Transportation Research Part C 86 (2018) 510–526

Contents lists available at ScienceDirect

Transportation Research Part C

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

A spatial econometric model for travel flow analysis and real-world applications with massive mobile phone data

Linglin Ni[a,b], Xiaokun (Cara) Wang[c], Xiqun (Michael) Chen[a,*]

[a] College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China
[b] Dongfang College, Zhejiang University of Finance and Economics, Hangzhou 314408, China
[c] Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA

A R T I C L E I N F O

Keywords:
Travel flow
Spatial autocorrelation model
Cellular signaling data
Marginal effect

A B S T R A C T

Cellular signaling data provide a massive and emerging source for acquiring urban origin–destination (OD) travel flow information, supporting decision making on large-scale mobility enhancement, and enabling the exploration of factors that influence travel demand. This study investigates the effects of population, facilities, and transit accessibility on travel flows between traffic analysis zones. A spatial econometric model is employed for the OD travel flow analysis by integrating massive mobile data with other explanatory features of urban regions. The results of real-world applications in Hangzhou, China, show that: (I) all coefficients of the origin dependence, destination dependence, and OD dependence are statistically significant, which verifies the consideration of the spatial interdependence in the OD flow modeling; (II) all of the permanent population, number of facilities, and transit accessibility have positive correlations with travel flows; and (III) travel time, as expected, is negatively correlated with the travel flow volume. Finally, policy implications are discussed, further contributing to the design of urban land use and transportation policies.

1. Introduction

Understanding, estimation, and prediction of origin–destination (OD) travel flows on urban transportation networks is essential for government agencies and decision makers to implement planning, operations, management, and allocation of urban resources. Traditional methods for the OD data collection often rely on household surveys and/or traffic counts. However, those approaches may involve expensive data collection efforts and therefore have limited sample sizes and low update frequencies (Ren et al., 2014). Recent advances in mobile devices offer opportunities to acquire the OD travel flow information with a low data collection cost, large sample size, high updating frequency, and broad spatial and temporal coverage (Calabrese et al., 2013). Along with the development of mobile sensing technologies, cell phone signaling data have emerged to be a widely-used resource to measure both individual travel behavior and network demand, e.g., individual human mobility patterns (Calabrese et al., 2013; González et al., 2008), estimation of OD matrices (Iqbal et al., 2014), and OD trip purposes (Alexander et al., 2015).

As indicated by the classic gravity model, OD flows are often proportional to the zonal trip generation, attraction, and inversely proportional to travel impedance. The trip generation and attraction can be further estimated as functions of the population size, employment size, and other land use variables. Overall, it is believed that population is the primary explanatory variable explaining the variation of OD travel flows. If an area's population increases, travel flows from/to this area usually increase. As a key land use

* Corresponding author at: B828 Anzhong Building, College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China.

E-mail address: chenxiqun@zju.edu.cn (X.M. Chen).

https://doi.org/10.1016/j.trc.2017.12.002
Received 23 April 2017; Received in revised form 6 December 2017; Accepted 6 December 2017

variable, the existence of facilities such as commercial centers, hospitals, and transportation hubs in an area is related to the trip attraction (Cervero and Kockelman, 1997). The construction of new roads tends to reduce travel impedance thus increases travel flows. In urban areas, transit accessibility is also a major factor influencing travel impedance. Generally, travel flow increases with the improvement of transit accessibility. The operation of new transit lines that connect two areas often increases trips between these areas.

Numerous empirical studies, mostly derived from the gravity model, have been developed to estimate parameters to quantify the relationship between OD flows and these influencing factors (Curry, 1971; Evans, 1973; Bar-Gera and Boyce, 2003; Calabrese et al., 2011). Spatial agglomeration indicates that similar activities tend to cluster. If people know that they have similar activities in neighboring zones, then they tend to choose the zones more than other zones, which induces spatial correlations. The self-selection theory also encloses that people who tend to use public transit instead of car trips tend to choose the neighborhood with good bus services (Handy and Clifton, 2001; Mokhtarian and Cao, 2008). Therefore, urban travel flow exhibits an extensive spatial autocorrelation theoretically. It is essentially caused by the origin dependence, destination dependence, and OD dependence. Origin dependence means that travel flows from the neighboring origin to the same destination are spatially influential to each other. Similarly, destination dependence means that travel flows from the same origin and to neighboring destinations are also spatially interdependent. OD dependence means travel flows that don't share the same origin or destination could even be spatially interdependent, as long as their origins or destinations are neighbors.

To the best knowledge of the authors, few existing studies have systematically investigated factors influencing urban OD travel flows between traffic analysis zones (TAZs) considering such spatial autocorrelations. This paper uses an OD spatial autocorrelation model (SAM) to quantitatively identify the effects of population, urban facilities, transit accessibility, and travel distance on travel flows, using massive cellular signaling data collected from millions of mobile phone users in Hangzhou, China, during September 7–13, 2015.

The rest of this paper is organized as follows: Section 2 summarizes methodologies and findings from the previous literature, followed by the description of the mobile phone data in Section 3. The spatial effect of travel flows and the spatial autocorrelation model used for this analysis are formulated in Section 4. Section 5 presents the results with the spatial effects of the population, facility, and transit accessibility. Section 6 discusses policy implications. Section 7 concludes the whole paper and discusses future research.

2. Literature review

2.1. Impacts of land use on urban travel flow

Ample evidence exists that land use and travel flow are strongly associated. Vichiensan et al. (2007) enclosed that the interaction between land use and transportation was quite strong in the developing countries, through analyzing the urban railway development of Bangkok. Pan et al. (2009) suggested that residents travel behaviors and mobility demands were influenced effectively by land use and urban design, based on 1709 individuals travel survey in Shanghai. Stead (2001) proposed that land use had a significant effect on travel patterns, though socioeconomic variables were more significant by using data from national and local travel surveys in Britain. Hong et al. (2014) enclosed that land use factors had highly significant effects on vehicle miles traveled (VMT), by building Bayesian hierarchical models controlling the travel attitude and spatial autocorrelation, using the 2006 Household Activity Survey and the 2005 parcel and building data in the Seattle metropolitan region.

Land use variables include the population, land use mix, land use balance, employment density, percent of commercial buildings, distance to the CBD and so on (Boarnet and Sarmiento, 1998; Ewing and Cervero, 2001).

First, the population is an important factor to influence the travel flow. More people mean the higher travel demand for work, shopping or leisure. Many studies found the population density was significant with travel flow (Holtzclaw, 1994). For example, Frank and Pivo (1994) found that the influence of the population density and land use mix entropy were consistently positive and significant for both work and shopping trips. Kitamura et al. (1997) examined the effects of land use on travel behavior for five diverse San Francisco Bay Area neighborhoods and concluded that the population density, public transit accessibility, mixed land use, and the presence of sidewalks were significantly associated with the trip generation. Saelens et al. (2003) proposed that travel mode would be influenced by the population density, connectivity, and land use mix. People from areas with higher density, greater connectivity, and more land use mix would be more inclined to walking or cycling travel than low-density, poorly connected, and single land use neighborhoods, through reviewing the transportation, urban design, and the planning literature between 1990 and 2002. Brownstone and Golob (2009) explored that the population density was lower, and the average automotive miles traveled of the household was longer through comparing two California household neighbors with different residential densities using the 2001 U.S. National Household Transportation Survey. Kim and Brownstone (2013) found that the higher-density population decreased the average vehicle usage. Boarnet and Sarmiento (1998) explored that the impact of population density on non-work automobile trips was rarely statistically significant, using the travel diary data for Southern California residents. Bento et al. (2005) enclosed that a 10% increase in population centrality, was estimated to reduce 1.5% annual VMT.

Second, the land use mix is associated with the travel flow (Kitamura et al., 1997; Saelens et al., 2003). According to the properties of land use, there are commercial, medical or transportation facilities, which are all important to affect the travel flow. Commercial centers and hospitals are the facilities to attract the travel flow. Handy (1994) analyzed 34 commercial centers in the San Francisco Bay Area and concluded that the high levels of both local and regional accessibility to commercial centers were associated with the shorter trip lengths, but not with the fewer trips, by using the 1981 Metropolitan Transportation Commission travel survey