Evaluating FMS
A preliminary comparison with a traditional travel survey

Carlos Carrion, Francisco Pereira, Rudi Ball, Fang Zhao, Youngsung Kim, Kalan Nawarathne*
Naijia Zheng †
Chris Zegras, Moshe Ben-Akiva ‡

1 August 2013

Abstract

This paper presents compares two different travel survey instruments that were administered to the same set of users, in the context of the 2012 travel survey data collection effort in Singapore. The Household Interview Travel Survey (HITS) of 2012 follows the traditional static *paper based* approach while the Future Mobility Survey (FMS) is a smartphone based travel survey being developed in Singapore and MIT. The paper describes the FMS technology as well as the survey implementation and its relationship with HITS. Moreover, we discuss the data collected and compare the survey results of 244 participants who completed both the HITS and FMS surveys. Overall, participants collected 739 days of data. Using the data, we discuss both the successes and challenges experienced with the two approaches.

Keywords *travel survey, smartphone, data analysis, comparison.*

Word Count: *3638*
Word Count incl. figures and tables: *6888*

*Future Urban Mobility IRG, Singapore-MIT Alliance for Research and Technology, Singapore, carlos@smart.mit.edu, camara@smart.mit.edu, rudi.ball@smart.mit.edu, fang.zhao@smart.mit.edu, youngsung@smart.mit.edu, kalan@smart.mit.edu

†National University of Singapore, a0091829@nus.edu.sg

‡Massachusetts Institute of Technology, Cambridge, MA, USA, czegras@mit.edu, mba@mit.edu

1 Introduction

The rise in the availability and use of location-enabled devices, particularly smartphones, has greatly expanded the means of collecting various forms of transportation data. While traditional self-reported travel surveys typically suffer from problems such as limited sample size, the under-reporting of total completed trips, an imprecision of trip start and end times (4), location-enabled smartphone based surveys present the opportunity to collect more detailed and precise data needed for emerging agent and activity-based behavioral models. Developments in this field (1, 3) suggest that location-enabled technologies can reduce the number of erroneous 'no travel' days and missed trips; improve accuracy of reported trip times, locations and paths; and reduce respondent burden.

The usage of mobile technologies for automatic surveying is not new. Traditionally, GPS-based logging surveys have been widely implemented worldwide and largely successful as a supplement to household travel surveys (2, 8). However, pure GPS logging suffers from some limitations. Financially, the agencies conducting travel surveys are required to purchase and distribute the GPS collection devices, which can be a significant investment. Also, the participants may forget to carry the GPS loggers with them for the duration of the travel survey, and they will face a recollection problem when completing their travel diary. In contrast, smartphones provide some clear benefits. For instance, users are accustomed to carrying their phones with them constantly and as such, there is a decreased likelihood of missing trips. They are almost always charged, and smartphones contain a combination of sensors not limited to positioning data. The sensors are capable of providing spatial, temporal and proximity data, which can be used to infer activity and mode information. These attributes make smartphones ideal 'life-loggers'.

In this paper, we present preliminary results of data collected between the Future Mobility Survey (FMS) and the Household Interview Travel Survey (HITS) of 2012, where FMS is a smartphone based approach and HITS a more traditional paper based survey. Our comparison set of 387 participants used both survey methods, first they filled the HITS questionnaire, then ran FMS on their phones for at least 14 days. The FMS is currently being field tested as a collaborative project between the Singapore-MIT Alliance for Research and Technology (SMART) and the Land Transport Authority of Singapore (LTA).

This paper is organized as follows: HITS and FMS (covers overview of both surveys, socio-demographic statistics of the sample, and data preparation for the sample), evaluation (presents a discussion comparing the HITS and FMS for distinct statistics related to activity and mobility), challenges (discusses main difficulties for FMS), and conclusions.

2 HITS and FMS

HITS and FMS present two alternative approaches in the collection of survey data. In this section we briefly outline the approaches, particularly enhancing their differences, and we discuss the socio-demographics of the comparison sample.

2.1 Overview

HITS is a paper based survey conducted within Singapore every four years. It is currently on the final stages for the 2012 effort. It is conducted by the Land Transport Authority of Singapore (LTA). The survey collects activity and mobility data for a typical weekday (Monday to Friday) for an individual. It also collects the socio-demographic characteristics of the households, and the individuals. The data is collected through face to face interviews. A local subcontractor is responsible for the recruitment and interviewing of participants.

The format of the survey follows the standard trip diary-based approach. Travel was defined as a one-way journey completed for a purpose. Walking or bicycle trips which connect other methods of travel are

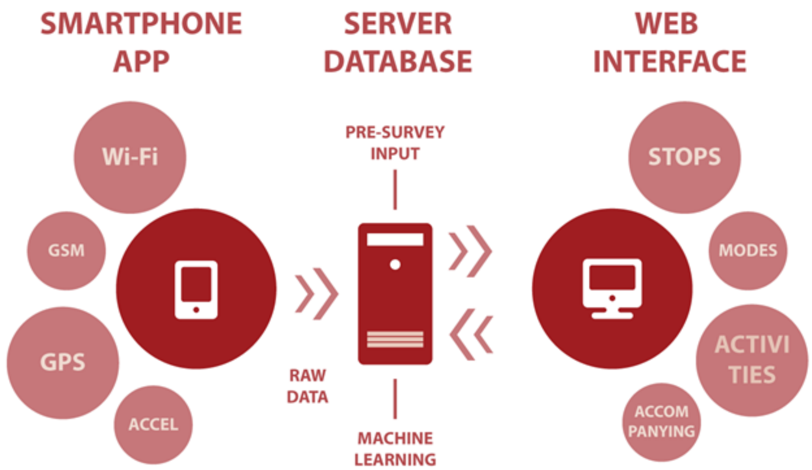

Figure 1: The FMS architecture.

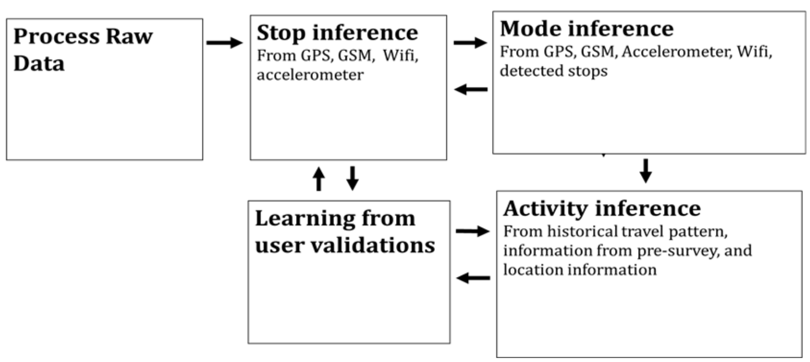

Figure 2: FMS flowchart.

also recorded for outbound and return trips. The survey includes only walks longer than 100 meters taken as part of a trip (e.g. walking to a bus stop), and walking trips before or after a trip with at least one motorized mode (e.g. walking to work, and leaving work using a taxi). Walking trips from home to work or school are recorded for workers and students. Furthermore, the sample size of this survey is targeted for roughly 10,500 households. This is about 1% compared to the total household number Singapore, which is roughly 1 million.

The HITS 2012 follows a similar format, methodology, and objective to other metropolitan-wide travel surveys found in the Metropolitan Travel Survey Archive (http://www.surveyarchive.org).

The FMS survey (6) is comprised of a smartphone application, available both for Android and iOS, a web-interface (front-end) and an analytics back-end (Figure 1). The long-term aim of the FMS smartphone application is that it silently collect data without user intervention. Participants would therefore not be influenced in any way by the smartphone application during their normal day. Spatial and temporal data is initially collected by a participant's smartphone. The data collected by the smartphone is mapped, filtered and analyzed by the back-end. This data is then analyzed (Figure 2) and presented to the user via a web-interface (Figure 3). The user is required to validate data collected via the web-interface. Validation entails that a participant check that their activity and mobility patterns accurately represent their actual activity

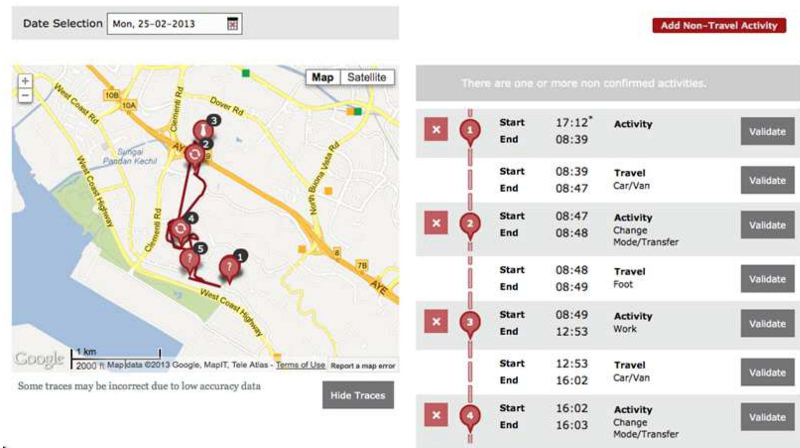

Figure 3: The FMS web-interface.

patterns for a specific day. Thus far, participants have been required to provide at least 14 days of collected data, of which at least 5 has to be validated in order to receive a monetary incentive. As the process of validation is fully dependent on user's choice, there is a need to provide clear *prompted-recall* visual and textual cues to remind him/her of the details of those past activities and travel. To reduce the burden placed on participants, analytics algorithms have been employed.

Figure 2 illustrates the progression of data between successive analytics components. Back-end components transform logged raw data into understandable information for review by the survey participant and end-users. Algorithms are particularly focused on inferring stops, modes and activities. The algorithms used in each step vary. The 'process raw data' step consists of a series of scripts for cleaning, composing and temporally aligning the incoming data for use in the subsequent analysis steps. The 'stop inference' applies a rule-based algorithm in two phases: first, it matches spatial/temporal windows (7) to the data to obtain candidate stops; subsequently, it uses wifi and GSM data to merge stops, particularly using accelerometer information to detect 'still' periods (where, although the GSM is 'jumping', the user should stay in the same place). It also uses past validation information to match users recurrent locations (e.g., home, work) with GSM signatures and adds/removes stops based on mode detection results (for example, there must be a stop for change mode/transfer between any two different modes). The 'mode inference' step applies a machine-learning algorithm, support-vector machine or SVM (5), to accelerometer and GPS data to identify the mode out of the set of car, bus, subway, walk, bicycle or motorbike. Finally, the 'activity inference' matches the historical data, namely the previous validations, to current stops to identify recurrent locations. Current development of this module also considers contextual information such as the Points of Interest or the mode interchange areas. The goal of the 'learning from user validations' step, under development, is to systematically update these algorithms in time, i.e. perform online learning.

2.2 Data

2.2.1 Data description

At the moment of the writing of this paper, 387 individuals participated in both HITS and FMS surveys. Data collection is still ongoing. A sample of 244 participants provided both weekday (Monday to Friday) and home-to-home travel data. These data comprise 739 days of 24 hour data i.e. 24 hours of travel and activity data. 143 participants were excluded as they either did not travel or they had collected insufficient data for comparison. As the number of days provided by FMS was larger than that provided by HITS, the HITS data of a participant was weighted by the number of days corresponding to the participant in FMS. Also, it should be noted that the days of collected data for both surveys are not the same.

Table 1: Socio-demographic statistics

			Sample	Singapore[a]
Gender	Male	129	53%	49%
	Female	115	47%	51%
Age	0 - 15 years	0	0%	17%
	15 - 19 years	11	5%	7%
	20 - 29 years	61	25%	14%
	30 - 39 years	69	28%	16%
	40 - 49 years	69	28%	17%
	50 - 59 years	28	11%	15%
	60+ years	6	2%	14%
Personal Monthly Income	No income	61	25%	-
	$1 - $1999	18	7%	-
	$2000 - $2999	30	12%	-
	$3000 - $3999	20	8%	-
	$4000 - $4999	15	6%	-
	$5000 - $5999	15	6%	-
	$6000 - $7999	11	5%	-
	$8000+	8	3%	-
	Refused	66	27%	-
Household Size	1	6	2%	12%
	2	23	9%	19%
	3	44	18%	20%
	4+	171	70%	49%
Employment Type	Full-time	156	64%	-
	Part-time	11	5%	-
	Self-employed	12	5%	-
	Full-time Student	31	13%	-
	Others	34	14%	-
Fixed workplace	Yes	164	67%	-
	Yes from home	9	4%	-
	No	6	2%	-
	No response	65	27%	-

[a] Statistics gathered from Statistics Singapore (www.singstat.gov.sg) on 1 August 2013.

Table 1 presents the socio-demographic characteristics of the comparison sample. Most participants are between the ages of 20 and 50 years. Another characteristic is that most household sizes are 4+. Lastly, most participants are full-time workers with fixed workplaces. The sample is similar to the population of Singapore in terms of gender distribution, and it is dissimilar in terms of household size distribution (sample has a higher proportion of participants from household size 4+). In addition, the sample has a low proportion of participants of 60 years or more.

2.2.2 Data preparation

While smartphones are capable of collecting increasingly accurate data - it is still possible that some geolocation points may be erroneous due to sensory errors and/or a limited sample of points. For example, GPS location accuracy is reduced when participants travel indoors. Another common problem occurs where participants' smartphones run out of battery power or devices are turned off. Such scenarios have the possibility of producing gaps in data (presenting non-continuous data). A *data flagging* process was applied to data validated by participants prior to data analysis. The process involved the application of logical checks and heuristics to identify unsatisfactory records and participants which did not meet the minimum requirements necessary for inclusion in analysis. These logical checks included:

- **temporal checks:** determine if the start and end times of activities are consistent and whether data is continuous between activities;
- **spatial checks:** determine the location of activities and their relative distance as compared with points of interest such as bus and train stations;
- **speeding checks:** flag activity sequences where a participant's mode of transport does not accurately reflect their resultant mobility. For example, a participant reached a destination faster than physically possible, given the reported mode of transport.

If data was deemed to fail a logical check, then the particular record was flagged. If a single flag was found in any participants' data pertaining to a day, then that day was excluded from the analysis.

3 Evaluation

This section presents a descriptive statistical analysis focusing on (i) activity and (ii) mobility related statistics of the participants. For the analysis, the authors present the following activity related statistics: time-scale distributions of activities for the HITS (only weekdays) and FMS (weekdays and weekends); and cumulative probability for the hours spent at work. For the mobility related statistics, the authors include: the cumulative probability for the total hours spent traveling; and the distribution of the number of trips per day.

Figure 4 presents the time-scale distribution of activities for HITS aggregated for week days. Most participants begin their day at home. Most participants transition to work between 06H00 and 07H00. The transition is reversed between 17H00 and 22H00. We observe out of home activities between 06H00 and 22H00. The home and work time scale distributions are asymmetric. Other activities are not as clearly defined.

In contrast, Figure 5 presents the time-scale distribution of activities for FMS aggregated for week days. We observe a similar set of activity patterns, however we observe a more refined transition from home to work. Other activity patterns are more clearly defined than in HITS. In addition, relationships between activity patterns are notable. For example, the relationship between work and meal-eating can be observed between the hours of 12H00 and 14H00. In these hours work decreases, while Meal Eating Break increases. Moreover the peaks of Meal Eating Break are more resolute in FMS rather than HITS. Another difference is the Education activity - HITS Education is reportedly higher than that of FMS. We observe that out of home activities occur between 06H00 and 23H00.

Figure 6 presents the cumulative probability of the hours spent at work. This sample excludes non-working days. In the HITS, participants report more work hours than in FMS. The range provided FMS is larger than that provided by HITS. It should be noted that participants may be over-reporting HITS work hours because they may add their meal eating break. This was previously discussed in Figures 4 and 5.

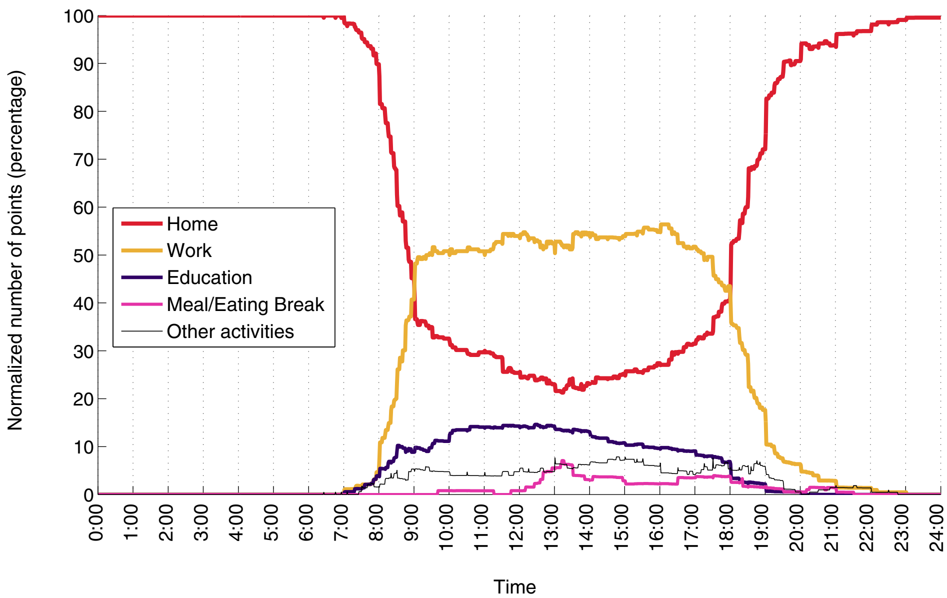

Figure 4: HITS: Week day hours activity distribution (Home, Work, Education, Meal/Eating Break and Other Activities).

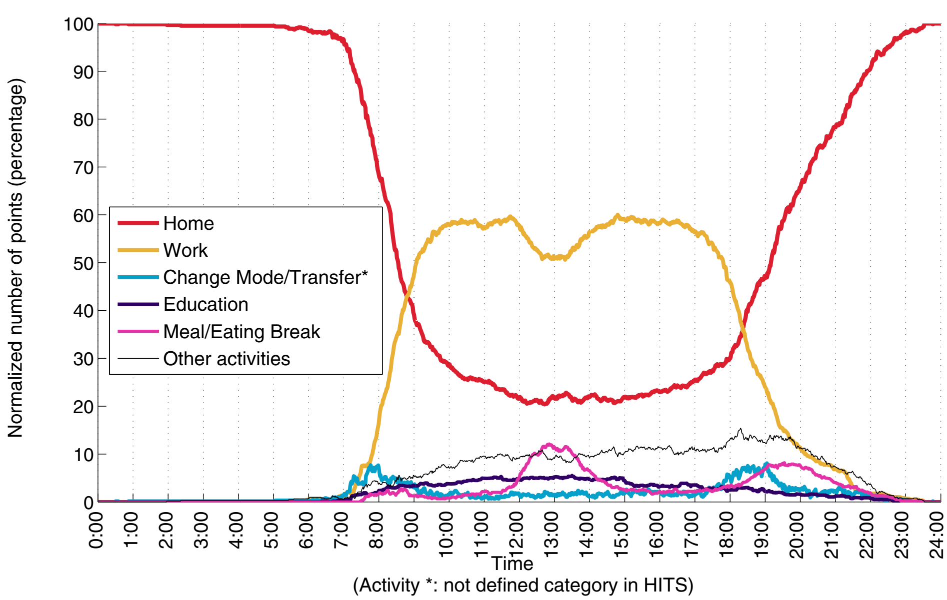

Figure 5: FMS: Week day hours activity distribution (Home, Work, Change Mode/Transfer, Education, Meal/Eating Break and Other Activities).

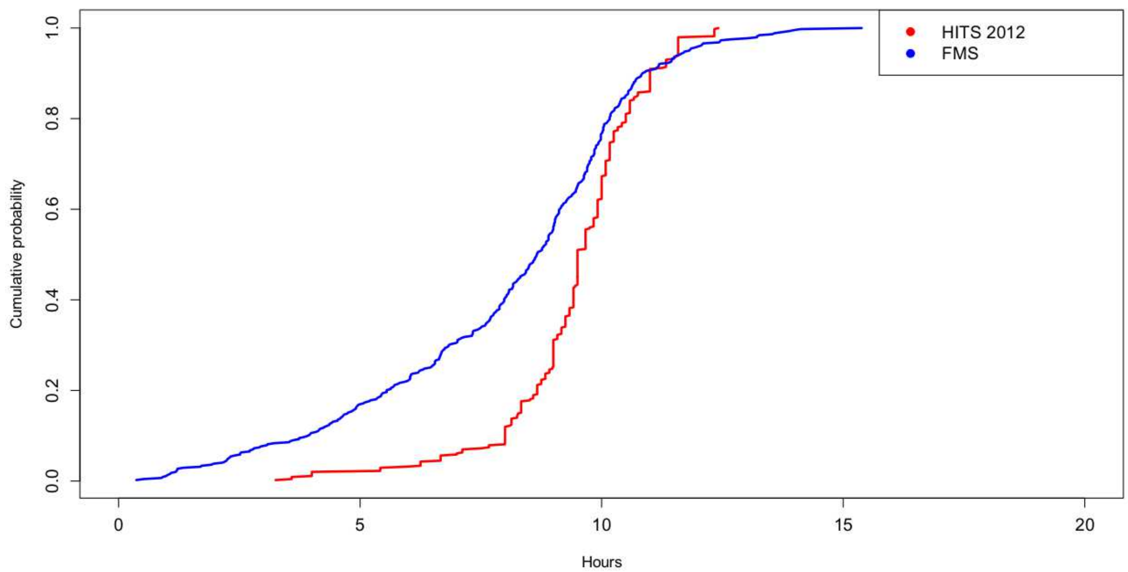

Figure 6: Total hours spent at work per day

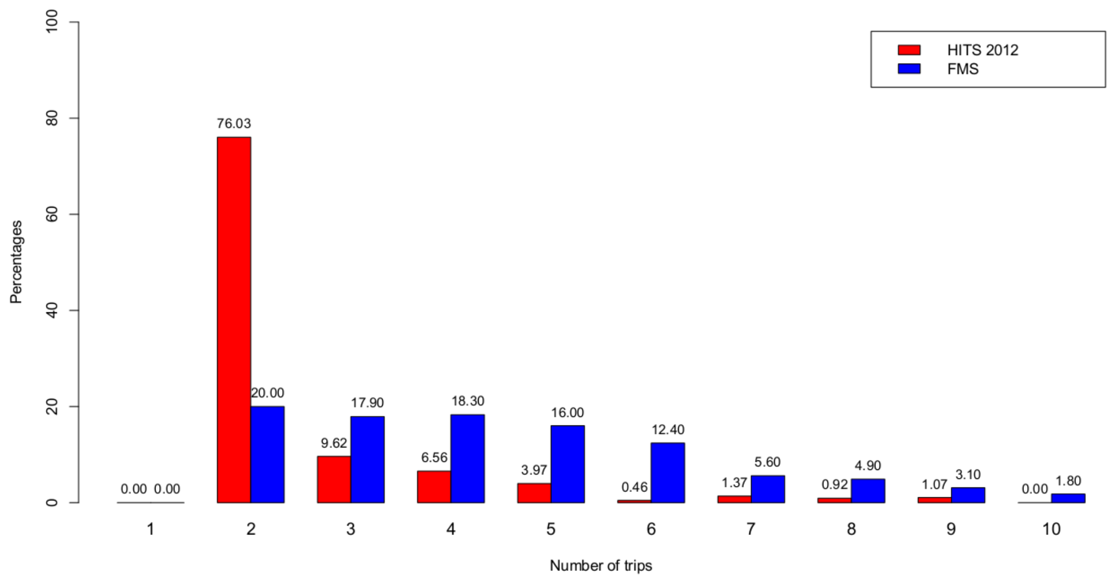

Figure 7: Total number of trips per day.

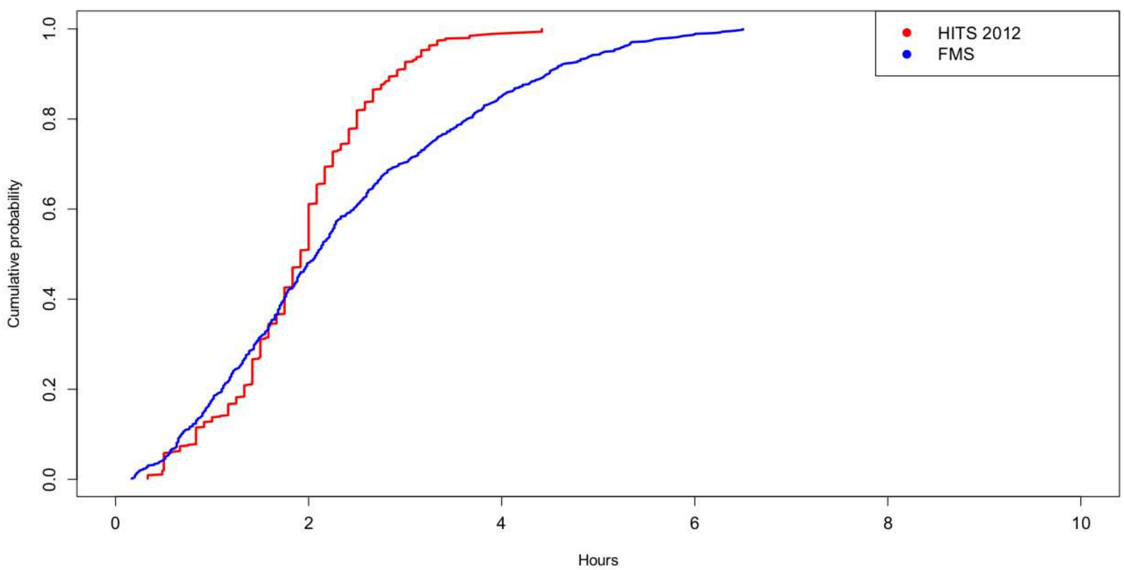

Figure 8: Total hours spent traveling per day

Finally, FMS is able to collect all days including weekends. HITS is focused on the collection of typical week days. Figure 9 indicates a clear difference of the activity travel patterns of the participants compared to Figure 5.

Figure 7 presents the distribution per day of the total number of trips for HITS and FMS. FMS collects more trips than HITS. It is interesting to note that most participants report 2 trips per day in the HITS (76%). In contrast, most FMS participants report the number of trips to be between 2 and 6 trips per day.

Figure 8 presents the cumulative probability per day of the total time spent traveling. We define this as the sum of the travel times across all modes between activities per day - this includes access times, egress times and waiting times. The FMS range is larger than that of HITS. The total time spent traveling is shorter than FMS before approximately 1.8 hours and significantly greater after that. The difference may be related to the perception of the participants (among other factors), because the travel times are reported by participants in the HITS. In FMS the smartphone device records the travel times.

4 Challenges

A constraint of FMS relates to how it affects the battery life of the user's smartphone. This may require the user to recharge the device more often than usual to avoid running out of power in undesired moments. Both Android and iPhone applications were designed to minimize such burden but the technological constraints are particularly limiting, especially considering the heterogeneity of hardware, phone usage habits, charge habits or battery health. Thus, the answer to the common question about FMS battery load has no trivial answer. However, we can provide general statistics about battery usage from our dataset.

In Figure 10, we show the observed hourly battery decay rate (in percentage). This measure is more meaningful than, for example, average daily autonomy, because it is not affected by initial battery level, or intermediate charging. In any case, it is still sensitive to phone usage, i.e. we could not measure battery

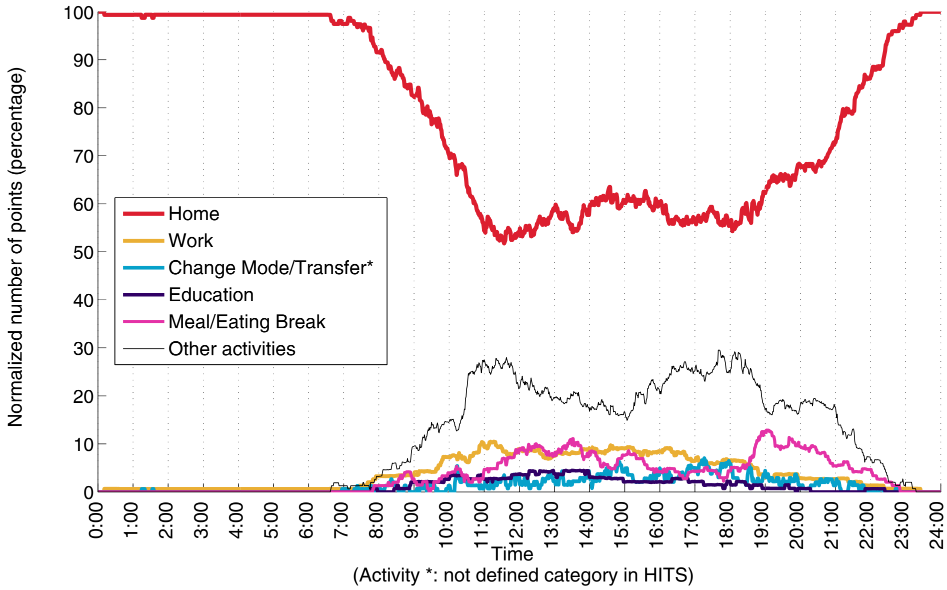

Figure 9: FMS: Weekend hours activity distribution (Home, Work, Change Mode/Transfer, Education, Meal/Eating Break and Other Activities).

usage exclusively from FMS. It would demand individualized phone calibration (e.g. measuring battery consumption without any FMS data collection for a few minutes). Even with such a calibration we could not control for battery usage from other applications during FMS execution.

There is noticeable difference between Android and iOS. For Android, we can see a much higher concentration on low hourly consumption, with mode on 5% per hour (equivalent to 20 hours from 100% to 0%, assuming linear decay) while iPhone will have a much fatter tail at the higher hourly consumptions, with mode on 6% per hour. iOS is well known to be more constrained than Android in terms of low-level programming API (application interface), and this has been affecting the capability to optimize the software. Fortunately, latest versions of iOS are becoming more flexible so the team is improving the iOS FMS application as we write this document.

We recall that, besides GPS, FMS also logs GSM, WiFi and Accelerometer information. In fact, it also does so to save battery life: it periodically switches off GPS, and collects only these lower precision data, a procedure we call *phased sampling*. In this way, the stop detection procedure may be affected but not ruled out. Taking the validated stops by the users as the *ground truth*[1], the true positive rate of the algorithm is 95.5% and the false negative rate is 4.5%. This is achieved with a relatively small proportion of GPS data collection times (see Figure 11).

The detailed analysis of battery consumption, sampling strategies, and their implications to data quality in FMS, is an extensive topic that demands an article by itself, therefore we defer further discussions to subsequent publications. For the purpose of this paper, suffice it to say that our applications were generally able to achieve reasonable battery consumption, but the heterogeneity of software, hardware and usage profiles makes it a serious technological constraint for some users.

Another relevant limitation concerns to the almost complete absence of personal interactions related to FMS. After the invitation and quick introduction to FMS by the recruiter, the participant is left alone to

[1]We know this is a strong assumption, as the users may be resistant to make corrections.

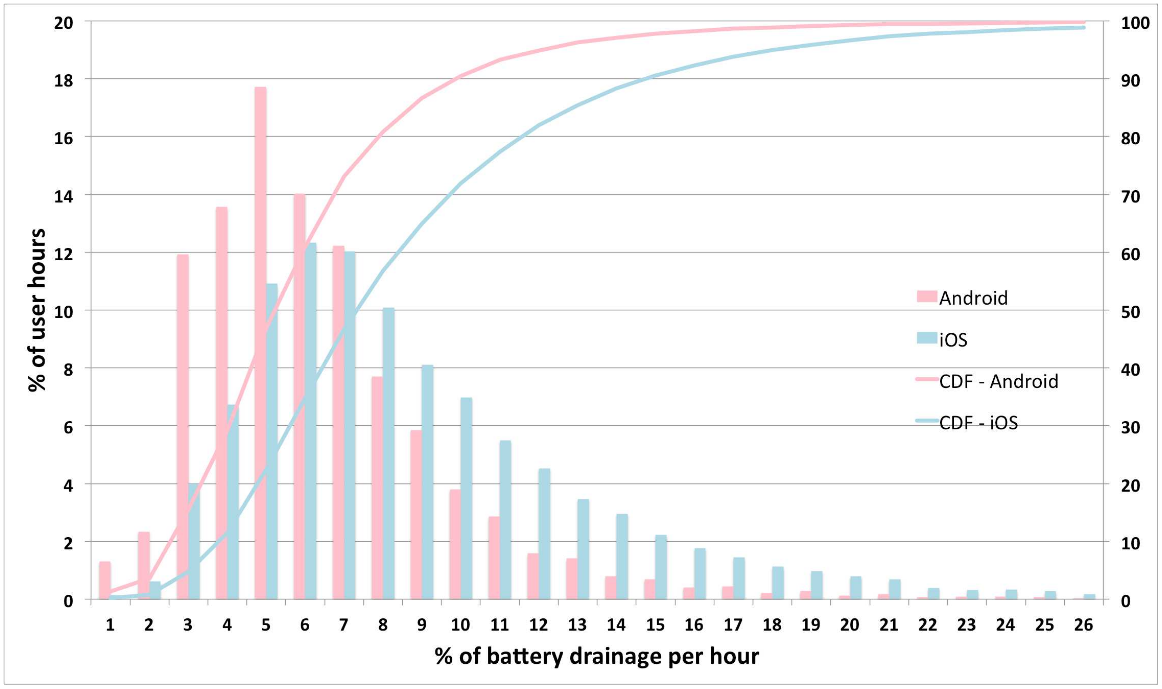

Figure 10: Distribution of battery performance per hour for the dataset, for Android and iOS users.

follow the work-flow summarized in Figure 12. This raises two important challenges: understanding the technology usage; understanding the task itself. While the former is facilitated by the tech-savviness of the user and, in contexts like Singapore, becomes less of a concern with a high smartphone penetration rate, the latter is entirely dependent on the user interface, both on the web client as well as on the phone.

To help users overcome the initial steep learning curve, we equipped FMS with a remote desktop platform called firefly®that allows a helpdesk assistant to remotely visualize the participant's interaction on his/her computer. This works on a single session basis and can only happen after user's explicit request and authorization. Despite the process being seamless and intuitive, this resource was not used as much as was desired and the team was only able to help about 20 users overall.

The web interface (Figure 3) was repeatedly redesigned and user tested throughout several iterations, and trailed in two pilot runs. As with the case of battery management, the heterogeneity of the participants' population makes it a particularly complex task. Some users prefer to interact with the map while others prefer otherwise; icons are generically intuitive, but some users prefer text; font size and available screen space also varies among users hardware. Overall, the details required for an activity diary for transportation modeling are not immediately understood by the layman. When they exist at all, these challenges are overcome in HITS through the interaction between the recruiter and the user.

A final limitation to mention relates to recruitment self-selection bias. Table 1 confirms the intuition that this type of survey finds difficulties in the even distribution across socio-demographic characteristics. This may relate not only to technology access and savviness as well as to attitude towards institutional surveys, privacy and trust concerns, relevance of the incentive, time availability to name a few.

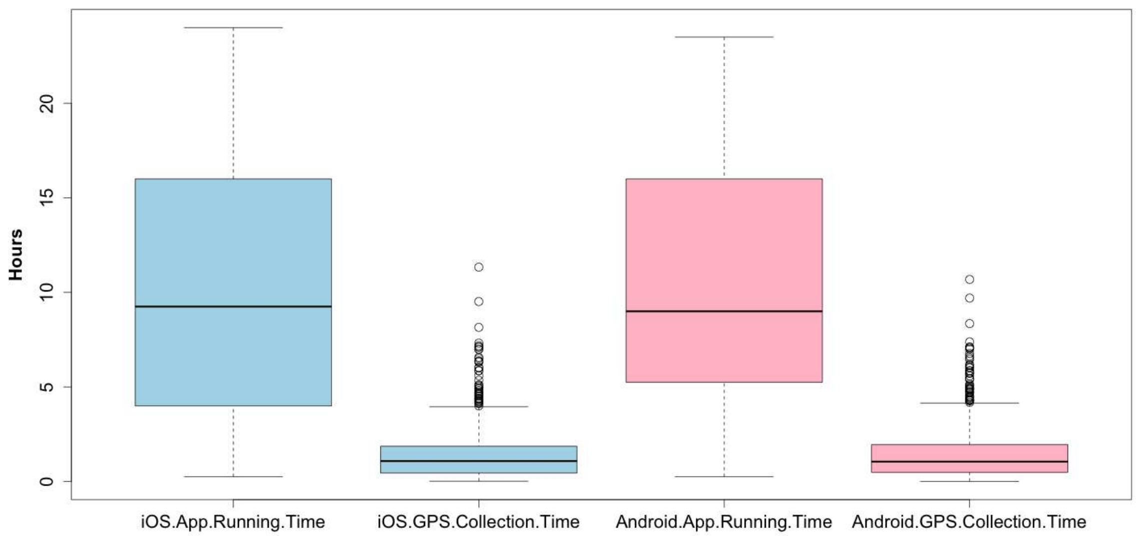

Figure 11: GPS an non-GPS data collection times, for iOS and Android.

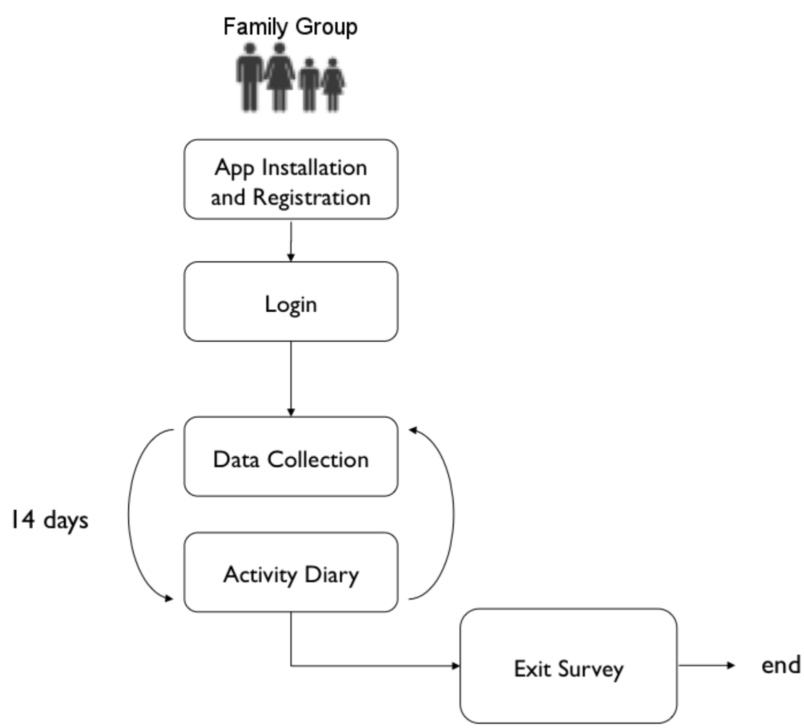

Figure 12: FMS workflow.

5 Conclusion

In this paper we have presented preliminary results for common HITS and FMS survey participants. The paper evaluated 739 days of data collected by 244 participants. We began by outlining the difference between the surveys, discussing the socio-demographic statistics of the comparison sample. In the following sections we compared various distributions related to the activity and mobility of the same participants from HITS and FMS. The distributions highlight several key differences such as: FMS captures more trips than HITS; FMS provides travel times as measured from smartphone devices; FMS has more detailed activity patterns within the day; and others. In addition, we note that both HITS and FMS present a number of trade-offs. HITS represents a sample for one weekday, but benefits from the addition of a human interviewer. FMS is a more lengthy process, but the inclusion of a smartphone in the survey improves the accuracy and broadens the range of data collected. FMS data is more resolute than that provided by HITS. FMS is dependent on the user validating their travel patterns sufficiently well.

Future work follows two directions. Firstly, the authors are implementing an analytical framework for comparing the HITS and FMS sample presented in this paper through econometric modeling. This framework features include: controlling the temporal effect (i.e. participants provide data for different days for HITS and FMS); modeling the attrition rate (i.e. participants choose the number of days to validate); panel periods (multiple days of data for FMS); and others. Secondly, we seek to improve the capabilities and accuracy of validated FMS data as well as further reduce the battery costs associated with FMS data collection on smartphones. The ubiquity of advanced technologies in the mobile environment reveals great potential for expanding data collection methods. Taking advantage of such potential, however, will require careful attention to competing needs.

Acknowledgements

The research described in this project was funded in whole or in part by the Singapore National Research Foundation (NRF) through the Singapore-MIT Alliance for Research and Technology (SMART) Future Urban Mobility (FM) program. We thank the FM group for their feedback.

References

[1] Auld, J. and Mohammadian, A. (2009). Framework for the development of the agent-based dynamic activity planning and travel scheduling (adapts) model.

[2] Bohte, W. and Maat, K. (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.

[3] Bricka, S. and Bhat, C. R. (2006). Comparative analysis of global positioning system-based and travel survey-based data. *Transportation Research Record: Journal of the Transportation Research Board*, 1972(1):9–20.

[4] Chen, C., Gong, H., Lawson, C., and Bialostozky, E. (2010). Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the new york city case study. *Transportation Research Part A: Policy and Practice*, 44(10):830–840.

[5] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

[6] Cottrill, C. D., Pereira, F. C., Zhao, F., Dias, I. F., Lim, H. B., Ben-Akiva, M., and Zegras, P. C. (2013). Future mobility survey: Experience in developing a smartphone-based travel survey in singapore. In *Transportation Research Board 92nd Annual Meeting*, number 13-4849.

[7] Hariharan, R. and Toyama, K. (2004). Project lachesis: parsing and modeling location histories. In *Geographic Information Science*, pages 106–124. Springer.

[8] Stopher, P., FitzGerald, C., and Xu, M. (2007). Assessing the accuracy of the sydney household travel survey with gps. *Transportation*, 34(6):723–741.