Chapter 9

A New Algorithm for Mode Detection in Travel Surveys: Mobile Technologies for Activity–Travel Data Collection and Analysis

ABSTRACT

New technologies offer various opportunities for advancing travel surveys. This chapter presents a new approach for automated identification of trip stages and travel modes as the core outcome from travel surveys and a key requirement for subsequent steps, such as the automated assignment of trips. Mode prediction of eight modes of transport is realized by two multinomial logistic regression models, based on only nine features from GPS and acceleration data. The algorithm achieved an overall detection rate of 79 percent. The authors found that motorcycle and moped, railway, bicycle, and pedestrian obtained better results, whereas urban public transport caused some difficulties in detection.

INTRODUCTION

Travel surveys are an appropriate means for collecting disaggregate data on people’s travel behaviour, in particular the number of trips, as well as their origin and destination, duration, length, mode, and purpose of travel. Traditional paper-pencil and telephone interviews are widely used and well established in person-based travel surveys. The main issues with these methods are the high response burden, the unknown extent of missing, wrong or inaccurate data and the
challenges in collecting data for more than one day (Stopher et al., 2008; Kelly et al., 2013). Kohla and Meschik (2013) showed that trips are underreported by as much as 25% on average in traditional trip diaries.

New technologies offer various opportunities for advancing methods of data collection on travel behaviour. Mobile devices such as GPS-trackers or Smartphones are used to record useful sensor data, such as positions over time to describe respondents’ movement patterns (Reddy et al., 2010; Schüssler, 2010; Stopher et al., 2011; Feng et al., 2011). User interfaces of devices provide active tracking so that respondents may annotate their trips even on-trip using Smartphone applications. Algorithms for automated detection of start and end of trips and trip stages as well as for the derivation of trip length, duration, and modes are required to reduce response burden. The goal is to extract complete and reliable information on travel behaviour based on passively collected information to support active tracking or as the basis for prompted recall interviews. Additional information (user habits, spatial data or public transport supply) may improve this information and give hints on trip purposes (Reddy et al., 2010; Rudloff & Ray, 2010; Feng et al., 2011; Rieser-Schüssler et al., 2011; Stopher et al., 2011; Widhalm et al., 2012).

Algorithms to detect trip stages and travel modes for each stage are currently tested by several researchers as the core outcome of travel surveys and a key requirement for subsequent steps, such as the automated assignment of trips (Bohte & Maat, 2008; Troped et al., 2008; Reddy et al., 2010; Feng et al., 2011; Rieser-Schüssler et al., 2011; Stopher et al., 2011; Yu et al., 2013). This chapter presents a new procedure for automated identification of trip stages and travel modes from sensor data. Our aim is to develop algorithms for all steps from data cleaning to mode detection. We have a unique dataset including reference data available for training models and testing algorithms. The developed procedure finally comprises algorithms for automated mode detection that are (1) based on data delivered by GPS-module, accelerometer and user information, (2) robust, so that they need not to be trained anew for new applications and (3) flexible for extensions and advancements when additional data (e.g. from gyroscopes or compasses) becomes available.

The remainder of this chapter is organized as follows: Section 2 describes the data basis including the method for data collection and main characteristics of the used mobile devices. Section 3 presents the hypotheses for developing the algorithms of mode detection. Each hypothesis is discussed including the relevant literature that fed into hypothesis building and a description of features derived from the hypothesis. These features are the basis for model training in section 5. Section 4 discusses the modes to be predicted with the developed models. Multinomial regression models are developed in section 5 as the core part of the method for automated mode detection. Two models are developed: A pedestrian model is used to identify trip stages. Identified stages are fed into a general model in the second step for detecting all modes of transport, included in the algorithm. Both models were embedded in an overall algorithm of automatic mode detection, which is described in section 6. The chapter ends with a discussion of the results and an outlook for further research in section 7.

DATA

The reference data for this paper comes from the MobiFIT pilot project (Herry et al., 2011), which combined different survey designs to analyse the applicability of GPS technology in travel surveys. The survey was conducted between November 2009 and February 2010 in two Austrian regions: one rural region (‘Tullnerfeld,’ about 41,000 inhabitants) and one urban region (city of Graz, about 260,000 inhabitants). Four groups of voluntary