Origin-destination pattern estimation based on trajectory reconstruction using automatic license plate recognition data

Wenming Rao\textsuperscript{a}, Yao-Jan Wu\textsuperscript{b}, Jingxin Xia\textsuperscript{a,⁎}, Jishun Ou\textsuperscript{a}, Robert Kluger\textsuperscript{c}

\textsuperscript{a}Intelligent Transportation System Research Center, Southeast University, No. 35 Jinxianghe Road, Xuanwu District, Nanjing 210096, PR China
\textsuperscript{b}Department of Civil Engineering and Engineering Mechanics, The University of Arizona, 1209E. 2nd St., Tucson, AZ 85721, USA
\textsuperscript{c}J.B. Speed School of Engineering – Civil & Environmental Engineering Dept., University of Louisville, 132 Eastern Parkway, Louisville, KY 40292, USA

ARTICLE INFO

Keywords:
- OD pattern
- Trajectory reconstruction
- ALPR
- Particle filter
- Time geography

ABSTRACT

Origin-destination (OD) pattern estimation is a vital step for traffic simulation applications and active urban traffic management. Many methods have been proposed to estimate OD patterns based on different data sources, such as GPS data and automatic license plate recognition (ALPR) data. These data can be used to identify vehicle IDs and estimate their trajectories by matching vehicles identified by different sensors across the network. OD pattern estimation using ALPR data remains a challenge in real-life applications due to the difficulty in reconstructing vehicle trajectories. This paper proposes an offline method for historical OD pattern estimation based on ALPR data. A particle filter is used to estimate the probability of a vehicle’s trajectory from all possible candidate trajectories. The initial particles are generated by searching potential paths in a pre-determined area based on the time geography theory. Then, the path flow estimation process is conducted through dividing the reconstructed complete trajectories of all detected vehicles into multiple trips. Finally, the OD patterns are estimated by adding up the path flows with the same ODs. The proposed method was implemented on a real-world traffic network in Kunshan, China and verified through a calibrated microscopic traffic simulation model. The results show that the MAPEs of the OD estimation are lower than 19%. Further investigation shows that there exists a minimum required ALPR sampling rate (60% in the test network) for accurately estimating the OD patterns. The findings of this study demonstrate the effectiveness of the proposed method in OD pattern estimation.

1. Introduction

Origin-Destination (OD) demands describe the distribution of trips between each origin-destination pair across a traffic network. Generally, OD demands are time-varying and influenced by stochastic fluctuation of traffic flow. However, many traffic applications (e.g. transportation planning) based on long-term OD demand require the stable distribution of OD demands in certain periods (e.g. a week, a month, or half a year). Therefore, OD patterns, sometimes called regular demand patterns (Mahmassani and Zhou, 2005; Zhou and Mahmassani, 2007), are used to represent the general distribution patterns in time-dependent OD demands. Since OD patterns are not affected by unique events, severe weather and noncurrent incidents, these patterns are often used as preference (prior) trip matrices for dynamic OD estimation (Mínguez et al., 2010; Lu et al., 2015).

Neither OD demand nor OD patterns are directly observable. Traditionally, the OD demands and patterns were derived from

⁎ Corresponding author.

E-mail addresses: raowenming@seu.edu.cn (W. Rao), yaojian@email.arizona.edu (Y.-J. Wu), xiajingxin@seu.edu.cn (J. Xia), jishun@seu.edu.cn (J. Ou), robert.kluger@louisville.edu (R. Kluger).

https://doi.org/10.1016/j.trc.2018.07.002

Received 27 December 2017; Received in revised form 28 June 2018; Accepted 4 July 2018

0968-090X/ © 2018 Elsevier Ltd. All rights reserved.
traffic surveys, a labor intensive and time-consuming process, making it only feasible in small networks. Subsequently, with the new development of traffic detection technology, many methods were proposed for OD demand or patterns estimation using traffic flow variables (e.g. link volume and turning movement) collected by various traffic sensors, but the estimation accuracy is limited because the origins and destinations of trips are difficult to observe using detectors. Depending on the data source, OD estimation methods can be generally classified into two categories: the fixed-sensor-based method and the trajectory-based method.

The fixed-sensor-based method estimates or predicts OD demands based on traffic flow variables collected from fixed sensors such as inductive loops, microwave detectors and video sensors. Some of these methods are non-assignment-based, meaning OD demands are estimated based on the relationship between inbound and outbound flows of traffic network and the law of traffic volume conservation (Chang and Wu, 1994; Lin, 2006). However, the non-assignment based methods are not able to describe the complex route choice behaviors, thus these methods only can be applied to a closed network (e.g. a simple freeway network). Comparatively, for a real-world urban network, studies are typically assignment-based, using a static or dynamic traffic assignment process to describe the relationship between OD demand and observed traffic flow. Commonly used models for OD demand estimation include the generalized least square model (GLS) (Cascetta et al., 1993; Sherali and Park, 2001), the maximum entropy model (Xie et al., 2011), Bayesian theory (Hazelton, 2008; Castillo et al., 2008a, 2008b, 2014), and the state-space model (Okutani and Stephanades, 1984; Ashok and Ben-Akiva, 2002; Zhou and Mahmassani, 2007; Alibabai and Mahmassani, 2009; Lu et al., 2015). The traffic assignment process needs to generate a set of potential paths based on a certain pre-determined path selection assumption. Usually it is assumed that the vehicles always select the shortest path. However, this shortest path assumption cannot fully represent the actual behaviors of travelers. Moreover, the fixed-sensor-based method could be considered as a mathematical optimization problem that minimizes the differences between the estimated and ground truth values of OD flows or traffic counts. But when in a large-scale network, the number of unknown OD pairs is much greater than that of known traffic counts, leading the fixed-sensor-based method to likely converge to a local optimal solution. However, this solution is not able to truly reflect the distribution of trip patterns on a road network. Moreover, previous studies found that congestion was another factor that may affect estimated OD pattern in the assignment-based methods. As Frederix et al. (2013) claimed, the assignment-based OD demand estimation method assumed the linear relationship between the link flow and OD flow. This assumption may lead to biased estimation in a congested network. Therefore, Frederix et al. (2013, 2014) derived the unbiased estimation of OD demands by calculating the sensitivity of the link flows to all OD flows, to incorporate the effects of congestion spillback on subareas of large-scale networks. Similarly, Shafiei et al. (2017) proposed a sensitivity-based method to relax the linear assumption for estimating OD demand in congested networks. Nevertheless, it is challenging to accurately define the relationship between the link flow and OD flow in the assignment-based method. Therefore, a novel method without using this relationship would be desirable.

Vehicle trajectories collected by vehicle identification or locating technologies, such as automatic license plate recognition (ALPR), cellular networks, and GPS-floating cars can also be used as data sources for OD matrix estimation. Unlike the traditional fixed sensors, these systems can accurately collect the movement information of individual vehicles. For instance, the ALPR sensors can capture the license plate numbers and re-identify them at other locations in the network. Some previous studies extracted traffic counts from observed vehicle trajectories. The extracted traffic counts were used as input data to the traditional fixed-sensor-based method for improving the accuracy of OD estimation (Dixon and Rilett, 2002; Zhou and Mahmassani, 2006). Dixon and Rilett (2002) formulated a Kalman filter model to dynamically estimate OD matrices for a freeway corridor, the traffic counts and link travel times were used as input variables. Zhou and Mahmassani (2006) proposed a non-linear least-squares based dynamic OD estimation method by fusing the traffic counts derived from point-to-point automatic vehicle identification (AVI) sensors with the observed link counts. The historical static demands were employed as prior OD patterns. Despite the fact that the aforementioned trajectory-based methods can achieve relatively reliable outcomes, the essence of these methods is to extract traffic link flow from vehicles’ partial trajectory. In other words, the route choice information included in the trajectories is not sufficiently utilized.

To make the estimated OD demand match correctly with the actual distribution of trips on the network, some studies (Antoniou et al., 2004; Kwon and Varaiya, 2005; Castillo et al., 2008a, 2008b; Sun and Feng, 2011; Feng et al., 2015; Yang and Sun, 2015) derived path flow information (i.e. the traffic volume that choose the same path) by analyzing features of the vehicle trajectories. Antoniou et al. used the path flows obtained from traffic data instead of the link flows, and mapped the path flows to the OD flows for improving the state-space model presented by Ashok and Ben-Akiva (2002). Kwon and Varaiya (2005) derived path flows from Electronic Toll Collection (ETC) systems, and then hourly OD matrices were estimated by an unbiased OD estimator that formulate the relationship between path flow and OD flow. These studies also found that the insufficient sensor coverage and resulting incompleteness of vehicle trajectories are the two main factors that impact the accuracy of OD estimation. Subsequently, the trajectory reconstruction approaches were studied to address trajectory incompleteness. Castillo et al. (2008a, 2008b) reconstructed the path flow by optimizing the potential path set (set of paths) based on the partial trajectories derived from ALPR data. Their results show that the method can effectively identify the origin/destination and path information of the trajectories not provided in link counts. Anderson and Farooq (2017) proposed a k-partite graph method for transportation data association, and implemented the method to reconstruct the trajectories of bicycles and agents. Results show that the method was efficient to handle large datasets and applicable to trajectory reconstruction of vehicles without ID tags. Sun and Feng (2011) used a Bayesian inference technique to estimate the selection probability of potential paths and calculated the posterior probability of a path via Monte Carlo simulation to reconstruct the vehicle trajectories. Due to the advantages in dealing with the nonlinear and non-Gaussian systems, the particle filter model was recently used to reconstruct partial trajectories. Feng et al. (2015) first proposed a particle filter-based vehicle path reconstruction method based on AVI-tags data and traffic counts, and tested it in a simulation environment. The method was further improved by combining the particle filter with a macroscopic path flow estimator. In the combined method, the vehicle path was reconstructed by updating the state-space probability curve (Yang and Sun, 2015). Test results showed that the improved method performs well even when the sensor coverage is as low as 40%. Overall, the studies of the trajectory-based methods
indicated that the performance of OD estimation was improved by the real route choice information carried in the vehicle trajectory and thus can act as an important data source for OD estimation. Nevertheless, most of current studies are limited to small-scale networks or closed freeway networks, because of the limitations of model itself such as calculation complexity, and the testing environment. For example, it is difficult to find a network with high sensor coverage in the real world. Therefore, the trajectory-based methods for large-scale urban networks with complex topology and route choice behaviors are still insufficient and challenging.

To put forward a new trajectory-based OD pattern estimation method for large-scale urban networks, the objectives of this paper are: (1) to propose an offline procedure to estimate OD patterns using the ALPR data; (2) to formulate a particle filter model for vehicle trajectory reconstruction, and (3) to estimate OD patterns from the reconstructed trajectories and learn which ALPR sampling rate is minimum required for OD pattern estimation. Building upon previous studies, including (Feng et al., 2015) and (Yang and Sun, 2015), our contributions include: (1) the particle filter-based trajectory reconstruction method, especially the initial particle generation process and the importance sampling process, is re-designed to accommodate large-scale network scenarios; (2) a path flow estimation approach is proposed through analyzing the complete trajectories of all detected vehicles. Unlike traditional previous analytical-based methods, our approach is data driven, and the detour and dwell can be identified in the proposed approach. (3) The OD patterns are extracted based on historical path flows and the impact of the ALPR sampling rate on OD pattern estimation is investigated. The paper is organized as follows, the methodology presented in the next section, followed by the implementation and performance evaluation, and finally conclusions are drawn.

2. Methodology

This study focuses on developing a method to estimate OD patterns using ALPR sensor data. Generally, ALPR data includes the following attributes: vehicle ID, time stamp, and sensor location. The raw trajectory of a vehicle can be easily extracted through connecting the sensor locations where the vehicle passed. However, the raw trajectory could be incomplete due to the low sensor coverage and detection errors. Thus, a method is required to rebuild complete trajectories first, then path flow and OD patterns can be estimated based on the complete trajectories. Fig. 1 shows the flowchart of the proposed method for OD pattern estimation. The proposed method consists of three components enclosed by the bold dashed lines in Fig. 1:

![Flowchart of the proposed method.](image-url)
1. Trajectory reconstruction with a particle filter

A particle filter is established to estimate and update the probability of all trajectory candidates (i.e. particles) for each vehicle. First, the initial particles are generated by searching potential paths between vehicle locations in road network, and these potential paths are viewed as trajectory candidates. Second, in the importance sampling process, the particle weights (i.e. the probability associated with each particle) are recursively updated using the designed measurement criteria based on ALPR observations. Finally, the partial trajectory is reconstructed selecting the particle with the maximum particle weight.

2. Path flow estimation.

The path flows are estimated through analyzing the complete trajectories of all detected vehicles at certain time intervals (e.g. morning peak). For all the complete trajectories, this step calculates the number of vehicles that shared the same trajectory as path flow of the trajectory.

3. OD pattern estimation.

In this component, the OD flow of each OD pair is calculated by adding up the path flows with the same origin and destination. The average OD flow of a specific interval is considered the OD pattern of this interval.

2.1. Trajectory reconstruction with a particle filter

The particle filter is used to estimate the hidden states of a nonlinear system based on Monte Carlo simulation. The core idea is to represent the probability distribution using random sampling and update the probability with given observed measurements. For a specific vehicle $l$, let $x_k = [x_k^l]_{k=1}^J$ denote the set of trajectory candidates (i.e. set of particles) at time $t_k$, where $J$ is the total number of trajectory candidates; and let $z_k = [z_k^l]_{j=1}^n$ denote the measurement state vector at time $t_k$ such as location, travel time, where $J$ is the total number of measurements. The particle filter can be formulated as:

\[ x_k = f_k(x_{k-1}, v_{k-1}) \tag{1} \]

\[ z_k = h_k(x_k, n_k) \tag{2} \]

where $f_k(.)$ is the state transition function, $h_k(.)$ is the system measurement function, $v_{k-1}$ is a process noise sequence, $n_k$ is a measurement noise sequence, $v_{k-1}$ and $n_k$ are both assumed to be independent and identically-distributed, random, with zero-mean noise.

The particle filter evaluates the posterior probability density function (PDF) of the state vectors (e.g. the potential trajectories) under the given measurements. Suppose $x_k$ is one element in the potential trajectories set $x_k$, and $z_k$ is one of elements in measurement vector $z_k$. Vector $Z^k = [z_1, z_2, \ldots, z_n]$ is defined as the observed measurements available at time $t_k$. Based on the recursive Bayesian framework (Ristic et al., 2004), the PDF $p(x_k|Z^k)$ of the trajectory $x_k$ given the measurements $Z^k$ can be calculated through the following two steps:

(1) The measurement update step

\[ p(x_k|Z^k) = \frac{p(z_k|x_k)p(x_k|Z^{k-1})}{p(z_k|Z^{k-1})} \tag{3} \]

where $p(z_k|x_k)$ is the conditional density of the measurements $Z^k$ when the trajectory $x_k$ is selected, $p(x_k|Z^{k-1})$ is a normalized constant, $p(x_k|Z^{k-1})$ is the PDF of the trajectory $x_k$ given the measurements $Z^{k-1}$ at time $t_{k-1}$.

(2) The state prediction step.

In this step, $p(x_k|Z^{k-1})$ is recursively predicted from the available PDF i.e. $p(x_{k-1}|Z^{k-1})$ at time $t_{k-1}$ using the state transition function (Eq. (1)) based on the Chapman-Kolmogorov equation (Ross, 2014):

\[ p(x_k|Z^{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Z^{k-1})dx_{k-1} \tag{4} \]

where $p(x_k|x_{k-1})$ is the state transition probability and relies on the state model in Eq. (1). It is noted that $p(x_k|x_{k-1}) = p(x_k|x_{k-1}, Z^{k-1})$ since the Eq. (1) is a one-order Markov process.

Since Eq. (4) is an integration process, it is computationally expensive especially when the state space is large (the number of potential paths between a specific OD is very large in urban network). The particle filter approximates $p(x_k|Z^k)$ by the empirical histogram corresponding to a set of particles, and uses the average value of these particles to instead the conditional probability. In this paper, $N$ potential trajectories $[x_i]_{i=1}^N$ are selected as initial particles, and a weight $w_i$ is associated to each particle to represent the posterior distribution at time $t_k$. The particles and their weights together define a histogram that approximates the conditional density function of the state vector $x_k$. Each time step has a new measurement $z_k$, and an importance sampling procedure is conducted to update the particle weights based on Eq. (5).
Algorithm of trajectory reconstruction.

**Step 1. Initialization:** k = 0
For i = 1 to N do
    Generate initial particles \(x_i^0\) using the algorithm shown in sub-section initial particle generation.
    Assume the initial weights follows a uniform distribution: \(w_i^0 = 1/N\)
End for
For k = 1, 2, ..., K do (K is the maximum times of iteration rest upon measurements)

**Step 2. Prediction step**
For i = 1 to N do
    Update the sample \(\{x_i^k\}\) for trajectory candidate i according to \(x_i^k \sim p(x_i^k|z_{k-1})\)
End for

**Step 3. Measurement processing step (Importance sampling)**
For i = 1 to N do
    Update the weights of trajectory candidate i: \(w_i^k = w_{i-1}^k p(z_k|x_i^k)\)
where the likelihood \(p(z_k|x_i^k)\) is calculated using Eqs. (10)-(13)
End for

**Step 4. Normalization**
For i = 1 to N do
    Normalize the weights of trajectory candidate i: \(\hat{w}_i^k = w_i^k/\sum_{i=1}^{N} w_i^k\)
End for

**Step 5. Resampling**
Resample N particles from \(\{x_i^k\}_{i=1}^{N}\) by replicating the larger weight particles and eliminating the lower weight particles based on residual resampling (Hol et al., 2006)
For i = 1 to N, set \(w_i^k = 1/N\), End for

**Step 6. Time step update**
\(k = k + 1\) and return to step 2.
End for

**Step 7. Output**
Select the trajectory candidate with highest weight as output: \(\hat{x} = \text{argmax}(\hat{w}_i^k)\)

\[
\frac{w_i^k \propto w_{i-1}^k p(z_k|x_i^k) p(x_i^k|x_{i-1}^k)}{q(x_i^k|x_{i-1}^k, z_k)}
\] (5)

where \(w_i^k\) is the weight of the trajectory candidate i at time \(t_k\); \(w_{i-1}^k\) is the prior weight of the trajectory i at time \(t_{k-1}\); \(x_i^k, x_{i-1}^k\) are trajectory candidates i at time \(t_k\) and and time \(t_{k-1}\), respectively; \(p(z_k|x_i^k)\) is the conditional density function of the measurement \(z_k\) given the state \(x_i^k\) (i.e. the probability that measurement \(z_k\) is observed under the condition that trajectory candidate i is hypothesized to be the actual trajectory); \(p(x_i^k|x_{i-1}^k)\) is the state transition probability; \(z_k\) is the measurement (e.g. vehicle location, travel time) in time \(t_k\); \(q(x_i^k|x_{i-1}^k, z_k)\) is the previous density distribution, in which the particles are sampled. To a specific trajectory candidate i, \(x_i^k\) is a static variable and keeps stable in any step, herein the transition probability \(p(x_i^k|x_{i-1}^k)\) can be regarded as a constant. Meanwhile, \(q(x_i^k|x_{i-1}^k, z_k)\) is also a known prior density distribution function, thus Eq. (5) can be rewritten as:

\[
\frac{w_i^k \propto w_{i-1}^k p(z_k|x_i^k)}{}
\] (6)

where the conditional density function \(p(z_k|x_i^k)\) is determined through an importance sampling step. After importance sampling, the sum of weights \(w_i^k\) must be normalized to 1, and to avoid the degradation of the particles, a resampling procedure needs to be implemented for improving the confidence of the particles. Finally, the proposed method considers the trajectory candidate with maximum conditional probability as the final reconstructed vehicle trajectory. The algorithm for trajectory reconstruction is described in Table 1.

**2.1.1. Initial particle generation**

The potential paths between an origin and destination are defined as the initial particles in the proposed method. Numerous path searching algorithms exist for determining potential paths between two different nodes on a road network such as Dijkstra’s algorithm (Dijkstra, 1959), and depth first search (DFS) (Cormen et al., 2001). However, these methods need to traverse through all nodes of the network to derive all potential paths. This process is time-consuming especially in networks with thousands of links and nodes. In this paper, a time-geography-based approach is proposed to delimit the searching process to a predetermined area.

Time geography (Hägerstrand, 1970) is an approach in geographic information science for analyzing individuals’ mobility using a space-time path and prism. Fig. 2 illustrates a vehicle’s movement between two activity locations, point i and point j, with respect to time. The red bold line is the space-time path which represents the spatio-temporal trajectory of a vehicle. The slope of each red line segment indicates the vehicle’s speed in a given time period. The two vertical lines mean that the vehicle dwells at the two locations. The space-time prism is a three-dimensional delineation including all possible space-time paths between two activity locations given
the maximum velocity and time budget. The spatial projection of the space-time prism on network space is an ellipse called potential path area (PPA) (Miller, 2005; Downs and Horner, 2012), in which all potential paths are covered.

The space-time prism has been adopted in some previous studies for GPS-based path estimation (Tang et al., 2016) and bus routing optimization (Tong et al., 2017) and achieved satisfactory results. Based on the space-time prism concept, this paper aims to delimit the path searching in the PPA and estimate the dwell locations according to the observed vehicle locations and time stamps. As represented by Miller (Miller, 1991), three basic types of input data are essential for generating the PPAs: the locations of travel origin and destination; the locations of activities; and a traffic network with link and turn impedance. In this paper, the points of intersection are treated as activity locations and the historical travel time and delay are viewed as impedances. To a specific vehicle, given two consecutive locations in the raw partial trajectory and the corresponding time stamps: time \( t(n) \), time \( t(n+1) \), then the initial particles of the vehicle can be generated based on the steps in Table 2.

2.1.2. Importance sampling and weight updating

To update the weight of a particle, the importance sampling (Hol et al., 2006) process is conducted based on three criteria: travel time consistency, the hierarchy, and the sensor measurability. These criteria are designed based on the ALPR observations and network topology. To quantify travel time consistency the conditional density function \( p(\tau_k | x_i) \) is computed by comparing the actual travel time with the average travel time between two consecutive sensors. Consider two consecutive locations \( \text{point}(n) \) and \( \text{point}(n+1) \). Let \( TT(n) \) denote the actual travel time and \( TT^i(n) \) denote the average travel time of the \( i \)th trajectory candidate between these two locations. The ratios of the actual travel time to the average travel time \( \frac{TT_i}{TT^i} \) of all trajectory candidates are calculated. The probability distribution of the ratios is assumed to be a piecewise linear function as shown Eq.(10). Thus, the weight of the particle can be updated by Eqs. (10) and (11):

\[
p_i = \begin{cases} 
0.01 & \frac{TT_i}{TT^i} \leq \beta TT^i \\
\frac{TT^i}{TT(n)} & \beta TT^i < \frac{TT_i}{TT^i} \leq TT \\
[\alpha - \frac{TT^i}{TT(n)}]/(\alpha - 1) & TT < \frac{TT_i}{TT^i} \leq T_{budget}
\end{cases}
\]

\[
p(\tau_k | x_i) = p^I_{\text{travel time}} = \prod_{n=1}^{n_{\text{total}}} p_i^n
\]

where \( p_i^n \) is the updated probability of the potential path in the \( i \)th trajectory candidate between location pairs \( \text{point}(n) \) and \( \text{point}(n+1) \), \( p(\tau_k | x_i) \) is the conditional probability of travel time measurement given the \( i \)th trajectory candidate, \( p^I_{\text{travel time}} \) is the updated probability of the \( i \)th trajectory candidate, and \( n_{\text{total}} \) is the total number of sensor locations in raw trajectory. \( \beta \) is a coefficient less than one. This value could be calibrated by analyzing the probability distribution of the ratios (the actual travel time/average travel time) and underpinning the network topology.
travel time) of the historically observed complete trajectories. The calibration process for $\beta$, is similar to the one for calibrating the value of $\alpha$ in Table 2.

In the urban area of a large city in China, there are many small, four-way stop or two-way stop intersections on the collector roads. Therefore, the desired speed of the minor and branch road is very slow, even during off-peak hours. Selecting the minor and branch roads is not a preferred way by road users to save time when higher functional roads are congested. Thus, travelers are more likely to select the higher functional classes of roads, such as arterials, to avoid the minor and branch roads on an urban network. Therefore, the criterion of road hierarchy updates the particle weights by analyzing the total length of arterials belong to the potential trajectory candidate using the following equation as

$$p(z_2|x_2^i) = p_{road}^i = L^i \sum_{i=1}^{N} L^i$$

where $p_{road}^i$ is the updated probability of $i$th trajectory candidate based on the road hierarchy criterion, $p(z_2|x_2^i)$ is the conditional probability of road hierarchy measurement given the $i$th trajectory candidate, $L^i$ is the total length of arterials belong to the $i$th trajectory candidate, $N$ is the total number of trajectory candidates.

The last criterion, sensor measurability, compares the number of sensors the vehicle passed through (i.e. the number of times vehicle was captured) in a raw trajectory with the number of sensors included in each trajectory candidate. Consider a certain trajectory candidate $i$. We calculate the number of sensors on each trajectory candidate and the raw trajectory covered as $N_{sensor}^i$, and the criterion of sensor measurability assigns a higher weight to the potential trajectory that covers more identical sensors with the observed measurement by the formulation as

$$p(z_3|x_3^i) = p_{sensor}^i = N_{sensor}^i / n^{total}$$

where $p_{sensor}^i$ is the update probability of $i$th trajectory candidate based on sensor measurability, $p(z_3|x_3^i)$ is the conditional probability of sensor measurability given the $i$th trajectory candidate, $N_{sensor}^i$ is the number of sensors both on the $i$th trajectory candidate and the raw trajectory, $n^{total}$ is the total number of sensors in raw trajectory.

### 2.2. Path flow estimation

Path flow estimation is usually conducted by solving a linear or nonlinear optimization model with constraints such as user equilibrium and link counts (e.g. Bell et al., 1996; Lu et al., 2013). Unlike the traditional analytical-based method, this procedure aims to estimate the path flow using a data driven approach by post-processing the complete trajectories, after all the vehicle trajectories are reconstructed. Regarding the path flow estimation based on trajectories, the traditional methods (Sun and Feng, 2011; Feng et al., 2015; Yang and Sun, 2015) usually estimate path flow by directly adding up the number of trajectories with the same
path, without any post-processing to the complete trajectories. Although these methods can derive satisfactory results in small networks, when in a large-scale network, the path flow may be underestimated for two reasons: (1) various long trajectories consist of multiple trips with different trip purposes; (2) the detours would cause the origin or destination of the complete trajectory to be inconsistent with actual origin or destination of travelers’ trips.

To improve the accuracy of path flow estimation, the reconstructed trajectory is viewed as a chain in this research, and then can be divided into multiple trips based on dwell time and dwell location estimated by the initial particles generation algorithm shown in Table 2. For each dwell location in the complete trajectory of a specific vehicle, if the dwell time at a dwell location is larger than a pre-determined threshold, the dwell location will be considered as a new destination/origin, and the trajectory is separated into two trips at this location. If the dwell time is less than or equal to the threshold, the dwell location is treated as a waypoint in the trajectory. The dwell time threshold should be calibrated in different districts of the network (e.g. the congested area, the uncongested area) empirically using observed historical complete trajectories. Through this process, various path flows with different trip purposes can be derived from the complete trajectory. After the traffic chain analysis, the path flow of each complete trajectory is estimated as the number of vehicles that shared the same trajectory.

2.3. OD pattern estimation

The OD pattern is a regular pattern of historical OD flows. To reduce the impact of OD flow fluctuation caused by harsh weather, non-recurrent congestion, etc., the average value of historical OD flows at the same time of day of a certain period (i.e. two weeks, a month, half year) is viewed as the regular OD pattern of the period. To extract OD patterns from path flow, the origin/destination of trips are first assigned to the closest upstream/downstream traffic zones. Then, the historical OD flows are calculated by adding up the path flows with the same origin and destination. Finally, the average of historical OD flows of the specific time of day (e.g. morning peak) is considered as the OD pattern of the period.

3. Implementation

3.1. Data description

To validate the proposed method, the urban road network around the downtown area in Kunshan, China was selected as a testbed. The road network is composed of 11 major arterials, 14 secondary roads and some collector roads. On the road network, 344 ALPR sensors are installed at the approaches of intersections to collect license plate numbers of vehicles passing by the devices. Fig. 3 shows the study area of the entire testbed and the locations of ALPR sensors.

Collecting ground-truth data is challenging particularly for large-network problems. As a result, to validate the proposed method a microscopic traffic simulation model was built by using Q-PARAMICS for the testbed. This model has been fully calibrated and tested in a previous study (Lu et al., 2015). The testbed was divided into 207 traffic zones following the zone IDs defined in this microscopic traffic simulation model (see in Fig. 3(b)).

The raw ALPR dataset was collected from January 11th, 2017 to January 25th, 2017 (two weeks data except weekends). The dataset includes about 175,000 records per hour with an average sampling rate of 88%. The sampling rate is the proportion of vehicles captured by the ALPR devices to the total number of vehicles. The fields collected by ALPR sensors are device ID, date, time stamp, and license plate number. The vehicle locations can be derived from a data table that associates device IDs with the latitude and longitude information. The travel time between two locations was estimated by comparing the adjacent time stamps and the method presented in our previous study (Li et al., 2018). To evaluate the OD pattern estimation results, the raw ALPR data was divided into three datasets: the morning peak (7:00–9:00 am), the evening peak (16:30–18:30 pm), and the midday non-peak (13:00–14:00 pm). The midday non-peak time interval typically has traffic volumes that are relatively low because of a nap time after lunch in China.

3.2. Trajectory reconstruction and path flow estimation

In this study, the proposed method reconstructs the trajectory of each detected vehicle and then estimates the path flow by dividing the complete trajectory into multiple trips based on dwell time and location. The accuracy of the trajectory reconstruction and path flow estimation is a critical factor to the performance of OD pattern estimation. Before implementing the proposed method, all parameters were calibrated. Based on our empirical studies, the calibrated coefficients $\alpha$ and $\beta$ are 2.5 and 0.35, respectively. The pre-determined threshold of dwell time is set to be five minutes in the uncongested area where recurrent congestion does not always happen, while in the congested area where recurrent congestion always happen, the threshold is set to be 10 min during the morning and afternoon peaks, and five minutes during the midday off-peak.

Fig. 4 shows the reconstructed trajectories of two typical vehicles in the morning peak and the distribution of particle weights of their potential paths. Vehicle 1, in Fig. 4(a) was captured at 12 locations, presented in time-order by number. Vehicle 2 in Fig. 4(b) was captured at 19 locations. It can be seen in Fig. 4(a), location 5 is identified to be an unreliable observation because the estimated travel time between locations 4 and 5 is less than the free flow travel time. The trajectory was divided into two trips: the red one from locations 1 to 6, and the blue one from location 6 to 12, since the dwell time at location 6 was 7.3 min, which is longer than the pre-determined threshold of 5 min through our empirical studies. Vehicle 1 stopped close to location 6 and then was re-detected at location 7 three minutes later; therefore location 6 was considered to be the destination of the red trip and origin of the blue trip.
Additionally, from locations 7 to 9, the vehicle made a detour passing by location 8 likely since the section between location 7 and 9 in main road was congested at that time. In comparison, vehicle 2 has a much longer trajectory which consists of three trips. Location 12 was eliminated in the reconstructed trajectory because travel time between locations 11 and 12 and that between locations 12 and 13 were both less than 1 min, which is unlikely to be realistic. Fig. 4(c) and (d) are the particle weights of vehicles 1 and 2.
Fig. 4. Reconstructed trajectories and particle weights.

Table 3
Experimental results of trajectory reconstruction.

<table>
<thead>
<tr>
<th>Time periods</th>
<th>Sampling rate (%)</th>
<th>Number of Trajectories reconstructed</th>
<th>Number of locations in raw trajectories</th>
<th>Number of particles per vehicle</th>
<th>Computation time per trajectory* (ms)</th>
<th>Average absolute error (veh/h)</th>
<th>Average relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning peak</td>
<td>88.3</td>
<td>63,472</td>
<td>7.75</td>
<td>67</td>
<td>122</td>
<td>23</td>
<td>12.2</td>
</tr>
<tr>
<td>Midday off-peak</td>
<td>86.0</td>
<td>41,820</td>
<td>8.33</td>
<td>43</td>
<td>96</td>
<td>17</td>
<td>11.4</td>
</tr>
<tr>
<td>Evening peak</td>
<td>89.5</td>
<td>63,537</td>
<td>7.46</td>
<td>54</td>
<td>110</td>
<td>28</td>
<td>9.8</td>
</tr>
</tbody>
</table>

* All the indices in the table are the average values in the testing days (two weeks).

** The proposed method was conducted on a computer with Intel Core i7 processor and 16 GB RAM.
Fig. 5. Distributions of trip attraction and generation.
respectively. One can observe that the vehicle 1 has 130 trajectory candidates and most of their weights are close to zero, only seven of them are larger than 0.05. The reason could be that in the particle filter method, the importance sampling process updates the particle weights and reassigned the lower weights to particles that do not meet the three measurement criteria. Comparatively, vehicle 2 has more than 1000 particles due to the existence of some adjacent points with long distances, which leads to more potential paths in the PPA. Meanwhile, it can also be seen that all the particle weights are smaller than those seen for vehicle 1 and the largest one is merely 0.098.

Since the actual path flow cannot be obtained in the real-world scenario, the estimated path flows were compared with the simulated path flows using the Kunshan traffic simulation model in Fig. 3(b). This model was well calibrated with the observed link counts captured by microwave sensors and turning movements captured by traffic image processing (Xia et al., 2013). The GEH statistic was used to measure the distances between the observed and the simulated traffic counts. More than 85% of the GEH values among all traffic counts were less than five as suggested by (De Villa et al., 2013). Besides the traffic counts, the arterial travel times were also used as indices to calibrate the simulation model, and the mean absolute percentage errors (MAPEs) of travel times were calibrated to be lower than 15%. The historical average OD matrices were implemented as the simulation inputs to ensure consistency. The historical average OD matrices were estimated using a Kalman filter-based method proposed by our previous study (Lu et al., 2015) and fine-tuned using the Estimator Tool in PARAMICS. This proposed Kalman filter-based method is based on both counts and assignment. The dynamic traffic assignment procedure is embedded in the Kalman filter framework. To ensure the distribution of the estimated OD matrices was consistent with the actual distribution of OD demands, an initial OD demand derived from a resident trip survey of Kunshan city conducted in the year of 2016 was used as the reference OD demand. Table 3 summarizes the experimental results of the trajectory reconstruction for three different periods, the morning, and evening peaks, and mid-day non-peak. The sampling rates show that at least 86 percent of vehicles were captured by ALPR sensors in the research area. The average computational time per trajectory is positively proportional to the number of particles and varies from 96 ms to 122 ms. Though that is not fast enough for real-time work, it is sufficient for the offline trajectory reconstruction. To evaluate the path flow estimation performance, the average absolute error and the average relative error were used as indices. The average absolute errors of path flow are between 17 vehicles per hour and 28 vehicles per hour, and the maximum average relative error is 12.2% at the morning peak. This error rate implies that the path flow estimation results are accurate for OD pattern extraction.
4. Results analysis and performance evaluation

4.1. OD pattern distribution

To qualitatively evaluate the performance of the proposed method, the temporal and spatial distribution of trip generation and trip attraction of the selected zones (where the total number of trips are larger than 50 vehicles per hour) are shown in Fig. 5. The area of the circle represents the total number of trips in the specific zone, while the yellow part of the circle is trip generations and the green part is trip attractions.

During the morning peak, the zones with more trips were mainly distributed in the outskirts of the study area (e.g. railway station, freeway ramps) and the downtown area with an abundance of commercial facilities. Trip generation in residential areas was higher than trip attractions, while in industrial areas and the downtown area trip attractions are more than trip generations. Additionally, we can also observe that the zones close to freeway ramps generate more trips than they attract, illustrating that there is more traffic demand moving into the urban area during the morning peak.

The results of the midday off-peak are presented in Fig. 5(b). The total number of trips in most zones decreased compared to the peak hours, especially in the residential areas and industrial areas. However, the zones close to the freeway ramps or railway station such as zones 83, 91, and 241, still have higher demand than other zones. During the evening peak shown in Fig. 5(c), the trips of almost all traffic zones matched the levels seen in the morning peak. Similar to morning peak, the zones with more trips are mainly in the downtown area and in the periphery of the research area. But on the contrary, the residential areas have more trip attractions while the industrial area has more trip generations. Despite the zones close to freeway ramps having the most trips as usual, the trip attractions take up a larger proportion than generations. This indicates that more traffic is leaving the research area in the evening peak. The trip distribution derived from the proposed method is consistent to what one would expect in real life, and can qualitatively and intuitively illustrate the changing of trip attraction and generation in a network in different time periods.

<table>
<thead>
<tr>
<th>OD pair</th>
<th>Time period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morning Peak</td>
</tr>
<tr>
<td></td>
<td>Estimated OD (veh/h)</td>
</tr>
<tr>
<td>27-25</td>
<td>126</td>
</tr>
<tr>
<td>33-34</td>
<td>107</td>
</tr>
<tr>
<td>41-56</td>
<td>116</td>
</tr>
<tr>
<td>46-48</td>
<td>117</td>
</tr>
<tr>
<td>56-61</td>
<td>226</td>
</tr>
<tr>
<td>80-117</td>
<td>159</td>
</tr>
<tr>
<td>83-33</td>
<td>320</td>
</tr>
<tr>
<td>83-117</td>
<td>133</td>
</tr>
<tr>
<td>91-34</td>
<td>20</td>
</tr>
<tr>
<td>92-126</td>
<td>106</td>
</tr>
<tr>
<td>99-117</td>
<td>141</td>
</tr>
<tr>
<td>114-116</td>
<td>233</td>
</tr>
<tr>
<td>114-211</td>
<td>189</td>
</tr>
<tr>
<td>117-194</td>
<td>65</td>
</tr>
<tr>
<td>124-127</td>
<td>137</td>
</tr>
<tr>
<td>126-95</td>
<td>358</td>
</tr>
<tr>
<td>126-175</td>
<td>202</td>
</tr>
<tr>
<td>127-163</td>
<td>336</td>
</tr>
<tr>
<td>163-189</td>
<td>67</td>
</tr>
<tr>
<td>182-177</td>
<td>163</td>
</tr>
<tr>
<td>188-189</td>
<td>23</td>
</tr>
<tr>
<td>189-163</td>
<td>427</td>
</tr>
<tr>
<td>189-188</td>
<td>109</td>
</tr>
<tr>
<td>194-117</td>
<td>234</td>
</tr>
<tr>
<td>194-116</td>
<td>89</td>
</tr>
<tr>
<td>198-92</td>
<td>111</td>
</tr>
<tr>
<td>198-123</td>
<td>125</td>
</tr>
<tr>
<td>198-126</td>
<td>643</td>
</tr>
<tr>
<td>211-114</td>
<td>118</td>
</tr>
<tr>
<td>241-124</td>
<td>53</td>
</tr>
<tr>
<td>Average</td>
<td>171</td>
</tr>
</tbody>
</table>

* Estimated OD is the average value of the specific observed OD pair in the testing days.
4.2. Overall performance

The OD patterns are represented by a matrix with the same size as OD demand matrices in this research. Traffic demand of each OD pair in the estimated OD pattern is compared with that in the observed OD pattern. It should be noted that the observed OD demands are considered to have 100% sampling rates, while the estimated OD demands are scaled by a factor determined by the sampling rate because the distribution of the trajectories is assumed to be homogenous over the network. Two statistical indicators: the mean absolute error (MAE) and the mean absolute percentage error (MAPE) (Simon et al., 2010) between the estimated OD patterns and the observed OD patterns are calculated to quantify the overall performance of the proposed method. The observed OD patterns are obtained by a resident and vehicle trip survey and are fine-tuned in PARAMICS based on the observed traffic volumes. The research area has a total of 42,849 OD pairs but most of them have no demand or a demand lower than 10 vehicles per hour. To avoid the disturbance of these small OD flows, a total of 30 OD pairs with larger OD demands were selected and the results are given in Table 4.

As shown in Table 4, during the morning peak the MAEs between the estimated and observed values are all less than 65 vehicles per hour, but the MAPEs vary from 5.07% to 38.10%. The maximum MAE at midday off-peak reduces to 36 vehicles per hour with the decrease of the estimated ODs. All the MAPEs are less than 34% except for OD pair 91-34 with 44.83%. The reason could be that a slight error can result in a large MAPE when the OD demand is small. In the evening peak, all but one OD pairs had MAEs less than 69 vehicles per hour. This may be due to that the coverage rate of ALPR sensors in the area near zone 211 is so limited, leading many vehicles are missed for re-detection. Moreover, for the OD pairs with heavy demand (e.g. OD pairs 198-126 and 189-163 in peak hours), their MAPEs are less than the average MAPEs of the same periods, some even less than 10%.

To further evaluate the performance of the proposed method, three scatter plots for the estimated demands with respect to the observed demands of all OD pairs are presented in Fig. 6.

W. Rao et al. 
Transportation Research Part C 95 (2018) 29–46

42
test network where ALPR coverage is limited.

It should be noted that the estimated OD patterns are based on an assumption that the distribution of the trajectories are homogenous in the test network. However, certain areas may not have a sufficient number of trajectories. The inhomogeneity of the distribution of trajectories is mainly caused by the distribution of the ALPR sensors and the recognition accuracy of the ALPR sensors. Moreover, the recognition rate of each ALPR sensor is also an important factor that results in inhomogeneous distributions. To overcome this limitation, several suggestions are made for future studies: First, one should evaluate the homogeneity of the trajectories and estimate the OD patterns in terms of different sub-areas. Second, the installation locations of the ALPR sensors should be optimized in order to achieve the maximum coverage of the observed trajectory data. Last, one should use the ALPR sensors with the high recognition rates and the raw data quality should be validated before vehicle trajectory construction.
4.3. Impact of ALPR sampling rates

In this study, the ALPR sampling rate is defined as the proportion of vehicles captured by the ALPR devices. Due to the low ALPR coverage rate and the recognition errors of ALPR sensor, it is challenging for all the vehicles to be captured by ALPR sensors in real-life scenarios. It was found in our study that the ALPR sampling rate has impact on the precision of OD pattern estimation. Therefore, the proposed method is tested with six datasets of morning peak traffic using six different sampling rates: 80%, 70%, 60%, 50%, 40%, and 30%. Fig. 7 shows the results at different sampling rates. As shown from Fig. 7(a) to (f), the points gradually move away from the 45-degree dashed line as the sampling rate decreases. This indicates that the precision of the estimated demands becomes much worse with the lower sampling rates, especially when the sampling rate is less than 60%. Moreover, the OD demand tends to be over-estimated as the sampling rate becomes lower. Additionally, one can also find that both the slopes and the \( R^2 \) values of the fitted lines (blue lines) trend downward from 1 as the sampling rate decreases, demonstrating that the OD estimation accuracy become worse in the lower sampling rates.

The MAPEs and Root Mean Square Errors (RMSEs) at all sampling rates are calculated and shown in Fig. 8. One can find that both MAPE and RMSE decrease with the increase of the sampling rates. The two indices change smoothly when sampling rate is between 90% and 60%. However, when the sampling rate is less than 60%, the MAPE becomes larger than 30% and the RMSE increases rapidly from 45 vehicles per hour to 130 vehicles per hour, which means that the estimated OD pattern cannot satisfy the practical applications such as traffic planning, traffic simulation, etc. Therefore, 60% is suggested as the minimum required sampling rate in this research based on the existing ALPR sensors installation status.

5. Conclusions and future work

This paper proposed a novel method for OD pattern estimation using ALPR data on a large-scale urban network. The proposed method consists of three components. First, a particle filter was established to estimate and update the probability of all trajectory candidates (i.e. particles) for each vehicle. The initial particles were generated by searching potential paths in PPA based on time geography theory, and these potential paths are viewed as trajectory candidates. In the importance sampling process, the particle weights were recursively updated using the designed measurement criteria and the particle with the maximum particle weight is viewed as the reconstructed trajectory. Second, the path flows were estimated through analyzing the complete trajectories of all detected vehicles at certain intervals. Finally, OD patterns estimation were estimated based on path flow. The OD flow of each OD pair was calculated by adding up the path flows with the same OD. The average historical OD flows of a specific interval were considered as the OD pattern of each interval.

The proposed method was implemented on a real-world traffic network in Kunshan, China. Implementation results showed that the proposed method can perform well in real-life traffic scenarios. The path flows estimated from the complete trajectories reconstructed by the proposed particle filter method match closely with the observed path flows in the simulation model. Moreover, the distribution of OD patterns can represent the variation of trip generation and attraction observed by time of day. Results show that the proposed method effectively uses ALPR sensors data to achieve high-accuracy trajectory reconstruction and historical OD pattern estimation and can be applied in real-world traffic management situations.

Comparative results in different ALPR sampling rates show that the accuracy of the proposed method decrease with the lower sampling rates, and the 60% is suggested as the minimum required sampling rate specifically in this study. More studies should be conducted to investigate the impacts of sensor locations and the spatial coverage of sampled trajectories.

Even though the proposed method achieved positive outcomes, several further research directions can be taken to expand this work. First, only one data source, the ALPR data, was used to reconstruct the trajectories. The probabilities of potential trajectories were updated based on limited observations. The proposed method could be enhanced by using multiple data sources to improve the estimation accuracy. Second, the proposed method may be further integrated with an online OD demands estimation method for real-time traffic management applications. Third, the historical impedances used in this study and how the variation of these impedances impact the outcome of the estimation results should to be verified in the future. Last, data quality is a key factor to our estimated
results. To measure the quality of the trajectory data in the coverage area, some potential measures, including the “penetration rate” and the “recognition rate” of ALPR sensors, can be used in the future studies.

Acknowledgements

The authors want to thank the anonymous peer reviewers for their useful comments and suggestions to help improve this paper. Moreover, the authors gratefully acknowledge the financial support received from the Key Research and Development Program of Jiangsu Province (No. BE2017027), and the Innovative Talents Training Fund for Postgraduates of Jiangsu province (No. KYLX 0272) since this research was supported by these programs. The authors also appreciate the technical support provided by the members of the Smart Transportation Lab at the University of Arizona. Last, the authors would like to acknowledge the Traffic Management Bureau of Kunshan for providing the traffic flow data for this study.

References

