Chapter 41
Real-Time Monitoring of Intercity Passenger Flows Based on Big Data:
A Decision Support Tool for Urban Sustainability

Zheng Li
Xi'an Eurasia University, China

Yan Wang
Xi'an Eurasia University, China

Qin Chen
Business Big Data (BBD), China

ABSTRACT

Managing population mobility is a key to urban growth and sustainable development. This study uses administrative and business data from a number of trustworthy and publicly-available websites for public transport to access passenger flows in a real-time manner. A case study is used to illustrate the application, with intercity passenger flows by public transport mode (rail or air), by rail service type and by time. Moreover, a model is developed for monitoring the implications of population movements, which can be a decision support tool for governments and policy makers to manage population mobility. The big-data approach to accessing public transport passenger movement has the following characteristics: (1) low cost, (2) a population scale, (3) instantaneous data collection/update, and (4) high quality.

1. INTRODUCTION

Urbanization plays an important role in economic growth, welfare, environment, and changes in social and human behaviour, which presents both opportunities and challenges to sustainability (Bettencourt et al., 2007). Governments across the globe keep a close watch on population movements. One traditional

DOI: 10.4018/978-1-5225-7311-1.ch041

source of data is the census. For example, in China, a national census is conducted every five years. A series of censuses illustrate that 1.17 billion Chinese moved to cities from rural areas between 2000 and 2010, where 35 percent moved to Beijing, Dongguan, Guangzhou and Shanghai metropolitan regions (International Organization for Migration, 2016). However, the cost of a national census is enormous, for example, the direct cost of the 2010 Census of China was 50 billion RMB (about 7.5 billion US$). Another source is research studies mainly based on travel diaries or surveys via phone, web or phone interviews; however, these studies are often associated with sampling issues. In addition to the high-cost and sampling issues of traditional approaches to population flow, other inherent limitations are statics and hysteresis. Therefore, these traditional approaches are incapable of continuously monitoring dynamics in population movements in a real-time manner.

Real-time monitoring of population mobility requires the following data characteristics: (1) low cost, (2) a population scale rather than a sample of the population, (3) instantaneous data collection/update, and (4) reliability. With technological development, big data provides such an opportunity. Big data has three dimensions (Daas and Puts, 2014): (1) Volume, that is the number of data records, attributes and linkages; (2) Velocity, that is the speed that data produced, changed processed and analyzed; and (3) Variety, that is the diversity of data sources, formats, media and contents. Wamba et al. (2015) add two "Vs": Veracity and Value, and defines big data as "a holistic approach to manage, process and analyze 5 Vs in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages" (p. 235). These characteristics are in line with the data needs for population mobility.

The existing big-data studies on population mobility and broader literature of urban studies have mainly used the following sources: mobile phones, transportation devices (e.g., public transportation smartcards and in-vehicle Global Position Systems (GPS)), and social-networking media (see Toole 2005; Williams et al. 2015; Sobolevsky et al. 2015; Tam and Clarke 2015; Hao et al. 2015 for detailed literature review). Compared to the existing studies, a unique property of this study is that administrative data and business data from trustworthy and publicly-available websites are captured in a real-time manner and on a population scale to establish reliable big data on intercity population flow. In addition, other significant advantages include reliability, consistency and accessibility. These merits, in turn, would ensure the quality of data analysis.

The paper is organized as follows. The next section introduces the procedures to generate big data sets on passenger flow of air and rail transportation. Then, Wuhan City is used as the case study to illustrate the applications of big data, investigates the changes in net inflow by mode (air and rail), by service type (intercity and longer-distance high speed rail (HSR)) and time (day and month), followed by a model for monitoring the condition of population mobility. The final section highlights the key conclusions of this study.

2. DATA

Among the existing studies of population mobility using big data, the dominant source is mobile phones (see e.g., Smoreda et al., 2013; Buckee et al., 2013; Smoreda et al., 2013; Zhang and Zhou, 2014; Wang and Tang 2014; Toole 2015 among others), where location data can be obtained by the tower that the phone is connected. Mobile phone data has these merits: (1) spatiotemporal precision, (2) low cost, and