Chapter 4

Identifying the Temporal Characteristics of Intra-City Movement Using Taxi Geo-Location Data

Wenbo Zhang
Purdue University, USA

Xinwu Qian
Purdue University, USA

Satish V. Ukkusuri
Purdue University, USA

ABSTRACT

In this chapter, the authors focus on temporal patterns of urban taxi trips and explore determinant factors in conjunction with geodatabase at aggregate level. Zero-Inflated Negative Binomial model is proposed in light of count data nature and excessive number of O-D pairs with zero trip. Three typical time slots on weekdays, as well as weekends, are introduced as case study to check temporal variations of intra-city movement. The results indicate that trip distance, land use, socioeconomics, and built environment are significant variables that affect the number of taxi trips between two locations. In particular, longer travel and worse economy conditions, such as low employment and average annual income and more population under poverty, may prevent more movements, which have more impacts during peak hours. A better transit system may reduce the taxi trips, except for areas with more subway stations. Developed area for instance more commercial or residential area is more likely to attract more visits by taxis, as well as dense public facilities but with more temporal variations.

DOI: 10.4018/978-1-5225-0827-4.ch004
INTRODUCTION

Intra-city movement is short-range, more frequent and random, and with various trip purposes, which is different from inter-city and international movements. It covers all human trips within a city, regardless of transportation modes. This chapter is to study its temporal characteristics and explore external factors (e.g. demographics, socio-economics, land use, and built environment) that influence amounts of intra-city movement among regions within a city. Characterizing intra-city movement needs much trip information that has good spatial and temporal coverage. However, the traditional survey methods, including Revealed Preference survey, Stated Preference survey, travel diary, and national or regional travel survey, are time-consuming and cost-ineffective, considering the large samples required by the problem (Bhat and Srinivasan, 2005; Yao and Morikawa, 2005). Moreover, some recorded information may be inaccurate during these post-event surveys. Fortunately, the emerging large-scale pervasive dataset are with better temporal and spatial information, which provides a new perspective for exploring intra-city movement. Many studies have confirmed the feasibility of using the pervasive datasets in various transportation problems (Asakura and Hato, 2004; Bohte and Maat, 2009; Chen et al., 2010; Stopher et al., 2008; Zheng et al, 2011). For example, in New York City (NYC), approximately 500,000 taxi trips that transport 0.6 million passengers are all automatically recorded by equipped GPS each day. For each trip, information including fares, trip distance, travel time, origin and destination (OD), start time, and end time are recorded in real time. Although number of passengers by taxis is only one tenth of those by transit (New York City Taxi and Limousine Commission, 2014), this emerging dataset has a better spatial and temporal coverage compared with traditional ones, which yields a better understanding of intra-city movement.

In this chapter, to explore temporal variances and causal factors of intra-city movement, one-week of taxi GPS data from NYC are sampled and aggregated at Zip Code Tabulation Area (ZCTA) level. Meanwhile, various attributes of ZCTAs (e.g. distance, demographics, socio-economics, built environment, land use and commuters) are collected as explanatory variables. Econometric models, such as the negative binomial model (NB) and zero-inflated negative binomial model (ZINB), are introduced and compared, considering the nature of trip data, potential overdispersion, and excessive zeros in the dataset. The whole sample is further divided into 6 subsamples to measure temporal variances in intra-city movement. Proposed models are estimated under each subsample. The remaining sections are listed as follows: the second section summarizes background and related studies; the third section presents proposed methodologies; the fourth section shows data preprocessing and variables selection; the fifth section discusses the estimation results and potential applications; and the last section summarizes our works, limitations, and directions for future study.

BACKGROUND

Knowing the characteristics of intra-city movement is very important in understanding complex space-time dynamics during urban study. The interactions between intra-city movement and land use have been discussing for many years. Waddell (2002) introduced a new model system, UrbanSim, to link planning of land use, transportation, and environmental quality. Han et al. (2011) explored human traveling patterns and found human traveling behaviors were strongly affected by the geographical structure. Thus, understanding the relationships between intra-city movement and land use can help us to develop