Transportation Research Part B 112 (2018) 73–88

Contents lists available at ScienceDirect

Transportation Research Part B

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

Detecting pattern changes in individual travel behavior: A Bayesian approach

Zhan Zhao [a], Haris N. Koutsopoulos [b], Jinhua Zhao [c,*]

[a] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States
[b] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, United States
[c] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

ARTICLE INFO

Article history:
Received 23 January 2018
Revised 28 March 2018
Accepted 30 March 2018

Keywords:
Pattern change detection
Travel behavior
Smart card data
Bayesian inference

ABSTRACT

Although stable in the short term, individual travel patterns are subject to changes in the long term. The ability to detect such changes is critical for developing behavior models that are adaptive over time. We define travel pattern changes as “abrupt, substantial, and persistent changes in the underlying patterns of travel behavior” and develop a methodology to detect such changes in individual travel patterns. We specify one distribution for each of the three dimensions of travel behavior (the frequency of travel, time of travel, and origins/destinations), and interpret the change of the parameters of the distributions as indicating the occurrence of a pattern change. A Bayesian method is developed to estimate the probability that a pattern change occurs at any given time for each behavior dimension. The proposed methodology is tested using pseudonymized smart card records of 3210 users from London, U.K. over two years. The results show that the method can successfully identify significant changepoints in travel patterns. Compared to the traditional generalized likelihood ratio (GLR) approach, the Bayesian method requires less predefined parameters and is more robust. The methodology presented in this paper is generalizable and can be applied to detect changes in other aspects of travel behavior and human behavior in general.

1. Introduction

There is a constant tension between dynamism and stability of travel behavior. Such tension can be examined at different levels. At the most basic level, travel behavior is described by a series of physical actions associated with individuals, e.g., a trip starting at 8 am going from point A to point B by car. We refer to these observable actions as incidences of travel behavior, or *travel incidences*. Travel incidences vary from day to day but the underlying behavior pattern is more consistent. At a higher level, a *travel pattern* describes the organization of travel incidences over a period of time; each pattern corresponds to a set of preferences and constraints that dictate the specific choices. Although not directly observable, travel patterns may be statistically represented as the distribution of observed travel incidences. Existing work on travel behavior modeling often makes the implicit assumption that a person’s travel patterns are stable, i.e., the distribution of the travel incidences observed in the past will not change in the future. While true in the short term, this assumption is less likely to hold over longer time periods. In the short term within the same pattern (e.g., in days or weeks) the dynamics of travel

* Corresponding author.
E-mail address: jinhua@mit.edu (J. Zhao).

https://doi.org/10.1016/j.trb.2018.03.017
0191-2615/

behavior are solely reflected through variations of travel incidences. This is the focus of numerous studies on intrapersonal variability (Hanson and Huff, 1988; Pas and Koppelman, 1987), and regularity (Williams et al., 2012; Goulet-Langlois et al., 2017). In the long term (e.g., in months or years), however, the travel patterns are also subject to changes, contributing to the long-term behavior dynamics. For example, when people move homes from the suburb to the central city, they may systematically increase the overall travel frequency, shorten the travel distance, shifting the commuting hours, increase the number of locations visited, and reduce car usage. The mechanism of such changes is less studied.

In this work, we define *travel pattern changes* as "abrupt, substantial, and persistent changes in the underlying pattern of travel behavior". This definition is adapted from the concept of "regime shift" in ecology (Biggs et al., 2009). Individual travel patterns may change when people move home, change work, shift work schedules, purchase a new car/bike, or any other events (e.g., a child starts school) that may alter their travel and activity routines. Typically, each change in travel patterns corresponds to a life choice related to housing, employment, education, household structure, car ownership, etc. Unlike travel incidences, *life choices* are long-term decisions that affect the preferences and constraints governing the daily travel and activity patterns. Since these life choices have implications across different aspects of life, changes in travel patterns may co-occur with changes in the behavior patterns in other domains. The strategic combination of these life choices over time also follows a certain pattern, referred to as *lifestyle* in the literature (Krizek and Waddell, 2002). Lifestyle describes preferences towards a particular way of living (Walker and Li, 2007). Over an even longer term (e.g., in years or decades), as one's life cycle and socioeconomic status evolve, the lifestyle can also change. In this paper, we consider the lifestyle as a higher-level construct, and more stable, than a travel pattern. Lifestyle changes always lead to travel pattern changes, but travel pattern changes do not always involve lifestyle changes. Whereas the concept of lifestyle is more difficult to operationalize, travel patterns are relatively easy to measure and analyze. In this paper we focus on travel pattern changes as situated in between lifestyle changes and travel incidence variations.

Prior research on travel pattern changes has been centered around the effect of social, economic, environmental, and attitudinal factors (Arentze and Timmermans, 2008; Albert and Mahalel, 2006; Cao et al., 2007; Verplanken et al., 2008), as well as how to utilize these factors to induce travel pattern changes (Meyer, 1999; Bamberg, 2006). In the literature, panel survey data were often used to model the change of individual travel behavior over time (Goulias, 1999). However, Kitamura et al. (2003) showed that discrete-time panels were not a dependable tool for observing dynamic behavior process, pointing to the need of continuous data. With the increasing prevalence of urban sensing technologies from transportation and other urban systems, individual travel incidences can be continuously captured by various data sources at a large scale and over a long term. A fundamental question is how to properly identify pattern changes from such data. Travel patterns and their changes are not directly observable; they are latent and need to be inferred. Travel incidences often exhibit substantial variability regardless of changes in travel patterns, making it a non-trivial task to infer pattern changes from the noisy stream of individual travel records. Furthermore, multiple behavior dimensions have to be taken into consideration since one may change travel pattern in certain dimensions but not others. The ability to automatically detect travel pattern changes from individual-level longitudinal travel records can provide important insights into the dynamic nature of personal travel demand, and is critical for developing behavior models that are adaptive over time.

The objective of this paper is to develop a methodology to detect changes in individual travel patterns. We formulate it as a change detection problem in the time series analysis, in which we aim to identify the time when the probability distribution of an individual's travel incidences changes. The problem concerns both detecting whether or not a pattern change has occurred, and, if yes, identifying the time points of such changes, referred to as *changepoints*. The proposed methodology considers three behavior dimensions in frequency, space, and time, and adopts a Bayesian approach to estimate the probability of pattern changes in each dimension over time. The methodology is general and can be applied to diverse human mobility aspects and various data sources. In this study, we demonstrate the method using pseudonymized smart card records from London, U.K.

2. Literature review

Change detection (or changepoint detection) in time series analysis is an important problem with a wide range of applications such as quality control (Lai, 1995), biomedical analysis (Bodenstein and Praetorius, 1977), financial prediction (Bai, 1997), and image processing (Radke et al., 2005). More related to transportation, it has also been applied to traffic flow analysis using loop detector data (Yan et al., 2018). However, unlike traffic flow data, individual travel behavior records are discrete, multi-dimensional, and relatively sparse. Therefore, we need a general and robust method to detect pattern changes from travel behavior data.

Over the years, a plethora of algorithms have been proposed for change detection. They generally follow two categories—online and offline methods. Online methods can process data in a sequential fashion; models are incrementally updated as time proceeds and new data arrive, making them suitable for real-time applications. Offline methods, on the other hands, take a global view of the whole data sequence, and as a result are usually more robust than the online methods (Keogh et al., 2001). Because of the dynamic nature of travel behavior, it is important to be able to detect travel pattern changes in a timely fashion, so that they can be useful for applications such as personal travel information, demand management, and dynamic mobility prediction (Zhao et al., 2017). For this reason, we choose to focus on the online methods in this paper.

Many online change detection methods are variations of the generic sliding window approach. Here, we denote a time series as $\boldsymbol{X} = \{x_1, x_2, \ldots\}$, where x_i represents the data point at timestep i. Let us use $x_{i:j}$ to denote a subsequence of the