Transportation Research Part C 37 (2013) 118–130

Contents lists available at ScienceDirect

Transportation Research Part C

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

Transportation mode recognition using GPS and accelerometer data

Tao Feng *, Harry J.P. Timmermans [1]

Urban Planning Group, Department of the Built Environment, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

ARTICLE INFO

Article history:
Received 29 May 2012
Received in revised form 12 August 2013
Accepted 27 September 2013

Keywords:
Accelerometer
Activity type
Bayesian Belief Network
GPS
Transportation mode

ABSTRACT

Potential advantages of global positioning systems (GPS) in collecting travel behavior data have been discussed in several publications and evidenced in many recent studies. Most applications depend on GPS information only. However, transportation mode detection that relies only on GPS information may be erroneous due to variance in device performance and settings, and the environment in which measurements are made. Accelerometers, being used mainly for identifying peoples' physical activities, may offer new opportunities as these devices record data independent of exterior contexts. The purpose of this paper is therefore to examine the merits of employing accelerometer data in combination with GPS data in transportation mode identification. Three approaches (GPS data only, accelerometer data only and a combination of both accelerometer and GPS data) are examined. A Bayesian Belief Network model is used to infer transportation modes and activity episodes simultaneously. Results show that the use of accelerometer data can make a substantial contribution to successful imputation of transportation mode. The accelerometer only approach outperforms the GPS only approach in terms of the predictive accuracy. The approach which combines GPS and accelerometer data yields the best performance.

1. Introduction

During the past years, GPS-based technology has proven its usefulness in collecting activity-travel diary data. Professional organizations are now discussing possible replacement of traditional travel survey methods by GPS data collection. Compared to conventional paper-based or phone-based data collection methods, GPS technology has been argued to reduce respondent and researcher burden, while the accuracy of the data would be better than that of conventional survey methods. The latter, however, does not necessarily apply to all facets of activity-travel diaries because transportation modes and activity types need to be imputed from the GPS traces. Such traces may contain errors and moreover (semi)-automated imputation algorithms are not perfect.

Previous studies have adopted different imputation methods, including ad hoc rules (Du and Aultman-Hall, 2007; Bohte and Maat, 2009; Stopher and Wargelin, 2010), regression models (Rudloff and Ray, 2010) and learning-based algorithms (Moiseeva et al., 2010). These studies have shown that all these methods are only partially successful in correctly identifying transportation modes and activity purpose. All these approaches rely on speed and time information which are extracted from the GPS traces. These traces may contain errors in some situations. For example, when traveling underground or in urban canyons, GPS signals may not be received accurately or may be even completely lost. As a result, GPS data may be

* Corresponding author. Tel.: +31 40 247 2301; fax: +31 40 243 8488.
E-mail addresses: t.feng@tue.nl (T. Feng), h.j.p.timmermans@tue.nl (H.J.P. Timmermans).
[1] Tel.: +31 40 247 3315; fax: +31 40 243 8488.

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.09.014

incomplete or inaccurate, causing problems for the correct imputation of particular facets of activity-travel patterns. In addition, the use of speed may in some cases result in misclassifications. When measured or calculated speed served in a range which is feasible for multiple transportation modes, no algorithm will be able to perfectly discriminate between transportation modes on this piece of information only. In practice, this often happens for fast walking and slow biking, for bus and car on a congested road and for tram, light rail, and bus. A prompted recall process may compensate for such incomplete or erroneous data, but it involves participation and additional respondent burden in confirming imputed activity-travel diaries. Moreover, prompted recall data are not necessarily error-free either (Feng and Timmermans, 2013). Therefore, a better imputation process, especially in contexts where speed information is not sufficiently discriminating between transportation modes and/or the GPS signal becomes a problem, would be beneficial to the data collection process.

Accelerometer data provide such an opportunity to enrich the data. This technology is not sensitive to the problems mentioned above and could therefore be used in a complementary fashion to GPS data. An accelerometer is a sensor that returns a real valued estimate of acceleration along the x, y and z axes from which velocity and displacement can be estimated. It can capture data independent of the exterior situation. Accelerometers have been used to identify the type of people's physical activity (Bao and Intille, 2004; Ravi et al., 2005), such as walking, running, sitting and relaxing, watching TV, scrubbing, brushing teeth and climbing. The technology has also been used as motion detectors for body-positioning and posture sensing (Ravi et al., 2005). Recent research has attempted to combine GPS and accelerometer data to recognize physical activities (Wolf et al., 2006; Troped et al., 2008; Cooper et al., 2010; Oliver et al., 2010). A few studies have also attempted to detect transportation modes using accelerometer data from smart phone sensors (Reddy et al., 2010; Wang et al., 2010; Xu et al., 2011). However, these studies either covered a limited number of physical activities or transportation modes (e.g. Troped et al., 2008; Cooper et al., 2010; Reddy et al., 2010), or have presented only illustrative findings (e.g. Oliver et al., 2010). Furthermore, none of these studies paid attention to the simultaneous detection of transportation modes and activity episodes.

Thus, this relatively scant knowledge about the potential advantage of the combined use of accelerometer and GPS data to infer transportation mode and activity episode suggests that research on this issue is timely and relevant. In this paper, therefore, we report the findings of a study which aimed at examining the identification of transportation mode by combining accelerometer and GPS data. We adopt a learning-based Bayesian Belief Network (BBN) model to investigate three different approaches: GPS data only, accelerometer data only and the combination of both types of data.

The remainder of the paper is organized as follows: Section 2 will briefly describe the data and the GPS device used in this study. Then, the possibility of accelerometer data to impute transportation mode will be discussed in Section 3. Next, the improved imputation model will be presented in Section 4. Section 5 will present the results. Finally, Section 6 will summarize and conclude this paper.

2. Data and the GPS device

The data was collected in the context of a project contracted between National Center for Social Research (Natcen) and Eindhoven University of Technology and funded by the Department of Transport (Dft) to assess the feasibility of GPS-only data for the new National Travel Survey (NTS) for Great Britain. According to our previous experience (Moiseeva et al., 2010), we adopted the Bayesian Belief Network approach to infer transportation modes and activity episodes. A small sample of volunteers working for these organizations was used for the present study. GPS data were recorded for every second for almost all transportation modes available in London area. Carefully recorded activity-travel diaries were considered as the ground truth. Activity-travel diaries are not necessarily error-free, but considering the critical importance of this pilot study for the organization, error is expected to be relatively small. In total, 80,670 data records were available as a training dataset, used for model development. We identify the conditional probability between input and output at the epoch level (in this case each second).

The device used for data collection was an accelerometer-enabled GPS device, named MobiTest GSL. The device is equipped with a broad range of sensors including three accelerometer sensors and a GPS sensor. The device has a battery life of up to 100 h, and its internal memory ranges between 128 Mb and 512 Mb for approximately 3 months storage of data (MGE, 2009). The GPS measurements include longitude, latitude, height, number of satellites, time, date, HACC and VACC. The accelerometer measurements represent accelerations along three axis (XACC, YACC and ZACC) and the state on whether the device is moving relatively (no acceleration detected in three directions). The accelerations are accurate in 10 Hz.

The speed information which is popularly available in many existing GPS devices however was not recorded by this device. On the other hand, the distance information was recorded only when certain threshold of the measurement accuracy was satisfied, which was rare in practice. Therefore, we propose an algorithm to approximate the instantaneous distance and speed for the imputation of transportation modes.

3. Accelerometer for transportation mode detection

In domains other than transportation, studies using accelerometer data have reported accuracies up to 80–90%, although most studies have been conducted in laboratory settings. Bao and Intille (2004) conducted an experiment using five accelerometers placed at different places of the body instantaneously to check the sensitivity of the accelerometer device for 20