Automatic segmentation and classification of movement trajectories for transportation modes

Filip Biljecki*, Wilko Quak, Hugo Ledoux, Peter van Oosterom

Department of GIS Technology
OTB Research Institute for Housing, Urban and Mobility Studies
Delft University of Technology, the Netherlands

1 Introduction

The knowledge of the used transportation mode (e.g., bike, train, car, and boat) while traveling is critical for disciplines such as travel behaviour research, transport planning, traffic management, and public transportation planning [1, 2, 7, 10].

In the past, researchers collected the information of the used transportation mode through paper diaries filled by participants or telephone surveys, which often resulted in underreporting of short trips and inaccurate data [3, 5, 8, 11]. With the recent advancements in the Global Positioning System (GPS), movement data can be collected with handheld GPS receivers which usually results in higher accuracy, and more data which can be integrated in a GIS environment for an analysis. Figure 1 shows an example of a trajectory layered over a satellite image, with the information for one sampled point. Series of positions and timestamps are acquired, from which additional data such as speed and heading can be derived.

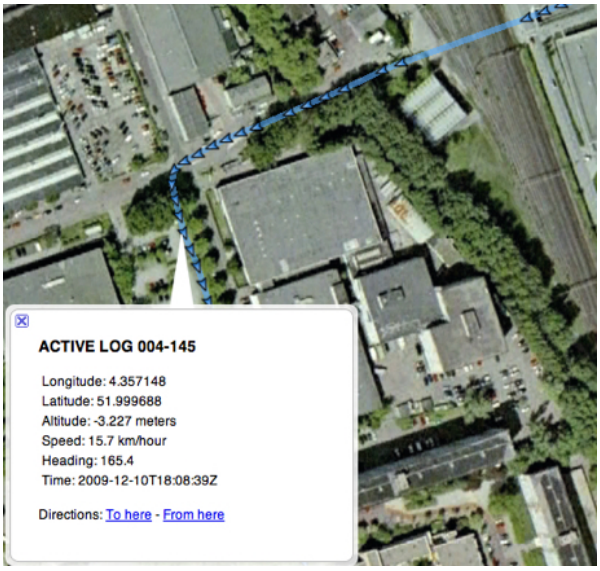

Figure 1: A GPS dataset with the information for one sampled point (©Google/Tele Atlas)

However, data such as transportation mode and trip purpose can never be acquired directly with a GPS receiver, unlike with travel diaries or telephone surveys. Due to massive datasets manual classification is not possible.

This on-going research concentrates on developing a method that automatically enriches the GPS logs with the information of the used transportation mode. We describe briefly in the following our approach.

Current methods classify only few modes, they are deterministic, and GIS data, which may be valuable, is seldom used.

2 Methodology

Since some datasets may contain trajectories made by more than one transportation mode, multimodal data should be first partitioned (segmented) for single-mode subsets by detecting the change in the behaviour of the statistical functions. Figure 2 shows the change in the acceleration after a mode was changed (from a train to a car).

One method to determine the used transportation mode is to assign a bijective function for various statistical descriptors derived from the acquired movement data, which may give a significant implication of the used transportation mode, for instance the maximum speed in the trajectory. However, since of overlapping characteristics of such descriptors for different transportation modes (e.g., train and car may have a comparable maximum speed), this method is limited since it may distinguish only between few modes.

The on-going project investigates the usage of the following properties of the movement trajectories for different transportation modes: acceleration, trip distance, speed, number of stops during travel, the length and the duration of the trajectory,

*email: f.biljecki@student.tudelft.nl

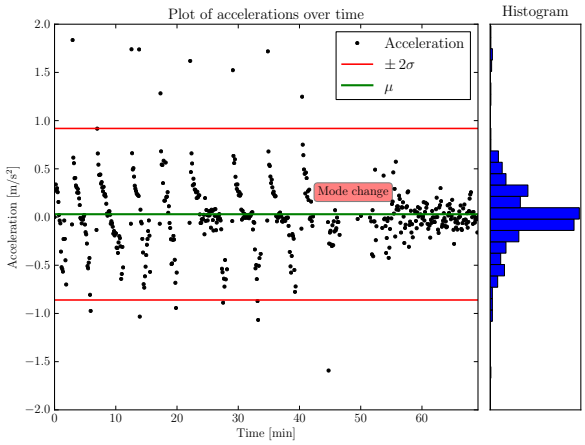

Figure 2: Different behaviour in the acceleration for two different transportation modes (train and car). Since the acceleration is different after the 50th min, the trajectory can be segmented there.

heading changes, and proximity of the trajectory to the GIS network data such as railways, highways, canals, and bus stops.

The latter is critical, since matching the trajectories with the corresponding GIS data infrastructure, such as bus lines, may assist distinguishing modes that have similar properties, such as the speed in car and bus in urban areas.

3 Approaches

This project involves the use of a fuzzy rule-based reasoning system based on analysing available knowledge (facts) with a degree of uncertainty [4, 6, 9]. In such rule-based system membership functions assign a certainty factor to each class, unlike deterministic solutions of classification.

For example, if the mean speed of a trajectory is 30 km/h, a rule could infer from this fact that a car was *probably* used, but might also imply a (faster) bicycle ride, therefore uncertainty is present and a unique (crisp) answer cannot be derived. Fuzzy logic reasoning should be introduced to help us to solve the problem. Next, an empirically derived membership function then assigns a certainty factor to each transportation mode, e.g., for a car 0.8, and for a bicycle 0.2. The combination of multiple membership functions for different facts conclude the final result.

Other investigated approaches are the classification with Support Vector Machines, and dynamic Bayesian networks (i.e., Hidden Markov model).

It should also be noted that there are several problems with the GPS data acquisition that may bias the results, such as GPS signal shortage which leads to gaps in the data, and GPS inaccuracy which involves noisy measurements. Such problems are addressed in the classification system.

References

[1] Y. Asakura, J. Tanabe, and Y. Lee. Characteristics of positioning data for monitoring travel behaviour. *7th World Congress on Intelligent Transport Systems, Torino*, Jan 2000.

[2] W. Bohte and K. Maat. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transport Res C-Emer*, 17(3): 285–297, Jan 2009.

[3] S. Bricka and C. Bhat. Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1972:9–20, Jan 2006.

[4] B. G. Buchanan and R. O. Duda. Principles of Rule-Based Expert Systems. *Advances in Computers*, 22:1–62, Jul 1982.

[5] G. Draijer, N. Kalfs, and J. Perdok. Global Positioning System as data collection method for travel research. *Transportation Research Record: Journal of the Transportation Research Board*, 1719:147–153, Jan 2000.

[6] J. M. Garibaldi. *Fuzzy Expert Systems*, volume 173, pages 105–132. Springer-Verlag Berlin Heidelberg, 2006.

[7] J. Lester, P. Hurvitz, R. Chaudhri, C. Hartung, and G. Borriello. MobileSense - Sensing Modes of Transportation in Studies of the Built Environment. *International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems - UrbanSense08*, pages 1–5, Oct 2008.

[8] P. McGowen and M. McNally. Evaluating the Potential To Predict Activity Types from GPS and GIS Data. *Transportation Research Board 86th meeting*, Jan 2007.

[9] R. A. Orchard. FuzzyCLIPS Version 6.04A User's Guide. *Institute for Information Technology, National Research Council Canada*, pages 1–96, Oct 1998.

[10] P. Ranjitkar, T. Nakatsuji, G. Gurusinghe, and Y. Azuta. Car-Following Experiments Using RTK GPS and Stability Characteristics of Followers in Platoon. *ASCE Conference Proceedings*, 245 (40632):608–615, Jan 2002.

[11] J. Wolf. *Using GPS data loggers to replace travel diaries in the collection of travel data.* PhD thesis, Georgia Institute of Technology, Jan 2000.