International Environmental Modelling and Software Society (iEMSs)
7th International Congress on Environmental Modelling and Software
San Diego, California, USA, Daniel P. Ames, Nigel W. T. Quinn, Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

Investigating mobility styles using smartphones: advantages and limitations according to a field study in Southern Switzerland

Andrea E. Rizzoli [a], **Roman Rudel** [b], **Anna Förster** [c], **Giorgio Corani**[a], **Francesca Cellina**[b], **Luca Pampuri**[b], **Roberto Guidi**[c] **and Andrea Baldassari**[c]

[a] *IDSIA USI/SUPSI, CH-6928 Manno, Switzerland ({andrea,giorgio}@idsia.ch)*

[b] *SUPSI ISAAC, CH-6952, Canobbio, Switzerland ({francesca.cellina, luca.pampuri, roman.rudel}@supsi.ch)*

[c] *SUPSI ISIN, CH-6928, Manno, Switzerland ({anna.foerster, andrea.baldassari, roberto.guidi }@supsi.ch)*

Abstract:

Diffusion of smart mobile devices offers unprecedented opportunities to monitor travel behaviour, by means of the GPS devices they are equipped with: using a suitable application, potentially every smartphone owner can produce huge, inexpensive quantities of data suitable to profile her mobility patterns. We take advantage of this opportunity within the e-mobiliTI project, which aims at analysing the main psychological and behavioural barriers affecting the transition to new mobility solutions. The project sets up a ”living lab” made up of around twenty families and gives them the opportunity to test electric cars and bikes, public transport season tickets and car and bike sharing. Beside traditional social research tools (questionnaires, interviews, focus groups), to analyse their mobility behaviour we use a specifically developed smartphone application.
In this paper we present the results of the first phase of our field trial and discuss the major challenges faced so far in the automatic gathering of mobility data: high battery consumption, limited performances of the GPS smartphone devices, problems in the Internet connectivity, limited reliability of the information the application asks to the users and risk that the users quit using the application, due to the lack of immediate compensation for the nuisance of being always monitored and for the daily effort of actively using the application.

Keywords: personal mobility tracking; smartphones; living lab; mobility transition.

1 INTRODUCTION

In recent years we have observed an increased need for solutions to the problems of congestion and pollution in city centres. The car industry is devoting more and more resources to the development of electric and hydrid vehicles, to overcome emission limits that are ever more stringent. Electric cars are also an optimal choice for car-sharing companies when they need to deploy their fleets: in cities they can avoid congestion charges and their operating costs are competitive. At the same time, public transport is evolving, allowing for an even tighter inter-modality, with denser and more frequent runs, which, together with the availability of online transportation information systems, are making public transport much more appealing to the users.

An effective transition to new mobility styles should anyway be approached from both sides: it can be

”pushed” by public authorities, by car manufacturers and public transportation companies, by expanding the range of their products and services, but at the same time it must be ”pulled” by the citizens, willing to explore new transportation alternatives and to change their mobility habits.

In this framework, at the University of Applied Sciences of Southern Switzerland (SUPSI) we are developing the e-mobiliTI project (Cellina et al. [2013]), aimed at elaborating policy guidelines, incentives, procedures and strategies for a transition in the mobility sector at the urban level. In order to achieve its aims, e-mobiliTI focuses on the ”pull” side, thanks to direct interaction with the end-users. The project in fact investigates if and how the mobility behaviour of end-users can be changed by offering them alternative mobility solutions, such as electric cars and bicycles and subscriptions to public transport systems and car sharing companies.

The empirical methodology of e-mobiliTI has been to set up a *living lab* experiment (Higgins and Klein [2011]), that is a direct interaction with real life end-users, in their actual living environments. We selected a group of around thirty participants (sixteen families) and monitored their mobility patterns for a period of two months. During this period we simply observed their behaviour, without making any attempt at influencing it. Smartphones were used to monitor their movements, thanks to a specifically developed *app*, able to gather data and to interact with the user to obtain additional information, such as the purpose of the trip.

We are now in the second phase of our experiment, during which we are still monitoring users behaviour, but under a completely new setting, as each user has been now endowed with alternative mobility options (e-car, e-bike, public transport tickets, and so on). During this three months period we will provide the users with feedback, especially about their CO_2 footprint, in order to stimulate them to choose more ecological travel options, and will organise discussions and focus groups with them, in order to understand their perceived barriers for change. The comparison between the mobility patterns monitored in the first and in the second phase will allow us to identify the potential for diffusion of the new mobility solutions in everyday life.

In this paper we report on the outcomes of the first phase of our living lab experiment, describing the outcomes of the statistical analysis of the data we collected and on the individual mobility profiles that emerged from that analysis. The paper is structured as follows: first we describe the state of the art in the research about persuasive technologies in the mobility sector. Then, in the Materials and Methods section, we describe and discuss the technological solution, based on smartphones, we have implemented to monitor the users mobility patterns and to provide them with timely feedback about their performances. In the Results sections we present the analysis of data we have collected during the first phase of our project. Finally, in the Conclusions we discuss the results and we also analyse the limitations and the shortcomings of the use of smartphones for tracking mobility patterns in the context of our project.

2 BACKGROUND AND STATE OF THE ART

The actor-based living lab framework offers an ideal test-bed to assess effectiveness of the persuasive technology approach by Fogg [2003] or the Thaler and Sunstein [2008] approach in enhancing sustainability transitions.

In the mobility sector, however, the idea of nudging behaviour change by providing an eco-feedback is not as developed as it is for the energy sector, where various projects (see for instance Darby [2010] and Hargreaves et al. [2010], Hargreaves et al. [2013]) demonstrated that increased awareness has a positive impact on energy consumptions: receiving a feedback on their (quasi) real time behaviour, end-users understand how to save more energy. Recent studies are also investigating if and how feedback from social pressure and virtual communities such as Facebook or Twitter can play a relevant role in influencing user behaviour (Foster and Lawson [2013]): one’s social network might in fact offer a stimulating and challenging environment where to compare one’s own energy consumption patterns with friends’ and families’ ones. An example is provided by the US power utility OPower, which recently developed a social game to challenge users to save energy[1].

Automatic monitoring of mobility patterns, instead, is much more complex, since a static monitoring system is not enough: a flexible tracking system, able to follow the users along their movements, is necessary. Recently, however, the availability of affordable GPS devices and the large scale diffusion of smart mobile devices opened new research perspectives: in the last couple of years a few pilot projects aimed at automatic travel data tracking were developed (Jariyasunant et al. [2011], Nitsche et al. [2012], Ythier et al. [2012], Yuan et al. [2012]).

Also commercial applications for smart mobile devices are now available, mainly developed for health promotion and sport reasons (personal trainers): among the most downloaded applications, we mention Moves[2] and Endomondo[3]. Both *apps* track travels made by bike and foot (walking and running), with increasing precision and accuracy and with a very communicative and user friendly graphical interface.

Large corporations such as Google have also been actively developing various location based services that can assist the user in its personal mobility activities. An example is Location History in Google Maps[4] that records major displacements of a smartphone. The information gathered by Location History is put to work in the location based personal assistance provided by Google Now[5].

Finally, in the mobility sector only a few small-scale attempts were made to explore the combined effects of smartphone-based travel tracking and virtual communities in triggering behaviour change, such as the work of Jylhä et al. [2013] and the EU funded project SUNSET[6] (Bie et al. [2012]).

3 MATERIALS AND METHODS

In order to track the mobility patterns in the e-mobiliTI living lab, we needed a technological solution that was both inexpensive and non-intrusive, therefore smartphones were a logical choice. Not only modern smartphones have a range of powerful sensors and a GPS receiver, but they are also a technology users are already familiar with.

To reduce development costs and to simplify maintenance and assistance, we chose to develop an Android app, targeting a specific smartphone hardware (a Huawei system). This choice had no practical counter-indications, given the small size of our sample, especially because the users received a dedicated phone for the project needs. For a larger scale experiment the development of apps for different operating systems would be needed. It has to be noted that different phones have different hardware, and especially the GPS receivers might be performing differently, producing data of different quality.

A major design requirement was for the phone to operate without user intervention as much as possible, leaving to the users only the responsibility to bring the device always with them, with the battery charged, and to validate a daily report. To this purpose, the application is mainly divided in two parts: back-end processes and interface.

The back-end processes are in charge to detect and store user positions and send them to the server for elaboration. Taking advantage of the system background tasks provided by Android, we were able to create background services collecting data from GPS and to store them, without asking for user intervention. The operation of the services must be continuous and guaranteed for the whole duration of the day. The Android platform permits to create services and run them as "sticky processes": if the service stops working for any reason, the system will restart it without any external intervention. Data

[1] https://social.opower.com, last accessed March 2014
[2] http://www.moves-app.com, last accessed March 2014
[3] http://www.endomondo.com, last accessed March 2014
[4] https://maps.google.com/locationhistory last accessed April 2014
[5] http://www.google.com/landing/now/ last accessed April 2014
[6] http://sunset-project.eu, last accessed March 2014

collected by the phones are sent to the server via phones 3G or Wi-Fi connection. After transmission, data are erased from the phone for both performance and privacy issues.

Once on the server, data are processed according to an algorithm that identifies the *segments* which constitute a *trip*. A segment is a path along a series of GPS points between two stops. A stop is an interruption of movement defined according to some parameters. A trip is defined as a set of segments departing home and returning to it. The algorithm, briefly described in Cellina et al. [2013], creates segments by connecting the GPS points collected by the smartphone. The algorithm also tries to automatically identify the transportation mean.

The segments which have been identified by the algorithm running on the server are then transmitted back to the user interface for validation. Every day, at a fixed time (usually in the evening, but the users are allowed to choose the time of the day), the users are required to check their movements elaborated by the server. They can check all the daily segments on a map, decide if they are valid or not and, if necessary, merge them. They are also asked to provide, for each segment, additional information on the places they visited, the transportation means (correcting a first guess made by the algorithm based on average speed), the reason for the trip and the number of people travelling together.

When users have finished to check their movements, all their modifications and notes are sent back to the server, which stores them for further elaboration. Finally, on a daily basis, the server creates a second data storage, called "exposed database", containing all the users movements, completely anonymized, used for statistical purposes.

4 CHALLENGES IN DATA COLLECTION AND EVALUATION

During the first phase of the project we encountered a number of problems in data collection activities.

The biggest problem using a smartphone which is intensively exploiting the accelerometer and the GPS is battery life. In order to minimise the impact on battery life, a background service uses the gyroscope and tries to detect if the user is moving or not, consequently enabling and disabling GPS data request. Another trick to preserve battery life is to delegate all computations to the back-end server: the algorithm runs only on the server and CPU cycles are saved on the mobile phones.

Another challenge is the hook up of the GPS signal, which usually happens about 30 seconds after leaving a building. This issue caused problems in detecting the precise locations from which certain trips originated. GPS signal is also shielded in many trains, thus segments travelled by train were in some cases hard to detect and/or the computed travelled distance was affected by a large error.

We also encountered a number of problems related to the generation of segments from GPS points, and consequently to the identification of trips: the wrong recording of GPS signals caused problems such as segments with negative duration (0.6% of a total of 57'796 segments), segments with unrealistic speeds (e.g. cars travelling in urban areas at more than 200 km/h or regional trains at speeds exceeding 300 km/h, for a grand total of 0.2% of the segments). Additionally, some segments were excessively long (greater than 100 km), though being registered in a limited geographical area, and in some other cases they were too short, as trips were composed of too many segments (the algorithm artificially generates spurious stopping points, for example when stopping at traffic lights or queueing in the traffic).

Finally, some trips were not recognised, as the return to home was not correctly identified, and in various cases the wrong transportation means was assigned by the algorithm and not corrected by the user.

Cleaning the database from all these defective data meant removing approximately 10% of the data, which is a significant percentage. However, the most critical consequence of the presence of these errors is that they might have limited a generalised effective and enduring usage of the application by the users. Since they had no direct reward for their effort in using the app, the presence of errors in the recognition of their movements might in fact have discouraged them from using it, leading to a general lack of care in the validation of the daily report and in the specification of the additional information requested by the app (especially regarding the means of transport and the reason for the trip). Even though we dedicated an important effort in providing assistance to the users (phone calls, emails, "SOS" meetings to check the operating conditions of both the smartphones and the app, and so on), we cannot be sure of the quality of the data entered by the users. Yet, the user participation during the study has been relatively high and frequent. Out of 30 participants, 27 have been actively using the app throughout the 2.5 month period, filling in the daily reports, and taking the smartphone with them on every trip.

Furthermore, in some occasions, the users lamented they did not even receive their daily report, due to poor Internet connectivity in the places where they were staying at the time scheduled for the report: in those cases they could not validate their segments and provide us with the additional information requested by the app - or, better, we simply received the default information set up in the app.

We had therefore two major sources of potential data errors: the app and the users themselves. In order to reduce the uncertainty on data quality, we decided to interview each single user, discussing with them the mobility patterns we had identified and exploring with them the suspicious patterns in the data. Though only on a qualitative basis, the interviews with the users confirmed the general validity of the analyses we had made and, consequently, of the data we had collected. They also helped us in understanding the reasons for the users mobility choices, providing us with a fundamental insight to understand their mobility behaviour and their potential for change.

5 RESULTS OF DATA EVALUATION

At the end of the first phase of the project we started to process and analyse the data we collected over the period 1 May - 14 July 2013. Our objective was to identify the current mobility patterns of the participants to our "living lab", therefore we performed the following analyses: (i) a computation of statistical indicators for the whole user sample, (ii) a clustering analysis of our users, to identify similar behavioural patterns, and finally (iii) a personalised analysis of each user (kilometres travelled, impacts on energy consumption and CO_2 emissions, spatial representations, identification of the potentials for change). The latter were summarised in a report provided and discussed in "face-to-face" meetings to prepare the users for the second project phase.
Here we focus on the clustering analysis, which provides a general overview of the potentialities to profile mobility styles using the smartphone approach.

In our study there are 27 participants: 13 adults without kids (*adults*), 11 adult with kids (*parents*) and 3 *kids*. During the experimentation period, a user travelled on average 1811 ± 1148 km. An average adult or parent travelled some 40% more kilometres than an average kid.
For all categories, car turned out to be the most frequently chosen means of transport (Table 1). In percentage, parents are the heaviest users of car; moreover, kids use train, bus and foot more than adults and parents. Each category uses the bike for less than 1% of the kilometres.
As for the travel reason (Table 1), *free time* causes the highest number of travelled kilometres, followed by *work* and *shopping*. The causes of displacement are similar between adults and parents, but significantly different for kids. As expected, kids have almost no displacements due to work, while instead they have significant displacements due to school.

5.1 Clustering analysis of the local mobility during the working days

The data set we have collected allows for a number of different analyses. As an example, we report on the analysis of *local mobility* during the *working* days. This analysis will be useful to assess the impact of the measures we will be employing in the second phase of the study to change the personal mobility patterns of our users.

% km	adults	parents	kids
car	55	**68**	45
bike	**1**	1	0
motorbike	**18**	2	13
train	1	8	**24**
bus	3	1	**5**
foot	5	8	**11**
ecar	16	11	0

% km	adults	parents	kids
free time	36	**39**	**52**
work	**38**	26	4
shopping	11	9	1
family	2	**5**	0
school	1	2	**16**
personal	1	1	1
sport	1	**3**	0
other	10	14	**25**

Table 1. Different usage of means of transport and travel reasons for the user categories.

In this analysis we ignored a) trips falling out from a bounding box of radius 150 km centred in Lugano (the main city in the study area) and b) trips accomplished during holidays and week-ends.

As a preliminary step we grouped the transport modes into two big categories: ECO and non-ECO. The ECO label stands for ECOnomic and ECOlogic transport. This is therefore a rather strong ”political” statement, which we motivate here. We assumed that slow mobility (by foot or by bike) and public transport were the most ecological and economical way to move around a limited area such as a conurbation. On the other hand, we assumed that motorized transport, especially by car, is neither ecological nor economical. From this point of view, use of electric cars is controversial: an electric car in fact does not directly pollute or emit CO_2, however it is just a displacement of the impacts, as electrical energy needs to be generated. For the European Electric Mix it has been estimated[7] that an electric car is actually generating 80 CO_2 g/km, which might be even higher than some hybrid cars. Therefore, adopting a precautionary approach, we included the e-car in the non-ECO group.
We then separated the users in two clusters, via the k-means clustering algorithm [Wu et al., 2007]. The two dimensions of clustering are the mean number of kilometres travelled per day and the % of usage of non-ECO means (Figure 1).

We identified two groups of users. The first group travels about 22 km/day, covering about 90% of such kilometres by a non-ECO mean; the average age of such users is about 42 years. The users of the second group have quite different characteristics: a) they travel only *one third* of the kilometres of the other group (on average, about 7 km/day); b) in percentage they use *much less* the non-ECO means (55% of the kilometres); c) they are *much younger* (23 years on average). Thus a higher age is significantly correlated with higher numbers of travelled kilometres and a higher probability of choosing a non-ECO means. The same conclusion is achieved if we change slightly the definition of the clustering dimensions, for instance if we consider the displacements accomplished only by car or e-car rather than by a non-ECO mean. A possible explanation is that more adult persons have to take care of both work and family duties. For this reason, they need to travel more than younger persons. Moreover, they might be more subject to more schedule constraints, which are easier to meet by using the private means of transport.

6 Conclusions and future developments

In this paper we have reported about the first phase of the e-mobiliTI project, a ”living lab” experiment with which we attempt to understand if and how personal mobility styles can be changed by providing direct access to new mobility options, in particular electric vehicles and public transport. Here we report how we have identified the current mobility patterns of the users in the sample, and the various problems we have encountered in the process. The collected data will be used to evaluate the impact of alternative mobility options on the mobility patterns during the second phase of the study, currently in progress.

[7] For more details, see http://www.thelma-emobility.net.

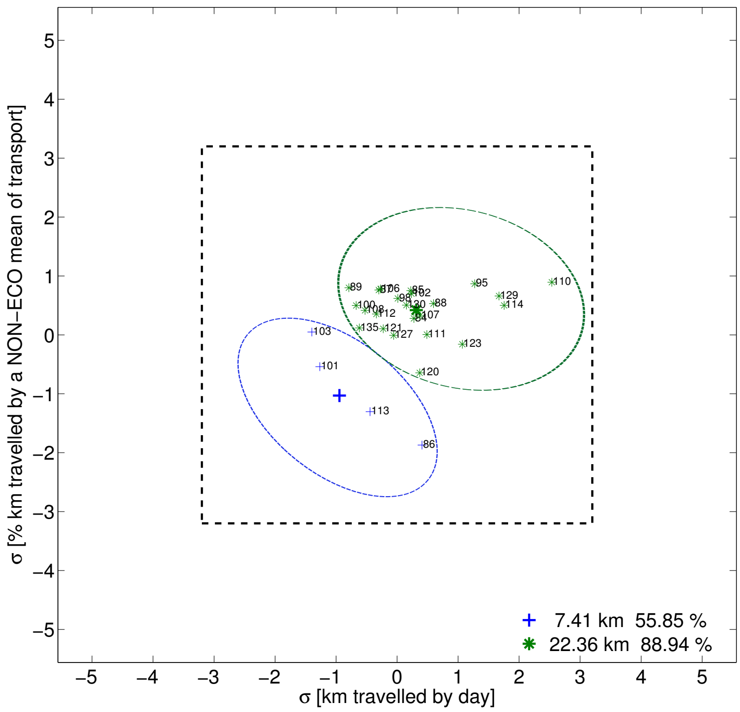

Figure 1. Clustering according to non-ECO means of transport. The X axis show the average number of km/day travelled by each user; the Y axis the % of km travelled by a non-ECO mean of transport. Since both X and Y variables are standardized, the unit of measure of both axis is the standard deviation from the mean of the sample. The right-bottom corner shows the coordinates of the centroids.

Within the project we developed a smartphone application capable of collecting huge quantities of data useful to profile the mobility styles of its users. Potentially, such an application might be used on a large scale, for instance in traffic planning processes, replacing the traditional mobility surveys based on interviews and devices for automatic counting of passages on the roads. Yet, we are working with an extremely reduced set of participants (27).This has the advantage of allowing us to analyse not only statistical data about their mobility behaviour, but also to have direct interactions with them during the examination of their profiles.This allows to understand subtler psychological motivations for travel choices, which might be otherwise buried in the data.

We are planning future studies which involve at least 800 participants, using their own smartphone, thus opening up new challenges in data processing and analysis. In that case in fact we will not be able to perform single-user interviews to validate the mobility styles identified through automatic analyses, not to mention the investigation of the qualitative motivations for change and the support during the use of the app. So far, we suppose we will replace the post-processing interview with each single user with a pre-processing written questionnaire, aimed at identifying their perceived general mobility patterns (typical trips during working/not working days, most frequently visited places, main reasons for the trips and so on), to be integrated in the automatic analyses.

The experience we have made in the first phase of the project has anyway allowed us to address the technical challenges and problems; in particular, with reference to the second monitoring phase, we have conducted several improvements of the smartphone application These improvements have been made to the user interface, to the on-board algorithm and to the segment recognition algorithm on the server.

In more detail, we have further simplified the graphical user interface to avoid confusion from the users.

Furthermore, we have implemented several additionalfunctionalities to automatically check for Internet connection and switch on the GPS (if switched off unintentionally by the user).We have then increased the GPS sampling rate of the on-the-phone data gathering algorithm; we have also re-designed the algorithm to allow for more efficient use of the battery. Furthermore, we have re-designed the server side segment recognition algorithm to deliver a more robust segment recognition. Finally, we have changed the daily surveys procedure to optimise usability.

The first experiences with internal tests with the improved version show very promising results. However, full evaluation and a detailed report on the improvements of the new system will only be possible after completion of the second phase with real users.

REFERENCES

Bie, J., Bijlsma, M., Broll, G., Cao, H., Hjalmarsson, A., Hodgson, F., Holleis, P., van Houten, Y., Jacobs, K., Koolwaaij, J., Kusumastuti, D., and Luther, M. (2012). Move better with tripzoom. *International Journal On Advances in Life Sciences*, 3–4:124–135.

Cellina, F., Förster, A., Rivola, D., Pampuri, L., Rudel, R., and Rizzoli, A. E. (2013). Using Smartphones to Profile Mobility Patterns in a Living Lab for the Transition to e-mobility. In *ISESS 2013*, pages 154–163, Neusiedl am See.

Darby, S. (2010). Smart metering: what potential for household engagement? *Building research and Information*, 38(5):442–457.

Fogg, B. (2003). *Persuasive Technology - Using Computers to Change What We Think and Do*. Morgan Kaufmann.

Foster, D. and Lawson, S. (2013). 'liking' persuasion: case studies in social media for behaviour change. In *CHI 2013 Changing Perspectives*. ACM.

Hargreaves, T., Nye, M., and Burgess, J. (2010). Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors. *Energy Policy*, 38.

Hargreaves, T., Nye, M., and Burgess, J. (2013). Special section: Transition pathways to a low carbon economy: Keeping energy visible? exploring how householders interact with feedback from smart energy monitors in the longer term. *Energy Policy*, 52.

Higgins, A. and Klein, S. (2011). Introduction to the living lab approach. In Tan, Y.-H. and et al., editors, *Accelerating Global Supply Chains with IT-Innovation*. Springer-Verlag, Berlin Heidelberg.

Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D., Kote, T., Sengupta, R., and Walker, J. L. (2011). The quantified traveler: Using personal travel data to promote sustainable transport behavior. Technical Report UCTC-FR-2011-10, University of California, Berkeley.

Jylhä, A., Sirén, M., Hemminki, S., and Jacucci, G. (2013). Augmenting a persuasive sustainable mobility application with social elements. In *CHI 2013 Extended Abstracts*, Paris, France.

Nitsche, P., Widhalm, P., Breuss, S., and Maurer, P. (2012). A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. In *Procedia-Social and Behavioral Sciences*, volume 48, pages 1033–1046.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., and Steinberg, D. (2007). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1–37.

Ythier, J., Walker, J., and Bierlaire, M. (2012). The role of the social network and the usage of communication in travel behavior measured with smartphone data. Technical Report TRANSP-OR 120801, Transport and Mobility Laboratory, ENAC, EPFL.

Yuan, Y., Raubal, M., and Liu, Y. (2012). Correlating mobile phone usage and travel behavior - a case study of harbin, china. *Computers, Environment and Urban Systems*, 36(2):118–130.