Travel Behaviour and Society 1 (2014) 69–78

Contents lists available at ScienceDirect

Travel Behaviour and Society

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

Zooming into individuals to understand the collective: A review of trajectory-based travel behaviour studies

Yang Yue [b,*], Tian Lan [a], Anthony G.O. Yeh [a], Qing-Quan Li [b]

[a] Centre of Urban Studies and Urban Planning, Department of Urban Planning and Design, The University of Hong Kong, Hong Kong Special Administrative Region
[b] Shenzhen Key Laboratory of Spatial Smart Sensing and Services, College of Civil Engineering, Shenzhen University, China

ARTICLE INFO

Article history:
Available online 31 December 2013

Keywords:
Human travel behaviour
Trajectory data
Complex network theory

ABSTRACT

Understanding travel behaviour is significant in travel demand management as well as in urban and transport planning. Over the past decade, with the advancement of data collection techniques, such as GPS, transit smart cards, and mobile phones, various types of travel trajectory data are increasingly complementing or replacing conventional travel diaries and stated preference data. Other location-aware data are used in studying human movement patterns, such as social network check-in data and banknote dispersal data. Abundance of the emerging trajectory data has driven a new wave of travel behaviour research, and introduced new research problems. This paper provides a state-of-the-art review of the travel behaviour studies categorised by trajectory data types. Based on the literature review, research challenges are discussed and promising research topics in this field are proposed for future studies.

Introduction

Human movements are the major dynamics of various spatial and temporal phenomena, such as urban commuting, transportation of goods, spread of influenza, and diffusion of automobile pollutants. Thus, the studies of human travel behaviour are crucial across a broad spectrum, including urban planning, transport forecasting, epidemiology, and ecology. Since the early boom of travel behaviour studies in the 1970s, new technologies, tools, and data sources have emerged recently. The objective of this paper is to provide the state-of-the-art review of the travel behaviour studies as categorised by trajectory data types.

Trajectory data is the basis of travel behaviour studies. A trajectory comprises a chronological series of location points at discrete time intervals associated with diverse context-varied attributes, which can be represented using a finite set of triples as follows:

$$Trajectory : \{(P_{T1}, T_1, A_{T1}), (P_{T2}, T_2, A_{T2}), \ldots (P_{Tn}, T_n, A_{Tn})\} \quad (1)$$

where P_{Tn} denotes the spatial position at time T_n and A_{Tn} denotes the associated attributes, such as travel speed, heading, and vehicle status. Among the triples, "location" and "time" are compulsory elements and are the syntactic properties of a trajectory, whereas "attribute" is optional and varies among different scenarios.

Numerous terms have been used in studies that discuss trajectory data, such as space–time path (Hägerstrand, 1970), trip chain (Kondo and Kitamura, 1987), geospatial lifeline (Hornsby and Egenhofer, 2002), tracking data (Asakura and Iryo, 2007), movement data (Long and Nelson, 2012), and spatiotemporal footprints (Cheng et al., 2011). All of these terms are in accordance with the definition provided in Eq. (1). Therefore, the studies reviewed in this paper are selected via their research datasets, which comply with the norm of the trajectory data in Eq. (1).

This study focuses on the trajectory data that are related to human travel rather than on the trajectories that track motions of wildlife or natural hazards, such as hurricanes. The collection, pre-processing, and warehousing of trajectory data are not intensively discussed in this review. For those who are interested in spatial trajectory computation may to refer to the work of Zheng and Zhou (2011).

The remainder of this paper is organised as follows. Section 2 discusses the reasons for the revival of travel behaviour research. The subsequent sections review the travel behaviour research that is categorised as traditional travel survey data, GPS log data, smart card data, mobile phone data, and other trajectory data. Section 8 presents the research challenges and opportunities. Section 9 concludes the paper.

Renaissance of trajectory-based travel behaviour studies

Before the ICT (information and communications technology) making large-scale trajectory collection possible, travel survey is the most commonly-used approach to obtain traveller trajectory. Based on the data collected from travel diary or similar approach, early practitioners and researchers in transport studies established a classical four-step model framework to simulate the human

* Corresponding author. Tel.: +86 755 26531934.
E-mail address: yueyang@szu.edu.cn (Y. Yue).

2214-367X/$ - see front matter © 2013 Hong Kong Society for Transportation Studies. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.tbs.2013.12.002

travel behaviour in transport studies across the United States, such as the Detroit Metropolitan Area Traffic Study, the Chicago Area Transportation Study (CATS), and Pittsburgh Area Transportation Study (PATS). Hägerstrand (1970) initiated a space–time prism approach to analyse human travel activities in his seminal paper. The space–time framework explicitly defines people's movement into a trajectory format. Several pioneers (Chapin, 1974; Cullen and Godson, 1975; Lenntorp, 1976) also contributed their work to this line of time geography literature. However, a gradually fading interest occurred after the 1970s and 1980s' research boom (Timmermans et al., 2002). This research boom was hindered by three factors: data source, analytic tool, and computation capability.

First, travel behaviour models faced a data-hungry but data-poor dilemma (Kwan, 2000; Timmermans, 2003; Miller, 2010). Datasets were often not available and not updated frequently enough for contemporary studies. Moreover, the paradigm had shifted from aggregated modelling to disaggregated modelling (Rasouli and Timmermans, 2013). There was once a research boom evolving from the gravity-based approach (Taaffe et al., 1996), to trip-based and activity-based approaches (Axhausen and Gärling, 1992), which consequentially required more longitudinal individual trajectory data with better spatial and temporal resolutions. Second, only a few effective tools were available to analyse the disaggregated trajectory data at that time. The deficiency of software that handles spatial and temporal data led to an unrealistic uniform assumption about the urban environment (Miller, 1991). Finally, the computation hardware for running these models was expensive and time-consuming. Transport models exhausted the first and second generations of IBM computers easily when regional transport studies were conducted in the 1950s (Weiner, 1999). Calculation and storage power were also limited and sometimes cumbersome (Miller, 2010).

Nonetheless, science and technology have advanced radically since the last series of the aforementioned travel behaviour studies. First of all, the computation hardware capacity has been improved exponentially, as depicted by Moore's law (Moore, 1965). Furthermore, GIS-assisted analytical tools of human spatial behaviour are more readily available and popular among social scientists and researchers in different disciplines. Kwan (2000) suggested that researchers should use their abilities to organise, visualise, and analyse the space–time data in GIS, particularly in three-dimensional models. Goodchild (2000) characterised the behaviour view as one of the three evolution stages of GIS for transportation. Last but not least, human society is experiencing an era of big data. With the development of ICT as well as location-aware technology, an unprecedented volume of data, termed big data by researchers and the industrial, has been harvested and is still streaming daily at an accelerating pace. Batty (2012) characterised big data as massive data streams that were produced in real-time and space by certain novel sensor technologies and new social media. New technological applications in travel survey data have emerged, particularly in human trajectory data (Wolf, 2004). People move around with cellular network-registered smartphones every day. They swipe smart cards to take public transit and buy goods with credit cards. They also check in at various locations and share the information on social networks. Science has moved from data-poor straits to a data-rich environment (Miller, 2010).

Thus, technical advancements have alleviated the three impedance factors mentioned. These technical advancements have also triggered another burst of travel behaviour research, which has led to the renaissance of travel behaviour study.

The trajectory data obtained through different means, such as traditional travel survey data, GPS log data, smart card data, mobile phone data, and other non-conventional sources, such as mobile social media and banknote data, have different characteristics that can be used in trajectory data research. The following sections review related studies according to these five data categories.

Trajectory data for travel behavior studies

Traditional travel survey data

In the early transport studies, traditional travel survey data or travel diaries were mostly derived from regional censuses or household interviews. Several data collection approaches have been developed, from paper-and-pencil interviews by mail to computer-assisted telephone interviews (CATI) and computer-assisted self-interviews. The travel trajectory properties, such as origins and destinations (OD), departure and arrival times, trip purposes, and travel modes, could be extracted from such survey data and then fed into transport models.

A number of related studies have generally paid significant attention to temporal issues rather than spatial ones (Timmermans et al., 1992), such as day-to-day time-use variability of travel activities (Hanson and Huff, 1981; Burnett and Hanson, 1982; Jones and Clarke, 1988; Schlich and Axhausen, 2003), time allocation (Kitamura, 1984b; Wissen and Meurs, 1989), and trip-chaining issues (Adler and Ben-Akiva, 1979; Hanson, 1979; Kitamura, 1984a).

A notable contribution came from Hägerstrand (1970) who created time geography and established a space–time framework for travel behaviour studies. In this seminal work, space–time path and prism were employed to explore individual travel behaviour under capability, coupling, and authority constraints. Chapin (1974) and Cullen and Godson (1975) also contributed to the establishment of the space–time framework by investigating time budgets and activity patterns. Miller (1991) and Kwan (1998) later implemented this space–time prism concept in GIS to measure its accessibility using household survey data. The space–time framework has provided an approach to visualise and explore the patterns of individuals (Shaw et al., 2008; Lam et al., 2012). The time geography empowered with the geo-computation and geo-visualisation capabilities of GIS is a powerful tool for human movement research (Kwan, 2004).

Timmermans et al. (2002) labelled the aforementioned time geography based travel behaviour studies as constraint-based models in their taxonomy of space–time behaviour models. Other space–time models using travel survey data include utility maximizing, computational process (Jiang et al., 2012), and micro-simulation models (Batty et al., 2003).

Although these traditional travel survey data have contributed significantly to classical travel behaviour studies, they are still labour-intensive, error-prone, and not that cost-effective. They are highly expensive to collect, which makes travel survey data not quite updated. In addition, some intermediate trip details are also missing. Criticism has also been against traditional survey data in the literature (Ettema et al., 1996; Axhausen, 1998). Thus, researchers and practitioners have been seeking new data sources, such as GPS logs.

GPS log data

GPS log data are the most widely used form of trajectory data and usually consist of time, longitude, latitude, altitude, direction, and speed, with other information depending on specific cases. For example, the passenger boarding information is collected from the taxis installed with GPS receivers.

The early applications of GPS trajectory data revolved in transport parameter estimation and model calibration. According to Pearson (2001), the first regional travel survey that used the GPS-assisted approach was conducted in Austin, Texas in 1997.