Learning Mobility Profiles: an Application to a Personalised Weather Warning System

Maximilian Leodolter
AIT Austrian Institute of Technology GmbH
Mobility Departement
Giefinggasse 2, 1210 Vienna, Austria
maximilian.leodolter@ait.ac.at

Christian Rudloff
AIT Austrian Institute of Technology GmbH
Mobility Departement
Giefinggasse 2, 1210 Vienna, Austria
christian.rudloff@ait.ac.at

ABSTRACT

Learning mobility profiles of citizens can play a crucial role in many applications, including traffic demand estimation, urban planning or personalized advertising. In this paper we demonstrate a framework for building and constantly readjusting mobility profiles using smart phone data coupled with manual user input and personalised discrete choice models. The methods are applied as weather warning service supporting the daily mode choice decisions of users of the system by supplying personalised information based on their mobility profile and current weather conditions. Since it is well known that weather conditions influence the traffic demand and the modal split of transport modes, the framework can also further the understanding of mobility patterns and their variability due to weather or traffic events.

Categories and Subject Descriptors

H.4.2 [**Information Systems Applications**]: Types of Systems—*Decision support*

General Terms

Algorithms, Experimentation, Human Factors

Keywords

Discrete choice model, mobility behaviour, profile learning, weather warning

1. INTRODUCTION

Learning mobility profiles of citizens can play a crucial role in many applications, including traffic demand estimation, urban planning or personalised advertising. Progress has been made both in the area of mobility profiles as well as the estimation of travel demand. For Mobility profiles this is done by applying data analysis techniques to data sources like mobile phone data or data collected using smart phones (GPS, accelerometer data). In [1] mobility profiles and patterns are learned from mobile phone log data to estimate location time distributions for the users. While the advantage of using mobile phone data is that large samples of users can be reached, it is difficult to attach important information like mode choice to the profiles. First steps to wards enriching such mobility profiles with activity data are taken e.g. in [7]. Some works exist on improving the quality of information of mobility profiles created from mobile phone data using GPS data [9].

Another approach to reach mobility profiles is that of applying mobility diaries. While the sample sizes are considerably smaller than in the case of mobile phone data the rise of smart phone availability makes the collection and enrichment of the data through mode detection easier and enables the collection of larger samples. In [5] such a system for collection and data analysis is described in some detail. While the collection of mobility diaries becomes simpler as technology improves, large scale mobility surveys are currently still collected mostly through interviews. As a result, while this data is useful to model mobility choice behaviour, for example under the influence of weather [6], it is usually collected for only one day for each participant and hence presents a snapshot of peoples' behaviour.

To learn more about the long term behaviour of people data needs to be collected more long term and new methodologies for the analysis of the data need to be developed. In this paper we demonstrate some missing steps to get from collected long term data to mobility profiles and personalised mode choice models. This framework can be applied to build and constantly readjust mobility profiles of users with data collected with smart phones (GPS-data, accelerometer data) coupled with some manual user input. Finally, this framework is applied as personalised weather warning service. The system supplies personalised information based on the mobility profile enhanced by individualized discrete choice models and current weather conditions to the user to support their daily mode choice decisions. Since it is well known that weather conditions influence the traffic demand and the modal split of transport modes, the framework can also further the understanding of mobility patterns and their variability due to weather or traffic events.

The paper will first present the data collection and the learning algorithms for the mobility profiles in section 2 before describing the personalisation of the mode choice models. In section 3 the final weather warning system will be presented. The final section will give some conclusions and an outlook to future work.

2. LEARNING THE PROFILE

As a basis for the learning process of the mobility profile, GPS and other cell phone data (e.g. accelerometer data, cell position) were collected on the users daily trips. For each recorded trip there is a data preprocessing step before the main learning process. This cleans the data and detects the transport mode by transforming the collected GPS, acceleration and cell phone signal data, following the methods of [5]. The following section 2.1 explains the actual data collection part on the user side. The learning methodology will be demonstrated in 2.2. The methodology is such that initial mobility profiles are calculated once at least four trips are collected and are subsequently refined with new data whenever trips are recorded.

2.1 Data Collection

Within a three week long experiment users were asked to record their daily trips with a smart phone application, which was developed especially for this very experiment. The application was easy to handle and designed to require minimum user interaction. Four steps had to be done per trip: (1) start the trip by opening the application and starting the recording, (2) state the travel purposes, (3) end the trip by opening the application again and finishing the recording, (4) adjust the detected modes instantaneously on the device. In between steps three and four trips were automatically cut into uni modal stages and the mode was detected as described above.

2.2 Learning Algorithm

The profile learning algorithm consists of three steps; (1) identify points of routine (POR), (2) identify routine trips (RT) connecting the PORs, and (3) learning about the user's travel behaviour by modelling their choice situations (see section 2.3). On a regular basis the user profiles are refined by applying the learning algorithm including newly collected records.

2.2.1 *Point of Routine, POR*

For every trip the GPS positions are recorded, so the origin and destination can be defined as the first and last GPS points of a trip. Consolidating these by an agglomerative hierarchical cluster method the PORs are defined as the clusters' centers. Since the main emphasis is detecting travel patterns and routine trips, only clusters containing a minimum number of four points are considered as POR, whereas smaller ones may build another one as soon as more trips start or end in the same region. The size of a cluster was defined by the clustering height, that is the agglomeration was stopped at a height of 500 meters.

The cluster analysis was performed in R, applying agnes [4] with the Euclidean metric and average distance concept. Due to the fact that users state the travel purpose these can be assigned to PORs and, for example, the home and work place could be located. Assuming a user starts her first trip per day at home, the algorithm used the location of the home POR for adjusting the weather warning (see section 3).

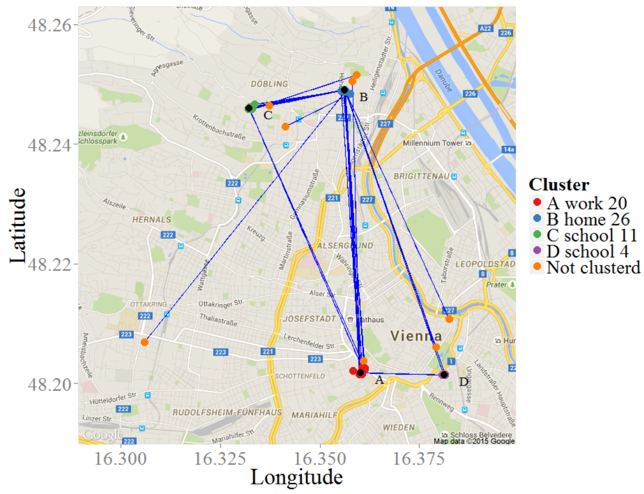

Figure 1: Mobility Profile of one user. PORs are labelled with letters from A to D. The legend also indicates a POR's purpose and cluster size. Orange dots are origin/destination points without any cluster assignment.

2.2.2 *Routine Trip, RT*

Once a profile includes at least two PORs, the connecting trips are aggregated to formulate routine trips. Each RT contains information about the usual travel time, main transport means and earliest and latest departure time for two PORs. A trip's main transport mean is defined as the one chosen for the biggest distance. In this work the following main transport means are considered: walk, bicycle, public transport and car. Consequently for one RT up to four alternatives can be learned. Knowing the available alternatives for a RT is crucial for formulating:

- the choice set in choice model estimation (see section 2.3), and
- the weather warning service (see section 3).

Figure 1 shows a user's profile in Vienna with the points of routine and all trips, most of which connect the PORs and some starting or ending in locations without a cluster assignment.

2.3 Personalised Choice Model

For the algorithm of personalising mode choice models we follow the methodology of [3], applying the idea of individual level parameters [8]. In [2] it was shown that the prediction quality of the personalised models clearly benefits from a base model for a large sample population. Hence, a mixed logit approach was applied to a combined data set. The data consisted of a large travel diary data set for the Vienna region (see [6] for details of the data set and data preprocessing) combined with the data collected by the App. The App data was added whenever a user travelled on a RT with at least two possible alternative modes for that trip. For trip i of user n, a utility to use mode m is given by

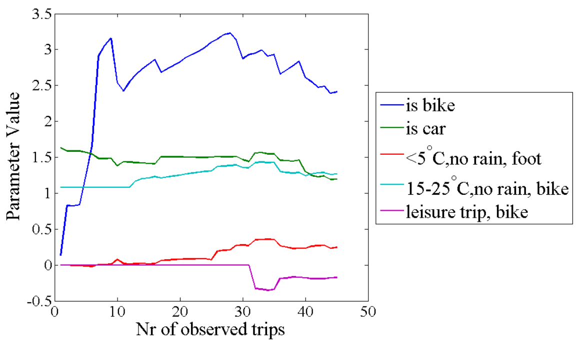

Figure 2: Changes in some parameter values for the personalised model of one user.

$U_{ni}(m) = X_{nim}\beta + \epsilon_{nim}$ where X_{nim} is a vector of decision variables for using mode m for that trip, $\beta \sim N(\mu, \Sigma)$ are the normally distributed parameters from the mixed logit model with mean μ and covariance Matrix Σ and ϵ is a extreme value distributed random error.

To get to the personalised models the choice situations of the user are used by

1. draw a sample β^R of size R from $N(\mu, \Sigma)$, where R is a large number ($R = 25000$ in our implementation)
2. calculate the personalised parameters $\check{\beta}$ as

$$\check{\beta} = \sum_{r=1}^{R} \beta^{(r)} \frac{P(y_m|U_{ni}(m), \beta^{(r)})}{\sum_{s=1}^{R} P(y_m|U_{ni}(m), \beta^{(s)})}$$

where y_m is one if m is the chosen mode and zero otherwise and $P(y_m|U_{ni}(m), \beta^{(b)}) = \frac{exp(U_{ni}(m))}{\sum_k exp(U_{ni}(k))}$, where the sum in the denominator is over all routes observed up to this point for the user.

An example of the changes in parameter values for one user can be seen in Figure 2. The user is an avid bike rider. As a result the alternative specific constant for bike rises quickly to a value between 2.5 and 3. This raises the likelihood that the model predicts the mode bike for that user. One can also see that at low temperatures without rain the person is more likely to walk than the average user and that biking becomes more likely for higher temperatures without rain. For some parameter values it can be seen that they stay at constant values until a trip of that category is observed before there are changes in that parameter value.

3. WEATHER WARNING SERVICE

For applying the learning algorithm in a real world experiment users were provided with a smart phone application that collects and transmits trip data and receives personalised weather messages. The main purpose was to provide the users with information whenever it would be helpful for their decision process, rather than sending redundant messages, so that the user stayed motivated to continue using the app and recording trips. Finally, this would sharpen the mobility profile and result in more relevant messages to prevent users from experiencing unpleasant weather situations

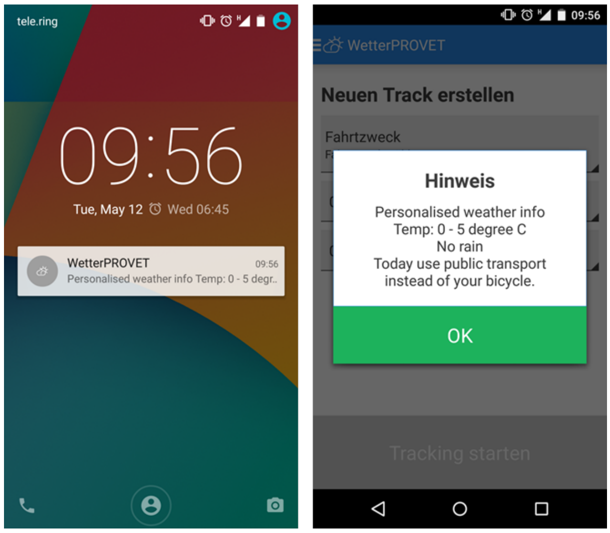

Figure 3: Example of a personalised weather warning message at the smart phone's lock screen (left) and after opening within the application with more information (right).

for their chosen modes, i.e. prevent them from choosing a mode that they would not have chosen with perfect information. Based on the personalised choice model the user's utilities for different transport modes were calculated, both for the current weather forecast and optimal weather conditions (no rain and 15-25 degrees C). For each RT only utilities of modes are compared, that the user has already recorded for this RT. In case of expected behaviour change the user was informed about the weather forecast and a favoured transport mode choice, otherwise a generic weather message was sent. Figure 3 depicts a personalised message, both on the lock screen of the smart phone and within the application. The experiment showed that the learning algorithm performed as expected. It learned the users' PORs and RTs as demanded and adjusted the personalised mode choice parameters continuously. Ten users took part and recorded on average 40 trips each. Two users left the experiment in an early stage, so that no PORs could be recognised. For each of the remaining eight users the home location could be identified. Further the algorithm found 6 education related PORs, 3 work PORs and one for leisure purpose. For 6 users the algorithm could learn on average 3.2 RTs, whereas for each of these one to four alternatives (walk, bicycle, car, public transport) with different main transport means were identified.

4. CONCLUSIONS AND OUTLOOK

In this project algorithms were developed that create and update a user mobility profile, estimate personalised mode-choice models and finally provide users with weather and mode information supporting their daily mobility decisions. The system was tested in a field test showing promising results. Due to the stable nice weather during the experiment time span of March 2015 as well as the limitations of the underlying mobility survey data-set that was collected in late spring, learning the weather sensitivity of the users' mobility behaviour was challenging. However the individual level parameters approach learning about sensitivities to weather events can be incorporated into the models quickly after observing such an event. Furthermore due to the short test period, the mobility profiles lack some grade of detail.

Further steps will deal with developing the algorithm to send weather messages dependent on the weekday. Therefor a longer data collection and learning period is necessary. Consequently for sharpening the profile also older trips could be skipped for keeping the input information for profiling up to date. For a more precise evaluation method user interaction would be required. This could either be achieved by a special survey or directly integrated in the smart phone application.

Furthermore, to improve the system, automated start and stopping of the data collection would limit the reliance on the user's willingness to record trips and would improve profiling and would better enable the next step of suggesting alternative routes rather than just alternative modes. Lastly, the integration of other data like traffic events would improve the system further.

5. ACKNOWLEDGMENTS

This research was partially supported by the research project Wetter-PROVET. The project was funded through the FFG - Austrian Research Promotion Agency program Ways2Go by the Austrian Ministry for Transport, Innovation and Technology (BMVIT). We also thank the project partners Fluidtime Data Services GmbH and UBIMET GmbH.

6. REFERENCES

[1] M. A. Bayir, M. Demirbas, and N. Eagle. Mobility profiler: A framework for discovering mobility profiles of cell phone users. *Pervasive and Mobile Computing*, 6(4):435–454, 2010. Human Behavior in Ubiquitous Environments: Modeling of Human Mobility Patterns.

[2] P. Campigotto, C. Rudloff, M. Leodolter, and D. Bauer. Personalized and situation-aware multimodal route recommendations: the favour algorithm. *IEEE Transactions on ITS*, sumbmitted, 2015.

[3] T. Eiter, T. Krennwallner, M. Prandtstetter, C. Rudloff, P. Schneider, and M. Straub. Semantically enriched multi-modal routing. *International Journal of Intelligent Transportation Systems Research*, pages 1–16, 2014.

[4] M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. *cluster: Cluster Analysis Basics and Extensions*, 2015. R package version 2.0.1 — For new features, see the 'Changelog' file (in the package source).

[5] P. Nitsche, P. Widhalm, S. Breuss, N. Brändle, and P. Maurer. Supporting large-scale travel surveys with smartphones âĂŞ A practical approach. *Transportation Research Part C: Emerging Technologies*, 43, Part 2(0):212–221, 2014. Special Issue with Selected Papers from Transport Research Arena.

[6] C. Rudloff, M. Leodolter, D. Bauer, R. Auer, W. Brög, and K. Kehnscherper. Influence of weather on transport demand: A case study from the vienna region. *Transportation Research Record: Journal of the Transportation Research Board*, To appear.

[7] J. L. Toole, M. Ulm, M. C. González, and D. Bauer. Inferring land use from mobile phone activity. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, UrbComp '12, pages 1–8, New York, NY, USA, 2012. ACM.

[8] K. Train. *Discrete Choice Models with simulation.* Cambridge University Press, Cambridge, UK, 2003.

[9] P. Wang, T. Hunter, A. M. Bayen, K. Schechtner, and M. C. Gonzalez. Understanding Road Usage Patterns in Urban Areas. *Sci. Rep.*, 2, Dec. 2012.