A SERIES OF THREE CASE STUDIES ON THE SEMI-AUTOMATION OF ACTIVITY TRAVEL DIARY GENERATION USING SMARPTHONES

Adrian C. Prelipcean
Corresponding Author
KTH Royal Institute of Technology
Department of Transport Science &
KTH Royal Institute of Technology
Department of Urban Planning and Environment (Geoinformatics)
acpr@kth.se

Győző Gidófalvi
KTH Royal Institute of Technology
Department of Urban Planning and Environment (Geoinformatics)
gyozo.gidofalvi@abe.kth.se

Yusak O. Susilo
KTH Royal Institute of Technology
Department of Transport Science
yusak.susilo@abe.kth.se

Word Count: 5207 words + 1 figure(s) + 8 table(s) = 7457 words

Paper was presented at a poster session at the TRB committee ABJ40 Standing Committee on Travel Survey Methods during the 96th Annual Meeting of the Transportation Research Board, Washington, D.C., January 2017, https://trid.trb.org/view.aspx?id=1438182.

Poster available at: http://adrianprelipcean.github.io/papers/TRB_poster.pdf

Preferred citation: Prelipcean, A.C., Gidófalvi, G. and Susilo, Y.O. (2017) A series of three case studies on the semi-automation of activity travel diary generation using smarpthones, TRB 96th Annual Meeting Compendium of Papers, 16 pages, paper number 17-02898.

ABSTRACT

The growing need of acquiring data that is useful for travel behaviour analysis led scientists to pursue new ways of obtaining travel diaries from large groups of people. The most promising alternative to traditional (declarative) travel diary collection methods are those that rely on collecting trajectories from individuals and then extract travel diary semantics from the trajectories. However, most studies report on routines specific to the post-processing of data, and seldom focus on data collection. Even the few studies that deal explicitly with data collection describe the final state of the collection system, but do not go at the lengths that are required to describe the decision that were taken to bring the system to its current state. This leads to a considerable amount of work that is needed for designing collection systems that are often undocumented, which impedes the reuse of the aforementioned systems. In light of the aforementioned problems, this paper presents a series of three case studies behind the continuous development of MEILI, a travel diary collection, annotation and automation system, in an effort to: 1) illustrate the utility of the developed system to collect travel diaries, 2) identify how MEILI and other semi-automatic travel diaries collection systems can be improved, and 3) propose MEILI as an open source system that has the potential of being improved into a widely available semi-automated travel diary collection system.

1. INTRODUCTION

The recent increase of population, accompanied by an ongoing expansion of urban areas induce a notion of urgency for planning the continuous development of transportation networks, infrastructure and policies. The traditional planning ecosystem does not have difficulties due to time pressure and, as such, relies on data collection strategies that are thorough but time consuming. In contrast, the modern ecosystem is time bound and relies on the promptness of response, which is generally linked to continuous data collection methods on a large scale. The type of data collected for understanding and predicting how people can react to changes in transportation networks, infrastructure and policies is commonly called travel behaviour data, and is obtained via travel diaries.

A travel diary is a sequential description of what a traveler has been doing during a predefined time frame. Common travel behavior data collection strategies rely on asking respondents (via phone interviews, letters or web forms) for a verbose description of their scheduled activities during a given day (*1–7*). The main issues of these traditional methods are their high organization and maintenance costs (some countries cannot allocate the required budget for organizing and performing the traditional surveys), and their continuously declining response rates, which vary on a per country basis, e.g., Norway had a response rate of 20% in 2013 (*8*), Sweden had a response rate of 45.4% in 2011 (*9*), and the United Kingdom had a response rate of 59% in 2014 (*10*).

Researchers studied different technologies to replace the traditional travel diary collection methods, in particular using GPS enabled devices (e.g., GPS receivers, smartphones) accompanied by software that allows respondents to annotate their collected data. However, most research does not offer a thorough and detailed description of the steps taken to obtain a system that collects travel diaries, but rather present a static snapshot of the system in its final stage. Furthermore, the focus of the relevant research is on methods that automatically extract travel diary semantics from trajectories, which is usually presented as a post processing routine and is independent of a system that collects travel diaries, which impedes the usage of the system in the real world. As such, at the moment of this writing, there is no widely accepted system for the collection of travel diaries that is preferred over the traditional travel diary collection methods.

This paper presents a series of case studies behind the continuous development of MEILI, a travel diary collection, annotation and automation system, in an effort to: 1) illustrate the utility of the developed system to collect travel diaries, 2) identify how MEILI and other semi-automatic travel diaries collection systems can be improved, and 3) propose MEILI as an open source system that has the potential to be improved into a widely available semi-automated travel diary collection system. The discussions that relate each of the case studies to one another emphasize the qualitative and quantitative aspects of the data collected by MEILI.

The remainder of this paper is organized as follows: Section 2 presents a literature review, Section 3 presents an overview of the methodology relevant for the current paper, Section 4 presents the case studies associated with this paper, Section 5 presents the conclusion of the paper and provides relevant discussions, and Section 6 presents the future work.

2. LITERATURE REVIEW

There is a large research corpus that is dedicated to inferring travel diary semantics from trajectories, i.e., trip and tripleg segmentation (*11–14*), travel mode detection (*15–19*), destination inference (*20*) and purpose inference (*21–27*). However, each task is usually performed in isolation to other tasks, which does not allow for any direct insight regarding how close a fully automated solution that generates travel diaries from trajectories is. This downside is accompanied by a lack of thorough analysis of performance and error measures for the aforementioned inferences, most research relying on the same default measures used to analyze the performance of simple inference tasks such as spam detection (*28*). In-depth investigation of error measures show that the traditional performance measures do not offer meaningful insights for more complex tasks such as travel behavior (*15, 16*). Due to the understudied measures of performance for the required inferences, this paper does not go into the depths of inferring travel diary semantics from trajectories, but focuses on analyzing the data collected with MEILI during three case studies.

The attempts that have been made for automated and semi-automated travel diary collection systems fit in one of the following cases: i) the data on which the analysis / inferences is performed on was collected in other studies and for other purposes, in which case the data collection details are not emphasized (*11*), ii) the data were collected for the study but the collection is not central and, as such, not emphasized (*17*), and iii) the data were collected for the study and the collection is central and emphasized in the study (*29*). However, even for the case studies that focus on data collection, the description of the strategy used for data collection is on the current state of the system, with little emphasis on the incremental updates that preceded the current state. This has the potential disadvantage of researchers trying to integrate technologies that have already been tested and found unfit, but the process has not been documented. Furthermore, the lack of continuation of research related to such travel diary collection systems makes it difficult to understand which systems are still being developed and improved, and which systems were only developed for the case study and interrupted.

This paper presents the evolution of the MEILI system during three case studies over a period of 3 years (2013, 2014 and 2015). The focus of the paper is analyzing the quality of the collected data (both raw, GPS data, and processed, annotated travel diaries), and the overall user experience. The analysis is presented comparatively between each of the case studies.

3. METHODOLOGY EMPLOYED FOR THE CASE STUDIES

The methodology relevant for the case studies has been documented and discussed previously (*16, 20, 30–33*) and, as such, this section presents a brief description on the employed methodology and specifies which publications contain more details. The methodology relevant for this paper touches on the architecture of MEILI – the system designed to collect travel diaries –, the strategy used for raw data collection, the annotation of trajectories into travel diaries, and the methods used to compare travel diaries collected by using different methods.

The MEILI system is designed as a typical, three-tier, Model-View-Controller that has two types of clients: a data collection component and an annotation component. First, the data collection component collects movement information (in the form of GPS trajectories) from a user's smartphone in a seamless and battery efficient fashion. The battery efficiency collection is due to two main strategies: disabling the GPS collection when a user is in-doors and not moving, and dynamically adjusting the GPS receiver's frequency based on the user's current speed. For more details on data collection and battery efficiency, the reader is directed towards (*20, 30*). Second, the primary task of the data annotation component is to allow users to annotate their movement information with travel semantics (i.e., trips, triplegs, travel modes, trip destinations and purposes) and to display them. To reduce the user's burden, MEILI performs inferences about the semantics, which the user can verify and correct. For more information on the system design and the used data models, the reader is directed towards (*32*) and (*33*).

Furthermore, the MEILI system's collection capabilities are tested by organizing in-tandem surveys (using the same respondents, where users first fill in the traditional surveys and then have access to annotate their data with MEILI) that collect travel diaries using traditional collection methods, such as the modern version of paper and pen surveys, i.e., web forms that people fill in with travel details. The traditional methods are further referred to as PP methods. The travel diaries collected by both PP and MEILI are compared to identify the collection strengths and weaknesses of each system. Matching the entities from both collection methods is done based on temporal co-occurrence and same purpose constraints. For more information on how to compare different travel diary collection systems, the reader is directed towards (*31*).

The spatial and temporal indicator values proposed in (*31*) allow for the extraction of sets containing different types of collected trips (e.g., noisy trips, unreliably collected trips, etc.), out of which the most important is the extraction of a ground truth candidate set. This paper makes use of the ground truth candidate set to propose penalties that reflect the ability of each system to collect trips with regards to the temporal and spatial domains.

For the time domain, one can compute errors for the start time, end time or duration discrepancy, further referred to as error entities, as shown in Equation 1, where P_t^j represents the penalty, $S_i^j.t$ represents the capture of the error due to time domain discrepancies (start time, end time or duration discrepancy) by a trialled system S_i that is associated with a travel diary entity j, $S_i^j.T\_idx$ represents the temporal indicator value of the capture of the travel diary entity j by S_i, and δT_idx is the maximum value for which the two indicator values are regarded as similar.

$$P_t^j = \begin{cases} 0 & \text{if } \mid S_1^j.T\_idx - S_2^j.T\_idx \mid \leq \delta T\_idx, \\ S_1^j.t - S_2^j.t & \text{if } S_1^j.T\_idx > S_2^j.T\_idx + \delta T\_idx, \\ S_2^j.t - S_1^j.t & \text{if } S_2^j.T\_idx > S_1^j.T\_idx + \delta T\_idx \end{cases} \tag{1}$$

A similar penalty is the space domain penalty, which can be computed for length differences, as shown in Equation 2, where P_s^j represents the spatial penalty, $S_i^j.s$ represents the capture of the error due to spatial domain discrepancies (length discrepancy) by a trialled system S_i that is associated with a travel diary entity j, $S_i^j.S\_idx$ represents the spatial indicator value of the capture of the travel diary entity j by S_i, and δS_idx is the maximum value for which the two indicator values are regarded as similar.

$$P_s^j = \begin{cases} 0 & \text{if } \mid S_1^j.S\_idx - S_2^j.S\_idx \mid \leq \delta S\_idx, \\ S_1^j.s - S_2^j.s & \text{if } S_1^j.S\_idx > S_2^j.S\_idx + \delta S\_idx, \\ S_2^j.s - S_1^j.s & \text{if } S_2^j.S\_idx > S_1^j.S\_idx + \delta S\_idx \end{cases} \tag{2}$$

The penalties can offer more insight regarding the collection pervasiveness and accuracy of multiple systems when compared to each other.

4. CASE STUDIES

This paper is based on three case studies, which are summarized in Table 1. Since these studies are interlinked, the problems that are considered in each case study are summarized in Table 2.

The first case study (CS I) was mainly performed to test the feasibility of the battery efficient data collection using the MEILI Mobility Collector (*30*). During CS I, an initial web interface for data annotation was implemented, and an initial data model and inference methods were proposed. For CS I, no travel diaries were collected, no feedback was asked for from the participants,

TABLE 1 : Overview of the case studies conducted for improving MEILI.

	CS I	CS II	CS III
Start date	14.11.2013	29.09.2014	02.11.2015
End date	24.11.2013	05.10.2014	09.11.2015
# PP part.	N/A	42	415
# MEILI part.	11	30	171
# PP & MEILI part.	N/A	28	83
# Feedback ans.	N/A	34	303
Median age	N/A	40	42
# Raw GPS	22,000	91,000	970,000
# Annot. GPS	15,000	66,000	322,000
# Annot. triplegs	165	1,307	5,961
# Annot. trips	156	718	2,132
Observations	Small sample	Respondent bias	Large field trial

TABLE 2 : Overview of the problems treated by each case study.

Problem	Case Studies
Battery efficiency and user experience	I, II, III
Travel mode inference	I, II, III
Destination inference	I, II, III
Purpose inference	I, II, III
Travel diary comparison	II, III
Ground truth candidates	II, III

and the number of participants was low. However, CS I revealed that it is feasible to use MEILI Mobility Collector to collect data for sufficiently long periods of time.

The second case study (CS II) was designed as a pre-large-trial run to verify the feasibility of using MEILI for collecting travel diary data. During this case study, the purpose and travel mode schemas have been adapted to mimic the schemas found of the Swedish National Travel Survey. Furthermore, travel diaries were collected in-tandem with MEILI and via a declarative Paper-and-Pen interface, PP. Different methods were used to assess the quality of the collected data via either system and then analyzed how well each system can collect data. Finally, user feedback regarding battery efficiency and experience with interacting with the MEILI Travel Diary has been collected. The respondents that used MEILI and PP during CS II are either transportation engineers or work in adjacent fields, which raises the issue of respondent bias.

The third case study (CS III) was the main case study and it collected data from a large number of users during the same period of time with the official Swedish National Travel Survey collection. Compared to previous case studies, the web interface of MEILI was reimplemented, the data storage model was changed, and the user annotation effort was reduced by modifying the lists presented to users to choose their trip's destination or purpose, or their tripleg's travel mode, from an alphabetical order to a probability order, as decided via machine learning. As such, the focus of the inference methods changed from providing the most probable inference to providing

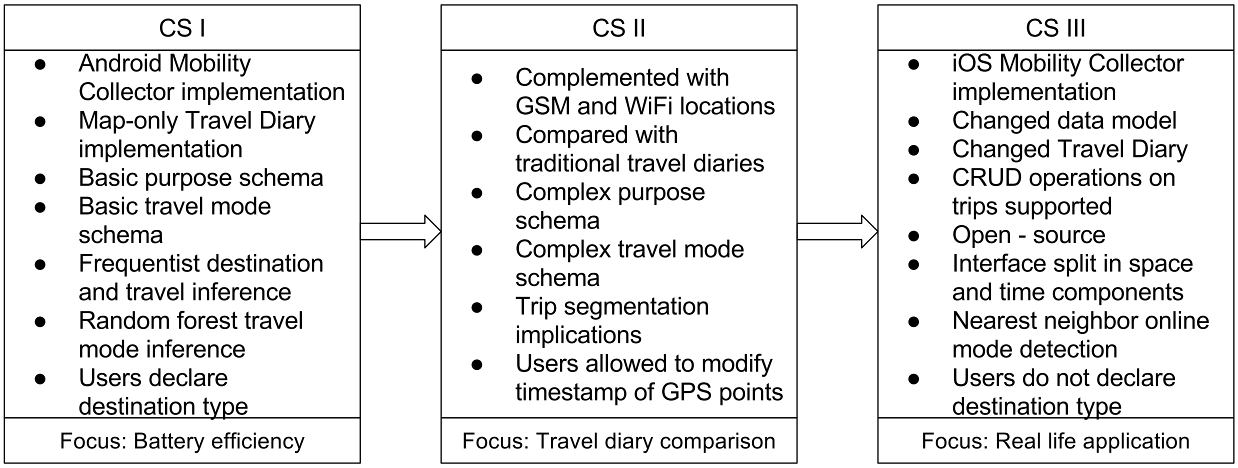

FIGURE 1 : The focus shift between the three case studies. While the initial purpose was the complete automation of the travel diary generation, CS I revealed that data can be collected in a battery efficient manner, but CS II revealed that the implications of automating the travel diary entities have not been studied sufficiently in the scientific literature. CS III was performed under the assumption of a semi-automated travel diary system, where the inferences are used to minimize user interaction with the system.

an ordered list of the schema items. CS III revealed the advantages and disadvantages of collecting travel diaries with MEILI from a non-biased user group.

4.1. Change of focus in between the case studies

This section briefly discusses the relation between the three case studies with regards to Table 1 and Table 2. The initial purpose was to develop a fully automated system that can take as input a trajectory covering the movement of a user during a predefined period and annotate it into a travel diary. However, the studies that were made for obtaining such a system (*16, 30, 31*) revealed that the studies of the automation of travel diary entities are built on assumptions that are not applicable to real life scenarios, such as having a trajectory perfectly segmented in triplegs, or a perfect trip end detection algorithm. Due to this, the initial focus of designing MEILI as a fully automated travel diary collection system, where users do not need to annotate their data, shifted to designing MEILI as a semi-automated travel diary collection system, where the users' annotation of their data is paramount. The purpose shift during the three case studies has been accompanied by major system redesigns concerning user interaction and data modelling (see Figure 1).

The most notable changes in between case studies are the schema changes for purposes and travel modes between CS I and CS II, the complemented GPS data collection with GSM and WiFi positioning between CS I and CS II, the *iOS* implementation of MEILI Mobility Collector between CS II and CS III, MEILI's web interface (i.e.,the MEILI Travel Diary) redesign between CS II and CS III, the change of employed inference methods between CS II and CS III, and the dropping of the type declaration for destinations between CS II and CS III (the type of a point of interest was used in CS I and CS II for purpose inference). The redesign of the system between CS II and CS III was performed based on the user feedback received at the end of CS II and on consulting with user design experts.

TABLE 3 : User experience in two case studies.

	CS II	CS III	Observations
Battery life			
Same as without	23.5%	61.9%	The experience of battery life
Little effect	50%	20.8%	consumption improved between
Major effect	26.5%	17.3%	the two case studies.
Ease of installation			
No problems	85.3%	86.5%	Most users do not have problems
Slight problems	8.8%	7.7%	installing the MEILI Mobility
Severe problem	5.9%	5.7%	Collector.
Annotation process			
Intuitive	55.9%	15%	The annotation user experience
Slight problems	38.2%	23%	drastically deteriorated between the
Severe problems	5.9%	62%	two case studies.

4.2. Battery efficiency and user experience

As shown in Section 4.1, the only case studies that collected user experience feedback are CS II and CS III. Table 3 shows the data collected from the users via a form designed for getting user feedback with regards to battery consumption, the difficulty of installing the MEILI Mobility Collector application and the difficulty of using the MEILI Travel Diary web page for data annotation.

One of the main differences between CS II and CS III is the additional implementation of the MEILI Mobility Collector for *iOS* devices, which enable the data collection of a larger user base. The same battery aware data collection method was implemented for *iOS*, since initial measurements showed that it outperforms the default battery efficiency considerations of tracking in *iOS* both in terms of battery efficiency and the spatial and temporal distribution of the collected GPS points. The perception the users have on the battery consumption is in general positive, and the majority of users in the larger CS III user pool do not notice any effect of having the MEILI Mobility Collector installed on their smartphone. The difference of battery consumption perception between CS II and CS III can be explained by the user sets, i.e., a highly biased transportation expert user base in CS II, and a non-expert user base in CS III.

Both user groups report few problems when installing the MEILI Mobility Collector, which can be due to the availability of the application in the official application web stores of both *iOS* (AppStore) and *Android* (Google Play).

There is a drastic shift when it comes to how easy it is for the users to annotate their trips, most CS II users reported an intuitive experience during the annotation process, and most CS III users reported several problems during the annotation process. The most probable causes are:

- the CS II user set was biased towards transportation professionals, and the CS III user set contained non-technical users
- the re-design of the web application decreased the user experience
- the occasionally occurring bugs prompted users into abandoning the case study and reporting an unpleasant experience
- users of different mobile operating systems have different user experience expectations

TABLE 4 : Overall spacing between consecutive GPS locations (in meters).

	Avg.	SD	Med.	Min.	Max.
CS I	137	577	59	0	19,515
CS II	124	1,685	58	0	242,184
CS III	98	3,397	58	0	1,286,541

TABLE 5 : An overview of the trips collected by the systems in CS II and CS III.

	#		%	
	CS II	CS III	CS II	CS III
PP Only	51	112	37.0%	24.0%
MEILI Only	44	166	31.9%	35.5%
Both	43	189	31.2%	40.5%
Total	138	467	100.0%	100.0%

4.3. Equidistance sampling

The sampling method used for data collection for all case studies is that of equidistance sampling, with the distance parameter set to 50 meters. Table 4 shows the descriptive statistics on the distance between consecutive locations. The average distance between consecutive locations is within 2-3 times the specified sampling distance when looking at the overall locations, and the median is close to the specified sampling distance. Most travel modes have both the average and median distance between consecutive GPS readings close to the distance parameter, i.e., 50 meters.

The modes whose average distance between consecutive GPS readings is larger than expected are either difficult to track due to lack of GPS coverage (e.g., subway, flight) or those modes that are used by few users, in which case the values are subjected to the regression to the mean phenomenon (e.g., bus in CS I, tram in CS II). One interesting observation is that the users annotate the readings that occur inside airports as flights, which skews the average and median values to unexpected values.

4.4. Travel diary data analysis

This section discusses the comparison between the collected travel diaries via traditional methods, i.e., PP, and via MEILI during CS II and CS III, respectively.

4.4.1. Travel diary collection and comparison

In this step, the correspondence between trips collected by MEILI and PP has been identified based on the methodology proposed in (*31*), for each of the case studies. Table 5 shows an overview of the collected trips. It is notable that the percentage of trips captured only by PP has decreased between CS II and CS III, which could suggest an improvement in data collection for MEILI. Similarly, the percentage of trips captured only by MEILI slightly increased. Overall, PP captures 68% of the trips in CS II and 64% of the trips in CS III. Similarly, MEILI captures 63% of the trips in CS II and 75% of the trips in CS III, which is consistent with the previous observation on the improvement of MEILI's data collection capabilities.

TABLE 6 : Reasons for failing to collect a trip by either system.

	PP Only		MEILI Only	
	CS II	CS III	CS II	CS III
Purpose difference	15.7%	21.4%	18.2%	14.5%
Trip chaining	17.6%	17.0%	4.5%	20.5%
No movement (MEILI) / forgot to declare (PP)	19.6%	48.2%	65.9%	65.1%
Other reasons	47.1%	13.4%	11.4%	65.1%

Since the percentage of trips captured by MEILI only or by both MEILI and PP has increased at the expense of trips captured by PP only in between CS II and CS III, it is important to investigate the plausible reasons for missing the capture of a trip. The reasons have been identified by analyzing the collected data and are summarized in Table 6, and they are analyzed in more detail in the following paragraphs.

The first reason presented in Table 6 is that of declaring a different purpose for the same trip that falls within the buffered time period as specified in Section 3. The most plausible reasons for different purpose specification are: i) multi-purpose trips, in which case the user declared one of the purpose in PP and another purpose in MEILI, ii) the perception of purpose changed once the trip was visualized on a map, and iii) the complex purpose schema confused the user.

The second reason presented in Table 6 is trip chaining, which can occur in two ways. First, in PP users merge two consecutive trips into one either unintentionally, by forgetting -, or intentionally, by disregarding- short trips with probably secondary purpose (such as shopping while on the way home). Second, in MEILI when the system fails to accurately segment a trajectory into trips and the user does not manually correct the error. The difference in between CS II and CS III, where the percentage of chained trips in MEILI only increased by 15% can be explained by the difference of expertise in between the two user groups, where CS II users are more likely to go through the extra effort of manually correcting MEILI segmentation errors.

Finally, the users might have forgotten to declare trips in PP, which results in missing trips. Similarly, the smartphone might not have recorded GPS locations for a period of time because of: i) GPS receiver errors, where fixes could not be obtained due to satellite visibility constraints, ii) not carrying the smartphone during a trip, which results in no data available for annotation, or iii) the battery efficiency algorithm prevented the GPS thread to run due to constant movement below threshold. In the lack of GPS data, users usually do not find the system friendly enough to add missing data on their own.

4.4.2. Travel diary quality assessment

The quality of the data collected by MEILI and PP is assessed by investigating the descriptive statistics of each subset of the collection, and the spatial and temporal quality indicators proposed in (*31*), as shown in Table 7.

There are two distinct types of observations that can be made on the data presented in Table 7: observations regarding the descriptive statistical summaries of travel diary entities and observations regarding the spatial and temporal indicator values of the collected entities.

TABLE 7 : Descriptive statistics on the trips gathered in the two case studies.

			PP and MEILI	PP Only	MEILI Only
Duration (min)		CS II	24±19	23±20	64±85 (20)
		CS III	25±23	26±26	173±429 (12)
Length (km)		CS II	6.3±6	4.5±5	3.8±5.1
		CS III	11.8±18	20±7	13±46
# triplegs		CS II	1.8±1	1.7±1.1	1.2±0.3
		CS III	1.9±1.3	1.6±1	1.4±0.8
Time indicator	MEILI Obs.	CS II	43%±30%	N/A	48%±30%
		CS III	62%±31%	N/A	48%±37%
	PP. Decl.	CS II	46%±32%	32%±36%	N/A
		CS III	55%±29%	47%±30%	N/A
Distance indicator	MEILI Obs.	CS II	63%±37%	N/A	73%±35%
		CS III	71%±33%	N/A	71%±35%
	PP. Decl.	CS II	68%±35%	47%±43%	N/A
		CS III	71%±34%	69%±34%	N/A

First, there are small differences between the duration of the collected trips during CS II and CS III, the most noticeable duration difference being the one concerning MEILI only trips, which can be explained by the relatively restrained variability of the characteristics of trips captured only by MEILI, where median values are low (20 minutes in CS II and 12 minutes in CS III) and the average and standard deviation values are high. One of the unexpected differences between CS II and CS III concerns the average length of trips, where all trips captured in CS III are longer than the trips captured in CS II. This can be due to the user group difference, a very narrow user group in CS II and a more representative user group in CS III, or to the difficulty encountered by users when annotating their trips after the redesign in CS II.

Second, there is a general trend in the increase of spatial and temporal quality indicators in between CS II and CS III, which is resonating with MEILI's data collection improvement between the two case studies. While (*31*) does not mention the calculation of the spatial and temporal indicators for PP only data since routes are unavailable in PP surveys, it is possible to use the start and end time of a trip as declared in PP and extract all the GPS points that fall within that period from the MEILI dataset, which allows for the computation of the indicators for the PP only data.

As explained in Section 3, the spatial and temporal indicators can be used to propose ground truth candidates and to compute spatial and temporal penalties that describe the weaknesses of each travel diary collection system, as shown in Table 8. While MEILI makes less mistakes in the time domain than PP (it has lower start time, end time and duration penalty values), it does have a higher value for the distance penalty than PP. This can be explained by the fact that MEILI captures most short trips better than PP, but the declared intervals (by users) for long trips are approximated to a comfortable time unit, and this time difference between MEILI and PP might coincide with the initial time period it takes for a GPS receiver to pick up locations after a cold fix.

TABLE 8 : The spatial and temporal penalties associated with each trip collection system.

		MEILI			PP		
		Avg.	SD	Med.	Avg.	SD	Med.
Time penalties (minutes)	Start Time	3.2	4.8	1.5	4.6	6	1.7
	End Time	2.7	4.9	0.6	4.7	5.7	2.9
	Duration	4	6.6	1.8	6.2	9.1	3.1
Distance penalty (m)	Length	1000	2900	88	700	2100	77

5. CONCLUSIONS AND DISCUSSIONS

This paper described the series of case studies that led to the proposal and improvement of MEILI, a travel diary collection, annotation and automation system. The decisions that were taken between the case studies to improve MEILI are presented in an incremental fashion, the logic behind the modifications is explained, and the evolution of the system is analyzed by analyzing travel diary data collected during the case studies.

One critical finding during the case studies regards the performance evaluation of the inference methods that are commonly being used in transportation science for travel diary entities and their attributes. The traditional methods rely on the fact that variables are independent and identically distributed, which also affects how the precision and recall measures are defined (*34*). While precision and recall measure the performance of a classifier given a labeled dataset, they do not offer enough information to assess the utility of the classifier in multiple-step processing scenarios such as travel diary extraction from trajectories. This problem is discussed in (*15*) and (*16*), and is the main reason why this paper does not present any in-depth discussion on the machine learning algorithms used by MEILI. This finding also produced a change of focus between case studies, which prompted from moving away from a fully automated travel diary collection system to a semi-automated one. This is a utilitarian change, which allows MEILI to be available for usage at the present time, given the semi-automated nature of collection, which could not have been achieved for a fully automated travel diary collection system given the time frame and resources.

The difference between the percentage of trips missed by MEILI in between CS II (37%) and CS III (24%) hints at the fact that the collection capabilities of MEILI are improving. However, the basis for this observation is a set of two case studies, which is not sufficient to derive a conclusion, but it would be noteworthy to look at these collection differences between present and future case studies. Furthermore, a more in-depth investigation should be made into whether MEILI missed trips because users forgot to carry their smartphone on a trip, or because of collection / segmentation issues in the MEILI system. While the statistics presented in Table 7 suggest the latter to be true (the low spatial and temporal indicator values for trips captured only by PP suggest that users have their smartphone with them for at least a part of their trip), it would be beneficial to try and pinpoint what are the most probable origins of the collection problems.

The case studies indicate a decrease in the user experience between CS II (56% found the system intuitive) and CS III (15% found the system intuitive). However, this observation can be challenged because of the biased user group that offered feedback in CS II, and because no in-depth interviews with users were made after the case studies to identify what constitute the problems in terms of user interface and user experience. There are difficulties with keeping users interested in annotating the data, which is can be seen in the user drop (from 171 users in the first day to 51 users in the 7th day). There are two critical improvement that can be made to MEILI: 1) improve the user interface by testing it on unbiased user sets and revising it based on feedback, and 2) propose and test different strategies for user incentives to assure that they return to annotating data.

Finally, although the MEILI system is a work in progress with several drawbacks, it is a viable option for researchers that are interested in collecting travel diaries in an inexpensive way or in the improvement of the travel diary collection capabilities of MEILI, either in a semi-automated or fully-automated way. The MEILI system is made available for free, and its source code is available at `https://github.com/Badger-MEILI`. This offers an unprecedented opportunity for researchers interested in this field for taking MEILI and modifying it to suit their needs, without having to implement their own collection system from the ground up (*32*).

6. FUTURE WORK

One of the future work directions is on improving MEILI, which was mentioned throughout the test, with an emphasis un improving the user experience and the user interface.

Another future work direction is collecting travel diaries by using multiple methods (not limited to traditional ones and MEILI) and identifying whether there is a system that can collect unbiased data, or if such data can only be collected by complementary systems. The importance of finding an answer to this question is given by the fact that most widely used models for travel behaviour or movement simulations have been built using travel diaries. If the traditional collection methods are proven to be biased, then new and more reliable models for travel behaviour and / or movement simulation would be needed.

At the same time, exploring the potential of the collected dataset to model and analyze multi-day travel behaviour is be the authors' next step. First, it is imperative to analyze groups of travelers to uncover any tendency and degree of variability and stability of travelers chosen modes, activity locations, individual's time allocations and trip chaining behaviours. Second, the collected data allows for the comparison of the theoretical routes chosen in modelling in the absence of real data (e.g., shortest path assumptions, multi-modal trip assumption, etc.) with the actual routes and modes used by travelers to perform their daily activities. Finally, the data can be used to reveal whether there are any patterns regarding how travelers make use of their time budget under a set of different constraint, e.g., spatial or temporal constraints.

7. ACKNOWLEDGMENTS

This work was partly supported by Trafikverket (Swedish Transport Administration) under Grant "TRV 2014/10422".

REFERENCES

[1] Shen, L. and P. R. Stopher, Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*, Vol. 34, No. 3, 2014, pp. 316–334.

[2] Kagerbauer, M., W. Manz, D. Zumkeller, M. Kagerbauer, W. Manz, and D. Zumkeller, Analysis of PAPI, CATI, and CAWI methods for a multiday household travel survey. *Transport Survey Methods. Best Practice for Decision Making*, 2013, pp. 289–304.

[3] Wolf, J., J. Wilhelm, J. Casas, and S. Sen, A case study: Multiple data collection methods and the NY/NJ/CT regional travel survey. *Transport Survey Methods: Best Practice for Decision Making*, 2013, p. 71.

[4] Lee-Gosselin, M. E., Scope and potential of interactive stated response data collection methods. In *Transportation Research Board Conference Proceedings*, 1996, 10.

[5] Griffiths, R., A. Richardson, and M. E. Lee-Gosselin, Travel surveys. *Transportation in the new millennium*, 2000.

[6] Chapleau, R. and C. Morency, Dynamic spatial analysis of urban travel survey data using GIS. In *25th Annual ESRI International User Conference, San Diego, California*, Citeseer, 2005, pp. 1–14.

[7] Stopher, P., C. FitzGerald, and M. Xu, Assessing the accuracy of the Sydney Household Travel Survey with GPS. *Transportation*, Vol. 34, No. 6, 2007, pp. 723–741.

[8] *Norwegian National Travel Survey 2013/2014*, 2015, found at `http://nsddata.nsd.uib.no/webview/index/en/Individdata/Other-classifications.d.243/Datasets-which-are-part-of-series.d.142/Travel-Surveys.d.33/Norwegian-National-Travel-Survey-2013_2014/fStudy/NSD2163`. Accessed: 2016-07-30.

[9] *Swedish National Travel Survey*, 2012, found at `https://www.kth.se/polopoly_fs/1.310699!/Menu/general/column-content/attachment/1.The%20problems%20with%20the%20current%20travel%20survey%20in%20Sweden_Andreas%20Holmstr%C3%B6m_Trafikanalys.pdf`. Accessed: 2016-07-30.

[10] *United Kingdom National Travel Survey 2014*, 2014, found at `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457688/nts2014-notes.pdf`. Accessed: 2016-07-30.

[11] Axhausen, K. W., S. Schönfelder, J. Wolf, M. Oliveira, and U. Samaga, 80 weeks of GPS-traces: Approaches to enriching the trip information. *83rd Transportation Research Board meeting*, 2003.

[12] Wolf, J., *Using GPS Data Loggers To Replace Travel Diaries In the Collection of Travel Data*. Dissertation, Georgia Institute of Technology, School of Civil and Environmental Engineering, Atlanta, GA, 2000.

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

[14] Safi, H., B. Assemi, M. Mesbah, and L. Ferreira, A Trip-detection Method for Smartphone-assisted Travel Data Collection. In *Transportation Research Board (TRB) 95th Annual Meeting*, 2016.

[15] Zheng, Y., L. Liu, L. Wang, and X. Xie, Learning Transportation Mode from Raw Gps Data for Geographic Applications on the Web. *Proceedings of the 17th International Conference on World Wide Web*, 2008, pp. 247–256.

[16] Prelipcean, A. C., G. Gidófalvi, and Y. O. Susilo, Measures of transport mode segmentation of trajectories. *International Journal of Geographical Information Science*, Vol. 30, No. 9, 2016, pp. 1763–1784.

[17] Biljecki, F., L. Hugo, and P. van Oosterom, Transportation mode-based segmentation and classification of movement trajectories. *International Journal of Geographical Information Science*, Vol. 27, No. 2, 2013, pp. 385–407.

[18] Shafique, M. A. and E. Hato, Travel Mode Detection with Varying Smartphone Data Collection Frequencies. *Sensors*, Vol. 16, No. 5, 2016, p. 716.

[19] Feng, T. and H. J. Timmermans, Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data. *Transportation Planning and Technology*, Vol. 39, No. 2, 2016, pp. 180–194.

[20] Prelipcean, A. C., *Capturing travel entities to facilitate travel behaviour analysis: A case study on generating travel diaries from trajectories fused with accelerometer readings*. Dissertation, KTH Royal Institute of Technology, Stockholm, Sweden, 2016, ISBN 978-91-7595-958-0.

[21] Bohte, W. and K. Maat, 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*, Vol. 17, No. 3, 2009, pp. 285 – 297.

[22] Deng, Z. and M. Ji, Deriving rules for trip purpose identification from GPS travel survey data and land use data: A machine learning approach. *Traffic and Transportation Studies*, Vol. 2010, 2010, pp. 768–777.

[23] Li, M., J. Dai, S. Sahu, and M. Naphade, Trip Analyzer Through Smartphone Apps. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, New York, NY, USA, 2011, GIS '11, pp. 537–540.

[24] Shen, L. and P. R. Stopher, A process for trip purpose imputation from Global Positioning System data. *Transportation Research Part C: Emerging Technologies*, Vol. 36, 2013, pp. 261–267.

[25] Montini, L., N. Rieser-Schüssler, A. Horni, and K. W. Axhausen, Trip Purpose Identification from GPS Tracks. *Journal of the Transportation Research Board*, 2014, pp. 16–23.

[26] Oliveira, M. G. S., P. Vovsha, J. Wolf, and M. Mitchell, Evaluation of Two Methods for Identifying Trip Purpose in GPS-Based Household Travel Surveys. *Transportation Research Record: Journal of the Transportation Research Board*, , No. 2405, 2014, pp. 33–41.

[27] Kusakabe, T., T. Tsubota, and A. Bhaskar, Validation Study of Naïve Bayes Probabilistic Model for Transit Passengers' Trip Purpose Estimation: Case Study Exploiting Detailed Brisbane Household Travel Survey Data. In *Transportation Research Board 95th Annual Meeting*, 2016, 16-2539.

[28] Wang, A. H., Detecting spam bots in online social networking sites: a machine learning approach. In *IFIP Annual Conference on Data and Applications Security and Privacy*, Springer, 2010, pp. 335–342.

[29] Wolf, J., R. Guensler, and W. Bachman, Elimination of the Travel Diary: An Experiment to Derive Trip Purpose From GPS Travel Data. *Transportation Research Board: Journal of the Transportation Research Board*, , No. 1768, 2001, pp. 125–134.

[30] Prelipcean, A. C., G. Gidófalvi, and Y. O. Susilo, Mobility collector. *Journal of Location Based Services*, Vol. 8, No. 4, 2014, pp. 229–255.

[31] Prelipcean, A. C., G. Gidófalvi, and Y. O. Susilo, Comparative framework for activity-travel diary collection systems. In *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, IEEE, 2015, pp. 251–258.

[32] Prelipcean, A. C., Y. O. Susilo, and G. Gidofalvi, MEILI: A Travel Diary Collection, Annotation and Automation System. *Submitted to Journal of Transport Geography*, 2016.

[33] Susilo, Y. O., A. C. Prelipcean, G. Gidófalvi, A. Allström, I. Kristoffersson, and J. Widell, Lessons from a trial of MEILI, a smartphone based semi-automatic activity-travel diary collector, in Stockholm city, Sweden. *Proceedings of World Conference on Transport Research – WCTR 2016 Shanghai*, 2016.

[34] Tong, S. and D. Koller, Support vector machine active learning with applications to text classification. *Journal of machine learning research*, Vol. 2, No. Nov, 2001, pp. 45–66.