

Available online at www.sciencedirect.com**ScienceDirect**

Transportation Research Procedia 11 (2015) 227 – 241


www.elsevier.com/locate/procedia

10th International Conference on Transport Survey Methods

Comparison of travel diaries generated from smartphone data and dedicated GPS devices

Lara Montini^{a*}, Sebastian Probst^b, Johann Schrammel^b, Nadine Rieser-Schüssler^a,
Kay W. Axhausen^a

^a Institute for Transport Planning and Systems, ETH Zurich, 8093 Zurich, Switzerland

^b Austrian Institute of Technology, Business Park Marximum, 1110 Vienna, Austria

Abstract

This paper provides further insight into the usefulness of smartphones and dedicated GPS devices for collecting current travel survey data. GPS and accelerometer time series for 33 European research project PEACOX study participants are available for analysis; these were tracked simultaneously with smartphones and dedicated devices for 8 weeks. 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.

© 2015 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of International Steering Committee for Transport Survey Conferences ISCTSC

Keywords: GPS; smartphone; travel diary; trip planner; ecological behaviour

1. Introduction and related work

In transportation research, GPS traces are used, along with other data sources, to construct travel diaries. This data is primarily collected using dedicated GPS devices that respondents must carry with them during the tracking period. Smartphones are a promising source for GPS 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

* Corresponding author. Tel.: +41-44-6333088; fax: +41-44-6331057.

E-mail address: lara.montini@ivt.baug.ethz.ch

opposed to dedicated devices, are often anyway carried by participants. 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, hand-held 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)).

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 possibility 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.

The goal of this paper is to provide further insight into the current usability of smartphones and dedicated GPS devices for collecting travel data. Data quality and usage of the two device types are compared, as are travel diaries generated from the two data sources. GPS and accelerometer time series of 33 study participants are available; these were tracked simultaneously with smartphones and dedicated devices for 8 weeks. The data was collected as part of the PEACOX project (www.project-peacox.eu), where a personalised journey planner application for smartphones to encourage ecological behaviour was developed. In the app, position data is collected to generate travel diaries; this is then used to personalise route suggestions. In this paper, data collected in the second field trial of the app (Vienna and Dublin from August to October 2014) is analysed.

The paper is structured as follows. First, smartphone applications used in the PEACOX project (journey planning app, as well as a prompted recall app) are presented. In the next section, the field trial is described. Section 4 outlines differences in travel diary construction for the different device types. In Section 5, results are reported, including quantitative analysis, as well as users' subjective perceptions. An interpretation of results and an outlook on continuing work concludes.

2. 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 CO₂ 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 the prompted recall applications are also introduced.

2.1. Journey planning app

The PEACOX journey planning app allows the user to perform a multi-modal search for a route tailored to the user's individual preferences and behaviour 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. Selected eco-friendly route options are promoted by adding a persuasive message (Fig. 1(a)). Other persuasive elements were implemented: challenges where users competed against each other, as well as comparing themselves on emissions rankings and, finally, a

user's own improvement, represented by a growing or shrinking tree. These and other potential persuasion strategies are discussed in Prost et al. (2013b). Implementation details of this android application can be found in Artukovic et al. (2013).

2.2. Prompted recall app

A prompted recall tool allows participants to review collected trip history and provide manual corrections. Normally, these tools are web-based, as shown in Auld et al. (2009) and as, e.g., implemented by Montini et al. (2013) for a GPS-based survey in Zurich. 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 (Fig. 1(b)). Each day could be selected from the menu, the activity type and transport mode could be changed given predefined lists and, for each day, a comment can then be left. 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.

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.

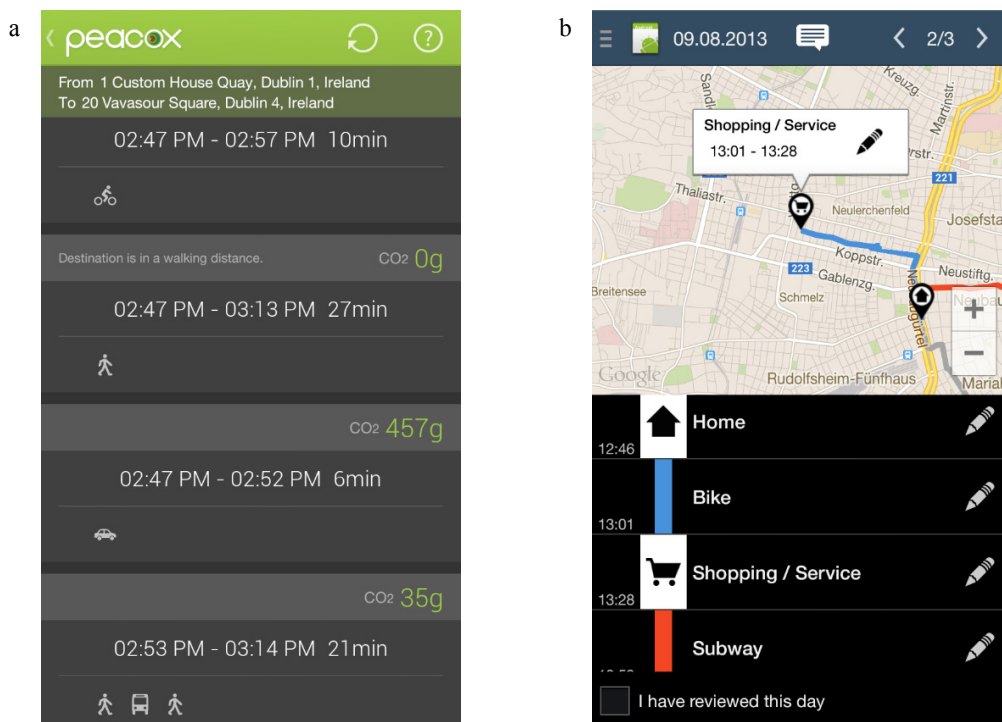


Fig. 1. PEACOX applications; (a) journey planner: routes with CO₂ information and persuasive message; (b) prompted recall: map and diary list.

3. 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). 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 financial compensation after completion of the survey.

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 sample described in Table 1.

Table 1. Characteristics of participants.

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 and 2 on parental leave
Main transportation means	6 users mainly use car or motorbike, 6 use bicycles, 11 public transport, 5 walking and 9 did not define
Usage of journey planning app	8 participants had never used a journey planning app prior to the study, 29 had

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.

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.

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 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 2. Smartphone types used in field trial.

Smartphone type	Number of 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. 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.

Routine configuration was calibrated on data collected with the same dedicated GPS loggers used in this survey (MobiTest GSL). The training data was collected in and around Zurich, for up to one week, by 150 different people (Montini et al., 2013).

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.

4.1. Uncorrected diaries from smartphones

The uncorrected diaries evaluated in this paper, 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. (2014), 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.

4.2. 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; accordingly, activities are also merged. Further, transport mode and trip purpose corrections by participants are considered. For this paper, no further corrections were made by the researchers.

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.

5. Results

Results of the field trials are organised 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.

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) after cleaning of invalid GPS points. In Fig. 2, 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 Fig. 2, 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.

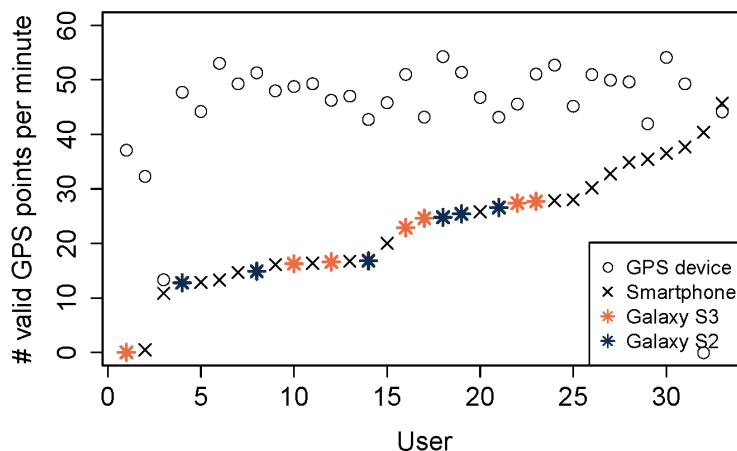


Fig. 2. Average GPS point frequency for detected stages for both devices carried simultaneously.

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, Fig. 3 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 of 10 days, which is above expectations for an 8-week-field-trial, but not too much. In any case, there is definitely movement not captured 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.

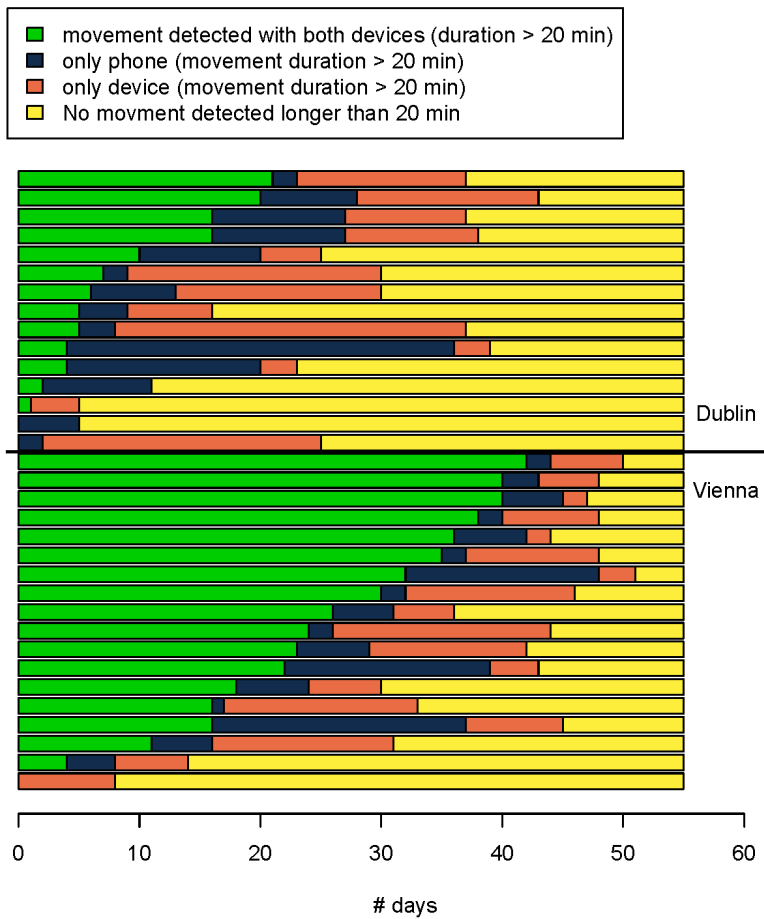


Fig. 3. Daily detected duration per person.

Fig. 3 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, Fig. 4 shows summed up movement durations for all days on the right 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 Fig. 4, 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 Fig. 4(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 around 9 hours; probably, an activity was wrongfully marked as 'no activity'. The movement patterns on the right show trips 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 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.

Fig. 4(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. A good example of this is the participant's data shown in Fig. 4(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. Fig. 4(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:

$$C_{s1,d2} = \frac{move_{d2,s1}}{dur_{s1}} \quad (1)$$

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, Fig. 5 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 Fig. 3. The sample users presented there are also highlighted in Fig. 5 and the order of participants is determined by the sum of all four shown criteria.

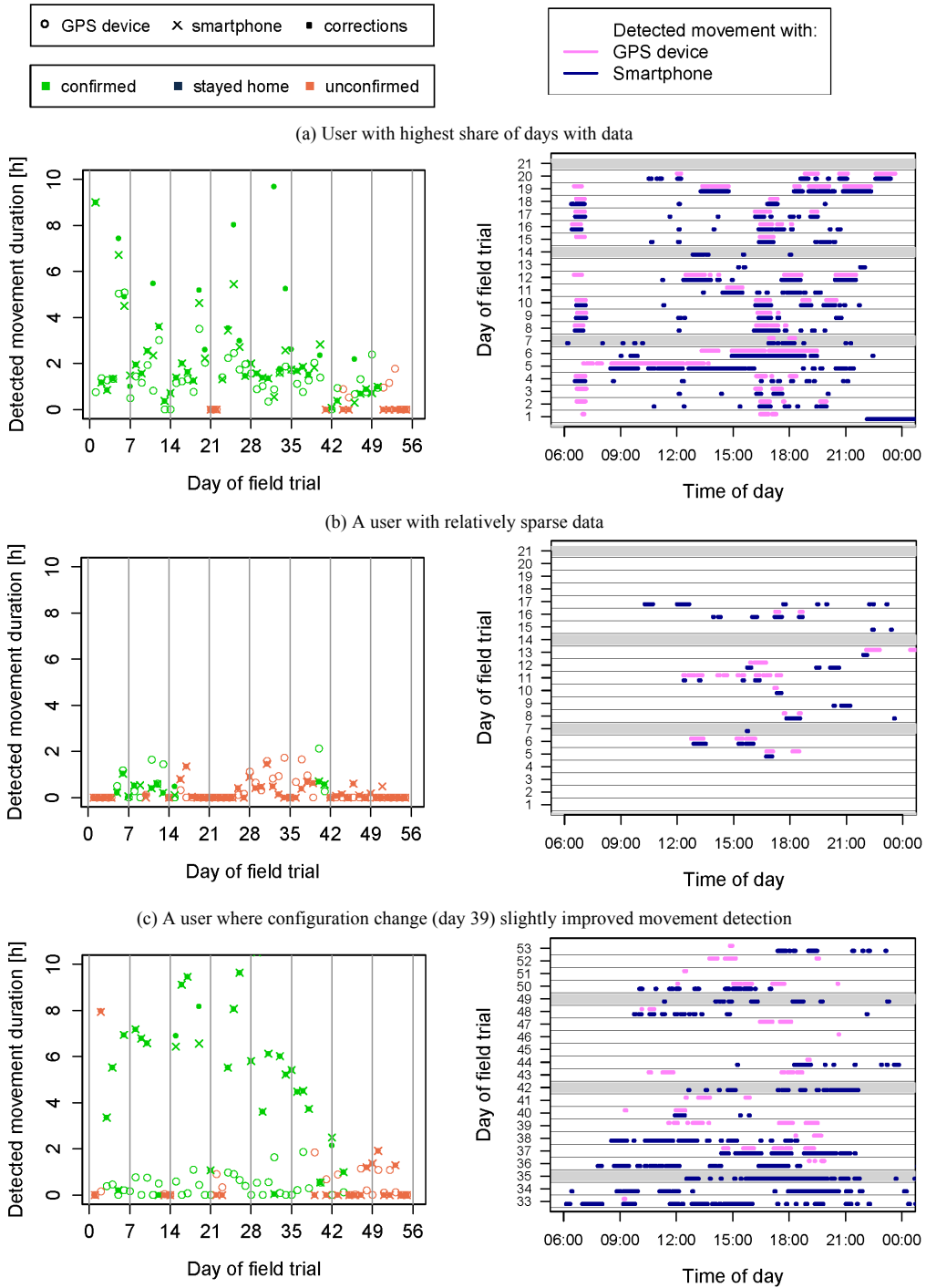


Fig. 4. Cumulated detected daily movement and sample of movements during the day for three selected users. Sundays highlighted in grey.

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 positive, as it can probably be improved by personalised configurations.

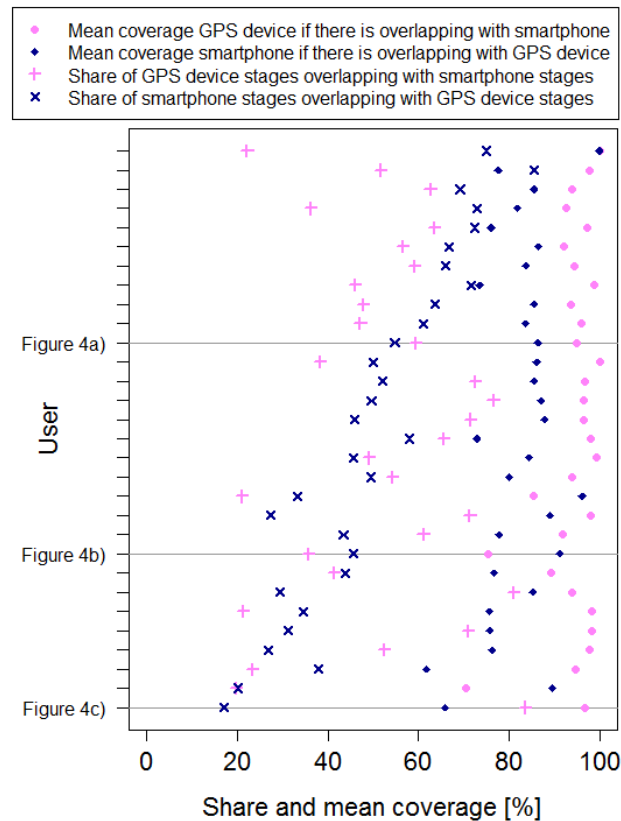


Fig. 5. Mean coverage of overlapping stages and share of overlapping stages. Only days where both devices were active (Fig. 3).

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, 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 Fig. 6, 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.

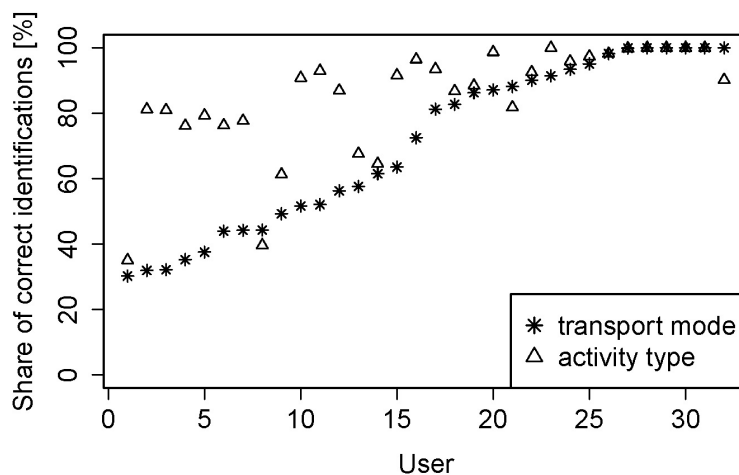


Fig. 6. Detection accuracy per user (based on corrections of smartphone-based diary).

5.4. 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 the logging off 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 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.

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 (Fig. 5); 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 researcher's 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 includes 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, will 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.

Acknowledgements

This research was funded by the European Union as part of the project “PEACOX - persuasive advisor for CO₂-reducing cross-modal trip planning” within the Seventh Framework Programme (FP7). We would like to thank all project partners for the smooth collaboration. Especially, all co-workers at AIT and the survey team from Trinity College Dublin for recruitment and support of survey participants. The smartphone apps were developed by Fluidtime Data Services GmbH and Telematix Group. We also thank the anonymous reviewers for their valuable input which helped improving the paper.

References

- Alam, S., McNabola, A., 2012. Peacox – door to door emissions models, Deliverable, D3.1, PEACOX project, Trinity College Dublin, Dublin.
- Artukovic, M., Abunaim, B., Montini, L., Bothos, E., Tvrzsky, T., McNabola, A., 2013. Peacox – server and client implementation, Deliverable, D6.2, PEACOX project, Fluidtime Data Services GmbH, Vienna.
- Asakura, Y., Hato, E., 2004. Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies*, 12, 273–291.
- Auld, J., Williams, C., Mohammadian, A. K., Nelson, P., 2009. An automated GPS-based prompted recall survey with learning algorithms. *Transportation Letters*, 1 (1) 59–79.
- Becker, H., Ciari, F., Brignoni, M., Axhausen, K. W., 2015. Impacts of a new free-floating car-sharing system in Basel. Presented at the 14th International Conference on Travel Behaviour Research (IATBR), Windsor, July 2015.
- Ben Abdesslem, F., Phillips, A., Henderson, T., 2009. Less is more: Energy-efficient mobile sensing with senseless. Presented at the MobiHeld’09 the 1st ACM workshop on Networking, systems, and applications for mobile handhelds, Barcelona, August 2009.
- Bothos, E., Apostolou, D., Mentzas, G., 2012. Recommending eco-friendly route plans. Presented at the Workshop on Recommendation Technologies for Lifestyle Change, Dublin, September 2012.
- Cottrill, C. D., Pereira, F. C., Zhao, F., Dias, I., Lim, H. B., Ben-Akiva, M. E., Zegras, P. C., 2013. Future mobility survey: Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record*, 2354, 95–67.
- Fan, Y., Chen, Q., Liao, C. F., Douma, F., 2013. UbiActive: A smartphone-based tool for trip detection and travel-related physical activity assessment. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., January 2013.
- Fan, Y., Wolfson, J., Adomavicius, G., Das, K. V., Khandelwal, Y., Kang, J., 2015. SmarTrAC: A smartphone solution for context-aware travel and activity capturing, Final Report, U.S. Department of Transportation, U.S. Department of Transportation.
- Gould, J., 2013. Cell Phone Enabled Travel Surveys: The Medium Moves the Message, in “*Transport Survey Methods: Best Practice for Decision Making*”. In: Zmud, J., Lee-Gosselin, M. E. H., Munizaga, M.A., Carrasco, J. A. (eds.). Emerald, Bingley, pp. 51-70.
- Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D., Kote, T., Sengupta, R., Walker, J., 2012. The Quantified Traveler: Using personal travel data to promote sustainable transport behavior. Presented at the 91st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2012.
- MGE DATA, 2012. MobiTest GSL, webpage, <http://www.mgedata.com/de/hw-und-sw-produkte/custom-produkte/mobitest>. Accessed on 14/08/2015.
- Montini, L., Rieser-Schüssler, N., Axhausen, K. W., 2013. Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area. Presented at the 13th Swiss Transport Research Conference, Ascona, April 2013.
- Montini, L., Rieser-Schüssler, N., Horni, A., Axhausen, K. W., 2014. Trip purpose identification from GPS tracks. *Transportation Research Record*, 2405, 16–23.
- Nitsche, P., Wildhalm, P., Breuss, S., Brändle, N., Maurer, P., 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*, 43, 212–221.
- Ohmori, N., Nakazato, M., Harata, N., Sasaki, K., Nishii, K., 2006. Activity diary surveys using GPS mobile phones and PDA. Presented at the 85th Annual Meeting of the Transportation Research Board, Washington, D.C., January 2006.
- POSDAP, 2012. Position Data Processing, webpage, <http://sourceforge.net/projects/posdap/>. Accessed on 14/08/2015.

- Prost, S., Röderer, K., Schrammel, J., Bothos, E., Montini, L., Alam, S., 2013a. Peacox – field trials I report, Deliverable, D7.4, PEACOX project, Center for Usability Research and Engineering, Vienna.
- Prost, S., Schrammel, J., Röderer, K., Tscheligi, M., 2013b. Contextualise! personalise! persuade!: a mobile hci framework for behaviour change support systems. Presented at the MobileHCI '13, 510–515, Munich, August 2013.
- Resource Systems Group (RSG), 2015. rMove, webpage, <http://rmove.rsginc.com/index.html>. Accessed on 10/08/2015.
- Safi, H., Assemi, B., Mesbah, M., Ferreira, L., Hickman, M. D., 2015. Design and implementation of a smartphone-based system for personal travel survey: Case study from New Zealand. Presented at the 94th Annual Meeting of the Transportation Research Board, Washington, D.C., January 2015.
- Shen, L., Stopher, P. R., 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews: A Transnational Transdisciplinary Journal*, 34 (3) 316–334.
- Studio Mobilita, 2015. Studio Mobilita, webpage, <http://www.studio-mobilita.ch>. Accessed on 10/08/2015.
- Vlassenroot, S., Gillis, D., Bellens, R., Gautama, S., 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.
- Wolf, J., 2004. Applications of new technologies in travel surveys. Presented at the 7th International Conference on Survey Methods in Transport, Costa Rica, August 2004.