DISS. ETH NO. 23531

EXTRACTION OF TRANSPORTATION INFORMATION FROM COMBINED POSITION AND ACCELEROMETER TRACKS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

LARA MONTINI

MSc ETH

born on 16.08.1985

citizen of
Basel BS and Clos du Doubs JU

accepted on the recommendation of

Prof. Dr. Kay W. Axhausen, examiner
Prof. Dr. Hillel Bar-Gera, co-examiner
Dr. Nadine Rieser-Schüssler, co-examiner

2016

Contents

List of Figures

List of Tables

Abstract

Travel surveys are increasingly taking advantage of global positioning system (GPS) data offering precise and objective route and time observations whilst potentially reducing response burden. However, there are still several open issues concerning the automated post-processing of these large datasets. Without a reliable post-processing, GPS-based studies require either a considerable amount of manual analysis, leading to costly surveys or extensive prompted-recall interviews with the respondents.

As part of this thesis a travel diary study was conducted in the Greater Zurich Area. 150 participants carried dedicated GPS devices for up to one week and corrected their diaries in a web-based prompted recall tool. Using the resulting data set, the existing POSition DAta Processing framework was extended by a trip purpose module. Random forests, a machine learning technique, is used for classification. For trip purpose a share of correct predictions between 80 and 85 % is achieved for different setups. High variability in accuracy between persons is observed. Hence, personalisation strategies are tested. It is shown that the classifier is improved if it is learned on data that includes some of the participant's annotations (median accuracy + 5.5 %). The updated processing tool, and also lessons learned from the GPS survey in Zurich are tested in the PEACOX project, a joint project with many partners where a smartphone cross modal trip planner was developed that encourages ecological friendly behaviour. GPS and accelerometer time series for 33 study participants in Vienna and Dublin are available for analysis; these were tracked simultaneously with smartphones and dedicated devices for 8 weeks. Therefore, further insight into the usefulness of smartphones and dedicated GPS devices for collecting current travel survey data is gained. Meaningful diaries can be extracted from both data sources. However, if high resolution data is needed, results suggest that dedicated GPS devices are still relevant; they have no battery issues, meaning that more data is recorded and that data quality is more stable.

High resolution data is particularly interesting to observe taken routes.

Two potential applications are shown here: route choice models are estimated for all travel modes (public transport, car, bicycle and walking) and parking search is shown to be hard to identify in our data.

Zusammenfassung

Befragungen zu Verkehrstagebüchern werden immer häufiger durch GPS (global positioning system) Daten ergänzt. Diese ermöglichen eine präzise und objektive Beobachtung von Zeiten und Routen, optimalerweise bei tieferem Aufwand für die Befragten. Dies bedingt eine vollständige, qualitativ hochstehende und verlässliche Verarbeitung der üblicherweise grossen Datenmenge, ansonsten ist für GPS basierte Studien weiterhin erheblicher, kostenintensiver Aufwand nötig, entweder bedingt durch manuelle Analyse aller Tagebücher oder durch persönliche Interviews mit den Teilnehmern.

Als Teil dieser Arbeit wurde im Grossraum Zürich eine GPS Studie mit 150 Teilnehmern durchgeführt. Diese trugen für bis zu einer Woche ein GPS Gerät bei sich und korrigierten die automatisch erstellten Tagebücher auf der Umfragewebseite. Der resultierende Datensatz wurde genutzt um für die bestehenden Auswertungsroutinen (POSDAP) ein Aktivitätenerkennungs Modul zu entwickeln. Die Klassifizierung erfolgt mit dem Random Forest Algorithmus, wobei je nach Setup 80 bis 85 % der Aktivitäten richtig erkannt werden. Die Genauigkeit variiert stark zwischen den Teilnehmern, daher werden Personalisierungsstrategien getestet, es wird gezeigt, dass die Erkennungsrate um 5.5 % steigt, wenn beim Lernen des Random Forests Daten des Teilnehmers genutzt werden können. Diese erweiterten Routinen wurden im PEACOX Projekt eingesetzt, ein Gemeinschaftsprojekt in welchem ein multi-modaler, ökolgisches Verhalten fördender Routenplaner für Smartphones entwickelt wurde. Ausserdem konnten Lehren, welche aus der Zürcher Studie hervorgingen, umgesetzt werden. In Wien und Dublin sammelten 33 Teilnehmer 8 Wochen lang GPS sowie Accelerometer Daten sowohl mit ihrem Smartphone als auch mit einem GPS Gerät. Dies ermöglicht einen Vergleich der beiden Geräte und gibt Einblick in deren Brauchbarkeit für Tagebuch Studien. Beide Geräte liefern brauchbare Tagebücher, GPS Geräte sind weiterhin relevant falls hochaufgelöste Daten benötig werden, da sie keine Probleme mit der Batterie haben und die Datenqualität gleich bleibend ist.

Hochaufgelöste Daten sind insbesondere interessant, wenn Routen beobachtet werden müssen. Zwei potentielle Anwendungen werden vorgestellt: Routenwahlmodell für alle vorhandenen Verkehrsmittel wurden geschätzt (Auto, öffentlicher Verkehr, Velo, zu Fuss) und es wird gezeigt, dass Parksuchverkehr in unseren GPS Spuren schwer zu finden ist.

Acknowledgments

First, I would like to thank Prof. Kay Axhausen, for giving me the opportunity to work on this thesis, for his advice and for providing me and my fellow doctoral students with such a good and friendly working environment. This thesis is to a great extent based on Nadine Rieser-Schüssler's work, you were my supervisor in the beginning and helped me to settle in real quickly. I very much appreciate and thank you for your advice, your corrections and your friendship, you definitely helped me a lot to improve my papers and in the end this thesis. I also thank Prof. Hillel Bar-Gera for taking the time to read this thesis and for his valuable comments.

This research was mainly funded by the European Union as part of the project "Peacox - persuasive advisor for CO_2-reducing cross-modal trip planning" within the Seventh Framework Programme. I would like to thank all project partners for the smooth collaboration. Billy Brazil and Brian Caulfield from Trinity College Dublin. Johann Schrammel and Sebastian Prost from AIT who were my main contacts and also the main organisers of the PEACOX survey. I would also like thank the rest of the survey team, who I do not know personally, for their efforts. Mirjana Artukovic and Bilal Abunaim from Fluidtime, they were always a great help whenever technical issues occurred and he was also the main responsible to implement the logging components of the PEACOX app, without which I would not have had any data to analyse. Tomas Tvrzsky and his team from Telematix, especially Tomas Markovic implemented the prompted recall app for smartphones. Efthimios Bothos from ICCS made sure that all GPS and accelerometer data was saved and backed up.

This thesis was partly funded by the Swiss State Secretariat for Education and Research as part of the project "Route choice in urban public transport systems" within the COST Action "TU0603 - Buses with high levels of service". As part of this project the GPS survey of the greater Zurich area was conducted. I would like to thank our student assistants: Ilona Imoberdorf, Nathalie Schenk, Sebastian Sele, Lara Carisch, Fabian Binder and especially Patrice Frei and Tobias Karrer who assisted us during the entire survey. The first version of the prompted recall tool of this survey was implemented by Christoph Dobler, thank you for helping me developing it further and for all other numerous times, when you helped promptly with database, Java, Eclipse, etc. issues.

This thesis would not exist without the many good colleagues that contributed to the friendly working environment. Boris Jäggi lectured me that acknowledgments are not there to be thankful and tearful, but to acknowledge the work of those helping and contributing to the work done during the last 5 years. You were also one of those listening to my complaints (without taking them too seriously :)), thank you for that. Ilka Ehreke, ups I mean Dubernet, I enjoyed to once in a while take a break in your office, sometimes talking about models and sometimes not. Javier Ortigosa, I am very glad you started the PhD journey the very same day as me.

Andreas Horni long time office mate and now godfather of my little one, I thank you for trying to help me even when I did not accept any help, I highly value your opinion and am looking forward to occasional lunch breaks, birthday parties and Ners-times. Reto, father of my little one, also trying to help me until he accepted that armi muhmaus was all I wanted to hear. Kimi's grandparents especially -mothers, for taking her out for walks and helping me several times a week, so that I could finish the thesis. And finally, Kimiko, for sleeping 10 to 12 hours at night almost since she was born :)

Chapter 1

Introduction

In order to understand and analyse travel behaviour, information on trips and activities have to be collected. Traditionally, this is done either by requesting people to fill in a paper and pen diary or by interviewing participants e.g. by phone and asking them about recent travels. These surveys are now often complemented by GPS data collections, as they allow to observe movement outside of buildings on a very detailed level. In theory, the data enables us to derive objective, precise and complete travel diaries as it does not depend on participants' memory. Further advantages are precise trip start and end times, as well as capturing of short walking trips. The goal of this thesis is to further develop the processing routines, to examine accuracy of the generated travel diaries and gain insights on usability of the collected data for applications in transportation research.

The dissertation is organised in three parts that cover different aspects when working with GPS data in transportation. First, collection of GPS and accelerometer data and its challenges are presented in Part I. Having GPS points as such is usually not the main goal, but collecting travel diaries is. Automatic extraction of those is beneficial or even necessary due to the large amount of data. Evaluation and optimisation of those routines is presented in Part II, main focus being on the implementation of a trip purpose identification module and optimisation of the mode detection module.

In Part III applications where observation of detailed routes is of interest are introduced, namely parking search behaviour and route choice models for all modes of transport.

The basis on which the thesis builds upon is the open source POSition DAta Processing project (POSDAP, 2012; Schüssler, 2010; Rieser-Schüssler et al., 2011). The framework was developed using an already existing Swiss GPS data set only consisting of raw GPS points (that is x, y, z and timestamp) described in Chapter 5, and consists amongst others of modules to extract stages and stops, a module for mode detection, one for map-matching as well as choice set generation functionalities.

In Chapter 11 this original data set is analysed with regards to parking search behaviour in the Swiss cities Zurich and Geneva. The drawback of this data set is mainly the missing annotations for modes and trip purposes, therefore there was some uncertainty as to the performance of e.g. the mode detection module.

Consequently, one major goal of this dissertation was to collect annotated data with dedicated GPS devices. Chapter 3 presents the survey which consists of three parts: the GPS travel diary, a socio demographic questionnaire as well as psychometric scales capturing variety seeking, risk propensity and attitude towards the environment. This was the first survey, where the processing framework was used to preprocess the data before presenting participants their travel diaries on a web-based prompted recall tool.

The annotated data was then used to further develop the processing framework. Performance of the mode detection module is analysed and a trip purpose detection module is introduced in Part II. The data set is also used to estimate route and mode choice models in Chapter 10, considering all available data, that is socio-demographics, attitude scales and trip purposes.

Within the last years GPS surveys became even more interesting as smartphones enabled us to collect data without even providing devices to participants. Chapter 4 presents a smartphone-based survey where processing routines could be tested on data from Vienna and Dublin. As participants were also equipped with dedicated devices the resulting travel diaries using the two devices are compared.

Part I

Surveys and data

Part I is partially based on the following papers:

- Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013a) Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013

- Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler and K. W. Axhausen (2015) Comparison of Travel Diaries Generated from Smartphone Data and Dedicated GPS Devices, *Transportation Research Procedia*, **11**, 227–241

Chapter 2

Introduction and related work

In transportation research, GPS traces are used, along with other data sources, to construct travel diaries (Murakami and Wagner, 1999; Wolf et al., 2001b; Bricka et al., 2009) as they promise higher accuracy of duration and distance and increased numbers of short trips and activities. In addition, it is often argued, that GPS-based surveys allow for longer survey periods with less fatigue effects as response burden is reduced by automatically generated diaries. The responsibility for data annotation is still given to participants and is mostly handled using self-guided web-based prompted recall approaches (Auld et al., 2009; Bohte and Maat, 2009; Doherty et al., 2006; Frignani et al., 2010; Giaimo et al., 2010; Oliveira et al., 2011a). Location data is primarily collected using dedicated GPS devices that respondents must carry with them during the tracking period. Smartphones are a promising alternative source for location data (see e.g., Gould, 2013), as they have been equipped with good-quality GPS, accelerometer and other potentially useful sensor functionality during the last years and, as opposed to dedicated devices, are often carried by participants anyway. In this thesis, both a more traditional as well as a smartphone-based survey are presented in the following two chapters.

A recent review of GPS-based travel studies and the required processing tools is given in Shen and Stopher (2014), who list representative studies using dedicated devices from 14 different countries, as well as four smartphone studies. The first GPS studies were undertaken in the late 1990s (Wagner, 1997). GPS devices were at first attached to cars. Later on, handheld devices were used to capture all modes of travel (Wolf, 2004). Initial solutions for mobile phones were implemented in the mid-2000s (Asakura and Hato, 2004; Ohmori et al., 2006). By now, the smartphone-based travel data collection is growing rapidly, as evidenced by several new applications implemented over the last few years by the research community (Nitsche et al. (2014), Quantified Traveler (Jariyasunant et al., 2012), UbiActive (Fan et al., 2013), Future Mobility Survey (Cottrill et al., 2013), CONNECT of the MOVE project (Vlassenroot et al., 2015), SmarTrAC (Fan et al., 2015), SITTS (Safi et al., 2015)). Already, commercial tools designed to be used in different mobility studies are implemented, e.g., rMOVE (Resource Systems Group (RSG), 2015) and Studio Mobilita (2015) (used in Becker et al. (2015)). An application with very similar goals as the PEACOX app (Section 4.1) is the GoEco! app launched in March, 2016 (Cellina et al., 2015) a tracking app where usage of sustainable travel modes is encouraged.

A clear advantage of smartphones is the large number of potential participants who do not have to be provided with devices. Further, smartphones are less likely to be left at home than dedicated GPS devices and the possiblity to provide immediate feedback, e.g., on emissions, can increase the willingness to participate for longer time periods (Jariyasunant et al., 2012). But, using smartphones as a survey tool - in addition to normal usage - also implies an important challenge: battery life. Another difficulty for survey use of smartphones is a large variety of different operating systems, brands and types, with antennas of differing quality that must be covered.

Part I is structured as follows. First, a survey in the Greater Zurich Area where GPS data was collected with dedicated devices is presented in Chapter 3. In Chapter 4 the PEACOX project is presented, where data was colleced with smartphones and dedicated GPS devices, therefore, further insight is provided into the current usability of the two device types for collecting travel data. Data quality and usage are compared, as are travel diaries generated from the two data sources. Chapter 5 briefly introduces external data used in this thesis, that is another GPS data set as well as network and elevation data.

Chapter 3

GPS survey Greater Zurich Area

Travel behaviour is commonly modelled using socio-demographic as well as mobility specific attributes. But there are less easily surveyed latent variables that influence behaviour. Examples of such variables are risk propensity, attitude towards the environment as well as search for variety. The goal of this study was to evaluate the influence of these attitudes on route choice behaviour with a special focus on public transport. To better observe the route choice behaviour, a person-based GPS travel survey was combined with the attitude questionnaire.

Following, the field report of this GPS study is structured as follows. First, the survey design is discussed including target group and tools used. The second part describes who actually participated, followed by some descriptive statistics of the travel diaries. Lessons learned conclude this chapter.

3.1 Survey design

This survey aimed at collecting one week of GPS data of participants living in and around Zurich. The exact study area contains all municipalities within 22 km of Zurich Bellevue, just including the cities Winthertur in the north and Zug in the south. Addresses including telephone numbers were bought from an address dealer. As age distribution of such address databases are in our experience not representative for the population, in particular older people are over represented, addresses were bought by age category.

The survey was implemented as online questionnaires with two major parts:

1. Psychometric scales for the attitudes towards risk, environment and change
2. A one week GPS-based travel diary

These parts are described in more detail in Section 3.1.2 and Section 3.1.3, respectively. The third questionnaire concerning person and household characteristics covered basics such as age, gender, income, education level, was enriched by mobility tool ownership (e.g. cars, bikes, public transport season ticket) and concluded with questions about typical locations (home, work as well as two main shopping addresses).

3.1.1 Survey process

The survey was conducted between August 2011 and December 2012, and the design of the survey was changed in early 2012. Following, both process designs and the reasons for the changes are described in detail. The six main steps of the survey process and the differences of the two designs are outlined in Table 3.1. Technical details of the survey are addressed in Section 3.1.4.

Table 3.1: Survey process

	August 2011 - January 2012	January 2012 - December 2012
Introduction letter	✓	✓
Recruitment phone call	✓	✓
Equipment delivery	in person	by mail
Introduction to survey	in person	help page and brochure
Assistance during survey	phone / e-mail on request	in person after collection period
Returning equipment	by mail	fetched in person

The first two steps, constituting the recruitment of participants, remained the same in both designs. First, an introduction letter was sent out explaining the aim of the survey and announcing the recruitment call. Typically, participants were called 2 to 7 days after receiving the letter. All recipients were called up to 5 times over several days. After 5 unsuccessful calls the person was categorised as non-responsive. If the call was successful it was first checked that the person answering the phone is the recipient of the letter. Referencing the introduction letter, the goals and design of the survey and particularly the contribution expected from participants were explained in detail. No incentives were offered. If the person refused participation the reasons for doing so was recorded whenever possible. If the person agreed to participate the one week survey period was scheduled.

In the original design, a time and location chosen by the participant was scheduled to deliver the survey equipment. In the alternative design, the equipment was sent by mail. The survey equipment included the GPS logger, a charging device, a self-addressed postpaid envelope (original design), access information for the survey website and a brochure explaining the motivation of the survey and the handling of the website elements.

The pre-survey meetings usually lasted 30 to 90 minutes depending on the interest of the participant. During the meeting, the handling of the GPS logger and the different parts of the website were introduced. The focus was on the usage of the GPS-based prompted recall diary that was demonstrated using an artificial example created for every participant. The example contains common errors like missing signal during rail trips or tunnels where stages have to be merged, wrong mode identification that have to be corrected and randomly occurring wrong points that were not filtered by the post-processing. At the end of the interview, the participants were given a phone number and an email address where they could reach the survey team in case of any problems or difficulties. From the side of the survey team they were only contacted again if necessary.

Our experience showed that the artificial example, even though constructed from real data, only helped participants partly in understanding their own tracks. Further, many participants had only little time and did not want the assistants to introduce them to the web-survey in depth. This was one of the reasons why the in person meeting was postponed to the end of the survey in our second design. Without the preliminary instructions participants fully relied on the information in the brochure. Further help was available through the web-site which contained content-related but also extensive technical information on how the survey had to be filled out. Still, the main advantage of meeting after the survey period was, that assistants were able to provide instructions on actual data of participants or they could even help filling out the complete diary. These meetings lasted 15 to 60 minutes, depending on time reserved by the respondent and on the status of the questionnaires. Another advantage was that devices could be fetched directly, as before, they were not always sent back immediately. And third, concluding the survey with a meeting resulted in less postponing of the survey period.

To support these changes in survey design, the survey material was slightly adapted as well. On the one hand, a few simple diary pages were added to the brochure and participants were encouraged to take notes. This was a precaution as the transmission of GPS data did not always work as desired. On the other hand, the design of the web-based prompted recall interface was modified as explained in detail in Section 3.1.3.

3.1.2 Psychometric scales

The development and implementation of the psychometric scales has been extensively described in Rieser-Schüssler and Axhausen (2011) in the context of a prestudy, and is summarised in this section.

Corresponding to the three attitude domains that are investigated in this study, three separate scales have been developed: one measuring the risk propensity of the respondents, one addressing their attitude towards the environment and environmental protection and one quantifying the level of variety the persons seek in their life. Each scale is presented to the respondents with a 5-point agree-disagree scale. To minimise effects resulting from the order of the scale items, their order is determined randomly with three different random orders for each scale.

Risk propensity

There is a growing understanding in risk propensity research that a person's degree of risk taking does not only depend on individual, group and cultural factors but also on the domain in which the risk occurs. While it is still an open research issue whether this is caused by variations in the attitude towards risk over different domains or by varying perceptions of risks, Weber et al. (2002) argue that for the modelling and prediction of risk behaviour this distinction is irrelevant and that it is sufficient to observe the person's risk behaviour in the domain of interest. The risk propensity scale used in this study is shown in Table 3.2. It combines a reduced version of the domain specific risk propensity scale by Weber et al. (2002) with seven additional items for transport related risks. Overall, the scale contains 42 items covering the domains social, ethical, recreational, financial, health/safety and transport-related risks.

Environmentalism
Due to the increasing awareness of environmental issues, a lot of work regarding the measurement of environmentalism has been published in recent years. One of the earliest and most well-known studies is the land-use and transport behaviour study by Kitamura et al. (1997) who measured environmentalism using a 10 item scale. Subsequently, Schultz (2001) argued that environmental concern has to be differentiated between concern for oneself, other people and the biosphere because different values and awarenesses of harmful consequences are attached to them. Gatersleben et al. (2002) investigated the relationship between environmental attitudes and believes, socio-economics, social science indicators of pro-environmental behaviour and measurements of direct and indirect energy consumption. Following the theory of planned behaviour (Ajzen, 1991), Anable (2005) developed a 105 item scale to examine the influence of habits, moral norms, environmental attitudes, felt efficacy and perception of other persons' behaviour on mode choice and showed that the mode choice behaviour of different attitudinal population segments is indeed very different. After reviewing, amongst others, the scales of these authors, the scales used by Gatersleben et al. (2002) and Kitamura et al. (1997) were judged to be most appropriate for the study at hand. To use the advantages of both scales, they were combined into the 25 item scale presented in Table 3.3 that takes into account general concern for the environment, awareness of consequences for oneself, others and the biosphere and the evaluation of measures for environmental protection.

Variety seeking
Compared to the variety of studies employing measures for environmentalism relatively little research has so far been directed towards the quantification of variety seeking and its incorporation in models for daily transport behaviour. The few studies aiming in this direction investigate the phenomenon of travel for its own sake, i.e. undirected travel or travel with unnecessary detours (e.g. Mokhtarian and Salomon, 2001). A wider recognition of the influence of variety seeking on travel behaviour can be found in the tourism literature (Bello and Etzel, 1985; Niininen et al., 2004). Since none of the scales reported in the literature was completely satisfactory,

Table 3.2: Scale items measuring the attitude towards risk

Code	Question
R1	I admit if my taste differs from that of my friends *
R2	I argue with a friend if we have different opinions *
R3	I ask my boss for a raise when I think that I earned it *
R4	I would date a coworker *
R5	I would openly disagree with my boss in front of my coworkers *
R6	I speak my mind about unpopular issues at social occasions *
R7	I wear unconventional clothes *
R8	I would cheat a fair amount on my income tax *
R9	I still drive home after I had three drinks in the last two hours *
R10	I would forge somebody's signature *
R11	I have used cable TV without paying for it *
R12	I use office materials provided by my employer for private purposes *
R13	I would shoplift a small item (e.g. a lipstick or a pen) *
R14	I have at least once used illegally copied software *
R15	I go camping in the wild *
R16	I ski down slopes that are too difficult for me *
R17	I would like to do a safari in Kenya *
R18	I would go whitewater rafting at high water in spring *
R19	I would go on a 2 week vacation in a foreign country without booking ahead *
R20	I engage in dangerous sports, e.g. paragliding *
R21	I tried out bungee jumping at least once *
R22	I eat food that is beyond its expiration date if is still looks good *
R23	I ignore pain as long as possible before consulting a doctor *
R24	I rarely use sunscreen before sunbathing *
R25	I rarely wear a seat-belt *
R26	I would engage in unprotected sex outside a relationship *
R27	I usually ride my bike without wearing a helmet *
R28	I smoke at least one packet of cigarettes per day *
R29	I would co-sign a loan for a new car for a friend *
R30	I would invest 10% of my annual income in a blue chip stock *
R31	I would invest 10% of my annual income in speculative stocks *
R32	I would invest 10% of my annual income in government bonds *
R33	I would lend my best friend an amount of money equivalent to one month's income *
R34	I would bet a day's income in a casino *
R35	I would accept a job that is paid solely based on commission *
R36	I always take the latest possible public transport connection to the train station
R37	I start earlier if I assume that there will be congestion on my route
R38	I prefer public transport connections with very short transfer times
R39	If I don't know the way I just start into the general direction and search step by step
R40	I avoid streets that are occasionally congested
R41	I start earlier if I have to drive an unfamiliar route
R42	I try to be at the airport at the latest possible time

(*) Source: Weber et al. (2002)

Table 3.3: Scale items regarding environmentalism

Code	Question
E1	I worry about environmental problems *
E2	Too much attention is paid to environmental problems *
E3	Environmental problems are exaggerated *
E4	The attention for the greenhouse effect is exaggerated *
E5	I am optimistic regarding the state and future of our environment *
E6	Environmental pollution affects my health *
E7	Environmental problems have consequences for my life *
E8	I can see with my own eyes that the environment is deteriorating *
E9	Environmental problems are a risk for the future of our children *
E10	Saving threatened species is unnecessary luxury *
E11	We should be careful with our environment because we depend on it *
E12	Vehicle emissions increase the expenses for health care **
E13	Environmental protection starts with myself *
E14	People who do not care about environmental protection avoid their responsibilities *
E15	Behavioural change requires more environmental friendly products *
E16	Behavioural change requires a right example by the government *
E17	Pro-env. beh. is only useful if everybody cooperates and I don't think this will happen *
E18	Environmental protection costs too much **
E19	Environmental protection is good for the economy **
E20	Jobs are more important than the environment **
E21	Stricter vehicle smog control should be enforced **
E22	The price of gas should be raised to reduce pollution **
E23	Using tax dollars to pay for public transport is a good investment **
E24	There should be incentives for using electric vehicles **
E25	Who causes environmental damage should pay to repair it **

(*) Source: Gatersleben et al. (2002)

(**) Source: Kitamura et al. (1997)

we constructed our own scale including some of the questions reported by Mokhtarian and Salomon (2001). The variety seeking scale reported in Table 3.4 contains 28 questions measuring the desire for variety in the daily routine in general and in shopping, eating, recreational activities and transport behaviour in particular.

3.1.3 Web-based prompted recall interface

The main parts of the prompted recall interface are visualisation of the collected data, presentation of the travel diary, and editing of activities and stages.

In this section, first, the original prompted recall interface is presented, which is an all-in-one stage-centred approach. Based on our experience with this GUI we implemented some improvements, at the time we changed the survey design. The result was an activity-centred approach that consisted of two consecutive steps.

All-in-one survey - stage-centred

In the all-in-one user interface (Figure 3.1) participants first choose the day they want to review from a drop-down menu. All GPS points - or for performance reasons every n-th point - of the chosen day are presented on an interactive map; the timestamp for each point can be accessed by clicking on the respective point on the map. GPS points of the same stages are depicted in the same colour, stop points are shown in green. In addition, the diary information is presented in a table below the map. Each diary entry consists of a stage and the subsequent stop point. The attributes for each diary entry contain the start and end time of the stage, the chosen mode, travel costs, the characterisation of the stop point - activity purpose or mode change - and, finally, a personal location can be picked from a list to describe the stop point. Respondents can add such personal locations, defined by a description and its complete and geocoded adress, before or while filling out the diary. If the address is not known it can be derived using reverse geocoding.

Participants use the table to review and confirm or correct the diary entries and their attributes. They can delete and add diary entries, i.e. pairs of stage and stop points and change all the attributes provided.

Two-step survey - activity-centred

In the two-step survey, choosing the day as well as the handling of the map is the same as in the all-in-one survey. The diary is still presented as table,

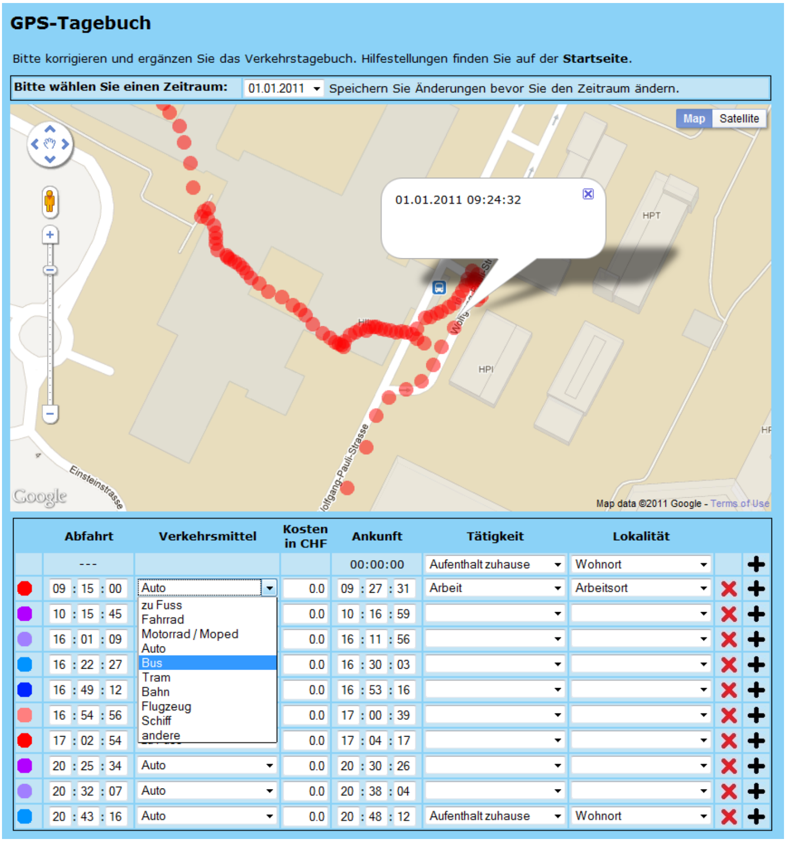
Table 3.4: Scale items evaluating the variety seeking tendency

Code	Question
V1	I like to experience novelty and change in my daily life *
V2	I sometimes look for ways to change my daily routine *
V3	I like to have lots of activity around me *
V4	I prefer a clearly structured, repetitive daily schedule
V5	Reoccurring rituals give me a feeling of control and security
V6	I love surprises
V7	A week in which all my evenings are similar bores me
V8	Shops with exotic herbs and fragrances fascinate me *
V9	When eating out I like to try unusual items *
V10	The content of my shopping cart looks pretty much the same all the time
V11	I buy only trendy clothes
V12	I prefer seasonal fruits and vegetables
V13	I actively search for bands whose music I do not yet know
V14	I always shop at the same supermarket
V15	I like to explore unknown towns or parts of my town
V16	I prefer to spend my holidays always at the same location
V17	I prefer having drinks always at my regular pub
V18	I like to try new types of sports
V19	Cultures completely different from my own fascinate me
V20	I prefer to organise my holidays spontaneously
V21	I always keep an open door for surprise visitors
V22	I like to meet new people
V23	I like to explore new places in my town or new towns **
V24	I like to try new routes to familiar destinations
V25	I sometimes take a longer route to see something new
V26	I like to drive around just for the fun of it
V27	When commuting I always take the same route
V28	I like to meet new people while travelling by train

(*) Source: Mehrabian and Russell (1973)

(**) Source: Weber et al. (2002)

Figure 3.1: Stage-based all-in-one GUI



but activity and stage variables are not corrected in the same step, therefore, more space is available to provide helpful input and interaction in the two separate steps.

In the first step, participants only correct activities (Figure 3.2). To help distinguish activities on the map, they are shown in different colours and stages in between are shown in grey. The following improvements were implemented: First, data is presented in a more intuitive way: starting with where the activity took place, followed by what was done from when to when and additionally the duration of the activity is displayed. The duration should help to understand the pre-processed diary e.g. if the duration is 3 minutes one might realise quicker that it is a mode transfer point, and it should also help to detect erroneous time specifications by participants. Second, adding locations was simplified by providing the possibility to create a locations for a specific activity, the location pop-up uses the GPS coordinates of the activity as a suggestion for the location coordinate.

In the second step, stages are corrected (Figure 3.3), therefore stages are depicted in colours and activities in grey. As times and locations are already specified in the first step the trip is presented as text containing the information where the trip started and where it ended, when it started and ended and again duration is included. Participants then only have to specify the travel mode, number of passengers and cost.

3.1.4 Technical details and data flow

To collect GPS data a dedicated device was used (MobiTest GSL (MGE DATA, 2012)). GPS points are sampled every second. Additional to the three-dimensional position and timestamp, vertical and horizontal accuracy and number of satellites in view were logged. Further, the three dimensional acceleration is measured at 10 Hz by the internal accelerometers. The devices were equipped with a SIM-card that enabled sending the data over the GSM network.

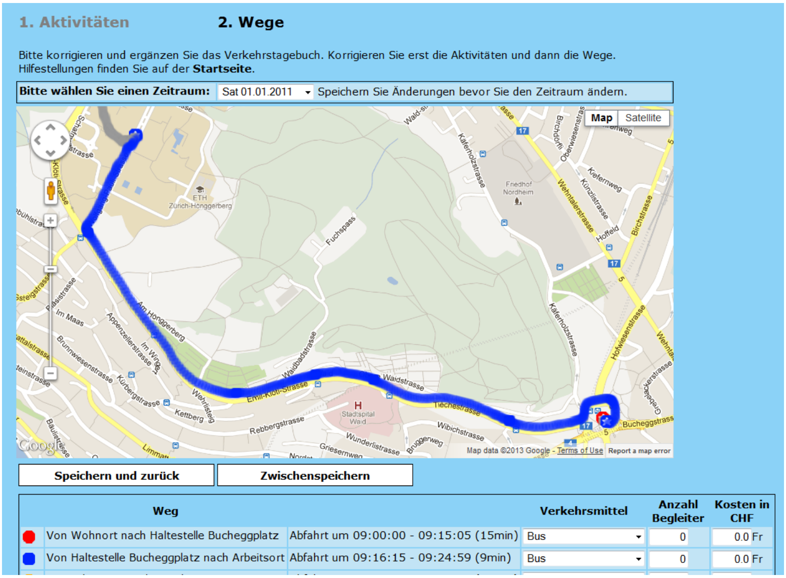
Figure 3.4 depicts the data flow detailed below. Participants were instructed to carry the GPS device for one week and charge it every night. When the device was charged transmission of the data to the ftp-server was triggered. Unfortunately, this did not always work as desired. If no data was sent, it was downloaded directly from the device and uploaded on the server. Every four hours, raw data was filtered and smoothed, which is the first step of typical GPS post processing routines, and then stored in a central MySQL database. The automated post-processing routines that were used are published open source (POSDAP, 2012) and are described in detail in Rieser-Schüssler et al. (2011) and Schüssler and Axhausen (2008). The three main steps executed are the filtering and smoothing, detection of stop points and stages followed by mode identification. Generation of the travel diary based on the GPS data was done once every night. As

Figure 3.2: Two-step GUI: activities

soon as post-processing was concluded the diary could be corrected by the participants on the survey homepage.

Apart from the survey homepage two administrative tools had to be developed: the telephone list to facilitate recruiting and a tool to observe the survey status of participants. Both were implemented as OpenOffice™Base user interface with direct connection to the central survey database. At first we worked with a non-central spreadsheet, which got very confusing even though only one person worked with it. It was therefore quickly replaced by the entry mask depicted in Figure 3.5. The telephone list had two main purposes, first, to provide all relevant information to make a call, and second, to keep track of how often, when and by whom potential participants had been called. The mask consists of the navigation-bar on top, information on when the introduction letter was sent out, recipient

Figure 3.3: Two-step GUI: stages



information (telephone obviously, age category as well as address, which is important for the assistant when scheduling the meeting) as well as recruitment information (number of calls, date of the last call, last callee, comments about the call or meeting schedule). The mask to track progress of participants simply consisted of some general information on the participant and if they logged in, and information on the completion level for each questionnaire.

Figure 3.4: Data flow: Integration of dedicated GPS device, ftp server, homepage and the central survey database.

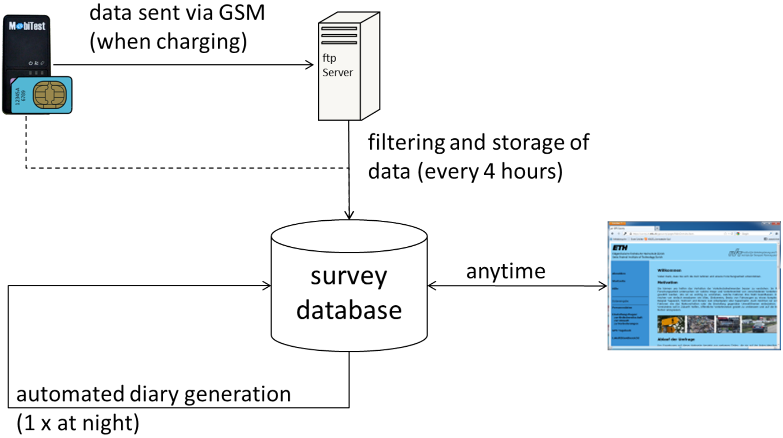
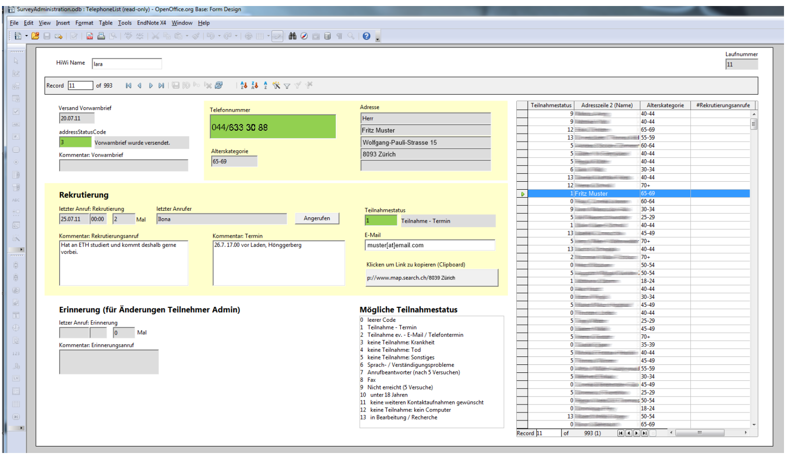


Figure 3.5: Administration tool: telephone list



3.2 Survey execution

Response burden as suggested in Axhausen and Weis (2010) sums up points assigned to each question depending on its type. For example, a yes/no question equates to 1 point, scales with 3 and more grades to 2 points and a half-open question with more than 8 possible answer to 4 points. For this survey, response burden was calculated to be approximately 1360 points split as follows on the three survey parts: 66 points for the socio-demographic questionnaire, 190 points for the three psychometric scales and the majority of the burden caused by the travel diary with around 1100 points. No incentives were offered and burden is comparatively high, therefore response rate were expected to be moderate.

In total, 1134 persons were contacted by telephone between 6 and 8 in the evening, of those 176 (16 %) agreed to participate, 133 persons (12 %) were not reached. Figure 3.6 reports, that young people are both less likely to be reached and less likely to participate. From the 176 persons agreeing to participate 156 (14 %) collected data for at least three days, and were therefore classified as valid. The timing when data was collected over the survey period is shown in Figure 3.7.

The peaks in October 2011 are due to the highest number of student assistants working for the survey at that point, and the peak in June 2012 is the period where one assistant worked almost full-time.

A comparison of respondents of this survey and the Microcensus (Swiss Federal Statistical Office (BFS), 2005) is given in Table 3.5. It was expected that GPS devices are more accepted by younger people, but this does not seem to be the case as people over 55 are well represented whereas younger people below 25 are highly underrepresented with 1.3 % in this study compared to 7.2 % in the Microcensus. The most interested group are the 45 - 54 year olds, who are well reached by phone and over 20 % accepted to participate. At first sight it looks like females were less willing to participate (42 %), but having a look at the addresses revealed that the share of addresses of females was also 42 %, therefore, willingness to participate is very similar. Our respondents were wealthier, better educated and lived in smaller households, than a representative sample of the Swiss population. However, this is a common finding in the institutes transport studies. Furthermore, the share of public transport ticket owners is higher than in the Microcensus, this could be due to the study area but also because of public transport goals of the study.

Figure 3.6: Response according to age category as given by address data file

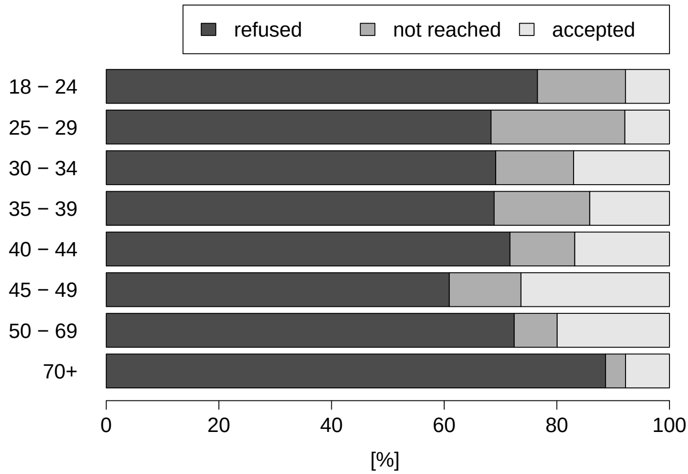


Figure 3.7: Number of participants starting the survey per month

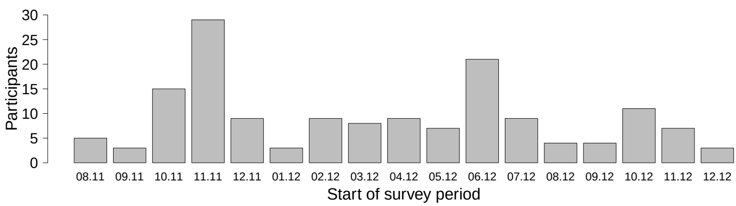


Table 3.5: Comparison of socio-economic attributes of the respondents with the Microcensus 2010

Attribute		Survey participants All [%]	Survey participants Valid [%]	MZ 2010 [%] (*)
Gender	Male	56.5	57.4	45.4
	Female	43.5	42.6	54.6
Age	18 - 24	1.7	1.3	7.2
	25 - 34	9.6	10.3	11.9
	35 - 44	11.9	13.5	17.3
	45 - 54	24.3	26.5	18.8
	55 - 64	10.2	11.0	17.6
	>= 65	14.7	16.1	27.2
	No answer	27.7	21.3	–
Education	Compulsory school	1.1	1.3	13.4
	Matur	4.0	3.9	4.7
	Apprentice	19.2	20.6	44.1
	Prof. diploma	24.3	27.1	9.1
	Univ. of appl. sc.	9.6	11.0	11.3
	University/ETH	19.8	21.3	11.6
	Other	3.4	3.2	5.2
	No answer	18.6	11.6	0.7
Employment status	In training	4.0	3.9	3.1
	Full time empl.	46.3	50.3	38.2
	Part time empl.	13.6	15.5	22.7
	Unemployed	1.1	1.3	1.9
	Houseworker	3.4	3.2	4.7
	Retired	12.4	13.5	27.0
	Other	0.6	0.6	2.4
	No answer	18.6	11.6	–
Household size	1	18.1	20.0	27.1
	2	31.1	32.9	36.4
	3	12.4	14.2	12.5
	4	11.3	12.3	13.7
	>= 5	4.5	5.2	5.8
	No answer	22.6	15.2	4.5
Monthly household income	< 4,000	4.0	4.5	17.3
	4,000 - 8,000	19.8	21.9	31.7
	8,000 - 12,000	20.9	23.2	15.5
	12,000 - 16,000	12.9	13.5	4.9
	> 16,000	7.9	8.4	3.1
	No answer	34.5	28.4	27.4
Car availability	Always	63.8	71.0	65.0
	Sometimes	10.7	12.3	11.9
	Never	5.6	5.2	4.6
	No answer	19.8	11.6	18.6
Public transport subscriptions (multiple selection possible)	Nationwide sub.	10.7	11.0	9.5
	Halbtax	51.4	56.1	41.5
	Other PT sub.	37.3	41.3	18.8
	None	33.3	27.7	42.0

(*) Microcensus 2010 based on 53025 people able to walk without help and age 18 or higher

3.3 Results

3.3.1 Psychometric scales

In order to examine the suitability of the psychometric scale results for subsequent choice modelling, a factor analysis was conducted for each of the three scales. The results of these factor analyses are presented in Tables 3.6, 3.7 and 3.8. To improve readability, only the factor scores with an absolute value of at least 0.4 are shown. The factor analysis was conducted with SPSS and the best results were achieved using a principal component analysis and a Varimax rotation with three factors for the risk propensity and variety seeking scales and four factors for the environmentalism scale.

Since the risk propensity scale covers such a variety of domains, the three main factors identified in the factor analysis explain only about 28% of the variance. They do, to a certain extent, follow the domains specified beforehand. The first factor mainly covers *health related risk* such as engaging in unprotected sex outside of a relationship, rarely wearing a seatbelt or habitual smoking. The second factor is a very mixed factor. On the one hand it entails items that are concerned with *recreation and transport related risks* – e.g. going on vacation without booking ahead or prefering risky public transport connections – on the other hand it contains the two items for taking a risk for a friend. The third factor summarises several of the *social risks* addressed by the scale items.

The four main factors for environmentalism shown in Table 3.7 explain about 52% of the variance in the data. The first factor describes the respondents' *agreement with measures to reduce car emissions*. The second factor characterises an *awareness of the negative consequences* of environmental pollution and our responsibility to restrict behaviour that is harmful to the environment. The third factor summarises more of an *overall concern about the environment* whereas the fourth factor shows a certain *expectations that others*, e.g. the government and companies, *step in* and provide better circumstances for environmental protection.

For variety seeking, about 37% of the variance is explained by the three main factors found in the factor analysis. The first factor describes the *interest in varying one's daily routine* through small changes such as trying out new routes to familiar destinations or trying out new food when eating out. The second factor is similar but puts a stronger emphasis on the *spontaneity, liking for surprises and for meeting new people*. The third factor captures a *desire for making new experiences* with other cultures but also at a smaller scale such as trying out new sports or music.

The factors identified in the three factor analyses all represent interesting approaches for an attitude based classification of the respondents. They are therefore suitable for the usage in latent variable and class models.

3.3.2 One-week travel diaries

For the analysis of the diaries, information reported by respondents is used. Data was extensively checked and manually corrected if necessary by the survey team. All information is available at the stage level. A stage is a building block of a trip that is covered by exactly one means of transport. For analysis, stages were consolidated into trips, that is consecutive stages which are connected through a mode transfer stop point are merged. Activities without annotated trip purposes are assumed to be mode transfers if shorter than 3 minutes, otherwise they were flagged as unknown. Trip purpose of trips leading home are assigned the purpose of the activity carried out longest after leaving home. Home to home trips without any activity in between are assigned to leisure.

The data set consists of 1039 person days. In total 7233 stages are observed that are part of 5284 trips. Figure 3.8 compares the number of trips per day and person for each trip purpose to the Microcensus (Swiss Federal Statistical Office (BFS), 2010). It can be seen that leisure trips are comparable. There are more work trips in the survey than in the Microcensus, this can be explained by the higher share of fulltime employees in our survey (50 % vs. 38 %, Table 3.5). Underrepresentation of education trips are probably due to the much higher share of young people in the Microcensus. A remaining issue is, that the purpose of many trips is not known (Other and Unknown), one can assume, that these are probably not work trips as such trips could be identified, even for participants providing very little information on their diary. It has to be kept in mind, that the sample size of our survey is small. From that perspective results are reasonable.

Comparing the mode shares shows even better results, both for number of stages (Figure 3.9) and for travel time (Figure 3.10). Travel time excludes waiting and mode transfer times. During manual processing it is easier to identify transport modes than trip purposes, therefore, the share of unknown

Table 3.6: Results of the factor analysis for risk propensity

Question (full version see Table 3.2)		Factor 1	Factor 2	Factor 3
R1	Admit taste differs	–	–	0.618
R2	Argue different opinions	–	0.474	–
R3	Ask for a raise	–	–	0.496
R5	Openly disagree with my boss	–	–	0.545
R6	Speak about unpopular issues	–	–	0.507
R9	Drive home after drinks	0.429	–	–
R13	Shoplift a small item	0.583	–	–
R15	Go camping in the wild	–	0.535	–
R16	Ski down too difficult slopes	0.493	–	–
R18	Whitewater rafting at high water	0.494	–	–
R19	Go on vacation without booking	–	0.532	–
R21	Tried out bungee jumping	0.412	–	–
R23	Ignore pain as long as possible	0.483	–	–
R24	Rarely use sunscreen	0.408	–	–
R25	Rarely wear a seat-belt	0.695	–	–
R26	Unprotected sex outside a relationship	0.714	–	–
R28	Smoking a lot	0.537	–	–
R29	Co-sign a loan for a friend	–	0.553	–
R31	Speculate 10% of annual income	0.525	–	–
R33	Lend best friend one month's income	–	0.543	–
R35	Job that is paid based on commission	0.533	–	–
R36	Latest connection to the station	–	0.589	–
R37	Start earlier if congestion expected	–	–	0.594
R38	Prefer short pt transfer times	–	0.592	–
R39	Search gradually on unfamiliar routes	0.466	–	–
R41	Start earlier if route unfamiliar	–	–	0.454
R42	Be at the airport as late as possible	–	0.535	–

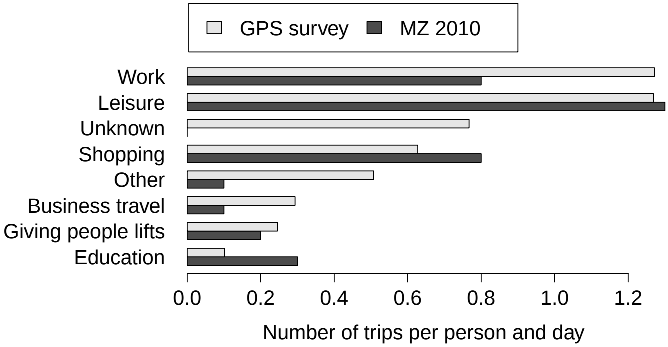
Table 3.7: Results of the factor analysis for environmentalism

Question (full version see Table 3.3)		Factor 1	2	3	4
E1	I worry about env. problems	–	–	0.591	–
E6	Pollution affects my health	–	0.683	–	–
E7	Env. probl. affect my life	–	0.502	–	–
E8	I see the environment is deteriorating	–	–	0.618	–
E9	Env. probl. are a risk for the future of our children	–	0.630	–	–
E11	We should care for our env. because we depend on it	–	0.708	–	–
E12	Vehicle emissions increase the need for health care	0.664	–	–	–
E13	A better environment starts with myself	–	0.775	–	–
E14	Not caring about the env. is avoiding responsibility	–	0.494	–	–
E15	Behav. change requires more env. friendly products	–	–	–	0.666
E16	Behav. change requires a right example by the government	–	–	–	0.719
E19	Env. protection is good for the economy	0.537	–	–	–
E21	Stricter veh. smog control should be enforced	0.583	–	–	0.444
E22	The price of gas should be raised to reduce pollution	0.613	–	–	–
E23	Using taxes to pay for pt is good	0.556	–	–	–
E24	Incentives for using electric vehicles	–	–	0.493	0.521
E25	Who causes env. damage should pay to repair it	–	0.608	–	–

Table 3.8: Results of the factor analysis for variety seeking

Question		Factor 1	Factor 2	Factor 3
V1	I like to experience novelty and change in my daily life	–	0.592	–
V2	I sometimes look for ways to change my daily routine	0.605	–	–
V3	I like to have lots of activity around me	–	–	0.435
V6	I love surprises	–	0.664	–
V7	A week in which all my evenings are similar bores me	–	0.507	–
V8	Shops with exotic herbs and fragrances fascinate me	–	–	0.591
V9	When eating out I like to try the most unusual items	0.450	–	–
V11	I buy only trendy clothes	–	–	0.500
V12	I prefer seasonal fruits and vegetables	–	0.601	–
V13	I actively search for new bands	–	–	0.643
V15	I like to explore unknown towns or parts of my town	–	–	0.624
V18	I like to try new types of sports	–	–	0.510
V19	Cultures completely different from my own fascinate me	0.440	–	0.452
V20	I prefer to organise my holidays spontaneously	0.402	–	0.470
V21	I always keep an open door for surprise visitors	–	0.576	–
V22	I like to meet new people	–	0.642	–
V23	I like to explore new places	–	0.487	0.535
V24	I like to try new routes to familiar destinations	0.586	–	–
V28	I like to meet new people while travelling by train	–	0.527	–

Figure 3.8: Comparison to the Microcensus 2010: Trip purpose, number of trips per day and person



modes is much smaller than was the case for trip purposes. Most modes have a slightly higher share in the GPS survey. Except for walk stages, where the share in number and the share of travel time is much smaller compared to the Microcensus.It has to be checked if many access stages were lost due to the cold start problem. Further, the share of tram and bus stages is slightly higher but interestingly their share of travel times is lower, this might be influenced by wrong travel time estimates and by the higher density of the public transport network in the survey area. Overall, mode share results are reasonable.

Figure 3.9: Comparison to the Microcensus 2010: Share of stages per mode

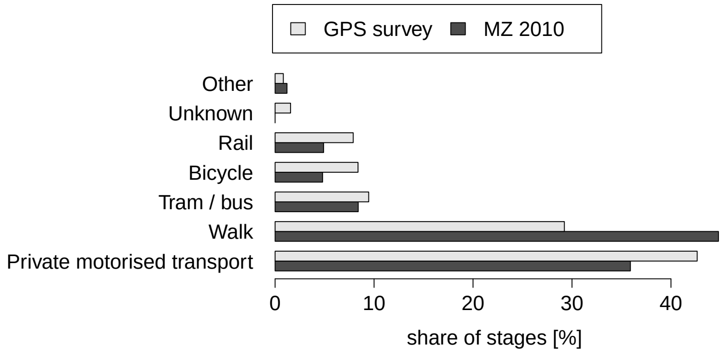
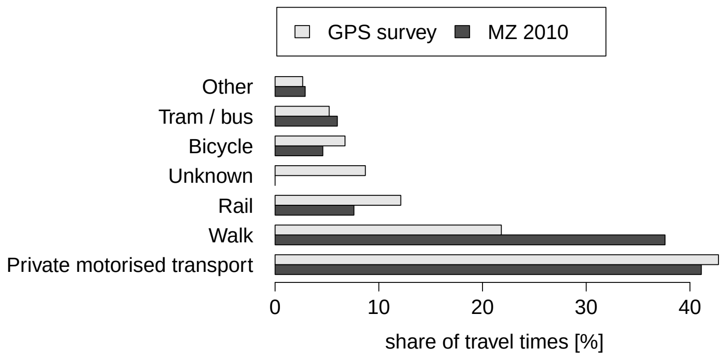


Figure 3.10: Comparison to the Microcensus 2010: Travel time share of modes



3.4 Discussion

Conducting such an extensive survey needs a lot of effort and attention. Realisation with part-time assistants and therefore spreading it over a long time period is not ideal. Our main finding, that might be obvious to experts, is that full-time workers are needed, and we consequently recommend to concentrate survey periods to predefined weeks. In such a setting, we assume it to be easier to maintain a hotline and actively attend participants, which sometimes requires trying to contact people every day. Small things like reminder calls are easily lost sight of, if not immediately successful and if other projects are worked on in parallel.

One of the goals when using GPS data is reduction of response burden, achieving this strongly depends on the display and handling of the prompted-recall diary. Our prompted recall diary has room for improvement both in design and in performance. For example, the reason for showing individual GPS points was to render it possible for respondents to correct start time and duration of activities. This was also the driving force of having all points of one day on the same map, as points before and after a detected activity or stage are needed for corrections. In retrospective, this is too much information for novice users. It made the map confusing in several situations, for example if activities outside consisted of many GPS points, or routes that are travelled several times a day covering each other. If the signal is bad, processing routines do not always produce reliable results, information in the diary does not correspond anymore to the diary in respondents' recollection; but unfortunately these are the situations where most input is needed from them.

To simplify correction of diaries several options can be considered. First, it has to be checked if letting people changing start and end times is necessary. If not, it might be better to show map-matched trips than tons of GPS points. Further, it might be reasonable to present only diaries that are evaluated to be reasonable (e.g. not too many trips), otherwise respondents are asked to reconstruct the diary from scratch, which might be less confusing than a badly prefilled diary. For this, but also in general, extension of our routines with an activity type detection module as presented in Part II is crucial. Stages could then be merged to trips and main activities might help people recognising their day quicker. Finally, the quality of the GPS tracks and the quality of the stage and stop point detection are key to the reduction of response burden. In the future probably more and more GPS data sets will be collected by smartphones, ensuring good quality will be a major challenge, especially if different smartphones and therefore different GPS sensors will be used (see Section 4.5.1).

It is perceived, that the quality was improved by changing the design (as summarised in Table 3.1), but admittedly, more people just handed in notes on their diary on paper and the student assistants had to complete the online GPS diary afterwards.

As overall the quality of the corrections was very diverse, all diaries were double checked by us. The majority of diaries needed some editions, which is definitely not cost-effective for large survey samples. Efforts and reminders during the survey period are therefore extremely important. However, with only a few well reported days and maybe some comments from participants a complete diary can be reconstructed manually with high certainty. With no input at all on the other hand corrections are rather uncertain. Especially with immobile days it is unclear if a device was just forgotten at home, or if the respondent actually stayed home. This aspect should definitely be incorporated in GPS-based surveys.

This data set is used to evaluate the processing routines in Part II as well as for the route choice models estimated in Chapter 10.

Chapter 4

PEACOX

The goal of the PEACOX project (www.project-peacox.eu) was to develop a personalised journey planner application for smartphones to encourage ecological travel behaviour. In the app, position data is collected to generate travel diaries; this is then used to personalise route suggestions. In this chapter, the main focus is on description of the second field trial of the app (Vienna and Dublin from August to October 2014) and analysis of the collected data. GPS and accelerometer time series of 33 study participants are available; these were tracked simultaneously with smartphones and dedicated devices for 8 weeks.

The Chapter is structured as follows. First, the smartphone applications used in the PEACOX project (journey planning app, as well as a prompted recall app) are presented. In the next section, a quick overview of the first field trial and its results is given. The main field trial is described in Section 4.3. Section 4.4 outlines differences in travel diary construction for the different device types. Next, results are reported, including quantitative analysis, as well as users‘ subjective perceptions. An interpretation of results and an outlook on continuing work concludes.

4.1 Study context: the PEACOX project and applications

PEACOX focuses primarily on the potential influence of the journey planning application, including its persuasive elements and how they affect users' travel behaviour and attitudes towards mobility. As part of this effort, GPS and accelerometer data was collected to inform users about past travel behaviour and CO_2 emissions. The application is a prototype and was tested in field trials, enabling us to enhance the data with questionnaires, a prompted recall tool and by giving participants dedicated GPS loggers (MobiTest GSL). In the following, the journey planning, as well as the prompted recall applications are introduced.

4.1.1 Journey planning app

The PEACOX journey planning app, developed by the PEACOX consortium, allows the user to perform a multi-modal search for a route tailored to the user's individual preferences and behavioural patterns. In general, it works like a common journey planner; an origin and a destination are specified and possible routes are then suggested. When routes are requested in PEACOX, available alternatives are enhanced with emission information (Alam and McNabola, 2012). The enriched results are then ranked and personalised by the recommender engine (Bothos et al., 2012). Recommendations are partially based on the trip history gathered from recorded GPS and accelerometer data, the trip history is the author's contribution to the application. When clicking on a route all details are displayed, that is walking-, driving- and waiting times, public transport line and schedule. Routes can then be viewed on the map, as they are multi-modal different transport modes are distinguished by colours (Figure 4.1(b)). Selected eco-friendly route options are promoted by adding a persuasive message (Figure 4.1(a)). Other persuasive elements were implemented: challenges where users competed against each other (Figure 4.1(c)), as well as comparing themselves on emissions rankings and, finally, a user's own improvement, represented by a growing or shrinking tree (Figure 4.1(d)). These and other potential persuasion strategies are discussed in Prost et al. (2013b).

The journey planner is implemented as a smartphone application for the Android platform version 4.0. and higher. Concerning data storage, all data is transmitted via mobile internet every few minutes to a central database. GPS and accelerometer data pose a particular challenge because of their huge numbers. Especially accelerometer data had to be backed up and deleted every day so that only two days of data are saved in the database at a time, otherwise querying the database (PostgreSQL) became slow. During the field trial 1.5 to 4 GB of accelerometer data were saved per day. More implementation details are given in Artukovic et al. (2013).

4.1.2 Prompted recall app

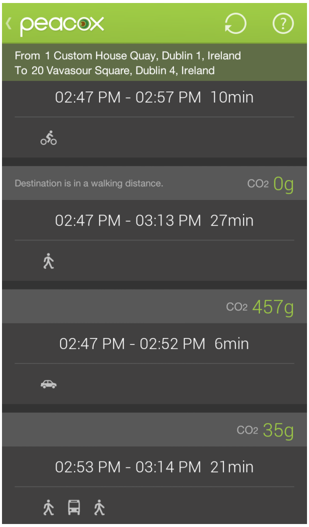
One goal of the PEACOX study is to observe behavioural changes during the 8 week field trial due to the journey planning app. The travel diaries are one data source to analyse such changes, it is therefore important that in particular the chosen transport modes are known. A prompted recall tool allows participants to review collected trip history and provide manual corrections. Often, these tools are web-based, as in our example described in Chapter 3. In this context, however, like the journey planner, the prompted recall tool is developed for smartphones.

For the field trial, a clearly laid out and user-friendly interface was developed, consisting of a map with GPS tracks and a prefilled list of transport modes and activity types representing the diaries (Figure 4.2(b)). There, to finish corrections, the checkbox 'I have reviewed this day' had to be checked. Each day could be selected from the menu (Figure 4.2(a)), the activity type and transport mode could be changed given predefined lists (Figure 4.2(c)) without any other restrictions from the system. The different transport modes were represented by different colours and all activity types had distinct icons. This supports processing the information displayed on the map, it particularly helps linking the map with the diary at the bottom. To explore the diary, on the right of the top bar, the number of activities was indicated, and one could quickly flip through them by clicking on the arrows. The map was automatically zoomed to the selected activity. Changing departure and arrival times was not allowed and instead of deleting activities or stages participants could select 'no trip' or 'no activity'. If no data was available, users could check the box 'I stayed home all day'. Unfortunately, it became clear only after the field trial that, in the list of days not yet corrected, the only days included were those where some data was available; most days without data were not confirmed by users. For every day, users could also leave a comment, e.g., if something was unclear or if trips or activities were missing (Figure 4.2(d)).

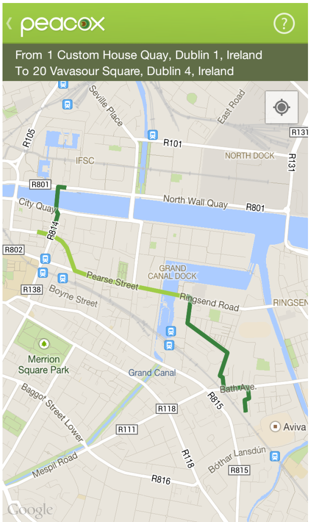
The app uses and shows very private data and is therefore login protected; login is the same as for the PEACOX journey planner. Overall, users stated that they were pleased with the handling of the trip diary app. They described it as easy to use and user-friendly. At least one user also found the app interesting for private use to check on the routes travelled during a day.

Figure 4.1: Screenshots of the journey planning app

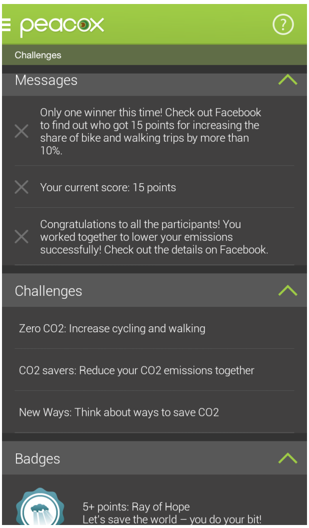
(a) Suggestions incl. CO_2



(b) Route visualisation



(c) Challenges and badges



(d) Persuasion: growing tree

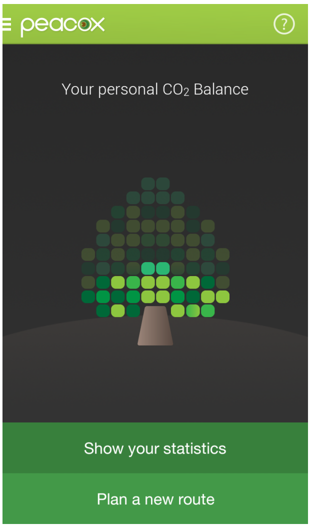
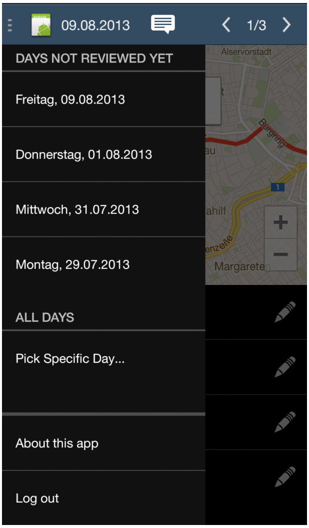
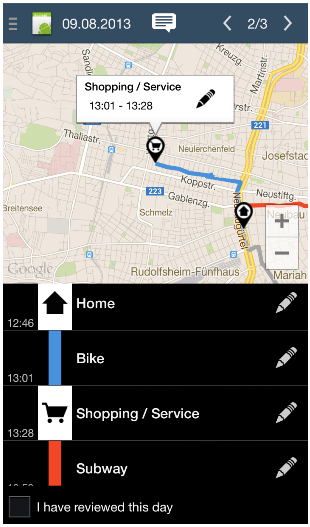


Figure 4.2: Screenshots of the prompted recall app

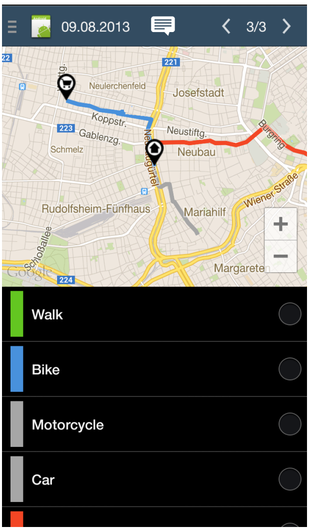
(a) Menu



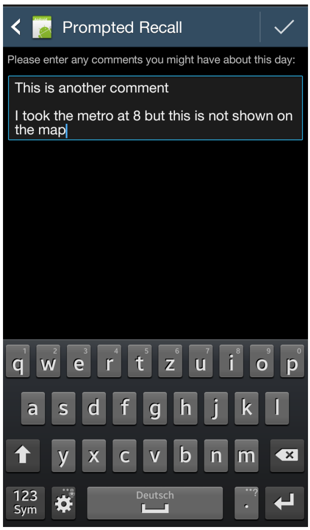
(b) Map and diary list



(c) Editing transport mode



(d) Comment screen



4.2 First field trial

In 2013, the first version of the trip planning application was tested in an initial field trial in Vienna (Prost et al., 2013a). Concerning the travel diaries, the goal was that the 25 participants collect 8 weeks worth of GPS data by smartphones and to get ground truth data using paper diaries for one week. Additionally 11 participants volunteered to simultaneously collect GPS and accelerometer data with dedicated devices.

4.2.1 Smartphone data collection

For the first 4 weeks of the field trial the logging of GPS data had a bug, that is the latitude and longitude of all GPS points were 0. Unfortunately, this issue was detected very late, consequently the data monitoring had to be greatly improved for the second field trial. For the last two weeks of the field trial, this issue was solved, but still very few valid GPS points were logged: that is 0 to 150 GPS points per day for smartphones, in comparison dedicated GPS devices logged 2400 to 19800 points per day. In both cases the intended logging frequency was 1 Hz. Considering that a 5 minute journey should produce 300 GPS points it is clear that the smartphone data could not be used to generate complete travel diaries. Logging of the accelerometer data on the other hand was much better. If data was logged, 2400 to 16400 points were logged per day (three phones did not have accelerometer at all), the dedicated devices logged approximately three times more that is 9400 to 29300 points per day.

4.2.2 Paper diaries

In order to collect actual travel data which is necessary for validation of the GPS imputation routines a paper diary was developed, as the later used smartphone application was not ready yet. The design goal was twofold, first, all travel data should be collected on a stage level, that is all transport modes used during a trip should be reported including start and end times. Second, participants should be able to fill it in shortly after the trips were made that is they should carry it around, therefore it should be kept as small as possible. The final design of the travel diary is compact enough to fit on an A6 page, and is depicted in Figure 4.3. On each page 3 trips can be reported, on the left the purpose of trip (Wegzweck) is asked, in the middle the times (Uhrzeiten) when a new transport means is used should be specified. And, on the right, the sequence of transport modes (Verkehrsmittel) should be indicated. Furthermore participants are asked if they used the route planner for a trip and there is also space to write down a comment.

Figure 4.3: Final design of the paper diary. Filled in example as given to participants.

Wegzweck	Uhrzeiten	Verkehrsmittel
☐ Zu Hause	Start: 18:30	1+4 Zu Fuss
☐ Arbeit/Ausbildung	Einstieg 1: 18:36	____ Fahrrad
☐ Einkaufen	Einstieg 2: 18:51	____ Auto
☒ Freizeit	Einstieg 3: __:__	3 Bus
☐ Dienstliches	Einstieg 4: __:__	2 U-Bahn
☐ Sonstiges	Ankunft: 19:05	____ Tram
		____ Sonstige
Routenplaner benutzt ☒ Ja ☐ Nein	Kommentar: Bar	
Wegzweck	**Uhrzeiten**	**Verkehrsmittel**
☐ Zu Hause	Start: 19:58	1 Zu Fuss
☐ Arbeit/Ausbildung	Einstieg 1: __:__	____ Fahrrad
☐ Einkaufen	Einstieg 2: __:__	____ Auto
☒ Freizeit	Einstieg 3: __:__	____ Bus
☐ Dienstliches	Einstieg 4: __:__	____ U-Bahn
☐ Sonstiges	Ankunft: 20:12	____ Tram
		____ Sonstige
Routenplaner benutzt ☐ Ja ☒ Nein	Kommentar: Kino	
Wegzweck	**Uhrzeiten**	**Verkehrsmittel**
☒ Zu Hause	Start: 22:55	1+3 Zu Fuss
☐ Arbeit/Ausbildung	Einstieg 1: 22:59	____ Fahrrad
☐ Einkaufen	Einstieg 2: __:__	____ Auto
☐ Freizeit	Einstieg 3: __:__	____ Bus
☐ Dienstliches	Einstieg 4: __:__	2 U-Bahn
☐ Sonstiges	Ankunft: 23:20	____ Tram
		____ Sonstige
Routenplaner benutzt ☐ Ja ☒ Nein	Kommentar:	

Datum: 21.05.2013

The paper diaries were filled in by 23 participants of which 21 reported all 7 days that were expected, 2 participants reported 5 days and 2 dropped out of the complete field trial. Overall they seemed to understand the concept well. The most challenging task was probably to insert the sequence of modes, which was done correctly by approximately 50 % of participants. The other half either made a cross on the modes they used or they indicated the number of times a mode was used during the trip. The missing sequence can usually be restored if GPS data is available, together with some knowledge of the network.

4.2.3 Automatically generated diaries

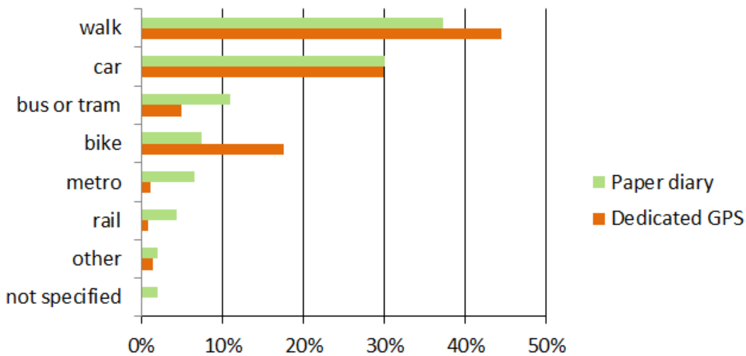
Diaries were automatically generated for the 11 participants that carried along a dedicated GPS device during Field Trial I. Unfortunately, due to the before mentioned problems with the smartphone data logging no diaries could be generated for the other participants.

Trip purposes were imputed using two random forest classifiers. The first one was learned on data from the GPS study conducted in Zurich (Chapter 3), using this classifier on a test data set from Zurich showed accuracies, that is the share of correctly classified purposes, of well over 80 %. But using this classifier on the paper diary data accuracy dropped to around 60 %. This is partly due to the small number of participants, as we showed that accuracies between participants vary a lot. On the other hand, there seem to be relevant differences between study areas. Therefore, a second random forest was learned on the paper diary data. When learning a random forest an error measure can be computed internally (out-of-bag error), which showed an accuracy over 90 % for this classifier. But as this was based on a very small data set, it was decided to combine the two classifiers. The scores for each trip purpose of both classifiers were added up and the trip purpose with the highest score was the classification result. Unfortunately, no test data set was available to get a final estimate of this combined classification. Running this classifier on the paper diary data showed an accuracy of 84 %. As this data is included in the learning process, this accuracy is probably a little higher than what is achieved on the 7 weeks of GPS extracted travel diaries, where no ground truth data was collected.

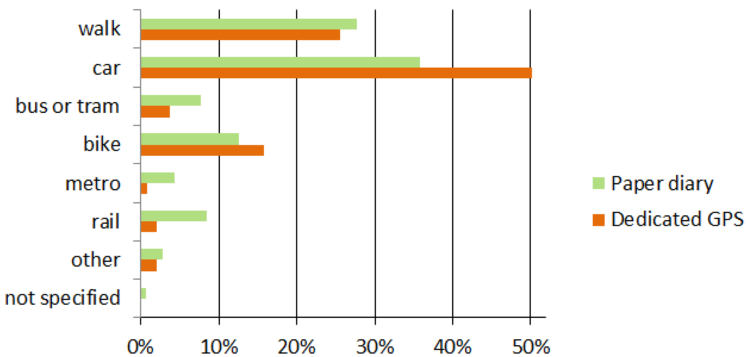
Transport modes were imputed using a fuzzy rule system, that was optimized ad hoc on data from Zurich. The most reliable data source for validation are the travel diaries. In Figure 4.4 the mode shares as reported in the paper diaries are compared to the mode shares of the automatically generated diaries from the dedicated GPS devices. It can be seen that the walk and car shares in terms of number of stages are very similar, the public transport shares on the other hand are underestimated by the automated travel diaries and the bike shares are overestimated. Looking at shares in terms of duration the bike share is similar, but car on the other hand is overrepresented and public transport is still underestimated.

Figure 4.4: Transport mode shares of the paper diaries compared to the automatically generated stages with the dedicated devices.

(a) Mode share according to number of stages



(b) Mode share according to duration



4.2.4 Analysis of travel diaries

The number of participants is too small to expect a representative sample. Anyhow, a comparison of the mode share as surveyed by Socialdata (Socialdata, 2009b) for Vienna in 2009 shows that results are reasonable: They observe for both car and public transport a share of 32 %, compared to 30 % and 21.5 % respectively observed with the paper diaries. Our share of bike is slightly higher (7.4 % compared to 6 %) and the share of walk is higher as well (37 % compared to 27 %) which might also be due to differences on the survey methodology.

Data collected by Socialdata (Socialdata, 2009a) regarding trip purposes show that the main share (33 %) are recreational trips, followed by work and educational trips (29 %) and shopping trips (26 %). A direct comparison of the numbers of Field Trial I data is not possible as home trips are given another purpose by Socialdata. Anyhow, trips as collected by the paper diaries are reasonable, as the sequence is the same. That is, recreational trips (29 %) are more common than work trips (16 %) as well as shopping trips (11 %). The remaining trips are home (27 %), other (9 %), business (5 %) and not specified (2 %). Even though direct comparison is not possible, the difference between the purposes is bigger in our sample, which suggests that work and shopping trips are underrepresented.

4.3 Second field trial

The second field trial was conducted in 2014, from August 11th to October 4th in Vienna, Austria and in Dublin, Ireland, where the trial started one week later.

A total of 37 test users (20 in Vienna, 17 in Dublin) participated in the field trial and tested the application on their own smartphones for eight consecutive weeks. The application accessed smartphone built-in sensors and logged their GPS as well as accelerometer data. Additionally, participants were equipped with a dedicated high-precision GPS positioning and logging device. As part of the trial, users were also asked to manually monitor their logged data (based on the smartphone GPS), using the prompted recall diary and to provide corrections to validate the automatically generated travel diaries. As an incentive, participants received 150 euros after completion of the survey.

4.3.1 Participants

Participants were recruited from a database of people interested in taking part in usability and user experience studies, by open calls for participation (promoted in university lectures) and through university mailing lists. Prospective participants had to fill in a screening questionnaire and could only be recruited when they fulfilled the following predefined criteria: age 18 or older, living and working, or studying, in the test area (Vienna, respectively Dublin metropolitan area), a smartphone (running Android OS 4.0 or newer) for at least three months, including an associated data plan with a minimum of 500 MB per month and, during the eight weeks of trial, no planned absence for more than one week (e.g. holiday outside of the study regions).

Overall, recruitment aimed at including a balanced representation of relevant mobility types (car users, cyclists, pedestrians, users of public transport), as well as demographic characteristics such as sex and education. This recruiting strategy resulted in the following sample (Table 4.1):

GPS data is available for 33 of the 37 participants, where for one person only smartphone and for another person only device data is usable.

Table 4.1: Characteristics of PEACOX participants (2nd field trial)

Age	Average age 33, oldest participant 69 and the youngest 19
Sex	14 female 23 male
Occupation	16 participants employed 12 students 4 unemployed or retired 3 self-employed 2 on parental leave
Main transportation means	6 users mainly use car or motorbike 6 use bicycles 11 public transport 5 are mostly walking 9 did not define
Usage of journey planning app	8 participants had never used a journey planning app prior to the study, 29 had

4.3.2 Procedure

After agreeing to take part in the trial, participants were invited to an introductory workshop instructing the users on the trial procedure, explaining the functionality and handling of the devices and apps and how participants were expected to use them. Participants were instructed to: carry the devices around at all times, turn on smartphone GPS sensor (to enable logging) and regularly charge the devices.

During the field trial, after about three and six weeks of usage, qualitative in-depth interviews about app usage - as well as experiences and their influence on transport mode decisions - were conducted with most users. Some participants were not reachable. At the end of the trial, participants were invited to focus groups to concentrate on collecting and reflecting on users' experiences during the trials.

Beside these face-to-face interactions with the participants, online questionnaires were also sent three times during the trial: at the beginning, in the middle and at the end. The questionnaires focused on demographic data, mobility behaviour and attitudes towards different transportation means and environmental issues. The second and third questionnaire also included questions on usage and experience with the apps.

4.3.3 Data collection

As described above two different approaches to data collection were used. The dedicated GPS was a MobiTest GSL device (MGE DATA, 2012); GPS data was collected with 1 Hz and accelerometer data with a frequency of 10 Hz. Data was stored locally on the device; after the end of the trials, when participants handed back the devices, the data was downloaded and made accessible for analysis. For smartphone data, as participants used their own devices, the sample consists of a variety of models, mainly Samsung devices, as shown in Table 4.2. Position data was collected in the background by the PEACOX app. GPS data was collected with a frequency of 1 Hz and uploaded to the server every minute. Accelerometer data was specified to use the sensor's standard frequency which is usually set to 5 Hz; data was uploaded every 70 seconds. Dedicated programming of the app ensured that the logging process was not stopped by the Android Task management, and that all available location information sources (GPS and WiFi network) were used for acquiring position information.

Table 4.2: Smartphone types used in Field Trial

Smartphone type	Nr devices
Samsung Galaxy S3	7
Samsung Galaxy S2	6
Motorola Moto G	3
Samsung Galaxy Nexus 2	2
Samsung Galaxy Nexus 4	2
Samsung Galaxy S3 mini	2
Sony Xperia Z1	2
Samsung Galaxy Note 2	1
Samsung Galaxy S4	1
Samsung Galaxy S4 mini	1
Alcatel One Touch 4030x	1
Huawei Ascend Y330	1
LG Nexus 5	1
LG P760 Optimus L9	1
UTime U100	1
Vodafone 875 Smart mini	1
Not reported	4

4.4 Travel diary generation

To process GPS and accelerometer data the software package POSDAP (2012) is used. The three most relevant steps when creating travel diaries are:

1. *Cleaning of raw data:* GPS points are filtered when too few satellites are accessible or accuracy measures are bad.
2. *Identification of activities and trips:* mainly based on point clouds, signal gaps and changes in the accelerometer signal if mode is changed to, or from, walk.
3. *Identification of transport mode and activity type:* done using either a fuzzy rule or a random forest classifier (see Part II).

Routine configuration was calibrated on data collected with the same dedicated GPS loggers used in this survey (MobiTest GSL). For classifier training, data collected in and around Zurich presented in the previous Chapter 3 was used.

In the following, differences in processing are described for the three travel diary types: (1) uncorrected diaries from smartphone data, (2) corrected diaries from smartphone data and (3) uncorrected diaries from dedicated device data. For all types in the subsequent analysis, stages were deleted if they were based on accelerometer only; that is, without any GPS point being part of that stage, as it turned out that most of those were unrealistic long.

4.4.1 Uncorrected diaries from smartphones

The uncorrected diaries evaluated in this chapter, created every night during the field trial, are the ones actually presented to the participants.

A random forest classifier for activity type identification is learned new every day, incorporating three data sources: (1) the data set collected in Zurich (around 7000 observations), (2) data collected in the first field trial (425 observations) and (3) all data collected and corrected during the second field trial. Using the freshly corrected data necessitates daily updating of the activity type classifier. As shown in Montini et al. (2014d), distance to home and work locations are important, but the PEACOX system does not know these locations, thus both locations must be learned as fast as possible. If corrected data is available, the locations most often annotated as home and as work are saved for that person. Otherwise, if GPS data was collected, but no corrections were available, a classifier not using distance to home and work was used to classify all activities. Locations predicted to be home and work are then used to extract an approximation of these two locations. Using these approximations, distance to home and work can be calculated and classification is run again, using a classifier that takes advantage of these distances.

After two thirds of the field trial (day 39 after start in Vienna), configuration of the processing routines was changed, because many stages were detected within point clouds. Hence, detection of point clusters was relaxed (radius for clouds increased from 10 to 35 meters) and the duration criteria were increased (minimum stage duration 3 minutes instead of 1 minute). Trip purpose detection stopped working due to an error when loading the freshly corrected data into a new classifier. Trip detection was rerun for the affected days (day 22 to 41).

At first, mode detection was implemented as fuzzy rule system. For the last third of the field trial period it was replaced by a random forest classifier. This classifier apart from considering the commonly used speed and accelerometer variables also included knowledge of self-reported mode shares.

4.4.2 Corrected diaries from smartphones

Corrected diaries are heavily based on the uncorrected ones, as users were not allowed to change start and end times, or add new activities. But users could add the flags 'no activity' and 'no trip'; thus stages are merged if 'no activity' occurred between; activities are merged if 'no trip' occurred between. Further, transport mode and trip purpose corrections by participants are considered. For this chapter, no further corrections were made by the researcher.

4.4.3 Uncorrected diaries from dedicated devices

Data collected by dedicated devices is processed all in one run, after the field trial. Processing used the first few weeks of the field trial's configuration. For trip purpose and mode detection, random forest classifiers were learned from the Zurich training data.

4.5 Results

Results of the field trials are organized as follows. In the first two subsections, data generated by and measures derived from both smartphones and dedicated devices are analysed. The following two questions about differences between the devices are tested:

1. Is the data quality of dedicated devices better and more stable than that of smartphones?
2. Are more days covered by smartphones because they are not often forgotten at home?

The first question is analysed looking at frequency of raw data. The second question is initially analysed by looking at daily levels of factors, but then more in depth, by investigating differences in the specific diaries.

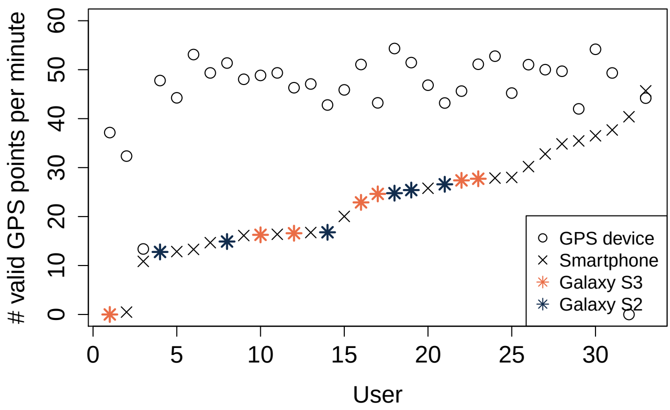
Third, users' usage and assessments of the trip diary app are presented. Trip purpose evaluation and transport mode detection based on users' changes are discussed. To conclude, user feedback on battery life is reported.

It is important to remember that all users' comments relate to data collected with their smartphones, as they had no access to data from dedicated devices during the field trial.

4.5.1 Data quality of raw data

To get a first proxy for data quality, GPS data sampling frequency is used. Both device types are specified to use a sampling frequency of one GPS point per second. The sampling frequency is computed for all detected stages (movement segments). In Figure 4.5, the average sampling is shown per user, ordered by smartphone sampling rate. For smartphones, the sampling is generally lower than that of the GPS device and between smartphones, bigger differences in sampling rates are observed than between GPS devices. This confirms our assumption and is not surprising, but the different frequencies must be considered when configuring GPS track segmentation routines. Types of smartphones used most in the study are also highlighted in Figure 4.5, indicating that sampling frequencies also differ within smartphone types. However, due to the small number of data points, it is unclear whether the differences are similar to the dedicated devices or greater.

Figure 4.5: Average GPS point frequency for detected stages (after cleaning GPS data) for both devices carried simultaneously



4.5.2 Comparison generated travel diaries from smartphone and dedicated device data

To compare usability of dedicated devices and smartphones for mobility studies, detected movement duration is chosen over the number of trips. The problem with number of trips as a comparison mode is that it relies on both trip segmentation and activity type detection, as stages are merged into trips if a mode transfer point is detected between them. Summing up the duration, on the other hand, should result in similar total durations, even if activities are not correctly classified or if the number of stages differs.

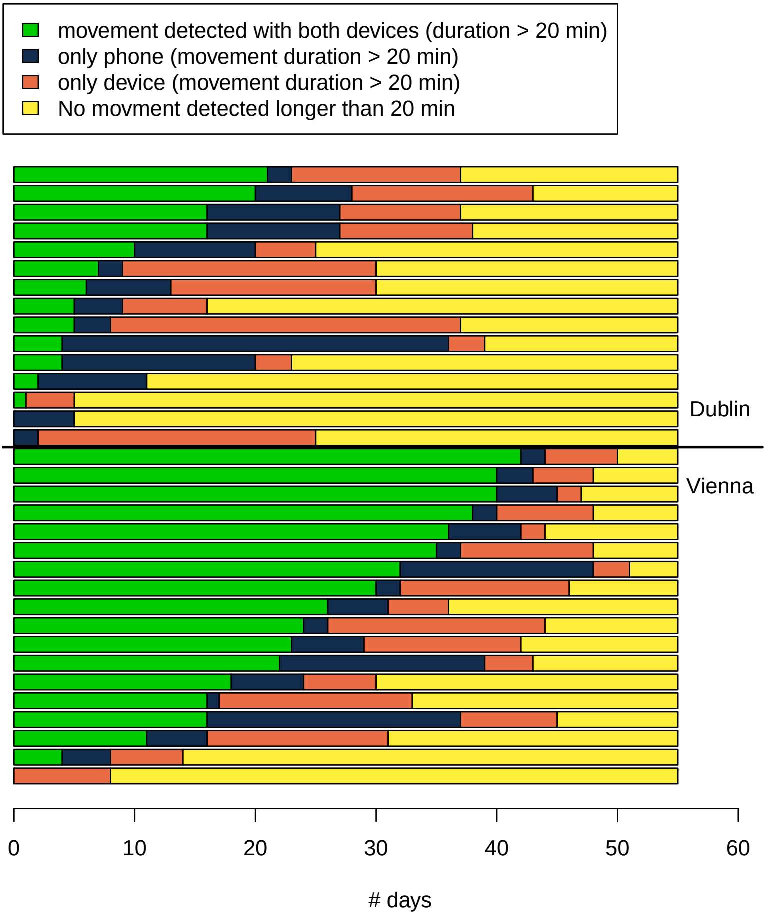
To get a general impression of the collected data amount, Figure 4.6 shows the number of days for which movement was detected with both devices, with one device or with no device, for each user. First, on many days, no movement was detected at all (yellow). Approximately half the Vienna users did not move, for up to 10 days, which is above expectations for an 8-week-field-trial, but not too much. Most likely, there was movement by participants that was not recorded by any of the devices. Interestingly, more movement is captured with dedicated devices; 312 days versus 227 days that were captured only by smartphones - this contradicts the assumption that smartphones cover more days. User feedback indicates that some days were not logged, even though the app was turned on; at least three users realised that full days of data were missing. In general, smartphones were probably not forgotten at home, but the app was turned off to save battery. Restarting the app when moving again can easily be forgotten. At least two users remembered that they left the GPS device in the car at some point.

Figure 4.6 also shows that the quantity of collected data differs for the two cities, possibly because: first, some of the Vienna users had already participated in the first field trial and knew what to expect and may also have had a special interest in the topic. Second, the main survey team was located in Vienna, possibly inducing more commitment in users living there. And third: users from Dublin were younger.

To get into more detail, Figure 4.7 shows summed up movement durations for all days on the left and an out-take of daily movement patterns on the right for three sample users. On the left, movement detected by smartphones, corrected movement (considering 'no trip', 'no activity' flags), as well as duration detected using dedicated devices data are shown. Further, colour indicates whether a user corrected the diary or not and also if s/he flagged 'stayed home'. It can be seen that no one flagged 'stayed home', probably because of an error in the trip diary due to which days with no GPS data were not listed under 'days to be reviewed'. In addition, almost no one made corrections during the last week and even though they were advised to make corrections the next day, most corrections were made after a week, with some only at the end of the field trial. On the right side of Figure 4.7, daily movement patterns extracted from dedicated devices are shown on top in violet and in blue, below, from smartphone data. Ideally, the two colours would match perfectly.

Details of the user with most days covered - with both devices - are depicted in Figure 4.7(a). On the left, it is clear, that the person collected data almost every day and also confirmed it using the trip diary. But there is also some evidence that not all corrections are correct; e.g. on day 32, the participant claims undetected movement of about 9 hours; probably, an activity was wrongfully marked as 'no activity'. The movement patterns on the right show a trip in the early morning around 07:00 on most weekdays and a trip home between 16:00 and 17:00. The patterns of smartphones and dedicated devices are similar; but, for example, the morning trip in this

Figure 4.6: Daily detected movement duration per person.



outtake is covered 12 times by the dedicated device and only seven times by the smartphone. In addition, there are still several shorter movements, perhaps erroneously detected, using smartphone data. This example is supported by above average data quality and quantity, as well as diary matching.

Figure 4.7(b) shows a user with relatively sparse data on all aspects. Of the 8 weeks, little more than 3 weeks are covered by data as shown on the left. It is also clear that the trip diary was only used for corrections at the beginning. The first three weeks depicted on the right show no work pattern, or any other pattern.

In several cases, too much movement was detected by smartphones, which was a reason for the configuration change mentioned in Section 4.4. A good example of this is the participants data shown in Figure 4.7(c), where the positive effect of changes is visible; movement detected for smartphone drops clearly after day 39 and is then similar to the dedicated device. The unrealistic movement durations are also clearly shown on the right, days 33 to 38, which were not the most extreme, according to the Figure on the left. But it must also be noted that, for some users, the original configuration was satisfactory (e.g. Figure 4.7(a)) and, for other users, the change in configuration had no clear effect.

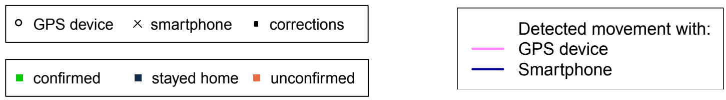
To compare detected daily movement, coverage criteria is introduced. Coverage c is defined here as the percentage of a stage detected with one device that is also covered by the other device. That is: when suffix $s1$ is a stage detected by device 1 and $d2$ is device 2 and $move$ is movement detected by the given device during the given stage and dur is the duration of a stage, coverage is given as:

$$c_{s1,d2} = \frac{move_{d2,s1}}{dur_{s1}} \tag{4.1}$$

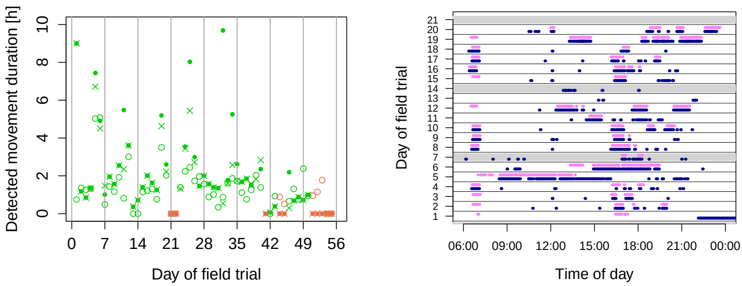
For example, if a stages of device 1 $s1$ is surrounded by a stage of device 2 $s2$ (which is twice as long), $c_{s1,d2}$ = 100% and $c_{s2,d1}$ = 50%. For every participant, Figure 4.8 shows - for both devices - the share of stages that overlaps with stages of the other device (cross symbols), as well as mean coverage of the overlapping stages (filled symbols). To compare the quality of the devices and not whether participants remembered using both (not always the case, as shown previously), analysis is done only for days where both devices registered movement as specified in Figure 4.6. The sample users presented in Figure 4.7 are also highlighted in Figure 4.8 and the order of participants is determined by the sum of all four shown criteria.

Stages detected from smartphone data tend to be longer; dedicated device stages are thus often completely surrounded, resulting in 100 % coverage for the GPS device and less than that for the smartphone stages. This shows in the higher values of mean coverage for the device stages. Smartphone stages are still mostly covered over 80 % on average. These are rather high and promising numbers; more problematic is, that many stages do not overlap at all. The range of values is rather high, varying between slightly less than 20 % and somewhat more than 80 %. Detecting too much movement is better than detecting too little, as it can probably be improved by personalised configurations.

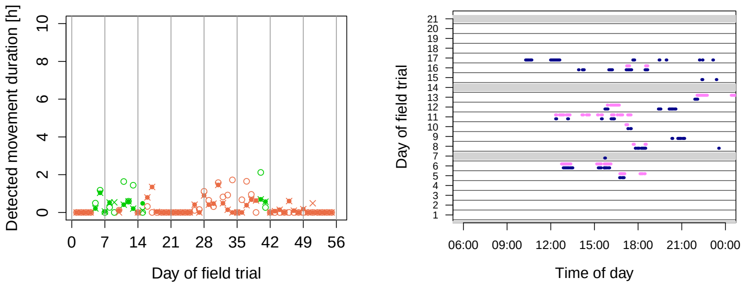
Figure 4.7: Detected movement for 3 selected users. Sundays in grey.



(a) User with highest share of days with data



(b) A user with relatively sparse data



(c) A user where config change (day 39) slightly improved movement detection

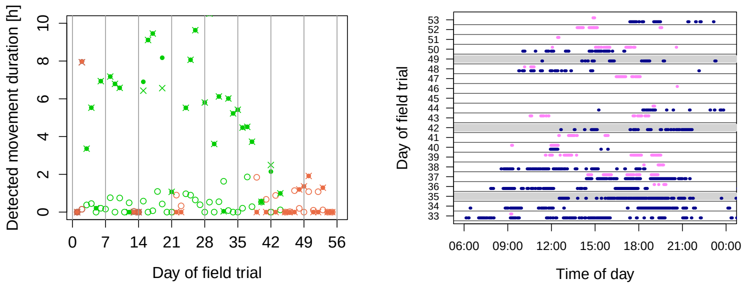
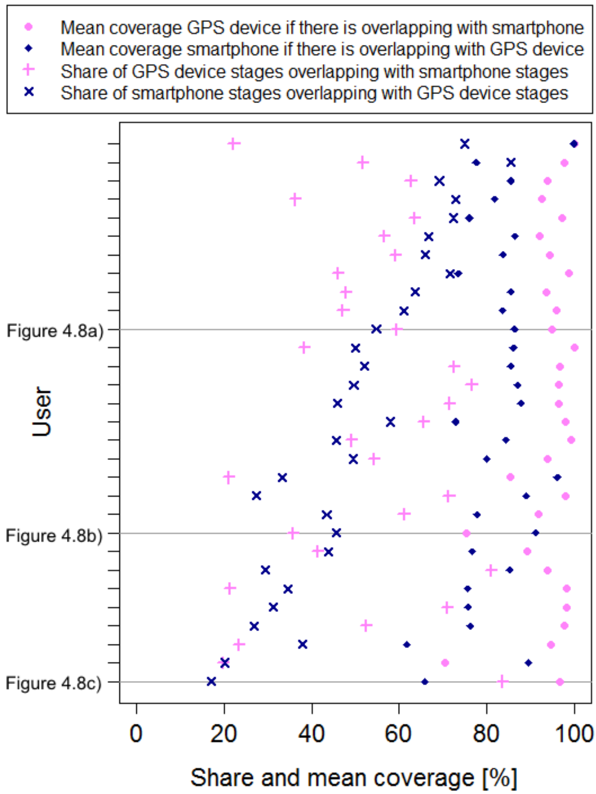


Figure 4.8: Mean coverage of overlapping stages and share of overlapping stages for each participant. Only days where both devices were active as defined in Figure 4.6 are considered here.



4.5.3 Evaluation of trip purpose and activity type detection

In total, 10415 stages were detected based on smartphone data presented to participants. Users made corrections in 41 % of all cases. There is, however, evidence that even more corrections would have been necessary. First, only 51 % of the days were confirmed by users. Second, 24 % of all stages were corrected to be no trips at all, and of those 67 % are stages detected as bike. Even when removing all 'no trip'-stages, the remaining bike share of 28.7 % seems too high, even though the study encourages ecological behaviour. And third, for these corrected diaries the share of stages marked with mode 'unknown' is still 8.2 %. The remaining reported mode shares are 39.7 % walking, 13.7 % car, 5.7 % bus or tram, 0.9 % rail and 2.8 % metro.

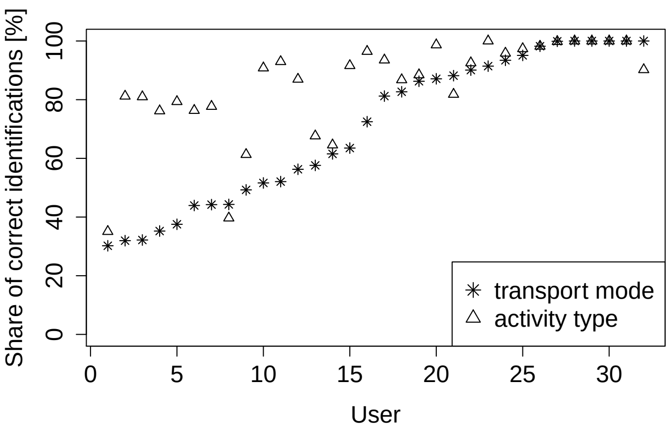
As explained in Section 4.4, trip segmentation was reconfigured during the trial due to problems with short stages detected within point clouds. Users noticed these problems and reported that this generally happened in situations where they were not moving at all; they complained that the many short segments were cumbersome to fix manually. At least one user reported that the system detected a lot of quick interchanges between different modes within a few minutes, including public transport. The configuration changes of both trip segmentation and mode detection had a positive effect. Slightly fewer stages were classified as 'no trip' (22 % compared to 25 % before). Overall, the share of correctly identified modes increased from a very low 55 % to 73 %.

Users' assessments of trip mode and purpose detection quality varied. Four users reported that they had no, or almost no, wrong trips, four others estimated a share of around 50 % to 75 % correct modes and activity types. Another four stated that the travel diary was inaccurate and they had much to fix. Users' assessments correspond well with corrections they made, as shown in Figure 4.9, where detection accuracies are shown per user, ordered by the share of correctly detected transport modes. Six users have 100 % accuracy, which indicates that no corrections were made. It is clear that activity type detection performed better than mode detection, which was either more influenced by the quality of the segmentation, or that participants tend to correct modes, but not activities.

Even though numerical activity-type detection performed better than mode detection, it was not perceived very well by users when asked about it: probably highly influenced by the three-week interruption where the classifier did not run. Overall, less than 20 % of all activities were corrected by users. Activity type detection accuracy was compared before and after configuration changes; no major differences could be found and therefore the following results cover the complete survey period. Of all corrections, 5 % are declared as 'no activity'. After removing those and merging activities with 'no trip' in between, the following shares are reported: 26.2 % at home, 27.7 % leisure activities, 15.6 % mode transfers, 12.0 % work or education, 10.7 % shopping, 1.2 % business activities, 0.9 % picking up someone and 5.7 % unspecified stops.

Travel diaries extracted from the dedicated devices have 7 % more activities compared to the corrected smartphone diaries, which is sensible as more days are covered by dedicated devices. The main difference is the share of activity type detected: over 4 times as many mode transfer points and approximately half as many home, work and leisure activities.

Figure 4.9: Detection accuracy per user (based on corrections of smartphone-based diary)



4.5.4 Trip diary app

Participants were advised to check the trip diary app every day as a fresh memory helps recall and correct recorded trips. As mentioned previously based on Figure 4.7 not all users corrected their data successfully or completely. Almost no one made corrections on the last week. In the first interview, participants were asked about how often they use the trip diary. Only three users stated that they used it every day, four used it once a week and seven users admitted not to have used it in the first two weeks of the trial. After the reminding interview this seems to have improved a bit. Figure 4.10 shows that 30 % of corrections were done within one day, and the majority within one week. But several entries were corrected more than 3 weeks after collecting the data.

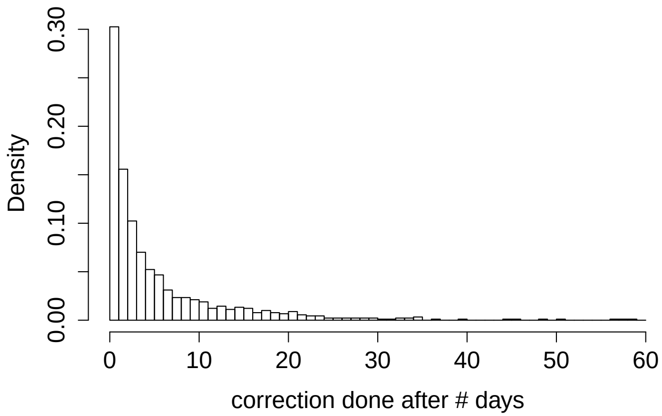
Overall, users were pleased with the handling of the trip diary app. They described it as easy to use and user friendly. A few users reported minor bugs and usability issues, some of them could be solved during the field trials. Problems with data quality of course also affected the user experience. Users were informed at the start of the study that detection accuracy rates of 60-80 % can be expected. However, at least one user expressed disappointment with the app, s/he would have expected more accurate results. In some instances, the predefined activity categories were not clear to users. There were a few grey areas. For example, it was not clear for some users, if they should categorise 'going for lunch' as 'leisure', 'shopping', or 'other'. Additional explanations or more activity types to choose from would have been beneficial for them.

4.5.5 Battery drain issues

A common and known issue with constant GPS logging is a considerable drain on the smartphone's battery. Accelerometer logging, on the other hand uses much less power (Ben Abdesslem et al., 2009) and is usually not an issue. To reduce impact on battery life, a scheduling mechanism was implemented that stopped any logging activity between 22:00 at night and 06:00 the next morning. However, several users reported that this did not work and that they had to turn off the logging manually at night.

In the introductory workshop, participants were advised to keep GPS antenna, the Google location services, WiFi and the PEACOX sensor logging on, whenever possible. Concerns about battery life were already being discussed, and leading to the consensus that turning off logging when

Figure 4.10: Number of days passing before corrections are done



not moving for some time, e.g., at home, is acceptable. Several users did so; however, at times logging had to be turned off to avoid an empty battery.

In the beginning, participants were also advised to carry a charger at all times and recharge whenever possible in the office (many users did that), or in the car (which was not always sufficient, depending on the length of trips). Two users even bought a second battery or a mobile charging device.

As expected, battery life was a problem (except for 4 users who specifically said otherwise), especially on the go, or outdoors without charging options. Several described it as quick and noticeable, one claimed that constant charging was needed; at least one user dropped out of the study because of these problems.

4.6 Conclusion and outlook

As expected, sampling frequencies of smartphones are lower and more diverse than for the same dedicated GPS device. Diversity between phones, as well as usage by participants, is also shown with the share of overlapping stages (Figure 4.8); this varies between 20 % and 80 %. In general, data quality collected with smartphones is sufficient, for example, to detect routes. Observation indicated that as much, or even more, movement is detected; missing data is not the problem. But if different smartphones are used, calibration of detection routines is a major challenge, particularly, when collecting information about short routes is important (often one of the reasons using GPS in travel surveys). In that case, from a researchers perspective, it is better to detect more and let users delete the wrongly detected trips and activities, which is much easier than adding trips. Following that reasoning, it might even be an option to detect 'no trip' and 'no activity'. This was actually done in this field trial; activity type classifier was learned on the actual field trial data, which includes these options.

The unexpectedly large differences in generated stages and activities made it almost impossible to compare the diaries on a fine-grained level considering detected transport modes and activity types. On an aggregated level, the type of activities detected for the dedicated device diaries are very different: four times as many mode transfer points and approximately half as many home, work and leisure activities.

For most users, more data is collected with dedicated devices, it is more likely that they will be taken along. On the other hand, users may carry their smartphones, but do not turn on the app. This is partly due to the heavy use of battery during high frequency data logging, which also renders such applications impractical beyond a dedicated study setup. For the sake of the study (and the financial compensation) users were willing to accept annoyances like carrying a charger with them and the occasional flat battery. However, for large-scale, long-term data collection, it is very unlikely users would be willing to compromise. Because of these issues - at the moment - if high resolution data is needed, dedicated devices are still relevant, as they last several days without charging, also thanks to a sleep mode of the used devices. It is expected, that battery life issues will be solved in the near future by better batteries, as well as optimisation of energy consumption and intelligent logging schemes.

Observation also indicated, that when users correct diaries, uncertainties remain about whether all incidents are corrected. On one hand, many entries are not confirmed at all; on the other hand, trip entries are confirmed as being corrected, where, obviously, too much movement was detected. For mode and activity type detection, several users made no changes at all, which is very suspicious. Thus, it is often obligatory that data is processed again, or cleaned manually after the end of the survey. The problem indicators mentioned above can be used as starting points to indicate where cleaning is necessary.

Despite issues with the quantity of corrections, implementation of this task as smartphone application instead of a (paper) diary was successful; users found the application easy to use. However, depending on data quality, this can impose additional workload to fix many false detections of trips or activities. While users expressed interest in reviewing their trips to learn about their own mobility habits, they have high expectations about quality of detection. So as not to discourage user activity in correcting detections, it must be carefully considered how much workload burden is put on users. In any case, an easy-to-use user interface is essential. Further, users should be actively triggered, or reminded, to regularly correct trips and changes.

The difference between cities in, for example, number of days with movement data, showed that even with passive collection methods of high resolution data, significant personal contact and effort by survey organisers are needed to ensure high-quality and comprehensive generated travel diaries.

Here, we described issues with GPS data processing when collected data was assumed to be the same in terms of specification, but different due to devices or people. Processing methods and data analysis, on the other hand, were identical. Given these preconditions, the extracted travel diaries are not the same. Future work should include configuration personalisation, as well as analysis of whether the method used is too sensitive about data specifications. Issues with mode and trip purpose detection, e.g. overrepresented cycling stages, should be investigated as well; does the problem stem from a previous erroneous detected stage, or does the mode detection itself need to be improved? It would also be most interesting to discover whether other processing methods yield the same results, or if they are more stable. Sharing issues about GPS data privacy and commercially-used processing routines, or the need for expert knowledge to apply the routines efficiently should be tackled.

Chapter 5

External Data Sources

In this chapter, a short overview of the data provided by external parties is given. First, a third GPS data set is introduced followed by a brief description of the network data as well as the elevation model.

5.1 External Swiss GPS data

An older GPS only data set collected in Switzerland is available, and is used to analyse parking search traffic in Chapter 11. The data was originally collected by a private company in order to analyse the potential success of places for advertisement. This is a nice example of GPS data being interesting for a wide range of analysis and purposes. This data set was also used to develop and test the original POSDAP framework and is described in more detail in Schüssler and Axhausen (2009c).

GPS tracks were collected from 2004 to 2006 using person-based trackers, meaning that all modes of transport are recorded. More than 30'000 days of data were recorded by 4892 volunteers living in the region of Zurich (including Winterthur) and Geneva. Only position data (x, y, z) and timestamps are available (no accuracy measures). As data was not collected as part of travel survey, also the diary information is missing, that is no annotations, neither start and end times nor transport mode or trip purpose are available for validation.

The socio demographics of the recruited respondents are given in Table 5.1 per main study area. For comparison, the values of the Microcensus 2005 are reported. Compared to this representative sample the data was collected by younger people; that is 25 - 34 year olds are over and people older than 65 are under represented, consequently also less people are retired compared to the Microcensus. As with many surveys the education level is slightly higher. Further, respondents seem to be more public transport oriented considering car availability and the high share of half fare card subscriptions. This is probably due to the urban study areas.

5.2 Network data

5.2.1 Open Street Map (OSM)

For map-matching of GPS tracks as well as for choice set generation network data is needed. The route choice analysis presented in Chapter 10 is restricted to the area around Zurich as depicted in Figure 5.1, the area was chosen such that most everyday travel of the survey participants was included. For walking, cycling and driving all network data is extracted from OpenStreetMap (2015). This includes geographical representation of links, speed limits and road type.

For the car network, pedestrian and cycling only links were excluded, almost all other links especially also tracks and residential roads are considered, resulting in a network consisting of 2.4 million links, that accurately describe the network geographically. Therefore, a curved road without intersections is described by several links.

Table 5.2 gives an overview of the road types, and details which highway tags were extracted from OSM. Speeds were specified according to the *maxspeed* tag of OSM whenever available, otherwise default speeds as given in Table 5.2 are used. Taking into account the oneway tag, a directed network in Matsim format (MATSim, 2015) is created. The conversion code can be found as part of the POSition DAta Processing project on sourceforge (POSDAP, 2012). Track and other roads cover 66 % (48000 km) of the total network length, different directions are counted separately, number of lanes are not considered.

For cycling and walking separate networks also in Matsim format are created, both consisting of approx. 3 million links (Table 5.3), not included are of course motorway and trunk links. Figure 5.2 shows an extract of the city of Zurich, where safe cycling paths are highlighted in green. It can be seen that outside the city many roads are considered to be safe cycling roads (as defined in Table 5.3), these are essentially smaller roads or dirt tracks within woods. Within the city, cyclists, cars and often public transport share the same road space.

Table 5.1: Socio-economic attributes of the respondents differentiated between regions compared to the Microcensus 2005

Attribute		Zurich and Winterthur [%]	Geneva [%]	MZ 2005 [%]
Number of participants		3527	1365	
Gender	Male	49.5	45.5	49.8
	Female	50.5	54.5	50.2
Age	< 25	15.3	18.8	20.3
	25 - 34	20.9	24.7	15.5
	35 - 44	25.2	23.3	18.3
	45 - 54	15.5	14.9	15.1
	55 - 64	11.8	11.1	13.5
	>= 65	11.4	7.1	17.2
Education	Compulsory school	11.5	10.4	12.9
	Matur	13.8	8.4	7.0
	Apprentice	48.9	38.7	49.1
	Prof. diploma	–	–	9.7
	Univ. of appl. sc.	14.0	11.6	7.0
	University/ETH	11.5	30.9	11.7
Employment status	In training	12.2	15.2	16.8
	Full time empl.	42.6	41.8	37.6
	Part time empl.	24.8	25.4	16.6
	Unemployed	1.9	5.3	2.9
	Houseworker	7.9	4.6	6.2
	Retired	10.5	7.6	18.8
	Other	0.1	–	1.2
Household size	1	15.7	22.3	32.9
	2	31.5	27.4	37.1
	3	18.7	18.5	12.1
	4	23.3	23.2	13.2
	>= 5	10.7	8.6	4.7
Monthly houshold income [CHF]	< 4500 || < 4000	13.6	14.6	20.6
	4500 - 9000 || 4000 - 8000	46.0	38.0	46.8
	9000 - 15000 || 8000 - 16000	20.0	18.2	28.5
	> 15000 || > 16000	3.7	4.5	4.1
	No answer	16.8	24.6	–
Car availability	Always	–	–	72.7
	Sometimes	–	–	20.8
	Never	–	–	6.5
	1	50.2	48.6	–
	more than one	35.7	37.9	–
	none	14.0	13.4	–
PT subscriptions	Nationwide sub.	–	–	8.6
	Half fare card (Halbtax)	50.2	36.2	26.3
	Other PT sub.	–	–	17.3
	None	–	–	38.0

Table 5.2: Properties of car network extracted from OSM, default speeds are only used if tag *maxspeed* not specified

Road type	OSM highway tag values	Average speed [km/h]	Default speed [km/h]	Nr. links	Length [km]
Motorway	motorway motorway_link	105	120	17305	846
Trunk	trunk trunk_link	87	100	3316	150
Primary	primary primary_link	58	50	60554	2137
Secondary	secondary secondary_link	57	50	68432	2299
Tertiary	tertiary tertiary_link	54	50	143530	4627
Residential	residential	31	30	505996	14597
Track	track	30	30	1117400	35795
Other	living_street unclassified road service	37	20 50 50 30	446775	12256
Total				2363307	72707

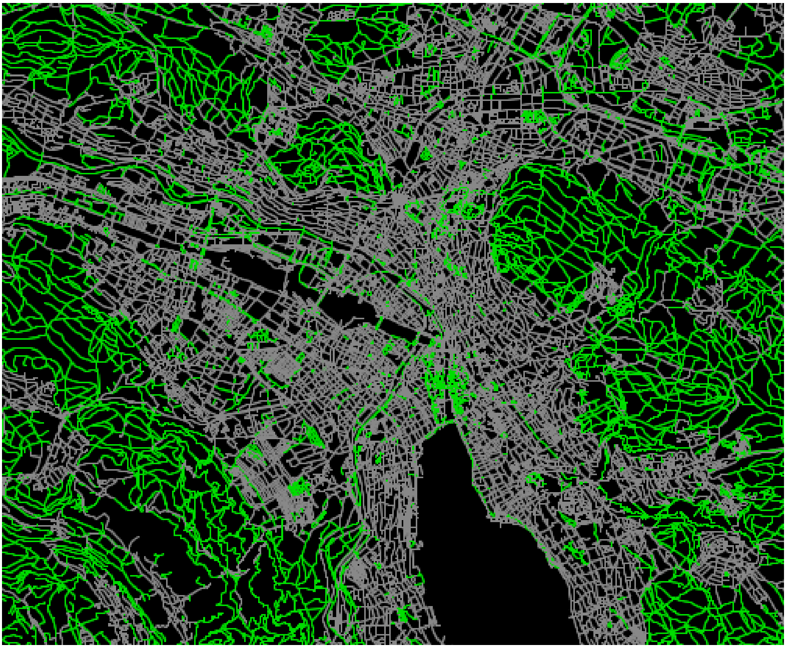
Table 5.3: Properties of bike and walk network extracted from OSM

Road type	OSM highway tag values	Nr. links	Length [km]
Bike lane	highway as in bike allowed with cycleway=(lane, opposite_lane, share_busway)	18496	606
Bike safe	footway with bicycle=yes, pedestrian, cycleway, path, track, highway as in bike allowed with cycleway=track	1462780	43833
Bike allowed	service, primary, secondary, tertiary, unclassified, residential, bridleway, living_street, road, path	1520214	41443
Total bike		3001490	85883
Pedestrian only	footway where bicycle = no, steps	274318	5440
Pedestrian / bike	footway where bicycle = yes, pedestrian, cycleway, path, track	1482128	44002
Sidewalk	see bike allowed	1271494	36788
Total pedestrian		3027940	86230

Figure 5.1: Area around Zurich included in analysis

Source: www.openstreetmap.org

Figure 5.2: Cycling network with safe cycling roads in green



Source: www.openstreetmap.org, Visualisation: Senozon Via

5.2.2 Public transport network and schedule

The public transport map matching routines were developed and tested on networks and schedules originating from the official transport model for the Canton Zurich (Rieser-Schüssler and Axhausen, 2013; Amt für Verkehr, Volkswirtschaftsdirektion Kanton Zürich, 2011). In this thesis, this network and schedule are used for all public transport map-matching tasks as well as for choice set generation (Chapter 8, Chapter 10).

The network and schedule covers all major train connections to and from Zurich as well as all public transport lines within the Canton, including ships and cable cars. The schedule represents the HAFAS-timetable of 2005. The network includes train and tram tracks as well as roads travelled by buses, and they are geographically correct, which is indispensable for map-matching.

5.3 Elevation model

Elevations for the canton of Zurich were released Open Source in 2015 under the GIS-ZH licence, the digital terrain model is available with a resolution of 0.5 meters (Office for Spatial Development of the Canton of Zurich, 2015). Outside the canton the digital elevation model with a resolution of 25 meters by swisstopo is used (Federal Office of Topography swisstopo, 2012). Each node of the network is assigned the elevation of the nearest measurement point.

The elevation measures for the bicycle and pedestrian route choice models, such as maximum and average rise as well as maximum and average fall are then calculated per route. For every link longer than 20 meters of a route the slope is calculated directly, if a link is too short it is joined with the next links until the sum of link lengths is greater than 20 meters, the slope is then calculated for the joined segment. The average rise is then calculated as the average of all positive segment slopes, the average fall is the absolute of the average of all negative segment slopes. Accordingly, the maximum fall is the absolute value of the most negative slope.

Part II

Automatically generated travel diaries

Part II is based on the following papers and documents:

- Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014d) Trip purpose identification from GPS tracks, *Transportation Research Record*, **2405**, 16–23

- Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2014c) Personalisation in multi-day GPS and accelerometer, paper presented at the *14th Swiss Transport Research Conference*, Ascona, May 2014

- Montini, L. and N. Rieser-Schüssler (2014a) Peacox – Implementation and evaluation of learning routines for mode and trip purpose detection, *Deliverable*, **D4.3**, IVT, ETH Zurich, Zurich

Chapter 6

Introduction and related work

In transportation research, GPS tracks often in combination with accelerometer data are used amongst other data sources to reconstruct diaries automatically, hence, complementing or replacing traditional travel diaries. These data allow to observe routes and times with high precision. Furthermore, respondent's recollection is supported by automatically generated travel diaries. If these generated diaries are accurate and complete, less input is needed from respondents and their burden is reduced. Consequently, GPS data is most interesting for longitudinal surveys to observe behavioural changes as well as routines.

Complete position data processing frameworks to construct travel diaries consist of the following parts:

- Preprocessing of raw data, e.g. cleaning and smoothing of GPS points
- Segmentation into stop points and stages
- Transport mode identification for stages
- Trip purpose identification for stop points

For further analysis, map-matching and choice set generation procedures are also key for transport research. All these mentioned modules are part of the Open Source project POSDAP (2012) that originated from Schüssler (2010) and is implemented in Java. The goal within this thesis is to add the so far missing trip purpose identification module, improve the mode detection and analyse personalisation strategies.

Two main groups for trip purpose imputation and mode detection routines can be found in the literature:

- rule-based systems, and
- machine learning approaches

For mode detection, rule systems are considering speed measures, stage criteria such as duration and also proximity to e.g. bus stops or roads (e.g. de Jong and Mensonides, 2003; Stopher et al., 2005; Chung and Shalaby, 2005; Bohte and Maat, 2008; Marchal et al., 2011). Further, fuzzy logic approaches were employed that also account for the fact that the boundaries between modes are overlapping (Tsui and Shalaby, 2006; Rieser-Schüssler et al., 2011). For trip purpose, rule systems rely mostly on the position of the activity, its timing, and GIS data on land uses (e.g. Moiseeva et al., 2010; Bohte and Maat, 2009; Stopher et al., 2008; Wolf et al., 2001a). One of the first rule-based trip purpose systems (Wolf et al., 2001a) considers land use data to narrow down possibilities to three purposes. In a second step, duration and time of arrival are used to refine classification. Overall, 67 % of trips were correctly identified. Most other rule-based approaches (Moiseeva et al., 2010; Bohte and Maat, 2009; Stopher et al., 2008) base the imputation solely on the location. Trip purpose is assigned by comparing the activity location to land use databases, points of interest and addresses provided by participants before or during the travel survey. In Wolf et al. (2004) trip purpose is imputed for clusters of activities. They included socio-demographics in the decision process, which was deterministic, but a probabilistic approach along the lines of a Bayesian network was proposed.

Several machine learning approaches were tested on mode detection e.g. Stenneth et al. (2011) use random forests with good success, that is overall more than 93 % of observations were correctly predicted and Zheng et al. (2008) as well as Moiseeva et al. (2010) use Bayesian inference models. For trip purpose detection the most common classifiers are decision trees (e.g. Gong et al., 2016; Oliveira et al., 2014; Lu and Zhang, 2014; Lu et al., 2012; Deng and Ji, 2010; Griffin and Huang, 2005). They are mostly based on variables computed for single activities, therefore the focus is more on the activity itself and less on position as was the case for rule-based systems. Deng and Ji (2010) report accuracies between 70 and 96 % on a rather homogeneous set of participants and Lu et al. (2012) achieve 60 to 73 % accuracies depending on the trip purpose. Liao et al. (2007) achieve good results (80 to 85 % accuracy) using hierarchical conditional random fields. But also other approaches are tested, Oliveira et al. (2014) compare a decision tree and a nested multinomial logit model. Liao et al. (2007) achieve good results using hierarchical conditional random fields. Lu and Zhang (2014) compare three algorithms on two datasets: decision tree, support vector machine and metalearner.

When comparing accuracies of the reported classifications, it has to be considered that the level of detail and therefore the number of classes differ and that the data sets used for classification are also very different in size as well as in the homogeneity of participants.

In this thesis, mainly random forests (Breiman, 2001) are used for trip purpose and transport mode classification. Random forests have been successfully applied to mode detection as mentioned before, and to other transport related classification problems (e.g. Ali et al., 2012; Greenhalgh and Mirmehdi, 2012; Rodrigues et al., 2012; Moreira et al., 2005).Similarly to the trip purpose detection problem, Wu et al. (2011) use random forests to classify activities based on GPS data. They distinguish between indoor, outdoor static, outdoor walking and in-vehicle travel.

The goal of this part is to analyse performance of transport mode and trip purpose detection routines using the data described in Chapter 3. It is investigated if and how personalisation during processing of multi-day GPS and accelerometer data can improve the quality of the produced travel diaries. For analysis, applications for GPS processing where some of the available data is annotated are considered. This is for example the case for travel surveys, where participants can be asked to at least correct some of their schedule. Having annotated data is not granted, as position data is more and more often collected as a side product, especially in smartphone applications, such as journey planners (like PEACOX in Chapter 4).

The remainder of this part is structured as follows. First, all methods used are introduced, including random forests, an optimisation algorithm for fuzzy rules, as well as computation of input variables (features). Following, results are presented first for mode detection and then for trip purpose imputation. Conclusions and recommendations on future work complete Part II.

Chapter 7

Method for transport mode and trip purpose detection

Transport mode as well as trip purpose identification are multi-class classification problems. For classification, three tasks have to be completed, first, features, that is the input variables, have to be computed from raw data. Then the best performing features have to be selected and finally, a classifier has to be learned or implemented and its performance has to be tested.

For trip purpose identification, it has to be noted that the method described is based on multi-day GPS and accelerometer data for survey respondents living in the same region. To exploit the multi-day nature of the data, activities are clustered into locations. Clustering is done for single persons but also for the complete set of activities as several respondents might frequent the same public locations. Classification variables, called features, can then be derived for location clusters. To impute trip purpose the following steps are performed:

1. extraction of activities and their locations (Section 7.2),
2. clustering of locations (Section 7.3),
3. computation of features (Section 7.4), and
4. learning and applying the classifier (here we use random forests Section 7.5).

This chapter is structured as follows. Before the above mentioned steps a quick introduction about the performance measures employed is given. To conclude, an optimization algorithm for the fuzzy rule system is introduced, which is the original mode detection of our processing framework.

7.1 Performance measures

The overall accuracy is defined as the percentage of correctly identified observations. Performance of classification algorithms are also presented for some cases as confusion matrices, which indicate for each truth / classification pair the number of data points falling in that bin. In that context, accuracy measures are also given per class (e.g. trip purpose, transport mode), that is recall and precision are defined as follows:

$$\text{recall (class)} = \frac{\text{correct classifications of a class}}{\text{all actual observations of the class}} \tag{7.1}$$

$$\text{precision (class)} = \frac{\text{correct classifications of a class}}{\text{all activities predicted to be of that class}} \tag{7.2}$$

In the case of perfect classification, both recall and precision values are 100 % for all classes. Recall alone is not sufficient, as 100 % could also be achieved for one class if the classifier always predicts this class. As standalone value it is therefore not as meaningful as in combination with precision values and with the according measures for all other classes.

7.2 Activity location calculation

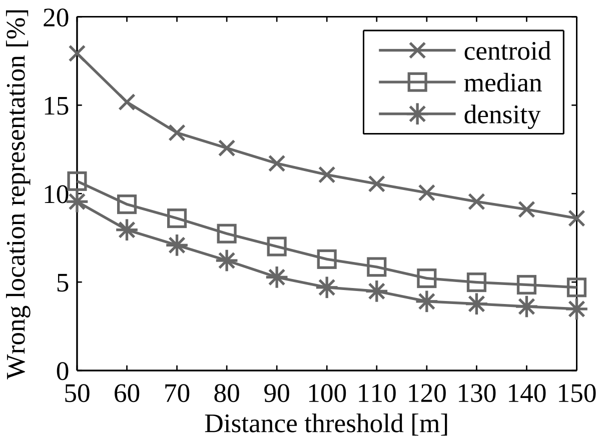
Activities are specified at least by start and end times and the GPS and accelerometer points recorded during this interval. Deriving these activities is either done by survey respondents or by an automatic activity detection module of a GPS processing framework. To assign each activity a location in space a representative coordinate is calculated from all GPS coordinates during the activity. Coordinates in our application are given in the metric National Swiss Grid representation. Using a metric grid is important when aggregating coordinates. To compare different aggregation approaches, only home activities were used for testing. Home locations are most suitable as they are known for all participants and they are visited several times within the survey period, which is important for subsequent clustering. Three approaches were tested:

- the unweighted point cloud centroid (which is the mean position of all points),

- median x and y coordinates (calculated separately) and,
- the coordinate with highest density (density was computed as number of coordinates within a radius of 20 meters around the candidate coordinate).

For comparison, distances between the calculated aggregated coordinates and the geocoded home coordinates were computed. Representations were categorized as wrong if the distance was higher than a given threshold. Figure 7.1 shows the erroneous coordinates for different distance thresholds. It can be seen that the centroid performs worst, the median coordinate is already much better, possibly because the influence of outliers is reduced. The best representation is achieved by the coordinate with the highest density. In this case for a threshold of 50 meters more than 90 % and for 100 meters already 95 % of home locations are well represented. Therefore, in the subsequent clustering, the coordinate with the highest density was used.

Figure 7.1: Comparison of centroid, median coordinate and densest coordinate as representation for an activity location.



7.3 Clustering

The total number of locations visited by respondents is not known. Therefore, clustering has to allow for as many clusters as needed. A well established approach that fulfills this condition is hierarchical clustering (Hastie et al., 2009). Agglomerative hierarchical clustering starts with each activity as one cluster. In each step the pair of clusters that is closest is merged into a new cluster. Clusters are merged as long as they are closer than a given cutoff distance. If two clusters consist of several activity locations, the distance between clusters is not straight forward. Thus, different linkage criteria were tested:

- single linkage: the distance between the nearest possible points of two clusters
- complete linkage: the maximum distance between two points of two clusters
- average linkage: the average of all distances of all possible point pairs of two clusters

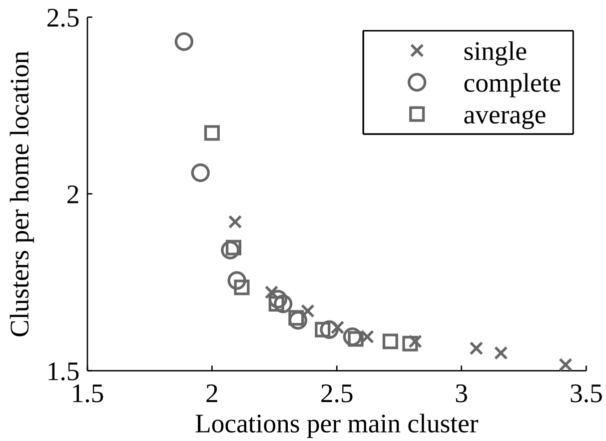
To determine optimal clustering, a location should be represented by one cluster and this cluster should only consist of activities of this location. These two outcomes were computed for all home locations, in order to compare the different linkage criteria. Home locations were again used, because they are most reliably known and they occur more than once per person. For each person in the data set the following procedure is executed:

1. cluster all available activities
2. count how many clusters contain home activities
3. the cluster with most home activities is assigned to be the main home cluster
4. count how many different locations are part of this main home cluster

For different cutoff distances, the mean values for all persons are calculated and plotted in Figure 7.2. The optimum would be at 1 for both, the number of clusters per home locations and the number of locations per main cluster. None of the clustering runs reaches the optimum value for either measure and it is unclear which of the two measures is more important. Overall, all linkage criteria perform similarly. Average linkage creates more compact clusters than single linkage, but is not as restrictive as complete linkage (Hastie et al., 2009). Therefore, average linkage was chosen for all subsequent analysis.

To decide on the best cutoff distance for the average clustering, the complete classification system was run with different cutoff distances. For each case 100 different forests were learned, the best accuracies were achieved with a cutoff distance of 100 meters, that is 72.3 % compared to 72.0 % with 10 meters, 68.2 % for 200 meters going down to 56.3 % for 1000 meters (mean accuracy as well as the 5th and 95th percentile for more cutoff distances are shown as part of the results in Chapter 8, Figure 8.7). Therefore, the 100 meter cutoff distance threshold is used in subsequent analyses.

Figure 7.2: Pareto efficiency of the mean number of clusters that cover home activities and the mean number of different locations within the main home cluster for several cutoff distances.



7.4 Feature computation

Features we use for trip purpose imputation can be divided into three groups: specific to persons, activities and clusters (per person as well as per data set). Similarly, for mode detection features are specific to person, stage and mode.

Appropriate person-based data are usually collected in the socio demographic questionnaire of a travel survey. Some examples are age, education level, income and mobility tool ownership. Moreover, the home address of respondents is often known, and sometimes even more addresses such as work place or favourite shopping centres are surveyed as well. For applications other than travel surveys, such data are probably optional and more sparse.

Simple activity-centred features are, for example, duration and start time. Using GPS data and the derived location representation, distances to important places such as home or work can be calculated. If a travel diary is available or can be automatically generated, the modes used to get to and leaving the activity can also be used. Further, during walking a very distinct accelerometer pattern can be observed. Therefore, the percentage that is spent walking during an activity can be computed. The underlying idea is that during shopping one might walk more than at home.

Cluster-specific features are associated with the location and not with single activities. For clusters that were computed for single persons, examples are statistical aggregates of the activity features such as mean and standard deviation of the duration. Clusters can also be created for the complete data set, which allows to extract the number of persons knowing a location. However, this feature has to be treated carefully as it is data set dependent, e.g. for a very homogeneous group working at the same university, work is a place everybody knows, for a more diverse data set of a region, train stations are more likely to be known to several people.

7.4.1 Feature selection

When building a random forest, the feature importance can be directly computed from the out-of-bag observations, that is the observations that were not used to construct a tree. The main idea is, that a feature is more important if the change of its value in an observation causes a misclassification. Therefore, for each feature the increase of misclassification is determined when the value of this feature is permuted in the out-of-bag observations. Hastie et al. (2009) show that this importance measure shows reasonable ranking but the distribution of the importance tends to be spread more uniformly than with other importance measures. This out-of-bag importance measure was used as basis for the selection of the features best suited for trip purpose and transport mode imputation. Furthermore, it was ensured, that all feature-groups were represented.

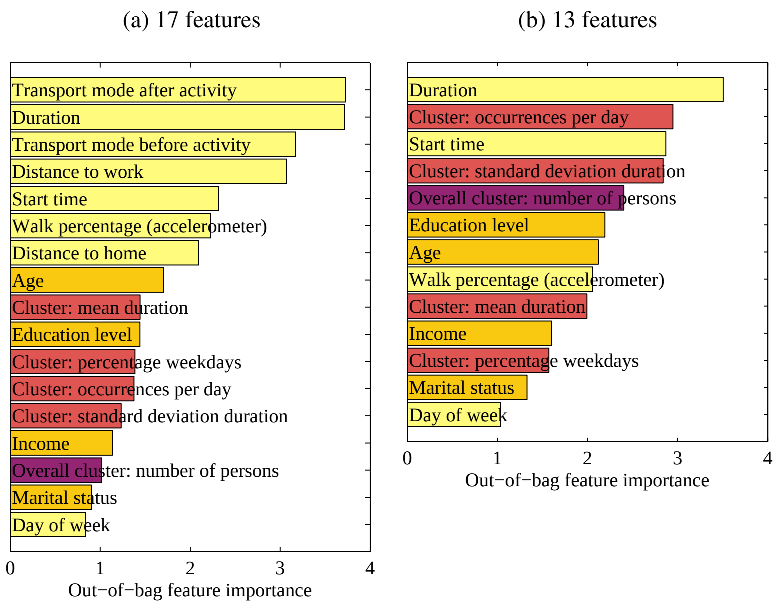
For trip purpose detection, around 40 potential features were tested. The feature sets selected to generate results are listed in Figure 7.3, where features of the same group are depicted in the same colour. Of the 17 features (the set with 13 features is a subset) the following are categorical features and treated as such in the classification process: day of week, education level, marital status and the transport modes. All other variables are continuous. Features that were not selected on the basis of the out-of-bag feature importance are mainly person-based features. Namely, all items related to mobility tools like possession of a driver's license or public transport season ticket were at the bottom of the importance ranking. The most important features are the activity-based features. The flag for weekday or weekend and the day of week, were not among the most important, but they are always available and reliable. Therefore, one of them, the slightly better evaluated day of week, was kept. Contrary to that, trip duration before the activity is left out, because feature importance was low and its computation is costly and the value is uncertain, especially in the context of uncorrected travel diaries. The modes before and after the activity, having the same drawbacks, are used anyway, as they are evaluated to be much more important.

7.5 Random Forest

For trip purpose imputation, decision trees, that is a set of rules learned by a machine and executed in a given order, were already used with good success (e.g. Griffin and Huang, 2005; Deng and Ji, 2010; Lu et al., 2012). Using a random forest, that is a set of decision trees also called ensemble of trees, is therefore, a natural step. Random forests were introduced by, and is a trademark of Breiman (2001).This classifier performs well in a variety of problems. It is also very popular, as it is easy to train and tune (Hastie et al., 2009). Breiman (2001) showed that random forest do not overfit even if more trees are added. A further advantage is that good results can be maintained even if data are missing, as they are estimated internally (Breiman and Cutler, 2013).

Technically, random forests work as follows. Each decision tree in the ensemble has one vote that counts for classification. The class with most votes, is the classification result. In a regular decision tree a data set is

Figure 7.3: Mean feature importance for 17 (a) and 13 (b) features respectively. Color coding for different feature types (activity, person, cluster and overall cluster features).



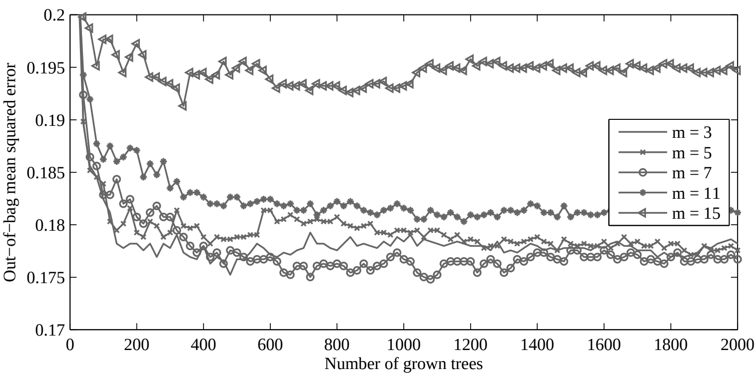
split using the feature that results in the best split. Using the same data to learn a tree, results in the same tree. But, in a random forest different votes are needed, and correlation between trees should be reduced to obtain best classification. To achieve that, on the one hand, each tree is learned from a different subset of the training data. On the other hand, at each split in the tree a random subset of features is considered. Each tree is fully learned, that is splits will be created until all training data are correctly classified.

To generate the subsequent results, either the random forest implementation of Matlab MathWorks (2012), called TreeBagger, or in the case of Java applications the FastRandomForest (Supek, 2012) based on the WEKA data mining tool (Hall et al., 2009) was used. The Java implementation is included in our open source position data processing framework (POSDAP, 2012).

7.5.1 Tuning and Stability Analysis

The main parameter that can be used for tuning random forests is the number of features (m) that should be randomly selected for each split decision. Figure 7.4 shows the out-of-bag error for a run using 17 features in total. It can be seen, that performance is similar for up to 7 randomly selected features at each split and then starts to drop slightly, as correlation between the trees starts to increase. For all runs with seventeen features $m = 7$ is selected. The recommended default value is the ceiling value of the square root of the number of features, which is $m = 5$ in the case of 17 features, as the default value performs well, all runs with less features use their respective default, which is $m = 4$. Furthermore, for all runs 500 trees are learned per classifier, as it can be seen that the error stays stable even if more trees are added.

Figure 7.4: Out-of-bag error for different numbers of randomly selected features (m).



7.6 Evolutionary algorithm to optimise fuzzy rule system

The original mode detection in POSDAP (2012) uses a fuzzy rules approach as described in Schüssler and Axhausen (2009c). The fuzzy rules systems is illustrated in Figure 7.6 and consists of linguistic input variables (e.g. median speed), each of which is split into a finite number of linguistic terms (e.g. *low*) that are described by the according membership functions. The shape of each membership is predefined in our case as trapezoidal and the functions of linguistic terms overlap and thus account for the fact that there is no sharp distinction between e.g., *low* and *medium* speeds.

Analysing the pretest data of the GPS study in Zurich, a set of rules, linguistic variables and according membership functions were derived by hand and are reported in Rieser-Schüssler et al. (2011). To further optimize the system automatically an evolutionary algorithm (see e.g., Michalewicz and Fogel, 2004), that is a biology inspired search heuristic, was implemented in MATLAB™. The idea of the algorithm is to start with an initial population of solutions (individuals) and evolve it generation by generation in order to find an optimal individual. New solutions are generated from old ones using variation operators that include some randomness (often crossover and mutation). The new population of equal size as the original (parameter *population size*) is then selected out of the old and new individuals based on a fitness criteria.

The following components of the classification system are defined beforehand and therefore not optimised:

- The variables
- The number of membership functions per variable
- The form of the membership functions (trapezoidal)
- The set of rules (Table 7.1)

For optimising the trapezoidal membership functions, the following building blocks of the algorithm are defined next:

- the individual, a representation of a solution (the fuzzy rule system)
- variation operators to create new individuals
- the fitness function, that is the measure that can be optimised, and
- a selection mechanism.

Table 7.1: Fuzzy rules for mode identification

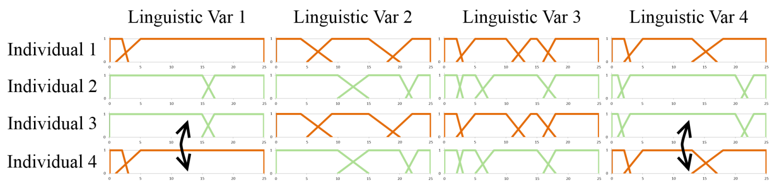
Median speed	Median variation accelerometer	95th-percentile speed	Mode
Very low	Very low	-	Urban PuT
Very low	Low	-	Bike
Very low	Medium	-	Bike
Very low	High	-	Walk
Low	Very Low	-	Urban PuT
Low	Low	-	Car
Low	Medium	-	Bike
Low	High	-	Walk
Medium	Very Low	-	Urban PuT
Medium	Low	-	Car
Medium	Medium	-	Bike
Medium	High	-	Bike
Medium	-	High	Car
High	Very low	-	Rail
High	Low	-	Car
High	Medium	-	Car
High	High	-	Car

Rules derived in Rieser-Schüssler et al. (2011)

7.6.1 The individual

For the problem at hand an individual has to represent the complete fuzzy rule system. The system is represented by all membership functions of each linguistic variable (Figure 7.5). The trapezoidal form is predefined as well as the value range of a fuzzy variable. Further, the number of membership functions of the fuzzy variable is also fixed. Therefore, an individual consists of the corner values of the trapezoidal functions.

Figure 7.5: Example of individuals consisting of 4 linguistic variables. Individuals 1 and 2 are the parents of individuals 3 and 4 generated by the crossover operation (arrows).



7.6.2 Variation operators: crossover and mutation

Inspired by crossover operators, two new individuals are created by switching the linguistic variables of two parent individuals as illustrated in Figure 7.5. It is randomly selected how many and which variables are flipped (at least one and less than all).

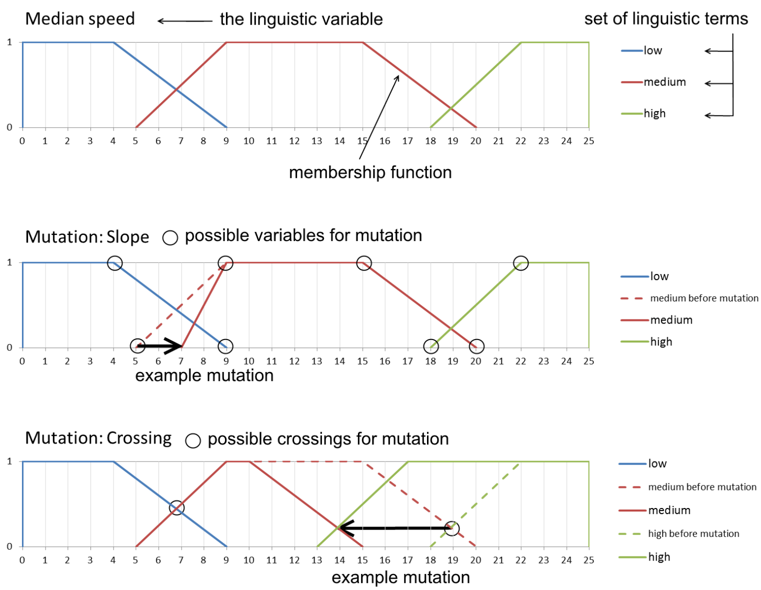
Two mutation operators are implemented: the slope and the crossing mutation (illustrated in Figure 7.6). In both cases the trapezoidal form, and the corners of the other membership functions is a constraint.

The slope mutation only changes one variable of one membership function. First, the variable to be mutated is randomly chosen, then one of the membership function parameters is selected and moved by a random amount, but such that the resulting membership function is still valid. This mutation only affect the slope of one membership function of one variable, therefore only small steps in direction of a potentially better solution are made.

Crossing mutation on the other hand, takes bigger steps. For the mutation, again a variable is selected randomly, followed by a random selection of the crossing of two membership functions as well as the direction of the move (left or right). The amount by which it is moved is defined as a proportion of the value range of the selected variable (parameter *range proportion*). E.g. if median speeds were constrained to be in the range of 0 to 200 km/h and the *range proportion* is set to 0.1 the crossing is shifted maximum by 20 km/h or if this shift is not feasible, the crossing is shifted by the maximum possible amount, that is constrained by the upper left or upper right corner of the left and right membership functions respectively (resulting in a triangle of one of the membership functions). As newer populations tend to be better, the influence of the crossing mutation is slowed down with each iteration, by reducing the *range proportion* linearly from *range proportion max* to *range proportion min* by *range proportion reduction*.

How many individuals are varied by each variation operator is specified by the run parameters *crossover probability*, *slope mutation probability* and *crossing mutation probability*. An individual can be drawn by all these operators at the same time.

Figure 7.6: Illustration of the elements of a fuzzy system and examples of mutation operations.



7.6.3 Fitness and selection

The fitness function is chosen to be the accuracy of the fuzzy rule system. In the selection step, a fixed number of fittest solutions (parameter *number of elites*) is kept and duplicated *elite multiplier* times, as it was found, that duplicates of good solutions speed up the search for an even better solution considerably. All other solutions are selected using weighted random draw with the fitness as weights.

Chapter 8

Results for transport mode and trip purpose detection

8.1 Transport mode

To analyse transport mode detection, first, accuracies of the optimised simple fuzzy rule system run on a subset of the Zurich GPS study (Chapter 3) are shown. Next, performance on the full Zurich GPS dataset using a random forest classifier is analysed. On the one hand, personalisation strategies are evaluated (Section 8.1.2) and on the other hand, performance using only accelerometer features is presented in Section 8.1.3.

8.1.1 Performance using optimised fuzzy rules

Before developing the optimisation algorithm, parameters were derived by hand based on the Institute's pretest data. In order to show feasibility of the method, the evolutionary algorithm was tested on a more diverse subset of data collected in the Zurich GPS study using median speed, 95th percentile of the speed as well as the median accelerometer measure as input features and classifying walk, bike, urban public transport, car and rail.

The introduction of an optimisation algorithm also allows to tune the fuzzy logic approach for different datasets.

Two different types of initial populations were used to test the algorithm. Completely random populations, where all individuals are different and generated as follows: the range of each linguistic variable is predefined, then random numbers are generated in this range and finally sorted such that the trapezoidal shape is kept. And a population consisting of 50 % randomly generated solutions needed for variation and 50 % identical solutions

Table 8.1: Run parameters for the optimisation algorithm

Parameter	
Selection	
Population size	50
number of elites	2
elite multiplier	4
Crossover	
probability	0.25
Slope mutation	
probability	0.25
Crossing mutation	
probability	0.25
range proportion max	0.1
range proportion min	0.001
range proportion reduction	0.00001 per iteration

derived from the pretest data (see Table 7.1). As this solution is better than a random solution this population is named *good*. All populations were evolved for 2000 iterations, using the run parameters as specified in Table 8.1 and described in Section 7.6.

Overall accuracy of the different runs for the training data as well as the test data is shown in Table 8.2. The training data (998 cases) was used to calculate the fitness of different solutions and to evolve the population, the test data (250 cases) on the other hand is only used at the beginning and the end to verify the results of the algorithm. In all cases the accuracy of the training and the test data are similar, therefore over-fitting is not a problem.

The good initial population (73.2 %) evolves considerably to an accuracy of 84.8 % and importantly this good result can also be reached by a random initial population. The three runs with random initial population start with the accuracy of the best individual between 40 and 51 % and all reach accuracies over 84.8 % for the test data. As depicted in Figure 8.1 evolution is very steep within the first 200 iterations and then slows down. For all runs the final result is stable, improvements after 500 iterations are small. Run time for 2000 iterations is around 20 minutes on a lenovo T440s Laptop running Windows 7 Enterprise.

The membership functions of the system derived by hand and the final result of the run with good initial population is shown in Figure 8.2. The median speeds are corrected towards lower values, or even 0 for *very low*, which shows that the fuzzy rule system could even be simplified, by removing this membership function and the 4 rules associated with it. In Table 7.1 it can be seen, that *very low* median speeds were introduced hoping to be able to distinguish between car and bike with similar accelerometer characteristics.

Analysing the confusion matrices of the two solutions for the test data, it is clear that no problem distinguishing car and bike existed in this dataset in the first place, therefore these rules probably could be dropped easily. The increase of accuracy is due to 10 of 13 bike stages originally misclassified as walk could be corrected. Further, distinction between public transport and car was shifted towards car, that is 6 car stages initially misclassified as rail and 19 misclassified as urban public transport are correctly detected as car with the best classifier, therefore, considerably increasing the recall of car (from 67 to 92 of 97 cases), at the expense of misclassifying 7 public transport trips and hence the decrease in recall of those. Depending on the objective, one has to consider that the overall accuracy of a classifier is not always the best fitness function, also because the number of trips per mode influences the overall accuracy, parameters are optimised towards those modes present most in the data set.

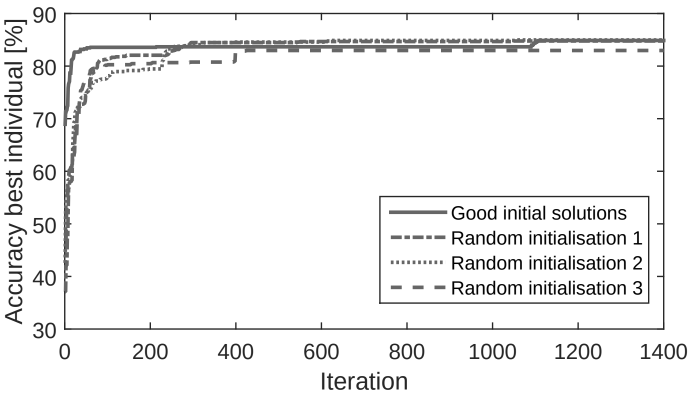
8.1.2 Personalisation using random forests

To show the influence of different features mode identification is performed with different feature sets. Training and test sets for the mode detection are built from random subsets of all stage observations. To be precise a 100-fold cross validation is performed, the results of which are shown in Figure 8.3. The first subset (run 1) includes the mode- and stage-specific features that are often used when performing mode detection. This run serves as minimum base line. The goal of adding the public transport-map-matching score (Rieser-Schüssler and Axhausen, 2013) is to improve distinction between car and public transport stages (run2). In Figure 8.3 it

Table 8.2: Evolution of accuracy of the training data for all runs

Initial population	Training Data		Test Data	
	It. 1	It. 2000	It. 1	It. 2000
Good	68.9 %	84.9 %	73.2 %	84.8 %
Random 1	37.1 %	84.8 %	40.4 %	84.8 %
Random 2	42.6 %	85.0 %	41.6 %	86.0 %
Random 3	47.0 %	83.5 %	50.4 %	85.2 %

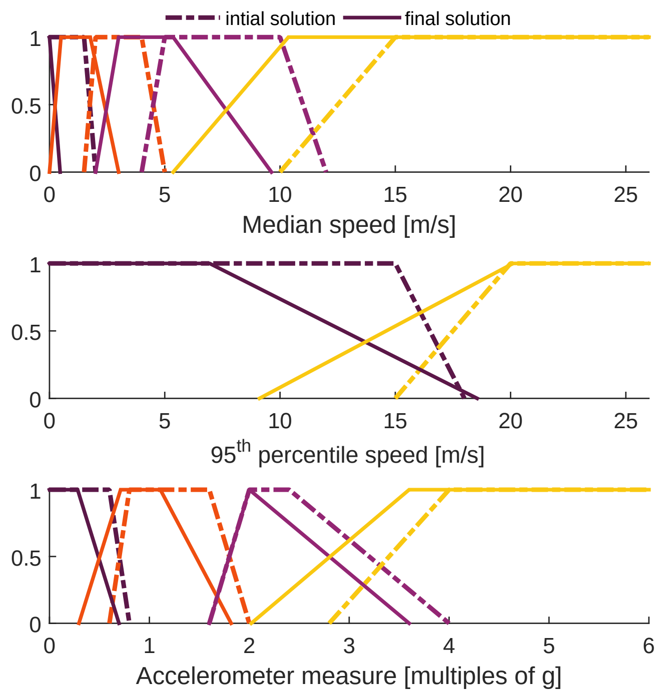
Figure 8.1: Evolution of accuracy of the training data for all runs.



can be seen that overall accuracy improves only slightly, but the percentage of bus stages correctly identified is increased from 20 % to 37 % (detailed results not presented). Therefore, the general effect is as expected, but the recall of bus stages is still very low.

Adding the personal mode shares (run 3) computed from the available one week of data as well as the mean bike and walk speeds of a person improves the classification with a median accuracy over 88 %. Importantly, also bus stages are better detected with a recall of 55 % (see confusion

Figure 8.2: Best membership functions of first and last iteration of the run with good initial population.



matrix in Table 8.3). Half of the misclassified bus stages are wrongly identified as car. However, more unexpectedly most of the other errors are bus stages misclassified as walk. Further adding socio-demographic information (gender, age, income, education, employment) has a negative effect on classification and is therefore not useful for mode identification (run 4).

For the best performing feature set, that is run 3, the feature importance measure for all used features is given in Table 8.4. The accelerometer measure is by far the most important followed by the mode shares that are observed for a person. Personal bike and walk speeds on the other hand are not that important and can therefore be neglected.

For the same feature set, a per person analysis is presented in Figure 8.4, where it can be seen that for mode identification the variation between participants is high, this is also true for trip purpose (Section 8.2.4). For one person only 30 % of the predicted modes are correct and for other persons 100 % are correct.

Figure 8.3: Accuracies for different feature sets, box plot for 100 runs each.

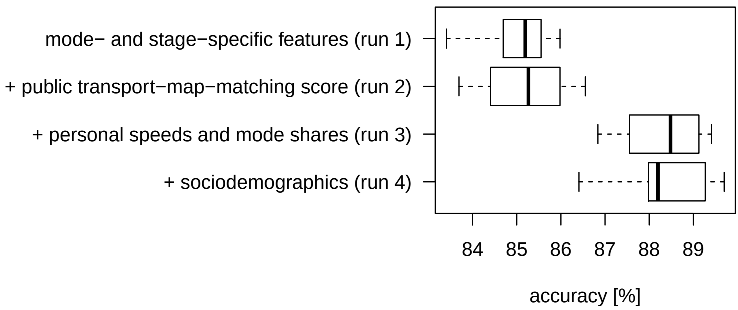
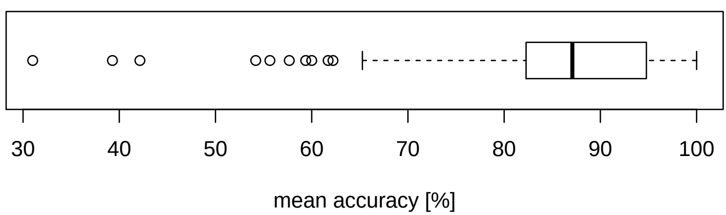


Figure 8.4: Per person mean accuracy of 10 runs. Same features as run 3.



Since including data collected by a participant has a positive effect on trip purpose detection for other data of this participant (Section 8.2.4) it

Table 8.3: Confusion Matrix: Run 3 including person-based features.

		Prediction							
		Walk	Bike	Bus	Tram	Rail	Car	Other	Recall [%]
Truth	Walk	1899	26	5	13	39	108	4	**90.7**
	Bike	73	504	4	0	2	22	0	**83.3**
	Bus	40	7	165	10	10	67	1	**55.0**
	Tram	38	3	7	300	19	12	0	**79.2**
	Rail	42	1	2	5	478	35	4	**84.3**
	Car	95	8	18	7	18	2780	5	**94.8**
	Other	14	0	0	5	1	37	57	**50.0**
	Precision [%]	**86.3**	**91.8**	**82.1**	**88.2**	**84.3**	**90.8**	**80.3**	

is also tested for transport mode classification. The number of days that are included in the training data is varied between 1 and 3 and the weights given to the person's training data is also varied (1, 10, 100). Unfortunately, compared to classification without including a person's data, there is no effect at all, as can be seen in Figure 8.5.

8.1.3 Analysis of accelerometer features and their importance using random forests

Collecting accelerometer data is cheaper in terms of battery consumption than GPS data, it is therefore often used to identify modes. Here it is shown how well mode can be identified given only features calculated from 10 Hz accelerometer data. Classification is done by random forests with 400 trees and results are reported for 10-fold cross validated runs.

Different statistical values describing the accelerometer series are tested

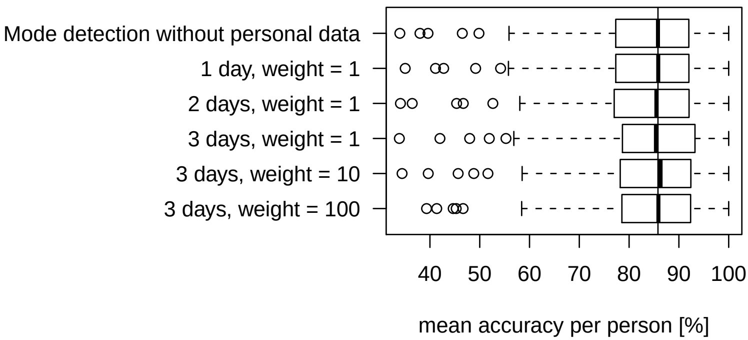
Table 8.4: Random forest out of bag feature importance measure (run 3), orange-coloured features are person-based, mode-specific features are dark red and light yellow features are stage-specific.

Median accelerometer measure	19.37
Bike share	4.22
Car share	3.92
Median speed	3.63
GPS quality (number of coords per second)	2.96
95th percentile speed	2.65
Public transport share	2.37
Duration	1.79
Public transport network-matching score	0.99
95th percentile accelerometer measure	0.96
Standard deviation of walk speeds	0.41
Start time	0.40
Mean of walk speeds	0.24
Mean of bike speeds	0.19
Standard deviation of speed	0.13
Standard deviation of bike speed	0.00

to find the best ones distinguishing different modes. The following values are tested:

- Percentiles (5th, 25th, 50th (median), 75th and 95th),
- mean value,
- standard deviation,
- variance,
- second moment,
- kurtosis, hinting at how tailed the distribution is,
- skewness, measuring the asymmetry of the distribution and
- RMS (root mean square) describing the energy of the signal

Figure 8.5: Mean accuracies of 10 runs per person including personal training data. Varying number of training days of the person to be classified as well as different weights.



All features are calculated for the complete stage but also for different window sizes, that is at every point in time a new value is calculated considering measurements of half the window size before and after that time. For classification then the mean, median or 95th percentile value of all windowed values is used.

Results are given in Table 8.5. Taking the medians resulted in best recall and overall accuracy (top part of Table 8.5). Further, the start and end of stages are cut assuming that the middle of the data is more typical and potential time rounding effects are reduced. This was not beneficial as all recall values and overall accuracy decrease with higher cut values (middle part of Table 8.5). Acceleration at the start and end are either important or all additional data is beneficial with the given amount of data. Further, as start and end times of the corrected data were extracted by algorithms no time rounding effects of participants distort the results, making a cut unnecessary. Window sizes between 10 seconds and 1 minute result in similar accuracies (77.4 % - 77.9 %), for a 2 minute window size accuracy slightly decreases (75.8 %), if no windowing is used an overall accuracy of 73.1 % is achieved (bottom part of Table 8.5).

Figure 8.6 shows the feature importance of all 12 features calculated

Table 8.5: Recall values for accelerometer-only mode detection

cut = 60 seconds, window size = 20 seconds				
aggregation		median	mean	95th percentile
Overall	6065 cases	77.9 %	75.6 %	69.5 %
Car	2602 cases	93.6 %	93.4 %	89.6 %
Walk	1794 cases	81.7 %	81.3 %	75.1 %
Bike	509 cases	67.0 %	63.9 %	48.1 %
Rail	463 cases	60.3 %	53.1 %	42.8 %
Tram	348 cases	58.3 %	35.1 %	26.7 %
Bus	257 cases	0.8 %	0.0 %	0.0 %
aggregation median, window size = 20 seconds				
cut [s]		30	120	300
Overall		78.3 %	76.9 %	75.5 %
Car		93.9 %	92.7 %	91.7 %
Walk		81.8 %	81.3 %	80.3 %
Bike		70.1 %	66.2 %	64.8 %
Rail		58.3 %	57.5 %	52.1 %
Tram		60.6 %	54.9 %	52.3 %
Bus		0.8 %	0.0 %	0.4 %
aggregation median, cut = 60 seconds				
window size [s]		10	60	120
Overall		77.4 %	77.4 %	75.8 %
Car		92.9 %	93.6 %	93.6 %
Walk		82.0 %	80.9 %	80.2 %
Bike		66.8 %	65.0 %	57.4 %
Rail		57.7 %	61.6 %	56.6 %
Tram		57.5 %	55.2 %	48.6 %
Bus		0.0 %	0.4 %	0.0 %

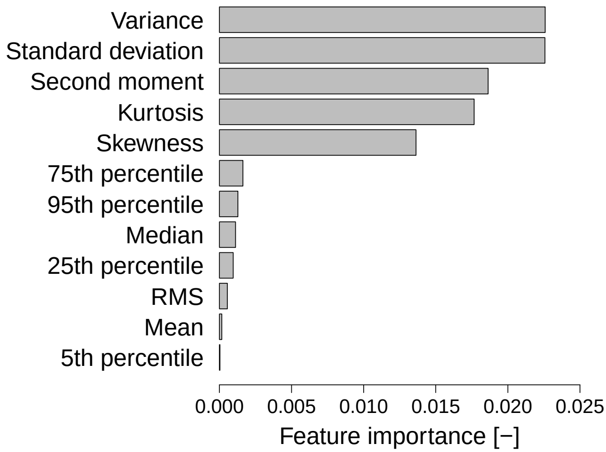
for a 10 second window without cutting start and end of stages (overall accuracy = 78.1 %). The mean, percentiles as well as the root mean square seem not important for classification, variance and standard deviations are most important. To exclude effects of similar measures stealing each others importance, classification was done for all combinations of feature pairs as well as triplets.

Using only two features resulted in lowest overall accuracies for mean combined with skewness of 51.2 % up to 74.9 %. For 15 pairs accuracies over 74.5 % are reached, achieved by combining any of the percentiles (including the median) with one of the following three: variance, standard deviation or second moment. All 27 combinations with those, except with each other, result in accuracies above 72.5 %, no other combination results in accuracies as high. Apart from those only the combinations of the RMS combined with percentiles achieve accuracies over 70 %. When analysing the confusion matrix of RMS and median compared to standard deviation and median, almost 5 % accuracy is lost by misclassifying bike as walk and vice versa, as well as misclassification between rail, tram and car.

When looking only at feature importance one would have expected skewness and kurtosis to be more relevant. Interestingly this is again the case when combinations of three features are used to learn the classifier. The best combination of three (only 95th percentile is used as a percentile, as pairing showed that they are interchangeable) is variance, skewness and kurtosis achieving 77.1 % overall accuracy which is only 1 % less than if all 12 features are used. On the other hand the worst combination is mean with skewness and kurtosis achieving only 63.8 % accuracy. In any case only 7 of the 84 tested combinations of three features result in an overall accuracy below 70 %.

Only considering accelerometer measures, reasonable results can be achieved.

Figure 8.6: Feature importance of accelerometer features.



8.2 Trip purpose

The tuning parameters of the classification procedure were discussed in the previous chapter. In this section, the accuracy and performance of random forests for trip purpose detection is analysed. To do so, different situations are compared and analysed in the follwing section:

1. A base situation for comparison is learned on a random subset of activities with 17 input features (*base setup*).
2. The features set is reduced, and changes in the classifier performance are discussed (*reduced features setup*).
3. The activity-based classification of the *base setup* is compared to a location-based approach (*location-based setup*).
4. The per person performance is analysed (*per person setup*).
5. The influence of person-specific input features is shown (*personalisation by features setup*)
6. Finally, four personalisation strategies are tested (*select best, grouping, includ person data, overrule*).

8.2.1 Base Setup

For the *base setup*, random forests are trained on randomly selected 75 % of all activities per activity type. The remaining data are used to test the performance. In total, 100 runs with 500 trees were employed. The mean accuracy, that is the share of correctly classified observations of these runs is 82.3 %. The highest accuracy reached is 84.4 %. The respective confusion matrix is reported in Table 8.6 together with the recall and precision values of each trip purpose.

It can be seen that mode transfers, being at home and work / education are recalled best. Almost all mode transfers (99 %) are identified as such and being home is recalled in 97 % of occurrences. Importantly, also the precision of the classifier for these three classes is good with around 85 % for working, 91 % for mode transfers and 93 % for being home, therefore only few activities are misclassified as one of these classes. Shopping and recreational activities are reasonably recalled with 74 % and 68 % respectively. The remaining trip purposes (pick-up / drop-off, business and other), that do not occur that often, are not well recalled (between 20 and 40 %) but there are also not too many misclassifications as pick-up / drop-off or business, as shown by high precision values of 83 %.

This *base setup* uses 17 features in total. In order to understand which features are important to distinguish trip purposes, their out-of-bag importance is shown in Figure 7.3(a). The 7 most important features are all activity-based (light yellow colored bars). The cluster or location-specific features (darker red colors) and the person-based features are slightly less important.

8.2.2 Performance with a Reduced Feature Set

Among the most important activity-based features are the transport modes used before and after the activity. In the context of travel surveys, when data have been corrected by participants this data are probably of high quality and it is beneficial to use it. Yet, they are not included in the reduced set, because in order to get good classification results reliable and accurate input data is needed, which is not always the case for transport modes. Particularly problematic can be cases where the mode is imputed and not double-checked by participants. In addition, distances to home and work are left out of the reduced set, in order to analyse, if these locations can also be imputed when their location is not known beforehand.

Mean accuracy of 100 runs for this *reduced feature setup* was 77.2 % and therefore, lower than the accuracy of the *base setup*. This was expected, as important features were left out. The confusion matrix of the best run (79.8 % accuracy) is reported in Table 8.6. Values are generally lower than compared to the *base setup* but the overall tendencies are similar. Note, that home and work locations are recalled as good as in the *base setup* with 95 % and 86 % respectively. Therefore, it would be possible to extract these location for each person in a first step, and to use the distance to the home and work location in a second step to improve classification. Recall of shopping and services on the other hand decreases substantially to 44 %. Most misclassifications are predicted as mode transfers, which consequently has a much lower precision than in the *base setup*.

The ranking of feature importance changes slightly compared to the *base setup* (Figure 7.3(b)). Especially some cluster-based features (occurrences per day, standard deviation of the duration, and the number of persons knowing a location) gain importance.

Table 8.6: Confusion Matrix: Best Run out of 100 for Two Different Feature Sets (Random Forest with 500 Trees).

Base setup with 17 features. Overall accuracy 84.4 %.

		Prediction: Mode transfer	Home	Work / edu.	Shop. / service	Recreation	Pick-up / drop-off	Business	Other	Recall [%]
Truth	Mode transfer	490	1	1	2	0	1	0	0	**99.0**
	Home	5	374	4	1	2	1	0	0	**96.6**
	Work / edu.	8	5	177	6	8	0	1	0	**86.3**
	Shop. / service	13	3	3	124	19	2	1	2	**74.3**
	Recreation	10	11	7	25	115	0	0	1	**68.0**
	Pick-up / drop-off	6	3	1	14	4	19	0	1	**39.6**
	Business	7	2	11	9	9	0	19	0	**33.3**
	Other	4	1	0	19	13	1	0	7	**20.0**
	Precision [%]	**90.6**	**93.3**	**85.1**	**62.6**	**69.3**	**82.6**	**82.6**	**69.2**	

Reduced setup with 13 features, leaving out potentially hard to obtain features. Overall accuracy 79.8 %.

		Prediction: Mode transfer	Home	Work / edu.	Shop. / service	Recreation	Pick-up / drop-off	Business	Other	Recall [%]
Truth	Mode transfer	479	3	0	9	4	0	0	0	**96.8**
	Home	8	369	5	1	4	0	0	0	**95.3**
	Work / edu.	10	8	177	1	7	1	1	0	**86.3**
	Shop. / service	58	2	7	74	18	2	5	1	**44.3**
	Recreation	17	6	5	19	121	0	0	1	**71.6**
	Pick-up / drop-off	21	2	1	6	5	10	2	1	**20.8**
	Business	14	1	9	3	8	0	22	0	**38.0**
	Other	16	0	5	5	14	1	1	3	**6.7**
	Precision [%]	**76.9**	**94.4**	**84.7**	**62.7**	**66.9**	**71.4**	**71.0**	**50.0**	

8.2.3 Performance of Location-Based Identification

So far, the classification has been done based on individual activities. However, a lot of the approaches for trip purpose identification in the literature rely only on the location of the activity, which corresponds more to a location-based identification of trip purposes. To compare these two approaches, the activity-based *base setup* is compared to a *location-based setup* learning a random forest classifier on clustered data using the same data source. This data source does not include land use data, which is typically used in location-based approaches, but it could be added to both classifiers if available. To learn the *location-based classifier*, the data have to be aggregated, that is locations have to be defined, and for each location the trip purpose has to be set, and features have to be computed. The location is represented by the cluster of activities as described in Section 7.3. The trip purpose for this cluster is derived from its activities where trip purpose is known. Most probably, not all activities are of the same purpose, hence, the one that occurs most often is chosen. Naturally, all cluster-based and person-based features are used, activity-based features on the other hand cannot be computed for clusters. The exceptions are the distance to work and home, as all activities in a cluster are per definition nearby and therefore, distances to other locations are still valid.

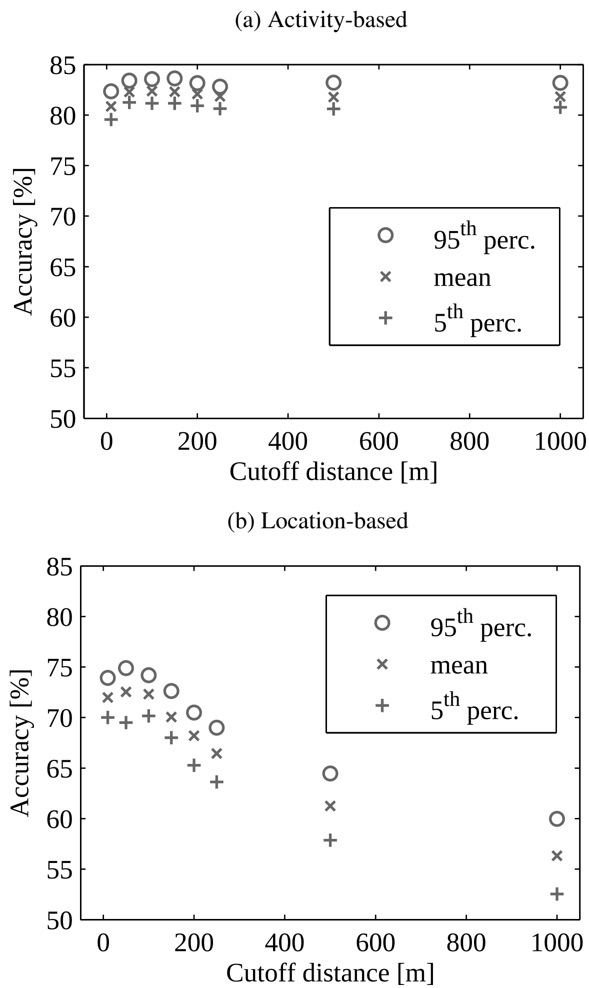
The *location-based classifier* is trained on 75 % of all clusters and classification is done for the remaining clusters. For the evaluation, all activities in a cluster are assigned the purpose imputed for this cluster, in order to compare accuracies of activity detection not location detection. Figure 8.7 compares the results of the *location-based setup* to the activity-based *base setup* for different cutoff distances. The influence of the cutoff distance is much higher for location-based identification, which is sensible, as almost all features are cluster dependent. Overall, the performance of the location-based system is worse (mean accuracy at 100 m cutoff 72.3 %).

8.2.4 Classifier Performance for Different Persons

When generating an automatic travel diary for survey respondents, the classifier is probably learned beforehand using a different data set or at least data from previous respondents. To simulate this situation, a classifier with all 17 features of the *base setup* is learned, but on a subset of persons not on a subset of activities.

For this *per person setup* the mean accuracy achieved is 78.4 % (5th

Figure 8.7: Accuracy of activity-based (a) and location-based (b) classification using different cutoff distances (hierarchical clustering with average distance).



percentile: 74.3 % ; 95th percentile: 82.1 %). Thus, the mean is 4 % lower than is the case for the activity-based *base setup* (mean: 82.3 %; 5th percentile: 81.2 %; 95th percentile: 83.2 %), also the variation within the 100 runs is higher for person-based training sets. This can be explained, if the variation of accuracies achieved for different persons is high. And indeed, the mean accuracy computed per person was 78.1 %, the 5th and 95th percentile are 50 % and 100 % respectively and the standard deviation is 16.6 %. One has to consider that, the quality of reporting for different respondents is very different. For some respondents, only a few and for others more than 100 valid activities were available, still the differences in prediction accuracy are uncomfortably large. Therefore, personalisation strategies are evaluated in the next two subsections.

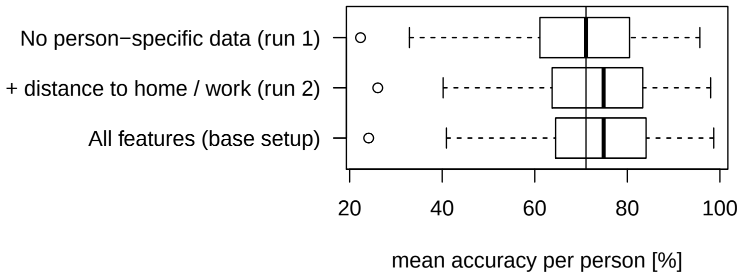
8.2.5 Personalisation Using Person-Specific Input Features

All personalisation results are presented as per person accuracies. For cross validation, 10 different classifiers are learned for each person, that is all other persons are randomly split into 10 groups, in each training set one of these groups is omitted. To show the effect of person-specific input features on the *per person setup*, a classifier is learned excluding all person-based features as well as the distance to home and work location respectively (full feature set see 17 features of *base setup* in Figure 7.3(a)). Results of this run 1 are shown in Figure8.8. In run 2 the above mentioned distances are included.

The mean of the accuracy of the 10 validation runs varies among participants between 25.3 % and 100 % with a median of the mean accuracies of the classifier without person-based data of 71.1 %. Including the person-specific features improves the results to 74.9 %. The per person standard deviation of accuracy are very similar for all runs. For the *per person setup* using all features the mean is around 2.3 % but goes up to 8.7 % for the person with highest variation within the 10 runs.

It has to be considered that the number of reported days, the validity of corrections and the trip purposes varies amongst participants. The split of trip purposes has probably the biggest influence on the spread of accuracies. As reconstruction of the diary is easier if a participant is e.g., only at home and at work. This influence is also illustrated in Figure 8.9 which depicts the mean accuracy per person versus the per person share of the three

Figure 8.8: Distribution of mean accuracies per person for different feature sets in trip purpose detection.



best predicted purposes, that is: being home, working (or studying) and changing the mode.

8.2.6 Personalisation Based On Corrected Data

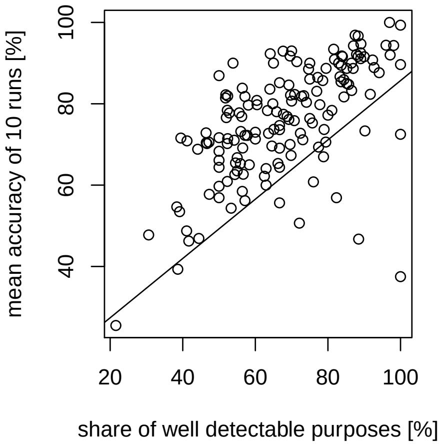
For trip purpose detection 4 strategies to personalise classification based on data corrected by participants are tested:

- Select best: selection of one classifier out of many based on performance on a subset of a person's data
- Group: group participants first and learn a different classifier for each group
- Include person data: include some of the person's data when learning the classifier
- Overrule: overrule the classifier when the location is already known

All strategies are subsequently described in more detail, but it is shown in Figure 8.11 that only inclusion of personal data improves predictions.

To ensure a fair comparison of the results of these strategies to the *base setup*, the influence of the amount of the training data is shown in Figure 8.10. It is clearly shown that classification is better the more data is used. The slope starts to flatten, therefore, the used data set is just about big enough to get a realistic estimation of how well purposes are detectable.

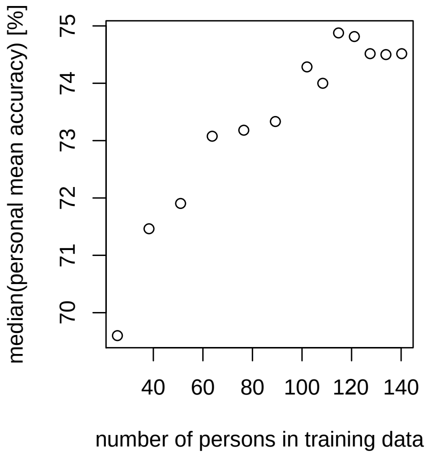
Figure 8.9: Mean accuracy of 10 runs for each person plotted against the share of easiest detected trip purposes



In the *select best* approach, for every person 10 random forests are evaluated on a subset of the person's data (selector data). The classifier that performs best on the selector data is then used to produce the results on the test data. The underlying hypothesis is, that some classifiers work better for one person's data than other classifiers and that this classifier is consistently better on this person's data. To create 10 different classifiers to select from, each classifier is learned on different person-subsets consisting of 100 persons. The selector data is fixed between the runs and consists of two randomly picked days. The mean standard deviation of the 10 test runs of the *base setup* with around 100 persons is 3.2 % therefore, there is some potential to select a better classifier. But as shown in Figure 8.11 accuracies are not increased. Furthermore, it was tested if only selecting weekdays would yield better results, assuming they contain more information, but it did not.

The idea behind *grouping* participants is that some of them have more similar diaries than others, hence, a classifier built using similar persons should be more successful. Three different criteria for grouping were tested.

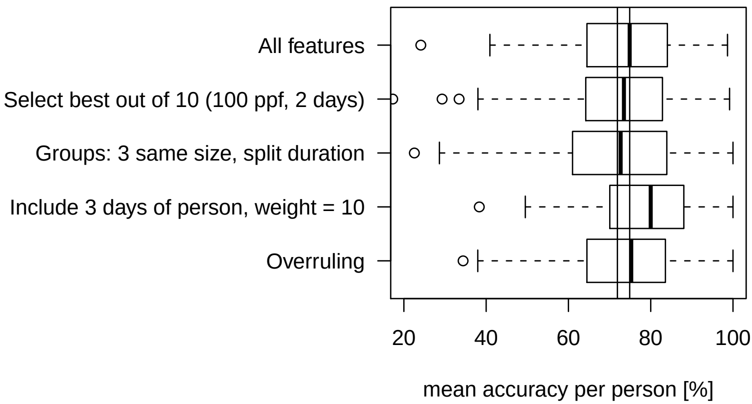
Figure 8.10: Median accuracy of the per person mean accuracy for different number of persons in the training set



First, participants are grouped based on all socio-demographic properties using hierarchical clustering. Then, for this first approach, classifiers are learned without the socio-demographic features. Second, participants are split into three groups of the same size based on the mean duration of all their activities. Finally, 86 participants were grouped as 'mostly using car', 35 as 'mostly using public transport or bike' and 32 as 'using both'. Neither of the groupings had any effect on the accuracies. In Figure 8.11, the results of the groups based on activity duration are shown. At first it looks like grouping even decreases accuracy, but it has to be considered that only 50 persons are used per group to learn the classifier. Hence, comparing it to the *base* classifier with 50 persons (left vertical line) shows that the grouping just does not influence results.

Including personal data when learning the classifier is straight forward. The person's data is split into a test and a training set. The training set consists of a given number of days that are randomly selected. As a part of one persons's data is not enough to learn a classifier, all other person's observations are added to the training set without special weighting. Results

Figure 8.11: Distribution of mean accuracies per person for all strategies. The vertical lines are the medians of the base runs with 50 (left) and 130 (right) persons respectively.



are compared for 1 randomly selected day and 3 randomly selected days. Whereas using 1 day does not improve classification, using 3 days increases the median accuracy to 80.0 %. Besides the number of days to be selected for training also the weights of the person's data was varied but did not have a relevant effect as shown previously for mode detection (Section 8.1.2).

To implement *overruling* each person's data is also split into test and training set. The classifier is learned on all training data (including the person's). But when classifying the test data, it is checked whether the person's training data includes an activity that was clustered into the same location. If this is the case the random forest is overruled and trip purpose is set to the one of the activity in the training set. If the training test set contains several activities at the same location with different purposes the purpose of the activity with the most similar duration is selected. Overall, overruling performs worse than the random forest. In total, 36627 classifications were made, of those almost 50 % (17671) are overruled and 79 % of those overruls were not necessary, 8 % were not helpful that is both the overrule and the random forest predict different but wrong purposes and

9 % of the overrules are counter-productive that is the random forest is correct. Especially home and mode transfer points are falsely corrected. Only 4 % (648) of the overrules are correct.

Chapter 9

Conclusion

The original fuzzy rules approach for transport mode detection was developed further, by employing an optimisation algorithm, which considerably improved the performance of the system. Compared to machine learning, the advantage of rule-based systems is that they can easily be shared with other research groups without having to share private GPS tracks. They can be employed without having to collect any data in advance. But the initial rules can be adapted using expert knowledge e.g. considering speed limits, and importantly, as shown they can automatically be improved with the evolutionary algorithm, as soon as data is available. The optimisation algorithm could possibly be further developed as many parameters are fixed, especially the rules themselves.

The fuzzy rule system performed well with 85 % overall accuracy. Random forests performed slightly better with 88 % using more features but also distinguishing between bus and tram. It is also shown, that good accuracies of almost 80 % can be reached if only accelerometer data is used. It is recommended to at least use one of the three best performing features: variance, standard deviation or second moment, these were calculated for 10 second windows, the median of those values over a stage was used for classification.

Random forests perform well with an overall mean accuracy above 80 % for trip purpose imputation on the data set at hand, which is relatively large and very diverse with respect to respondents and purposes. It is shown that quite a lot of data is necessary to achieve good results. For a classifier learned on data of 20 persons, which corresponds to approximately 100 person days, the median accuracy is around 4 % lower than for a classifier learned on 100 persons. Trip purposes that occurred more often in the data set also had higher recall values. This might be because random forests tend to favour the classes that appear more frequently in the training data if no good classifier can be built (Speed, 2003).But, the two most frequent activities home as well as mode transfers were also expected to give better recall values as they have more distinctive characteristics, e.g., mode transfers are typically very short and people tend to be home every night, thus they are easier to classify.

It was shown in Figure 8.7 that the activity-based classifier performed better than the location-based classifier. This, as well as the feature importance measures given Figure 7.3 are an indication that the use of activity-based features is important for good classification and therefore, these should not be neglected. An important input feature, that we did not use in neither of the two classifiers is land use data, which is usually associated with location-based imputation. We argue that inclusion of this data would improve both classifiers similarly. In the future, land use data should be considered if available, but this can be challenging depending on the application. For example if an application concerns different regions it is less likely that common land use data is found, than for a survey with clearly defined regional boundaries.

Considering applications where only GPS tracks and accelerometer data and no personal information is available, one can assume to get reasonable results with the approach presented in this part. On the one hand, because person-specific features are not as important as those computed for activities. On the other hand, concerning distance to the home and work location, it is shown, that these activities can be derived with good precision. Therefore, to impute a home or work coordinate, one could take all GPS coordinates of a person that are part of a predicted activity and then calculate the densest coordinate for those. A second classifier could then be run, that uses the distance to home and work as input feature, which was shown to be important (3.8 % increase of median accuracy).

For the most traditional application in travel surveys, it has to be considered that the variance of results for different persons is rather high for both, trip purpose and mode detection. Therefore, when presenting survey respondents with automatic diaries, some respondents will be much more annoyed than others and the burden varies as well. In practice, when conducting a GPS-based survey (Chapter 3) we also made the experience that it was participant-dependent how well the daily schedule was recognised. Hence, different personalisation strategies were tested.

The main conclusion concerning personalisation is, that it is worth collecting annotated data from participant's in a multi-day or even multiweek survey, as a median accuracy of 80.0 % is achieved if three days of annotated personal data are included in the training set. This corresponds to an increase in accuracy of 5.5 % compared to the base scenario. If the processing of the collected data is done after the survey, this is straight forward. For continuous processing during surveys, the classifier should be updated whenever newly corrected data is available, which was done in the PEACOX project (Chapter 4).

Interestingly, this strategy did not work for mode detection, but it was shown that using the mode shares of participants as input feature the mean classification accuracy is increased by 3 %. All other personalization strategies tested did not have an effect on accuracies. Grouping participants seemed like a good idea, but in essence when thinking in rule-based systems, this is just adding another rule at the beginning of the decision process. This contradicts the idea of decision trees, where the best possible split at any point is found automatically. Instead of grouping people according to a new variable, probably the easiest and most successful way is to add it to the feature set and make sure that it is not counter-productive.

The goal when selecting the best classifier out of many is that for each person a random grouping is found that performs better than an average classifier. However, first results were not promising. Maybe more classifiers with higher diversity would be necessary, but to achieve that, more training data is needed. A similar approach, that could be tested, is to use subsets of activities instead of creating classifiers from a subset of persons.

To conclude, one week of position data is not enough to highly personalise trip purpose detection routines. Especially the variance between participants is still very high and therefore, per person analysis of automatically processed data is problematic.

For mode detection, another step in direction of learning would be to impute the mode for a group of matched trips. Trip matching might be especially interesting for more sparse data, where gaps could be filled with knowledge from similar trips.

Part III

Applications

Chapter 11 of Part III is based on the following paper:

- Montini, L., A. Horni, N. Rieser-Schüssler and K. W. Axhausen (2012) Searching for parking in GPS data, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012

Chapter 10

Route and mode choice models

10.1 Introduction

The goal of this chapter is to describe and model route choice for car, public transport, cyclists and pedestrians observed in our multi-day multi-modal GPS study of the Greater Zurich Area (see Chapter 3).

When working with revealed preference data and real networks one of the problems is the huge number of possible path alternatives. Further, these paths overlap and models have to deal with resulting similarities. Both problems are analysed in Schüssler (2010). For choice set generation she introduces the Breadth First Search on Link Elimination algorithm (BFS-LE), which is used here, and compares it to the random walk (Frejinger et al., 2009), a branch & bound algorithm (Prato and Bekhor, 2006) as well as a stochastic choice set generation where link costs are randomised given a probability distribution before each least cost path computation, similar to (e.g., Ramming, 2002; Dugge, 2006; Prato and Bekhor, 2007; Bovy and Fiorenzo-Catalano, 2007). To account for route overlaps she compared different adjustment terms: two formulations of the C-logit model (Cascetta et al., 1996), two path size models (Ben-Akiva and Bierlaire, 1999), which is used here, as well as the path size correction term (Bovy et al., 2008) and two trip part specific path size factors (Hoogendoorn-Lanser and Bovy, 2007). This last factor is used here for the public transport models, as it performed best on the PEACOX trip planner data as shown in Fischer (2015). Route choice models for car on Swiss data were estimated e.g. in Bierlaire et al. (2006), and also the above mentioned comparisons were based on car stages observed with the external GPS data set described in

Section 5.1. This data set was also used to estimate bicycle route choice in Menghini et al. (2010) using the Path Size Logit approach as was done by Hood et al. (2011) and Broach et al. (2012). The same approach was also successfully used recently for pedestrian route choice (Rodriguez et al., 2015; Broach and Dill, 2015).

The remainder of this chapter is structured as follows. Section 10.2 briefly introduces the applied map-matching, modelling and choice set generation methods. Following, results for car, public transport, bicycle and pedestrian route choice models are presented. A combined mode and route choice model completes the results section before concluding with summary remarks.

Figure 10.1: Map-matching walk stages: GPS points on roundabout, weighting ensures that side-walk and crossing are chosen

10.2 Method

10.2.1 Map-matching and network-based variables

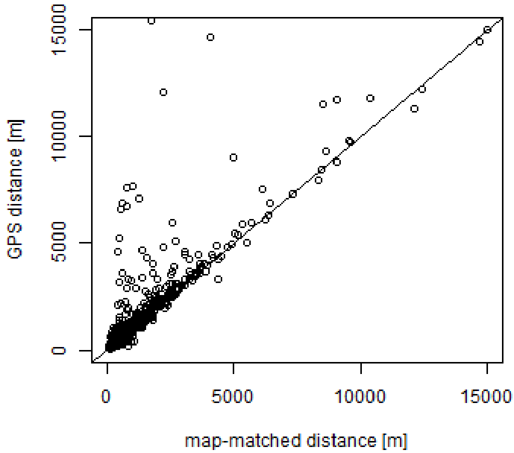
Matching the GPS points to a given network is done as described in Schüssler and Axhausen (2009b), the implemented module is part of the POSDAP (2012) framework. The map-matching routines were originally developed for navigation networks. Networks used here on the other hand are extracted from OSM and are spatially correct but therefore have many small links to e.g. represent a curved street. Hence, some constraints like number of points per link to get a valid match had to be loosened. Further, for map-matching of walk stages links that are pedestrian only were weighted slightly more, to ensure that people walked on side-walks and not in the middle of the street or on the roundabout as illustrated in Figure 10.1.

To ensure valid map-matching, the map-matched distance is compared to the distance computed from the GPS points. Figure 10.2 shows the comparison for car and pedestrian stages. Most distances are the same, and if they do not correspond, the map-matched distance is lower than the GPS distance, which is preferable, as invalid coordinates can cause high deviations and it shows that no detours are generated due to missing links.

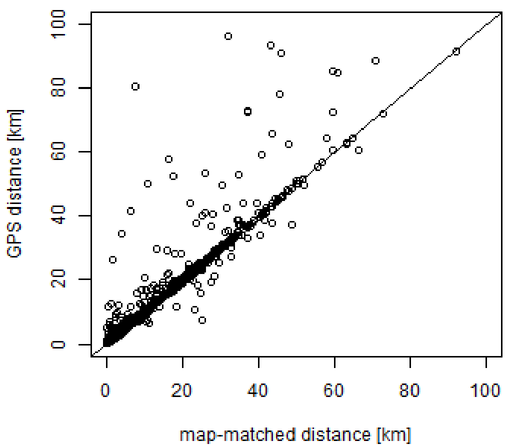
Routes are enriched with information on traffic signals and crossings that were extracted from OSM. Turn information is calculated for each route alternative as follows. A U-turn is easily identified if the link back to the origin node is chosen. Otherwise a node is identified as crossroad if

Figure 10.2: Map-matched distance compared to GPS distance

(a) Pedestrian



(b) Car



more than one outgoing link exists not counting the link back. To define if the chosen outlink is a left or right turn, first a straight link is defined. An outgoing link that is on the same OSM way (which is usually a named road) as the incoming link of the route is chosen as going straight. If there is no link on the same OSM way the link closest to an 180° angle (minimum 135°, maximum 225°) is chosen to be the straight link. A right turn is added to the route if the outgoing link is on the right of the incoming and straight link, otherwise a left turn is counted.

10.2.2 Route choice model

For route choice modelling the path size logit formulated by Ben-Akiva and Bierlaire (1999) is used, a multinomial logit (MNL) model (McFadden, 1974) with the path size (PS) as adjustment term to correct for overlapping routes. The MNL model is a discrete choice model where utility U is the sum of a deterministic part V and an identically and independently (i.i.d.) Gumbel distributed error term ϵ, given this definition the probability of alternative i being chosen out of the choice set C is:

$$P(i|C) = \frac{e^{V_i}}{\sum_j e^{V_j}} \tag{10.1}$$

The path size is given in Equation (10.2), it is 1 if a path does not overlap with any other in the choice set and it is very small if there is a lot of overlap. The general form of the deterministic part of the utility function is given in Equation (10.3), the path size is either transformed logarithmically such that it is very negative for very overlapping routes and 0 for routes without overlap or more generally a box cox transformation is applied.

$$\text{PS}_{\text{route,set}} = \sum_{\text{link} \in \text{route}} \left(\frac{\text{length}_{\text{link}}}{\text{length}_{\text{route}}} \right) \frac{1}{\text{\# routes in set via link}} \tag{10.2}$$

$$\begin{aligned} V =& \beta_{\text{TT}} * \text{travel time} + \beta_{\text{RT}} * \text{proportion road type}+ \\ & \beta_{\text{ELEV}} * \text{elevation measure} + \beta_{\text{PS}} * \ln(\text{path size}) \end{aligned} \tag{10.3}$$

Socio-demographic attributes as well as the available attitude scores could be introduced into the model via interaction terms. As it is not straight forward with which attributes especially attitudes should be interacted, alternatively a latent class approach is chosen. The idea is, that different behaviour is captured in Q classes, and for each class different parameters are estimated. Here, class membership is based on person attributes as indicated in Equation (10.4).

$$V_q = \beta_{CONST} * 1 + \beta_{\mathrm{SD}} * \text{socio demographic} + \beta_{\mathrm{SC}} * \text{scale score} \quad (10.4)$$

The probability of an alternative being chosen depends on the degree of class membership. If Q classes should be estimated, the sum of class membership probabilities has to be 1, therefore, $Q-1$ membership models and for each class a behavioural model has to be specified. The class membership models are also MNL models, the probability of being in a given class is hence given in Equation (10.1), with i being a class and C the set of all classes. Finally, the choice model is a discrete mixture of logit, the probability of alternative i being chosen is:

$$P(i|C) = \sum_{q=1}^{Q} P(i|\text{Class } q) * P(i|\text{Model } q) \quad (10.5)$$

For model estimation the python implementation of BIOGEME version 2.3 was used (Bierlaire, 2003). To correct for panel effects, the loglikelihood in all models is defined as sum of the conditional probabilities of a person, and only robust error measures computed with the sandwich estimator are reported (Daly and Hess, 2011).

10.2.3 Choice set generation

For car, bike and walk stages the Breadth First Search on Link Elimination algorithm (BFS-LE) as described in Rieser-Schüssler et al. (2012a) is used to generate route alternatives. The general working of the algorithm is as follows. Given a cost function the shortest path of the network is computed. To generate the next alternative one link of this shortest path is eliminated from the network, then the shortest path on the resulting subnetwork is computed. The procedure is repeated until the desired number of paths or all possible paths are generated, or until a timeout is reached. This algorithm was shown to be computationally very efficient while as well producing relevant routes, e.g. for bicycle routes in Halldórsdóttir et al. (2014).

Most important for the work with high resolution OSM networks is the performance optimisation of the BFS-LE implementation, where a topologically equivalent network is created before choice set generation. That is, vertices that are not a junction, intersection or a dead-end are removed and links are joined per direction. For the car network this means instead of 2'363'307 links 499'928 segments are processed.

For public transport the network is much more sparse instead schedules have to be considered when generating alternatives. Two of the approaches described in Rieser-Schüssler et al. (2014) are used here: the basic and the via point choice set generation. For the basic version connections are generated for all combinations of start and end stops that are within acceptable walking distance around origin and destination. The via point choice set generation further acknowledges that some stops provide well known transfer opportunities, also to other stops. A good example is the main station. In the case of Zurich, there are several tram and bus stops that all serve the station, all these are combined into a transfer set. When generating alternatives, additionally to selecting start and end stops, potential transfer sets are deterministically defined, and a subset is drawn randomly. Routing is then also done via the selected transfer sets.

For the following models up to 200 alternatives were generated for car routes and up to 100 for bike and walk stages. Additionally, the chosen path was added if not generated by the algorithm. These choice sets were reduced using similarity distribution-based reduction described in Schüssler and Axhausen (2009a). That is, as a similarity measure the path size is computed for the full choice set. Then path size bins are defined, and alternatives added such that each bin contains the same amount of alternatives if possible, otherwise they are spread uniformly over the remaining bins. For public transport all alternatives are generated given the run parameters as defined in Rieser-Schüssler et al. (2014).

10.3 Results

For model estimation some observations are excluded: Car stages where the map-matched distance is more than 2 kilometers longer or 1.5 times the GPS recorded distance are excluded. Bike and walk trips are removed if they are roundtrips or if the chosen route is more than 2.5 times the shortest route, assuming that these are sports trips that can not be explained by the available explanatory variables. Further, very short bike trips below 500 meters are removed as well as walk trips with an average speed above 5 m/s. For the mode choice experiment access and egress walk stages are not considered either, as these are part of the public transport connections.

10.3.1 Car route choice model

For car routes, following models are estimated:

1. A simplified model to determine choice set size.
2. Model 1 with road type specific free-flow travel times and time proportions.
3. Model 2 adding turns and traffic lights to Model 1.
4. Model 3 replacing travel times per road type of Model 2 by the logarithm of free-flow travel time.
5. Finally, Model 3 split by trip purposes (work, leisure, remaining)

The influence of the choice set size on the parameter estimates is shown in Table 10.1 given the most simple model considering all types of input features available. That is the travel time, the sum of turns and traffic lights as well as the average speed, which is a rough indicator for the road types, as these are to a great extent distinguishable by their speed limits. To ease comparison, results are presented as values scaled to the time parameter of the choice set with 80 alternatives. Parameter estimates start to stabilise with 60 alternatives in the choice set. The path size parameter is the least stable, it even changes sign at a choice set size of 60 and is constantly increasing with choice set size.

For the following car route models a choice set size of 80 is used. Parameter estimates of three models are given in Table 10.2.

Travel times per road type, the time proportions on these road types as well as path size correction factor are the explanatory variables of Model 1, similar variables are used in Bierlaire et al. (2006) and Schüssler and Axhausen (2009a), two models estimated on Swiss data (different datasets).

Travel time parameters have the expected sign, and time on residential roads is least attractive. The two before mentioned models also value small and local roads respectively least, for these models also the proportion of those is most negative. Here small roads are represented by the road types residential and track, where proportions of track is the most negative, but the parameter for residential road is not as negative as e.g. motorways. In Models 2 and 3, the residential road proportion parameter even becomes positive, but values are not significant. Adding turn and traffic light counts (Model 2), seems to explain the negative valuation of residential roads between Model 1 and 2. Adding this information significantly improves the model fit from 0.161 to 0.241. Left turns are perceived worse than right turns, and both are more negative than traffic lights.

The best model fit of 0.324 is obtained with Model 3 where the free-flow travel times per road type are replaced by the natural logarithm of total travel time. The same difference in travel time for short trip leads to bigger difference in utility as for a long trip, which is sensible.

Model 3 is split into three models for work, leisure and remaining trips. Parameter estimates are given in Table 10.3. For work trips the time parameter as well as the left turn and traffic light parameters are lower than for the other two models, it is sensible that these often recurring trips are more optimised than leisure trips. The proportion of motorways on the other hand is more negative and the proportion of track less negative, which is unexpected.

Table 10.1: Scaled parameter values of car route MNL model for different choice set sizes

Size	Scale	Time[1] [min]	Turns and lights[2] [/km]	Speed[3] [km/h]	ln(PS[4])	$\bar{\rho}^2$
10	1.794	-1.15	-1.18	-0.080	-1.07	0.30
20	1.363	-1.15	-1.18	-0.055	-0.57	0.22
40	1.106	-1.15	-1.13	-0.039	-0.15	0.17
60	1.036	-1.15	-1.12	-0.034 *	0.04	0.15
80	1	-1.15	-1.12	-0.030	0.15	0.14
100	0.991	-1.15	-1.13	-0.029	0.23	0.13
120	0.975	-1.15	-1.13	-0.028	0.26	0.13
140	0.975	-1.15	-1.13	-0.028	0.30	0.13
all	0.958	-1.15	-1.13	-0.026	0.33	0.13

[1] free flow travel time

[2] sum of traffic lights, left and right turns

[3] average speed (distance / free flow travel time)

[4] path size of the unreduced choice set based on link travel times

* value not significant on a 95-% level

Table 10.2: Car routes

	Model 1		Model 2		Model 3	
ln(free-flow travel time [min])	-		-		-7.51	
Robust std err | Robust t-test	-	-	-	-	0.681	-11.02
FF time motorway / trunk [min]	-1.07		-0.985		-	
Robust std err | Robust t-test	0.173	-6.2	0.164	-5.99		
FF time extra-urban [min]	-0.869		-0.795		-	
Robust std err | Robust t-test	0.163	-5.32	0.154	-5.16		
FF time urban main [min]	-1.01		-0.863		-	
Robust std err | Robust t-test	0.122	-8.29	0.113	-7.66		
FF time track / other [min]	-1.01		-1.07		-	
Robust std err | Robust t-test	0.211	-4.77	0.214	-4.97		
FF time residential [min]	-1.3		-1.16		-	
Robust std err | Robust t-test	0.196	-6.65	0.166	-6.95		
T(prop. motorway)	-2.31		-2.34		-2.47	
Robust std err | Robust t-test	0.797	-2.90	0.785	-2.99	0.418	-5.91
T(prop. extra urban)	-3.10		-3.11		-2.54	
Robust std err | Robust t-test	0.869	-3.57	0.774	-4.01	0.720	-3.53
T(prop. urban main)	0 (fixed)		0 (fixed)		0 (fixed)	
T(prop. residential)	-1.90		0.727 *		0.460 *	
Robust std err | Robust t-test	0.562	-3.38	0.552	1.32	0.409	1.12
T(prop. track / other)	-6.80		-3.9		-4.29	
Robust std err | Robust t-test	0.760	-8.96	0.724	-5.39	0.735	-5.84
$\sqrt{\text{left turns / km}}$	-		-2.53		-2.74	
Robust std err | Robust t-test	-	-	0.135	-18.76	0.202	-13.53
$\sqrt{\text{right turns / km}}$	-		-2.32		-2.6	
Robust std err | Robust t-test	-	-	0.139	-16.76	0.195	-13.3
$\sqrt{\text{traffic lights / km}}$	-		-1.23		-1.29	
Robust std err | Robust t-test	-	-	0.238	-5.15	0.281	-4.58
ln(path size)	0.36		0.227		0.36	
Robust std err | Robust t-test	0.0565	6.36	0.0552	4.1	0.0625	5.75
Init log-likelihood $\mathcal{L}(\beta_0)$	-8950.6		-8950.6		-8950.6	
Final log-likelihood $\mathcal{L}(\hat{\beta})$	-7498.172		-6781.998		-6041.941	
$\bar{\rho}^2$	0.161		0.241		0.324	

T(): arcus sinus transformation $\tilde{y} = \arcsin(\sqrt{y})$
* value not significant on a 95-% level

Table 10.3: Car route choice models for different trip purposes based on Model 3

	Work		Leisure		Remaining	
ln(free-flow travel time [min])	-10.0		-7.39		-8.69	
Robust std err | Robust t-test	1.29	-7.74	1.55	-4.78	0.615	-14.12
T(prop. motorway)	-3.00		-2.44		-2.33	
Robust std err | Robust t-test	0.844	-3.56	1.03	-2.37	0.425	-5.48
T(prop. extra urban)	-3.76		-0.353 *		-2.72	
Robust std err | Robust t-test	1.56	-2.41	0.991	-0.360	0.726	-3.75
T(prop. urban main)	0 (fixed)		0 (fixed)		0 (fixed)	
T(prop. residential)	-0.203 *		-0.023 *		0.0545 *	
Robust std err | Robust t-test	0.796	-0.25	0.720	-0.030	0.504	0.11
T(prop. track / other)	-2.55		-6.72		-5.18	
Robust std err | Robust t-test	1.11	-2.29	1.07	-6.29	0.638	-8.11
$\sqrt{\text{left turns / km}}$	-1.82		-1.13		-1.44	
Robust std err | Robust t-test	0.319	-5.71	0.290	-3.88	0.153	-9.44
$\sqrt{\text{right turns / km}}$	-1.10		-1.1		-1.4	
Robust std err | Robust t-test	0.340	-3.22	0.209	-5.30	0.149	-9.3
$\sqrt{\text{traffic lights / km}}$	-1.16		-0.236 (*)		-1.09	
Robust std err | Robust t-test	0.442	-2.63	0.563	-0.420	0.201	-5.42
ln(path size)	0.639		0.274		0.404	
Robust std err | Robust t-test	0.110	5.79	0.137	2	0.0668	6.05
Sample size	537		419		1111	
Init log-likelihood $\mathcal{L}(\beta_0)$	-2335.19		-1815.59		-4799.88	
Final log-likelihood $\mathcal{L}(\hat{\beta})$	-1621.79		-1314.75		-3159.73	
$\bar{\rho}^2$	0.302		0.271		0.340	

T(): arcus sinus transformation $\tilde{y} = \arcsin(\sqrt{y})$
* value not significant on a 95-% level

10.3.2 Public transport route choice model

For public transport, model results are presented in Table 10.4 for the basic as well as the via choice set generation. Tendencies of the two models are the same. With the exception of the very small and not significant but positive access travel time parameter, all parameters have appropriate signs. Positive transfer time parameter makes sense if tight connections are perceived negative, furthermore, the parameter is very small. The most important parameters are number of transfers, in-vehicle travel time as well as the share of tram stages. One transfer is as bad as 11 minutes of in-vehicle travel time, respectively 10 minutes for the via model. In order to have only tram stages instead of only bus stages, almost one transfer is acceptable.

Table 10.4: Public transport route choice models for different choice set methods

	Basic CSG		VIA CSG	
Access travel time [min]	0.00477 *		0.000975 *	
Robust std err | Robust t-test	0.00925	0.52	0.00752	0.13
Egress travel time [min]	-0.00487		-0.00410	
Robust std err | Robust t-test	0.00224	-2.18	0.00204	-2.01
In-vehicle travel time [min]	-0.238		-0.229	
Robust std err | Robust t-test	0.0577	-4.13	0.0516	-4.45
Transfer time [min]	0.00456 *		0.00345 *	
Robust std err | Robust t-test	0.00250	1.82	0.00198	1.74
Number of transfers	-2.13		-2.30	
Robust std err | Robust t-test	0.252	-8.47	0.226	-10.17
Prop. bus stages [-]	0 (fixed)		0 (fixed)	
Prop. rail stages [-]	-0.451 *		-0.339 *	
Robust std err | Robust t-test	0.556	-0.81	0.597	-0.57
Prop. tram stages [-]	1.84		2.02	
Robust std err | Robust t-test	0.325	5.67	0.372	5.43
Path size (stage times)	1.73		1.37	
Robust std err | Robust t-test	0.513	3.38	0.474	2.89
Sample size	268		273	
Null log-likelihood $\mathcal{L}(\beta_0)$	-667.780		-800.181	
Final log-likelihood $\mathcal{L}(\hat{\beta})$	-299.889		-317.795	
$\bar{\rho}^2$	0.539		0.593	

* value not significant on a 95-% level

10.3.3 Bicycle and pedestrian route choice

For bicycle the following models are estimated:

1. A basic bicycle model, considering bike path, maximum rise and traffic lights to determine choice set size.
2. A latent class model, considering gender and attitudinal scores.
3. A MNL model with gender-specific parameter estimates.

Similarly, for pedestrians the models specified are:

1. A basic model, considering distance, maximum rise and maximum fall.
2. A latent class model, considering gender and attitudinal scores.
3. A MNL model for access and egress stages and one for all other walks.

In Menghini et al. (2010) the Swiss GPS data set described in Section 5.1 was used to estimate a bicycle route choice model for Zurich. The basic Model 2 of that paper is estimated here using the new data (Chapter 3). Results for several reduced choice set sizes are given in Table 10.5. For the different choice set sizes, estimates are consistent except for the path size parameter which decreases with growing choice set. For subsequent analysis a choice set size of 40 is chosen. Comparison with the above mentioned model, is more difficult as parameter estimates are very different. On the one hand maximum rise, traffic lights as well as the interaction of maximum rise and distance are not significant for choice set sizes bigger than 40. And contrary to expectations the proportion of bike paths is highly negative. Even though the very negatively perceived roads of type *track* were excluded from bike paths for this analysis. In Figure 5.2 they are included as safe roads.

A positive bike path parameter is only estimated for class 1 of the latent class model presented in Table 10.6. On the other hand in all models the proportion of residential roads is positive. Suggesting that participants perceive those as safer than main roads with a painted bike lane.

As gender is one of the significant parameters of the class membership model, for comparison an MNL model is estimated where parameters are distinguished by gender (Table 10.6). The model fit is much higher for the latent class model therefore capturing the data better.

First, it has to be noted that degree of membership for class 1 is lower than that for class 2 in the bike model as depicted in Figure 10.3. The Figure distinguishes participants for whom a bike stage was observed, and therefore influenced the membership model. They do not differ from the rest of the participants. The significant variables for class membership are gender, as well as variation score 1 mainly covering *interest in daily routine variation* and variation score 2 that emphasises *spontaneity* the risk score 1 covering *health related* risks was significant in the previous step and therefore kept in the model. Other variables that were tested that were not significant are: age, public transport subscription, number of bikes, as well as the remaining attitude scores. Female participants that are not that interested in varying their daily routine but are still spontaneous are most influenced by the class 1 model. In the class 1 model the distance parameter is extremely negative, the trade off with proportion of residential road is -1.3 m/%, for class 2 it is -47.4 m/%, for male -36.4 m/% and for women -7.1 m/%.

Table 10.5: Parameter values of bicycle route MNL model for different choice set sizes

Size	Dist. [km]	Prop. bike path [-]	Max rise [-]	Dist. * max rise	Nr. traffic lights	β_{PS} [1]	λ_{PS} [1]	$\bar{\rho}^2$
10	-1.70	-6.99	-8.87	-1	0.176 *	9.83	2.45	0.46
20	-1.56	-7.29	-15.1	0.328 *	0.202 *	12.2	1.92	0.46
40	-1.93	-7.86	-11.1 *	-0.794 *	0.208 *	8.75	1.21	0.49
60	-1.92	-8.02	-11.4 *	-0.769 *	0.200 *	7.13	0.95	0.50
80	-1.78	-8.03	-11.6 *	-0.827 *	0.188 *	5.12	0.68	0.50

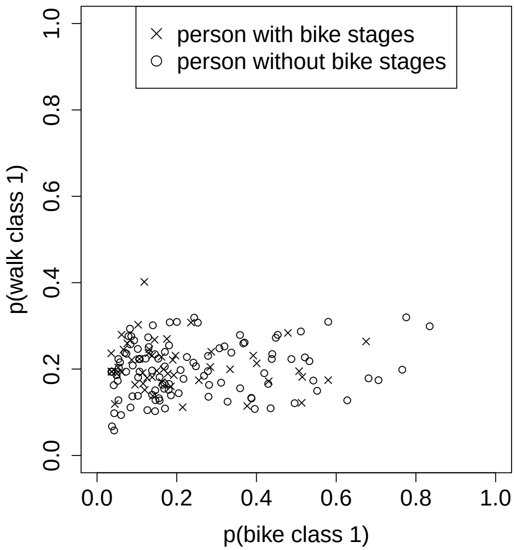
[1] Path size box cox transformed: $\beta_{PS}\frac{PS^{\lambda_{PS}}-1}{\lambda_{PS}}$

* value not significant on a 95-% level

For pedestrian route choice the influence of choice set size is analysed in Table 10.7. Values are scaled by the distance parameter. Again it is path size that changes sign as for the car model. Other than that, parameters are similar.

As for cycling a latent class model as well as a MNL model are estimated. The MNL model distinguishes between access/egress stages and all others. Model fit of the latent class model is again substantially better even though the degree of membership for class 1 is rather low, that is between 0 and 0.4 and is less spread than was the case for the bicycle model (Figure 10.3).

Figure 10.3: Class membership probabilites for the bike as well as pedestrian latent class models



Membership of class 1 is highest for those who are in *agreement with CO_2 reduction measures* (Env. score 1). Apparently they do prefer flat routes even if more turns are necessary. Class 1 seems to capture going for a walk while class 2 is just optimising distance.

In the model for access and egress stages distance is valued more negative then for other walks, this is as expected. The only other significant parameter is number of turns / km, which has the wrong sign but it is very small. For the remaining walks also maximum rise and fall as well as proportion of residential roads are significant and have the appropriate sign.

Table 10.6: Bicycle route choice model (choice set size 40)

	LC class 1	LC class 2	MNL (male)	MNL (female)
Distance [km]	-41.6	-1.41	-1.18	-5.74
Robust std err | Robust t-test	13.0 -3.19	0.473 -2.98	0.460 -2.56	0.744 -7.71
Average rise [-]	-4.41	-0.924 *	-1.74	-1.26
Robust std err | Robust t-test	1.74 -2.53	0.521 -1.78	0.647 -2.69	0.373 -3.38
Prop. flat [-]	-	-	0.232 *	2.35
Robust std err | Robust t-test			0.508 0.46	0.350 6.72
Prop. bike path [-]	8.64	-8.52	-7.84	-4.52
Robust std err | Robust t-test	3.42 2.53	2.10 -4.05	1.62 -4.83	1.98 -2.28
Prop. residential [-]	5.38	6.69	4.29	4.05
Robust std err | Robust t-test	2.33 2.31	1.31 5.12	0.880 4.88	0.909 4.46
Nr. crossings / km	0.581 *	0.500	-0.356	-0.256 *
Robust std err | Robust t-test	0.429 1.35	0.153 3.27	0.116 3.07	0.229 -1.12
Nr. left turns / km	-	-	-0.188 *	-0.253 *
Robust std err | Robust t-test			0.116 -1.61	0.135 -1.88
Nr. right turns / km	-	-	-0.101 *	-0.428
Robust std err | Robust t-test			0.0816 -1.24	0.176 -2.43
ln(path size)	1.42 *	4.24	1.33	1.25
Robust std err | Robust t-test	0.784 1.90	0.411 10.34	0.199 6.68	0.240 5.21
	Membership model			
Constant	-1.00			
Robust std err | Robust t-test	0.407 -2.46			
Male	-1.14			
Robust std err | Robust t-test	0.410 -2.78			
Variation score 1	-0.695			
Robust std err | Robust t-test	0.244 -2.84			
Variation score 2	0.417			
Robust std err | Robust t-test	0.168 2.49			
Risk score 1	0.0799 *			
Robust std err | Robust t-test	0.0589 1.36			
Sample size	411		411	
$\mathcal{L}(\beta_0)$	-1231.195		-1513.422	
$\mathcal{L}(\hat{\beta})$	-329.187		-697.357	
$\bar{\rho}^2$	0.719		0.527	

* value not significant on a 95-% level

Table 10.7: Scaled parameter values of pedestrian route MNL model for different choice set sizes

Size	Scale	Dist. [km]	Max rise [-]	Max fall [-]	ln(PS)	$\bar{\rho}^2$
10	1.266	-6.18	-2.86	-3.62	-2.115	0.27
20	1.077	-6.18	-3.68	-3.96	-0.978 *	0.20
40	1	-6.18	-3.80	-3.94	2.280	0.12
60	1.053	-6.18	-4.93	-3.90	4.106	0.10
80	1.140	-6.18	-5.35	-3.55	5.154	0.09

* value not significant on a 95-% level

Table 10.8: Pedestrian route choice model (choice set size 20)

	LC class 1	LC class 2	MNL (Access/ Egress)	MNL (Rest)
Distance [km]	0.151 *	-13.7	-7.59	-3.67
Robust std err | Robust t-test	0.357 0.42	1.28 -10.75	1.13 -6.7	1.42 -2.59
Rise max [-]	-11.7	-0.0924 *	-0.0455 *	-0.0449
Robust std err | Robust t-test	3.42 -3.42	0.967 -0.10	0.0263 -1.73	0.0228 -1.97
Fall max [-]	-6.71	-0.749 *	-0.0338 *	-0.0445
Robust std err | Robust t-test	3.03 -2.21	1.92 -0.39	0.0216 -1.56	0.0191 -2.33
Prop. residential [-]	1.10 *	0.405 *	0.0815 *	1.25
Robust std err | Robust t-test	0.867 1.27	0.408 0.99	0.392 0.21	0.508 2.46
Nr. crossings / km	1.52 *	0.881 *	0.0264 *	0.0544 *
Robust std err | Robust t-test	1.49 1.02	0.570 1.55	0.0257 1.03	0.0363 1.5
Nr. turns / km	0.382	-0.196	0.0796	-0.0329 *
Robust std err | Robust t-test	0.0603 6.34	0.0479 -4.09	0.0357 2.23	0.0478 -0.69
ln(path size)	2.01	-0.778	-0.202 *	-2.33
Robust std err | Robust t-test	0.359 5.59	0.192 -4.06	0.710 -0.28	0.978 -2.38
	Membership model			
Constant	-1.3			
Robust std err | Robust t-test	0.199 -6.55			
Env. score 1	0.322			
Robust std err | Robust t-test	0.125 2.58			
Env. score 2	-0.156 *			
Robust std err | Robust t-test	0.0924 -1.69			
Env. score 4	-0.119 *			
Robust std err | Robust t-test	0.131 -0.90			
Male	-0.178 *			
Robust std err | Robust t-test	0.253 -0.70			
Variation score 3	-0.00956 *			
Robust std err | Robust t-test	0.0530 -0.18			
Risk score 1	0.0332 *			
Robust std err | Robust t-test	0.0318 1.04			
Sample size	945		590	355
$\mathcal{L}(\beta_0)$	-2823.925		-1763.116	-1060.809
$\mathcal{L}(\hat{\beta})$	-1753.080		-1368.156	-876.736
$\bar{\rho}^2$	0.372		0.220	0.167

* value not significant on a 95-% level

10.3.4 Combined mode and route choice model

Parameter estimates for the combined mode and route choice model are given in Table 10.9 and 10.10. The influence of person-based variables is fixed for car. The alternative specific constants for the other modes are highly significant and very negative. Having no car or only seldom access to one has a positive effect on the other modes, as expected. Having a pt subscription on the other hand has no significant effect. Having a bike has a significant positive effect on choosing it. Of the attitude scores only being in *agreement with CO_2 reduction measures* (Env. score 1) has a significant positive effect on choosing to cycle or walk. The other attitude scores, gender, age, income, household size as well as trip purpose were tested but not significant.

The route choice parameters for car, bike and walk are similar to the separate model estimations. For public transport the estimates are a bit problematic, as number of transfers is very negative and in-vehicle travel time is not significant.

Table 10.9: Mode and route choice model (6 alternatives per mode)

	Car	PT	Bike	Walk
ASC	0 (fixed)	-8.39	-8.21	-3.86
		0.832 -10.09	0.887 -9.26	0.84 -4.59
No car or rarely	0 (fixed)	3.67	3.01	3.11
		0.628 5.84	0.694 4.34	0.702 4.43
PT subscription	0 (fixed)	0.931 *	-0.202 *	-0.235 *
		0.691 1.35	0.774 -0.26	0.539 -0.44
Nr. bikes > 0	0 (fixed)	-0.195 *	1.94	-0.507 *
		0.448 -0.44	0.761 2.55	0.528 -0.96
Env. 1 score	0 (fixed)	0.0755 *	0.489	0.305
		0.117 0.65	0.176 2.78	0.103 2.97
Risk 1 score	0 (fixed)	-0.0852 *	0.102 *	-0.017 *
		0.0454 -1.88	0.0813 1.25	0.0477 -0.36
Var. 3 score	0 (fixed)	-0.00865 *	-0.189 *	-0.0652 *
		0.0843 -0.1	0.11 -1.71	0.0937 -0.7
Access travel time [min]		-0.00634 *		
		0.00463 -1.37		
Egress travel time [min]		-0.119		
		0.0355 -3.36		
In vehicle travel time [min]		-0.0119 *		
		0.0194 -0.61		
Transfer time [min]		0.0147		
		0.00363 4.06		
Number of transfers		-42.3		
		3.54 -11.97		
Share rail stages [-]		-0.579 *		
		0.355 -1.63		
Share bus stages [-]		0 (fixed)		
Share tram stages [-]		2.28		
		0.364 6.26		

Table 10.10: Mode and route choice model (6 alternatives per mode) (cont.)

	Car		PT	Bike		Walk	
ln(f-f time [min])	-3.08						
	0.312	-9.89					
Distance [km]				-0.246		-2.63	
				0.0794	-3.1	0.437	-6.01
Prop. motorway [-]	-1.99						
	0.41	-4.86					
Prop. extra urban [-]	-1.66						
	0.609	-2.73					
Prop. residential [-]	0.868 *			3.01		1.71	
	0.509	1.7		0.613	4.9	0.509	3.35
Prop. track / other [-]	-4.15						
	1.09	-3.82					
Prop. safe bike paths [-]				-3.05			
				0.684	-4.46		
Maximum rise [-]				-11.8		-2.58	
				2.42	-4.89	1.07	-2.4
Maximum fall [-]						-1.1 *	
						1.26	-0.87
Nr. turns / km				0.00213		0.0276 *	
				0.000649	3.28	0.0259	1.07
Nr. left turns / km	-0.949						
	0.122	-7.8					
Nr. right turns / km	-0.827						
	0.118	-7.03					
Nr. traffic lights / km	-0.793						
	0.105	-7.57					
Sample size	2861						
$\mathcal{L}(\beta_0)$	-8245.42						
$\mathcal{L}(\hat{\beta})$	-4995.834						
$\bar{\rho}^2$	0.389						

10.4 Conclusion

The latent class models for cycling as well as walking outperform the MNL models, with relatively high $\bar{\rho}^2$ values of 0.719 and 0.372 respectively. Significant attitude scores were found for both class membership models. Class membership probabilities for the participants show that they can not be distinguished into 2 clear groups. Therefore, determining the class first selecting the one with highest probability and then applying this class model would not yield good results. In the case of the pedestrian models all participants would belong to class 2.

Model fit of the public transport connection choice models are also good and parameter estimates have the appropriate signs. The most important variables are number of transfers, in-vehicle time as well as the proportion of tram stages. Values estimated as part of the combined mode and route choice model on the other hand seem to be extreme.

For the car route choice model, number of turns and traffic signals proved to be a very important variables. Also logarithmic transformation of free-flow travel time improved the model significantly, reducing the number of variables at the same time.

All significant person-based variables in the combined mode choice model have the appropriate sign. The route choice parameters have the same tendencies and therefore also the same problems as the separate models. Comparing bike and walk, distance is punished more for walks and maximum rise more for bikes, which is sensible.

Chapter 11

Searching for parking in GPS data

11.1 Introduction and related work

Parking search is regarded as a significant contributor to congestion in city centres (see e.g., Shoup, 2005). Understanding, modelling and managing it, e.g., with parking policies, are thus important tasks (see e.g., Marsden, 2006; Topp, 1991; Feeney, 1989; Baier et al., 2000; Glazer and Niskanen, 1992; Miller and Everett, 1982; van der Waerden et al., 2009). However, parking search behaviour is complex as it depends on traffic circumstances, trip purpose, individual strategies, the driver's knowledge of the area and more. Hence, search start is latent and even the driver may not know it exactly. To survey and quantify search behaviour is thus difficult (Kipke, 1993; Arnott and Inci, 2005), especially time and distances reported in interviews are biased as estimations are probably influenced by traffic conditions, trip purpose but also the frustration level of drivers.

Survey approaches used so far were laboratory experiments (e.g., Bonsall et al., 1998), stated preference surveys (Axhausen and Polak, 1991; Weis et al., 2011; Golias et al., 2002; van der Waerden et al., 2006, 1993; Widmer and Vrtic, 2004) or field observations such as riding with a searcher (Laurier, 2005) or following a car until it is parked (Wright and Orram, 1976). Modeling approaches range from discrete choice models, numerical models, Possibility Theory to simulations (Gillen, 1977, 1978; Hensher and King, 2001; Arnott et al., 1991; Arnott and Rowse, 1999; Anderson and de Palma, 2004; Benenson et al., 2008; Gallo et al., 2011; Thompson and Richardson, 1998; Dieussaert et al., 2009; Kaplan and Bekhor, 2011; Axhausen, 1988; Young, 1986; Young and Thompson, 1987; Maley and

Weinberger, 2011; van der Waerden et al., 1998; Young and Weng, 2005; van der Waerden et al., 2002).

A very rich data source to complement these surveys are GPS data. It is mainly collected to get more complete and accurate travel diaries (e.g., Yalamanchili et al., 1999; Draijer et al., 2000; Wolf et al., 2001a; de Jong and Mensonides, 2003; Auld et al., 2009; Marchal et al., 2011; Oliveira et al., 2011b; Rieser-Schüssler et al., 2011). Lately, GPS data are also used to observe more specific travel behaviour; e.g. Moiseeva and Timmermans (2010) focus on activity patterns in retail areas. The work most related to this Chapter is Kaplan and Bekhor (2011) who investigate the joint decision of parking type and parking-search route. To observe the actual route taken they intend to use GPS data collected in Tel Aviv. Using GPS data to observe parking search data has the advantage over interviews and questionnaires that time and distance calculations are objective and not estimated. Very recently, Karlin-Resnick et al. (2016) present results of a controlled GPS sample where 70 tracks were recorded and the test drivers stated explicitly when they started searching.

This chapter's goals are to develop a parking search analysis module for GPS travel data and to provide a descriptive analysis of the parking search found in the Swiss GPS data for the cities Zurich and Geneva (see Section 11.2.1) by applying this module.

Parking search traffic is a widely discussed and very political issue also in Zurich. Planungsbüro Jud (2010) shows that on Saturdays in the inner city of Zurich parking occupancy is around 97 %. Kipke (1993) indicates that searching for parking gets potentially a problem for occupancies higher than 95 %. Therefore, it is reasonable to assume that parking search can be problematic in Zurich.

The latent but very important *search starting point* is not known in GPS data. The distinction between the search and the rest of the journey is therefore not straight-forward, different from, e.g., stated preference experiments. Thus, as detailed in Section 11.2.2 a spatial proxy is developed, and indications for the start point are given.

The descriptive analysis provides numbers for driving times and distances in a certain area around the parking location. Furthermore, walking times and distances from parking to activity location are given. Route choice in relation to shortest path and loops are analyzed. An initial analysis of the parking type, that is on-street or garage parking, is included.

The results of this analysis could be used for calibration of parking simulations as e.g. described in Horni et al. (2012).

The remainder of the chapter is structured as follows. Section 11.2 describes definitions and methods used. Section 11.3 discuss the findings and in 11.4 conclusions and future work are described.

11.2 Method

11.2.1 GPS data and processing

For this analysis, the GPS data set introduced in Chapter 5 is used. To analyse parking search relevant characteristics, the two subsets of residents of Zurich and of Geneva are used. The first is concentrated on Northeast-Switzerland and the other one on West-Switzerland (Figures 11.1(a) and 11.1(c)). Centre areas for Zurich and Geneva are defined with a diameter of 3 kilometers as shown in Figures 11.1(b) and 11.1(d). Further analysis is focused on Zurich and its twelve districts, the location of those is depicted in Figure 11.2 and some descriptors are summarised in Table 11.1. Additionally, GPS locations of public on-street parking spaces and garages are available for the city of Zurich.

Using person-based as opposed to car-based GPS data complicates the post-processing, but it has the advantage that not only the car stages but also the subsequent walking stages or activities can be detected. For the processing the open-source POSDAP (2012) routines developed at the institute are used. In short, the GPS traces are first cleaned and smoothed to ensure reasonable speed and acceleration calculations. Later, the traces are split into stages and stop points, that is mode transfer points and activities. Then, using a fuzzy logic approach, all stages are assigned a mode.

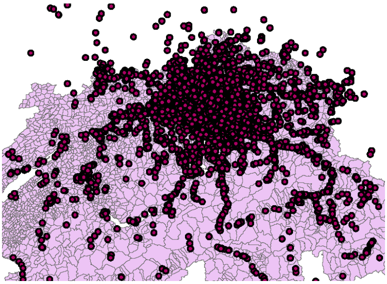
For the parking search analysis only car stages longer than 10 minutes are considered. This decreases the probability of erratic signals being interpreted as car stages. Car stages are further categorised in:

(i) car stages followed by an activity shorter than 15 minutes and then by a stage faster than walk,
(ii) car stages followed by an activity shorter than 15 minutes and then followed by walk,
(iii) car stages followed by an activity of at least 15 minutes.

The car stages of category (i) are not considered for further analysis as signal gaps longer than 3 minutes are interpreted as stop points possibly due to tunnel usage. Another possibility is that the short activities are mode transfer points and the detected car stage might be a bus or a rail stage. For category (ii) the stop point after the walk stage is assumed to be the activity that induced car driving. This is a first approximation as trip purposes are not known and the trip detection module was not available at the time of analysis. For category (iii) the immediate stop point is assumed to be the

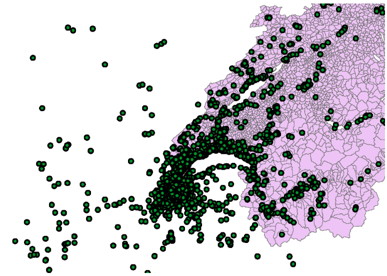
Figure 11.1: GPS data sets and centre definition for Zurich and Geneva

(a) Parking spaces Northeast-Switzerland data set



(b) Zurich centre

(c) Parking spaces West-Switzerland data set

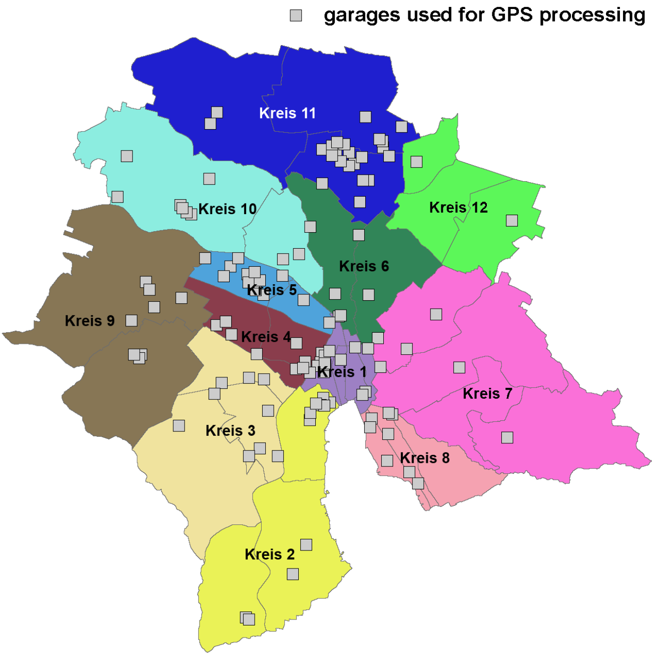


(d) Geneva centre

main activity. As a consequence, the walk stage to this activity is assumed to be zero meters and minutes.

The last GPS point of a detected car stage is used as an approximation of the parking space location. Using the available public parking location data, parking types are assigned to each parking space. Spaces that are closer to a public garage location than to an on-street parking space are assigned garage, parking spaces that have a garage within 50 meters are classified as uncertain and the rest is assigned on-street parking. It is important to note that on-street parking also includes private parking. The activity location is approximated by the median of its x and y coordinates, which is mostly

Figure 11.2: 12 districts (Kreis) of Zurich with garages used for GPS analysis.

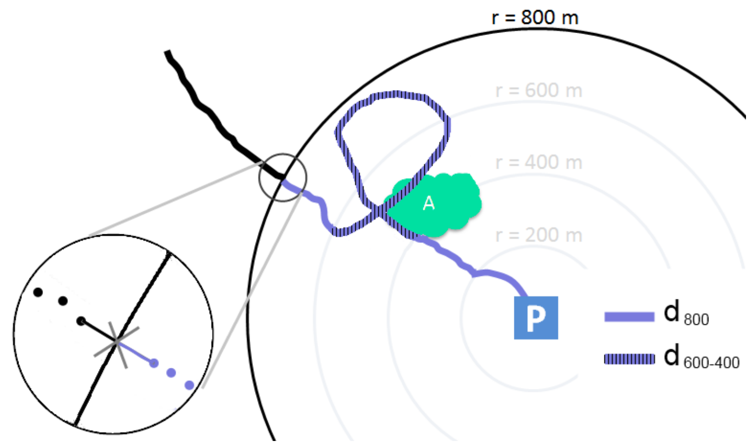


reasonable but does e.g. not work for long signal gaps that start and end at different locations.

11.2.2 Parking search path and strategies

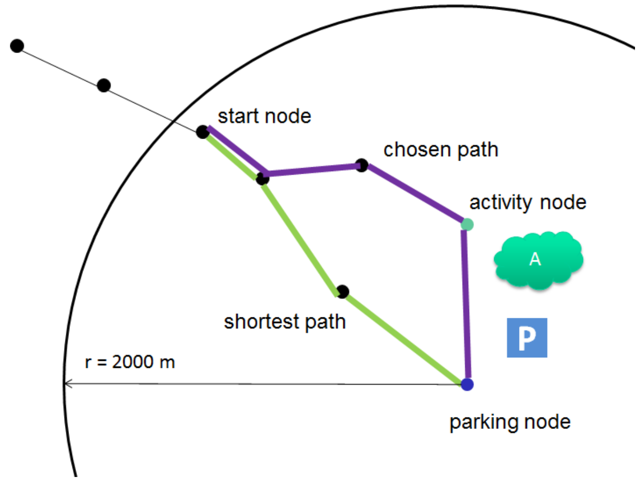
Several definitions for parking search start point, and consequently, the parking search path exist. Kipke (1993) suggests that the search starts as soon as the activity location is passed. This definition is problematic as this location does not have to be passed during parking search, e.g., if an activity location in a pedestrian only area or the driver finds a parking space beforehand. The second uncertainty is how well this activity location is definable e.g. eating at a friends place is easier to capture than shopping in the inner city. Birkner (1995) suggests that the search starts as soon as the first parking space is passed that would have been accepted if free. Using this definition, it is not possible to extract a start point from raw GPS data, as not only the drivers thoughts are unknown but also traffic conditions or parking occupancy, influencing the search start, are usually not available.

Figure 11.3: Path segmentation



Unfortunately neither of these definitions can be used to extract the parking search start point and the actual search path. Therefore, we decided to use the path after entering an 800 meter radius around the parking space as a measure to analyse parking search (d_{800} in Figure 11.3). Analogous, the time after entering an 800 meter radius t_{800} is used as a means to analyse parking search time. It is very likely that this simple measure, representing an upper bound of search effort, includes the search path. The underlying assumptions are that walking distances acceptable for the majority of car drivers in Zurich are below 600 meters (Planungsbüro Jud, 1990) and that searching for parking usually takes place between the actual found and the aspired parking space. The radius criteria is also used to split the path into segments that start when the driver enters a circle around the parking space and end when she enters the next smaller circle (e.g.

Figure 11.4: Chosen and shortest path



$d_{600-400}$ as illustrated in Figure 11.3). These segments are used to analyse the progress of the search path. As distances between two successive GPS points are not negligible, the cutting point with a circle is interpolated and the distances are corrected accordingly to ensure comparability of all segments and paths.

The distance difference between the chosen and the shortest path to the parking space is, as mentioned by Birkner (1995), another possible indicator for parking search traffic but it is not the search effort itself. To calculate this difference the GPS points of the car stage are first map-matched (Schüssler and Axhausen, 2009b). The last node of the resulting path is defined to be the parking node. The start node is defined to be the first node within 2 kilometers around the parking node (see Figure 11.4). This start and parking node are then used for shortest path calculations using Dijkstra's algorithm with distance as cost. Using the difference between the paths of the complete journey would lead to differences due to the chosen route into the city. But we are only interested in the last part of the journey that could be influenced by parking choice.

Having extracted the chosen path it is used to get an indication for the underlying search strategy. (Polak and Axhausen, 1990) identified seven strategies that are briefly described here:

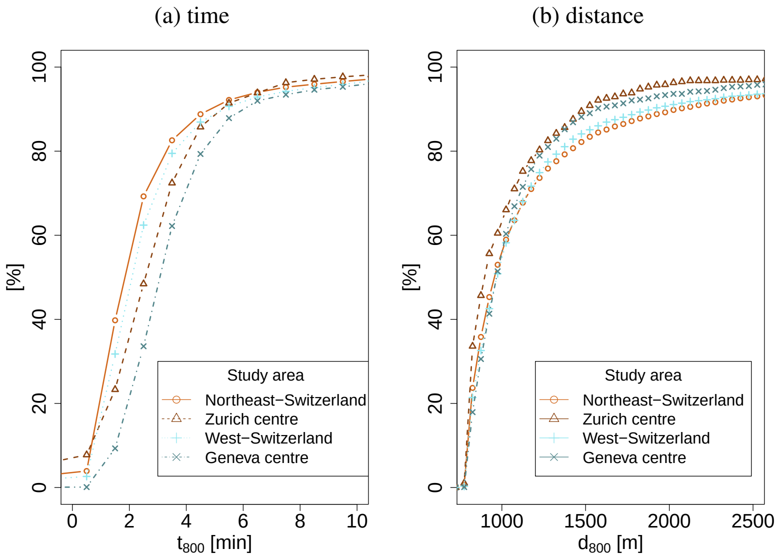
(i) Drivers drive directly to an almost guaranteed 'inside tip' parking place that is not officially for them (e.g. customer parking spaces).
(ii) Drivers know a fixed number of opportunities which almost always lead to no search time (e.g. garage, facilities around the core) and drivers are willing to accept long walking distances.
(iii) Drivers drive in direction of a garage but use on-street facilities if available.
(iv) Drivers have a fixed sequence of on-street and cheaper off-street opportunities and accept long walks.
(v) Drivers adapt search according to trip purpose and duration, illegal parking is an option. Search time might be long.
(vi) Drivers circle around their destination and long searches are accepted to ensure short walks.
(vii) Drivers accept illegal parking for short stays.

Strategies (i) and (vii) are undetectable in GPS data as illegal or customer parking spaces are most likely near legal public parking spaces and the resulting short search times can not be assigned to these strategies, as short searches also result from private parkers or parking during unproblematic times where no search strategy is needed. Only strategies (ii) and (iii) use garages, information that was extracted from GPS travel and parking location data. Strategy (iii) can also lead to on-street parking and can therefore easily be misinterpreted - GPS data of several weeks might help identify such drivers if parking spaces are often near or in garages. Strategies (iv), (v) and (vi) are all on-street parkers with possibly long search times and are therefore hard to distinguish. Strategy (iv) might be extracted if several weeks of data is available as the drive patterns stay the same. Driving in circles (vi), is detected by inspecting the map-matched chosen path for network nodes traveled several times. GPS data can consequently hint at strategies (ii), (ii) and (vi) which is investigated in the next section.

11.3 Findings

Results are provided for Northeast- and West-Switzerland. As Geneva is more densely populated than Zurich the hypothesis is that searches are longer in Geneva. This was confirmed for search times (t_{800}) but not for search distances (d_{800}) as can be seen in Figure 11.5. Search times were also higher in the centres, interestingly this does not hold for distances. This is influenced by lower speeds in the centres, but maybe also points to different search strategies.

Figure 11.5: Time and distance driven within a radius of 800 meters around parking space for Geneva and Zurich.



For the remainder of this chapter analysis is focused on car stages ending in the city of Zurich. Two districts with different characteristics are highlighted. Kreis 1 is the historic inner city, that stretches from the lake to the main station including shopping streets, commercial and very expensive residential buildings. It has many more employees than residents and the share of parking spaces is lowest with 0.13 spaces per resident and employee (Table 11.1). Kreis 9 on the other hand has the highest share of parking spaces (0.45). It is much larger and it is a commercial but also residential district.

In total, 4086 car stages longer than 10 minutes are detected. Approximately 20 % of those are filtered as they are not followed by a walk or a long activity as shown in Figure 11.6. Still for each district at least 130 cases are left after filtering. The figure also shows that the share of car stages followed by a long activity, that is where a parking spaces was found very close to the activity location, is highest in Kreis 9 and lowest in Kreis 1 which corresponds to the ratio of parking space to residents and employees.

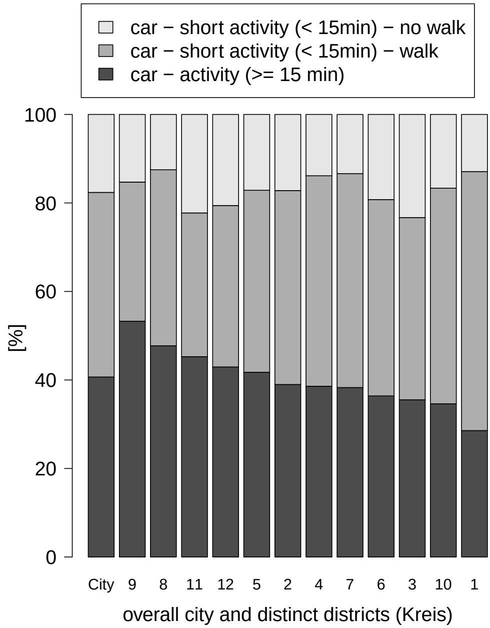
Table 11.1: Zurich city data by district (Stadt Zürich Präsidialdepartement, Statistik Stadt Zürich, 2011)

District (Kreis)	**Residents**	**Area** [ha]	**Parking spaces**	$\frac{\textbf{Parking}}{\textbf{(res. + empl.)}}$	**Car cases**
1	**5563**	**180**	9087	**0.13**	**294**
2	29878	1106	24931	0.39	372
3	46699	865	25805	0.32	442
4	27429	280	18005	0.31	368
5	12764	209	16351	0.34	321
6	31464	511	16838	0.35	239
7	35447	1502	24833	0.42	269
8	15518	481	14899	0.39	176
9	**48494**	**1207**	**39504**	**0.45**	**458**
10	36879	907	20705	0.41	312
11	65796	1343	42666	0.40	665
12	29537	597	13374	0.39	170
City	**385468**	**9189**	**266998**	**0.36**	**4086**

11.3.1 Driving times and distances

For the applied analysis method, the minimum driving distance is of course 800 meters. The additional driving distance is influenced by the network,

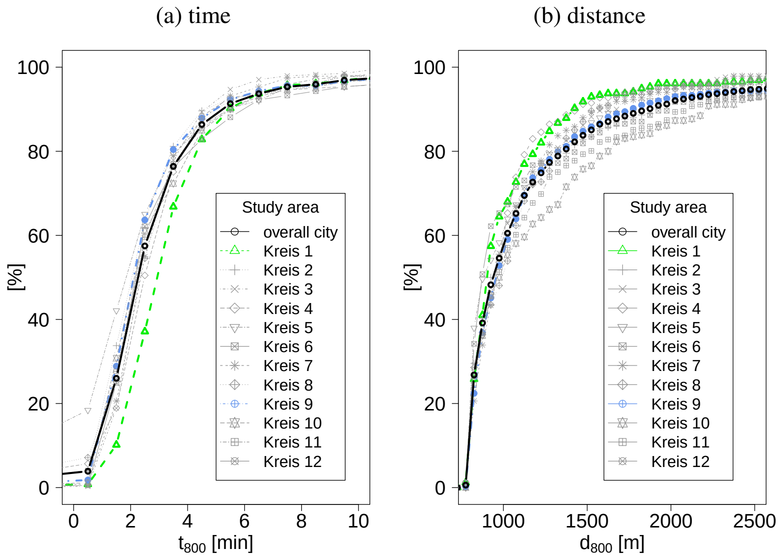
Figure 11.6: Categorisation of detected car stages.



that is by the shortest possible path, considering one way streets, speed limits but also by the drivers knowledge of the city. The driving times are additionally influenced by traffic conditions. Driving times in the inner city (Kreis 1) are highest but driving distances are shorter (Figure 11.7), as this area is more congested and speeds are lower. In general driving times are less than 4 minutes for 80 % of cases in the overall city; distances driven range from 1100 to 1400 meters for 80 % of cases, which indicates that parking search substantially varies for districts. Possible remaining processing errors such as misinterpretation of bus or rail stages wrongly identified as cars do not include search paths and thus lower the distance and time estimates. Consequently the share of low estimates is too high and has to be corrected if used as upper limits in parking models.

In the city, the distance difference of the chosen and shortest path is

Figure 11.7: Time and distance driven after entering an 800 m radius around parking space.

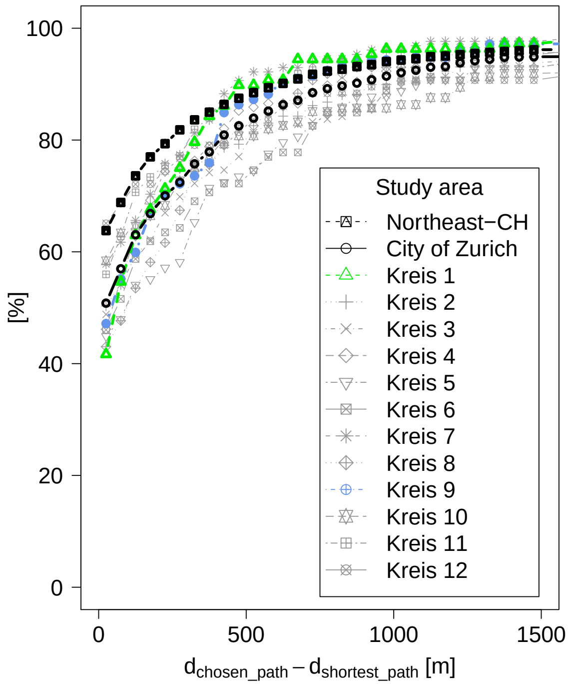


below 500 meters for more than 80 % of cases (Figure 11.8). As expected, for the complete data set the distance differences are lower. It is mostly due to the fact that over 60 % of stages in the Northeast-Switzerland data set were shortest paths. In the city, this share is considerably lower but still around 50 %.

11.3.2 Walking times and distances

Times and distances walked after parking are depicted in Figure 11.9 for all districts and the overall city. Walking stages are potentially underestimated as they are ended by stop points of 3 minutes which might be a short stop on the way to the actually planned main activity. In Kreis 9 over 65 % of car stages end at the activity. For another 25 % of observations the subsequent walk is less than 5 minutes or less than 400 meters respectively. In Kreis 1, considerably less but still 40 % park at the activity location

Figure 11.8: Difference chosen and shortest path on the last 2 km to the parking space.

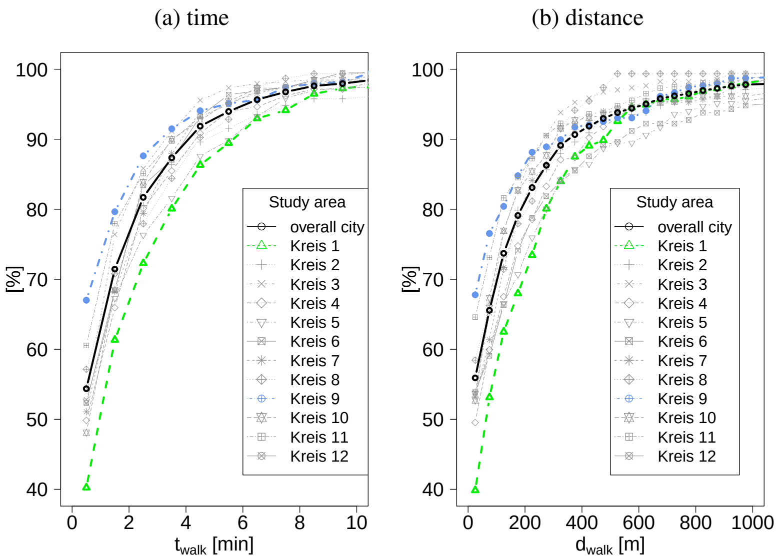


and for another 40 % of observations walk is less than 400 meters. The difference of the 90th percentile is around 2 minutes showing that parking success substantially depends on location.

11.3.3 Speed distribution

The distribution of the average speeds in path segments (illustrated in Figure 11.3) are depicted in Figure 11.10 for Kreis 1 and 9, which are chosen as previously. Kreis 1 is interesting as it is the district in which parking searches are longest and Kreis 9 has a similar speed distribution as the city overall. As expected speeds in Kreis 1 are generally lower than

Figure 11.9: Time and distance walked for all districts (Kreis) of Zurich.

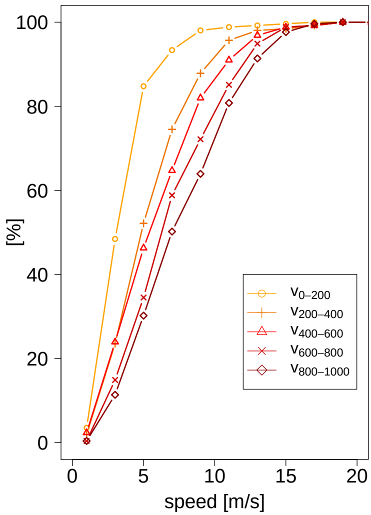


in the overall city and Kreis 9. In both districts, the average speed on the last 200 meters are considerably lower than 1000 meters away from parking space. For all districts it was found that aggregated speeds strongly decrease when approaching the parking space. The differences between the segment speeds are very small for Kreis 9 but for Kreis 1 they are more pronounced.

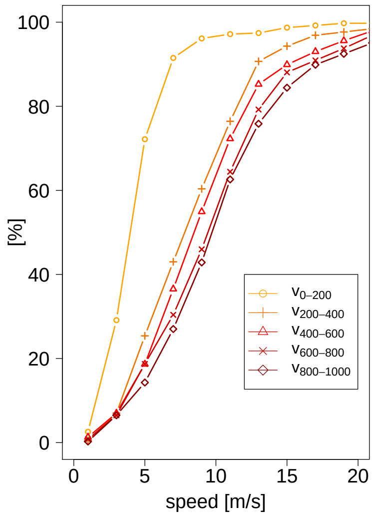
This is probably due to two influences: first, traffic slows down when approaching city centre. But as parking spaces are spread all over Kreis 1 (Figure 11.10(c)), speeds of the same path segment are not speeds of the same area, the second possible influence is therefore drivers slowing down because they start searching. Consequently speed distributions can point to the start of parking search, e.g., parking search starts earlier in Kreis 1 than in Kreis 9.

Figure 11.10: Speed distribution for path segments and parking spaces Kreis 1

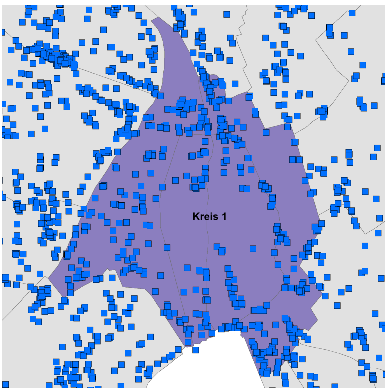
(a) Kreis 1



(b) Kreis 9



(c) Distribution of parking spaces in Kreis 1.



11.3.4 Search path

To identify drivers circling while searching as described in the strategy (vi) of Section 11.2.2, loops of the extracted chosen path are counted. Less than 10 % of the paths contain one loop. And only very few paths, that is less than 1 %, contain more than one loop. This indicates that circling is not necessary or favoured by drivers. The highest share of potential circling drivers are found in Kreis 5 a former industrial district, least in Kreis 9, this maybe due to more private parkers there.

Search strategies considering public garage parking (ii and iii) are used in 5 - 15 % of cases (Figure 11.11). The public garages used for classification are shown in Figure 11.2. All districts have garages still no surveyee parked in a garage in Kreis 3, a residential district. This this might be because garages there are public but mostly for very specific trip purposes, potentially not performed by the respondents in the survey period. In the entire city for 98 persons more than 3 parking activities are identified. Of those 27 used garages and on-street parking and and only 1 person used garages in all cases.

Interestingly, garage and on-street parking strategies lead to very similar distributions of walking distances (Figure 11.12).

11.3.5 Dynamics

As trip load curves commonly show clear peaks, assuming peaks in parking search effort is natural. However, Figures 11.13(a) and 11.13(b) show that there are no pronounced peak days or hours for parking search traffic in Zurich. Peaks might level out as the analysis is performed on the whole city, which is necessary here due to small sample size. Furthermore, parking assumedly has to be seen as a cumulative phenomenon smoothing the peaks in demand. In other words, parking search is not necessarily easier for trip off-peak times than for trip peak hours. Having no peaks can also mean, that either parking demand is very low or that it is always in saturation range.

Figure 11.11: Parking type - on-street vs. garage.

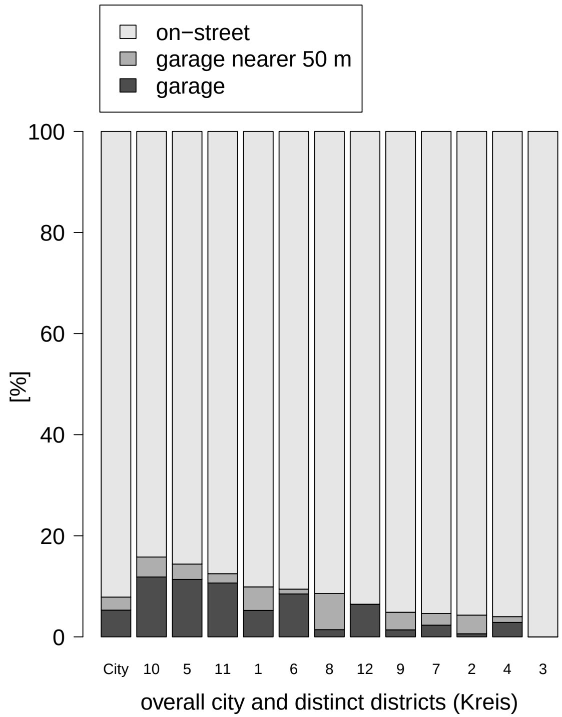


Figure 11.12: Walking distances for garage and on-street parking.

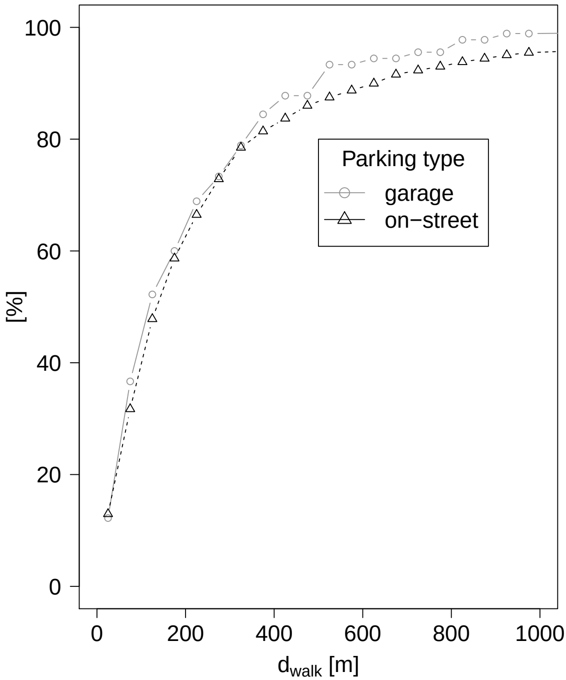
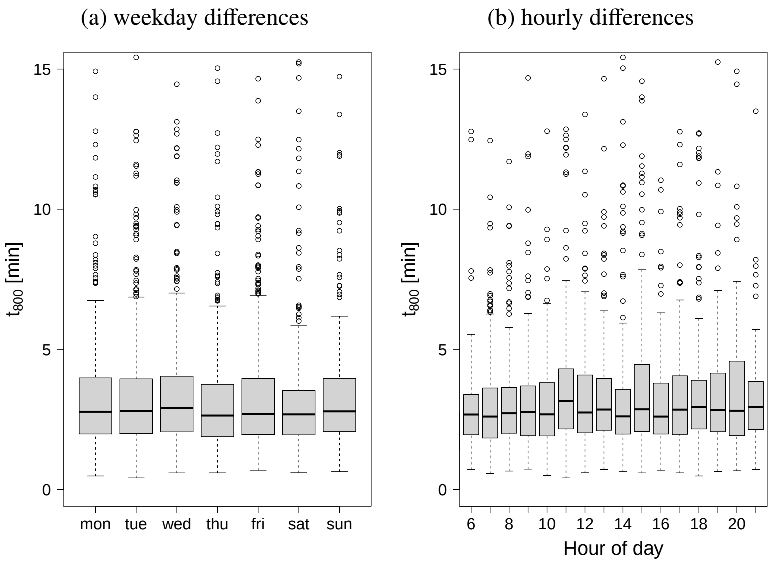


Figure 11.13: Driving times within 800 m of parking for Zurich city.



11.4 Conclusion and outlook

Quantification of parking search effort is difficult and results found are controversial as a large range of values is found for different studies and locations (see e.g., Shoup, 2005). Clearly, situation strongly differs from city to city. However, differences are also due to diverging definitions, latency of parking search and bias in reporting parking search effort. Usually, parking search is regarded as important. However, parking search for Zurich seems to be undramatic based on indicators computed in this study. The time driven with an 800 meters radius around parking space is less than 4 minutes for 80 % of cases in the overall city; distances driven range from 1100 to 1400 meters for 80 % of cases, which indicates that parking search substantially varies by districts. For a plausibility check, number of parking spaces per resident and employee per district (overall city: 0.36) and share of parking spaces on private ground (overall city: 81 %) have been computed (see Table 11.1). Both numbers are relatively high inducing low parking search being inline with the computed results.

However, this is a first explorative analysis with the main goal to develop and improve parking search extraction methods. The data set used here does only contain raw data, no accuracy and accelerometer data—important for high-quality processing—are available. Results have to be interpreted as indications. It has to be noted that driving times are not corrected for congestion as no such information is available. Further, distance and driving times are potentially underestimated due to wrongly detected public transport stages. Walk times are also potentially underestimated by misinterpretation of the following main activity.

The findings concerning parking choice can be compared to stated preference studies conducted in Zurich. Results can be used to calibrate parking search models e.g., in the agent-based transport simulation MATSim or in the authors' simulation (Horni et al., 2012).

Assuming that parking behaviour heavily depends on type of activity, the trip detection module could be run for the data set used here and the annotated data collected in Zurich (Chapter 3) could be analysed as well. Processing routines might also reveal pseudo- or intermediate activities (e.g., window shopping) while walking to the actually planned main activity. It can be assumed that in questionnaires only this main activity is reported while in GPS data intermediate activities are recognisable, leading to a biased estimation of walk distances (underestimation).

As private parking is usually associated with only little or no search, an important future task is to distinguish the analysis between private and public parking.

Loop analysis is in the first instance based on network node counting. As this is highly dependent on network resolution, analysis should in the future be refined to areas (instead of using single network nodes). In other words, loops are counted if a certain area is crossed multiple times.

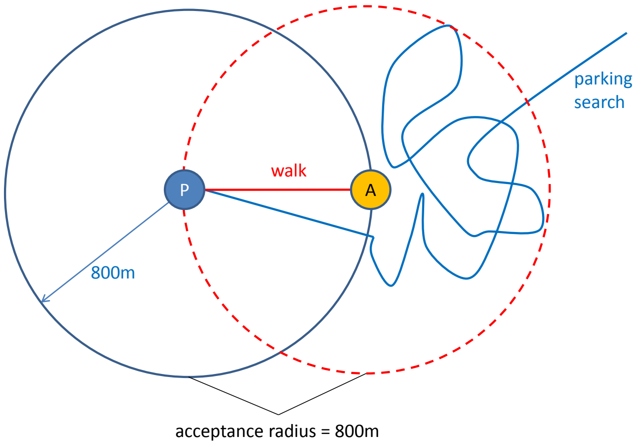
To reveal if higher travel times (and potentially detours with respect to free-flow minimal path) are due to search or due to traffic conditions a time-dependent analysis for persons without activity or parking in the respective area should be performed, in particular for inner city areas, and compared with the speed distributions given in Section 11.3.3.

There is a discrepancy between average search effort found in GPS data and subjective parking effort estimation reported in personal communication, where higher search times are expected. Thus, an additional analysis should focus on the high efforts including outliers.

Considering a circle around parking location, as done in this work, but also around the activity, harbors the danger of missing part of the parking search in situations where a long walk is followed (see Figure 11.14). Combining analysis of a region around both the parking location and the activity should be tested.

Indications for the latent but very important *actual search starting point* are revealed. In Karlin-Resnick et al. (2016) none of the recorded searches started more than 350 meters away from parking space, suggesting that our 800 meter buffer is well on the safe side and includes parking search. They state, that determination of the actual starting point is on-going research, but they suggest rules to identify if parking search takes place considering the following variables: "Excess travel distance within 400 meters of the final parking space, the ratio of actual to shortest-path travel distance within 400 meters of the parking space, repeated roadway segments toward the end of a trace, and out-of-direction travel at the end of a trace." These rules could be applied to the available GPS data sets.

Figure 11.14: Circle around parking or activity location: parking search right to activity location is missed in this case.



Chapter 12

Discussion

One of the main goal of GPS processing in transportation is to obtain complete and precise travel diaries, especially in terms of trip times and short trips that tend to be forgotten. When looking at the tracks of my own movement, this goal seems to be easily achievable, as my cycling route to work was usually well captured both with the dedicated device (MobiTest) as well as with my smartphone (Samsung S3). But it also happens quite often that the trip to the mensa, a very short trip between buildings on the campus, was missing. This did not bother me enough, as I usually knew the reason for it, such as forgetting the device, low battery or recording of too few points. It shows that there is still a lot of uncertainty when dealing with GPS tracks. Neither with paper and pen diaries nor with GPS-based diaries you can be confident that all travel is captured. One major advantage with GPS is that at least it is certain that the recorded route (fragments) did take place, even if some participants deny it until you explain them their data on the map.

Researchers are used to the data and its visualisation and are well aware of potential errors, automatically filtering or adding to the data in their minds. Therefore, it is easily forgotten that it is not for granted that study participants can read a map and interpret traces, especially in the more complicated cases where we actually need their corrections most. Response burden therefore not only depends on the accuracy of the diaries but also on the quality of the prompted recall tools.

Different ways to collect ground truth data are presented in the thesis. A paper and pen diary (Section 4.2) as well as two prompted recall applications with pre filled diary information, one a web-based version (Section 3.1.3) and the other one a smartphone application (Section 4.1). The paper diaries were filled in well, a clear disadvantage is, that data has to be digitalized, another potential source for errors (or corrections). In the case of the first PEACOX field trial, the final travel diaries were obtained by manually comparing the automatically generated ones with the paper and pen diary. This was possible for the low number of participants but would not be practical for a larger scale survey. For both prompted recall instruments, stages and activities were presented as a list as well as on a map. Colour coding and/or icons were used in order to help participants understand the diaries and connect the list of items with the map. The web-based interface has the advantage (if only used on desktop computers) of a large map that can be shown, compared to that the map on smartphones is small but on the other hand it can be filled in anywhere. To use that advantage fully, diaries would need to be prepared immediately which was not the case, as processing was done over night. Further, it was shown that many participants did not fill in the diaries immediately but mostly when they were reminded or even just at the end of the eight week survey period. Considering this reluctance and also evidenced by data it can be assumed that at least some of the diaries were confirmed without actually correcting them. A more in-depth analysis of how true the ground truth data really is in the line of Stopher et al. (2015) would be interesting. They even propose and tested a new way of collecting ground truth data by usage of a life-logging camera, that takes a picture every 30 seconds. In that case privacy is even more of an issue than with GPS as strangers might be on the pictures. But it is also a new source for interesting future research, e.g. on image processing to extract mode and trip purposes, but also new features that can not be observed with GPS only, like crowding, traffic, scenery or companions.

The amount of collected data is already an issue with GPS and accelerometer measurements. E.g., for the PEACOX project study participants were required to have mobile abos with data flat rates or at least 1 GB included. Accelerometer data had to be backed up every day, otherwise the available database would not perform any more. These are not necessarily research issues, but practical problems concerning data storage and processing, that have to be considered much more, when planning surveys with more participants and longer survey periods, which is one of the main goals of smartphone-based travel surveys.

In Section 4.5.2 travel diaries generated from the same routes captured with two devices (Smartphone and dedicated GPS device) are shown to be different in many cases, that is movement is captured with one device but not with the other, which suggests that some movement was not captured at all. Figure 4.7(b) also shows how sparse the useful data is, even though a huge amount is collected.

The PEACOX project also enabled us to test the processing routines developed on Swiss data in other European cities, namely Vienna and Dublin. The achieved accuracies were not as high as expected, mode detection was correct in around 70 % of cases compared to over 85 % reported in Part II. Activity detection was good in approx. 80 % of cases which is as expected, uncertainties remain as of how well participants corrected their diaries. Considering discrepancy of performance between data sets it would be interesting to compare classification performance or better travel diary generation of different research groups on a common data set. Most processing frameworks are not open source and data sets often cannot be shared due to privacy reasons. Currently, also the random forest classifiers of POSDAP cannot be shared, as data is stored within them. In the future, in order to share these, the random forest should be stored without data. Other than that, potential future work concerning the processing framework are inclusion of indoor movement using additional (smartphone) sensors like wifi positions, pressure (e.g. detecting lifts) and temperature (e.g. detecting going inside). For smartphone surveys, on the phone pre-processing has to be considered in order to reduce data traffic. It should be investigated what data resolution is actually needed for the problem under investigation, as reduced collection frequency saves data as well as battery. Battery saving strategies are also a research issue, even if better batteries can be expected, smartphones are used for more and more things, therefore battery consumption of other apps possibly increases as well. Up to now the framework was used to process whole days. Real time processing would be beneficial, e.g., in order to prompt for immediate corrections.

Bibliography

Ajzen, I. (1991) The theory of planned behavior, *Organizational Behavior and Human Decision Processes*, **50** (2) 179–211.

Alam, S. and A. McNabola (2012) Peacox – door to door emissions models, *Deliverable*, **D3.1**, PEACOX project, Trinity College Dublin, Dublin.

Ali, S. S. M., N. Joshi, B. George and L. Vanajakshi (2012) Application of random forest algorithm to classify vehicles detected by a multiple inductive loop system, paper presented at the *15th IEEE International Conference on Intelligent Transportation Systems (ITSC)*, Anchorage, September 2012.

Amt für Verkehr, Volkswirtschaftsdirektion Kanton Zürich (2011) Gesamtverkehrsmodell, webpage, March 2011, URL `http://www.afv.zh.ch/internet/volkswirtschaftsdirektion/afv/de/home.html`.

Anable, J. (2005) "Complacent Car Addicts" or "Aspiring Environmentalists"? Identifying travel behaviour segments using attitude theory, *Transport Policy*, **12** (1) 65–78.

Anderson, S. P. and A. de Palma (2004) The economics of pricing parking, *Journal of Urban Economics*, **55**, 1–20.

Arnott, R., A. de Palma and R. Lindsey (1991) A temporal and spatial equilibrium analysis of commuter parking, *Journal of Public Economics*, **45**, 301–335.

Arnott, R. and E. Inci (2005) An Integrated Model of Downtown Parking and Traffic Congestion, *Working Paper*, **11118**, National Bureau of Economic Research, Cambridge, February 2005.

Arnott, R. and J. Rowse (1999) Modeling parking, *Journal of Urban Economics*, **45** (1) 97–124.

Artukovic, M., B. Abunaim, L. Montini, E. Bothos, T. Tvrzsky and A. McNabola (2013) Peacox – server and client implementation, *Deliverable*, **D6.2**, PEACOX project, Fluidtime Data Services GmbH, Vienna.

Asakura, Y. and E. Hato (2004) Tracking survey for individual travel behaviour using mobile communication instruments, *Transportation Research Part C: Emerging Technologies*, **12**, 273–291.

Auld, J., C. Williams, A. K. Mohammadian and P. Nelson (2009) An automated GPS-based prompted recall survey with learning algorithms, *Transportation Letters*, **1** (1) 59–79.

Axhausen, K. W. (1988) Eine ereignisorientierte Simulation von Aktivitätenketten zur Parkstandswahl, Ph.D. Thesis, University of Karlsruhe, Karlsruhe.

Axhausen, K. W. and J. W. Polak (1991) Choice of parking: Stated preference approach, *Transportation*, **18** (1) 59–81.

Axhausen, K. W. and C. Weis (2010) Predicting response rate: A natural experiment, *Survey Practice*, **3** (2).

Baier, R., C. Hebel, C. Peter and K. H. Schäfer (2000) Gesamtwirkungsanalyse zur Parkraumbewirtschaftung, *Research Report*, Bundesamt für Strassenwesen (BAST), Büro für Stadt- und Verkehrsplanung Dr.Ing. Reinhold Baier GmbH, Aachen.

Becker, H., F. Ciari, M. Brignoni and K. W. Axhausen (2015) Impacts of a new free-floating carsharing system traced with a smartphone app, paper presented at the *14th International Conference on Travel Behaviour Research (IATBR)*, Windsor, July 2015.

Bello, D. C. and M. J. Etzel (1985) The role of novelty in the pleasure travel experience, *Journal of Travel Research*, **24** (1) 20–26.

Ben Abdesslem, F., A. Phillips and T. Henderson (2009) Less is more: Energy-efficient mobile sensing with senseless, paper presented at the *MobiHeld '09 the 1st ACM workshop on Networking, systems, and applications for mobile handhelds*, Barcelona, August 2009.

Ben-Akiva, M. E. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.) *Handbook of Transportation Science*, chap. 2, 5–34, Kluwer, Dordrecht.

Benenson, I., K. Martens and S. Birfir (2008) PARKAGENT: an agent-based model of parking in the city, *Computers, Environment and Urban Systems*, **32** (6) 431–439.

Bierlaire, M. (2003) BIOGEME: A free package for the estimation of discrete choice models, paper presented at the *3rd Swiss Transport Research Conference*, Ascona, March 2003.

Bierlaire, M., E. Frejinger and J. Stojanovic (2006) A latent route choice model in Switzerland, paper presented at the *European Transport Conference*, Strasbourg, September 2006.

Birkner, T. (1995) Innerstädtischer Parksuchverkehr: Eine unbekannte Grösse?, *Straßenverkehrstechnik*, **95** (7) 323–327.

Bohte, W. and K. Maat (2008) Deriving and validating trip destinations and modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, paper presented at the *8th International Conference on Survey Methods in Transport*, Annecy, May 2008.

Bohte, W. and K. Maat (2009) Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, *Transportation Research Part C: Emerging Technologies*, **17** (3) 285–297.

Bonsall, P. W., I. A. Palmer and P. Balmforth (1998) Parkit - a simulated world for parking choice research, paper presented at the *8th World Conference on Transportation Research*, Antwerp, July 1998.

Bothos, E., D. Apostolou and G. Mentzas (2012) Recommending eco-friendly route plans, paper presented at the *Workshop on Recommendation Technologies for Lifestyle Change*, Dublin, September 2012.

Bovy, P. H. L., S. Bekhor and C. G. Prato (2008) The factor of revisited path size: Alternative derivation, *Transportation Research Record*, **2076**, 132–140.

Bovy, P. H. L. and S. Fiorenzo-Catalano (2007) Stochastic route choice set generation: Behavioral and probabilistic foundations, *Transportmetrica*, **3** (3) 173–189.

Breiman, L. (2001) Random forests, *Machine Learning*, **45** (1) 5–32.

Breiman, L. and A. Cutler (2013) Random forests, webpage, URL `http://stat-www.berkeley.edu/users/breiman/RandomForests/cc_home.htm`. Accessed on 29/07/2013.

Bricka, S., J. Zmud, J. Wolf and J. Freedman (2009) Household travel surveys with GPS: An experiment, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

Broach, J. and J. Dill (2015) Where do people prefer to walk? pedestrian route choice model developed using gps data, presentation, Active Living Research Conference, San Diego, February 2015.

Broach, J., J. Dill and J. P. Gliebe (2012) Where do cyclists ride? a route choice model developed with revealed preference GPS data, *Transportation Research Part A: Policy and Practice*, **46** (10) 1730–1740.

Cascetta, E., A. Nuzzola, F. Russo and A. Vitetta (1996) A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, in J. B. Lesort (ed.) *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, 697–711, Pergamon Press, Oxford.

Cellina, F., V. de Luca, N. Kovacs, A. E. Rizzoli, R. Rudel, D. Bucher, P. Weiser and M. Raubal (2015) Goeco! a smartphone application leveraging eco-feedback and gamification techniques to nudge sustainable personal mobility styles, poster presentation, 2nd annual conference SCCER mobility, Zurich, 2015.

Chung, E.-H. and A. Shalaby (2005) A trip bases reconstruction tool for GPS-based personal travel surveys, *Transportation Planning and Technology*, **28** (5) 381–401.

Cottrill, C. D., F. C. Pereira, F. Zhao, I. Dias, H. B. Lim, M. E. Ben-Akiva and P. C. Zegras (2013) Future mobility survey: Experience in developing a smartphone-based travel survey in singapore, *Transportation Research Record*, **2354**, 95–67.

Daly, A. J. and S. Hess (2011) Simple approaches for random utility modelling with panel data, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

de Jong, R. and W. Mensonides (2003) Wearable GPS device as a data collection method for travel research, *Working Paper*, **ITS-WP-03-02**, Institute of Transport Studies, University of Sydney, Sydney.

Deng, Z. and M. Ji (2010) Deriving rules for trip purpose identification from GPS travel survey data and land use data: A machine learning approach, paper presented at the *7th International Conference on Traffic and Transportation Studies*, Kunming, August 2010.

Dieussaert, K., K. Aerts, T. Steenberghen, S. Maerivoet and K. Spitaels (2009) SUSTAPARK: an agent-based model for simulating parking search, paper presented at the *AGILE International Conference on Geographic Information Science*, Hannover.

Doherty, S. T., D. Papinski and M. E. H. Lee-Gosselin (2006) An internet-based prompted recall diary with automated GPS activity-trip detection: System design, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.

Draijer, G., N. Kalfs and J. Perdok (2000) Global Positioning System as data collection method for travel research, *Transportation Research Record*, **1719**, 147–153.

Dugge, B. (2006) Ein simultanes Erzeugungs-, Verteilungs-, Aufteilungs- und Routenwahlmodell, Ph.D. Thesis, Technical University Dresden, Dresden.

Fan, Y., Q. Chen, C.-F. Liao and F. Douma (2013) UbiActive: A smartphone-based tool for trip detection and travel-related physical activity assessment, paper presented at the *92nd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2013.

Fan, Y., J. Wolfson, G. Adomavicius, K. V. Das, Y. Khandelwal and J. Kang (2015) SmarTrAC: A smartphone solution for context-aware travel and activity capturing, *Final Report*, U.S. Department of Transportation, U.S. Department of Transportation.

Federal Office of Topography swisstopo (2012) DHM25, webpage, February 2012, URL `http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/height/dhm25.html`.

Feeney, B. P. (1989) A review of the impact of parking policy measures on travel demand, *Transportation Planning and Technology*, **13** (4) 229–234.

Fischer, J. (2014) Wie fragt man nach dem Wegezweck auf dem Smartphone?, Final Thesis DAS Verkehrsingenieurwesen, IVT, ETH Zurich, Zurich.

Fischer, R. (2015) Modell zur kombinierten Routen- und Verkehrsmittelwahl aus multimodalen Fahrplanauskünften, Master Thesis, IVT, ETH Zurich, Zurich.

Frejinger, E., M. Bierlaire and M. E. Ben-Akiva (2009) Sampling of alternatives for route choice modeling, *Transportation Research Part B: Methodological*, **43** (10) 984–994.

Frignani, M. Z., J. Auld, A. K. Mohammadian, C. Williams and P. Nelson (2010) Urban travel route and activity choice surveys (UTRACS): An internet-based prompted recall activity travel survey using GPS data, paper presented at the *89th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2010.

Gallo, M., L. D'Acierno and B. Montella (2011) A multilayer model to simulate cruising for parking in urban areas, *Transport Policy*, **18** (5) 735–744.

Gatersleben, B., L. Steg and C. Vlek (2002) Measurement and determinants of environmentally significant consumer behavior, *Environment and Behavior*, **34** (3) 335–362.

Giaimo, G., R. Anderson, L. Wargelin and P. R. Stopher (2010) Will it work? Pilot results from the first large-scale GPS-based household travel survey in the United States, *Transportation Research Record*, **2176**, 26–34.

Gillen, D. W. (1977) Estimation and specification of the effects of parking costs on urban transport mode choice, *Journal of Urban Economics*, **4** (2) 186–199.

Gillen, D. W. (1978) Parking policy, parking location decisions and the distribution of congestion, *Transportation*, **7** (1) 69–85.

Glazer, A. and E. Niskanen (1992) Parking fees and congestion, *Regional Science and Urban Economics*, **22** (1) 123–132.

Golias, J., G. Yannis and M. Harvatis (2002) Off-street parking choice sensitivity, *Transportation Planning and Technology*, **25** (4) 333–348.

Gong, L., T. Yamamoto and T. Morikawa (2016) Comparison of activity type identification from mobile phone GPS data using various machine learning methods, *Asian Transport Studies*, **4** (1) 114–128.

Gould, J. (2013) Cell Phone Enabled Travel Surveys: The Medium Moves the Message, in J. Zmud, M. E. H. Lee-Gosselin, M. A. Munizaga and J. A. Carrasco (eds.) *Transport Survey Methods: Best Practice for Decision Making*, 51–70, Emerald, Bingley.

Greenhalgh, J. and M. Mirmehdi (2012) Traffic sign recognition using MSER and random forests, paper presented at the *20th European Signal Processing Conference (EUSIPCO 2012)*, Bucharest, August 2012.

Griffin, T. and Y. Huang (2005) A decision tree classification model to automate trip purpose derivation, paper presented at the *18th International Conference on Computer Applications in Industry and Engineering*, Honolulu, November 2005.

Hall, M., E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten (2009) The weka data mining software: An update, *SIGKDD Explorations*, **11** (1) 10–18.

Halldórsdóttir, K., N. Rieser-Schüssler, K. W. Axhausen, O. A. Nielsen and C. G. Prato (2014) Efficiency of choice set generation methods for bicycle routes, *European Journal of Transport and Infrastructure Research*, **14** (4) 332–348.

Hastie, T., R. Tibshirani and J. Friedman (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2. edn., Springer Series in Statistics, Springer, New York.

Hensher, D. A. and J. King (2001) Parking demand and responsiveness to supply, pricing and location in the Sidney central business district, *Transportation Research Part A: Policy and Practice*, **35** (3) 177–196.

Hood, J., E. Sall and B. Charlton (2011) A GPS-based bicycle route choice model for San Francisco, California, *Transportation Letters*, **3** (1) 63–75.

Hoogendoorn-Lanser, S. and P. H. L. Bovy (2007) Modeling overlap in multi-modal route choice by inclusion of trip part specific path size factors, *Transportation Research Record*, **2003**, 74–83.

Horni, A. and L. Montini (2013a) A Glimpse at Emergence in Agent-Based Microsimulations, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Horni, A. and L. Montini (2013b) ApplauSim: A Simulation of Synchronous Applause, presentation, Quantitative Sociology Colloquium, Zurich, April 2013.

Horni, A. and L. Montini (2013c) ApplauSim: Rhythmischer Applaus als emergentes Phänomen?, *Arbeitsberichte Verkehrs- und Raumplanung*, **853**, IVT, ETH Zurich, Zurich.

Horni, A., L. Montini and K. W. Axhausen (2013) An Agent-Based Cellular Automaton Cruising-For-Parking Simulation, *IATBR Special Issue of Transportation Letters*.

Horni, A., L. Montini, R. A. Waraich and K. W. Axhausen (2012) An agent-based cellular automaton cruising-for-parking simulation, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Jariyasunant, J., A. Carrel, V. Ekambaram, D. Gaker, T. Kote, R. Sengupta and J. L. Walker (2012) The Quantified Traveler: Using personal travel data to promote sustainable transport behavior, paper presented at the *91st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2012.

Kaplan, S. and S. Bekhor (2011) Exploring en-route parking type and parking-search route choice: Decision making framework and survey design, paper presented at the *2nd International Choice Modelling Conference*, Leeds, July 2011.

Karlin-Resnick, J., R. R. Weinberger, A. Millard-Ball, R. Hampshire, T. Dykstra and K. Lucas (2016) Parking search caused congestion: What we learn from a controlled GPS example with applications to big data, paper presented at the *95th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2016.

Kipke, H. (1993) Theoretische Überlegungen zum Parksuchverkehr, *Straßenverkehrstechnik*, **93** (4) 246–249.

Kitamura, R., P. L. Mokhtarian and L. Laidet (1997) A micro-analysis of land use and travel in five neighbourhoods in the San Francisco Bay Area, *Transportation*, **24** (2) 125–158.

Laurier, E. (2005) Searching for a parking space, *Intellectica*, **2-3** (41-42) 101–115.

Liao, L., D. Fox and H. Kautz (2007) Extracting places and activities from GPS traces using hierarchical conditional random fields, *International Journal of Robotics Research*, **26** (1) 119–134.

Lu, Y. and L. Zhang (2014) Trip purpose estimation for urban travel in the u.s.: Model development, NHTS add-on data analysis, and model transferability across different states, paper presented at the *93rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2014.

Lu, Y., S. Zhu and L. Zhang (2012) A machine learning approach to trip purpose imputation in GPS-based travel surveys, paper presented at the *4th Conference on Innovations in Travel Modeling*, Tampa, May 2012.

Maley, D. W. and R. R. Weinberger (2011) Food shopping in the urban environment: Parking supply, destination choice, and mode choice, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Marchal, P., J.-L. Madre and S. Yuan (2011) Post-processing procedures for person-based GPS data collected in the French National Travel Survey 2007-2008, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Marsden, G. (2006) The evidence base for parking policies - a review, *Transport Policy*, **13** (6) 447–457.

MathWorks (2012) MATLAB R2012b (Version 8.0), software, URL `http://www.mathworks.com/products/matlab/`.

MATSim (2015) Multi-Agent Transportation Simulation, webpage, URL `http://www.matsim.org`.

McFadden, D. (1974) Conditional logit analysis of qualitative choice-behaviour, in P. Zarembka (ed.) *Frontiers in Econometrics*, 105–142, Academic Press, New York.

Mehrabian, A. and J. A. Russell (1973) A measure of arousal seeking tendency, *Environment and Behavior*, **5** (3) 315–333.

Menghini, G., N. Carrasco, N. Schüssler and K. W. Axhausen (2010) Route choice of cyclists in zurich, *Transportation Research Part A: Policy and Practice*, **44** (9) 754–765.

MGE DATA (2012) MobiTest GSL, webpage, May 2012, URL `http://www.mgedata.com/de/hw-und-sw-produkte/custom-produkte/mobitest`.

Michalewicz, Z. and D. B. Fogel (2004) *How to Solve It: Modern Heuristics*, Springer, Heidelberg.

Miller, G. K. and C. T. Everett (1982) Raising commuter parking prices - An empirical study, *Transportation*, **11** (2) 105–129.

Moiseeva, A., J. Jessurun and H. J. P. Timmermans (2010) Semi-automatic imputation of activity travel diaries using GPS-traces, prompted recall and context-sensitive learning algorithms, *Transportation Research Record*, **2183**, 60–68.

Moiseeva, A. and H. J. P. Timmermans (2010) Imputing relevant information from multi-day GPS tracers for retail planning and management using data fusion and context-sensitive learning, *Journal of Retailing and Consumer Services*, **17** (3) 189–199.

Mokhtarian, P. L. and I. Salomon (2001) How derived is the demand for travel? Some conceptual and measurement considerations, *Transportation Research Part A: Policy and Practice*, **35** (8) 695–719.

Montini, L. (2012) Verkehrstagebücher aus GPS Daten, presentation, IVT Seminar, IVT, ETH Zurich, Zurich, December 2012.

Montini, L. and K. W. Axhausen (2015) Preliminary results: Route choice analysis from multi-day GPS data, paper presented at the *15th Swiss Transport Research Conference*, Ascona, April 2015.

Montini, L., A. Horni, N. Rieser-Schüssler and K. W. Axhausen (2012) Searching for parking in GPS data, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler and K. W. Axhausen (2014a) Comparison of travel diaries generated from smartphone data and dedicated GPS devices, paper presented at the *10th International Conference on Transport Survey Methods*, Leura, November 2014.

Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler and K. W. Axhausen (2015) Comparison of Travel Diaries Generated from Smartphone Data and Dedicated GPS Devices, *Transportation Research Procedia*, **11**, 227–241.

Montini, L. and N. Rieser-Schüssler (2014a) Peacox – Implementation and evaluation of learning routines for mode and trip purpose detection, *Deliverable*, **D4.3**, IVT, ETH Zurich, Zurich.

Montini, L. and N. Rieser-Schüssler (2014b) Peacox – Implementation and pretest of the trip purpose detection, *Deliverable*, **D4.2**, IVT, ETH Zurich, Zurich.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013a) Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013b) Handling GPS signal loss using accelerometer data, presentation, Mobile Ghent '13, Ghent University, Ghent, October 2013.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2014b) One-Week GPS-based Travel Survey in the Greater Zurich Area, presentation, TRB 2014 Workshop, Challenge of Documenting Complex Traveler Response Scenarios and Benefits of Having Such Data to Inform Transportation Policies, Washington, D.C., January 2014.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2014c) Personalisation in multi-day GPS and accelerometer, paper presented at the *14th Swiss Transport Research Conference*, Ascona, May 2014.

Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014d) Trip purpose identification from GPS tracks, *Transportation Research Record*, **2405**, 16–23.

Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014e) Trip Purpose Identification from GPS Tracks, paper presented at the *93rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2014.

Moreira, J. M., A. Jorge, J. F. d. Sousa and C. Soares (2005) Trip time prediction in mass transit companies. a machine learning approach., paper presented at the *10th Meeting of the Euro Working Group Transportation*, Poznan, September 2005.

Murakami, E. and D. P. Wagner (1999) Can using global positioning system (GPS) improve trip reporting?, *Transportation Research Part C: Emerging Technologies*, **7** (2–3) 149–165.

Niininen, O., E. Szivas and M. Riley (2004) Destination loyalty and repeat behaviour: An application of optimum stimulation measurement, *International Journal of Tourism Research*, **6** (6) 439–447.

Nitsche, P., P. Wildhalm, S. Breuss, N. Brändle and P. Maurer (2014) Supporting large-scale travel surveys with smartphones - A practical approach, *Transportation Research Part C: Emerging Technologies*, **43**, 212–221.

Office for Spatial Development of the Canton of Zurich (2015) DTM GIS ZH, webpage, February 2015, URL `http://maps.zh.ch/?topic=LidarZH&offlayers=dom2014hillshade&over=UpBackgroundZH`.

Ohmori, N., M. Nakazato, N. Harata, K. Sasaki and K. Nishii (2006) Activity diary surveys using GPS mobile phones and PDA, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.

Oliveira, M., P. Vovsha, J. Wolf, Y. Birotker, D. Givon and J. Paasche (2011a) GPS-assisted prompted recall household travel survey to support development of advanced travel model in Jerusalem, Israel, *IEEE Transactions on Power Systems*, **2246**, 16–23.

Oliveira, M., P. Vovsha, J. Wolf, Y. Birotker, D. Givon and J. Paasche (2011b) GPS-assisted prompted recall household travel survey to support development of advanced travel model in Jerusalem, Israel, paper presented at the *90th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2011.

Oliveira, M., P. Vovsha, J. Wolf and M. Mitchell (2014) Evaluating two methods for identifying trip purpose in GPS-based household travel surveys, paper presented at the *93rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2014.

OpenStreetMap (2015) The Free Wiki World Map, webpage, URL `http://www.openstreetmap.org`.

Planungsbüro Jud (1990) Parksuchverkehr Zürich, *Research Report*, Stadtplanungsamt der Stadt Zürich, Planungsbüro Jud.

Planungsbüro Jud (2010) Belegung und Verkehrsaufkommen von Parkfeldern in der Stadt Zürich, *Research Report*, Tiefbauamt der Stadt Zürich, Planungsbüro Jud.

Polak, J. W. and K. W. Axhausen (1990) Parking search behaviour: A review of current research and future prospects, *Working Paper*, **540**, Transport Studies Unit, University of Oxford, Oxford.

POSDAP (2012) Position Data Processing, webpage, URL `http://sourceforge.net/projects/posdap/`.

Prato, C. G. and S. Bekhor (2006) Applying branch & bound technique to route choice set generation, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.

Prato, C. G. and S. Bekhor (2007) Modeling route choice behavior: How relevant is the composition of choice set?, *Transportation Research Record*, **2003**, 64–73.

Prost, S., K. Röderer, J. Schrammel, E. Bothos, L. Montini and S. Alam (2013a) Peacox – field trials I report, *Deliverable*, **D7.4**, PEACOX project, Center for Usability Research and Engineering, Vienna.

Prost, S., J. Schrammel, K. Röderer and M. Tscheligi (2013b) Contextualise! personalise! persuade!: a mobile hci framework for behaviour change support systems, paper presented at the *MobileHCI '13*, 510–515, Munich, August 2013.

Ramming, M. S. (2002) Network knowledge and route choice, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.

Resource Systems Group (RSG) (2015) rmove, URL `http://rmove.rsginc.com/index.html`. Accessed on 10/08/2015.

Rieser-Schüssler, N. and K. W. Axhausen (2011) Combining GPS travel diaries with psychometric scales, paper presented at the *9th International Conference on Survey Methods in Transport*, Termas de Puyehue, November 2011.

Rieser-Schüssler, N. and K. W. Axhausen (2013) Identifying chosen public transport connections from GPS observations, paper presented at the *92nd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2013.

Rieser-Schüssler, N., M. Balmer and K. W. Axhausen (2012a) Route choice sets for very high-resolution data, *Transportmetrica*.

Rieser-Schüssler, N., L. Montini and C. Dobler (2011) Improving automatic post-processing routines for GPS observations using prompted-recall data, paper presented at the *9th International Conference on Survey Methods in Transport*, Termas de Puyehue, November 2011.

Rieser-Schüssler, N., L. Montini, M. Rieser and K. W. Axhausen (2012b) Preparations for estimating the influence of attitudes on public transport connection choice, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Rieser-Schüssler, N., M. Rieser, L. Montini and K. W. Axhausen (2014) Exploring choice set generation approaches for public transport connection choice, *Arbeitsberichte Verkehrs- und Raumplanung*, **1023**, IVT, ETH Zurich, Zurich.

Rodrigues, F., F. C. Pereira, A. Alves, S. Jiang and J. Ferreira (2012) Automatic classification of points-of-interest for land-use analysis, paper presented at the *4th International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2012)*, Valencia, January 2012.

Rodriguez, D. A., L. Merlin, C. G. Prato, T. L. Conway, D. Cohen, J. P. Elder, K. Evenson, T. L. McKenzie, J. L. Pickrel and S. Veblen-Mortenson (2015) Influence of the built environment on pedestrian route choices of adolescent girls, *Environment and Behavior*, **47** (4) 359–394.

Safi, H., B. Assemi, M. Mesbah, L. Ferreira and M. D. Hickman (2015) Design and implementation of a smartphone-based system for personal travel survey: Case study from new zealand, paper presented at the *94th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2015.

Schmutz, S. (2015) Effect of analytical units and aggregation rules on mode choice models, Master Thesis, IVT, ETH Zurich, Zurich.

Schultz, P. W. (2001) The structure of environmental concern: Concern for self, other people, and the biosphere, *Journal of Environmental Psychology*, **21** (4) 327–339.

Schüssler, N. (2010) Accounting for similarities between alternatives in discrete choice models based on high-resolution observations of transport behaviour, Ph.D. Thesis, ETH Zurich, Zurich.

Schüssler, N. and K. W. Axhausen (2008) Processing GPS raw data without additional information, *Working Paper*, **515**, IVT, ETH Zurich, Zurich.

Schüssler, N. and K. W. Axhausen (2009a) Accounting for route overlap in urban and suburban route choice decisions derived from GPS observations, paper presented at the *12th International Conference on Travel Behaviour Research (IATBR)*, Jaipur, December 2009.

Schüssler, N. and K. W. Axhausen (2009b) Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique (MHT), *Working Paper*, **568**, IVT, ETH Zurich, Zurich.

Schüssler, N. and K. W. Axhausen (2009c) Processing GPS raw data without additional information, *Transportation Research Record*, **2105**, 28–36.

Shen, L. and P. R. Stopher (2014) Review of GPS travel survey and GPS data-processing methods, *Transport Reviews: A Transnational Transdisciplinary Journal*, **34** (3) 316–334.

Shoup, D. (2005) *The High Cost of Free Parking*, Planners Press, Chicago.

Socialdata (2009a) Aktivitäten, webpage, URL http://www.socialdata.de/daten/modechoice.php. Accessed on 10/12/2015.

Socialdata (2009b) Zeitreihen Verkehrsmittelwahl, webpage, URL http://www.socialdata.de/daten/modechoice.php. Accessed on 10/12/2015.

Speed, T. (ed.) (2003) *Statistical Analysis of Gene Expression Microarray Data*, Chapman & Hall/CRC, Boca Raton.

Stadt Zürich Präsidialdepartement, Statistik Stadt Zürich (2011) Quartierspiegel Gesamtreihe aller 34 Stadtquartiere.

Stenneth, L., O. Wolfson, P. S. Yu and B. Xu (2011) Transportation mode detection using mobile phones and GIS information, paper presented at the *19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Chicago, November 2011.

Stopher, P. R., C. FitzGerald and J. Zhang (2008) Search for a global positioning system device to measure person travel, *Transportation Research Part C: Emerging Technologies*, **16** (3) 350–369.

Stopher, P. R., Q. Jiang and C. FitzGerald (2005) Processing GPS data from travel surveys, paper presented at the *2nd International Colloqium on the Behavioural Foundations of Integrated Land-use and Transportation Models: Frameworks, Models and Applications*, Toronto, June 2005.

Stopher, P. R., L. Shen, W. Liu and A. Ahmed (2015) The challenge of obtaining ground truth for GPS processing, *Transportation Research Procedia*, **11**, 206–217.

Studio Mobilita (2015) Studio mobilita, webpage, URL http://www.studio-mobilita.ch. Accessed on 10/08/2015.

Supek, F. (2012) FastRandomForest, webpage, URL https://code.google.com/p/fast-random-forest/.

Swiss Federal Statistical Office (BFS) (2005) Mikrozensus Verkehr 2005, URL http://www.bfs.admin.ch/bfs/portal/de/

index/infothek/erhebungen__quellen/blank/blank/mz/01.html.

Swiss Federal Statistical Office (BFS) (2010) Mikrozensus Verkehr 2010, URL http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/mz/01.html.

Thompson, R. G. and A. J. Richardson (1998) A parking search model, *Transportation Research Part A: Policy and Practice*, **32** (3) 159–170.

Topp, H. H. (1991) Parking policies in large cities in Germany, *Transportation*, **18** (1) 3–21.

Tsui, S. Y. A. and A. Shalaby (2006) An enhanced system for link and mode identification for GPS-based personal travel surveys, *Transportation Research Record*, **1972**, 38–45.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (1998) The impact of the parking situation in shopping centres on store choice behaviour, *GeoJournal*, **45** (4) 309–315.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (2006) Attitudes and behavioral responses to parking measures, *European Journal of Transport and Infrastructure Research*, **6** (4) 301–312.

van der Waerden, P., A. W. J. Borgers and H. J. P. Timmermans (2009) Consumer response to the introduction of paid parking in a regional shopping center, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.

van der Waerden, P., H. Oppewal and H. J. P. Timmermans (1993) Adaptive choice behaviour of motorists in congested shopping centre parking lots, *Transportation*, **20**, 395–408.

van der Waerden, P., H. J. P. Timmermans and A. W. J. Borgers (2002) PAMELA: Parking analysis model for predicting effects in local areas, *Transportation Research Record*, **1781**, 10–18.

Vlassenroot, S., D. Gillis, R. Bellens and S. Gautama (2015) The use of smartphone applications in the collection of travel behaviour data, *International Journal of Intelligent Transportation Systems Research*, **13** (1) 17–27.

Wagner, D. P. (1997) Lexington area travel data collection test: GPS for personal travel surveys, *Final Report*, Office of Highway Policy Information and Office of Technology Applications, Federal Highway Administration, Battelle Transport Division, Columbus, September 1997.

Weber, E. U., A.-R. Blais and N. E. Betz (2002) A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors, *Journal of Behavioral Decision Making*, **15** (4) 263–290.

Weis, C., M. Vrtic, J.-P. Widmer and K. W. Axhausen (2011) Influence of parking on location and mode choice: A stated choice survey, *Working Paper*, **684**, IVT, ETH Zurich, Zurich.

Widmer, J.-P. and M. Vrtic (2004) Einfluss von Änderungen des Parkierungsangebots auf das Verkehrsverhalten, *Research Report*, **VSS 1997/46**, Schweizerischer Verband der Strassen- und Verkehrsfachleute (VSS), Büro Widmer, Frauenfeld.

Wolf, J. (2004) Applications of new technologies in travel surveys, paper presented at the *7th International Conference on Survey Methods in Transport*, Costa Rica, August 2004.

Wolf, J., R. Guensler and W. Bachman (2001a) Elimination of the travel diary - experiment to derive trip purpose from Global Positioning System travel data, *Transportation Research Record*, **1768**, 125–134.

Wolf, J., R. Guensler and W. Bachman (2001b) Elimination of the travel diary: An experiment to derive trip purpose from GPS travel data, paper presented at the *80th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2001.

Wolf, J., S. Schönfelder, U. Samaga, M. Oliveira and K. W. Axhausen (2004) Eighty weeks of Global Positioning System traces, *Transportation Research Record*, **1870**, 46–54.

Wright, C. C. and H. C. Orram (1976) The Westminster route choice survey: A new technique for traffic studies, *Traffic Engineering and Control*, **17** (8/9) 348–351, 354.

Wu, J., C. Jiang, D. Houston, D. Baker and R. Delfino (2011) Automated time activity classification based on global positioning system (GPS) tracking data, *Environmental Health*, **10**, 101.

Yalamanchili, L., R. M. Pendyala, N. Prabaharan and P. Chakravarty (1999) Analysis of Global Positioning System-based data collection methods for capturing multistop trip-chaining behavior, *Transportation Research Record*, **1660**, 58–65.

Young, W. (1986) PARKSIM/1: a simulation model of driver behaviour in parking lots, *Traffic Engineering and Control*, **27** (12) 606–613.

Young, W. and R. G. Thompson (1987) PARKSIM/1: a computer graphics approach for parking lot layouts, *Traffic Engineering and Control*, **28** (3) 120–123.

Young, W. and T. Y. Weng (2005) Data and parking simulation models, in R. Kitamura and M. Kuwahara (eds.) *Simulation Approaches in Transportation Analysis: Recent Advances and Challenges*, 235–267, Springer, New York.

Zheng, Y., L. Liu, L. Wang and X. Xie (2008) Learning transportation mode from raw GPS data for geographic applications on the web, paper presented at the *17th World Wide Web Conference*, Beijing, April 2008.

LARA MONTINI

PERSONAL INFORMATION

Address:	Schlierenstrasse 17 8142 Uitikon Waldegg Switzerland
Phone:	+41 44 633 30 88
E-mail:	lara.montini@ivt.baug.ethz.ch
Nationality:	Swiss
Date of birth:	16th of August 1985
Place of birth:	Basel, Switzerland

EDUCATION

Mar 11 – present	Doctoral studies, ETH Zurich, Switzerland
Dissertation title:	*Extraction of Transportation Information from Combined Position and Accelerometer Tracks*
Advisor:	Prof. Dr. Kay W. Axhausen
Oct 03 – Aug 08	MSc ETH in Mechanical Engineering, ETH Zurich
Master thesis:	*Modeling of Biological Processes*
Bachelor thesis:	*CFD analysis of a vortex-flow solar reactor for the combined calcination reforming process*
Aug 02 – Jul 03	Diplom Direktionsassistentin, Neue Sprach und Handelsschule (NSH) Basel
Aug 98 – Mar 02	Eidgenössische Matura Typus C, BZB Basel

WORKING EXPERIENCE

Mar 11 – May 16	Institute for Transport Planning and Systems, ETH Zurich, Switzerland Research associate
Sep 08 – Dec 10	Helbing Technik AG, Aarau, Switzerland Software Developer

SKILLS

Programming	Java, R, Matlab, SQL, Ruby (on Rails), C++, VBA
Language skills	Native: German Fluent: English Good: French

PAPERS & PUBLICATIONS

REFEREED JOURNAL PAPERS

Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler and K. W. Axhausen (2015) Comparison of Travel Diaries Generated from Smartphone Data and Dedicated GPS Devices, *Transportation Research Procedia*, **11**, 227–241.

Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014d) Trip purpose identification from GPS tracks, *Transportation Research Record*, **2405**, 16–23.

Horni, A., L. Montini and K. W. Axhausen (2013) An Agent-Based Cellular Automaton Cruising-For-Parking Simulation, *IATBR Special Issue of Transportation Letters*.

REFEREED CONFERENCE PAPERS

Montini, L., A. Horni, N. Rieser-Schüssler and K. W. Axhausen (2012) Searching for parking in GPS data, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Rieser-Schüssler, N., L. Montini, M. Rieser and K. W. Axhausen (2012b) Preparations for estimating the influence of attitudes on public transport connection choice, paper presented at the *13th International Conference on Travel Behaviour Research (IATBR)*, Toronto, July 2012.

Rieser-Schüssler, N., L. Montini and C. Dobler (2011) Improving automatic post-processing routines for GPS observations using prompted-recall data, paper presented at the *9th International Conference on Survey Methods in Transport*, Termas de Puyehue, November 2011.

CONFERENCE PAPERS

Montini, L. and K. W. Axhausen (2015) Preliminary results: Route choice analysis from multi-day GPS data, paper presented at the *15th Swiss Transport Research Conference*, Ascona, April 2015.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2014c) Personalisation in multi-day GPS and accelerometer, paper presented at the *14th Swiss Transport Research Conference*, Ascona, May 2014.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013a) Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Horni, A. and L. Montini (2013a) A Glimpse at Emergence in Agent-Based Microsimulations, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

WORKING PAPERS AND RESEARCH REPORTS

Montini, L. and N. Rieser-Schüssler (2014a) Peacox – Implementation and evaluation of learning routines for mode and trip purpose detection, *Deliverable*, **D4.3**, IVT, ETH Zurich, Zurich.

Montini, L. and N. Rieser-Schüssler (2014b) Peacox – Implementation and pretest of the trip purpose detection, *Deliverable*, **D4.2**, IVT, ETH Zurich, Zurich.

Rieser-Schüssler, N., M. Rieser, L. Montini and K. W. Axhausen (2014) Exploring choice set generation approaches for public transport connection choice, *Arbeitsberichte Verkehrs- und Raumplanung*, **1023**, IVT, ETH Zurich, Zurich.

Artukovic, M., B. Abunaim, L. Montini, E. Bothos, T. Tvrzsky and A. McNabola (2013) Peacox – server and client implementation, *Deliverable*, **D6.2**, PEACOX project, Fluidtime Data Services GmbH, Vienna.

Horni, A. and L. Montini (2013c) ApplauSim: Rhythmischer Applaus als emergentes Phänomen?, *Arbeitsberichte Verkehrs- und Raumplanung*, **853**, IVT, ETH Zurich, Zurich.

Prost, S., K. Röderer, J. Schrammel, E. Bothos, L. Montini and S. Alam (2013a) Peacox – field trials I report, *Deliverable*, **D7.4**, PEACOX project, Center for Usability Research and Engineering, Vienna.

PRESENTATIONS

Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler and K. W. Axhausen (2014a) Comparison of travel diaries generated from smartphone data and dedicated GPS devices, paper presented at the *10th International Conference on Transport Survey Methods*, Leura, November 2014.

Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014e) Trip Purpose Identification from GPS Tracks, paper presented at the *93rd Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2014.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2014b) One-Week GPS-based Travel Survey in the Greater Zurich Area, presentation, TRB 2014 Workshop, Challenge of Documenting Complex Traveler Response Scenarios and Benefits of Having Such Data to Inform Transportation Policies, Washington, D.C., January 2014.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013b) Handling GPS signal loss using accelerometer data, presentation, Mobile Ghent '13, Ghent University, Ghent, October 2013.

Horni, A. and L. Montini (2013b) ApplauSim: A Simulation of Synchronous Applause, presentation, Quantitative Sociology Colloquium, Zurich, April 2013.

Montini, L. (2012) Verkehrstagebücher aus GPS Daten, presentation, IVT Seminar, IVT, ETH Zurich, Zurich, December 2012.

SUPERVISED STUDENT THESES

Fischer, R. (2015) Modell zur kombinierten Routen- und Verkehrsmittelwahl aus multimodalen Fahrplanauskünften, Master Thesis, IVT, ETH Zurich, Zurich.

Schmutz, S. (2015) Effect of analytical units and aggregation rules on mode choice models, Master Thesis, IVT, ETH Zurich, Zurich.

Fischer, J. (2014) Wie fragt man nach dem Wegezweck auf dem Smartphone?, Final Thesis DAS Verkehrsingenieurwesen, IVT, ETH Zurich, Zurich.

R E V I E W I N G

Transportation
Transportation Research Part D
International Journal of Geographical Information Science
Computers, Environment and Urban Systems
European Journal of Transport and Infrastructure Research