Transportation Research Part C 95 (2018) 29–46

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

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

Wenming Rao[a], Yao-Jan Wu[b], Jingxin Xia[a,*], Jishun Ou[a], Robert Kluger[c]

[a] Intelligent Transportation System Research Center, Southeast University, No. 35 Jinxianghe Road, Xuanwu District, Nanjing 210096, PR China
[b] Department of Civil Engineering and Engineering Mechanics, The University of Arizona, 1209E. 2nd St., Tucson, AZ 85721, USA
[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), yaojan@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

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 trajectory 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