# Understanding travellers' preferences for different types of trip destination based on mobile internet usage data

Yihong Wang[a,*], Gonçalo Homem de Almeida Correia[a], Bart van Arem[a],
H.J.P. (Harry) Timmermans[b,c]

[a] Department of Transport and Planning, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands
[b] Department of the Built Environment, Section of Urban Systems and Real Estate, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
[c] Department of Air Transportation Management, Nanjing University of Aeronautics and Astronautics, China

A B S T R A C T

New mobility data sources like mobile phone traces have been shown to reveal individuals' movements in space and time. However, socioeconomic attributes of travellers are missing in those data. Consequently, it is not possible to partition the population and have an in-depth understanding of the socio-demographic factors influencing travel behaviour. Aiming at filling this gap, we use mobile internet usage behaviour, including one's preferred type of website and application (app) visited through mobile internet as well as the level of usage frequency, as a distinguishing element between different population segments. We compare the travel behaviour of each segment in terms of the preference for types of trip destinations. The point of interest (POI) data are used to cluster grid cells of a city according to the main function of a grid cell, serving as a reference to determine the type of trip destination. The method is tested for the city of Shanghai, China, by using a special mobile phone dataset that includes not only the spatial-temporal traces but also the mobile internet usage behaviour of the same users. We identify statistically significant relationships between a traveller's favourite category of mobile internet content and more frequent types of trip destinations that he/she visits. For example, compared to others, people whose favourite type of app/website is in the "tourism" category significantly preferred to visit touristy areas. Moreover, users with different levels of internet usage intensity show different preferences for types of destinations as well. We found that people who used mobile internet more intensively were more likely to visit more commercial areas, and people who used it less preferred to have activities in predominantly residential areas.

## 1. Introduction

There is a recent trend in complementing or even replacing traditional travel survey data with new mobility-related data sources, such as GPS data, mobile phone traces and smart card transaction data (Chen et al., 2016; Demissie et al., 2013a; Iqbal et al., 2014; Ni et al., 2018; Toole et al., 2015; Wang et al., 2017; Wolf, 2006; Yue et al., 2014; Zhao et al., 2018). These trajectory-based data are getting popular for travel analysis because (1) they are inexpensive to collect; (2) they are usually up to date; and (3) most of them contain a large sample with observations that are longitudinal in time (Calabrese et al., 2013; Demissie et al., 2013b; Morency et al.,
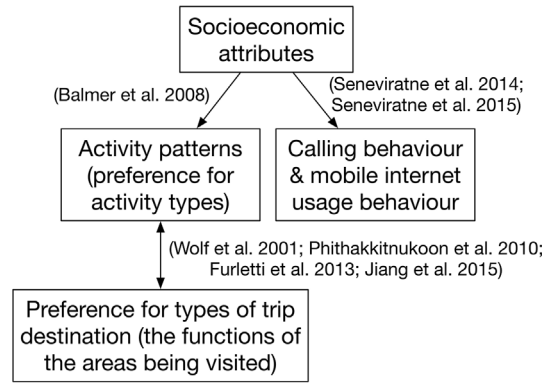
---

Fig. 1. The conceptual framework.

2007).

However, despite the potential advantages, these sources of information only include the spatial-temporal traces describing people's movements. If the aim is to understand travel behaviour from an activity-based perspective (Chen et al., 2016; Rasouli and Timmermans, 2014; Zhao and Zhang, 2017), the information of these data sets is usually very limited. For example, activity purpose of the trips is typically missing (Calabrese et al., 2013). Moreover, in traditional travel demand models, socioeconomic information is used to segment the population, and better explain the heterogeneity of activity-travel behaviour, including, but not limited to, activity patterns (Balmer et al., 2008) and location choice (Sivakumar and Bhat, 2007). However, in anonymous big data, socio-economic information is unavailable mainly due to privacy reasons (Calabrese et al., 2014).

To deal with such problems, researchers have tried to combine different types of data in order to fill the gaps (Anda et al., 2017). In attempting to derive activity purpose information from trajectory data, there have been several applications fusing trajectory data with land use data, OpenStreetMap data or point of interest (POI) data (Dashdorj et al., 2013; Demissie et al., 2015; Wolf et al., 2004; Yuan et al., 2012). This geo-coded background knowledge can help estimating the function of an area, which can tentatively be connected to the type of activity that a visitor performed in that area (Furletti et al., 2013; Jiang et al., 2015; Phithakkitnukoon et al., 2010; Wolf et al., 2001). We referred to the main function of an area being visited as "type of trip destination" in this paper. The left chain in Fig. 1 shows how we derive the dependency of one's preference for destination types on socioeconomic attributes, based on literature review. Intuitively, such dependency exists in most cities. For example, it is common that some specific urban areas are more frequented by young people.

To partition the population using mobile phone data, Arai et al. (2014) and Bwambale et al. (2017) suggested utilizing calling behaviour such as calling frequency and duration to predict one's personal attributes. However, mobile phones are less used for calls today, making calling behaviour less useful, while simultaneously people are spending more time on services provided by mobile internet such as mobile apps (Richmond, 2012). Therefore, mobile internet usage behaviour, if available, could have a greater potential to reflect individuals' traits, such as gender and age (Seneviratne et al., 2015, 2014). The right chain in Fig. 1 shows the dependency of mobile internet usage behaviour on socioeconomic attributes.

As a whole, Fig. 1, which can be regarded as a conceptual framework, shows the relationship between mobile internet usage behaviour and preference for types of trip destination. Since they are both dependent on the socioeconomic attributes, even if the socioeconomic attributes are unobserved, they are still likely to be correlated with each other. Based on this hypothesis derived from the conceptual framework, our study aims to understand travellers' preferences for types of trip destination by means of segmenting them based on the preferred type of sites and applications visited through mobile internet as well as the level of visiting frequency, by fusing mobile phone traces and mobile internet usage data. We are allowed to do this study because of the data provided by the Shanghai Unicom WO+ Open Data Application Contest.[1]

Furthermore, mobile internet usage behaviour might sometimes be able to reflect even more information about a person, such as one's specific interests and lifestyles, than the traditional socioeconomic attributes do. At the same time, one's interests and lifestyles are regarded as the determinants of location choice through preference for different types of non-work activities (Wen and Koppelman, 2000). A more specific interest or lifestyle might be related to a more specific travel preference especially for non-work activities. For example, a foodie would visit more sites and applications about food, and meanwhile, he/she would also like to visit more restaurants in real life. We see the potential to explore such relationships by fusing mobile internet usage data and mobile phone traces, and we especially focus on the types of destinations for out-of-home non-work activities, designated herein as secondary activities for simplicity. Many studies have used mobile phone data to analyse users' home and workplace locations as well as commuting trips (Ahas et al., 2010; Alexander et al., 2015; Calabrese et al., 2011; Isaacman et al., 2011). However, trips for sec-ondary activities have not often been analysed using this type of data, except in only a few studies (e.g., Järv et al., 2014; Huang and Levinson, 2015), which does not mean that they are not an important part of urban travel demand. In fact, they are taking a larger share than ever before, especially in large metropolitan areas (Wang et al., 2017).

---

[1] https://www.kesci.com/woplus/.