Transportation Research Part C 89 (2018) 96–112

Contents lists available at ScienceDirect

Transportation Research Part C

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

Multi-day activity-travel pattern sampling based on single-day data

Anpeng Zhang[a], Jee Eun Kang[a,*], Kay Axhausen[b], Changhyun Kwon[c]

[a] Department of Industrial and Systems Engineering, University at Buffalo, SUNY, United States
[b] Institute of Transport Planning and Systems, Swiss Federal Institute of Technology, Switzerland
[c] Department of Industrial and Management Systems Engineering, University of South Florida, United States

ARTICLE INFO

Keywords:
Activity-travel patterns
Day-to-day variability
Interpersonal variability
Sampling multiday activity-travel patterns

ABSTRACT

Although it is important to consider multi-day activities in transportation planning, multi-day activity-travel data are expensive to acquire and therefore rarely available. In this study, we propose to generate multi-day activity-travel data through sampling from readily available single-day household travel survey data. A key observation we make is that the distribution of interpersonal variability in single-day travel activity datasets is similar to the distribution of intrapersonal variability in multi-day. Thus, interpersonal variability observed in cross-sectional single-day data of a group of people can be used to generate the day-to-day intrapersonal variability. The proposed sampling method is based on activity-travel pattern type clustering, travel distance and variability distribution to extract such information from single-day data. Validation and stability tests of the proposed sampling methods are presented.

1. Introduction and background

Intrapersonal variability, also known as day-to-day variation, of activity-travel patterns is found to show strong repetitions, yet with considerable variations (Hanson and Huff, 1981; Hanson and Huff, 1988; Pas and Sundar, 1995; Pendyala and Pas, 2000; Chikaraishi et al., 2009, 2011). Observations of day-to-day variation of activity-travel patterns have been studied to understand activity-travel behavior of adaptation, habit, and symmetry. Both stability and variability have been observed at intrapersonal levels as well as at both spatial and temporal levels (Buliung et al., 2008; Koppelman and Pas, 1984; Pendyala et al., 2001; Pas and Koppelman, 1986; Pas and Sundar, 1995; Susilo and Axhausen, 2014). Variations of travel behavior have also been explained by day-of-week factors. In previous studies, Pas and Koppelman (1986) utilized daily trip generation rates to measure the intrapersonal variability. According to their observations, employment status, household role, social class and daily travel resource could all affect intrapersonal variability; thus different population groups are likely to have huge differences in day-to-day travel activity. Later, Pas (1988) categorized activity-travel patterns into five types with cluster analysis and calculated the probability of selecting each pattern type for day-of-week. They mentioned that day-of-the-weeks differences are highly related to sociodemographic characteristics, while day-of-week would not affect weekday travel behavior for workers. Then, by including trip chaining and daily travel time, Pas and Sundar (1995) extended trip generation rate day-to-day variation analysis with similar formulations of the total sum of squares in travel behaviors. Their results indicated that intrapersonal variability could vary according to different sample data, but it significantly affects the total variability in day-to-day travel behaviors of individuals. Elango et al. (2007) introduced delta trips as the measurement of day-to-day trip making variability. Their experiment results showed that intrapersonal variability based on household trip number is greatly affected by demographic variables, including income, person number, etc. without considering seasonal affects. In Table 1, we show whether previous works conclude that intrapersonal variability occupies a large proportion of

* Corresponding author.
E-mail address: jeeeunka@buffalo.edu (J.E. Kang).

https://doi.org/10.1016/j.trc.2018.01.024
Received 13 June 2017; Received in revised form 27 January 2018; Accepted 27 January 2018

Table 1
Variability proportion and measurements in previous studies.

Large proportion of total variance	Paper	Intrapersonal variability measurements			
		Trip frequency	Travel time	Actvity location	Activity-travel pattern
Yes	Pas and Koppelman (1986)	✓			
	Pas (1988)				✓
	Pas and Sundar (1995)	✓	✓		
	Kang and Scott (2010) (Weekdays)				✓
	Chikaraishi et al. (2011)		✓		
No	Susilo and Kitamura (2005)			✓	
	Chikaraishi et al. (2011)				
	Kang and Scott (2010) (Weekends)				✓
	Moiseeva et al. (2014)				✓
Not mentioned	Schlich and Axhausen (2003)	✓	✓		
	Elango et al. (2007)	✓			
	Buliung et al. (2008)			✓	
	Susilo and Axhausen (2014)				✓

the total individual travel variability. We include four major factors (trip frequency, travel time, activity location and activity-travel pattern) in variability measurements while detailed measurements related to the same factor could vary depending on different studies. A more detailed summary of variability measurements in previous studies is given in Appendix A. However, how the new generations of travelers will change in terms of activity-travel pattern remains uncertain according to some recent works. Lyons (2015) focused on the interactions between motor age and digital age based on a socio-technical conceptualization of society, and they showed the future of transport is uncertain in the digital age. On the other hand, Garikapati et al. (2016) introduced the potential changes of activity pattern as well as time usage from a generation of people. According to their work, the activity pattern might change based on the age of travelers instead of the generation. They showed that millennial tend to mimic the activity-time use patterns of prior generations and fundamental shift in travel demand in the future should not be expected based on similar work (McDonald, 2015) with earlier household travel survey data. Given diverse measurements and distinct numerical results in previous work, intrapersonal variability has been proved to be closely related to the variation of people's activity-travel patterns even if the effect of new generations of travelers remains uncertain.

Despite this evidence, day-to-day intrapersonal variability is often ignored in studies analyzing activity-travel behaviors and estimating travel demand due to the unavailability of multi-day data. Several studies showed the limitations of relying on single-day data from longitudinal travel pattern data. Since single-day data contain little time-related information and various measurements of intrapersonal variability are applied, the results based on those data range from 20% to 80% comparing to each other (Chen et al., 2016). In case of applications or case studies in reality, working with only single-day data could result in distinct conclusions.

However, data collection is an expensive task and data sets with multi-day activity-travel patterns are rarely available. National-, state-, and regional-level household travel surveys collect detailed information of activity-travel along with household socio-demographics. Governments, industry, and researchers rely on these data sets for travel forecasting, planning, traffic management, etc. These surveys generally include only one weekday activity-travel information. Recently, with various types of IT technologies, collecting multi-day data has become more readily available and affordable. These data sets enable us to understand intrapersonal variability of certain travel choices. However, these data are often passively collected and therefore miss information such as travel/activity purpose, specific travel modes (carpooling, specific services used), cost of travel and accompanying passengers.

There is limited work focusing on the generation of multiple-day travel dataset. Recently, Medina (2016) presented two discrete choice models to generate multiple-day travel activity types based on the revealed sample trips according to the likeliness of the activity, using large AFC data and survey. We intend to generate multi-day data with consideration of day-to-day variation via sampling from a single-day dataset, which is easier to collect comparing to smart card data. *Cross-sectional data* (single-day data) in statistics and econometrics is usually collected by observing many subjects (such as individuals or regions) without regard to differences in time (day-of-year). For contrary *panel data* (multi-day data), researchers conduct several observations of the same subjects over a period of time in a longitudinal study. Large-scale cross-sectional datasets contain detailed information of various aspects of activity-travel decisions and interpersonal variability. We intend to extract intrapersonal variability given the rich travel activity information in such datasets. We assume that single-day data contains a diverse set of activity-travel patterns that is sufficient to be used as a surrogate for multi-day activity-travel patterns. Suppose several people have similar travel activities, they may travel for work-only on some days and travel for work and shopping on other days. When we collect the single-day travel-activity pattern data of these people, we are likely to observe work-only activity-travel pattern on some days and work-and-shopping activity-travel pattern on other days. The distribution of work-only pattern and work-and-shopping pattern among the chosen single-day samples could be similar to the distribution of work days as well as work-and-shopping days of each person on average. In other words, multiple observations over a large collection of presumed homogeneous observations can be used as a surrogate for repeated observations over a single individual.

We also need a well-defined measurement of intrapersonal variability in order to indicate the travel activity pattern of a person