Towards Volunteered Digital Travel Demand Data

M. Zilske, K. Nagel

Transport Systems Planning and Transport Telematics, Technische Universität Berlin, Salzufer 17-19, 10587 Berlin, Germany
Email: {zilske; nagel}@vsp.tu-berlin.de

1. Introduction

Behavior-based transport modeling relies on travel diaries from a sample of the population of the study area. A travel diary is an account of all activities that a person performs within a day, with locations and times, and of all the trips taken between locations, with mode of transport.

Such travel diaries are conventionally acquired by means of pen-and-paper questionnaires. This method carries the problem that many trips are not reported because they are deemed too unimportant by the subject to warrant reporting. This can be improved by using specialized data collection devices equipped with GPS receivers, which also allow data entry at the time activities or trips are commenced or finished (Battelle Transport Division 1997; Wolf et al., 2004; Auld et al., 2009).

There have also been studies using completely non-interactive collection of GPS data with the aim of estimating all considered attributes of trips and activities from the trajectory. In a large-scale study (Axhausen et al., 2004), trip attributes were estimated based on land-use indicators and demographic data. Since no additional information about the trips was available, the results were statistically compared with classical mobility surveys.

Both approaches have drawbacks: Being required to enter data before and after each trip is considered a burden, even if it is assisted by an intelligent device, while with the analysis of passively obtained trajectories, important attributes of trips and activities can only be statistically estimated. A synthesis of these approaches is the prompted recall survey. Passively collected trajectories are analyzed, visualized, and presented to the subject for validation and completion. In newer such studies (Auld et al., 2009; Lee-Gosselin et al., 2006), this has been done by means of Internet applications. Survey participants upload the acquired trajectories on their own and are immediately prompted to answer follow-up questions about them. Delays from letter delivery or manual preprocessing are eliminated. The Internet application determines times and locations of activities from the trajectory. They are presented to the participant, who can then remove badly recognized activities and add missed activities. There are also algorithms for an automatic recognition of the means of transport (Chung & Shalaby, 2005) and for the trip purpose (Wolf et al., 2001; Bricka & Bhat, 2006).

2. Travel Surveys with Commodity Hardware

The GPS devices commonly used in travel surveys are not consumer products. They must be acquired from the study budget, distributed, and explained to the participants. The massive proliferation of smartphones seen in recent months raises the question if this is still necessary. In contrast to specialized GPS devices, smartphones are already available and familiar to a rising number of people. The familiarity extends to using location-based services (Mascolo 2010). Smartphones are often quipped with GPS receivers, but can also locate the user by the alternative means of WiFi network signatures and the mobile phone network, which conserves the battery and, in contrast to GPS, also works inside buildings. Moreover, with Google Latitude (http://www.google.com/latitude), there is an application platform which lets users acquire their

movement history, share it, and (in a limited way) even edit and analyze it. The client software is available for all major smartphone platforms. To the knowledge of the authors, there is only one publication on the acquisition of behavioral data from Google Latitude (Ferrari & Mamei, 2011), where it was shown that daily routines ("usually goes to sports club X in their lunch breaks") could be reconstructed from data acquired over a large timeframe.

Our work in progress is to develop a survey process for the acquisition of travel diaries, based on smartphones and the Google Latitude platform, which can be used to acquire large data sets with low costs per user. The Latitude application is pre-installed on Android handsets. It records the current location of the user at an adaptive sampling rate (Figure 1). Running in the background, it still admits the typical smartphone charge cycle of 24 hours.

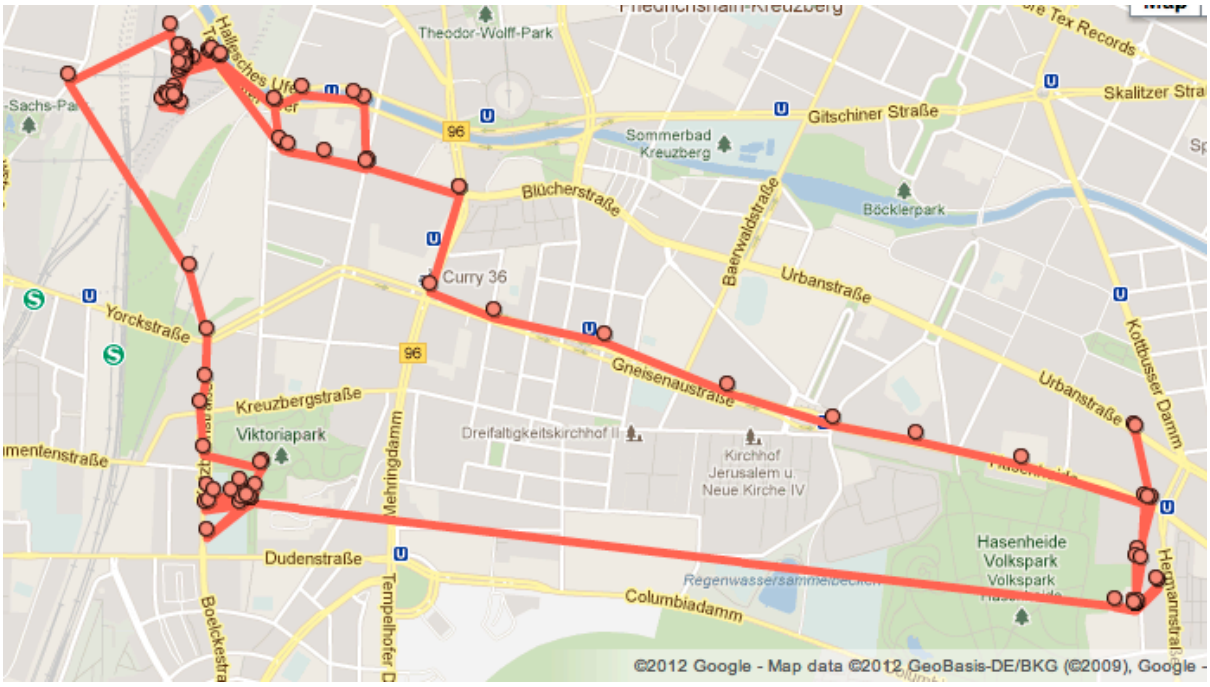

Figure 1. One-day trajectory of one of the authors as recorded by Google Latitude.

To this end, we are developing an Internet application to create annotated travel diaries from such trajectories. Survey participants choose the times at which they use the application at their own discretion. They hand in their recorded Latitude trajectory and are given the possibility to edit and annotate activities and trips. The goal is that the participant will in many cases only have to acknowledge the correctness of the items in a schematic representation of a travel diary (Figure 2) without making changes.

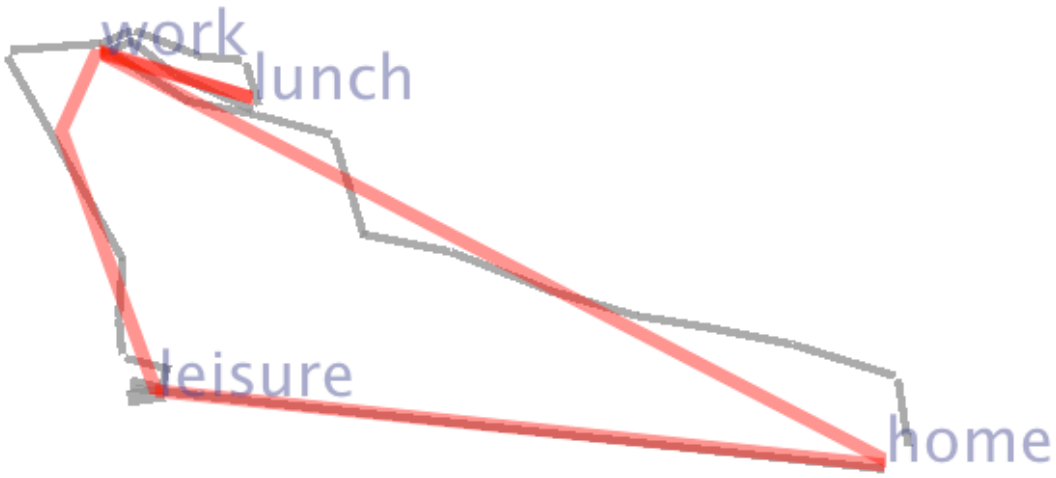

Figure 2. The spatial structure of a travel diary, overlayed with the trajectory.

In this figure, the red polyline connects the activity locations, which were determined by simply clustering consecutive recordings which lie within a small area, and then tagging clusters where more than five minutes were spent as activity locations. The activity types were added manually. Note that while all activity locations of this day were indeed identified, the trace contains one "false positive" (the leftmost vertex, without label).

Our proposed system consists of a user interface which admits efficient editing of such a suggested travel diary, and a supervised learning algorithm which quickly improves the quality of suggestions. It has been demonstrated, for example by Liao et al. (2007), that a machine learning model trained with a set of labeled trajectories can be used to label trajectories from other people with high accuracy. Note that for the purpose of travel diaries, a detailed reconstruction of routes through the road network is not necessary. The more modest but still challenging task of determining activity locations and modes of transport is sufficient.

When used in place of a pen-and-paper survey, the approach described here may lead to better data since there will be less reporting bias: The user will not have to recollect the travel diary from memory, but will be presented with an initial suggestion which may contain trips which would otherwise have been forgotten. In that way, it resembles a GPS-based prompted recall survey. For government-initiated surveys, building trust may be an issue. This may better be achieved using a smartphone and a transparent Internet platform with which the user is already familiar, than with a GPS device that most people have not seen before. In particular, in the medium term it will be necessary to address privacy concerns by fuzzifying certain locations directly at the source. This process would be made visible to the participant by presenting a preview of the final data before it is handed in for the survey.

3. Voluntarily Contributed Travel Diaries

Apart from being used as a surveying tool, we plan to further develop the prototype into a system which can continuously acquire travel diaries from the public at large. The idea is to leverage the existing familiarity with the smartphone platform and create game-like incentives (McKenzie

2011) for the voluntary donation of mobility data. Web-based tracking services for aspects like sports activities, nutrition and even mood are becoming popular, and they are being used because these platforms create value for their users. In our case, we would begin with giving users a structured presentation of their own mobility behavior, including amounts of time spent for different types of trips, and offer alternatives, like switching to public transit.

Voluntarily contributed trajectories will be biased: It is to be expected that only a certain type of personalities will participate, and that they will select data only from certain days. At present, this is not found to be a sufficient deterrent to abandon this line of research. Data from voluntarily contributed trajectories could be fused with other data, such as cellphone data from mobile phone operators, or traffic counts, or aggregated time use surveys. The time use survey provides unbiased but non-localized activity patterns, while the cellphone data would lead to an initial origin/destination pattern. The data investigated in the present paper could presumably be used to enrich *some* of these patterns, and traffic counts could be used to correct for biases (see Flötteröd et al. (2010) for the theory, and Moyo & Nagel (2011) and Flötteröd et al. (2011) for first applications).

4. Conclusion

We outlined a prompted recall mobility survey method based on smartphones and an Internet application which could be used in place of a conventional pen-and-paper survey with selected participants by a city council or a transport agency. The same method can be used to continuously acquire travel diaries from volunteers from the public at large. While we do not expect that travel diaries collected from voluntary samples alone will be sufficient to drive a realistic travel model, we believe that sufficiently powerful methods exist to usefully integrate them with other data sources to create a model without a formal travel survey.

References

Battelle Transport Division (1997): Global Positioning Systems for Personal Travel Surveys. Lexington Area Travel Data Collection Test. Final Report to Office of Highway Information Management, Federal Highway Administration, Washington, D.C., USA.

Wolf, J., Bricka, S., Ashby, T., Gorugantua, C. (2004): Advances in the Application of GPS to Household Travel Surveys. Presented at the Transportation Research Board National Household Transportation Survey Conference, Washington D.C., USA.

Auld, J., Williams, C., Mohammadian, A., Nelson, P. (2009): An Automated GPS-Based Prompted Recall Survey With Learning Algorithms. *Transportation Letters*, 1(1), 59-79.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., Timmermans, H. (2010): Quality assessment of location data obtained by the GPS-enabled PARROTS survey tool. *Journal of Location Based Services*, 93-104.

Wolf, J., R. Guensler and W. Bachman (2001): Elimination of the Travel Diary: An Experiment to Derive Trip Purpose From GPS Travel Data. *Proceedings of the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Axhausen, K.W., Schönfelder, S., Wolf, J., Oliveira, M., Samaga, U. (2004): Eighty Weeks of GPS Traces: Approaches to Enriching Trip Information. *Proceedings of the 83th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Lee-Gosselin, M.E., Doherty, S.T., Papinski, D. (2006): An Internet-based Prompted Recall Diary with Automated GPS Activity-trip Detection: System Design. *Proceedings of the 85th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Chung, E-H., Shalaby, A.S (2005): Development of a Trip Reconstruction Tool for GPS-Based Personal Travel Surveys. *Journal of Transportation Planning and Technology*, 28 (5), 381-401.

Bricka, S. Bhat, C.R. (2006): Comparative Analysis of GPS-Based and Travel Survey-Based Data. *Transportation Research Record: Journal of the Transportation Research Board*, 1972, 9-20.

Ferrari, L., Mamei, M. (2011): Discovering Daily Routines from Google Latitude with Topic Models. *IEEE Workshop on Context Modeling and Reasoning*, 432-437.

Liao, L., Fox, D., Kautz, H. (2007): Extracting Places and Activities from GPS Traces. *International Journal of Robotics Research*, 26.

McKenzie, G. (2011): Gamification and Location-based Services. *Workshop on Cognitive Engineering for Mobile GIS*, Belfast, ME, USA.

Flötteröd, G., Bierlaire, M., Nagel, K. (2010): Bayesian demand calibration for dynamic traffic simulations. *Transportation Science*, 45, 541-561.

Moyo Oliveros, M., Nagel, K. (2011): Automatic Calibration of Microscopic, Activity-Based Demand for a Public Transit Line, *Proceedings of the 91st Annual Meeting of the Transportation Research Board*, Washington D.C., USA.

Flötteröd, G., Chen, Y., and Nagel, K. (2011): Behavioral calibration of a large-scale travel behavior microsimulation, Working Paper, Technische Universität Berlin.