



An Overview of Technologies for Individual Trip History Collection: Mobility Decision Science Pillar SMART Mobility Consortium

Ted Kwasnik, Scott P. Carmichael, and Steven Isley

National Renewable Energy Laboratory

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

**Technical Report
NREL/TP-6A20-70331
January 2019**



An Overview of Technologies for Individual Trip History Collection: Mobility Decision Science Pillar SMART Mobility Consortium

Ted Kwasnik, Scott P. Carmichael, and Steven Isley

National Renewable Energy Laboratory

Suggested Citation

Kwasnik, Ted, Scott P. Carmichael, and Steven Isley. 2019. *An Overview of Technologies for Individual Trip History Collection: Mobility Decision Science Pillar SMART Mobility Consortium*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-70331. <https://www.nrel.gov/docs/fy19osti/70331.pdf>.

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-6A20-70331
January 2019

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by [applicable Department of Energy office and program office, e.g., U.S. Department of Energy Office of Energy Efficiency and Renewable Vehicle Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Executive Summary

The WholeTraveler Transportation Behavior Study is an initiative within the Mobility Decision Science pillar of the SMART (Systems and Modeling for Accelerated Research in Transportation) Mobility consortium. The study will survey residents in the San Francisco Bay Area and employ state-of-the-art multidisciplinary methodological approaches to identify and analyze complex changes in travel choice patterns, traveler preferences, and the traveler decision-making processes over various time scales, particularly with relation to emerging transportation technologies and services (such as electric vehicles, mobility on demand, shared mobility options, connected and automated vehicle technology, and e-commerce). Moreover, stated preferences or characteristics will be correlated with observed travel behaviors to measure the relative importance of personality characteristics and circumstantial constraints (e.g., the travel modes available within one's neighborhood) in travel choices. Innovative tools, capabilities, and scientific findings produced through this effort are anticipated to help improve future transportation and energy systems solutions by generating fundamental research and insights that can inform the design of pathways toward more efficient and secure transportation systems.

Fine-grained location data (FGLD), which for the purpose of this study is a record of an individual's location collected at a resolution sufficient to determine both the unique destinations a traveler visits each day and the modes used to travel between them, are a component of the WholeTraveler study. The methods for collecting FGLD and the means by which it is associated with a traveler's personal attributes have significant implications for the burden on research participants, the cost and complexity of the study, as well as the confidence with which hypotheses can be validated. This paper provides a broad overview of existing technologies for FGLD collection, categorizing them as either network or onboard approaches. An onboard system, such as smartphone or Global Positioning System (GPS) device sends signals to external networks to capture its own location. Conversely, network systems, such as cell towers, are comprised of sensing nodes with known locations that record when individually identifiable assets come within their range of detection. Tables ES-1 and ES-2 summarize the characteristics of prominent technologies within each category. Table ES-1 focuses on implications for researchers and categorizes technologies by their relative burden to participants and researchers, their potential to influence who will be willing and able to participate in the study, the risks associated with data collection, and the degree to which researchers will be able to communicate with and identify participants. Table ES-2 details key technical specifications of each technology including its temporal and spatial resolution, the physical environments for which it is suitable, and the modes it is equipped to classify.

Our analysis indicates that the most sensible approach to FGLD collection for the purposes of WholeTraveler is to leverage onboard devices, and in particular the existing Moves or Google Maps applications for Android and iOS smartphones. About two-thirds of U.S. adults currently own a smartphone, and analysis of key demographics suggests a sufficiently broad pool from which to draw a representative study population (Smith 2015). Use of existing smartphone applications offloads the burden of developing and maintaining sophisticated software that provides a simple user experience and yields standardized travel diary data. Both Google Maps and Moves implement background processing to classify mode, and they also allow users to generate export files. Instructions for how to generate and upload these files (or even automated

data transfer tools in the case of Moves, which supports an application programming interface [API] for software developers) can be integrated into a larger web-based survey. Collecting FGLD through Moves or Google Maps would likely need to engage users at least twice. The first interaction would instruct users how to install and activate location history collection through the Google Maps or Moves application. After a period of time in which sufficient data have been collected, a second interaction would provide means for participants to transmit extracts of their location history data to researchers.

Of these two applications, *Google Maps* is a strong choice because it already serves a substantial user base and comes with extensive user support resources. Accordingly, participants are expected to have at least some familiarity with *Google Maps* and to face low barriers to participation in the study. However, the robust activity tracking functionality of Moves holds significant value to the WholeTraveler initiative specifically. The Moves application also has the potential to facilitate a less burdensome participant experience through programmatic API data collection.

Independently developing a custom smartphone application is not recommended at this stage of the WholeTraveler research initiative. Such an approach is attractive in that it enables the coupling of data collection for both traveler location and personal characteristic data. It also affords the most customization, as data can be collected at intervals and in formats specified by the researcher. However, there are notable challenges in establishing minimal viable functionality, as well as in ensuring compatibility across an ever-growing combination of hardware devices and software versions. Moreover, mode inference is not inherent in raw data, so advanced machine learning, algorithmic or statistical approaches would need to be employed to derive mode classifications with confidence. Furthermore, behavioral barriers to user adoption pose unique challenges. Overall, the technical requirements and the rapid development cycles inherent in smartphone application development pose challenges and costs that are beyond the scope and primary research interests of the WholeTraveler study.

Dedicated GPS devices, which have been widely implemented by the transportation research community, represent an alternative and acceptable form of FGLD collection to smartphones. This method of FGLD collection, however, also poses unique challenges. Studies based on this approach typically furnish participants with car-mounted or handheld devices that passively collect FGLD. At the end of the study, participants return the devices via mail. Participants also conduct mail-in, web-based, or telephone surveys to provide supplemental information as necessary. While dedicated GPS devices have low barriers to use and collect data at comparable spatial and temporal resolutions to smartphones, these devices are less suitable for the purposes of WholeTraveler. To minimize study participant attrition rates, studies using dedicated GPS devices typically have relatively short FGLD collection periods, commonly no more than a week. Smartphones, on the other hand, are commonly deeply entrenched in their owners' everyday routines and are thus less likely to be forgotten than a separate FGLD device that does not directly serve the need of the participant. GPS devices also do not provide FGLD with travel modes already classified. Furthermore, smartphone costs also do not scale in the same way as do GPS devices. Once a mechanism for porting smartphone derived FGLD has been developed, it can be made available to a vast number of users at low marginal costs. For GPS-based surveys, however, each new participant adds the same amount of hardware and shipping costs.

Accordingly, larger sample sizes are expected to be more readily obtained through smartphone-based applications.

Alternative FGLD collection approaches outlined in Tables ES-1 and ES-2 (below) are not recommended primarily because of either the cost of developing a sensor network or the difficulty of linking location data to supplemental data sets. Bluetooth readers and radio-frequency identification (RFID) cards, for example, require the development of distributed sensor networks whose cost and complexity prohibit large area coverage. Also, these approaches lack convenience, as they may require participants to carry additional hardware (introducing the risk of location tracking assets being forgotten or abandoned). Likewise, network-based approaches, including cell towers and optical sensors are well equipped to record FGLD for large numbers of individuals, yet they lack the ability to interact with users in order to obtain key individual identifiers and attributes.

No matter the technical means of data collection, the secure handling of FGLD is of critical importance. For instance, location information can be analyzed to reveal an individual's residence, place of work, and typical daily routine. Ensuring location data remains secure from the time of collection to the time the data are warehoused is a significant responsibility.

FGLD that have been collected frequently and over long durations afford powerful insights for the design, operation, and planning of energy and transportation systems. One primary benefit of integrating these data within the WholeTraveler project will be to understand how this information can provide transportation managers with innovative services that save travelers time and money, reduce overall network congestion, and contribute to energy efficiency initiatives. The opportunities to improve traveler experiences with transportation and energy systems are only beginning to be imagined, yet personal location data clearly have an important role to play in these transformations.

Table ES-1. Methods and Qualitative Attribute Assessments

Method	User Burden	Potential Sampling Biases to Overcome	Privacy and Security Risks	Researcher Burden	Researcher-Participant Communication
Portable GPS Device	High	Self-selection	Moderate	<ul style="list-style-type: none"> - Costs per device - Mode Inference - User engagement/retention 	Acceptable Methods include mail-in surveys, web/phone applications, interviews
Vehicle GPS	Moderate	Self-selection Non-drivers (young, seniors, low income)	Moderate	<ul style="list-style-type: none"> - Device cost - User engagement 	Acceptable Methods include mail-in surveys, web/phone applications, interviews

Method	User Burden	Potential Sampling Biases to Overcome	Privacy and Security Risks	Researcher Burden	Researcher-Participant Communication
Existing Smartphone Application	Low	Self-selection 50+ No high school degree Low income (<\$30,000/yr) Rural	Moderate to High	- User adoption/retention - Device compatibility - Preservation of battery life	Acceptable Methods include mail-in surveys, web/phone applications, interviews
Custom Smartphone Application	Moderate	Self-selection 50+ No high school degree Low oncome (<\$30,000/yr) Rural	High <i>(Application is less mature and the researcher is more accountable for it.)</i>	- Software Development - Device compatibility - Preservation of battery life - Mode inference - User adoption/ experience/retention	High Continuous real-time feedback possible through application Other methods include mail-in surveys, web/phone applications, interviews
Cell Tower	Low	50+ No High School Degree Low Income Rural	Low	- User engagement - Data acquisition	Not feasible
BlueTooth/ RFID	Low	50+ No High School Degree Low Income (<\$30,000/yr) Rural	Moderate	- Cost of tracking devices and sensing network - Scaling of study area - User engagement/ retention	Marginally Acceptable Methods include mail-in surveys, web/phone applications, interviews

Method	User Burden	Potential Sampling Biases to Overcome	Privacy and Security Risks	Researcher Burden	Researcher-Participant Communication
License Plate Reader	Low	Non-Drivers (Young, Seniors, Low Income)	Low	- User engagement	Very Low Difficult to engage participants vehicle registrations

Table ES-2. Methods and Technical Attributes

Method	Type	Temporal Resolution	Spatial Resolution	Spatial Coverage	Mode Classification
Portable GPS Device	Onboard	Adjustable—about one second intervals	7–100 m	Moderate Most outdoor locations	Identifiable Modes: All Accuracy: Moderate
Vehicle GPS Device	Onboard	Adjustable—about one second intervals	~ 10 m–100 m	Moderate Vehicle-accessible outdoor locations	Identifiable Modes: Car Accuracy: Not Applicable
Smartphone Application	Onboard	Adjustable—about one second intervals	8–100s m Reception Dependent	High Most indoor and outdoor locations	Identifiable Modes: All Accuracy: Moderately High
Cell Tower	Network	Inconsistent Depends on phone use habits	100s m	High Range of cell network	Identifiable Modes: All Accuracy: Very Low
BlueTooth	Network	Inconsistent Whenever user is within range	10 cm	Very Low Range of beacon network	Identifiable Modes: Variable Accuracy: Low (Unless network is mode specific)
RFID	Network	Inconsistent Whenever user is within range	< 1 cm	Very Low Range of beacon network	Identifiable Modes: Variable Accuracy: Low (Unless network is mode specific)
License Plate Reader	Network	Inconsistent Whenever user is within range	10's m	Low Range of camera network	Identifiable Modes: Car Accuracy: High

Table of Contents

Introduction	1
Onboard Systems	5
GPS Devices	5
Smartphones	7
Network Systems	12
Cell Tower Networks	12
Bluetooth	12
Radio Frequency Identification (RFID)	13
License Plate Readers	14
Recommendation	15
Conclusion	17
References	18
Appendix	21

List of Figures

Figure 1. Types of FGLD system.....	4
-------------------------------------	---

List of Tables

Table ES-1. Methods and Qualitative Attribute Assessments	v
Table ES-2. Methods and Technical Attributes	viii
Table 1. Common Location-Based Smartphone Applications.....	Error! Bookmark not defined.

Introduction

The WholeTraveler Transportation Behavior Study is an initiative within the Mobility Decision Science pillar of the SMART (Systems and Modeling for Accelerated Research in Transportation) Mobility consortium. The study will survey residents in the San Francisco Bay Area and employ state-of-the-art multidisciplinary methodological approaches to identify and analyze complex changes in travel choice patterns, traveler preferences, and the traveler decision-making processes over various time scales, particularly with relation to emerging transportation technologies and services (such as electric vehicles, mobility on demand, shared mobility options, connected and automated vehicle technology, and e-commerce). Moreover, stated preferences or characteristics will be correlated with observed travel behaviors to measure the relative importance of personality characteristics and circumstantial constraints (e.g., the travel modes available within one's neighborhood) in travel choices. Innovative tools, capabilities, and scientific findings produced through this effort are anticipated to help improve future transportation and energy systems solutions by generating fundamental research and insights that can inform the design of pathways toward more efficient and secure transportation systems.

This paper is primarily concerned with assessing the modern technologies suitable for recording fine-grained location data (FGLD), which represent a travel history of sufficient spatial and temporal resolution to determine the unique destinations a traveler visits and the modes used to travel between them. Detailed route locations and trajectories may be beneficial to mode classification efforts and technologies will accordingly be assessed by their spatial and temporal resolution. Moreover, this paper will discuss the suitability of technologies to the priorities of the WholeTraveler research initiative, including low participant burden and the ability to effectively associate FGLD with supplemental traveler profiles. Since the information in this report is expected to be of interest to researchers similarly exploring the societal implications of transportation decisions, we also propose categorization of the modern FGLD technologies to frame distinct approaches to FGLD collection.

Furthermore, while this report focuses on automated FGLD solutions, the use of self-reported travel diaries as a primary source of data bears discussion. Researchers have employed such methods extensively for decades, and their known limitations include participant burden, cost, accuracy, and the inability to record precise fluctuations in speed and trajectory during trips (Jan, Horowitz, and Peng 2000; Raza et al. 2015; Vij and Shankari 2015). Houston, Luong, and Boarnet (2014) suggest self-reported diaries typically fail to record between 10% and 80% of trips, and a summary of eight transportation studies engaging 1,900 households or more (Bricka and Bhat 2006a) suggests that underreporting rates tend toward about 30% of trips on average. Short trips (often within a chain of trips) and those occurring later in the day are also most likely to be neglected in trip accounting (Bricka and Bhat 2006a; Wolf et al. 2004). Though self-reported travel diaries are not of primary interest for this study, they may be indispensable for collecting mobility related information from individuals for time periods prior to participation in a study. Historical information regarding duration of car ownership, transit usage, and frequently visited locations adds context and richness to personalized analyses, even if it is subject to human reporting error.

Priorities for WholeTraveler

The technologies that underlie FGLD collection yield distinct implications for the success of WholeTraveler in capturing high-quality records of origin and destination pairs and mode choices that can be associated with traveler attitudes and preferences. This paper assesses the suitability of each technology according to the following attributes:

Technical Considerations

- **Temporal Resolution:** Technologies are assessed according to the frequency at which spatial coordinates are collected. Ideal methods collect data frequently enough to reveal all origin-destination pairs throughout a day. Acceleration and velocity estimates may also be necessary for fine-grained travel mode inference.
- **Spatial Resolution:** An ideal source collects spatial coordinates at fine enough resolution to reveal all origin-destination pairs. This research effort does not require a detailed understanding of the exact route taken during trips, so long as travel mode can be acquired or imputed through alternative methods.
- **Spatial Coverage:** An ideal FGLD approach facilitates data collection in all urban and rural locations outside of buildings.
- **Mode Classification:** An ideal FGLD approach facilitates data collection across walking, biking, public transit, and automotive modes. The ability to distinguish between driving alone, driving with others, and using ride-sharing or car-sharing/rental services is also of interest.

Other Considerations

1. **Participant Burden:** An ideal FGLD collection process would be entirely passive, requiring no user intervention beyond opt-in consent. It would pose no costs to participants, financially or otherwise. In practice, however, participants will likely need to be prompted for information at various stages of the study. Additionally, the study may require participants to carry a device (e.g., a smartphone). Further, participation in the study could result in participants incurring costs related to smartphone data plan usage or battery drainage. A user experience designed to minimize unnecessary hassles or annoyances will be critical in retaining participants throughout the entire duration of the study. For instance, an FGLD solution that frequently prompts the participant to enter their transportation mode would be expected to have a higher attrition rate than a solution that can discern the user's transportation mode passively.
2. **Sampling Bias:** The decision to collect FGLD through a particular technology may inherently bias the study population. Vehicle-mounted Global Positioning System (GPS) devices, for example, will exclude those who do not drive or otherwise have access to a vehicle. This demographic segment includes the young, physically disabled, very old, and low-income populations. While sampling biases can never be eliminated, the selection of this study's FGLD collection method should include the most equitable representation of age, race, ethnicity, gender, income, and education levels possible. Recruitment outreach efforts may be necessary to include demographic groups that may otherwise be inadequately represented in the study's sample population. Unavoidable sampling biases

should be accounted for by employing appropriate bias reduction techniques (e.g., weighting the survey sample).

3. **Privacy and Security Risks:** Maintaining control and security over all location data throughout all phases of data collection and analysis is a critical responsibility for researchers. An optimal FGLD collection method will allow data to be encrypted at the time of collection and remain encrypted until analysis requires decryption.
4. **Researcher-Participant Engagement:** The study requires the ability to link FGLD with survey data collected from a given participant. The technology used to collect FGLD has notable implications for the ability of researchers to regularly engage with study participants, and individual identifiers of some form are necessary to facilitate this interaction. An ideal solution would provide opportunities for researchers to prompt users for feedback at any point throughout the study. Further, the solution should allow these interactions to occur in response to changes in the participant's location.
5. **Researcher Burden:** Each technology poses unique challenges to researchers in study design and implementation. An ideal solution would furnish researchers with the data needed to answer the essential research question at minimal costs in terms of time and money. For the purposes of this study, existing high-quality FGLD methodologies and technologies that enable researchers to focus on analysis rather than collection will be prioritized over the development of innovative approaches.
6. **Third Parties:** Each technology will be evaluated in terms of the ability for researcher and participant burden to be mitigated through the involvement of third parties. Involvement may take the form of partnerships with similarly aligned entities, including state, federal, and regional planning agencies that are conducting their own FGLD collection efforts. Researchers may also be able to use FGLD streams already being collected by users, such as the location history on smartphone devices, or locations captured by popular activity tracking apps.

Location Information Collection Systems

The field of possible FGLD collection technologies is broad enough to benefit from a taxonomy. We have chosen to distinguish between *onboard* and *network* locational systems (Figure 1). An *onboard* system sends signals to one or more external networks to capture its own location. Examples of such systems include GPS devices that triangulate location from communication with satellites, and smartphones that interact with GPS, wireless, and Bluetooth networks to position themselves. Conversely, *network* systems are comprised of sensing nodes with known locations that record when individually identifiable assets come within their range. Network systems include cell towers, Bluetooth beacons, and license plate readers. Network and onboard systems can be coupled such that the individual assets a network interacts with are aware of their own location (e.g., as is the case with cell towers and smartphones).

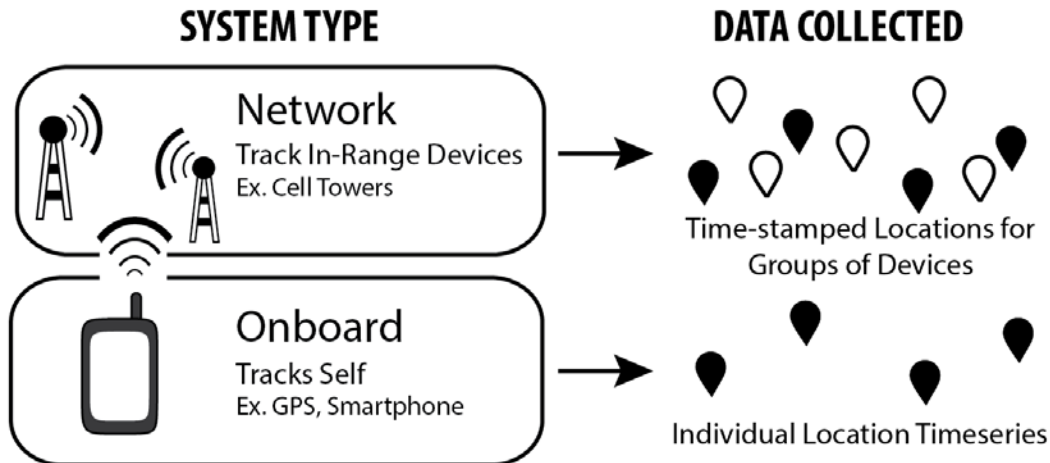


Figure 1. Types of FGLD system

As illustrated in Figure 1, the onboard and network system approaches yield distinct forms of raw locational data. Onboard systems can continuously track an individual's location, thus providing detailed trajectories. Onboard systems may also directly interact with participants to classify modes and trips, and to collect supplemental traveler attribute information. Network systems, conversely, record only the intermittent times at which devices interact with them, for example, when an individual comes within range of a Bluetooth beacon or license plate reader, or when he or she places a call over a cellular network. The passive nature of individual interactions with network systems poses minimal burdens to participants, many of whom are unaware these data are being generated. However, such systems do not typically afford opportunities to obtain consent from users or mechanisms to further engage with them for more information. Location data obtained from third parties are typically anonymized, and thus correlations between travel patterns and traveler characteristics would need to be made for populations aggregated by geographic boundaries, such as Zip codes or census statistical areas.

Onboard Systems

GPS Devices

The Global Positioning System (GPS) is a network of 24 satellites that orbit the Earth. It was developed by the U.S. Department of Defense and is now freely available for the public to use.¹ GPS receivers on the Earth's surface process signals from this network and triangulate their latitude and longitude after having established connection to four or more satellites. GPS receivers are typically either sold as lightweight handheld battery-powered devices or as devices that plug into a passenger vehicle's on-board diagnostic (OBD) port. GPS trackers have been used extensively to inform transportation studies since the 1990s (Shen and Stopher 2014). The 2010–2012 California Household Travel Survey is one of the largest U.S. travel surveys conducted using the technology. It engaged 5,715 households with either wearable or vehicle-mounted GPS devices (NuStat Research Solutions 2013).

Technical Considerations

The *spatial accuracy* of GPS positioning varies according to the quality of the receiving device, the number of satellites with which it can communicate, atmospheric conditions, and interference on the ground (e.g., from deep canyons or clusters of tall buildings). Conventional GPS device positioning is typically accurate to within 7–100 meters (Houston, Luong, and Boarnet 2014). Vehicle-mounted devices typically offer slightly less accurate location readings because of signal blocking from the metal in the car body and potentially from window tinting (GPS Technologies 2017). Similarly, the *spatial coverage* is limited to outdoor locations with reception to satellites (and vehicle accessibility for OBD devices). The *temporal resolution* of conventional receivers is typically sufficient to record multiple position readings per second, though higher sampling frequencies do come at the cost of reduced battery life.

Mode classification is not innate to GPS-derived location data, yet post-processing techniques to infer this information are advancing. Zhu et al. (2016) recently demonstrated machine learning algorithms to distinguish between walking, cycling, bus, subway, and driving modes with greater than 90% accuracy.

Research Implications

A primary *burden posed to participants* by dedicated GPS-devices is the need for participants to remember to carry and charge the device. Numerous studies have shown that owners of GPS-based travel diaries can fail to record trips, especially those occurring later in the day, likely because they forget to carry the devices (Houston, Luong, and Boarnet 2014; Montini et al. 2015; Wolf et al. 2004). Bricka and Bhat (2006b) report that of those who elect to participate in GPS studies, those most likely to underreport trips are households that own three or more vehicles, households with incomes under \$50,000, and those 25 years old and younger.

Beyond reading use instructions, charging device batteries, and shipping devices back to researchers at the end of the survey (possibly requiring a trip to a shipment location), GPS data collection requires little effort on a daily basis. Some *sampling bias* may be introduced among older demographics that face physical challenges, skepticism, or difficulties learning new

¹“What is GPS?” Garmin, <http://www8.garmin.com/aboutGPS/>.

technologies (Smith 2014). Researchers may consider making dedicated support staff available via email, phone, or in-person communication to mitigate such risks.

Regarding *privacy*, GPS-derived location data hold notable privacy implications. Supreme Court Justice Sotomayor has stated, “GPS monitoring generates a precise, comprehensive record of a person’s public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations” (New York Times 2012). Accordingly, such information should be handled with utmost care. While the risk of cyberattack on dedicated GPS devices is limited in that they are not inherently connected to the internet, once data are warehoused for post-processing and analysis, industry best practices should be implemented to restrict access to and anonymize data.

Regarding *researcher burden*, undertaking a large-scale FGLD collection effort primarily through GPS devices is a significant challenge. The 2012 California Travel Survey, for example, cost about \$10 million over two years (NuStat Research Solutions 2013). Costs scale proportionally with the size of the study population; as more participants are brought onboard, more devices will need to be acquired, mailed, and retrieved. The same GPS can be used consecutively by multiple participants, though this adds to the duration of the data collection phase of the research initiative.

Development of a separate platform to solicit auxiliary surveys and diaries also adds to project complexity. When using dedicated GPS devices to obtain FGLD, researchers commonly supplement GPS data collection with written mail-in surveys, computer-assisted telephone interviewing (CATI), or web-based surveys (NuStat Research Solutions 2013).

Travel surveys already being conducted by state, regional, and national transportation planning agencies provide opportunities to collaborate with *third parties* to derive new GPS-derived data sets. Such partnerships naturally reduce the control over the manner in which FGLD are collected, as well as the types and number of supplemental questions to which participants respond. However, long-term partnerships provide greater access to funding and opportunities to reduce researcher burdens that may help facilitate longitudinal studies over many years. The Transportation Secure Data Center² at the National Renewable Energy Laboratory (NREL) hosts such data from previous transportation studies, using anonymization and controlled access techniques that extract additional value from these publicly funded studies in a way that preserves the privacy of study participants.

Summary

Dedicated GPS devices provide an acceptable means of gathering FGLD for the WholeTraveler research effort. Data can be collected at more than sufficient temporal and spatial resolutions, and over large enough areas to capture travel patterns. While mode classification is not inherent in data collection or viable through device notifications, plentiful literature exists for those willing to invest time in developing post-processing schemes.

The greatest limitation inherent in dedicated GPS device data collection is participant burden, particularly in terms of the need to carry an additional device but also the need to supply

² Transportation Secure Data Center, NREL, <https://www.nrel.gov/transportation/secure-transportation-data/>.

additional information through an alternate data collection platform. Costs also scale proportionally to the number of participants engaged in the study, inhibiting widespread deployment.

Smartphones

Over two-thirds of adults in the United States owned smartphones as of 2015 (Smith 2015), and such widespread adoption has spurred a suite of location-aware applications related to navigation, ridesharing, fitness, activity, and destination reviews. These applications have become essential everyday tools, and they collect rich data sets pertaining to travel histories. Numerous researchers have also developed custom applications for smartphones to collect FGLD and make personal travel patterns more visible to travelers (Bie et al. 2012; Cottrill et al. 2013; Froehlich et al. 2009; Holleis et al. 2012; Jariyasunant et al. 2015; Meloni et al. 2014; Jylhä et al. 2013; Schrammel, Busch, and Tscheligi 2013; Shankari et al. 2014).

Technical Considerations

Smartphones collect location through a suite of instruments, including GPS, accelerometer, Bluetooth, and wireless sensors. While the *spatial accuracy* of GPS sensors are equivalent to those within dedicated GPS devices, in areas of poor GPS reception, smartphones can use Bluetooth and wireless signals to position themselves. The received signal strength indication (RSSI) methodology for wireless signals is accurate to within tens of meters, and matching wireless fingerprints of known locations further refines accuracy to within a few meters (Lawson 2012; Lee et al. 2016; Swangmuang and Krishnamurthy 2008). Moreover, smartphones are capable of multiple readings per second. The *spatial coverage* of GPS-equipped smartphones includes nearly all indoor and outdoor locations.

Detailed GPS traces in coordination with precise tilt and positioning attributes provided by accelerometers provide information that facilitates *mode inference*. Reddy et al. (2010) report that the use of GPS and accelerometers together can distinguish between stillness, walking, running, biking, and motorized transport with an accuracy of over 93%, and that eliminating either sensor would decrease accuracy between 10% and 20%. Users can also be prompted through notifications to ground truth predicted mode classifications, though this increase in classification accuracy comes at cost of increased user burden and may be dangerous for users (e.g., by prompting drivers for mode classification while they are at a stop light) (Vij and Shankari 2015; Reddy et al. 2010).

Research Implications

The *burden for participants* to provide smartphone derived FGLD includes smartphone ownership, battery drainage, and data usage fees (Shen and Stopher 2014). The installation process of a new application, no matter how simple, is likely to dissuade users from completing the study and introduces a sampling bias toward more technologically savvy users. While users are less likely to forget a smartphone than a dedicated GPS device, users will also need to remember to make sure a smartphone application is running and configured appropriately throughout the study duration to avoid underreporting trips.

Reliance on smartphone ownership is expected to introduce some *sampling bias*. The Pew Research Group reports that there is no difference among racial or ethnic groups in smartphone

ownership but that ownership rates are as low as 30% for those 65 years and older and those without a high school degree (Smith 2015). Those earning less than \$30,000 a year and those living in rural areas are also more likely not to own a smartphone than more affluent and urban participants, though adoption rates in these groups still exceed 50% (Smith 2015). Sampling bias can be diminished through dedicated efforts to solicit participation among underrepresented groups, or by following the model of the RAND American Life panel which provides devices to those who could not otherwise participate in the study (RAND Corporation 2017).

Collecting FGLD from smartphone applications poses greater *privacy* and *security* concerns than from dedicated GPS devices because poor programming, unanticipated cyberattacks, inadequate encryption schemes, and the possibility of exposed passwords introduce additional vulnerabilities to the data collection pipeline. Moreover, beyond exposing location data, another significant concern with smartphones is the disclosure of user credentials that may be necessary to access applications that collect FGLD. Some security comes from hosting apps on the Android and iOS app stores which enforce basic integrity requirements. Moreover, third party user account services like *Oauth* leverage existing user credentials on popular site like Google or Facebook, and in doing so incorporate latest industry best practices.

A clear advantage of smartphones is that they facilitate rich *researcher engagement with participants* and the ability to couple data collection efforts for both location information and traveler characteristics. For example, notifications prompting users to answer questions about trip purpose and travel attitudes can be either scheduled for certain times of day or triggered by changes in one's location. Commonly, after a user has remained at a location for a notable amount of time, smartphone application developers will tag the local area with a "geofence." Entering or exiting a geofence serves as the basis for identifying distinct trips, initiating location collection algorithms, and providing information to users.

The *burden for researchers* depends largely on whether a third-party application is used to collect FGLD or the researcher decides to build a custom application. Custom development requires a handle on effective user experience strategies, deep technical knowledge, and an understanding of compatibility across hardware, software, and data protocols. There exist software development kits such as HyperTrack³ that reduce the burden of undertaking the development of a location-tracking smartphone application. This particular software (HyperTrack) allows for raw data error correction, real-time tracking, and trip purpose classification (but not inherent mode classification as of the time of publication).

Fortunately, many *third-party* applications already adopted by users passively collect trip information and provide access to detailed trip histories. Table 1 (below) organizes these apps by vendor and category, and also summarize key app limitations for the purposes of WholeTraveler. Researchers may consider partnering with these providers, or more simply, inviting new users to use the applications and then periodically send extracts of their trip logs.

The *Google Maps* application, which is available for both iOS and Android devices, has a Timeline feature through which a full export of trip histories (inclusive of route and mode) may be downloaded as a raw data file. While the process can be completed in little more time than it

³ Developer homepage, HyperTrack, <https://developer.hypertrack.com/>.

takes to download the file, users must have an account with Google, be signed into this account on the device, and activate the location history for the *Google Maps* application before the FGLD application will record travel histories. Participants must also have access to both a smartphone and an internet-connected desktop to retrieve FGLD via *Google Maps*. Access to a desktop is not expected to be a significant barrier to participation, as a recent study (Smith 2015) found that only 7% of smartphone users have a smartphone as their only form of internet connection. It is still noteworthy, however, that the rates of smartphone dependence for internet are nearly twice as high for blacks, Hispanics, those between 18 and 29 years old, and those with incomes less than \$30,000 annually (Smith 2015). For these participants, special accommodations (such as researcher-provided devices) may be necessary to overcome the burden of accessing publicly available computers. Overall, use of *Google Maps* will introduce some selection toward more affluent participants who already own and operate multiple internet-connected devices.

Genre	App	Android	iOS	User Access to FGLD	Limitations
Maps and Navigation	Google Maps	Yes	Yes	Export file	Google account sign in; Enabling location history
Maps and Navigation	Apple Maps	No	Yes	View only	Point locations only, no trip details
Maps and Navigation	Waze	Yes	Yes	Export file	Car trips when location services enabled only
Transportation	Uber	Yes	Yes	View/ API	Rideshare trips only
Transportation	Lyft	Yes	Yes	View / API	Rideshare trips only
Activity Tracker	Moves	Yes	Yes	Export file/API ^a	Account sign in; Enabling location history
City Guide	FourSquare	Yes	Yes	Export file/API	Point locations only, no trip details
Fitness	RunKeeper	Yes	Yes	Export file	Pedestrian trips only, manual trip start/stop
Fitness	Strava	Yes	Yes	Export file	Pedestrian/Bike Trips Only Manual Trip Start/Stop
Fitness	MapMyRide	Yes	Yes	Export file	Pedestrian/Bike trips only, Manual trip start/stop

Table 1. Common Location-Based Smartphone Applications

^a application programming interface

Apple Maps on iOS devices with location history enabled similarly record location data, but they are not available for export to an external file. Information about the route and mode of trips between frequently visited locations is also not retained.

Moves, an activity tracking application, is compatible with iOS and Android smartphones and makes location histories available to users for download. The application passively captures all the destinations one visits throughout the day and classifies the modes for the trips that connect them. It also provides relevant fitness metrics, including steps taken and calories burned. For researchers, Moves is advantageous because it provides an API through which participants can consent to grant researchers access to their location history with just a few button clicks. The most notable burden for participants for this application will likely be new user installation and registration.

Foursquare is social application through which users record visits to and review popular locations, and sometimes receive special discounts. As with Moves, location histories are available for download to users as an external file or to developers through an API. The application does not capture trips between destinations, which poses challenges to mode inference. Also, the app requires manually “checking in” to destinations, further posing barriers to usability over long study durations.

Fitness trackers, including RunKeeper, Strava, and MapMyride likewise can capture FGLD and fitness metrics and can make personal data readily available to users. While these are intended for activities such as running and biking, they could be used by participants to track commutes as well. However, this approach would be burdensome for users who would need to manually start, stop, and tag each trip. New user installation and registration also pose user challenges.

Rideshare applications, including *Uber* and *Lyft* make trip histories available to users through web dashboards. Moreover, these vendors operate API services that make this data accessible programmatically. As of the time of publication, the data returned from these vendors included each trip’s origin, destination and duration; additionally, Uber provides distance, and Lyft provides fare information. Once user credentials are secured, these major rideshare providers afford low burden access to rich rideshare datasets.

Beyond the third-party applications already discussed, commercial location data aggregators, such as Uber Retail Media⁴ and PlaceIQ⁵ sell location information collected from smartphones. However, this information is anonymized, which eliminates the possibility of directly engaging with users for supplemental information.

⁴ Vista homepage, Vista by UberMedia, <https://uber-retail.com/>.

⁵ PlaceIQ homepage, PlaceIQ, <http://placeiq.com/>.

Summary

Moves and Google Maps provide the most promising means of FGLD collection through smartphones. They provide researchers access to robust destination and trip route and mode information in a way that requires minimal user intervention. *Google Maps* has a more substantial user base than *Moves*, though participants could be required to download and use either application as a precondition to involvement in the study. The advantage of *Google Maps* is the extensive existing user base, while the advantage of *Moves* is the API through which the burden of users porting data to researchers can be greatly reduced.

Custom smartphone application development is not recommended, and an application that satisfies just the basic FGLD requirements would largely duplicate existing functionality in third-party applications. This approach may be considered if additional functionality is deemed necessary, in which case the use of existing frameworks are recommended, as is keeping the scope bounded to that of minimal functionality.

Network Systems

Cell Tower Networks

A common network-based approach to interpolating individual travel patterns is to use maintenance records created as cell towers record completed calls made by mobile phones (not strictly smartphones). These call detail records (CDRs), are collected by telecommunication companies, contain timestamps as well as triangulated coordinates.

Technical Considerations

Call detail records are spatially accurate to within a few hundred meters (Toole et al. 2015), and their *temporal resolution* is variable, which is reflective of the intermittent nature of user cellular activity. Accordingly, a CDR may indicate the coordinate of a daily destination or it may reflect a traveler's location while traveling. Post-processing algorithms use timestamps to help refine data points and make assumptions about a user's true travel history throughout a day. Probabilistic models, reflecting the likelihood of a user taking a certain mode between two destinations, have also been the most promising means of *mode classification* (Toole et al 2015).

The *spatial coverage* of cell towers permits FGLD collection throughout most urban areas, though breaks in cellular activity will result in incomplete trip logs.

Research Implications

CDRs pose low *participant burden* and *privacy* concerns relative to GPS and smartphone derived coordinates. Call Detail Records are foremost less accurate, but also importantly are commonly provided in aggregated and anonymized formats by *third-party* data aggregators, such as AirSage⁶ (Calabrese et al. 2013). The potential *sampling biases* are low as well. The least represented groups, including those with less than a high school education, those over 65 years of age, those earning less than \$30,00 per year, and those living in rural areas all have mobile phone adoption rates in excess of 80% (and the majority of these phones are smartphones) (Pew Research Center 2017).

The *burden for researchers* is high in that sophisticated statistical models typically are required to infer travel diaries from sparse low-quality data. Additionally, the anonymous nature of CDRs means that *engagement* with individual participants is not feasible. Instead, location data trends need to be aggregated by Zip code or U.S. census block group to make correlations with aggregate demographic attributes.

Summary

The poor resolution of CDRs and the inability to associate them with individuals render this approach unsuitable to the goals of WholeTraveler.

Bluetooth

Bluetooth-enabled devices, such as smartphones, wireless controllers, car stereos, and laptops, periodically broadcast a unique identifier code, known as a media access control (MAC) address. Leveraging this identifier, networks of Bluetooth sensors can triangulate the position of a device

⁶ AirSage homepage, AirSage, <https://airsage.com/>.

and record changes in its location over time. Use of these technologies for marketing and advertising is growing, and there have been some promising demonstrations within transportation research, including Barceló et al. (2010), who used the technology to record traffic flows along a highway in Spain.

Technical Considerations

Meshes of Bluetooth sensors can triangulate a Bluetooth-enabled device to a *spatial resolution* within 10 cm (Lawson 2012) and at a *temporal resolution* on the order of ten readings per second (Barceló et al. 2010). The *spatial coverage* of this approach is limited by the number and distribution of mesh nodes, each having a range of about 70 meters.⁷

Strategies for dense urban Bluetooth network design have been proposed (Isukapati, Barlow, and Smith 2014), though real world applications are typically contained to indoor locations, such as shopping malls. Reddy et al. (2010) suggest that this technology is not suitable for *mode classification* because its coverage does not extend outdoors; some modes can be inferred from their proximity to network nodes, but confidence and validation opportunities related to such data are low.

Research Implications

As passive data collection systems, the *participant burden* through network-based Bluetooth signals is low. The resolution and access mechanisms also present *privacy* concerns similar to those of smartphone-derived FGLD. *Researcher engagement with participants* would likely require users to either register the MAC address of their smartphone or be provided a dedicated Bluetooth device for location tracking. With either approach, a separate web application, smartphone app or survey device would need to be developed and linked to the Bluetooth identifier for the purposes of WholeTraveler.

While developing the infrastructure for a Bluetooth data collection network is conceivable at small scales, the cost and timeframe associated with implementing a system at the size necessary for the WholeTraveler project poses insurmountable *researcher burdens*. Costs scale both with network coverage and user participation (assuming dedicated Bluetooth devices with known MAC addresses need to be provided to users to facilitate FGLD collection). Moreover, *third parties* that have experience in Bluetooth network development typically collect data for commercial enterprises that have not historically made such data publicly available.

Summary

The use of Bluetooth technology is not suitable to the WholeTraveler research initiative.

Radio Frequency Identification (RFID)

Radio-frequency identification (RFID) is the technology behind many access cards, including ones swiped at kiosks to gain access within college or corporate campuses, or at ski resorts. Public transportation and highway toll collection agencies also commonly implement this technology to collect fares from regular commuters.

⁷ “Beacons: Everything You Need to Know,” Pointr Labs, <http://www.pointrlabs.com/blog/beacons-everything-you-need-to-know/>.

Technical Considerations

RFID affords very precise *spatial* (<1 cm) and *temporal resolutions*. However, *coverage* and *mode classification* capabilities, much like Bluetooth networks, are limited by the number and distribution of network nodes. Accordingly, this technology would be very good at capturing entry to and egress from public transit infrastructure or toll roads. Oberli, Torres-Torriti, and Landau (2010) found “Commercial off-the-shelf RFID technologies can effectively be used to recognize individual passengers as they board and alight buses in public transportation systems.” Outside these systems, however, the technology gives little insight into traveler movements.

Research Implications

As with Bluetooth-based networks, the cost and timeframe associated with developing the infrastructure for a robust RFID data collection system exceeds the scope of the WholeTraveler project. The technology would be better suited to transportation studies of public transit or highway usage.

Summary

The use of RFID technology is not suitable to the WholeTraveler project.

License Plate Readers

Distributed optical surveillance systems in urban environments coupled with machine learning technologies provide additional mechanisms for obtaining FGLD. Since the 1990s, researchers have demonstrated the use of automated license plate recognition (ALPR) scanning systems to conduct origin-destination pattern analyses (Shuldiner, D’Agostino, and Woodson 1996). Many state transportation agencies rely on such technologies for tolling, and use among police agencies is common.

Technical Considerations

The *resolution* and *coverage* of FGLD derived from ALPR systems is variable and dependent on the number and distribution of nodes (nodes can notably be fixed cameras or mounted to mobile assets such as police cars) and user’s intermittent interaction with them and *mode classification* is inherently limited to passenger vehicles. Distributed surveillance systems aided through artificial intelligence could theoretically interpret images to identify travelers and classify any mode; however, the development of such sophisticated systems pose very high privacy and security concerns, and the systems are thus well outside the scope of the WholeTraveler project.

Research Implications

As passive and unobtrusive data collection systems, optical sensor signals present a low *participant burden*. The *privacy* risks are highly variable, depending on the distribution, location, and quality of sending nodes. The low FGLD quality of ALPR systems, which stems from the intermittent nature of user interaction with nodes, might seem to suggest the risks are lower than those of high-resolution GPS or smartphone devices. However, additional information captured in images pose notable privacy and security concerns that post-processing tasks (obscuring all but license plate information) may help mitigate.

Traveler mailing addresses could conceivably be inferred from license plate readings of ALPR systems and be used to deliver mailed invitations for engagement in the WholeTraveler study.

In practice, however, the channels through which these data would be obtained are heavily restricted. Moreover, engagement through mail communications is not expected to result in high participation rates or to easily enable study sizes to scale as rapidly as would web-based approaches.

Summary

The use of ALPR technologies is not suitable to the WholeTraveler project.

Recommendation

Smartphone Application

Leveraging an existing smartphone application is the most promising means of FGLD collection. Smartphones themselves afford the greatest coverage and high spatial and temporal resolutions. More importantly, existing applications meet the needs of researchers and participants are already deriving personal benefit from them. *Google Maps* and *Moves* notably provide navigation, activity tracking, and health metrics services.

For researchers, the use of existing applications offloads the deeply technical components of FGLD collection (including user experience, battery life, and cross-platform compatibility) that would otherwise pose considerable costs. Also, costs do not dramatically scale with growing population sizes. Once instructions (or possibly automated tools) have been established to guide participants through installation, use, and data retrieval, the cost of adding more participants does not increase dramatically. Other methods require more physical devices or network infrastructure to be developed to bring more participants into the study group and lead to escalations in overall project cost and complexity.

Development of a custom smartphone application to collect FGLD should be undertaken with caution and an awareness of the human factors inherent in software development. Recent evidence suggests that (1) smartphone user's interest in new applications has been saturated and (2) user fatigue stemming from anxiety over privacy, information overload, and frustrating user experiences dissuades users from installing new applications.⁸ Moreover, costs and efforts spent customizing an application may also distract research efforts away from more pertinent objectives. Existing applications are successful in part because they have undergone many stages of iterative development, and the WholeTraveler project timeline does not lend itself to this development approach.

Google Maps and *Moves* already capture standardized data with modes classified with more than sufficient spatial and temporal resolution to answer the research questions of the WholeTraveler initiative. These entities also have their own incentives to fix bugs and continually improve offerings.

The choice of an application does have direct implications for the success of WholeTraveler, yet little published literature exists from which to compare *Google Maps*, *Moves*, and other applications that collect FGLD. Accordingly, a preliminary study to assess willingness to install the application, the user experience, and the ease with which users can export copies of their data

⁸ "How App Fatigue is Taking a Toll on Smartphone Owners," Barb Darrow, Fortune, <http://fortune.com/2016/08/16/app-fatigue-is-taking-a-toll-on-smartphone-owners/>.

across representative demographics is recommended to make an informed decision about which application will yield the greatest chance of success.

Dedicated GPS Devices

There is significant precedent for using dedicated GPS devices in transportation studies, and it should be noted that the devices do pose a viable alternative to smartphone applications. GPS devices provide high-quality passively collected data. They are lightweight and portable. And, they are unlikely to induce sampling biases if distributed effectively.

GPS devices are not as well suited to the objectives of WholeTraveler as smartphone applications primarily because they are burdensome to participants who must remember to charge and carry GPS devices with them. Participants must also spend time mailing these devices back to researchers at the conclusion of the study. For researchers, as noted previously, study costs escalate proportionally to the study size making large-scale data collection unfeasible, given budget and time constraints. Furthermore, the raw data GPS traces provided require data processing techniques to reduce errors, classify mode, distinguish trips, and make associations with known locations.

Other Alternatives

Network FGLD systems are collectively unsuitable for the WholeTraveler project because of the cost and complexity of distributing sensors over a wide enough area to capture the daily travel patterns of large populations. Moreover, legal and related data access constraints prohibit direct communication with travelers to understand key characteristics, attitudes, and behaviors.

Conclusion

Detailed travel histories, correlated with socioeconomic, life trajectory and other contextual information, provide a rich dataset for energy and transportation analysis. Collecting quality data of this kind is the first step in establishing current baselines for mode shares and transportation related energy impacts, assessing the energy and mobility impacts of emerging transportation mega-trends (electric vehicles, ridesharing services, connected and autonomous vehicles), and revealing the traveler preferences that will drive future travel decisions. Innovative analytic tools that can compare and forecast mobility trends are anticipated to improve traveler experiences, better elucidate the impact of emerging and transformative transportation technologies and services, and inform the design of pathways towards more efficient and secure transportation systems.

This analysis of the technical aspects and research implications of existing locational technologies suggests that smartphone applications provide the most promising means of fine-grained location data collection. Leverage existing applications, moreover, affords a more practical solution rather than independently building a data collection application. Deploying dedicated GPS devices provides a less scalable, but compelling alternative as a data collection approach. The cost and complexity of building network based FGLD collection systems, and their limited opportunities for direct communication and identification of participants render them unsuitable for the purposes of WholeTraveler.

As location data and other big data resources becoming increasingly embedded in the design, operation and maintenance of complex transportation and energy systems, future research into collection methods that increase the resolution and quality of the data and allow for more sophisticated traveler interaction will help to promote more robust scientific understandings.

References

- Barceló J, L. Montero, L. Marqués, and C. Carmona. 2010. “Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring.” *Transportation Research Record: Journal of the Transportation Research Board* 2175:19–27.
- Bie, J., M. Bijlsma, G. Broll, H. Cao, A. Hjalmarsson, F. Hodgson, and Luther. 2012. “Move Better with tripzoom.” *International Journal on Advances in Life Sciences* 4(3&4): 125–135.
- Bricka, S., and C. Bhat. 2006a. “Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data.” *Transportation Research Record: Journal of the Transportation Research Board* (1972): 9–20.
- Bricka, S., and C.R. Bhat. 2006b. “Using GPS Data to Inform Travel Survey Methods.” In *Innovations in Travel Demand Modeling Conference*. Austin, TX: Transportation Research Board.
- Calabrese F, M. Diao, G. Di Lorenzo, J. Ferreira Jr., and C. Ratti. 2013. “Understanding Individual Mobility Patterns from Urban Sensing Data: A Mobile Phone Trace Example.” *Transportation Research Part C* 26: 301–313.
- Cottrill C., F. Câmara Pereira, F. Zhao, H. B. Lim, M. Ben-Akiva, and P. C. Zegras. 2013.) “Future Mobility Survey: Experience in Developing a Smartphone-Based Travel Survey in Singapore.” *Transportation Research Record: Journal of the Transportation Research Board* 2354: 59–67.
- Froehlich, J., T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison, and J. A. Landay. 2009. “UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1043–1052. ACM.
- Holleis P., M. Luther, G. Broll, H. Cao, J. Koolwaaij, A. Peddemors, P. Ebben, M. Wibbels, K. Jacobs, and S. Raaphorst. 2012. “TRIPZOOM: A System to Motivate Sustainable Urban Mobility.” *International Conference on Smart Systems, Devices and Technologies (SMART '12)*.
- Houston D., T.T. Luong, and M.G. Boarnet. 2014 “Tracking Daily Travel: Assessing Discrepancies between GPS-Derived and Self-Reported Travel Patterns.” *Transportation Research Part C: Emerging Technologies* 48: 97–108.
- Isukapati I.K., G.J. Barlow, and S.F. Smith. 2015. *Cost-Effective Network Topology for Ubiquitous Bluetooth Reader Deployment in Urban Networks*.
- Jan O, A.J. Horowitz, and Z-R Peng. 2000. “Using Global Positioning System Data to Understand Variations in Path Choice.” *Transportation Research Record* 1725 (Paper No. 00-0322): 37–44.

- Jariyasunant, J., M. Abou-Zeid, A. Carrel, V. Ekambaram, D. Gaker, R. Sengupta, and J. L. Walker. 2015. “Quantified Traveler: Travel Feedback Meets the Cloud to Change Behavior.” *Journal of Intelligent Transportation Systems* 19(2): 109–124.
- Jylhä, A., P. Nurmi, M. Sirén, S. Hemminki, and G. Jacucci. 2013. “Matkahupi: A Persuasive Mobile Application for Sustainable Mobility.” *2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, pp. 227–230.
- Lawson, S. 2012. “Ten Ways Your Smartphone Knows Where You Are.” *PCWorld*. http://www.pcworld.com/article/253354/ten_ways_your_smartphone_knows_where_you_are.html.
- Lee, J. G., B. K. Kim, S. B. Jang, S. H. Yeon, and Y. W. Ko. 2016. “Accuracy Enhancement of RSSI-Based Distance Estimation by Applying Gaussian Filter.” *Indian Journal of Science and Technology* 9(20).
- Meloni, I., B. Sanjust G. Delogu, and E. Sottile. 2014. “Development of a Technological Platform for Implementing VTBC Programs.” *Transportation Research Procedia* 3.
- Montini L., S. Prost, J. Schrammel, N. Rieser-Schüssler, and K.W. Axhausen. 2015. “Comparison of Travel Diaries Generated from Smartphone Data and Dedicated GPS Devices.” *Transportation Research Procedia*, pp. 227 – 241.
- New York Times. 2012. “When GPS Tracking Violates Privacy Rights.” <http://www.nytimes.com/2012/09/23/opinion/sunday/when-gps-tracking-violates-privacy-rights.html>.
- NuStat Research Solutions. 2013. *2010-2012 California Household Travel Survey, Final Report* California Department of Transportation.
- Oberli, C., M. Torres-Torriti, and D. Landau. 2010. “Performance Evaluation of UHF RFID Technologies for Real-Time Passenger Recognition in Intelligent Public Transportation Systems.” *IEEE Transactions on Intelligent Transportation Systems* 11(3): 748–753.
- RAND Corporation. 2017. *American Life Panel*. <https://alpdata.rand.org/>
- Pew Research Center. 2017. *Mobile Fact Sheet*. <http://www.pewinternet.org/fact-sheet/mobile/>
- Raza, A., L. Knapen, K. Declercq, T. Bellemans, D. Janssens, and G. Wets. 2015. “Diary Survey Quality Assessment Using GPS traces.” *Procedia Computer Science* 52: 600–605.
- Reddy S., M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. 2010. “Using Mobile Phones to Determine Transportation Modes.” *ACM Transactions on Sensor Networks* 6(2).
- Schrammel, J., M. Busch, and M. Tscheligi. 2013. “Peacox-Persuasive Advisor for CO₂-Reducing Cross-Modal Trip Planning.” *PERSUASIVE (Adjunct Proceedings)*.
- Shankari, K., Y. Mogeng, S. Shanmugam, D. Culler, and R. Katz. 2014. *E-Mission: Automated Transportation Emission Calculation Using Smart Phones*. University of California at Berkeley.

Shen, L., and P.R. Stopher. 2014. “Review of GPS Travel Survey and GPS Data-Processing Methods.” *Transport Reviews*.

Shuldiner, P., S. D’Agostino, and J. Woodson. 1996. Determining detailed origin-destination and travel time patterns using video and machine vision license plate matching. *Transportation Research Record: Journal of the Transportation Research Board* 1551: 8–17.

Smith, A. 2014. *Older Adults and Technology Use*. The Pew Research Group.
<http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/>

Smith, A. 2015. *US Smartphone Use in 2015*. The Pew Research Group.
<http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>.

Swangmuang, N., and P. Krishnamurthy. 2008. “An Effective Location Fingerprint Model for Wireless Indoor Localization.” *Pervasive and Mobile Computing* 4(6): 836–850.

Toole, J. L., S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C González. 2015. “The Path Most Traveled: Travel Demand Estimation Using Big Data Resources.” *Transportation Research Part C: Emerging Technologies* 58: 162–177.

Vij A., and K. Shankari. 2015. “When is Big Data Big Enough? Implications of Using GPS-Based Surveys for Travel Demand Analysis.” *Transportation Research Part C* 56:446–462.

Wolf, J., S. Bricka, T. Ashby, and C. Gorugantua. 2004. “Advances in the Application of GPS to Household Travel Surveys.” In *National Household Travel Survey Conference*, Washington D.C.

Zhu, X., J. Li, Z. Liu, S. Wang, and F. Yang. 2016. “Learning Transportation Annotated Mobility Profiles from GPS Data for Context-Aware Mobile Services.” In *2016 IEEE International Conference on Services Computing*.

Appendix. Summary Tables

The following tables summarize the key attributes of FGLD technologies described in this report.

Table A-1. Methods and Qualitative Attribute Assessments

Method	User Burden	Potential Sampling Biases to Overcome	Privacy and Security Risks	Researcher Burden	Researcher-Participant Engagement
Portable GPS Device	High	Self-selection	Moderate	<ul style="list-style-type: none"> - Costs per device - Mode Inference - User engagement/retention 	Acceptable Methods include mail-in surveys, web/phone applications, interviews
Vehicle GPS	Moderate	Self-selection Non-drivers (young, seniors, low income)	Moderate	<ul style="list-style-type: none"> - Device cost - User engagement 	Acceptable Methods include mail-in surveys, web/phone applications, interviews
Existing Smartphone Application	Low	Self-selection 50+ No high school degree Low income (<\$30,000/yr) Rural	Moderate to High	<ul style="list-style-type: none"> - User adoption/retention - Device compatibility - Preservation of battery life 	Acceptable Methods include mail-in surveys, web/phone applications, interviews
Custom Smartphone Application	Moderate	Self-selection 50+ No high school degree Low oncome (<\$30,000/yr) Rural	High <i>(Application is less mature and the researcher is more accountable for it.)</i>	<ul style="list-style-type: none"> - Device compatibility - Preservation of battery life - Mode inference - User adoption/experience/retention 	High Continuous real-time feedback possible through application Other methods include mail-in surveys, web/phone applications, interviews
Cell Tower	Low	50+	Low	<ul style="list-style-type: none"> - User engagement 	Not feasible

Method	User Burden	Potential Sampling Biases to Overcome	Privacy and Security Risks	Researcher Burden	Researcher-Participant Engagement
		<p>Less than High School Degree</p> <p>Low Income</p> <p>Rural</p>			
BlueTooth/RFID	Low	<p>50+</p> <p>Less than High School Degree</p> <p>Low Income (<\$30,000/yr)</p> <p>Rural</p>	Moderate	<ul style="list-style-type: none"> - Cost of tracking devices and sensing network - Scaling of study area - User engagement/retention 	<p>Marginally Acceptable</p> <p>Methods include mail-in surveys, web/phone applications, interviews</p>
License Plate Reader	Low	<p>Non-Drivers (Young, Seniors, Low Income)</p>	Low	<ul style="list-style-type: none"> - User engagement 	<p>Very Low</p> <p>Difficult to engage participants vehicle registrations</p>

Table A-2. Methods and Technical Attributes

Method	Type	Temporal Resolution	Spatial Resolution	Spatial Coverage	Mode Classification
Portable GPS Device	Onboard	Adjustable—about one second intervals	7–100 m	Moderate Most outdoor locations	Identifiable Modes: All Accuracy: Moderate
Vehicle GPS Device	Onboard	Adjustable—about one second intervals	~ 10 m–100 m	Moderate Vehicle-accessible outdoor locations	Identifiable Modes: Car Accuracy: Not Applicable
Smartphone Application	Onboard	Adjustable—about one second intervals	8–100s m Reception Dependent	High Most indoor and outdoor locations	Identifiable Modes: All Accuracy: Moderately High
Cell Tower	Network	Inconsistent Depends on phone use habits	100s m	High Range of cell network	Identifiable Modes: All Accuracy: Very Low
BlueTooth	Network	Inconsistent Whenever user is within range	10 cm	Very Low Range of beacon network	Identifiable Modes: Variable Accuracy: Low (Unless network is mode specific)
RFID	Network	Inconsistent Whenever user is within range	< 1 cm	Very Low Range of beacon network	Identifiable Modes: Variable Accuracy: Low (Unless network is mode specific)
License Plate Reader	Network	Inconsistent Whenever user is within range	10's m	Low Range of camera network	Identifiable Modes: Car Accuracy: High