

# Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore

Fang Zhao (Corresponding Author) (fang.zhao@smart.mit.edu),  
Francisco Câmara Pereira (camara@smart.mit.edu),  
Rudi Ball (rudi@smart.mit.edu),  
Youngsung Kim (youngsung@smart.mit.edu)  
Singapore-MIT Alliance for Research and Technology, Future Urban Mobility  
1 CREATE Way, #09-01/02 CREATE Tower, Singapore  
138602  
Tel: (65) 6601-1547, Fax: (65) 6778-5654

Yafei Han (yafei@mit.edu),  
Christopher Zegras (czegras@mit.edu)  
MIT, Department of Urban Studies & Planning  
77 Massachusetts Avenue, Room 10-403, Cambridge MA 02139  
Phone: (617) 452-2433, Fax: (617) 258-8081

Moshe Ben-Akiva (mba@mit.edu)  
MIT, Department of Civil & Environmental Engineering  
Room 1-181, 77 Massachusetts Avenue, Cambridge, MA, 02139  
Telephone: (617) 253-5324

Submitted 30 March 2015 for publication in *Transportation Research Record*

Word count with figures and tables:  $4429 + 12 * 250 = 7429$

## **ABSTRACT**

Future Mobility Sensing (FMS) is an innovative smartphone-based travel survey system that was field tested in 2012/2013 together with the Household Interview Travel Survey (HITS) in Singapore. This paper presents findings of exploratory analysis of the data collected in this test. Clustering of day patterns from FMS data reveals large day-to-day variability of user behavior, which cannot be captured by a snapshot with a one-day survey. Even the high cross-sectional variability from the larger sample size of a traditional survey would not have achieved the comprehensive set of heterogeneous patterns as provided by FMS. Some common problems in traditional surveys – such as under-reporting of trips, over-estimation of travel times, and inaccuracy of locations and times – can apparently be reduced by FMS. The FMS data, as compared to HITS, have higher resolution and better accuracy. In addition, FMS is well suited to collect multi-day data as additional costs are marginal and user burden reduces over time. FMS offers a promising technology for next generation travel data collection.

## INTRODUCTION

The unprecedented and increasing penetration rate of smartphones together with advances in mobile sensing technology have greatly expanded the means of collecting various forms of personal transportation data. Traditional self-reported travel surveys typically suffer from problems such as limited sample size, under-reporting of total completed trips, imprecision in reported trip start and end times (1). 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 (2, 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. GPS-based logging surveys have been widely implemented worldwide and largely successful as a supplement to traditional household travel surveys (4, 5, 6, 7). However, pure GPS logging suffers from some limitations. Financially, the agencies conducting travel surveys must 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 may still 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, with adequate battery life, thus decreasing the likelihood of missing trips. Furthermore, smartphones contain a combination of sensors – such as Bluetooth, WiFi, accelerometers, and GPS – which expand upon pure positioning data, providing a richness which can be used to infer activity and mode information. These attributes make smartphones ideal “life-loggers.”

The Future Mobility Sensing (FMS) system capitalizes on these “life-logging” capabilities in a next-generation travel survey system, leveraging increasingly pervasive smartphones, advanced sensing and communication technologies and a machine learning architecture (8). With a web-based prompted recall user interface, FMS delivers a previously unobtainable range of data, more closely reflecting what people do, not what they say they do. FMS was field-tested in Singapore in conjunction with the Singapore Land Transport Authority’s (LTA’s) Household Interview Travel Survey (HITS) 2012. The pilot recruited more than 1500 users and produced a large set of rich and detailed travel/activity data, validated by the respondents (9). This paper reports on the exploratory analysis of this unique dataset, demonstrating the capabilities of this kind of survey platform to reveal interesting and diverse user day patterns and overcome some of the known issues of traditional travel surveys.

The remainder this paper has five parts. Section 2 gives an overview of the FMS system. Section 3 describes the field test with the HITS survey in Singapore and the data collected. Section 4 presents an example to illustrate the difference between the data collected in HITS and FMS. Section 5 presents the exploratory analysis results and the final section concludes.

## FMS SYSTEM

FMS consists of three separate, but inter-connected, components: the smartphone app that collects the sensing data; the server that includes the database as well as the data processing and learning algorithms; and the web interface that users access to view and validate the processed data and answer additional questions to supplement the validated data. Figure 1 shows three components and the data flows among them.

### **Smartphone App**

The smartphone app, available for both Android and iOS platforms, collects data from a multitude of the phones' sensors, including GPS, GSM, accelerometer, and WiFi. A major design objective of the FMS app is non-intrusiveness, i.e., the app runs in the background of the phone, silently collecting sensor data without user intervention. This aims to minimize the app's influence on participants during their normal daily activities. In addition, the application is designed to be lightweight (in terms of memory use), easy to use, and energy efficient, using various approaches to minimize battery consumption (10), a major concern for location-based applications. The sensor data collected on the phone are transferred to the back-end server through either the cellular network or WiFi, based on the user's preference.

### **Backend Server**

Raw data collected via the app are uploaded to a database where a series of algorithms are used to process the data and make inferences about stops, travel modes and non-travel activities (11). To minimize the user's interaction burden, the backend algorithms translate raw data into trips and activities. The first round of stop detection is made based on location and point-of-interest (POI) data. GSM, WiFi and accelerometer information are used to merge stops that would otherwise be interpreted as distinct stops. Travel modes are detected based on GPS and accelerometer features, as well as public transit location information. Short duration stops that are unimportant from a data validation standpoint (such as stops in traffic) are deleted for the purposes of presentation in the web interface. Travel destinations (e.g., home, work, shopping, drop-off) are also inferred based on previous validations by the user, POI data, and other contextual information.

### **Web-interface**

The web interface provides a platform that enables users to review and "validate" their processed data in the form of a daily timeline or activity diary (Figure 2). Validation involves filling in missing information and amending incorrectly inferred data about modes of travel used for particular trips and specific activities engaged in at inferred "stop" locations (destinations). The validated data are uploaded and the algorithms learn from the user validations to subsequently make better inferences. The website is flexibly designed to enable supplementary data collection, such as information pertaining to a specific trip (e.g., how many people the user traveled with or what, if any, fee was paid for parking), during the activity diary validation stage. The LTA pilot involved a helpdesk made available to users through a web-chat or phone call; users were encouraged to have a session with a helpdesk representative for assistance during their first data validation.

### **FMS FIELD TEST**

Between October 2012 and September 2013, FMS was field-tested together with Singapore LTA's HITS 2012 survey. HITS is a traditional household travel survey conducted quadrennially in Singapore since 2004. The survey collects activity and mobility data for a typical weekday (Monday to Friday) for individuals. It also collects the socio-demographic characteristics of the households and the individuals. The data are collected through face-to-face interviews, carried out by a local subcontractor also responsible for recruitment. The format of the survey follows the standard trip diary-based approach, with travel defined as a one-way journey completed for a purpose. The survey includes walk segments 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). Walk-only trips are recorded if they are longer than 10 minutes. HITS 2012 aimed for a sample size of roughly 10,500 households, approximately 1% of Singapore's resident population households. The HITS 2012 follows the format, methods, and objectives similar to other metropolitan-wide travel surveys.

The FMS recruitment process piggybacked on that for HITS. After a HITS interview, the surveyor introduced, and invited the participant to take part in, FMS. Unlike in HITS, which required participants to register in complete households, FMS users could join as individuals, to increase the participation rate in the pilot. An FMS participant is considered to have completed the FMS survey after collecting at least 14 days of data and validating at least 5 days of them. Of the 1,541 recruited users, 793 completed the FMS survey. The pilot implementation collected a total of 22,170 user-days, with 7,856 of them validated by the users.

As the FMS pilot required smartphone ownership and familiarity with web-browsers, participants were expectedly biased towards a younger population (Figure 3). In future implementations, this can be rectified by distributing GPS loggers or smartphones to users who do not have smartphones and by providing validation help (over the phone or in person) for less tech-savvy users. Such approaches fit naturally into the FMS platform and do not require additional development work. Some evidence suggests that younger and highly educated people may be more inclined to participate in web-based travel surveys while traditional methods may work better for older adults (12). Therefore, FMS may provide a good complement to traditional travel surveys for achieving a balanced overall sample.

Despite the volume and relative spatiotemporal precision of smartphone-based data collection, some geo-location points may be erroneous due to sensor errors and/or limitations. For example, GPS location accuracy is reduced when participants travel indoors. Close examination of the collected data revealed two main types of errors:

1. **Data gaps.** Data gaps may occur when participants' smartphones run low on battery power, the user logs off from the app, or devices are turned off. The loss of continuous data can hamper a user's ability to fully reconstruct her day, despite the possibility to add activities during validation.
2. **Validation errors.** Unlike in traditional surveys, where an interviewer is present to perform some quality control during data collection, FMS' web-based validation is totally controlled by the user. Mistakes can happen especially when the user is unfamiliar with the interface.

Various real-time checks were tested for the collected and validated data – including maximum gaps in data to allow for validation, ranges of realistic travel speeds, activity durations, changes of location, as compared to collected raw data – to minimize the above errors. In addition, post-processing was performed to further clean the collected data.

## **AN ILLUSTRATIVE EXAMPLE**

A comparison of HITS and FMS data collected from the same user highlights some of the properties of FMS data. The travel/activity information of the user is presented in two formats: a map showing her stops, activities at each stop, and her traces; and, a timeline below the map, showing the times associated with each activity.

The HITS data (Figure 4 (a)) suggest that the user left home for work at around 7am and returned home before 7pm. On both ways, she travelled by bus; since HITS does not capture the exact routes, the Figure shows likely bus routes based on Google Maps. Notably, for both trips, the reported travel times are exactly the same, 35 minutes. In fact, HITS reveals many cases where users apparently round the travel times to the nearest 5 or 10-minute blocks, with a large spike in travel time of 60 minutes for bus trips. In this example, the user reported a simple working day, suggesting considerable data uncertainty due to lack of detailed information. Figure 4(b) shows a much richer set of data for the same user collected in FMS on four weekdays. In the first one, the user only worked half-day, and went out for dinner in the evening. FMS captured the exact bus routes taken. On the second day, although the activity sequence is the same as for HITS (Home-Work-Home), the work hours are very long, from around 7am to past 9pm. Also, she apparently took different bus routes to and from work, taking 18 and 20 minutes, respectively. The third FMS map reveals, she did not go to work on this day, instead, carrying out sports activities in the morning, and social/meal activities in the afternoon. The fourth figure reveals yet another day pattern, with shopping in the morning and work in the afternoon/evening. In fact, in these four days, none of her trips to or from work took more than 20 minutes. Of course, for the day she reported in HITS, the traffic conditions may have been exceptionally bad, but the results likely reflect, at least partially, the fact that people tend to over-estimate travel times (5).

This example reveals some typical issues with traditional travel surveys:

- people tend to report a simple (typical) day;
- short activities are under-reported;
- travel times are over-estimated for short trips; and
- people have large day-to-day variability, which cannot be captured by a one-day survey.

FMS shows the potential to overcome some of these shortcomings through the capability for collecting more detailed and accurate data and capturing the variability in user's travel/activity patterns across days. The next section examines the overall FMS dataset to demonstrate these points in more detail.

## DATA ANALYSIS

The exploratory analysis uses post-processed and cleaned data from a subset of 319 FMS users, who have, on average, a larger number of validated days. With a total number of 2350 days, we have on average 10.5 validated days from each user, useful for studying intra-user travel/activity pattern variability.

### Clustering of User Day Patterns

Analysis of user's day-to-day behavioral variability is based on clustering FMS user day patterns for the 233 employed (full-time or part-time) participants. The day pattern is generated by dividing each day into 5-minute slots and assigning to each slot its associated activity (including travel). This transforms each user day into a vector of 288 elements. A weighted clustering algorithm is used to group the day patterns into 5 clusters.

Figure 5 shows the day patterns belonging to each cluster. Cluster 1 contains working days with lunch breaks at mid-day and many non-work activities after work. Cluster 2 has days of mainly just working, with few other activities. Cluster 3 has shorter work hours and many work-related activities. No work shows up in Cluster 4, which has many non-work activities,

likely for weekends or off-days. Finally, Cluster 5 represents days primarily spent at home or traveling. As most errors in FMS data are related to data gaps, which can lead to excessive travel times (since all non-stop segments were marked as traveling) or home activities (as home stops before and after a gap were merged, assuming users were home all the time), many of the days with errors fall into Cluster 5. The relationship between days of the week and clusters (Table 1) is consistent with expectations. Clusters 1, 2 and 3 are mostly on weekdays, whereas Cluster 4 is mainly on weekends. Cluster 5 also has a higher concentration on weekends, but the day patterns with errors throughout the week also contribute to the high number of days in this cluster.

Focusing only on weekdays provides a glimpse at people's day pattern variability for days consistent with HITS. Figure 6 shows the number of users having distinct day pattern clusters across weekdays. Almost all users (97%) have at least two clusters and a majority of them (73%) have at least three distinct day patterns. Traditional surveys, like HITS, typically only sample one or two days, which prevents the identification of an individual's day pattern heterogeneity. Linking the day patterns generated by these 223 users' HITS data to the 5 FMS-derived clusters shows that 208 of them (93%) correspond to Cluster 2, with simple work tours. This is consistent with the example from the previous section. In fact, none of the 223 users' HITS days fall into Cluster 1, supporting the hypothesis that HITS under-reports out-of-work activities.

### **Distance-from-Home Diagrams**

Distance-from-home diagrams visualize the travel/activity pattern of a user over multiple days, showing a user's distance from his/her home across the time covered by the survey. Markers represent non-home activities the user performed at each stop. If the activity lasts for a long time, the marker is placed at the start time of the activity. These graphs reveal interesting patterns for different types of users. Figure 7 shows three user examples, representing different socioeconomic and demographic backgrounds.

Figure 7(a) is the distance-from-home diagram for a full-time employed user who works long hours. Although she goes to work almost everyday, large day-to-day variability appears in her work hours and in the other activities she performs on these days. Also, there are large variances in her time spent in each type of activity and each mode of transportation, which cannot possibly be captured by a one-day survey.

Figure 7(b) shows data collected by a part-time employed user. He goes to work on some days, and performs other activities on the others. His main modes of transportation are car and trains, with a large variation in the travel times.

In contrast, a very different activity pattern can be seen in Figure 7(c) for a homemaker. One marked difference between her data and the previous two users is that she typically makes multiple short home-based tours per day. Again, a one-day snapshot would not capture the full spectrum of her patterns.

### **Time of Travel**

Figure 8 compares the time of travel for different activities captured in HITS and FMS for weekdays, plotting the percentage of users traveling for Work, Home, or Meal/Eating Break at different times during the day. For travel to work and to home (Figures 8 (a) and (b)), as expected, most of the trips take place around the morning and evening peak hours. In these two graphs, HITS shows a narrower travel distribution than FMS, supporting the hypothesis that people tend to report a "typical" day in self-reported surveys, when in reality travel times have a

wider spread. In addition, in HITS most users report arriving home in the evening by 8pm, while, in fact, a significant portion of the users reaches home after 9pm, indicating under-reporting of activities towards the end of the day.

Figure 8(c) shows the time of travel to Meal/Eating-Break activities. Three clear peaks emerge in the FMS curve, corresponding to trips to breakfast, lunch and dinner. However, this trend is not clear from HITS data. In fact, the lower values for the HITS curve indicate a much smaller percentage of users reporting Meal/Eating Break activities in HITS. This is also reflected in Table 2, which lists the top three purposes of home-based tours reported in HITS and FMS for users of different employment status. For all categories of users, except those self-employed, Meal/Eating Break activity is among their top three purposes in FMS. In contrast, this activity only appears once in the list for HITS. This supports the observation that short trips tend to be under-reported in self-reported travel surveys. On the other hand, the rank of Pick-up/Drop-off activities appears to be lower in FMS. This is due to the fact that these activities may have very short durations, and the modes of transportation before and after the stops are typically the same. When this kind of stop is detected in FMS, it is normally deleted in order to reduce the erroneous detection of “stops” related to traffic lights or congestion. The detection accuracy will improve in this aspect with more Points of Interest (POI) information incorporated in the underlying map and improved algorithms for better learning from users’ past validations.

### **Travel Time by Mode**

As the analysis of travel time is more sensitive to errors in the collected data, a subset of 1100 days from the 319 users were selected and manually checked to ensure data quality. This dataset is used in this and the next subsection.

Figure 9 compares the histograms of travel times by car in HITS and FMS. While the distribution of travel time in FMS is smooth and continuous, the HITS graph is more rugged, with spikes at 20 minutes, 30 minutes, and 60 minutes. As we know, survey participants tend to round their travel time estimations. The resulting data may not reflect the true shape of the distribution of travel times, and may pose problems when used in modeling. FMS data, on the other hand, derives the real “continuous” travel times from sensory data. Moreover, the average travel times recorded in FMS are lower than those in HITS, which again shows that people’s perceived travel times tend to be longer than reality, as mentioned. Similar issues emerge for public transportation modes, such as bus or train.

Figure 10 shows the box plots of travel times for each mode in FMS. The box shows the range of values between the first and third quantile, while the top and bottom whiskers represent  $1.5 \times \text{IQR}$  (Inter-Quantile Range) from the upper and lower quantiles, respectively. On average, trips by train (MRT/LRT) are the longest, and those by foot or bicycle are the shortest. Also, a large number of walk trips are recorded in FMS, which are not registered in HITS.

## **CONCLUSIONS AND FUTURE DIRECTIONS**

Exploratory analysis of the data collected by FMS in Singapore demonstrates its capability to produce detailed and rich data for travel surveys. It aims to sense what people do, rather than asking them to report what they do. This eliminates many problems that traditional self-reported surveys face, such as under-reporting of short trips, reporting inaccurate locations and times, and reporting on a “typical” day rather than the actual day. Also, large intra-user day-to-day variabilities in the travel/activity patterns have been observed across different types of users, a picture of day patterns which is inadequately captured in the snapshot provided in a

one-day survey. An advantage of FMS is that the marginal cost of collecting additional days of data is minimal, and, in fact, decreases over time: as a user provides validations and familiarizes herself with the web-interface, the user burden declines; at the same time the FMS backend learns from the validation, improving inferences from the data and further reducing validation burden. Smartphone based travel surveys, such as FMS, are now viable and will likely prove to be superior alternatives to traditional approaches.

The Singaporean field test of FMS provided a unique, large, and rich dataset on individual travel behavior. This paper provides an initial exploratory analysis only hinting at the possibilities for better understanding behavioral patterns and influencing factors. Additional research could aim at discovering and classifying user patterns over multiple days, performing high resolution space-time analysis, and associating travel behavior with other context information, such as weather, events, incidents, etc. The authors are also implementing an analytical framework for comparing the HITS and FMS samples presented in this paper through econometric modeling. This framework's features include: controlling possible temporal effects (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); implementing panel periods (multiple days of data for FMS); and others.

At the same time, the capabilities of the FMS system can be enhanced, especially to reduce user burden and improve data quality – goals which can have competing needs, and require careful consideration in design and development. Two key aspects of reducing user burden are battery consumption and the user-friendliness/intuitiveness of the interface. Improving data quality can partly be achieved with machine-learning techniques based on context information and user history. Ultimately, however, it remains to be seen, whether, and by how much, data such as that produced by FMS can be used to actually improve models and their capabilities for analyzing system interventions.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Land Transport Authority of Singapore for their support and help throughout this field test. This research was supported by the National Research Foundation Singapore through the Singapore MIT Alliance for Research and Technology's Future Urban Mobility IRG research program.

## REFERENCES

1. Chen, C., H. Gong, C. Lawson, and E. Bialostozky, 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*, Vol. 44, No. 10, 2010, pp. 830–840.
2. Auld, J. and A. Mohammadian, Framework for the development of the agent-based dynamic activity planning and travel scheduling (ADAPTS) model. *Journal of Transportation Letters: The International Journal of Transportation Research*, Vol. 1, No. 3, 2009, pp. 245–255.
3. Bricka, S. and C. R. Bhat, Comparative analysis of global positioning system-based and travel survey-based data. *Transportation Research Record: Journal of the Transportation Research Record*, Vol. 1972, No. 1, 2006, pp. 9–20.
4. 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.
5. 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.
  6. Oliveira, M. G. S., P. Vovsha, J. Wolf, Y. Birotker, D. Givon, and J. Paasche, Global Positioning System-Assisted Prompted Recall Household Travel Survey to Support Development of Advanced Travel Model in Jerusalem, Israel. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2246, 2011, pp. 16–23.
  7. Stopher, P. and L. Wargelin, Conducting a household travel survey with GPS: reports on a pilot study. In *12th World Congress on Transport Research*, Lisbon, Portugal, 2010.
  8. Cottrill, C., F. C. Pereira, F. Zhao, D. Ines, H. B. Lim, M. Ben-Akiva, and C. C. Zegras, Future Mobility Survey: Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Record*, Vol. 2354, 2013, pp. 59–67.
  9. Carrion, C., F. C. Pereira, R. Ball, F. Zhao, Y. Kim, K. Nawarathne, N. Zheng, M. Ben-Akiva, and C. C. Zegras, Evaluating FMS: a preliminary comparison with a traditional travel survey. In *Transportation Research Board 93rd Annual Meeting*, 2014.
  10. Nawarathne, K., F. Zhao, F. C. Pereira, C. C. Zegras, and M. Ben-Akiva, The impact of GPS based outdoor activity detection on smartphone battery life. In *10th International Conference on Transport Survey Methods*, Australia, 2014.
  11. Zhao, F., A. Ghorpade, F. C. Pereira, C. C. Zegras, and M. Ben-Akiva, Stop detection in smartphone-based travel surveys. In *10th International Conference on Transport Survey Methods*, Australia, 2014.
  12. Christensen, L. The Role of Web Interviews as Part of a National Travel Survey. In *Transport Survey Methods: Best Practice for Decision Making* (J. Zmud, M. Lee-Gosselin, eds.), Emerald Group Publishing Limited, Bingley, UK, 2013, pp. 115-153.

## List of Tables

- 1 Breakdown of user days in each cluster and each day of week
- 2 Top three purposes of home-based tours

## List of Figures

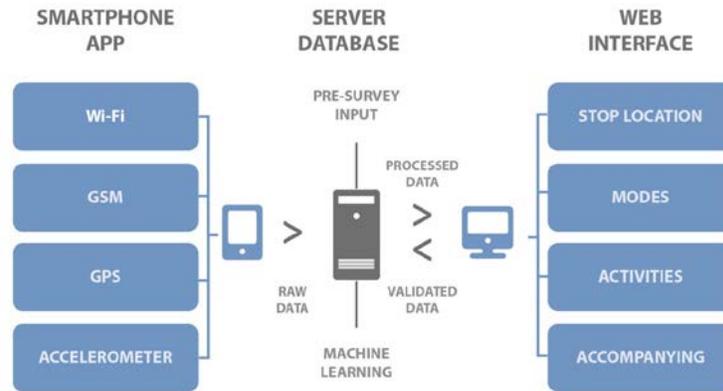
- 1 FMS architecture
- 2 FMS web-interface – activity diary
- 3 Age distribution of FMS users and Singapore population
- 4 HITS and FMS data from one user on different days
- 5 Results of clustering FMS user day patterns into 5 clusters (employed users)
- 6 Distribution of users with different numbers distinct clusters
- 7 Distance-from-home diagram for (a) a full-time employed user who works long hours; (b) a part-time employed user; and (c) a homemaker
- 8 Percentage of participants traveling to (a) Work, (b) Home, (c) Meal/Eating Break at different times of day.
- 9 Histogram of trip times by car for HITS and FMS
- 10 Boxplot of travel times for each mode in FMS

**TABLE 1 Breakdown of User Days in Each Cluster and Each Day of Week**

| Clusters (Day Pattern)                               | Mon. | Tue. | Wed. | Thur. | Fri. | Sat. | Sun. | <b>Total</b> |
|--|------|------|------|-------|------|------|------|--------------|
| Cluster 1 (work, lunch, after work activities)       | 113  | 110  | 112  | 121   | 94   | 26   | 11   | 587          |
| Cluster 2 (work only)                                | 86   | 89   | 85   | 85    | 66   | 14   | 9    | 434          |
| Cluster 3 (shorter work, many after work activities) | 17   | 28   | 28   | 33    | 20   | 7    | 6    | 139          |
| Cluster 4 (many non-work activities)                 | 39   | 38   | 48   | 28    | 52   | 145  | 132  | 482          |
| Cluster 5 (home)                                     | 95   | 88   | 83   | 90    | 94   | 106  | 137  | 693          |

**TABLE 2 Top Three Purposes of Home-Based Tours**

|                           | <b>HITS</b>                                       | <b>FMS</b>   |
|---------------------------|---|--|
| <b>Employed Full-time</b> | Work<br>Pick-up/Drop-off<br>Work-related          | Work<br>Meal/Eating Break<br>Personal Errand       |
| <b>Employed Part-time</b> | Work<br>Pick-up/Drop-off<br>Shopping              | Work<br>Meal/Eating Break<br>Personal Errand       |
| <b>Self-employed</b>      | Work-related<br>Work<br>Pick-up/Drop-off          | Work<br>Personal Errand<br>Shopping                |
| <b>Homemaker</b>          | Pick-up/Drop-off<br>Shopping<br>Meal/Eating Break | Meal/Eating Break<br>Shopping<br>Personal          |
| <b>Full-time student</b>  | Education<br>Shopping<br>Work                     | Education<br>Work<br>Meal/Eating Break             |
| <b>Retired</b>            | Other's Home<br>Pick-up/Drop-off<br>Social        | Meal/Eating Break<br>Personal Errand<br>Recreation |



**FIGURE 1 FMS architecture.**

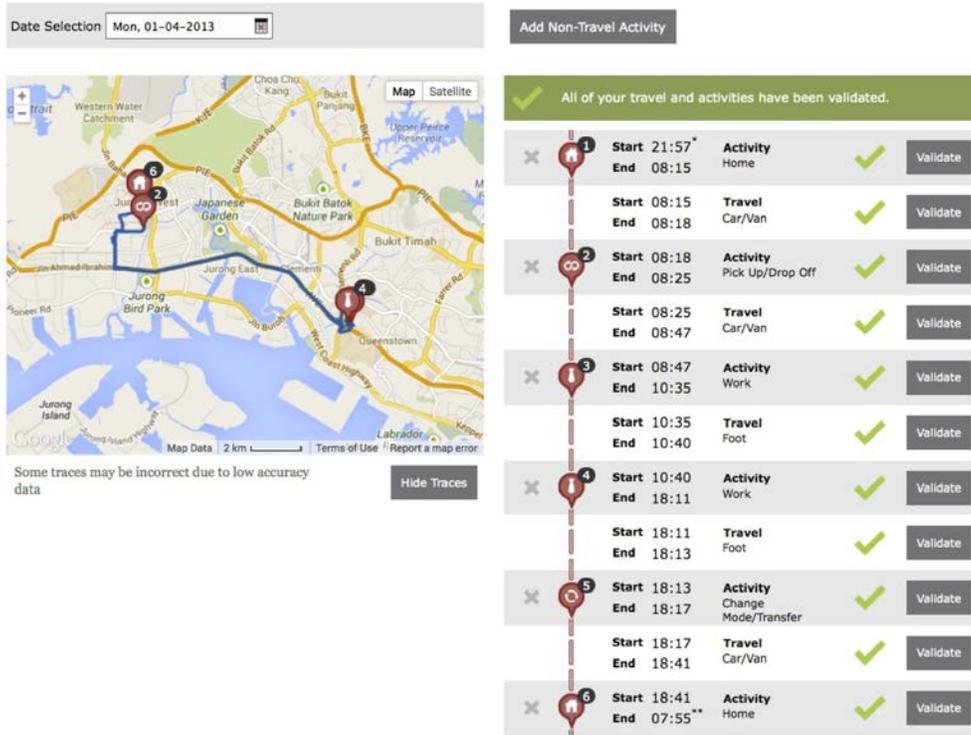
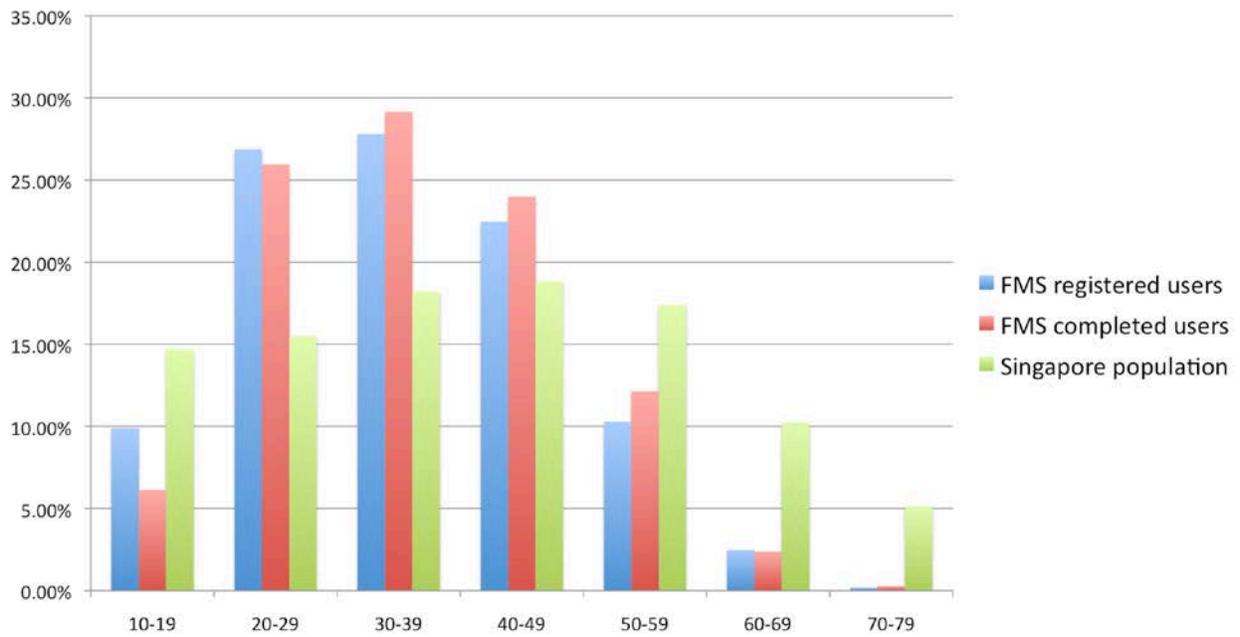


FIGURE 2 FMS web-interface activity diary.



**FIGURE 3 Age distribution of FMS users and Singapore population.**

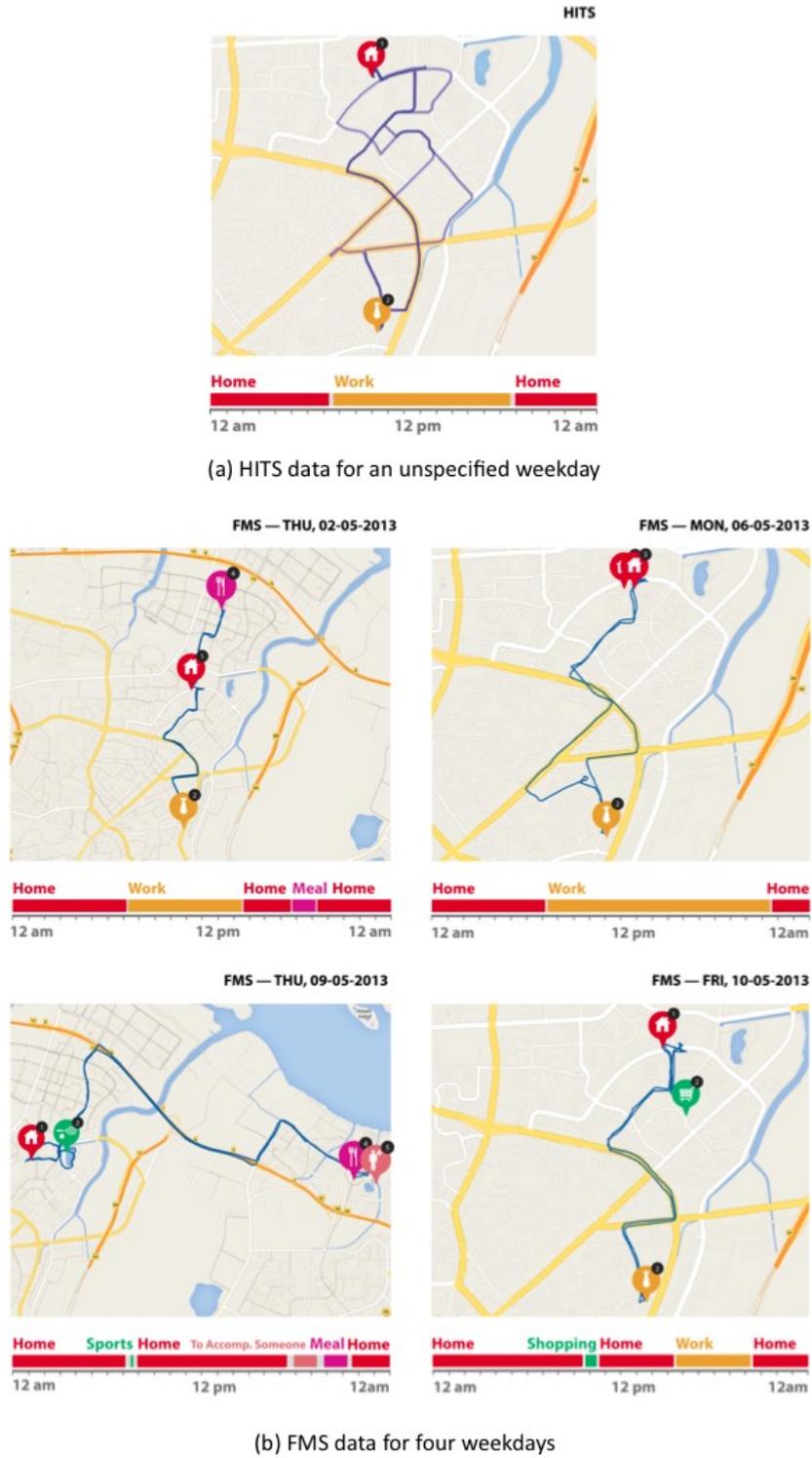
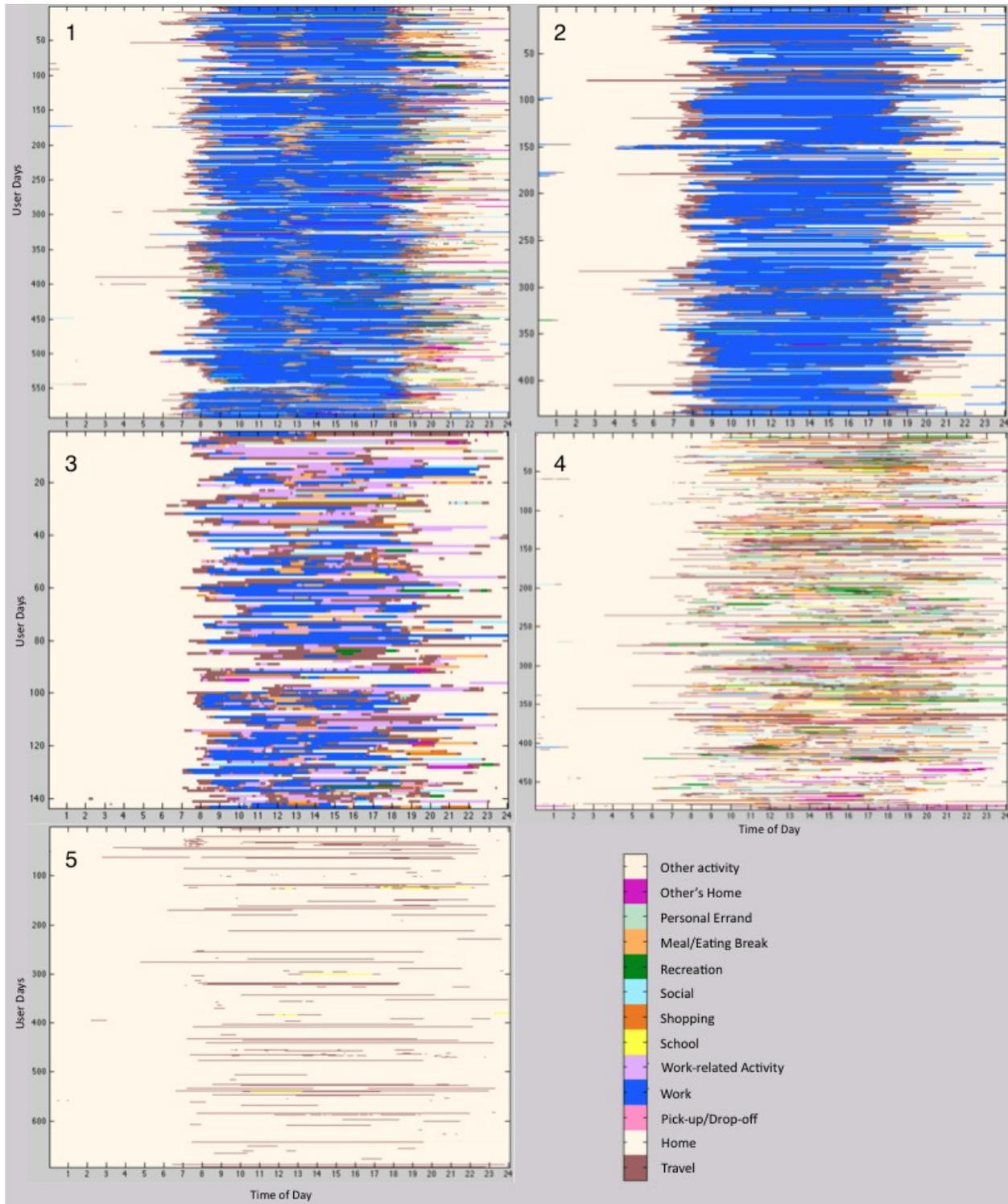
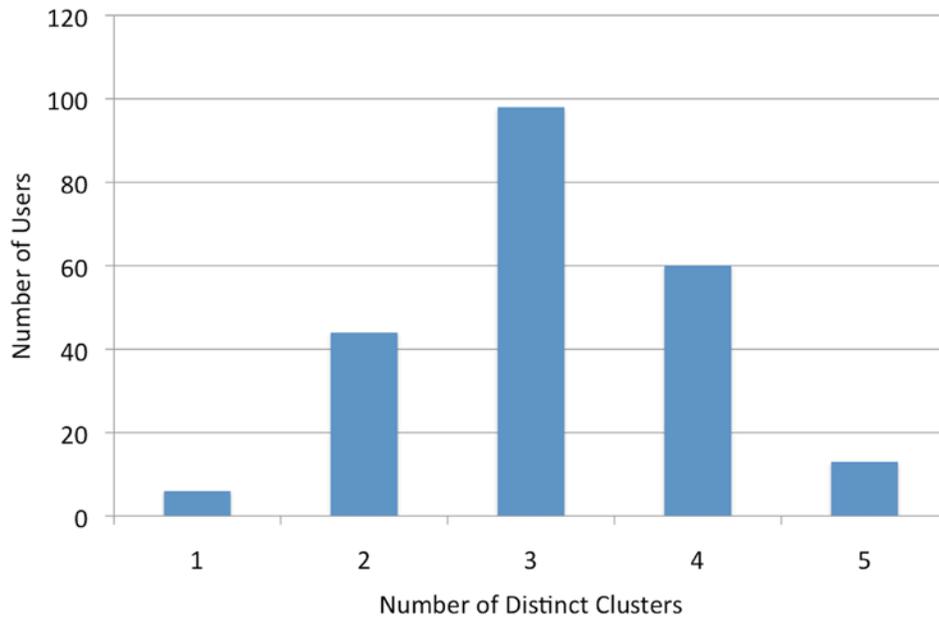


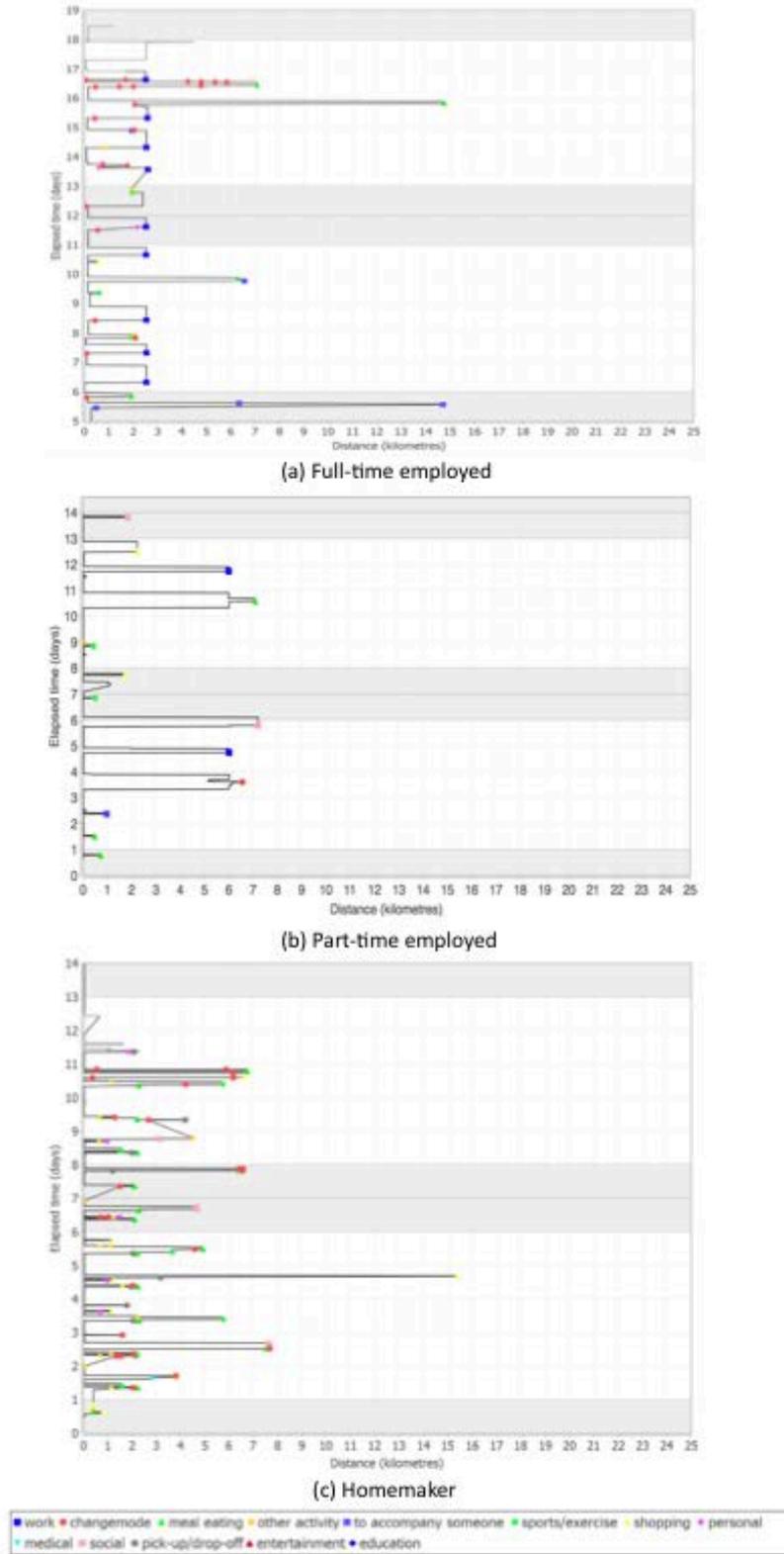
FIGURE 4 HITS and FMS data from one user on different days.



**FIGURE 5 Results of clustering FMS user day patterns into 5 clusters (employed users).**

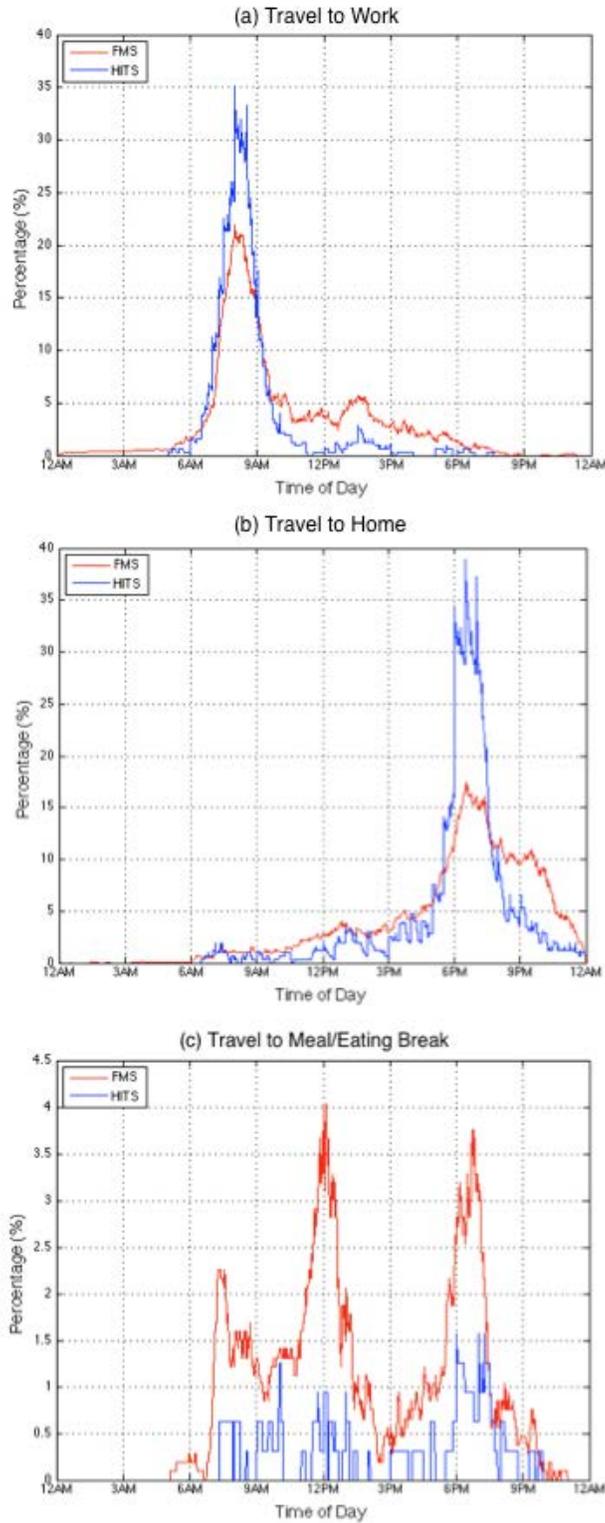


**FIGURE 6** Distribution of users with different numbers distinct clusters.

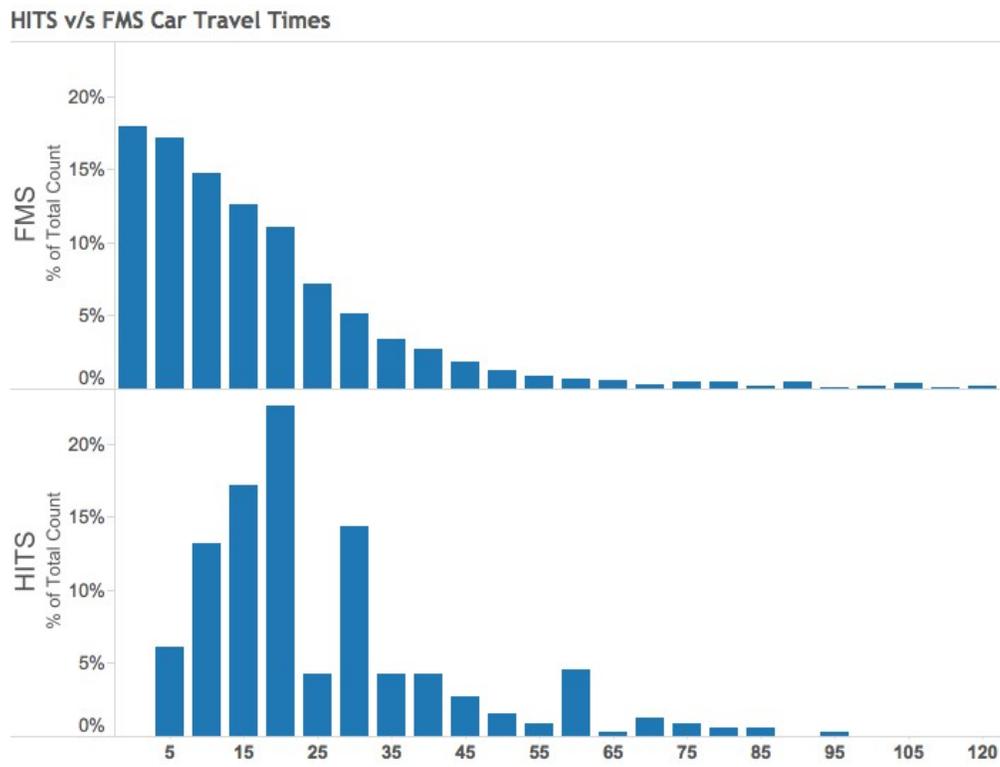


**FIGURE 7** Distance-from-home diagram for (a) a full-time employed user who works long hours; (b) a part-time employed user; and (c) a homemaker.

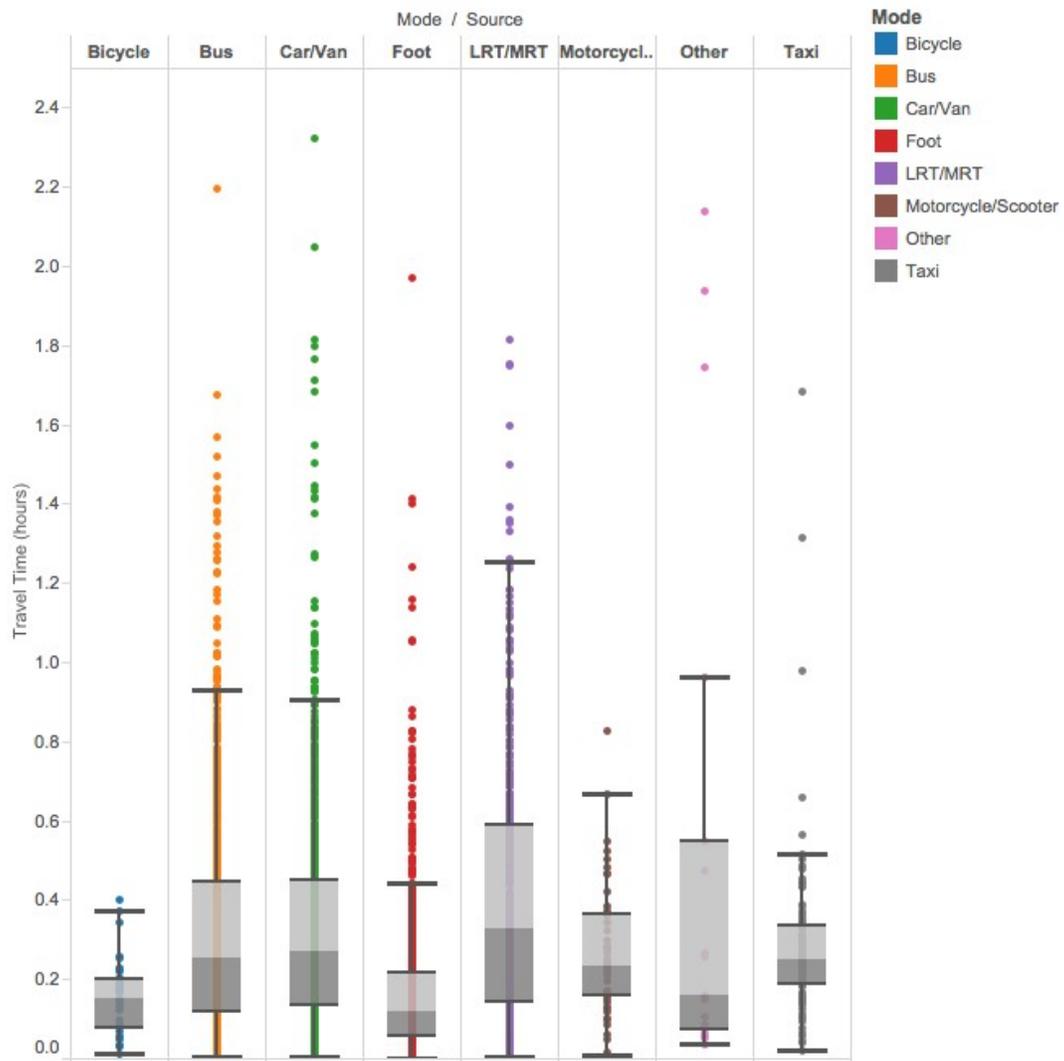




**FIGURE 8** Percentage of participants traveling to (a) Work, (b) Home, (c) Meal/Eating Break at different times of day.



**FIGURE 9 Histogram of trip times by car for HITS and FMS.**



**FIGURE 10** Boxplot of travel times for each mode in FMS.