



Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges [☆]



Taha H. Rashidi ^{a,*}, Alireza Abbasi ^b, Mojtaba Maghrebi ^d, Samiul Hasan ^c, Travis S. Waller ^a

^a School of Civil and Environmental Engineering, UNSW, Australia

^b School of Engineering and Information Technology, UNSW, Australia

^c Department of Civil, Environment, and Construction Engineering, University of Central Florida, United States

^d Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad, Khorasan Razavi, Iran

ARTICLE INFO

Article history:

Received 30 May 2016

Received in revised form 12 December 2016

Accepted 15 December 2016

Keywords:

Travel diary survey

Social media

Travel demand modelling

Mobility behaviour

ABSTRACT

In the past few years, the social science literature has shown significance attention to extracting information from social media to track and analyse human movements. In this paper the transportation aspect of social media is investigated and reviewed. A detailed discussion is provided about how social media data from different sources can be used to indirectly and with minimal cost extract travel attributes such as trip purpose, mode of transport, activity duration and destination choice, as well as land use variables such as home, job and school location and socio-demographic attributes including gender, age and income. The evolution of the field of transport and travel behaviour around applications of social media over the last few years is studied. Further, this paper presents results of a qualitative survey from travel demand modelling experts around the world on applicability of social media data for modelling daily travel behaviour. The result of the survey reveals positive view of the experts about usefulness of such data sources.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The digital age accelerated the evolution of online social networks. Social media has become an emerging industry with massive input and output cash flow. As a result, massive data sources have been created as a result of such massive market. Harnessing such big data has become an interesting topic for researchers, scientists, practitioners and governments. Fields such as computer science, mathematics, social sciences, economics and management have invested considerable effort in developing understanding about various aspects of social networks and media data. It has been only recently that transport engineering, urban planners and travel demand modellers have noticed the richness of such big data and have started exploring the capacity of such data source for planning, management and operating purposes.

Initial movements towards understanding social media and their impact on the transport system started with descriptive analysis on mobility using location based social networks (Onnela et al., 2011). As the potentials of such data sources further explored, transport modellers pushed the frontiers of applications of social media data for modelling transport related issues (Hasan and Ukkusuri, 2015; Hasan et al., 2016). Nonetheless such efforts are still at their infancy and the community is not

[☆] This article belongs to the Virtual Special Issue on "Social Network".

* Corresponding author.

E-mail addresses: rashidi@unsw.edu.au (T.H. Rashidi), a.abbasi@unsw.edu.au (A. Abbasi), Mojtabamaghrebi@um.ac.ir (M. Maghrebi), samiul.hasan@ucf.edu (S. Hasan), s.waller@unsw.edu.au (T.S. Waller).

yet convinced about full potential of such cheaply available but costly to prepare dataset (Wu et al., 2014), where privacy of users must be maintained through aggregate or anonymized parsing analysis (Smith et al., 2012).

There are several complications associated with using social media data, especially if analysing the content of such a big data is of importance in understating the observations. For example, Twitter¹ data (tweets) typically contain normal text, hash-tag(s), and/or check-in data. Check-in data include location of tweets, making it associated with activities happening at that location (e.g., all tweets linked to a stadium, by users who provided checked-in data, are more likely to be related recreational activities). Similarly, hash-tag (#) messages are associated with an activity, event, location, etc. Therefore, it is relatively easier to work with check-in and hash-tag data as they are already associated with an event or location (Katakis et al., 2008). In particular, when check-in data is used for analysis of the destination/origin of the activity, determining trip purpose is relatively easy (Cheng et al., 2011). More information about applications of Twitter data can be found in a review paper by Steiger et al. (2015) where transport is completely excluded from their study. If check-in data or hash tag data is not of interest and more general information is used, extracting meaningful information can be challenging. More importantly, there are several biases and issues highlighted affecting the research on human mobility behaviour in different ways for some of which solutions have been proposed in fields such as epidemiology, statistics, and machine learning (Ruths and Pfeffer, 2014).

This study presents an overview of transport related studies which used social media for transportation planning and management. A special focus is given to the application of social media data in travel demand modelling studies. Relevant studies focusing on applications of social media on the following categories that are related to transport research are discussed in Section 2: (i) travel demand modelling, (ii) mobility behaviour (iii) individuals' activity pattern, (iv) assessing public transport and (v) traffic condition, (vi) and incidents and natural disasters. Section 3 presents a discussion about the evolution of evolution of social media use for transportation applications. This section is followed by a more detailed discussion about the capacity of social media data through results of an online survey in which travel demand modelling experts declared their opinions about usefulness of different social media data sources for planning, management and operation purposes. Finally a summary of the discussion and recommendations for future directions of using social media data in the field is discussed.

2. Use of social media in transport research

2.1. Travel demand modelling studies

The history of planning the transport system infrastructure goes back to the time the wheel was invented followed by the construction of the first paved road in Sumer in 500 BCE. At the same time, Darius I the Great, 500 BCE, started construction of an extensive road system for Persia including the famous Royal Road which was one of the first highways. About the same time, Roman roads were constructed with advanced technologies of stone-paved and metaled, cambered for drainage and were flanked by footpaths, bridleways and drainage ditches. Same road structure was later used by the Great Britain in the 18th century to establish the first toll system which included 250 miles of road and 40 bridges. All of these early transport system planning and network design efforts inspired transport engineers of the 20th century to develop a systematic procedure for policy appraisal and network design purposes. It was in the 1950s when the first prototypes of the conventional four-step models developed in Chicago and Detroit in USA. Since then, many metropolitan areas adopted a similar structure to evaluate the short, medium and long term consequences of different designs and policies. The 4-step modelling paradigm, which is a trip-based approach, led to the tour-based scheme in which individual level travel information is regarded for modelling purposes. Tour-based models were later evolved to activity based model in which individual/household level data is used to model individual/household level travel attributes (Rashidi and Kanaroglou, 2013).

Travel demand modelling techniques target modelling the mobility (movement) of people and vehicles (including passenger and commercial vehicles) in cities to understand their (mainly short distance) travel behaviour. Models are developed based on individual level data sources, in which behaviour of travellers is reflected, have been argued to dominate aggregate level models in terms of policy appraisal (Rashidi and Kanaroglou, 2013).

The evolution of travel demand modelling techniques developed the need for high resolution databases in which socio-demographic and economic attributes of people are used to model their day-to-day travel behaviour. Such data sources encompass travel diary of a sample of people representing the population. Having access to such an individual level travel diary is crucial to develop several components of the advanced behavioural modelling frameworks like tour-based and activity-based. The most important travel attributes considered in these modelling frameworks are: (a) trip purpose, (b) departure time, (c) mode of transport, (d) activity duration, (e) activity location, (f) travel route, (g) party composition, and (h) traffic condition

Other than travel data, information about long-term household decisions should be collected and modelled to be used as an important input to travel demand models. The major household decision for which commonly data is collected and models are developed are: residential location, job location and vehicle ownership. Among these three, vehicle ownership has been modelled more in travel demand frameworks. Housing and job search behaviour have been mainly considered exogenously in the travel modelling structures (Rashidi et al., 2012).

¹ www.twitter.com.

Download English Version:

<https://daneshyari.com/en/article/4968660>

Download Persian Version:

<https://daneshyari.com/article/4968660>

[Daneshyari.com](https://daneshyari.com)