CURRENT STATE OF SMARTPHONE SURVEY METHODS

2015-09-30

Transportation Tomorrow Survey 2.0

Mohammed Rashed, Siva Srikukenthiran

Khandker Nurul Habib & Eric J. Miller

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

Table of Contents

1 INTRODUCTION

Travel surveys have been used for a very long time to gather data on people's travel patterns, behaviours and choices. The traditional way to collect this information has been either in-person interviews, telephone interviews (for example using Computer Aided Telephone Interviews aka CATI), or mail surveys. While in the past these were more effective in getting responses and maintaining engagement, today they are now costlier to implement and require a significant amount of time to set up and conduct.

When the Transportation Tomorrow Survey (TTS) was first implemented in 1986, the main method of conducting these types of surveys was through telephone interviews as the geographical area of interest is too vast to use in-person door-to-door interviews. Due to the popularity and penetration of landline telephones in households, using the landline directory as a sample frame was an effective way of obtaining a representative sample of the population. The household was also chosen as the basic sample unit as it allowed for obtaining data on allocation and sharing of resources between individuals in a household.

Due to the proliferation of newer technologies and mobile devices, progressively fewer households are using landlines (Statistics Canada, 2013). According to a 2013 report, 27% of households in Ontario do not have a landline, meaning almost a third of the target population does not have a chance to be selected in the sample.

As a result, using landlines as a sampling frame has become less effective in obtaining a representative sample of the general population. Studies have shown that these households tend to consist of younger, lower-income adults. This is consistent with the under-representativeness of this demographic in recent TTS surveys. Specifically, starting in 2006, TTS researchers noticed that certain demographics, mainly younger adults aged 18-35, were under-represented in the survey (Data Management Group, 2011).

Conversely, the proportion of households with at least one active cellphone has been rising, with 85% of Ontario households fulfilling this criteria in 2013 (Statistics Canada, 2013). Additionally, smartphone penetration in Canada has been rising very quickly, from 33% in 2012 to 56% in 2013 (Google Canada, 2013) to 73% in 2014 (J.D. Power and Associates, 2014). It therefore makes sense adapt to these trends and technologies to reach out to these demographics.

Web-surveys have been used all over the world as an alternative method of surveying. For example, this method was used as a supplement to traditional CATIs in the 2011 TTS survey. The main benefit is that interview personnel do not need to be hired and trained, thus lowering costs as well as time for conducting the survey. However, different challenges are raised, including the fact that these are self-administered by the respondent, thus the questionnaire design and layout needs to be easy to comprehend and navigate.

All these aforementioned methods are active survey methods, and, thus, respondent burden is high. These methods all employ a prompted recall method, where the respondent is asked to recall all the trips they made on a particular day. With the TTS, respondents are asked to provide all trip information about the previous day, including time, duration, purpose, mode, and whether the trip included other members of the household. Unfortunately, this can result in respondents being unable to recall exact trip information such as the end time of the trip, often rounding to nearest 15-minute interval, as well as over-estimating the duration of the trip. Additionally, respondents often forget to include minor walking trips or short trips to the grocery for example, where they might think that this information is not vital for the survey. The problem is enhanced due to proxy bias, since the respondent who answers the telephone is asked to answer on behalf of other members of the household. This requires an additional step of communicating with other household members, leading to the potential of missing trips or incorrectly reporting certain trip details.

To overcome this, efforts have been made for more than a decade to move towards more passive methods of data collection. Numerous studies and surveys have been carried out using wireless dedicated GPS devices to automatically detect the location of the user, thus providing trip times and locations without requiring user involvement. However, this method does not automatically provide trip mode and trip purpose. There are many additional issues related to using GPS dedicated devices. These include cold-start issues, the need for users to carry an additional device with them at all times, and the need to post-process the collected data to validate its accuracy. These will be discussed in greater detail in **Section 2.**

There are some key benefits of using smartphones in place of dedicated GPS devices. First, users are already used to carrying these devices around everywhere with them; second, they have more sensors, features and processing power that can be utilized. Ideally, these capabilities could be used together to create a purely passive data collection device, requiring absolutely no input from the user to validate or provide information, including trip mode and trip purpose. This would reduce user burden considerably, as well as ensuring that every trip, no matter how short, would not be missed. While no such purely passive solution has yet been implemented, significant amounts of research is being done in this area as smartphone adoption rapidly increases.

Many of the current solutions involve using the smartphone's features and internet connectivity to make an estimate of the trip information and details, such as the trip purpose using surrounding land-use information, and then asking the users to validate or correct this estimation on a desktop device or on the smartphone itself. While not purely passive, this helps the respondent in remembering all trips and trip details by displaying those detected on a map. This ensures that trips are not missed as well as reducing user burden by filling in most of the details.

This report summarizes the current state of smartphone-based survey methods in literature, how these findings could be used in the TTS context, and provides recommendations for future TTS pilots and field tests. **Section 2** provides a brief literature review of smartphone-based survey methods overall, with subsequent sections going into specific areas that are of interest and need to be addressed. **Section 3** discusses smartphone technology and how it is being used to implement passive surveys. **Section 4** explores issues that require careful consideration when designing a survey on a mobile device, including user interface design requirements and ways to reduce instrument bias. **Section 5** discusses ways to minimize impact on battery life, specifically with regards to GPS usage. This also includes a discussion on the exchange of data collected from the smartphone to the server in order to minimize impact on data usage for the user. **Section 6** explores various issues with regards to whether and how a representative sampling frame could be obtained. In Canada, no cellphone directory exists, making it very hard to obtain a sampling frame of cellphones, therefore making it even more difficult to obtain one of smartphones. **Section 7** explores the potential for smartphones to be used in surveys that take place over multiple days or even weeks, instead of a single day survey used currently in the TTS. Since smartphones are inherently individual, **Section 7** also discusses whether smartphone-based surveys can still be used to carry out household-level surveys or whether a shift to more individual-level surveys and models is required, as well as the potential for smartphone surveys to be used as core or satellite surveys. **Section 8** explores various topics important to smartphone-based survey methods, including privacy concerns and how they may be addressed, ways to drive up response rates and retain participation of users, as well as other miscellaneous concerns. Finally, **Section 9** provides some recommendations for TTS 2.0 pilots to be carried out in the summer of 2016.

2 LITERATURE OVERVIEW

Surveys and studies on smartphone-based surveys in current literature can be broken down into 3 large categories (see the annotated bibliography in the Appendices for more detail). The first, and smallest, category consists of the real world, large-scale smartphone studies that cover many of the topics of interest for this report, and that have been conducted with participants from the general population. There are four such large-scale smartphone-based surveys that were found in the literature (Cottrill, et al., 2013; Safi, et al., 2015; Thomas, et al., 2014; Hathaway, et al., 2015).

The survey that provided the most lessons was the Future Mobility Survey (FMS) in Singapore (which was done in collaboration with MIT). The survey had 1,500+ participants, who were recruited from the regular traditional (face-to-face interview) household surveys done in Singapore (called HITS). It is the most comprehensive study in terms of covering all the topics that are of interest in this report, including passive detection, validation, battery optimization, and exploring the potential for smartphones to conduct multi-day surveys. This study will be cited frequently within this report.

The other three large scale surveys are not as comprehensive as the FMS but do cover many topics of interest. The rMove application from RSG and the ATLAS II project also recruited from the regular household travel survey, but had a much smaller sample size of 300 and 76 participants, respectively, and did not attempt to passively detect mode and activity (Hathaway, et al., 2015; Safi, et al., 2015). The Dutch Mobile Mobility Panel survey recruited from a panel and had almost 600 participants. An especially interesting feature to note about this survey was that approximately 60% of the participants did not own their own smartphone and were loaned one for the purpose of the survey (Thomas, et al., 2014).

The second category consists of studies or surveys that discuss and explore the potential of the technology within smartphones. This is accomplished by implementing new methods and algorithms to improve the passive detection accuracies, or battery life, thereby reducing participant burden, and may include discussions on other topics as well to various degrees. These studies mainly involve pilots that were devised to test the accuracy of the various algorithms and systems. They use smaller sample sizes (<50) that were recruited via a variety of methods, but mainly through convenience sampling. Two of the more notable studies that are comprehensive include those by Jariyasunant, et al. (2012) and Fan et al. (2015). The Future Mobility Survey also has specific studies that address different topics like mode, stop and activity detection.

The final category consists of studies and surveys that include improving response rates, tips for creating a well-designed and easy-to-use app, instrument biases that can occur, various methods of recruitment that can be used and how effective they may be to reach out to certain demographics, and privacy considerations. The studies by Greaves, et al. and Millar and Dillman, are especially insightful with respect to improving response rates by the use of incentives and various recruitment methods (Greaves, et al., 2015; Greaves, et al., 2014; Millar & Dillman, 2011; Millar & Dillman, 2012). These studies are not limited to only smartphone surveys, but the methods can still provide insight into user behaviour. Other studies focus on issues with survey design on smartphones and how functionality can be improved (Bruijne & Wijnant, 2013; Jue, Aaron, 2014; Liebe, et al., 2015).

3 PASSIVE DATA COLLECTION

Ideally, the smartphone data collection approach should eliminate manual input by the respondents almost entirely. Additionally, trips that are inadequately captured in prompted recall methods could be accounted for more reliably in a passive approach. For example, in traditional surveys such as mail-back or telephone surveys, short walking or cycling trips are often missed as the respondent is either more prone to forget about them or thinks that these trips are not important enough to mention (Safi, et al., 2015). With a purely passive approach, this recall would not be required.

However, the utilization of the technology on smartphones has not yet reached a level of sophistication where all the required information can be inferred from passively collected data, and as such, the respondent is required to validate predictions generated by algorithms to various extents. While not purely passive, this still has to potential to avoid under-representation of shorter trips as it helps respondents with recalling all trips made by presenting the information collected.

This section will examine the current state of smartphone technology in travel data collection as well as the evolution of the technology, methods, and processes.

In the literature, three main components have been focused on in order to obtain a passive collection system. These three components are the location, mode and purpose of each trip. Other pieces of information such as time and duration of each trip are also of interest, but these are included in the discussion of these components, such as location in this case.

3.1 Location and Trip Start/End

3.1.1 Location

As discussed in **Section 2**, dedicated GPS devices have been used to detect the location of users at various intervals of time. However, due to cold-start[1] issues and reduced effectiveness indoors, they are limited in their applications. A key advantage of using dedicated GPS devices is that data can continuously be retrieved without having to worry about battery issues. On the other hand, however, users are not accustomed to carrying them around everywhere, leading to the possibility of trips not being logged because the device was forgotten (Montini, et al., 2014).

Table 1 shows studies that explore various ways to obtain location data from smartphones, as well as the accuracy results of the methods used. These studies are discussed in this and subsequent sub-sections.

While GPS data quality in smartphones can be similar to dedicated GPS devices, in practice this is not the case because of battery life concerns on smartphones. In order to minimize the impact on GPS, developers may choose not to log GPS data continuously, and sometimes not at all, thus leading to a loss of location data. Dedicated GPS devices usually log data points continuously during a trip and thus do not have this issue. GPS units in smartphones are capable of providing accuracy to 10 or 30 meters unless there are urban canyon effects or signal issues, and GPS readings of high accuracy themselves often drift between 20-50 m (Jariyasunant, et al., 2012).

TABLE 1 LOCATION AND TRIP START/END

[1] Cold-start means that the GPS device does not begin recording valid positions at the beginning of a new trip when it is just switched on, as it takes a little bit of time to acquire a satellite signal. This shows up as a gap in the data between the end of one trip and the beginning of the next.

CITATION	LOCATION ACCURACY	TRIP START/STOP
(Abdulazim, et al., 2013) Using Smartphones and Sensor Technologies to Automate Collection of Travel Data	(median) **100 m** accuracy in **urban** areas. **600 m** in **low density** areas.	
(Vlassenroot, et al., 2015) The Use of Smartphone Applications in the Collection of Travel Behaviour Data	Gap lengths are **shorter than 60 s** in **85%** of trip legs; the gaps are repairable through interpolation using the real road network	
(Zhao, et al., 2014) Stop Detection in Smartphone-Based Travel Surveys		**95.5%** accuracy of stop detection – vital in subsequent accuracy for mode, activity. **12.3%** of detected stops were false positive (e.g. at traffic lights)
(Ghorpade, et al., 2015) An Integrated Stop-Mode Detection Algorithm for Real World Smartphone-Based Travel Survey	**Filters away data** from network sources (i.e. when GPS unavailable) if accuracy reading is **more than 500 m**	
(Safi, et al., 2015) Design and Implementation of a Smartphone-based System for Personal Travel Survey: Case Study from New Zealand		May miss up to 400 m of the initial sections of the trip, since relies on GSM/accelerometer to detect large movements before GPS is switched on. Logs GPS data every 10 meters
(Pei, et al., 2013) Human Behaviour Cognition Using Smartphone Sensors	Location accuracy is **1.9 meters in corridor environments** and **3.5 meters in open spaces** Works in a very specific environment only	
(Fan, et al., 2012) UbiActive: A Smartphone-Based Tool for Trip Detection and Travel-Related Physical Activity Assessment	GPS receiver was programmed to generate outputs every 30 seconds when moving	Only 6 participants were satisfied by the app's ability to detect trip starts and ends
(Jariyasunant, et al., 2012) Overcoming Battery Life Problems of Smartphones when Creating Automated Travel Diaries	The average error for a hotspot location was 25m ± 101m. This error increases to **197m ± 599m** for origin/destination locations which were not hotspots	**16%** of trips could not be identified within a 6 minute window of their true start time
(Thomas, et al., 2014) Dutch Mobile Mobility Panel		**15-20%** of trips not detected
(Fan, et al., 2015) SmarTrAC: A Smartphone Solution for Context-Aware Travel and Activity Capturing		Detects a transition between trip and activity within ± 30 seconds **88%** of the time

3.1.1.1 SENSORS USED TO COLLECT LOCATION DATA:

In the current smartphone travel data collection literature, the following features and sensors present in the smartphone are used for collecting location data:

- GPS
- GSM detection, cellular network triangulation
- Wi-Fi network-based location detection
- Accelerometer (for motion and location change detection)

GPS is still the main method of collecting location data, but it is usually supplemented by the other sensors. While GPS is the most accurate of these methods, the drain on battery is significant; therefore, it cannot be kept switched on at all times (Jariyasunant, et al., 2012).

The other sensors have minimal or a much diminished impact on battery life. Therefore, they are usually used during the times when GPS is turned off; however, they are also less accurate on their own except in certain situations. They can also be used simultaneously with GPS to improve location accuracy, but this is not usually required as GPS data quality, if logged continuously, can be accurate to 10-30 meters. The specific ways in which battery life can be optimized are discussed in further detail in **Section 5**.

3.1.1.2 COLD-START ISSUES

The other main issue with GPS (even in dedicated GPS devices) is the issue of cold-start. There are significant gaps in data especially at the start of trips that require handling after the data are collected, and require assumptions to be made by the research team as to where the trip started and ended. For longer trips, this may not be an issue as it could be safely assumed that the destination of the previous trip is the origin of the current trip. But for shorter trips, or detours in a longer trip, there is a risk that they could be missed entirely depending on the time it takes for the GPS to accurately obtain a signal, as well as the frequency of GPS polling used. The "urban-canyon" effect is also an issue, where GPS signals are less accurate in high-density urban areas due to the presence of taller buildings. (Greaves, et al., 2014)

These issues are present with smartphones as well, although the cold-start delay is not as significant as GPS technology improves. Vlassenroot et al. (2015) note that delays at the start of the trip are shorter than 60 s in 85% of trip legs and that these gaps are repairable through interpolation using the real road network. This was done with the use of continuous GPS polling. However, many solutions described in the literature reduce the frequency of polling GPS in smartphones as a battery-saving method or to reduce the amount of data transfer required from the mobile device to the server, requiring more sophisticated methods to re-create trips.

3.1.1.3 UTILIZATION OF GSM AND WI-FI FOR LOCATION DETECTION

As a solution to the problems inherent in using GPS alone (cold start, battery drain, etc.), other features and sensors in smartphones are utilized in various ways. One such method is using GSM (cellular network towers) to triangulate the position of the user; however, there is considerable variability in accuracy, often less accurate than GPS (Abdulazim, et al., 2013; Zhao, et al., 2014).

Another method is the use of Wi-Fi networks to detect location. Information about these networks can be stored on the smartphone app so that if a particular Wi-Fi SSID is detected frequently (called "hotspots"), and the user is at this location during weekdays for a certain period of time, the location can be quickly determined as one that the user has traveled to before (Jariyasunant, et al., 2012; Fan, et al., 2015; Abdulazim, et al., 2013; Zhao, et al., 2014). This can also be used for determining trip purposes, especially for those trips that are made regularly and for the same purpose.

This way, the accuracy of location data can be improved, as well as providing help in processing the data. If any erroneous "jumps" are detected, and the user is at one of these hotspots, these anomalies can easily be corrected by the system. Jariyasunant et al. (2012) obtained an average error of 25m ± 101m for a hotspot location. This was an improvement over an average error of 197m ± 599m for origin/destination locations that were not hotspots.

Both cellular networks and Wi-Fi networks are especially useful for detecting the user's location while indoors as well as in dense areas, where Wi-Fi is more common. This allows their use to be complementary to GPS, which does not do well in these situations (Zhao, et al., 2014). Many studies have supported this, finding that accuracy was higher in these areas as opposed to suburban or low-density areas, sometimes even more than GPS (Greaves, et al., 2014; Jariyasunant, et al., 2012). Abdulazim et al. (2013) noted accuracies of 100 m in high-density areas but 600 m in low-density areas.

Pei et al. (2013) were able to obtain location accuracies of 1.9 m in corridor environments and 3.5 m in open spaces (both indoors), using Fingerprinting Based Wireless Positioning. In this method, the received signal strength indicators of Wi-Fi networks are used for determination of position using the routers as reference points in an indoor location. This only works in a specific location (it was tested only at a university location), and requires prior knowledge of the locations of the Wi-Fi access points.

As mentioned before, on average, both of these technologies are not as accurate as GPS. Sometimes, erroneous location readings are made when a user is indoors for a long time, but the GPS or GSM signal "jumps" while the user is not moving (Zhao, et al., 2014). Jariyasunant et al. (2012) found that under 1% of GPS points lay outside the circle defined by the latitude, longitude and horizontal accuracy provided, compared to 51% of network points for locations sourced from either cell towers or Wi-Fi beacons.

3.1.1.4 GSM, WI-FI AND ACCELEROMETER FOR LOCATION CHANGE AND MOTION DETECTION

Due to the lack of accuracy in these sensors for location detection by themselves, most studies have focussed on using GSM, Wi-Fi and accelerometer readings as indicators of movement and for detecting large changes in location. When these large changes occur, the GPS is then switched on to get more accurate location data. For example, Jariyasunant et al. (2012) use a progression of sensors. First the accelerometer gathers data at 5 Hz for 3 seconds once per minute. If movement is detected, Wi-Fi beacons are used to detect whether the location change is significant (this data are gathered once per minute as well). If the user is detected to be moving faster than walking speed, only then is the GPS sampled. This continues until the person begins moving at walking or slower speeds.

While this saves battery, it can affect data quality which then requires interpolation or cleaning. In the Jariyasunant et al. (2012) study, 16% of trips could not be identified within a 6-minute window of their true start times.

Finally, once the GPS is turned on, there is a delay before the signal is first obtained. Therefore, even if a change in location is detected within 60 seconds, there will be an additional delay for this before the trip will start being recorded. For example, Safi et al. (2015) noted that up to 400 m of GPS data at the start of trips could be missed because a large change in location has not been detected by the algorithms that use the other sensors.

3.1.2 Trip Start/End Detection

Another important part of location detection is determining the exact moment when a trip starts or ends, as this can affect subsequent mode and activity inference.

One study based on the Future Mobility Survey in Singapore focused solely on this topic (Zhao, et al., 2014). First, stops were detected based on location and Point-of-Interest (POI) data (like bus/train stations, malls, pick-up/drop-off points, etc.). GSM, Wi-Fi and accelerometer information were used to merge stops that would otherwise be interpreted as distinct stops. The study described 6 steps to the algorithm:

- filter raw data
- generate candidate stops (generated if a user has been within an area of 50 m for one minute)
- check against frequent location signatures
- merge stops
- detect still mode (using accelerometer data to revise start and end times of candidate stops or merge them)
- and remove extra stops after mode detection (to remove false positives such as traffic lights, traffic jams, etc.).

Based on user validations, the true positive rate of the algorithm was 95.5% and 12.3% percent of the detected stops were false positives.

Fan et al. (2012) use GPS and accelerometer data to infer this information, but only 6 participants (out of 23) were satisfied by the app's ability to detect trip starts and ends, even though the GPS data was polled every 30 seconds while in motion. A similar method is used by the Dutch Mobile Mobility Panel (Thomas, et al., 2014) – 21% of users were dissatisfied with accuracy of automatic detection of destination and arrival, and up to 20% of the trips were not detected.

Fan et al. (2015) also paid special attention to identifying the exact moment when a user transitions from a trip to an activity. This was done by comparing the distances between the points logged around a particular point and seeing when the points start clustering together – if the points were greater than 5 m apart, those were considered as part of a trip leg, otherwise it was determined to be a "dwelling episode" (i.e. the user is at their destination performing an activity). Through iteration, the exact point of transition from travel to dwelling was obtained. A similar method was used to find the end of a dwelling and a new travel episode. This method predicted a transition within 30 seconds 88% of the time. This was achieved without continuous logging through GPS, but instead obtained GPS data over 30-second segments at 30-second intervals.

3.3 Mode

Figure 1 shows the overall mode detection accuracy achieved in various studies. It should be noted that the number of modes detected is not the same for all these studies.

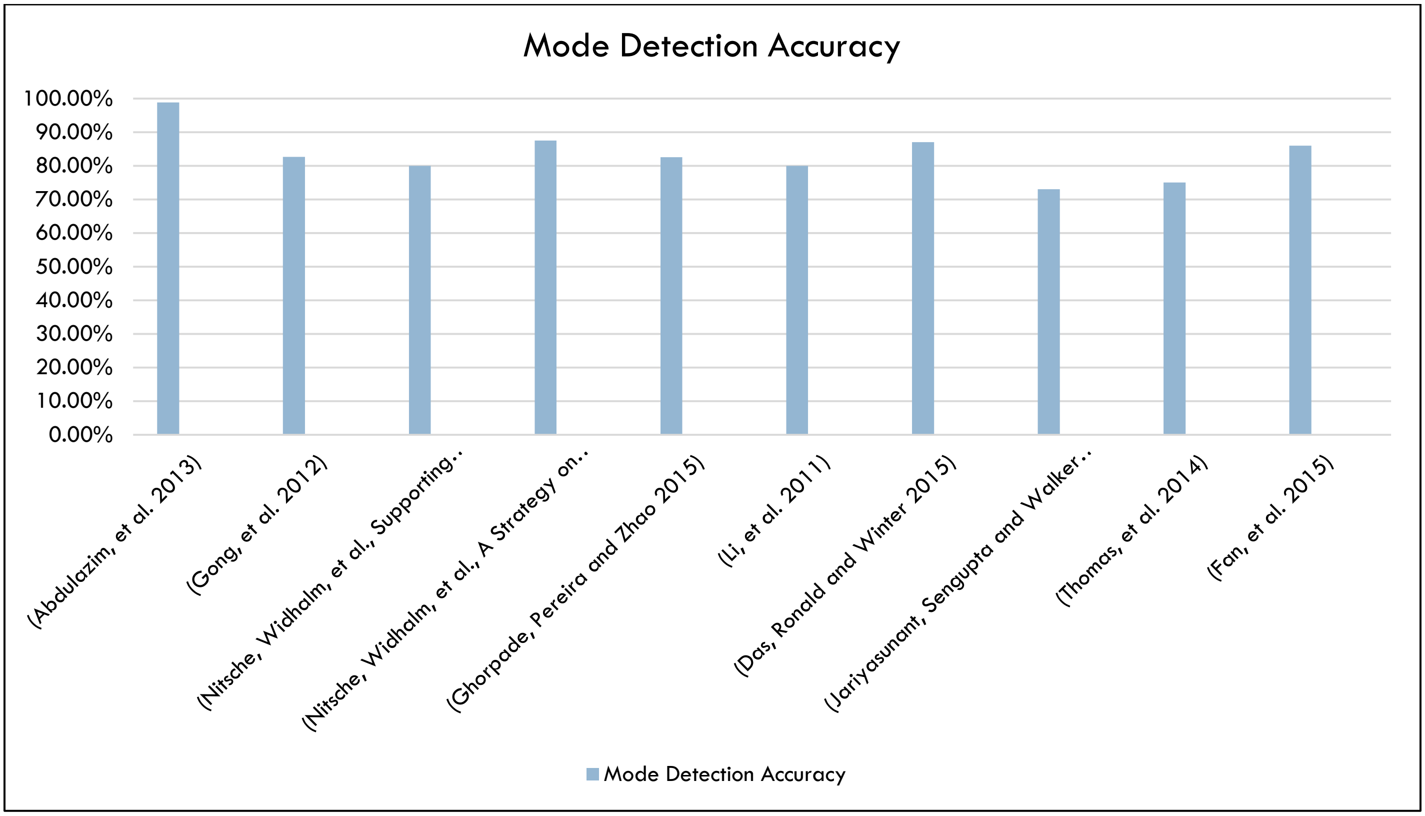

FIGURE 1 MODE DETECTION ACCURACY

3.3.1 Mode detection with dedicated GPS devices

Before smartphones became prevalent, when dedicated GPS devices were seen as the main way to collect passive travel information, the GPS data collected by these devices was the only way to estimate what mode of travel was used for each trip segment. This was performed via post-processing by utilizing GIS (geographic information system) data from existing road networks and by the development and application of heuristic rules (Gong, et al., 2012). One example of these rules would be assuming the mode was walking if the speed of travel detected was less than 5 km/hr. This simple method can produce good results when differentiating between motorized and non-motorized modes. While it would appear difficult at first to differentiate between a vehicle stuck in traffic and a person walking, much higher speeds on other sections of the trip data would indicate the use of a vehicle. However, detecting mode changes and differentiating between motorized modes such as cars and buses is much more difficult using simple rules.

This could be handled by using GIS information of the travel network, which includes transit information such as bus routes and stop locations, and overlaying the GPS data collected. For example, if the speed is less than 5 km/hr, but the location of a mode change point is detected near a bus stop, it could be assumed that there is a mode transfer from bus to walking or vice versa, as opposed to a driver in traffic. This method is not without flaws and has varying accuracy for different modes or places where the urban canyon effect causes the GPS signal to be weak, like high-density areas (Gong, et al., 2012). In their work, Gong et al. obtained a high accuracy result (82.6%), although the authors found it was difficult to differentiate between bus, car or walk modes.

3.3.2 Mode detection with smartphone devices in current literature

Smartphones on the other hand have more sensors that can be utilized. One such sensor, the accelerometer, present in all modern smartphones, has attracted significant research attention. This sensor is capable of detecting the smartphone's motion in different directions.

In combination with the GPS, the user's speed as well as acceleration can be determined. Even in cases where the speed for two different modes are the same, for example someone walking on the street or a driver stuck in traffic, the acceleration profiles for each are different. Through the use of machine learning, software can be used to identify the correct mode with greater accuracy.

Nitsche et al. (2012 and 2014) attempted to use GPS and accelerometer data to classify 8 different modes based on the motion trajectory recorded by these sensors. This was done by using probabilistic classifiers combined with Discrete Hidden Markov Model (DHMM). The classification results ranged from 65% (train and subway – since these are similar in terms of velocity and acceleration) to 95% (bicycle) accuracy. However, since it was constantly logging GPS data points, the impact on battery life was likely high.

Using similar methods, Li et al. (2011) also processed GPS and accelerometer data to detect mode. However, they used the accelerometer data to detect when there was a large change in location to determine when to turn on the GPS, thus reducing battery drain. This method was still able to achieve an 80% overall accuracy. The data sent to the server was also used to try to infer activity.

Similarly, Fan et al. (2015) obtained GPS and accelerometer data over 30-second and 120-second segments at 30 second intervals, using them as inputs for classifiers in random forests, to achieve a mode detection accuracy of 86%. They also performed a *correction to point-wise mode prediction,* where they used a backward-looking smoothing technique to correct mode predictions that were highly unlikely. For example, if a 30-second interval prediction showed "Car, Car, Car, Bus, Car, Car, Car", the "Bus" prediction was assumed incorrect.

Thomas et al. (2014) harnessed GPS data to achieve a 75% mode detection accuracy across 7 mode categories using pattern recognition via Bayesian statistics.

Das et al. (2015) used GPS data to obtain kinematic data (e.g. average speed, acceleration, variance of speed, etc.) and fused this with spatial information, such as average road/railroad proximity, variance of road/railroad proximity, relevance score for bus/train/traffic stop. They were able to achieve an 87% accuracy over five modes (car, walk, bus, train, cycling).

Jariyasunant et al. (2012) attempted to detect mode of transport after route determination, which was accomplished via the Google Directions API. This API returns a set of shortest paths by time between the origin and destination of the trip that run through the waypoints detected – as such, it is not an accurate solution since sub-optimal routes will not be considered. The mode detection was done using two classifiers: one to determine between motorized modes, bicycles and walking, and a second that used a map-matching algorithm to differentiate between public transit and driving. While a trip may consist of many mode segments (for example, a user takes a walk to a bus stop, rides the bus and then walks to their destination), only the main mode, bus in this example, was able be detected. Instead of using instantaneous data, the algorithm only used features that were defined over the course of the trip, such as median speed and trip start and end times. The main focus of this study was to optimize battery life, and while a mode accuracy of only 73% was achieved, the battery life was quite high at 33 hours, under conditions which assumed 2 hours of travel each day and average smartphone usage.

Ghorpade et al. (2015) collected GPS data for 3 minutes at 1 Hz frequency, then two minutes of no data collection, in order to preserve the battery. They also constantly obtained accelerometer readings at 2 Hz, and Wi-Fi data. Because of the gap in GPS data, a more sophisticated method for mode detection was required. Raw location data was segmented such that every location point within 5 minutes or 200 m from the first point was included in the same segment. In the first stage, each segment was classified into a stop, walk or motorized mode using random forest classifiers. This stage also included GIS features of the transport and transit network and user history (such as whether this user rides bikes, is more likely to walk than drive, etc.). These segments were then merged to remove erroneous stop segments. Finally, the homogenous trips of motorized modes were classified into car, bus or train. This achieved a relatively high mode accuracy of 82.5% across 5 modes.

Smartphones contain other sensors that also be utilized, such as gyroscopes. One study explored the potential of using the accelerometer, gyroscope and orientation sensor to detect the mode (Abdulazim, et al., 2013). The idea was that each mode has a unique signature of these three sensors associated with it, and thus no reliance on GPS data is required, which is especially useful for underground travel modes like subways, but would potentially have greater accuracy for other modes as well. The results obtained were very accurate at 98%. However, the sample size was very small in this study (6 people), and the method produces a lot of data that needs to be processed, potentially presenting challenges for practical application.

3.3.3 Areas for improvement in mode detection

One barrier to using accelerometer or other sensor data is that a lot of machine learning is required for different modes. Additionally, while most smartphones have these sensors, they do not necessarily function exactly the same way. If data used from training the application is used from a particular device, the accuracy of detections could be very high when used in practice with the same model of device, but not necessarily with another model. Therefore, data from each different model of smartphone is required. Furthermore, depending on where the smartphone is kept, the sensors readings are different. For example, if a user keeps their smartphone in their pocket or in their backpack while on the subway, the sensor readings will be different for both cases (Abdulazim, et al., 2013; Nitsche, et al., 2014). Therefore, all cases need to be accounted for with the training dataset before it can be used, which is a considerable task and has not been done in current literature as yet.

3.4 Activity/Trip Purpose

3.4.1 Activity inference in transportation

While a lot research is being done to determine trip purposes passively, this component has still not reached the level of accuracy of trip end detection and mode inference, and more frequent validations are required. Many different methods for passive inference are being employed, with the more successful efforts combining two or more of these methods (see

Table 2). These include using land use information from various sources, mainly third-party and freely available data sources, like Foursquare (Abdulazim, et al., 2013) or Google API (Fan, et al., 2015).

TABLE 2 ACTIVITY DETECTION/INFERENCE ACCURACY

CITATION	ACTIVITY DETECTION ACCURACY	NOTES
(Montini, et al., 2014)	**80%** median accuracy if 3-day data used. Collected second-by-second GPS and accelerometer data using dedicated GPS device	Uses person data as variables such as age, education level, distance to work, etc. Uses random forest classifiers.
(Montini, et al., 2014)	**80%** overall accuracy	
(Abdulazim, et al., 2013)	Not provided	Uses location-based networks like Foursquare to determine land use and thus infer trip purpose
(Cottrill, et al., 2013)	78 %	Uses a combination of methods, including acceleration and speed variables to detect stop segments and mode transfers, user history, distance to home and work, etc.
(Kim, et al., 2014)		
(Tanaka, et al., 2015)	F-measure of activity context recognition: 0.954 overall at a tavern and 0.920 overall a university	Uses machine learning by extracting features from GPS data to recognize several different contexts (activities) at the same location. Only works at two specific locations.
(Liao, et al., 2015)	**90%** detection accuracy of 4 different motion types: sit, walk, drive, and run – used to differentiate between different activities at one location	Analyzes mobile sensors to infer people's daily activities, which is then used to create a digital diary of the user
(Li, et al., 2011)	N/A	Uses rule-based methods to identify trip purposes
(Pei, et al., 2013)	Motion states are recognized with an accuracy rate up to **92.9%.** Limited to specific location and environment (workplace scenario)	Uses motion sensors and location data to determine what specific activities a user might be doing at a location
(Fan, et al., 2015)	The predicted activity matches the true activity type **25-40%** of the time. Correct activity type is among the **top three** predicted activities **80-95%** of the time	Tried to infer the purpose for unvisited locations using Google API to list the nearest points as the training dataset for activity classifiers

Wi-Fi network information is also used to infer activity as well as location, as these can provide a clue to the purpose by providing some more land use information. For example, a restaurant may have a Wi-Fi network with the company's name. This could be used to determine the user's trip purpose if that network is identified by their smartphone (Abdulazim, et al., 2013).

Many advanced rule-based methods are also used. For example, if a user is at a particular location during the weekdays for at least a certain amount of hours, then that is likely their work location, while if they are consistently at a particular location at 4 am, then that is likely their home location (Harding, et al., 2015). Li et al. (2011) provide some other examples of rules that could be used although they do not specifically state the accuracy of these methods. They use the end points of the last trip of the past 5 days as the home location for each individual. Work and school locations are identified if the duration of the activity lasts between 3-9 hours or is visited more than 3 times a day. If that location is a school as identified in location Points-of-Interest (POI) data, then it is labeled as a school trip, else it is a work trip. Other trip activities are identified using information like arrival time, day, stop time, nearest POI, distances between origin/destination and home/work/school. User feedback is then collected to validate the trip purposes, and this, in-turn, is used as further training data to build the trip purpose classifier.

Montini et al. (2014) collected second-by-second GPS data (from a dedicated GPS device) and user-specific information ("personalization") for trip purpose detection using random forests. In addition to information like trip start time, trip end time and day-of-the-week they also used personal data like age, education level, distance to work, etc. as they observed, in previous studies, that accuracy of trip purpose detection depended on the participant and their daily schedules. This way, they differentiated between location-based approach (where no personal data are included) and an activity-based approached (including personal data). They found that while adding more information increased accuracy, there was a point of diminishing returns. When using an activity-based approach, they obtained a median accuracy of 80% for trip purpose, a 5.5% increase from the base case where no personal data was included. Interestingly, they found that simply including distance of the trip from a user's home or work location (and not including any other personal data like age, income, etc.) noticeably improved the accuracy from the base case. They also found that this personal data had no effect in improving trip mode detections.

One way to improve accuracy of purpose detection at frequently visited locations by a user is through identifying and saving the set of Wi-Fi networks at that location. For example, if a user goes to their work location and validates the activity, the set of all Wi-Fi networks that are detected at that location are saved on the app. This way, if any one of these or more are detected in the future, the app can safely assign that location with that activity (Abdulazim, et al., 2013). The SmarTrAC report mentions that this method was used for home and work locations, and achieved a 95% purpose detection accuracy for these locations (Fan, et al., 2015). On the other hand, when purpose was inferred for unvisited locations by using Google API to list the nearest points as the training dataset for activity classifiers, only a 25-40% accuracy was acheieved.

The Future Mobility Survey used a combination of these approaches, to achieve an accuracy of 78%. This was without the collection of continuous GPS data in order to optimize battery life, unlike some of the other methods which used continuous GPS data from dedicated devices (Kim, et al., 2014). The Future Mobility Survey used acceleration and speed variables to detect stop segments (for example mode transfer, stop at traffic lights, etc.) as well as past activity information (i.e. validations about previous activities) and distance to home and work locations for each user.

3.4.2 Activity inference in other fields

Outside of the transportation field, there have also been efforts to determine activity passively, for example for health monitoring applications and personal diaries. In the former's case, these studies are focused more on the user's motion and physical activity, and have made great strides in determining whether a person is walking, running, cycling, or sitting – achieving a 90% across these 4 motion types using motion sensors and decision trees (Liao, et al., 2015). In their work, Liao et al. (2015) use location-based services like Google Maps and Yelp to determine a user's favourite restaurant – although this really only works for locations that are frequently visited. It would likely not be able to accurately detect a trip purpose if a location is being visited by the user for the first time.

(Pei, et al., 2013) used motion sensors and location data (via GPS and Wi-Fi) to determine exactly what activity a user might be doing at a specific location (they only tested a university workspace environment) – i.e. whether the user is at their desk, fetching coffee, in a meeting room, etc. To determine that the user is fetching coffee, the location of the user in the lunch room standing in front of the coffee machine can be determined (all of this requires prior knowledge of where all these rooms and devices might be within the room), as well as sensing the movements to and from the lunch room and pouring the coffee. Each different activity has its own set of features, and so it is difficult to generalize the model to classify activities for which it has not yet been trained.

Similarly, (Tanaka, et al., 2015) used machine learning by extracting features from GPS data to recognize several different contexts (activities) at the same location. Again, the application was limited to two specific locations (a tavern and a university location), but they achieve high accuracies (F-measure of 0.954 for the tavern and 0.920 for the university location).

While (Pei, et al., 2013), (Tanaka, et al., 2015) and (Liao, et al., 2015) are not directly applicable to travel surveys, they show the potential that smartphones have to differentiate between different activities at a particular location by detecting specific motion patterns. This could be used in conjunction with land-use determination methods to, for example, differentiate between someone walking on a sidewalk to get to a particular destination or shopping and going from store to store, or whether someone is at a shopping mall eating a meal during a lunch break or are shopping.

3.5 Recommendations

Location:

- Use different sensors like GSM, Wi-Fi and accelerometer to detect change in location before turning on GPS in order to conserve battery.
- Use GSM and Wi-Fi sensors for location detection while indoors and in high-density urban areas. Accuracy of these methods can be higher than GPS in these locations.
- Use hotspots – save the set of Wi-Fi access points available at frequently visited locations for each user in order to quickly detect arrival at these locations.
- Adopt a trip start/end detection algorithm similar to either Zhao et al. (2014) or Fan et al. (2015). This can help increase accuracy of subsequent mode detection and activity inference.
- Pilot recommendations: finding the right balance in polling other sensors and then turning on GPS – want to avoid delaying the turning on of GPS since this leads to loss of data, but do not want to constantly switch on GPS for false starts as well.
- Pilot recommendations: test trip start/end detection algorithms.

Mode:

- Sensors such as accelerometers, in addition to GPS data, should be used to detect travel mode, as it can be done without using up too much of the processing power or the battery of the smartphone.
- Incorporate GIS spatial information, such as the road and transit network information to help increase mode detection.
- Incorporate user history and information, such as whether the user ever takes transit, or has a bike, etc. to help in increasing mode detection as well.
- In terms of machine learning, random forests seem to be the most successful and widely used in the relevant literature.
- Sampling less GPS data to maximize battery life while still obtaining a relatively high accuracy can be done, similar to (Ghorpade, et al., 2015).

- Pilot recommendations: testing out and optimizing the above methods without affecting battery life significantly.

Activity:

- Use advanced rule-based methods for activity inference developed in literature, as discussed in Section 3.3.2.
- Use an activity-based approach, similar to Montini et al. (2014), using personalized data to help determine trip activity. Simply including distance of the trip from a user's home or work location (and not including any other personal data like age, income, etc.) noticeably improved the accuracy from the base case.
- Also include previous user validations for determining trip activity.
- Pilot recommendations: test out and optimize the above methods without affecting battery life significantly.
- Pilot recommendations: test methods and algorithms utilizing motion sensors and other smartphone data to differentiate different activities at a particular location, as a continuation of the works described in Section 3.3.3.

4 ACTIVE DATA COLLECTION, USER INTERFACE DESIGN AND INSTRUMENT BIAS

As mentioned previously, most travel surveys using smartphones employ some sort of validation instrument or have a separate prompted recall survey; therefore, they are not purely passive.

There are two types of methods:

- A prompted recall survey (such as in the Future Mobility Survey and the ATLAS II project) is usually in the form a web-based interface that is designed to be accessed from desktops, laptops or the smartphone itself. This requires the respondents to log on regularly (ideally at the end of the day) to validate trips detected by the application, correct any errors, and to add trips that may have been missed (Cottrill, et al., 2013; Safi, et al., 2015).
- Real-time validation, like in the rMove, UbiActive and other applications, where users are asked to validate and provide additional details about the trip right after it is made. Depending on the sophistication of the app, respondents may be required to input certain information themselves, such as the start and end of a trip, mode and purpose (Hathaway, et al., 2015; Fan, et al., 2012; Vlassenroot, et al., 2015).

The latter is more burdensome as the respondent has to continuously interact with the smartphone app. Additionally, the user has to fill in details like mode and purpose manually (to lessen the burden, usually a drop-down list is provided which can be personalized for each user). In the prompted recall case, on the other hand, respondents can pick a time during the day to correct and validate trip details. Mode and purpose are usually inferred and pre-filled, although these are not always accurate.

Discussion regarding the design of the web interface has not been included in this report as that is covered in the Web-based Survey Report – this report is limited to survey and app design on the mobile device itself.

4.1 Prompted recall via the app instead of web

Respondents tend to use the device that they are most comfortable with and that they access regularly, regardless of whether the survey asks them to use a specific device (Bruijne & Wijnant, 2013; Millar & Dillman, 2012). If respondents try to access the validation tool through their smartphones but the tool has not been designed for smaller screens, this may be off-putting for them and could result in non-response (Jue, Aaron, 2014). As a way to reduce respondent burden, it may be better to have the prompted recall survey accessible from the mobile device itself.

There have been successful efforts to move the prompted recall/validation component to the smartphone device itself so that users can conveniently record, recall and report their travel behaviour (Safi, et al., 2015). The app was designed with five tabs. The first tab, “Today,” allows users to visualize detailed trip information using Apple Maps about trips made so far that day. They can make corrections or add any trips missed by the app on this tab. The second tab, “History,” contains all previous trips and allows users to upload data to the server. The “Profile” tab contains the user’s account details as well as a section where various survey questions can be asked, perhaps as an exit survey. “Help” contains tutorials and FAQ to help users learn how to operate the app. “Info” contains information regarding the research team and goals of the research (Safi, et al., 2015). With this design, all the information is contained within the app and there is no reason to burden the user to use any other device or mode, thus making it more likely that the respondent is engaged and responds as required.

One of the most important points stressed in many studies is that the readability based on colour and font should be maintained on any device (Cottrill, et al., 2013; Jue, Aaron, 2014). In addition , if a user views a page from their smartphone that was actually intended to be accessed from a PC, they should automatically be re-directed to a mobile-optimized version instead (Millar & Dillman, 2012). In many places, it has been recommended to create the mobile version first, and then use that for the PC version. The argument is that for these purposes it is much easier to scale up than scale down – readability and navigation designed for a mobile version can easily be ported to devices that use larger screens, but doing it the other way round necessitates many compromises and changes. One study found that the respondents of a hybrid design survey (which is the mobile-optimized design on a desktop survey) completed the survey slightly quicker than the web/desktop version (Bruijne & Wijnant, 2013).

4.2 App survey and user interface design

The user interface for smartphone apps requires special consideration and design, whether it is for prompted recall or real-time validation. There are many studies and reports which provide recommendations for surveys on smartphones and app design in general, four of which are summarized in **Table 3**. Because of the smaller screen size and lack of physical input controls like a keyboard, apps on smartphones are designed to be simplistic and streamlined. Methods and survey design that have been tested and proven to work on desktops and laptops do not, in general, work for mobile devices.

For surveys on the app itself, there are many considerations to keep in mind. The screen size is the biggest factor, and the input method, which is touch-based for most smartphones nowadays, is important to keep in mind too.

Studies in the literature recommend the following practices:

- The number of questions should be limited to one per page, in order to minimize the need for scrolling (Bruijne & Wijnant, 2013; Wells, et al., 2014).
- Use large, readable fonts with clear spacing (Jue, Aaron, 2014)
- Grid designs for response options are discouraged, as a lack of spacing between choices can cause input errors and frustration for the user. If grid formats are to be used, the text should be placed above the individual attributes as opposed to the side, as one study showed that this provided higher correlation between mobile and non-mobile responses (Lattery, et al., 2013).
- Instead of radio buttons, button boxes should be used, such that the entire row is selectable and not just a single point (Jue, Aaron, 2014)
- Large buttons for navigation and response inputs should be used so that these can easily be accessed without requiring users to attempt to make a selection multiple times (Jue, Aaron, 2014).
- In order to save time and decrease response burden, questions should be designed as multiple choice where possible (Safi, et al., 2013). While smartphone users are becoming increasingly used to providing short text-based responses on mobile devices due to the prevalence of texting and using various services on these devices, they are still more likely to provide shorter responses compared to desktop users, thereby potentially not providing as much detail (Wells, et al., 2014).
- Questions and sets of responses should be made as short as possible to minimize scrolling. This may not be feasible for the TTS survey – the questions used in the study were very short, while the TTS questions tend to be more detailed (Wells, et al., 2014).
- The best option to use when making rating scales is to use the 5 point scale – if more point scales are used, mobile respondents tend to rate lower than PC respondents (Lattery, et al., 2013).
- Using large-sized background images should be avoided as this takes longer to load and can cause completion rates to drop (Lattery, et al., 2013).

TABLE 3 COMPLETION RATES

TITLE / REFERENCE	COMPLETION RATES	NOTES
(Lattery, et al., 2013) Optimizing Surveys for Smartphones: Maximizing Response Rates While Minimizing Bias	Completion rates: Mobile: 79.%, Non-mobile: 89.5% Completion time: Mobile: 15.3 min, Non-mobile: 14.1 min	Purpose of research was to identify differences in response rates and responses themselves between mobile and PC respondents. Tested wide range of different designs and features such as grid format, radio button spacing and label alignment, etc.
(Wells, et al., 2014) Comparison of Smartphone and Online Computer Survey Administration	7.8 % response rate. Completion rates: mobile: 58% Web: 61% Median completion time (min): mobile: 5.5, Web: 5.8	Tested whether mobile app survey respondents are sensitive to particular experimental manipulations similar to other modes and analyzed differences between modes
(Bruijne & Wijnant, 2013) Comparing Survey Results Obtained via Mobile Devices and Computers: An Experiment with a Mobile Web Survey on a Heterogeneous Group of Mobile Devices Versus a Computer-Assisted Web Survey	Completion rates: computer: 61% mobile: 47% hybrid: 64% Average Completion time (s): computer: 339 mobile: 542 hybrid: 321	Compares results of survey between a survey done on mobile device and computer. Used three version of the survey: web/computer, mobile, and hybrid (which used a mobile layout but for use on a computer)
(Jue, Aaron, 2014) Trends Report: Mobile Participation in Online Surveys	Desktop: 73% Tablet: 74% Smartphone 64%	Provides an overview of mobile survey design issues and provides recommendations for mobile survey design

One of the biggest issues with creating an app is that there are multiple brands and models of smartphones in the market, and each have different operating system designs and inputs. For example, iOS devices do not have a dedicated physical "back" button, while Android devices do. These type of differences need to be accounted for, either by having a different back button in the app itself that would be common to all versions of the app, or by designing around the differences for each version. It might be more advisable to go for the former option such that the app is standard across different devices in order to minimize instrument bias.

Whether on PCs or on the smartphone, many useful features can be implemented in the application to help lessen the amount of effort required by the user. Features like auto-fill, auto-complete for addresses or popular locations, prompting (for example for public transportation access mode to capture incidental walking and cycling), save and recall favourite trips so that users do not have to fill in details for a frequent trip every time, and providing maps of their daily travel to help with recall (Greaves, et al., 2014; Cottrill, et al., 2013). All these features can greatly help with the input process to reduce respondent burden in the absence of a purely passive solution. Some of these features, such as the favourite and frequent trips saved, can help to create a contextual knowledge base for each user in order to improve detection accuracy and to help with trip purpose identification (Cottrill, et al., 2013).

4.3 Instrument bias

Acquiescence is the phenomenon where respondents simply choose the top-most response in a survey without reading the question or all the options properly in order to quickly finish the survey. A study found that while respondents take longer to finish surveys on mobile devices than PC respondents, they are not more prone to acquiescence (Liebe, et al., 2015; Wells, et al., 2014). The only caveat to this is that the study was done on relatively short surveys with response options that were not overly long, so this result can only be applied in such cases, at least till further study proves otherwise.

It should be noted the study used a mobile-optimized version for the smartphone survey, as did the other studies cited in this sub-section.

Wells et al. (2014) also found that while mobile respondents are willing to provide short responses to open-ended questions, their responses are consistently and significantly shorter in length. It is for this reason that it is recommended to provide multiple-choice questions with sufficient responses where possible.

Bruijne et al. (2013) compared the results between a survey done a mobile device and a PC. They designed three versions of the survey: a web/computer version (Condition 1), mobile (Condition 3) and a hybrid (Condition 2). Condition 2 used the mobile layout of the survey but for use on a desktop. They found that the mobile group users took longer to complete the survey compared to the other two conditions, and their perception of the length of the survey was also longer. Interestingly, they also found that those who completed the hybrid versions of the survey completed the survey quicker than those who used the web version.

Jue et al. (2014) also noted that completion times are longer for mobile respondents, and that smartphone users almost always abandon surveys at a greater rate than devices with larger screens. However, completion rates are improving over the years as the surveys are becoming more mobile-friendly. Many of the issues and differences in completion rates can be mitigated by using a well-designed, mobile-optimized version of the survey. (Jue, Aaron, 2014)

4.4 Tutorials and FAQs

The main goal of smartphone surveys is to make the instrument as passive as possible. Even if smartphone travel survey apps become purely passive, however, users will still require some sort of instruction guide. This would need to show them how to use the app, and what they should or should not do in order to make the app efficient and to capture quality data. Otherwise, the research team will be required to address individual cases which, if the number of respondents is very large, could be a significant effort.

Quite a few studies, such as the Future Mobility Survey (Cottrill, et al., 2013), ATLAS II (Safi, et al., 2015) and SmarTrAC (Fan, et al., 2015), have employed the use of tutorials to help users understand how to use the app as well as to decrease the effort required by research team to answer each respondent individually. SmarTrAC has 9 tutorials that are posted online and are not on the app itself. The Future Mobility Survey and ATLAS II apps both contain tutorials that show the user everything they need to know in order to make sure the phone is recording data, as well as the steps required to validate the data. Both sets of respondents were very receptive of the tutorials and did not require much extra information from the research team. As such, both reports highly suggest the implementation of in-app tutorials (Cottrill, et al., 2013; Safi, et al., 2015). The Future Mobility Survey even had call center support during business hours to respond to any queries from respondents.

4.5 Recommendation

- Use the prompted recall method instead of real-time validation.

- Have a mobile-optimized version of the prompted recall survey, or have it as part of the app itself, such that users do not need to shift between different devices, and can have a seamless user experience.
- There should still be a web-based (desktop) prompted recall survey if the user prefers to answer that way. Design should be based on the mobile version as it is easier to scale up from smaller screens to larger ones than the other way round. The interface and login process should be the same for all devices so as provide a seamless user experience.
- Employ features like auto-fill and auto-complete based on prior responses and current trip logic, prompting, save and recall favourite trips, and provide maps to help with recall.
- Provide tutorials and FAQs for the respondent to guide them with respect to using the app, and the do's and don'ts. These can be on the app itself or on the web.
- Use one question per page, large fonts, large button boxes instead of radio buttons.
- Use multiple-choice questions where possible.

5 BATTERY, PERFORMANCE AND DATA ISSUES

Two of the biggest issues plaguing the use of smartphone-based travel survey apps that are being tested today are that of battery life and the effect on the smartphone's performance. A few methods that have been tried to various degrees of success from current literature are discussed in this section, along with a summary of results in **Table 4.**

5.1 Battery Life

The constant use of GPS for location data collection has a significant impact on a smartphone's battery. While GPS is highly accurate when a signal is obtained, its use needs to be managed in order to make the most of the battery life. If kept on constantly, the battery can drain within 5 hours (Jariyasunant, et al., 2012), making it very impractical. The UbiActive application polled the GPS every 30 seconds during motion, resulting in a significant drop in battery life of between 25% - 83% (Fan, et al., 2012).

Greaves et al. (2014) used a companion smartphone app to assist users in recalling trips that they made. This app was an optional component of the survey and it only detected location. Due to concerns about battery usage, it did not employ GPS – only Wi-Fi and cellular network information. While these sensors can have a wide range of accuracy, they found that in high-density and built-up areas where Wi-Fi is more common, the accuracy was quite high and sufficient for their needs. As a result, their app did not have a significant impact on battery. Since this method was mainly used to help users recall their trips, and was not meant to necessarily significantly reduce user-input burden, this satisfied their objectives.

As a comparison, another app that only tracked location attempted to use GPS, Wi-Fi and accelerometers sensors for location detection and try to minimize impact on battery life (Vlassenroot, et al., 2015). Despite their efforts, the app required a recharge of the battery at least once a day.

Abdulazim et al. (2013) also used Wi-Fi networks and cellular triangulation, in addition to GPS, in order to collect location data and to infer trip purpose for a more passive application. Of the 6 participants, 5 reported no significant change in battery life. As mentioned earlier, identifying location through Wi-Fi networks can have a wide range of accuracy, from 10 m to over 1000 m, which is not very useful, and this cannot provide any information about the user's speed of movement. In dense urban areas like downtown Toronto, however, the location accuracy was about 100 m (Abdulazim, et al., 2013). As such, these additional sensors are not useful for detecting location during trips, but are instead most useful to detect user locations when they have arrived at their destinations. They may also be helpful for short walking trips where Wi-Fi availability is very high.

Many studies used other sensors, not necessarily for location detection itself, but to detect when the GPS should be switched on and off based on if the user was moving or not or whether the user is indoors. A "phased sampling" approach was used in the Future Mobility Survey in Singapore (Cottrill, et al., 2013; Ghorpade, et al., 2015), in the ATLAS II project (Safi, et al., 2015), in the SmarTrAC project (Fan, et al., 2015), as well as in (Jariyasunant, et al., 2012), where they had cycles of GPS awake and sleep periods.

In the Jariyasunant et al. (2012) study, during the sleep periods, other sensors such as GSM (cellular network detection), Wi-Fi and accelerometer data were collected. If a change in GSM location was detected beyond a certain threshold distance, GPS was turned on, and if the user was in stasis for a threshold period of time, the GPS was turned off. The Future Mobility Survey even attempted to account for likely user activities as indications of awake/sleep periods. For example, longer sleep times were used when participants were detected to be at the office or at home. (Jariyasunant, et al., 2012) and SmarTrAC (Fan, et al., 2015) focussed more on movement detection through changes in Wi-Fi and accelerometer data, adjusting the duty cycle of these and the GPS sensor accordingly.

TABLE 4: BATTERY LIFE

Title / Reference	Battery Performance
(Greaves, et al., 2014) A Web-Based Diary and Companion Smartphone App for Travel/Activity Surveys.	Used Wi-Fi and network information for location. Since no GPS used, no large impact on battery
(Abdulazim, et al., 2013) Using Smartphones and Sensor Technologies to Automate Collection of Travel Data.	Avoids use of GPS for mode and location where possible. 5 out 6 participants reported no significant change in battery life.
(Vlassenroot, et al., 2015) The Use of Smartphone Applications in the Collection of Travel Behaviour Data.	Requires recharging phone at least once a day. Users required to input the start and end of each trip, as well as activity and mode
(Cottrill, et al., 2013) Future Mobility Survey.	Range of 10 to 24 hr, depending on user's travel pattern. To preserve battery, has cycles of GPS data collection: collects GPS data for 3 minutes at 1 Hz frequency, then two minutes of no collection
(Ghorpade, et al., 2015) An Integrated Stop-Mode Detection Algorithm for Real World Smartphone-Based Travel Survey.	
(Liao, et al., 2015) Smart Diary: A Smartphone-Based Framework for Sensing, Inferring, and Logging Users' Daily Life.	Battery life stays more than 11 hours before dropping down to 30%
(Safi, et al., 2015) Design and Implementation of a Smartphone-based System for Personal Travel Survey: Case Study from New Zealand	Average working duration of the app was 36.6 hours
(Fan, et al., 2012) UbiActive: A Smartphone-Based Tool for Trip Detection and Travel-Related Physical Activity Assessment	Battery life shortened by a range between 25% - 83%.
(Jariyasunant, et al., 2012) Overcoming Battery Life Problems of Smartphones when Creating Automated Travel Diaries	33 hours of energy consumption, assuming 2 hours of traveling each day and average smartphone usage. On average, the app accounts for 3% of total battery consumption at any given time
(Thomas, et al., 2014) Dutch Mobile Mobility Panel (Hoe mobiel zijn we nu eigenlijk? Eerste inzichten uit het Mobiele Mobiliteitspanel)	50% of users said impact on battery was significant
(Fan, et al., 2015) SmarTrAC: A Smartphone Solution for Context-Aware Travel and Activity Capturing	47% had a battery life longer than 8 hours.

This phased sampling approach usually requires trade-offs between resource efficiency and data accuracy. This is because data quality and location accuracy are reduced during GPS sleep times, while battery consumption is still high if the user makes long trips. As a result, the battery life for the Future Mobility Survey varied from 10-24 hrs, depending on the user's travel activity. The ATLAS II project had better success, with an average working duration of the app at 36.6 hrs.

A follow up to the Future Mobility Survey issue of high battery consumption during trips was addressed by Ghorpade et al. (2015). Even during awake periods, the GPS was polled for 3 minutes at 1 Hz, and then stopped for the next two minutes, repeating this cycle till the next sleep period was detected. While this improved the battery life, it had an impact on data accuracy. Mode detection especially required further sophistication, but a relatively high accuracy of 82.5% was still achieved.

Jariyasunant et al. (2012) also addressed the battery consumption issue by sampling the GPS only once per minute during motion. This allowed them to achieve a battery life of 33 hours on average, which was quite high. However, this method resulted in location, route and mode detection accuracy (73% accuracy across 4 modes) being not particularly high; it did not predict activity. They also found that battery performance could vary significantly across phone models, and sometimes even on two different phones of the same model. On one phone of a particular model, the battery life was not affected, but on a different phone of the same model, the battery life dropped by 2 hours.

SmarTrAC tried a slightly different approach, where GPS and accelerometer data were sampled for 30-second and 120-second segments at intervals of 30 seconds. The result on battery life was not good, as only about half of their users' phones had a battery life longer than 8 hours when using the app. Additionally, activity detection was quite poor (25-40%). On the other hand, mode detection was much higher at 86% over 6 modes. (Fan, et al., 2015)

The Dutch Mobile Mobility Panel used a mobile application that attempted to reduce the number of GPS samplings for longer trips, and also collected Wi-Fi and GSM data. However, whether and how these other sensors were used to detect when GPS should be switched on or off was not described. Half of the users said the impact on battery was significant, but it should be kept in mind that 60% of participants were loaned a smartphone for the survey, and as a result, may not be familiar with the generally shorter battery life of smartphones compared to regular cellphones. (Thomas, et al., 2014)

Li et al. (2011) used an adaptive sensing policy that reduced the sensing frequency of the GPS automatically with minimal impact on accuracy, but without using any of the other sensors. In essence, they detected the speed of the user through GPS, increased the interval of the GPS polling sensor when speeds were low, and decreased the interval when the speeds were high. This way the GPS receiver was not very active when people were stationary, but quickly started polling and collecting data when the user started to move. No specific data was provided for comparison.

Two studies of interest were found outside of the transportation field. Pei et al. (2013) used a combination of GPS and Wi-Fi networks to determine when the user was indoors or outdoors. If the number of GPS satellites and their signal-to-noise ratio was sufficiently high, this was recognized as the outdoor case. Conversely, the indoor case was assumed when the GPS signals were sufficiently weak, but Wi-Fi signal strengths were high. This assumed that Wi-Fi signals and networks were abundant where users are travelling. This may not be the most efficient use of the battery since a user could be outdoors but not moving, for example waiting at a bus stop or spending time at a park.

Liao et al. (2015) attempted to create a personal digital diary (note: not a travel diary), and monitor the smartphone's motion, location data, app usage, phone call and message history, etc. They were able to identify 4 different types of motion: sit, walk, run and drive with 90%, while maintaining battery life of greater than 30% for more than 11 hours. The app, however, was not necessarily interested in tracking the trips and routes taken by the user, but, instead in the physical, social and entertainment activities of the user.

5.2 Performance and Data Transfer

As smartphones are getting more and more powerful, the discussion around performance is not as big of an issue as before. Problems can arise when accelerometer or other motion sensors are continuously being logged, creating a huge amount of data that needs to be processed by the smartphone.

If the accelerometer is being used simply to detect whether the amount of movement of the user is high or not, Li et al. (2011) suggest using a low-pass filter algorithm that checks whether the sum of absolute difference between subsequent accelerometer data points (in the x, y, and z readings) is higher than a certain threshold. If so, then the user is in motion.

There is a trade-off between the data processing being done on the phone itself versus in the cloud. In the former case, the battery can be drained very quickly, while in the latter case, data bandwidth for users is limited. Most of the studies listed here prefer to maximize battery life and instead upload the data to the research team's central server, as all calculation and processing for the various algorithms for predicting location, mode, activity, trip start and end points, etc. can be more easily handled by a central server. On the other hand, this creates huge amounts of data that needs to be sent to the server. While data caps are increasing, this can still create an issue, especially for users who are unaware of how much data are being generated.

To mitigate this issue, some of the studies offered the user a choice of uploading the data only when Wi-Fi was available or using their data plan if they gave permission (Cottrill, et al., 2013; Safi, et al., 2015). Others attempted to compress the data before it was sent to the server. Nitsche et al. (2014) implemented a compression protocol that reduced the amount of data from 30 MB per hour to 6 MB per hour. SmarTrAC also used a compression protocol – normally, 50 MB of data are produced per day, but their compression algorithm was able to reduce it to approximately 21 MB per day, or 150 MB over the 7-day survey period. This, unfortunately, was still a considerable amount of data to upload.

5.3 Recommendations

- Use a phased sampling approach in order to improve battery life – use sensors like Wi-Fi, GSM and accelerometer to detect when a user might be making a trip. The accelerometer can be used to detect motion. If motion is detected for a period of time, the GSM and Wi-Fi sensors can be polled to detect if a large change in location is being made, beyond some threshold distance that will need to be determined. If so, the GPS sensors can then be switched on. If location changes of less than 50 m are detected for a few minutes, the GPS can be switched off, and the Wi-Fi, GSM and accelerometer data can then be monitored again to detect movement.
- During motion, the GPS does not need to be continuously sampled. Different approaches could be used. The Future Mobility Survey uses 3 minutes of GPS sampling at 1 Hz followed by 2 minutes of no sampling. This leads to a decrease in accuracy for detections, and increases the level of sophistication required for the algorithms. Still, relatively high accuracies can be achieved.
- Implement a compression protocol similar to (Jariyasunant, et al., 2012) or (Fan, et al., 2015) to reduce the amount of data being transferred to servers to processing. A low-pass filter can be used for accelerometer readings if only whether the user is in motion or not is required to be known, but not if it is needed for mode or purpose detection.
- **Pilot recommendations:** determining the right combination and use of sensors for different scenarios, such detection/inference of trip start/end, mode change, high density urban locations, rural areas, etc.

- **Pilot recommendations:** determining the threshold distance after which the GPS should be switched on. Make it too large and the location data will not be accurate, while having the distance too small will decrease battery life by sampling the GPS often.
- **Pilot recommendations:** similarly, determining the best frequency for polling GPS while in motion while maintaining data accuracy, and battery life – or implementing algorithms that improve detection accuracy with reduced data.
- **Pilot recommendations:** testing the app developed on different phone models and brands since performance can vary significantly.

6 SAMPLING FRAME

The discussion surrounding a suitable sampling frame for travel survey purposes has been spreading throughout the transportation community in the last decade, as landline use is dropping and cellphone adoption is rapidly increasing. While much research is being done to assess the viability and to explore the potential of smartphones for conducting travel surveys, the problem of finding a suitable sampling frame has not been satisfactorily resolved. Indeed, many of the smaller pilots in literature consist of samples obtained from targeting specific groups or recruited through convenience sampling, and have thus not specifically addressed a way of obtaining a representative sample of smartphone users in the general population, or the sample size required to do so.

6.1 Prior sampling frames for larger-scale smartphone surveys

Three of the larger-scale studies that have been found are the Future Mobility Survey done in Singapore, the rMove application by RSG, and the ATLAS II survey done in New Zealand (Cottrill, et al., 2013; Hathaway, et al., 2015; Safi, et al., 2015). All three of these surveys contacted people who had participated in a previous traditional household survey, with the Future Mobility Survey obtaining a sample size of 1,500 participants (Zhao, et al., 2015).

The surveys did not necessarily explore in depth how representative these samples were of the general population. For the RSG app, only the households that showed an interest in a smartphone survey were invited. Furthermore, households that were not eligible either because all members did not have smartphones were also not invited. From these, it was found that biases would clearly be introduced towards households in which all members have smartphones, and as such, were not necessarily representative of all households. This may change in the future as smartphone adoption is increasing at a very rapid rate.

The ATLAS II survey was biased towards only iPhone users as they found that around half of the participants in the households had iOS devices. As a result, only an app for that platform was developed.

6.2 Cellphone sample frames in the Canadian context

In Canada, there are only a couple of marketing firms that provide some sort of cellphone sampling frames. One such company is ASDE Survey Sampler. They create cell phone samples by generating random numbers from the list of existing dedicated cellphone exchanges. There is no definitive way to know the geographic information for each number as the information is only limited to area codes given by the first three digits of the number. In reality, this does not mean that the user is residing within that area, as users do not need to change their numbers if they move. Additionally, there is no information about whether this cellphone number belongs to an individual or a corporation, or whether the number is active at all. According to ASDE, a connection rate in the range of 50-55% is achieved using this method, but the response rate will likely be significantly lower. ASDE does assist in determining working cellphone numbers and cellphone only (CPO) households. However, this comes at a much higher cost. In the U.S., random digit dialing (RRD) of cellphone samples can be 2 to 3 times more expensive to field than RRD landline samples (Baker, et al., 2011).

According to the report titled Secondary Research into Cell Phones and Telephone Surveys 2012, at least half of all 18-34 year olds use only a cellphone, so it is vital to find a way to obtain a proper cellphone sampling frame, as this is the demographic that is underrepresented in the current TTS. Dual frame surveys using both landline and cellphone lists have become common in the U.S. when surveying the general population. The report goes into more detail regarding overlapping and non-overlapping design (the latter is where the cellphone sample is screened to obtain cellphone only households to ensure no overlap between the two frames), the differences with respect to cost, and ease of analysis. It also discusses the possible differences in data quality. A brief summary of this report can be found in the annotated bibliography. (Phoenix Strategic Perspectives Inc., 2012)

Given the difficulty in obtaining a proper cellphone sampling frame, it is even more difficult to figure out how many of those cellphones are smartphones. The percentage of smartphones in Canada is increasing rapidly, however, so it may be that the majority of cellphones will be smartphones in a few years, alleviating the additional step required to obtain a smartphone sample frame.

6.3 Panels used for smartphone sampling frames

Many of the smartphone-based surveys and studies done, apart from the ones mentioned in Section 6.1, use a panel that is meant to be representative of the general population. One of the large-scale smartphone surveys, the Dutch Mobile Mobility survey uses the Longitudinal Internet Studies for the Social Sciences (LISS) panel (also known as CentERdata) that consists of 8000 ongoing participants, from which 600 were recruited for the survey. As mentioned before, around 60% of the participants were loaned a smartphone as they did not own one of their own. This same panel was also used in a study that compared the results between surveys done on a mobile device and a computer, although only those with smartphones were recruited. (Bruijne & Wijnant, 2013)

Another similar survey used the U.S. probability-based online panel KnowledgePanel (Wells, et al., 2014). This panel offered a pre-screened group of willing online panelists who owned smartphones. Out of 25,221 panelists, 10,156 (42%) responded, and 2,443 were identified as smartphone owners who were willing to complete the survey on their smartphone. Recruitment for the panel is done using an address-based sample frame (obtained from the US Postal Service's Computerized Delivery Sequence File), which can yield more cell-phone only households than using random digit dialing sampling. This panel also has an office in Canada (Anon., 2015).

Greaves et al. (2015) also used an online panel as one of the ways of recruitment, to study whether the method of recruitment had any effect on the responses themselves. Invitations were sent to several hundred eligible members of the panel, but only 87 were recruited this way. The authors deduced that this poor response rate was mainly due to the nature of the survey being different than the ones the panelists were accustomed – this survey involved a substantial commitment over (potentially) several years, while typically the surveys were 10-15 minutes long.

Apart from convenience sampling or recruiting from the traditional household survey, this appears to be the method used in literature to recruit for smartphone-based surveys. Convenience sampling is usually not representative of the entire population and can be biased towards certain demographics. For traditional household surveys, there is an underrepresentation of younger demographics when landline sample frames are used, as is the case in the current TTS. As such, for the upcoming TTS pilots, it may be best to recruit from a Canadian market research panel.

6.4 Section 6: Recommendations

- If a representative cellphone sample frame becomes available and is used for the core survey, a smartphone survey could be done on a subset of those recruited for the core survey, similar to the Future Mobility Survey and the ATLAS II project
- Pilot recommendation: obtain a sample from a Canadian market research panel for the purposes of a smartphone-based survey pilot

7 MULTI-DAY OR –WEEK, HOUSEHOLD VS. INDIVIDUAL DATA COLLECTION, AND CORE VS SATELLITE SURVEYS

TABLE 5 CORE/SATELLITE, HOUSEHOLD/INDIVIDUAL AND MULTI-DAY SURVEYS

Title / Reference	Core / Satellite	Household/ Individual	Length of Survey or Data Collection
(Greaves, et al., 2014) A Web-Based Diary and Companion Smartphone App for Travel/Activity Surveys	Core (web + optional smartphone app)	Individual	7 days
(Montini, et al. 2014) Personalisation in multi-day GPS and accelerometer data processing	N/A	Individual	7 days (collection)
(Vlassenroot, et al., 2015) The Use of Smartphone Applications in the Collection of Travel Behaviour Data	N/A	Individual	2 weeks
(Nitsche, et al. 2012) A Strategy on How to Utilize Smartphones for Automatically Reconstructing Trips in Travel Surveys	Suggests smartphones as satellite	Individual	266 hours of travel data were collected
(Cottrill, et al., 2013) Future Mobility Survey **(Zhao, et al., 2015) Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore**	Satellite	Individual, recruited from household survey	2 weeks, validation required for at least 5 days
(Houston, et al., 2014) Tracking Daily Travel: Assessing Discrepancies between GPS-Derived and Self-Reported Travel Patterns	Suggests smartphones as satellite	Individual, recruited from household addresses	7 days (dedicated GPS device)
(Safi, et al., 2015) Design and Implementation of a Smartphone-based System for Personal Travel Survey: Case Study from New Zealand	Satellite	Individual, recruited from household survey	3 days
(Phoenix Strategic Perspectives Inc., 2012) Public Works and Government Services Canada Report – Secondary Research into Cell Phones and Telephone Surveys	N/A	Discussion of household surveys	N/A
(Fan, et al., 2012) UbiActive: A Smartphone-Based Tool for Trip Detection and Travel-Related Physical Activity Assessment	N/A	Individual	3 weeks
(Greaves, et al., 2015) Participation and Diligence in a Multi-Day, Multi-Year Survey: Impact of Recruitment Methods and Demographics	Core (web + optional smartphone app)	Individual	7 days (2 waves)
(Jariyasunant, et al., 2012) Overcoming Battery Life Problems of Smartphones when Creating Automated Travel Diaries	N/A	Individual	Field test: 3 months
(Thomas, et al., 2014) Dutch Mobile Mobility Panel (Hoe mobiel zijn we nu eigenlijk? Eerste	Core (recruited	Individual	2 weeks

inzichten uit het Mobiele Mobiliteitspanel)	from panel)		
(Fan, et al., 2015) SmarTrAC: A Smartphone Solution for Context-Aware Travel and Activity Capturing	N/A	Individual	7 days

7.1 Multi-day Data Collection

Numerous studies have been performed on the viability of using smartphones to infer travel diaries and collect location data, and many of them have been for multi-day or even multi-week periods, as seen in **Table 5** (Vlassenroot, et al., 2015; Cottrill, et al., 2013; Safi, et al., 2015; Fan, et al., 2012; Jariyasunant, et al., 2012; Thomas, et al., 2014; Fan, et al., 2015; Hathaway, et al., 2015). Some of these studies have been conducted over 2 weeks or more, but participants have still shown a willingness to engage regularly (Cottrill, et al., 2013), continue collecting data beyond the required days (Safi, et al., 2015), and even participate in future studies (Greaves, et al., 2014). As such, from the user's side, using smartphones appears to be the best instrument to collect data over multiple days, especially if these apps become more passive. With the increase in data that can be obtained to analyze patterns and behaviours, the functionality of the app and the accuracy of the predictions would also increase.

As a follow-up to the Future Mobility Survey (FMS), Zhao et al. (2015) analyzed the collected data and explored the various travel patterns that could be obtained from multi-day smartphone surveys. They categorized trip patterns into five clusters. Cluster 1 consisted of working days with lunch breaks in mid-day, and many non-work activities after work. Cluster 2 were days that were mainly working, with few other activities. Cluster 3 had days with shorter work hours and many work-related activities. No work days were placed in Cluster 4, with many non-work activities, likely for weekends or off-days. Finally, Cluster 5 were days that were mainly spent at home or traveling.

They found that none of the HITS (the in-person household survey to which the FMS data are compared) days fell in Cluster 1, highlighting the low reporting rates of out of work activities in regular household surveys. They also noted that for any one user, they observed a lot of variability from day to day.

A one-day travel survey cannot capture this pattern and behaviour. While this level of detail may not be necessary for aggregate travel demand modeling, it is required for more sophisticated agent based microsimulations.

HITS also showed narrower departure and arrival time trip distributions than FMS, which agrees with the idea that people tend to report a "typical" day in self-reported surveys, but in reality, travel times have a wider spread. Specifically, the HITS data shows spikes at 20, 30 and 60 minute intervals, which are the times up to which people normally round. Another observation was that the average travel times recorded in FMS was lower than in HITS, confirming that people tended to report travel times longer than they actually were.

The FMS data showed three clear peaks during the day that corresponded to breakfast, lunch and dinner, but this trend was not clear in the HITS data. There was much less reporting of these trips made by HITS respondents. This clearly shows the advantage of smartphones in overcoming the issue of trip under-reporting that is present with other survey methods.

The study also showed that there is under-reporting of trips towards the end of the day in self-reported household studies. Most user reported that they arrived home by 8pm but the FMS data showed that a significant portion of the users reached home after 9pm.

Similarly, the Dutch Mobile Mobility Panel also did a basic analysis of trips made throughout the week. They found, for example, that more people go running or biking on Sundays, and that more passenger trips are made on weekends/holidays. The latter was presumably due to more family trips being made on these days.

All of these trends and variability in travel could only be confirmed because of the wealth of data that a multi-day survey can provide, and shows that a one-day travel survey does not capture the full picture. There is a minimal marginal cost of collecting additional days with smartphone surveys, and user burden decreases over time as they become more familiar with the system. Additionally, the back-end systems can learn the user's patterns and can make more accurate predictions over time, especially if the user validates their trips (Cottrill, et al., 2013; Greaves, et al., 2014; Li, et al., 2011; Jariyasunant, et al., 2012).

As a side note, it should be kept in mind that users' willingness to participate could be due to interest in the potential application of new technologies. As smartphones become more powerful and widespread, and once the novelty of these apps wear off, response and interest could decline. A similar effect was observed with response rates to email surveys (Bailey, et al., 2011). Nitsche et al. (2012) also note that smartphone surveys may be biased towards those with a technical interest.

Furthermore, while multi-day data can easily be obtained, using a smartphone for capturing travel data alao ensures that short trips, usually walking, are not missed – something which is quite prevalent in self-reported household surveys. Most of the cited studies in this section report that more trips are detected through smartphones, especially shorter walking trips, which are normally missed. Additionally, the duration of trips is consistently over-estimated in self-reported trips and that people round up their travel times to the nearest 5 or 15-minute interval. Houston et al. (2014) found that the difference between GPS and self-reported walking trips was about 11.8 minutes.

For many of these studies, for example in the Future Mobility Survey, the respondents were assigned specific dates between which they were asked to use the app and validate their data. This bypasses the issue that may occur with convenience sampling surveys, where respondents start using the app in the middle of the day, thus making the data for that day unusable. Otherwise, at least two full days of data are needed for a given user in order to make a travel diary for them. If people only run the app for one day, the data are not usable.

7.2 Household vs. Individual Data Collection and Core vs. Satellite Surveys

Since a large number of trips are made jointly with other members of the household, the trip information for all members is required for use in household travel models. This is used, for example, in analyzing the implementation and pricing policy of HOV (High Occupancy Vehicle) lanes, understanding trip chaining and travel behaviour of multiple-person households, and for a multitude of other purposes.

Surveys done via smartphone apps are usually individual-based, as seen from **Table 5**. Data collected through various sensors and processes will relate only that particular smartphone, which is normally associated with only one person. As such, the information collected will not be on a household level as trips made by other members of a household will not be captured by one smartphone.

7.2.1 Options for capturing household-level interaction from individual survey tools

The Future Mobility Survey (FMS) and the ATLAS II project are individual surveys, but both were conducted as extensions of their traditional household travel surveys – as satellites. The FMS survey (Cottrill, et al., 2013) and the Dutch Mobile Mobility Survey (Thomas, et al., 2014) had sections where users could input supplementary information regarding trips made, such as whether other household members accompanied on the trip. The surveys also asked other general questions, either before or after the data collection period, to acquire household level information such as the number of vehicles, number of people in the household, number of drivers, and whether members have transit passes. There were all questions that would normally be asked in a household survey regarding the characteristics of members and the makeup of the household but without getting the trip details of each member. The FMS survey asked these questions at the end of the data collection period in order to avoid off-putting respondents with the length of the survey if it was encountered first.

Some inference about household characteristics and the effect on travel behaviour can be made from these questions, but it does not provide complete household trip information. The FMS used the satellite smartphone survey to determine the amount of under-reporting of trips and the difference in reported trip duration versus actual trip duration present in regular household surveys; the results were used to correct the data from household surveys. They also mentioned that the rationale for continuing with an individual survey was because this was the first implementation of large-scale smartphone-based survey, and they did not want to limit the number of respondents. However, such a use of a smartphone survey (to correct issues in responses from a more traditional method) is one that should be explored.

As mentioned earlier, another vital component that could be captured through this method is the variability in day-to-day trip activities of a particular individual. The FMS found that there was a lot of variability in trip patterns from day to day, none of which was captured in the regular household survey. This information would help in providing more context to the collected data, such as a better idea of clusters of activities and activity patterns. This could in-turn help in understanding and improving transportation models, especially agent-based microsimulations.

Another alternative that could be viable in the future is to simply survey everyone in the household through smartphones. Since smartphone penetration is rising exponentially (33% to 73% from 2012 to 2014 in Canada - (Google Canada, 2013)), it might the case that by 2021, most, if not all, cellphones are actually smartphones. As such, if one person in a household is recruited, this could be extended to all members of the household through the recruited member. This was the method used in the rMove survey conducted by RSG (Hathaway, et al., 2015). They recruited participants from the traditional household survey, but only households that were eligible were selected.

7.2.2 Core versus satellite recommendation in literature for smartphone surveys

To date, smartphone surveys have not been used as core surveys due to sampling frame issues and getting hold of a large enough sample of willing individuals people that have smartphones. Nitsche et al. (2012) and Houston et al. (2014) recommend that for now, smartphones are suitable as satellite surveys mainly because a sample representative of the general population cannot be obtained easily from smartphone users. Even the larger-scale studies cited in this report (the Future Mobility Survey, rMove and ATLAS II), as discussed in the previous sub-section, use smartphones as a satellite survey with a landline or face-to-face interview core.

7.2.3 Suitability of smartphones as a core survey

Smartphones could potentially be used as the core survey instrument in the future. According to the Transportation Association of Canada (TAC) report (November 2012), characteristics of a core survey are that it:

1. includes key data that are fundamental to the agency's primary policy/planning needs

2. includes attributes of the respondents that permit core data to be linked to common variables in satellite surveys
3. has a sample size large enough to make statistical inferences concerning variables of interest
4. is expandable to make statements about the full population
5. is consistently applied over a large geographical region
6. is stable (but not necessarily static) over time and is applied relatively frequently
7. is relatively short, so as to minimize response burden and to permit large sample sizes to be cost-effectively collected

Currently, smartphone-based surveys can satisfy most of these requirements depending on how they are designed and implemented. One of the main advantages of using smartphones as the core survey is the scalability and low cost of distribution of the survey. While up-front costs for development and testing might be high, as the app matures, incremental costs should be reduced, allowing for longer data collection periods and recruitment of more households (Hathaway, et al., 2015). Other advantages include the lower response burden for participants, the ability to be used for multiple day or week surveys, and the improvement in data quality by reducing recall bias and error, as discussed in previous sections.

The only current issue with smartphone-based surveys as the core are with respect to requirements 3 and especially 4 listed above, which have to do with the question of whether smartphone users are representative of the general population.

If smartphones were to be used as the core survey today, the inverse of the issues faced by TTS could happen. Currently, younger demographics are underrepresented in the TTS, while older generations are overrepresented. If smartphones were to be used as a core survey, younger generations might be overrepresented (Vlassenroot, et al., 2015), requiring satellite surveys to cover the demographics missed in order to balance this issue.

However, smartphone penetration is increasing at a very rapid rate in Canada, from 33% in 2012 (Google Canada, 2013) to 74% in 2014 (J.D. Power and Associates, 2014). At this rate, most of the population could potentially own a smartphone device in time for TTS 2.0. While the 18-24 year old generation were the first to adopt using smartphones, the share of smartphone owners in the 25-34 and 45-54 demographics increased most noticeably in 2015 (Catalyst, 2015). Even older generations (55+) are adopting smartphones, as regular feature phones are becoming increasingly more difficult to buy. However, the 55+ demographic do not necessarily use all the features available or download apps at the rate of other demographics (Deloitte, 2013).

As such, by TTS 2.0, smartphone users could provide a representative sample of the population. While smartphones may not be perfectly suitable as a core survey today, they could become so in the near future.

7.3 Recommendations

Multi-day:

- Conduct a 7-day smartphone-based survey to capture variability of travel patterns and behaviour over the week.
- If convenience sampling is used, at least 2 full days of data collection are needed from the user in order to compile a travel diary for them.
- If longer surveys are needed to help improve the accuracy of location, mode and activity detection/inference of the app, this should not be an issue as other studies have gone for longer than a week and participant engagement has still been high.

Household/Individual:

- If it is possible to obtain sample frames of households in which all eligible members have smartphones, recruit all members of the household in the smartphone survey, thereby still maintaining a household level survey.
- If conducting an individual survey, ask for household characteristics (such as number of vehicles, people with driving licenses/transit passes, whether they take transit, whether they ever bike, etc.) before the main data collection period.
- While it is ideal to acquire this information prior to the data collection period for detection and inference accuracy purposes, it can overwhelm the respondent. Therefore, it might be better to only include questions that are necessary for background processing, and move the remaining questions to the exit survey when the respondent is already fully committed.
- Additionally, provide personalized, pre-set responses for capturing household interaction. For example: if a user is validating a trip they made during the day, provide options for selecting if other members of the household accompanied on the trip, and which owned vehicle they used (similar to rMove app).

Core/Satellite:

- Use smartphone-based surveys as a satellite instead of a core. Based on smartphone adoption over the next few years, however, smartphones may become a viable tool for a core survey. If smartphones are used for the core survey, satellite surveys would be required to cover the underrepresented demographics; this would simply be the inverse of the current situation.

8 PRIVACY, RESPONSE RATES AND RECRUITMENT

8.1 Privacy

Due to the nature of the data collected, even if the names of respondents are not obtained and stored, it is possible to infer a substantial amount of information about the respondents, such as their home and work locations, and other more personal information.

There are significant privacy concerns surrounding smartphone usage in general, and this can affect response rates (Phoenix Strategic Perspectives Inc., 2012; Safi, et al., 2015). It is, therefore, of vital importance to communicate with potential respondents about how their data will be used, stored and shared. Studies that have looked at privacy specifically with smartphone surveys have shown that clear and simple communication can increase response rates significantly (Bouwman, et al., 2013).

Bouwman et al. (2013) specifically address this concern and provide ways to minimize its impact on survey response rates. Other studies cited in this report also provide similar suggestions. They suggest that:

- information should be stored securely using a password-protected login
- no personally identifiable information is directly collected
- only the researchers involved with the project can have access to disaggregate information (if required)
- disaggregate information should not be allowed to be published or shared with anyone else
- data needs to be collected and analyzed by separate entities, and once the information is passed onto the researchers, it should be deleted by the company collecting the information after a prescribed period of time

Finally, they also recommend documenting these policies and privacy regulations on a website related to the study, and creating a user-agreement and privacy-related FAQ. Potential respondents should be directed to these before being asked to provide responses so as to ensure their participation is done with informed consent. In the exit survey of their study in the Netherlands, most respondents were satisfied with the privacy warranty and trusted the researchers, and more than 70% of respondents indicated that they would be willing to participate again (Bouwman, et al., 2013). Of potential respondents, 12% declined to participate in the study due to privacy concerns.

8.2 Recruitment, Response Rates, and Incentives

In general, recruitment of smartphone users tends to skew towards younger demographics with lower income or towards people with an interest in technology (Nitsche, et al., 2012; Zhao, et al., 2015; Phoenix Strategic Perspectives Inc., 2012). Since landline surveys tend to underrepresent these very demographics, satellites that consists of smartphone-based surveys could be used to balance this.

TABLE 6 RECRUITMENT METHOD, INCENTIVES AND RESPONSE RATES

Title / Reference	Recruitment Method and Incentives	Response RAtes
(Greaves, et al., 2014) A Web-Based Diary and Companion Smartphone App for Travel/Activity Surveys	Online sources, telephone, face-to-face intercept / **$50 AUD**	**89%** completed all 7 days
(Cottrill, et al., 2013) Future Mobility Survey	**$25** upon completion. Recruitment for pilot via social networking sites, posted flyers, personal contacts – convenience sampling	74 people completed presurvey, 50% of which installed the app, 36% actually validated the data
(Zhao, et al., 2015) Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore		1541 recruited, 793 completed (**51%** response rate)
(Houston, et al., 2014) Tracking Daily Travel: Assessing Discrepancies between GPS-Derived and Self-Reported Travel Patterns		1 % response rate overall (27,275 household addresses obtained, 279 completed)
(Safi, et al., 2015) Design and Implementation of a Smartphone-based System for Personal Travel Survey: Case Study from New Zealand	186 eligible participants of regular household survey were asked to participate in the smartphone survey	41% response rate
(Hu, et al., 2011) The Impact of a Mixed-Mode Data Collection Design on Response and Non-Response Bias on a RDD Landline Telephone Survey		Boosted response rate by 10 %
(Millar & Dillman, 2012)	$2 incentive. Contacted 5 times, alternating between mail and email	Similar response rates for online and mobile: 50%. Group offered choice of online or paper had 56% response rate.
(Millar & Dillman, 2011)	$2 advance cash incentive	Web plus email augmentation group (contacted and reminded via postal mail and email alternatively) produced the highest response rate of all: 59.7%
(Greaves, et al., 2015) Participation and Diligence in a Multi-Day, Multi-Year Survey: Impact of Recruitment Methods and Demographics	$50 AUD. Recruitment done using a professional survey firm, cold-calling (main method), electronic circulation lists, social media, face-to-face, mailbox drop	Obtained an attrition rate of 32% (i.e. 32% did not return for Wave 2 of survey). highest retention rate were those who were recruited by mailbox drop

(Murphy, et al., 2014) Social Media in Public Opinion Research: Report of AAPOR Task Force on Emerging Technologies in Public Opinion Research		Review of social media and how they are being used to conduct surveys, limitations and potential
(Thomas, et al., 2014) Dutch Mobile Mobility Panel (Hoe mobiel zijn we nu eigenlijk? Eerste inzichten uit het Mobiele Mobiliteitspanel)	Incentive amount not mentioned. Assumed standard for panel.	
(Fan, et al., 2015) SmarTrAC: A Smartphone Solution for Context-Aware Travel and Activity Capturing	Convenience recruitment	

8.2.1 Response rates using mixed modes

Multiple modes of response and recruitment are usually possible, and can be used for the purposes of increasing response rates. Millar and Dillman (2012) suggested that the most effective strategy is the combined use of multiple response-inducing techniques, that is contacting and reminding through both mail, email, telephone, etc. *and* offering different modes of response or choice of modes at different stages. Millar and Dillman (2011) argued that the mode of response should be the same as the mode of recruitment. Asking respondents to switch to another mode creates an additional burden. For example, asking respondents via mail to respond to a web survey requires them to take an additional step by typing in the address for the website. This burden can lead to non-response, even if web is their preferred method of answering surveys (in this example).

They suggested that first an invitation should be sent by mail offering only a response by mail in return. If no response is obtained, then the invite should be sent through email for a web survey. This would take advantage of multiple contact modes by establishing memorability and presence and can reduce burden of responding by providing a link to the website directly when contacted through email (Millar & Dillman, 2011). They also provided a $2 incentive in advance via mail, and showed that this could increase response rates by 17%. They cited previous studies that showed that this advance incentive was more effective than raffle draws or gift certificates. Hu et al. (2011) also reinforced that using a mail-follow up survey to non-respondents of a landline telephone survey could boost response rates (by 10% in their study). The drawback of this method is that all these methods for contacting an individual need to be known, such as home address, email address, etc.

As an example of using only household addresses as a recruitment method, Houston et al. (2014) purchased such a list from InfoUSA, a marketing information firm. They mailed 27,275 households with an invitation letter, and all households who indicated that they were interested in participating (651) were invited to the study. Only 279 household provided a complete response – a 1% overall response rate (Houston, et al., 2014). The drawback with this method is that the cost per respondent could be quite high in this case, even before incentives come into play.

8.2.2 Recruitment methods in large-scale surveys in the literature

Various different recruitment methods have been utilized in the smartphone-based survey literature. The three main surveys that attempt to poll the general public, the Future Mobility Survey (FMS), the rMove app and the ATLAS II project, all recruited from participants of a traditional household survey. For example, in the FMS, once the regular household survey is completed, the interviewer informs the participants about the FMS survey and whether they would be interested in this survey. Of the 1,541 who were recruited in this way, only 793 completed the smartphone survey (51% response rate). For their pilot survey, they used social networking sites, posted flyers, personal contacts, and included a $25 incentive upon completion. Seventy-four people completed the pre-survey, but only 36% of them validated the data.

Safi et al. (2015) and the rMove app (Hathaway, et al., 2015) also recruited participants from their traditional household survey. The former found that almost half the participants had iPhones, so they targeted iPhone users specifically and did not develop an Android app. Of the 186 eligible participants that were recruited, only a response rate of 41% (76 respondents) was achieved.

Apart from these three studies, one of the biggest smartphone-based surveys is the Dutch Mobility Mobile Survey, which used participants recruited from the LISS panel. This panel has a pool of 5000 households/8000 panelists, and is meant to provide a representative view of the Dutch population. Nearly 800 participants who had previously expressed interest in a smartphone based research were contacted (therefore it was not an unbiased, random sample). The final sample size was 591 participants, who were also given a stipend for their participation (although the exact amount is not provided).

The other studies that involved smartphones were mainly pilot studies that had much smaller sample sizes that were recruited through convenience sampling, and this seems to be the most common way of conducting pilots since there is no readily available sample frame of smartphone users. (Fan, et al., 2015; Fan, et al., 2012; Abdulazim, et al., 2013; Fan, et al., 2015; Jariyasunant, et al., 2012)

8.2.3 The effect of recruitment method on participant responses

Greaves et al. (2015) analyzed the impact of various recruitment methods on participation, diligence and retention rates on a multi-day, multi-year survey. Even though they provided a $50 AUD incentive, recruitment was a challenge. They used different methods, first by using a professional survey firm. The main method of recruitment was cold-calling spatially-targeted home telephone number, where respondents were asked for their emails if they agreed to participate. In this way, 20,410 numbers were used, but only 415 (~2%) participants completed the questionnaire phase. They also used electronic circulation lists (136 participants recruited), social media (39 recruited), face-to-face recruitment around breakfast events (105 recruited), and mailbox drop.

They found that the age of respondents corresponded heavily with recruitment method – younger recruits more likely from electronic recruitment (89%), while older recruits more likely to come from cold-calling (79%). Reported mode of travel varied significantly by recruitment method – cyclists were more likely found through intercept and social media recruitment methods.

One important goal was to determine whether completion and retention rates differed solely because of different demographics obtained by the corresponding recruitment methods, or whether the recruitment methods themselves were responsible in some way, affecting the completion and response rates. They studied and surveyed those who completed both waves and those who did one but not the other; they also examined diligence (i.e. whether the participants would fill all 7 days of the diary). This analysis resulted in a finding that none of the recruitment methods explained participant attrition (i.e. not returning for the second wave) or diligence.

The findings show that, in terms of diligence and attrition, how people were recruited into the study was not a subsequent factor in their participation in the study. This means that recruitment efforts can focus on how to reach different demographics through various recruitment methods without worrying about the effects of said recruitment methods on the survey results.

8.2.4 Recruitment through social media

Social media can also be used for recruitment, but these can be biased towards certain demographics, and full research has not been done on the specific characteristics of these demographics (Murphy, et al., 2014). This cited report reviews how social media is currently being used for recruitment purposes, discusses the demographics of social media users, and what future potential this platform might hold.

Currently, researchers are using social media to obtain qualitative insights. More research is needed to determine whether a sampling frame can be obtained where individuals, who are social media users, can be sampled with a known and non-zero probability. It is still yet to be ascertained whether social media may represent a subset of the general population, and more importantly, what the demographic makeup of that subset is specifically.

One method of recruitment is to create a group, on Facebook for example, and target specific groups and users of interest to join that group. It is easy way to reach a specific demographic without using much resources – but is not a probabilistic sample.

Another way to get a more diverse set of respondents is a pay-per-click ad. The more the researcher pays, the more heavily featured their ad will be to "active" users (i.e. those individuals who click on ads with great regularity). This can also be used to do a targeted campaign, for example if looking for people between the ages of 18-35 only, the ads will be targeted more towards these demographics. Other factors that can be targeted include education, hometown. Previously, ads were only shown in the full, desktop version of the Facebook website, but now these are also shown in the mobile versions.

Some of the other challenges with social media include that the popularity of certain social media websites rises and falls over time, and it is difficult to predict whether one will be the most popular at some future point in time. Each has their own set of rules and protocols in terms of data availability and privacy, which are subject to change any time. The demographics on each website also vary. There are issues concerning cases when sometimes users can have multiple accounts, or many accounts belong corporations and businesses as opposed to individuals. There can be also be biases within the social media frame – individuals who never or rarely post may be missed by certain sampling techniques or might be systematically under-sampled. This can bias results of data collection towards the heaviest users. Little progress has been made to show how data collected through social media can represent the general population – so far only non-probability samples have been obtained.

8.3 Recommendations

- If possible, offer users the same mode of recruitment as the intended mode of response. If email address are somehow already available, this should be used as the recruitment mode directly. However, if these are not readily available, a “pre-survey” could be done via mail or telephone, where a participant is asked for their email address. Then, the link to the survey or the app should be sent to the user through their email address such that they can open it directly on their smartphone.
- In the latter case, the method of first contact with the participant could be through any mode, as this does not affect the responses of the participant (note: this is not the response rate but whether participants answer differently depending on the recruitment mode). However, different demographics could be obtained using different modes of recruitment, thus reaching out to a potentially broader population.
- Pilot recommendation: Incentives amounts vary over a large range in current literature, from $2 to $50 per participants. Due to the size of the TTS, larger amounts may not be feasible as this would incur a considerable cost. The Future Mobility Survey use $25 incentive upon completion per participant. This could be the topic of further study in a pilot, with perhaps a test range of $2 to $25.

9 PILOT RECOMMENDATIONS

Section 3 Recommendations:

- Testing out and optimizing these methods without affecting battery life significantly:
  - Sensors such as accelerometers, in addition to GPS data, should be used to detect travel mode, as it can be done without using up too much of the processing power or the battery of the smartphone.
  - Incorporating GIS spatial information, such as the road and transit network information, can help increase mode detection.
  - Incorporating user history and information, such as whether the user ever takes transit, or has a bike, etc. can help in increasing mode detection as well.
  - In terms of machine learning, random forests seem to be the most successful and widely used in the relevant literature.
  - Sampling less GPS data to maximize battery life while still obtaining a relatively high accuracy can be done, similar to (Ghorpade, et al., 2015).
  - Use advanced rule-based methods for activity inference (see Section 3.3.2).
  - Use an activity-based approach, similar to Montini et al. (2014), using personalized data to help determine trip activity. Simply including distance of the trip from a user's home or work location (and not including any other personal data like age, income, etc.) noticeably improved the accuracy from the base case.
  - Also include previous user validations for determining trip activity.
- Finding the right balance in polling other sensors and then turning on GPS – one should avoid delaying the turning on of GPS since this leads to loss of data, but, at the same time one does not want to constantly switch on GPS for false starts.
- Developing and testing trip start/end detection algorithms.
- Test methods and algorithms utilizing motion sensors and other smartphone data to differentiate different activities at a particular location, as a continuation of the works described in Section 3.3.3

Section 5 Recommendations:

- Determine the threshold distance after which the GPS should be switched on. Make it too large and the location data will not be accurate, while having the distance too small will decrease battery life by sampling the GPS too often.
- Similarly, determine the best frequency for polling GPS while in motion while maintaining data accuracy, and battery life – or implementing algorithms that improve detection accuracy with reduced data.
- Test the app developed on different phone models and brands since performance can vary significantly.

Section 6 Recommendation:

- Obtain a sample from a Canadian market research panel for the purposes of a smartphone-based survey pilot.

Section 8 Recommendations:

- Incentives amounts vary over a large range in current literature, from $2 to $50 per participants. Due to the size of the TTS, larger amounts may not be feasible as this would incur a considerable cost. The Future Mobility Survey use $25 incentive upon completion per participant. This could be the topic of further study in a pilot, with perhaps a test range of $2 to $25.

10 BIBLIOGRAPHY

Abdulazim, T., Abdelgawad, H., Habib, K. M. N. & Abdulhai, B., 2013. Using Smartphones and Sensor Techologies to Automate Collection of Travel Data. *Transportation Research Board,* Issue No. 2383, pp. 44-52.

Anon., 2015. *GFK KnowledgePanel.* [Online]
Available at: http://www.gfk.com/us/Solutions/consumer-panels/Pages/GfK-KnowledgePanel.aspx
[Accessed 31 August 2015].

Bailey, J. et al., 2011. *Can Your Smartphone Do This: A New Methodology For Advancing Digital Ethnography,* s.l.: AAPOR.

Baker, R., LaFrance, J. & Peterson, G., 2011. *Executing Household Telephone Surveys in a Wireless World.* [Online]
Available at:
http://www.marketstrategies.com/user_area/content_media/Telephone_Surveys_in_a_Wireless_World.pdf
[Accessed 25 August 2015].

Bouwman, H., Reuver, M. d., Heershcap, N. & Verkasalo, H., 2013. Opportunities and problems with automated data collection via smartphones. *Mobile Media Communications,* Issue 1, pp. 63-68.

Bruijne, M. d. & Wijnant, A., 2013. Comparing Survey Results Obtained via Mobile Devices and Computers: An Experiment With a Mobile Web Survey on a Heterogeneous Group of Mobile Devices Versus a Computer-Assisted Web Survey. *Social Science Computer Review,* 31(4), pp. 482-504.

Catalyst, 2015. *With Growth Comes Change: The Evolving Canadian Mobile Landscape.* [Online]
Available at: http://catalyst.ca/2015-canadian-smartphone-market/
[Accessed 28 August 2015].

Cottrill, C. D. et al., 2013. Future Mobility Survey. *Transportation Research Board,* Issue No. 2354, pp. 59-67.

Das, R. D., Ronald, N. & Winter, S., 2015. *A Simulation Study on Automated Transport Mode Detection in Near-Real Time using a Neural Network.* [Online]
Available at: http://ceur-ws.org/Vol-1323/paper24.pdf
[Accessed 20 August 2015].

Data Management Group, 2011. *Transportation Tomorrow Survey,* Toronto: Transportation Information Steering Committee.

Deloitte, 2013. *2013 Global Mobile Survey.* [Online]
Available at: http://www2.deloitte.com/global/en/pages/technology-media-and-telecommunications/articles/2013_GlobalMobileSurvey.html
[Accessed 20 August 2015].

Fan, Y., Chen, Q., Liao, C.-F. & Duoma, F., 2012. *UbiActive: A Smartphone-Based Tool for Trip Detection and Travel-Related Physical Activity Assessment.* s.l., s.n.

Fan, Y. et al., 2015. *SmarTrAC: A Smartphone Solution for Context-Aware Travel and Activity Capturing,* s.l.: Volpe Center at U.S. Department of Transportation in support of the Intelligent Transportation Systems Joint Program Office.

Ghorpade, A., Pereira, F. C. & Zhao, F., 2015. *An integrated stop-mode detection algorithm for real world smartphone-based travel survey*. [Online]
Available at: http://www.researchgate.net/publication/275949958_An_integrated_stop-mode_detection_algorithm_for_real_world_smartphone-based_travel_survey
[Accessed 20 August 2015].

Gong, H., Chen, C., Bialostozky, E. & Lawson, C. T., 2012. A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems,* Volume 36, pp. 131-139.

Google Canada, 2013. *Our Mobile Planet: Canada - Understanding the Mobile Customer*. [Online]
Available at: http://services.google.com/fh/files/misc/omp-2013-ca-en.pdf
[Accessed 10 August 2015].

Greaves, S. et al., 2014. *A Web-Based Diary and Companion Smartphone App for Travel Activity Surveys.* Leura, Australia, s.n.

Greaves, S. P., Ellison, A. B. & Ellison, R. B., 2015. *Participation and Diligence in a Multi-Day, Multi-Year Survey: Impact of Recruitment Methods and Demographics,* s.l.: s.n.

Harding, C., Zhang, Y. & Miller, E. J., 2015. *Multiple purpose tours and efficient trip chaining in Toronto: an analysis of the effects of land use and transit provision on mode choice and trip chaining using smartphone data,* s.l.: s.n.

Hathaway, K. et al., 2015. *"In the Moment" Travel Study: A GPS Smartphone Application.* Atlantic City, New Jersey, RSG.

Houston, D., Luong, T. T. & Boarnet, M. G., 2014. Tracking Daily Travel: Assessing Discrepancies between GPS-Derived and Self-Reported Travel Patterns. *Transportation Research Part C: Emerging Technologies,* Volume 48, pp. 97-108.

Hu, S. S., Pierannunzi, C. & Balluz, L., 2011. *The Impact of a Mixed-mode Data Collection Design on Response and Non-response Bias on a RDD Landline Telephone Survey,* s.l.: AAPOR 2011.

J.D. Power and Associates, 2014. *2014 Canadian Wireless Total Ownership Experience Study*. [Online]
Available at: http://www.jdpower.com/press-releases/2014-canadian-wireless-total-ownership-experience-study

Jariyasunant, J., Sengupta, R. & Walker, J., 2012. *Overcoming battery life problems of smartphones when creating automated travel diaries.* Toronto, s.n.

Jue, Aaron, 2014. *Trends Report: Mobile Participation in Online Surveys,* s.l.: FocusVision.

Kim, Y. et al., 2014. *Activity Recognition for a Smartphone Based Travel Survey Based on Cross-User History Data.* s.l., s.n.

Lattery, K., Bartolone, G. P. & Saunders, T., 2013. *Optimizing Surveys for Smartphones: Maximizing Response Rates While Minimizing Bias,* s.l.: Maritz Research.

Liao, J. et al., 2015. Smart Diary: A Smartphone-Based Framework for Sensing, Inferring, and Logging Users' Daily Life. *IEEE Sensors Journal,* 15(5).

Liebe, U., Glenk, K., Oehlmann, M. & Meyerhoff, J., 2015. Does the use of mobile devices (tablets and smartphones) affect survey quality and choice behaviour in web surveys?. *The Journal of Choice Modelling,* Volume 14, pp. 17-31.

Li, M., Dai, J., Sahu, S. & Naphade, M., 2011. *Trip Analyzer through Smartphone Apps.* Chicago, IL, ACM SIGSPATIAL GIS '11.

Millar, M. M. & Dillman, D. A., 2011. Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly, 75*(2), pp. 249-269.

Millar, M. M. & Dillman, D. A., 2012. Encouraging Survey Response via Smartphones: Effects on Respondents' Use of Mobile Devices and Survey Response Rates. *Survey Practice, 5*(3).

Montini, L., Rieser-Schussler, N. & Axhausen, K. W., 2014. *Personalisation in multi-day GPS and accelerometer data processing.* Ascona, STRC.

Montini, L., Rieser-Schussler, N., Horni, A. & Axhausen, K. W., 2014. Trip Purpose Identification from GPS Tracks. *Transportation Research Record: Journal of Transportation Research Board,* Issue No. 2405, pp. 16-23.

Murphy, J. et al., 2014. *Social Media in Public Opinion Research: Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research,* s.l.: American Association for Public Opinion Research (AAPOR).

Nitsche, P. et al., 2014. Supporting Large-Scale Travel Surveys with Smartphones – A Practical Approach. *Transportation Research Part C,* Volume 43, pp. 212-221.

Nitsche, P., Widhalm, P., Breuss, S. & Maurer, a. P., 2012. A Strategy on How to Utilize Smartphones for Automatically Reconstructing Trips in Travel Surveys. *Procedia - Social and Behavioral Sciences,* Volume 48, pp. 1033-1046.

Pei, L. et al., 2013. Human Behavior Cognition Using Smartphone Sensors. *Sensors,* 13(2), pp. 1402-1424.

Phoenix Strategic Perspectives Inc., 2012. *Secondary Research into Cell Phones and Telephone Surveys,* s.l.: Public Works and Government Services Canada.

Safi, H. et al., 2015. Design and Implementation of a Smartphone-based System for Personal Travel Survey: Case Study from New Zealand. *Transportation Research Board 94th Annual Meeting* , Issue No. 15-1462.

Safi, H., Mesbah, M. & Ferreira, L., 2013. *ATLAS Project - Developing a mobile-based travel survey.* Brisbane, Australia, s.n.

Statistics Canada, 2013. *Residential Telephone Service Survey 2013,* s.l.: s.n.

Tanaka, G., Okada, M. & Mineno, H., 2015. GPS-Based Daily Context Recognition for Lifelog Generation Using Smartphone. *(IJACSA) International Journal of Advanced Computer Science and Applications,* 6(2).

Thomas, T., Chargeurs, K., Bijlsma, M. & Douhou, S., 2014. *The Dutch Mobile Mobility Panel.* [Online] Available at: http://www.nwo.nl/en/research-and-results/research-projects/10/2300170410.html [Accessed 20 August 2015].

Vlassenroot, S., Gillis, D., Bellens, R. & Gautama, S., 2015. The Use of Smartphone Applications in the Collection of Travel Behaviour Data. *International Journal of Intelligent Transportation Systems Research,* 13(1), pp. 17-27.

Wells, T., Bailey, J. T. & Link, M. W., 2014. Comparison of Smartphone and Online Computer Survey Administration. *Social Science Computer Review,* 32(2), pp. 238-255.

Zhao, F., Ghorpade, A. & Pereira, F. C., 2014. *Stop detection in smartphone-based travel surveys.* Australia, s.n.

Zhao, F., Pereira, F. C., Ball, R. & Kim, Y., 2015. Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore. *Transportation Research Board 94th Annual Meeting,* Issue No. 15-4343.