frac{P_{\mathrm{w}}(k,i,w)P_{\mathrm{s}}(l_{k,i,j},w)}{Q_{s}(k,i)}, \tag{13}$$

where τ_{s} is the total stopping time on path $r_k(t_{k,i},t_{k,i+1})$ and is calculated as shown in Equation (14).

$$\tau_{\mathrm{s}}=t_{k,i+1}-t_{k,i}-\sum_{j=0}^{|J(k,i)|}\left\{\tau_{\mathrm{f}}\left(l_{k,i,j}\right)\right\}-\tau_{\mathrm{c}}. \tag{14}$$

The other symbols in Equation (10) are defined as shown in Equation (15) to Equation (20).

$$\delta_{k,i,j}=\frac{\tau_{\mathrm{f}}\left(l_{k,i,j}\right)}{\sum_{j=0}^{|J(k,i)|}\{\tau_{\mathrm{f}}(l_{k,i,j})\}}, \tag{15}$$

$$\tau_{\mathrm{c}}=\frac{w}{1-w}\sum_{j=0}^{|J(k,i)|}\{\tau_{\mathrm{f}}(l_{k,i,j})\}, \tag{16}$$

$$P_{\mathrm{w}}(k,i,w)=\min\left[1,\frac{1}{w}\frac{T_{\mathrm{c}}\left[k,I_p(i)\right]+T_{\mathrm{c}}(k,i)}{t_{k,I_p(i)+1}-t_{k,I_p(i)}+t_{k,i+1}-t_{k,i}}\right], \tag{17}$$

$$P_{\mathrm{s}}\left(l_{k,i,j},w\right)=H_{\mathrm{s}}(l_{k,i,j},w)\prod_{l\neq J}\left\{1-H_{\mathrm{s}}(l_{k,i,j},w)\right\}, \tag{18}$$

$$H_{\mathrm{s}}\left(l_{k,i,j},w\right)=w(1-w)(1-\mathrm{e}^{-w}), \tag{19}$$

$$Q_{\mathrm{s}}(k,i)=\int_0^{I_{\mathrm{TS}}}P_{\mathrm{w}}(k,i,w)\sum_{j=0}^{J(k,i)}P_{\mathrm{s}}(l_{k,i,j},w)\mathrm{d}w. \tag{20}$$

The calculation of the travel time on each section j in the route $r_k(t_{k,i},t_{k,i+1})$ is given in Equation (21).

$$t\left(l_{k,i,j}\right)=\tau_{\mathrm{f}}\left(l_{k,i,j}\right)+\tau_{\mathrm{s}}\left(l_{k,i,j}\right)+\tau_{\mathrm{c}}\left(l_{k,i,j}\right),\quad j=1,2,3,\cdots,|J(k,i)|. \tag{21}$$

The following introduces the modeling of the correlation between upstream and downstream traffic flow on urban roads using the Copula function.

The joint distribution $F(X_1,X_2,\cdots,X_n)$ of any n variables can be described by a series of marginal distributions and a deterministic Copula function, with the probability density equation given in Equation (22).

$$f(x_1,x_2,\ldots,x_n;\theta)=C[F_1(x_1,\theta_1),F_2(x_2,\theta_2),\cdots,F_n(x_n,\theta_n);\theta_n]\prod_{i=1}^{n}f_i(x_i;\theta_i), \tag{22}$$

where, $f_i(x_i)$ is the edge distribution function; $F_i(x_i)$ is the cumulative probability distribution function corresponding to the edge distribution; θ is the set of parameters of the Copula function; $C(\cdot)$ is the Copula density function, which is given in Equation (23).

$$C[F_1(x_1),F_2(x_2),\cdots,F_n(x_n)]=\frac{\partial C[F_1(x_1),F_2(x_2),\cdots,F_n(x_n)]}{\partial F_1(x_1)\partial F_2(x_2),\cdots,\partial F_n(x_n)}. \tag{23}$$

Constructing a Copula function is generally a two-step process. Firstly, the road section travel time distribution needs to be determined and fitted to the travel time distribution of each road section. Experimentally, some sections of Hangzhou Xiaoshan District travel times follow the LogLaplace distribution, Beta distribution, StudentT distribution, or Gamma distribution. Secondly, an appropriate Copula function is selected to model the correlation between the traffic flow on the upstream and downstream sections. The Gaussian and Frank Copula functions fit the adjacent section journey time distribution better[12].

Finally, the actual travel times and Copula functions are used for sampling. Vehicle time trajectory filling results matching upstream and downstream traffic characteristics are obtained. At this point, this chapter presents a complete framework for filling vehicle spatial and time.

4 Case study

In order to calibrate the effectiveness of the framework proposed in this study for filling the spatial and temporal trajectories of vehicles in real road networks, experiments were conducted using the AVI dataset from Xiaoshan District, Hangzhou, China. The time interval for the data was from 1 to 7 March 2019, and the experimental area is shown in Figure 1(b), with a total of 1,412 traffic cameras installed within the area. This chapter first introduces the pre-processing data methods and then tests the effect of filling in the vehicle's spatial and temporal trajectories, respectively. The data source for the traffic situation index and traffic event factor is AMAP.

4.1 Data pre-processing

This study uses software such as pandasql to pre-process the data. The raw data consists of vehicle information tables and camera information tables. The vehicle information table includes the license plate number, the time of passing the intersection, the vehicle type, and the camera ID. The camera information table includes the camera ID and coordinates of the intersection where the camera is located. The trajectory of each vehicle, as captured by the traffic cameras, is first obtained from the vehicle information sheet. The coordinates of the cameras are then matched to the nodes of the road network to obtain a list of intersection numbers through which the vehicles have passed.

Due to the limitations of image recognition technology, the number plates of some vehicles cannot be captured clearly, resulting in possible errors in plate number matching and abnormal path travel time data. The paper eliminates paths where the time interval between vehicles passing through 2 intersections is more than 10 min.

4.2 Spatial trajectory filling experiments

The complete vehicle trajectory is first extracted from the AVI dataset. To verify the effectiveness of the vehicle spatial trajectory filling algorithm, it is assumed that some of the intersections need to be equipped with camera devices. For example, when the coverage is 50%, 50% of the intersections are without camera equipment, which in this experiment is 706 intersections. These intersections without cameras were removed from the complete vehicle track to generate test samples. The trajectory filling of the test sample is then completed using the vehicle spatial trajectory filling method.

To verify the stability of the vehicle spatial trajectory filling method, a total of 7 days of data from March 1st to March 7th were selected for this study. One hundred vehicle trajectories were selected for testing each day, and 700 sets of test results were obtained. To test the performance of the algorithm at different camera coverage rates, this study completed experiments at 75%, 65%, 55%, and 45% coverage rates, with accuracy rates of 97.351%, 92.159%, 89.595%, and 85.919%, respectively. The results are shown in Figure 2. The horizontal coordinate indicates the trajectory number. The vertical coordinate indicates the accuracy of the trajectory filling.

To verify the effectiveness of the algorithm, each experiment was set to the same conditions, and calculations were performed using the shortest path algorithm and the static particle filtering algorithm proposed by Feng[2]. The path-filling accuracy of the shortest path algorithm and the static particle filtering algorithm was 87.092% and 81.88%, respectively, for 75% camera coverage, and the results are shown in Figure 3. The results illustrate that the particle filtering algorithm proposed in this study using dynamic correction factors for resampling has higher accuracy and stability and works better in practical applications.

4.3 Time trajectory filling experiments

The experimental results of vehicle time trajectory filling are presented below. The experimental results of the vehicle time trajectory filling algorithm proposed in this paper are compared with the real travel time, and the time error of the trajectory to be filled is obtained.

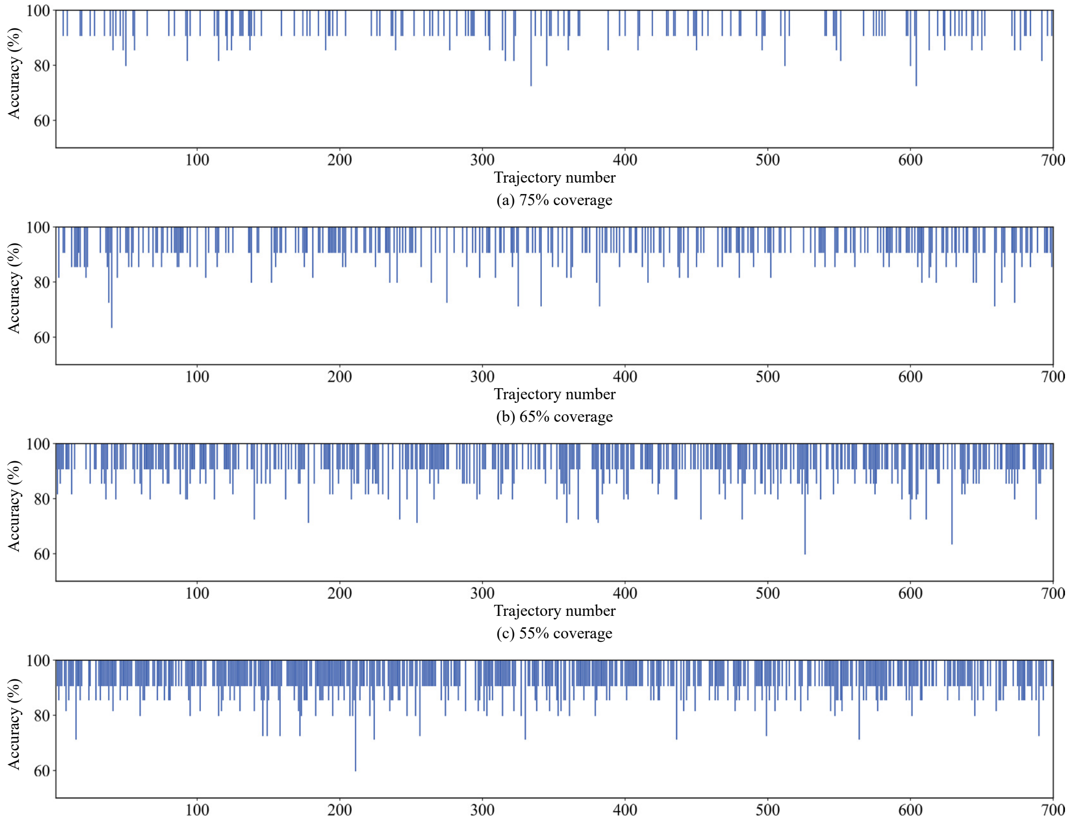
$E_{l(n_a,n_b)}$ is used to measure the accuracy of the time trajectory filling results and is defined in Equation (24).

$$E_{l(n_a,n_b)} = \frac{1}{T_{\mathrm{AT}l(\mathrm{n_a},\mathrm{n_b})}} \sqrt{\frac{\sum_{r=1}^{N}(T_{\mathrm{A}r} - T_{\mathrm{T}r})^2}{N}}. \tag{24}$$

The test data use 514 tracks filled out completely at 75% camera coverage. The calculations were performed using the Hellinga algorithm and the time trajectory filling method using the Copula function with an average error of 0.676 and 0.433, respectively. The experimental results are shown in Figure 4. The horizontal coordinate is the trajectory number. The vertical coordinate is the result of the calculation of $E_{l(n_a,n_b)}$.

It should be noted that the image is divided into four sections, representing the four intervals of the track travel time. The travel time of the tracks in the four zones is less than 120 s, 240 s, 360 s, 480 s, and 591 s, respectively. When the travel time is less than 120 s, the error increases rapidly with the increase in the travel time.

The experimental results show that compared with the Hellinga algorithm, the average error of the time trajectory filling method using the Copula function is reduced by more than 35%, which greatly improves the accuracy of the algorithm. At the same time, the physical mean-



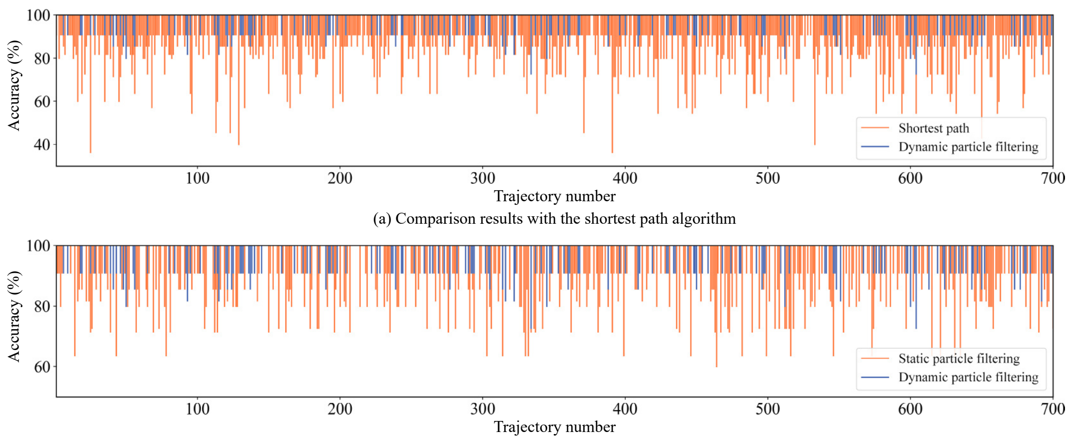
(a) 75% coverage

(b) 65% coverage

(c) 55% coverage

(d) 45% coverage

Fig. 2 Vehicle spatial trajectory filling accuracy



(a) Comparison results with the shortest path algorithm

(b) Comparison results with the static particle filtering algorithm

Fig. 3 Comparison of vehicle spatial trajectory filling accuracy

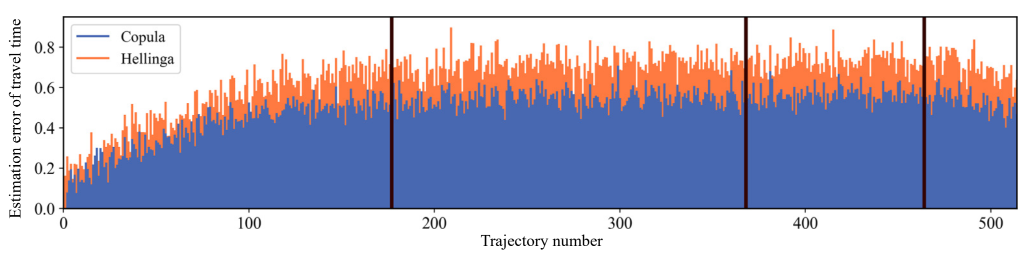


Fig. 4 Time trajectory filling error

ing of this method is clear and has a good interpretation. It not only considers the parking probability of vehicles in each section but also considers the degree of congestion in the section. At the same time, the method is simple to calculate and does not depend on additional data such as traffic conditions and signal timing. Therefore, it has good adaptability to travel time modeling and can be used to estimate the travel time of large, medium, and small urban road networks.

5 Conclusions and future work

(1) This study proposes a framework for filling vehicle trajectories in spatial and time for AVI data. In terms of vehicle spatial trajectory filling, the resampling process of the particle filtering algorithm was completely redesigned using the traffic situation index and traffic event factor as dynamic correction factors. The experimental results show that the particle filtering algorithm using dynamic correction factors is extremely accurate in filling in the spatial trajectories of vehicles. In an open road network environment with 75% camera coverage, the average accuracy exceeds 97%, which is a substantial improvement compared to the static particle filter model.

(2) In terms of vehicle time trajectory filling, the probability of stopping and congestion on a road section is first modeled based on the operational characteristics of the dynamic road network. The most probable travel time for each road section is obtained by solving the maximum likelihood function. This study proposes to use the Copula function to describe the correlation between the upstream and downstream traffic flows to complete the filling of the vehicle time trajectory. The experimental results show that the error is reduced by more than 35% compared to the results using the Hellinga algorithm. Vehicle time trajectory filling using the Copula function is simple and does not rely on data such as signal timing. This method is suitable for filling urban vehicle time trajectories of all sizes.

(3) The main shortcoming of the framework for filling trajectories proposed in this study is demonstrated in the area of time filling. The test results show a decrease in accuracy as the number of vehicles passing through the intersection increases. Future research could combine the construction of smart cities and digital twins to optimize time trajectory-filling algorithms using data-driven techniques.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgements

This study was supported by National Natural Science Foundation of China Key Project (52131202); National Natural Science Foundation of China General Project (52272314); Humanities and Social Sciences General Project of the Ministry of Education (21YJCZH116); Zhejiang Public Welfare Technology Research Program (LGF22E080007).

List of symbols and abbreviations

$W_1^{(i)}$	Non-normalized weights of possible trajectory $i$after path consistency update.
$W_0^{(i)}$	Initial a priori weights for possible trajectory i. If no historical trajectory data are available, the weights follow a uniform distribution.
$y_{t:t+\Delta t}$	Data collected from time t to time $t+\Delta t$, where t is the time at which data collection begins and Δt is the dynamic value of the vehicle path filling calculation interval.
$x_0^{(i)}$	Particles with an initial possible trajectory of i.
$P[y_{t:t+\Delta t}\vert\overset{\frown}{x}_1^{(i)}]$	The probability of a possible trajectory $i$after importance sampling based on path consistency.
$P\left[\left\vert\overset{\frown}{x}_1^{(i)}\right\vert x_0^{(i)}\right]$	The probability of transformation from prior data to path consistency.
$q\left[\overset{\frown}{x}_1^{(i)}\vert x_0^{(i)},y_{t:t+\Delta t}\right]$	A priori probability density function.
$N_1^{(i)}$	Number of particle aggregates for possible trajectory $i$after path consistency calculation
N_{pt}	The set of nodes in a partial trajectory
N_{ct}	Nodes set in a complete trajectory
N_{sct}	Serial number of the complete track
I_{TS}	Road segment traffic situation index
C_4	The four traffic conditions, namely unimpeded, slowdown, congestion and heavy congestion
δ_j	Traffic situation rating as a percentage of the overall roadway
S_j	Traffic situation factor
P_{in}^i	The posterior probability of possible trajectory i after resampling of the traffic event factor
N_{in}	The road number where the traffic incident occurred
P_{si}^i	The posterior probability of a possible trajectory i after resampling of the traffic situation index
S_{sD}	The size of the path dimension of possible trajectory i

Γ_i	The set of all sections comprising possible trajectory i
l_k	The length of road segment k
L_i	The length of section i of the possible trajectory
P^i_{ps}	The posterior probability of a possible trajectory i based on the path size criterion
S^i_{ps}	The size of the path size of possible trajectory i
M	The set of all possible trajectories
$\lvert J(k,i)\rvert$	Total number of sections that make up path $r_k(t_{k,i},t_{k,i+1})$
n_a	Start of section n
n_b	End of section n
$S_{\mathrm{f}}(n_a,n_b)$	The free-flow velocity on the road segment $l(n_a,n_b)$
$\lvert l(n_a,n_b)\rvert$	The length of the road segment $l(n_a,n_b)$
$T_{\mathrm{AT}l(n_a,n_b)}$	True path average travel times
T_{Ar}	Calculated journey times for the section
T_{Tr}	Real road journey times
$P_{\mathrm{w}}(k,i,w)$	Congestion probability
$P_{\mathrm{s}}(l_{k,i,j},w)$	Parking probability
$I_{\mathrm{p}}(i)$	The largest integer smaller than i
τ_{s}	Total stopping time on path $r_k(t_{k,i},t_{k,i+1})$

References

[1] Nieuwenhuijsen, J., de Almeida Correia G.H., Milakis, D., et al. Towards a quantitative method to analyze the long-term innovation diffusion of automated vehicles technology using system dynamics[J]. Transportation Research Part C: Emerging Technologies, 2018, 86: 300–327.

[2] Feng, Y., Sun, J., Chen, P. Vehicle trajectory reconstruction using automatic vehicle identification and traffic count data[J]. Journal of Advanced Transportation, 2015, 49(2): 174–194.

[3] Hellinga, B., Izadpanah, P., Takada, H., et al. Decomposing travel times measured by probe-based traffic monitoring systems to individual road segments[J]. Transportation Research Part C: Emerging Technologies, 2008, 16(6): 768–782.

[4] Tang, J.J., Wang, K., Ye, J.Q. Vehicle trajectory reconstruction based on multi-criteria particle filter algorithm[J]. Journal of Railway Science and Engineering, 2022, 19(8): 2407–2416.

[5] Qi, X.Y., Ji, Y.J., Li, W.H., et al. Vehicle trajectory reconstruction on urban traffic network using automatic license plate recognition data[J]. IEEE ACCESS, 2021, 9: 49110–49120.

[6] Chen, P., Wei, L., Meng, F.F., et al. Vehicle trajectory reconstruction for signalized intersections: a hybrid approach integrating kalman filtering and variational theory[J]. Transportmetrica B Transport Dynamics, 2020, 9(1): 1–20.

[7] Xiao, Z., Qian, X., Jiang, H.B., et al. Bidirectional RNN-based private car trajectory reconstruction algorithm[J]. Journal on Communications, 2020, 41(12): 171–181.

[8] Zhang, Y.H., Zheng, J.L., Pu, L., et al. Vehicle trajectory reconstruction method driven by traffic camera data[J]. Journal of Chongqing University of Technology(Natural Science), 2022, 36(8): 193–202.

[9] Chen, P., Zeng, W.L., Chen, M., et al. Modeling arterial travel time distribution by accounting for link correlations: a copula-based approach[J]. Journal of Intelligent Transportation Systems, 2019, 23(1): 28–40.

[10] Tang, K., Chen, S.Y., Liu, Z.Y. Citywide spatial-temporal travel time estimation using big and sparse trajectories[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(12): 4023–4034.

[11] Tang, J.J., Wang, Y.W., Hao, W., et al. A Mixed path size logit-based taxi customer-search model considering spatio-temporal factors in route choice[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(4): 1347–1358.

[12] Yu, Y. Urban traffic state evaluation methods based on automatic number plate recognition data[D]. Hangzhou: Zhejiang University, 2022.