Travel Behaviour and Society xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Travel Behaviour and Society

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

Applying mobile phone data to travel behaviour research: A literature review

Zhenzhen Wang, Sylvia Y. He *, Yee Leung

Department of Geography and Resource Management and the Institute of Future Cities, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

ARTICLE INFO

Article history:
Received 26 July 2016
Received in revised form 2 February 2017
Accepted 28 February 2017
Available online xxxx

Keywords:
Big data
Mobile phone data
Mobility
Travel behaviour

ABSTRACT

Travel behaviour has been studied for decades to guide transportation development and management, with the support of traditional data collected by travel surveys. Recently, with the development of information and communication technologies (ICT), we have entered an era of big data, and many sources of novel data, including mobile phone data, have emerged and been applied to travel behaviour research. Compared with traditional travel data, mobile phone data have many unique features and advantages, which attract scholars in various fields to apply them to travel behaviour research, and a certain amount of progress has been made to date. However, this is only the beginning, and mobile phone data still have great potential that needs to be exploited to further advance human mobility studies. This paper provides a review of existing travel behaviour studies that have applied mobile phone data, and presents the progress that has been achieved to date, and then discusses the potential of mobile phone data in advancing travel behaviour research and raises some challenges that need to be dealt with in this process.

1. Introduction

We have entered an era of big data (Manyika et al., 2011), where companies and organizations are capturing tremendous volumes of data from their customers and partners, and millions of sensors built into various devices are continuously sensing and collecting data from individuals and groups. Big data is believed to have the potential to make our cities smarter by facilitating the discovery and explanation of urban development and its dynamics, which will enable city managers and citizens to make more informed decisions and to enjoy better city life (Steenbruggen et al., 2015). Similarly, in the transportation field, big data has promoted intelligent transportation systems by providing a better understanding of where, when, and how people travel around.

For decades, in order to provide long-term guidance and short-term strategies for urban planning and transportation development, many studies have been conducted to identify, understand and predict human travel behaviour (Buliung and Kanaroglou, 2007; Yue et al., 2014; French et al., 2015). Traditionally, data that support travel behaviour research largely came from travel surveys, and such data are costly to collect and out of date. These shortcomings have restricted data collection and further deterred travel behaviour research progress to some extent (Mitchell, 2014; Liu et al., 2015). In the age of big data, various novel sources of data can be applied to supplement or substitute for traditional survey data to support travel behaviour research. Examples are smartcard records data, GPS-enabled taxi trajectory data and roadside sensor data, and among these, mobile phone data are the most widely applied and promising type (Yue et al., 2014).

Owing to their increasing penetration in the population, and their built-in location and motion sensors, mobile phones are becoming useful tools to collect extensive and dynamic data for human travel behaviour research (Wesolowski et al., 2014). Compared with traditional travel survey data, mobile phone data shows many unique attributes and advantages, such as its unprecedented coverage of population and geographic area. This attracts scholars in various areas to apply mobile phone data to travel behaviour research, and certain achievements have been made to date (Deutsch et al., 2012; Steenbruggen et al., 2013). However, this is only the beginning, and mobile phone data still have great potential that needs to be exploited to further advance travel behaviour studies. At the same time, it should be noted that along with the opportunities it brings, there exist many remaining challenges that need to be dealt with when taking the application of mobile phone data further.

Thus far, many papers have reviewed the application of mobile phone data in travel behaviour studies. For example, Choujaa and Dulay (2009) provided a review of human activity recognition based on mobile phone data. Deutsch et al. (2012) surveyed the

* Corresponding author.
E-mail address: sylviahe@cuhk.edu.hk (S.Y. He).

http://dx.doi.org/10.1016/j.tbs.2017.02.005
2214-367X/

collection of different types of data by various sensors built in smartphones and analysed sensor frequency, activity inference and battery drain. Jiang et al. (2013) presented existing applications of triangulated mobile phone data in spatiotemporal analysis and urban modelling. Calabrese et al. (2014) summarised the use of network-based mobile phone data for urban sensing. Steenbruggen et al. (2015) summarised existing spatial studies based on mobile phone data and explored the possibility of achieving smart city goals with mobile phone data. Yue et al. (2014) reviewed how different types of trajectory data have been applied to travel behaviour studies, including mobile phone data. Liu et al. (2015) reflected on the problems in the collection, processing and analysis of big data applied into spatial information sciences and related fields. In spite of considerable progress, there is still a paucity of review papers that focus on mobile phone data and travel behaviour studies, particularly for the new emerging smartphone sensor-based data.

Therefore, on the basis of an extensive overview of the literature, this paper aims to promote the application of mobile phone data in travel behaviour research by reviewing how different types of mobile phone data have been applied in travel behaviour research, how they may further facilitate travel behaviour research and what remaining challenges need to be dealt with in this process. The remainder of this paper is organized as follows. Section 2 gives a panorama of traditional travel behaviour research and data collection efforts. Section 3 makes a detailed introduction to mobile phone data that can be applied to travel behaviour research, including data collection systems and techniques as well as data sources and corresponding attributes. Section 4 reviews the current progress of applying different types of mobile phone data to travel behaviour research from three aspects, including the identification of travel patterns, the exploration of influencing factors, and the modelling and prediction of travel behaviour. In Section 5, the potential of and challenges to further applying mobile phone data to travel behaviour research are summarised and discussed. The last section presents a conclusion of this paper.

2. Traditional travel behaviour research and data collection efforts

2.1. Traditional travel behaviour research

For decades, considerable effort has been devoted to identifying and characterizing the dynamics of human travel patterns, which are measured as daily trip frequency, trip purposes, departure time, travel duration, travel distance, travel modes, trip sequences or complex trip-chains, trip destinations, travel companions, and so forth (Jackson and Jucker, 1982; McGuckin and Murakami, 1999; Noland and Polak, 2002; Murray-Tuite and Mahmassani, 2003; Scott and He, 2012; He, 2013a,b; He and Hu, 2015; He and Giuliano, 2017). Dynamics in travel patterns are mainly explored in two ways: intra-personal dynamics and inter-personal dynamics (Dharmowijoyo et al., 2014). Intra-personal dynamics include: short-term dynamics in travel behaviour during day and night or on weekdays and weekends; medium-term dynamics in travel behaviour such as before and after residential relocation, change of jobs, owning a car, etc.; long-term dynamics in travel behaviour in different stages of life, including marriage, giving birth, retirement, etc. (Pas and Sundar, 1995). Regarding inter-personal dynamics, gender, race, occupation, education level, income, lifestyle, and household structure have been extensively identified as influential factors (Senbil and Kitamura, 2009; Kang and Scott, 2008, 2011).

In order to explore the underlying rules governing the observed travel behaviour, a wide range of influencing factors have been examined. Demographic and socio-economic attributes of travellers are believed to be key determinants of their travel behaviour (Mauch and Taylor, 1997; Lu and Pas, 1999; Ryley, 2006; Marquet and Miralles-Guasch, 2014; Boarnet and Hsu, 2015); built environment, such as urban morphology, land use and neighborhood design, is regarded as another important factor affecting human travel behaviour (Boarnet and Crane, 2001; Ewing and Cervero, 2001, 2010; He, 2011). Some other factors, such as lifestyle, attitudes, perception and preference, have been extensively explored in the literature (Lanzendorf, 2002; Chen et al., 2009; Etminani-Ghasrodashti and Ardeshiri, 2015; Van Acker et al., 2016). Normally, descriptive methods are applied first to identify the differences in travel behaviour in various settings, and then multivariate techniques are used to find the statistical associations between travel behaviour and each factor.

On the basis of characterizing and understanding human travel patterns, many studies tend to model and predict travel behaviour. Earlier studies established aggregated models to forecast traffic flows across TAZs (traffic analysis zones), such as the trip-based four-step model (Lam and Huang, 1992; McNally, 2007; Yue et al., 2014). Later, micro-simulation of individual travel choices became popular, and various disaggregated models were established based on particular theories and hypotheses, such as the geography constraining theory and the utility maximization theory (Hägerstrand, 1970; Axhausen and Gärling, 1992; Batty et al., 2003; McNally and Rindt, 2007; Malleson and Birkin, 2012; Yue et al., 2014).

2.2. Travel behaviour data collection efforts

For decades, data used in travel behaviour research largely came from traditional travel surveys, which have evolved from questionnaires to recalled travel diaries, and the equipments used have also been improved from papers and pencils, to telephones, emails, computers and the Internet. In the meantime, with the development of travel behaviour research, more and more comprehensive and accurate data are needed to support increasingly complex research (Curtis and Perkins, 2006). However, these traditional travel surveys have obvious shortcomings that restrict the data collection and further deter the progress of travel behaviour research, such as costly to collect, small sampling rates, short survey duration, under-reporting, and quickly out-of-date (Shen and Stopher, 2014; Wolf et al., 2001). Recently, the portable GPS-aided travel surveys are becoming popular, which can collect data of higher quality by continuously and accurately recording respondents' travel trajectories without adding burdens on respondents (Feng and Timmermans, 2014). Nevertheless, the GPS-aided travel surveys still have many shortcomings that have restricted their applications. For example, it is quite expensive to provide GPS devices and the fully productive response rate is low (Wolf et al., 2014; Shen and Stopher, 2014).

Parallel to the evolution of conventional travel surveys, various sources of data with unprecedented volume, termed "big data", have emerged and been applied to support travel behaviour research, such as smartcard records data, GPS-enabled taxi trajectory data, and roadside sensor data. Among these, mobile phone data are the most widely used data source. This is actually a massive data stream collected about the real-time location and displacement of users (Yue et al., 2014). As a type of big data, mobile phone data are also characterized by "5V", including volume, variety, velocity, veracity, and added value[1] (Manyika et al.,

[1] Volume means data systems need to handle large volumes of data; Variety refers to the capability of processing data of different types and sources; Velocity means the ability to handle regularly or irregularly data; Veracity concerns the uncertainty of data, indicating raw data characteristics should be preserved when being processed; Added value means big data can help make informed decisions to create more value (Manyika et al., 2011; Normandeau, 2013).