Transportation Research Part B 78 (2015) 280–300

Contents lists available at ScienceDirect

Transportation Research Part B

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

Statistical approach for activity-based model calibration based on plate scanning and traffic counts data

Treerapot Siripirote [a], Agachai Sumalee [b,c,*], H.W. Ho [c], William H.K. Lam [c]

[a] *Department of Civil Engineering, Faculty of Engineering, Srinakharinwirot University, Ongkharak, Nakhon Nayok, Thailand*
[b] *Department of Civil Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand*
[c] *Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China*

ARTICLE INFO

Article history:
Received 30 June 2013
Received in revised form 15 March 2015
Accepted 7 May 2015
Available online 23 May 2015

Keywords:
Maximum-likelihood estimation
Plate scanning
Statistical model calibration

ABSTRACT

Traditionally, activity-based models (ABM) are estimated from travel diary survey data. The estimated results can be biased due to low-sampling size and inaccurate travel diary data. For an accurate calibration of ABM parameters, a maximum-likelihood method that uses multiple sources of roadside observations (link counts and/or plate scanning data) is proposed. Plate scanning information (sensor path information) consists of sequences of times and partial paths that the scanned vehicles are observed over the preinstalled plate scanning locations. Statistical performances of the proposed method are evaluated on a test network using Monte Carlo technique for simulating the link flows and sensor path information. Multiday observations are simulated and derived from the true ABM parameters adopted in the choice models of activity pattern, time of the day, destination and mode. By assuming different number of plate scanning locations and identification rates, impacts of data quantity and data quality on ABM calibration are studied. The results illustrate the efficiency of the proposed model in using plate scanning information for ABM calibration and its potential for large and complex network applications.

1. Introduction

Activity-based models (ABM) have been developed to overcome the limitations from the conventional 4-step models by deriving travel demand from activity participations and activity behavior sequences/patterns (Bhat and Koppelmen, 1999). ABMs can either be formulated by utility maximization-based (e.g. Bowman et al., 2006; Bifulco et al., 2010) or rule-based approach (e.g. Arentze and Timmermans, 2004). The utility maximization-based ABMs have been widely developed to evaluate traffic policies in many cities such as Portland, Columbus, Sacramento, and Jakarta (Bradley et al., 2010; Vovsha et al., 2004; Yagi and Mohammadian, 2010).

Traditionally, activity-based models are estimated from travel diary survey data (TD) of which estimated results can be biased due to low-sampling size and inaccurate TD. For example of the ABM with complex activity-travel decisions, only 1% of the population was used to estimate the ABM parameters (Bowman et al., 2006). In addition, trip under-reporting, due to response burdens or incomplete/inaccurate trip memory, will also lead to erroneous travel diary data (Bricka and Bhat, 2006). Consequently, the predicted travel demands based on TD may be inconsistent with actual roadside data (e.g. link

* Corresponding author at: Department of Civil Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. Tel.: +66 89280 1692; fax: +66 2274 0698.

E-mail address: asumalee@gmail.com (A. Sumalee).

http://dx.doi.org/10.1016/j.trb.2015.05.004
0191-2615/

counts). In order to calibrate ABM parameters that consist with roadside data, Bowman et al. (2006) developed a comprehensive model calibration approach. According to their approach, some ABM parameters in the utility function (e.g. constant terms) are heuristically adjusted to reproduce predicted traffic flows that fit with the collected traffic counts. In recent years, Cools et al. (2010) also conducted ABM calibrations by a heuristic method. In their study, the demand of activity-travel patterns was adjusted by randomly weighting their activity-travel pattern choices to reproduce external trip matrix information. The ABM calibration from this method, however, still lacks statistical efficiency to measure how well the model calibration results fit the actual roadside data (e.g. link counts).

Ideally, the data used to calibrate the parameters in ABM should be collected by GPS-based travel surveys such as activity-travel data collected by GPS equipment attached to probe vehicles or carried by travelers. (e.g. Cottrill et al., 2013; Frignani et al., 2010; Axhausen et al., 2003). The GPS-based data can provide a probe of a sequence of activity/time spent/travel decisions, which is an RP-typed data for the activity-based model. To maintain the high accurate TD data and minimize response burden, learning machine technique was conducted to pre-determine the activity-travel choices of the respondents from the historical record of travel diary made (Cottrill et al., 2013). Nevertheless, the sample sizes of the activity-travel data collected by GPS-based travel surveys may be small due to low response rate. As an alternative, it is possible that number plate scanning (PS) could be used to identify travel paths. Information, in the context of vehicle tracking, obtained from plate scanning, which does not require the installation of GPS equipment, is similar to that of the GPS-based data collection method. In PS data collection, vehicle information is obtained at pre-determined locations on the road network.

Furthermore, plate scanning process can identify the same vehicles traveling along a series of plate scanning locations by matching their license plate numbers. With this method of data collection, plate scanning is considered to be one of the methods to collect vehicle identification (VI) data. The data from plate scanning consists of: (i) the vehicle passing time at plate scanning locations and (ii) the sequence of scanned vehicles along a series of plate scanning locations on road networks. The accuracy of collecting the above data from the plate scanning method is determined from detection rate and identification rate. Detection rate is the proportion of the number of detected vehicles among those vehicles passing sensor locations (i.e. the vehicle is known to pass the sensor and the license plate number of the vehicle can be detected by the sensor). In addition, identification rate is the proportion of detected vehicles that their license plate numbers are correctly identified. To achieve a high identification rate, Ozbay and Ercelebi (2005) proposed a license plate recognition system with more than 90% of license plate samples correctly identified.

Further to the related studies on model calibrations with VI data, Castillo et al. (2008b) developed trip table reconstruction/estimation framework using plate scanning. In recent years, Siripirote et al. (2014) attempted to update travel behavior model parameters and to estimate trip chains by using individual vehicle's timestamps from plate scanning. In this study, a statistical framework for the calibration of activity-based model parameters using sensor path flows from plate scanning information is proposed that is based on the hierarchical activity-travel decision model. This model calibration approach was motivated by similar works done in the area of traffic flow estimations, in which one observes traffic in some links and tries to estimate OD flows or flows on remaining links (e.g. Bell, 1991; Castillo et al., 2007; Castillo et al., 2008a; Castillo et al., 2010; Castillo et al., 2011; Castillo et al., 2013; Maher, 1983; Ng, 2012; He, 2013; Watling and Maher, 1992; Watling, 1994; Yang et al., 1992; Yang, 1995; Shen and Wynter, 2012). The ABM calibration proposed in this paper is formulated as a maximum likelihood estimation problem for reproducing the collected link counts and sensor path flow information. The remainder of the paper is organized as follows. Firstly, basic components of the proposed model, including notation and information collected from plate scanning, are described in Section 2. The structured relations between activity-travel demands and traffic flows are then described in the third section. Also, the proposed calibration model derived from these structured relations is presented in Appendix A. In Section 4, numerical example for evaluating the proposed calibration model is setup, solved and discussed. Finally, conclusions are drawn in the last section.

2. Assumptions and definitions

In this section, the necessary assumptions and definitions for setting up the proposed ABM calibration model (see Appendix A) are introduced. The definitions in network, activity chain representations, and collected plate scanning data could be respectively found in Sections 2.1 and 2.2, while assumptions of the proposed model will be given in Section 2.3.

2.1. Network and activity chain representation

For traffic network $(\mathbf{N}, \mathbf{L})$, $\mathbf{N}$ is the set of nodes and $\mathbf{L}$ is the set of links. Activity location lo is located in each traffic zone where $\mathbf{N}_z$ is the set of zone centroids and $\mathbf{L}_z$ is the set of links in traffic zone z ($\mathbf{N}_z \subset \mathbf{N}$ and $\mathbf{L}_z \subset \mathbf{L}$). In this study, the activity location (lo) is assumed to be virtually located at the zone centroid, which is one of the nodes in a network, and represents all real activity locations in each traffic zone ($lo \in \mathbf{N}_z$).

As we consider on daily activity-travel participations, user i (ith observed vehicle) makes a plan to perform *activity pattern* y. Let $\mathbf{A}_y$ denotes the daily scheduled activity pattern that consists of an ordered set of the activities:

$$\mathbf{A}_y = \{a_1, \ldots, a_n, \ldots, a_{N_y}\}, \quad \forall y \in \{1, \ldots, Y\} \tag{2.1}$$