

161407-2



# Strategic Research **PROGRAM**

Emerging Data Collection Techniques for Travel Demand  
Modeling: A Literature Review

October 2014



1. Report No. <b>TTI/SRP/14/161407-2</b>	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle <b>EMERGING DATA COLLECTION TECHNIQUES FOR TRAVEL DEMAND MODELING: A LITERATURE REVIEW</b>		5. Report Date <b>October 2014</b>	6. Performing Organization Code
7. Author(s) <b>Richard J. Lee, Ipek N. Sener, and James A. Mullins III</b>		8. Performing Organization Report No.	
9. Performing Organization Name and Address <b>Texas A&amp;M Transportation Institute College Station, Texas 77843-3135</b>		10. Work Unit No. (TRAIS)	11. Contract or Grant No. <b>10727</b>
12. Sponsoring Agency Name and Address <b>Texas A&amp;M Transportation Institute College Station, Texas 77843-3135</b>		13. Type of Report and Period Covered	14. Sponsoring Agency Code
15. Supplementary Notes Supported by the State of Texas. Project Title: Model Development with Passively Collected Data			
16. Abstract This report provides a review of the literature gathering information on the collection, processing, and use of passively collected data in travel demand modeling. Traditional methods of travel data collection are often limited by high costs, infrequent updates, or small sample sizes. Several emerging technologies now allow the ability to quickly and easily acquire large trip data sets over long periods of time. In this technical memorandum, five of the most promising new technologies (mobile phone positioning, GPS tracking, Bluetooth re-identification, social networking data, and smart card data) are reviewed in the context of travel demand modeling. These technologies are growing in use but come with their own set of considerations, particularly when it comes to sampling bias and data pre-processing. When possible, it appears that integrating new technologies with traditionally collected data can lead to enhanced accuracy. It remains to be seen whether these methods of data collection will be able to fully supplant traditional travel diaries or sensor data.			
17. Key Words <b>Travel Survey; Mobile Phone; GPS; Bluetooth Re-identification; Social Networking; Smart Card</b>		18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service Alexandria, Virginia <a href="http://www.ntis.gov">http://www.ntis.gov</a>	
19. Security Classif. (of this report) <b>Unclassified</b>	20. Security Classif. (of this page) <b>Unclassified</b>	21. No. of Pages <b>26</b>	22. Price

# **Emerging Data Collection Techniques for Travel Demand Modeling: A Literature Review**

by

Richard J. Lee  
Student Technician II  
Texas A&M Transportation Institute

Ipek N. Sener  
Assistant Research Scientist  
Texas A&M Transportation Institute

and

James A. Mullins III  
Research Scientist  
Texas A&M Transportation Institute

Strategic Research Program  
Final Report  
Project Number 161407-2

Texas A&M Transportation Institute  
College Station, Texas 77843-3135

October 2014

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

## **ACKNOWLEDGMENTS**

Support for this research was provided by the State of Texas. This technical memorandum constitutes the first phase of the overall project. The authors would also like to acknowledge Johanna Zmud, Kevin Hall, and Thomas Williams for their review of an earlier version of this document. Finally, the authors thank Linda Cherrington and Katie Turnbull for their guidance during the course of this research project.

## TABLE OF CONTENTS

<b>Introduction.....</b>	<b>1</b>
<b>Mobile Phone Positioning.....</b>	<b>2</b>
Trip and Traffic Parameters .....	2
Trip Distribution Modeling.....	3
Activity-Based Modeling and Other Applications .....	3
Considerations .....	4
<b>GPS.....</b>	<b>6</b>
GPS-Based Surveys .....	6
Mode and Trip Purpose Imputation .....	7
Other Applications .....	8
Considerations .....	9
<b>Bluetooth Re-identification .....</b>	<b>10</b>
<b>Social Networking .....</b>	<b>12</b>
<b>Smart Card .....</b>	<b>13</b>
<b>Other Considerations Regarding the Use of Passive Data .....</b>	<b>14</b>
<b>Summary.....</b>	<b>15</b>
<b>References.....</b>	<b>16</b>

## INTRODUCTION

In recent years, several technologies have provided new sources of low-cost and efficient data collection at a scale that was once impractical. These technologies are promising for their ability to generate large quantities of sophisticated traffic measures at micro- and macro-levels. Some of these techniques are now being employed in travel demand modeling applications, but whether they can entirely replace traditional methods of data collection remains to be seen. The following sections review the current state of several travel data collection technologies for travel modeling:

- Mobile phone positioning.
- Geographic Positioning System (GPS).
- Bluetooth® re-identification.
- Social networking.
- Smart card.

## MOBILE PHONE POSITIONING

With the increasing pervasiveness of mobile phones across the globe, mobile phone positioning data are enticing for their potential to supplement or even supplant traditionally collected data in travel demand models. Sometimes termed wireless location technology, floating phone data, or floating car data, *mobile phone positioning data* can refer to any georeferenced data captured from mobile phones. In this review, the term will be used specifically to refer to data received from base transceiver stations (BTSs) collected through the cellular network. Researchers primarily investigated the following applications of these data: estimation of trip and traffic parameters, trip distribution modeling, and activity-based modeling. The results from some of the reviewed studies suggest the potential for mobile phone positioning data to play a larger role in travel demand modeling, though their use comes with its own set of considerations.

Mobile phone positioning data are sporadically acquired by cellular carriers as a result of user-initiated events. Recorded events typically include calls, Internet use (e.g., email or web browsing), short message service (SMS) activity, or handovers/handoffs (which are triggered when transferring between wireless cells). The data are anonymously collected through cellular networks by signals sent from phones to base stations, each of which corresponds to a cell—the standard unit of aggregation for a cellular network—allowing for the identification of a phone within that cell. Positional accuracy is low because the level of precision depends on cell size, but rough vehicle trajectories can be approximated, allowing for the estimation of travel time, speed, and trip origin-destination (OD) matrices. It is possible to better estimate position within a cell based on signal strength or through triangulation, though these techniques require the use of SMS data or Internet protocol traffic (Steenbruggen et al. 2013). Such techniques become increasingly important in rural areas where cells are larger, resulting in uncertainty errors as high as several hundred meters. Mobile phone positioning data sets typically include the time each event occurred, the coordinates of the corresponding BTS and cell dwell time (CDT), which refers to the duration a phone is associated with each BTS. CDT can be a useful measure of traffic congestion and can also be used to distinguish home and work locations for trip purpose imputation.

### Trip and Traffic Parameters

The first major assessments of the utility of mobile phone data in transportation planning began in 1994, with techniques being developed to estimate travel time and speed (Steenbruggen et al. 2013; University of Maryland 1997). These early studies were limited by poor accuracy and small sample sizes, but established a base for future research into more advanced modeling and refined methodologies. A number of studies have now been conducted using cellular network data to derive aggregate traffic parameters. In 2007, researchers used AirSage data in Minnesota to estimate highway travel time and speed, finding that the cellular measures were somewhat inconsistent, depending on levels of congestion, but were generally within 10 mph when compared to those derived from loop detectors and license plate matching (Liu et al. 2008). Velocities derived from mobile phone positioning data also have the benefit of being measures of average velocity, which can be more useful than the instantaneous velocities recorded by loop detectors (Caceres et al. 2008). To evaluate mobile-phone-based distance and flow measures, Huntsinger and Donnelly (2014) compared cellular data to a traditional survey-based model in

the Raleigh-Durham, North Carolina region. Trip length and highway traffic volume were generally comparable, with the mobile-phone-based model outperforming the traditional model for rural highway volume estimations.

## **Trip Distribution Modeling**

Beyond the estimation of basic traffic parameters, mobile phone positioning data have also been used in trip distribution modeling, which has traditionally relied on costly and labor-intensive surveys or traffic counts, requiring an existing OD matrix and a large network of sensors. Calabrese et al. (2011) developed a methodology for OD estimation using AirSage cellular data and eastern Massachusetts as a case study. Findings indicated improved accuracy at greater levels of aggregation; correlating results to county-level Census-derived OD data resulted in an  $R^2$  value of 0.76, but just 0.3 at the Census tract level. Iqbal et al. (2014) used call detail records to develop an OD matrix with the aid of video-generated traffic counts and a traffic simulator to calibrate scaling factors based on cellular penetration rate, mobile phone usage, and vehicle usage. Only evaluating call events may have undersampled safer drivers who may be less likely to talk on their phones while operating a vehicle, though it is unclear whether this would have had an impact on trip distribution.

In what is claimed to be the first major application of cell phone data in a regional/national travel demand model, Gur et al. (2009) employed cellular-derived OD matrices for Israel as part of a four-step model. To extract trip lengths from the raw cellular data, the researchers defined trip ends using a 20-minute BTS connection threshold and chained all trips with dwell times under 15 minutes. A general survey was also conducted to characterize the distribution of cell phone users in Israel in order to develop expansion factors accounting for sample bias. Although the mobile phone data were suitable for district-level modeling, accuracy issues existed at the traffic analysis zone (TAZ) level, which required the assignment of approximately 2,200 BTSs to 600 TAZs. Huntsinger and Donnelly (2014) also demonstrated reliable results when modeling mobile-phone-based district flows in the Triangle Region of North Carolina. The findings from these studies, and the work of Calabrese et al. (2011), suggest that mobile phone-derived OD matrices may only be suitable for coarse-scale zonal analysis.

## **Activity-Based Modeling and Other Applications**

Mobile phone positioning data have been infrequently studied in activity-based modeling with somewhat limited success due to location imprecision and the often infrequent sampling of data points compared to GPS data. Using a machine learning method, Freund and Bar-Gera (2013) distinguished between activity and travel events based on the distance and duration of trips with a reported classification error of 13 percent. Given the problem of distinguishing activities and travel itself, differentiating between trip purposes is even more challenging. Phithakkitnukoon et al. (2010) attempted to determine daily activity patterns based on land use and Massachusetts AirSage data, but it is doubtful that a cell resolution of 500 meters could result in accurate activity classification.

Improvements in location accuracy will likely be necessary before it becomes feasible to accurately determine activity and trip purpose information from cellular data. At the moment, mobile phone positioning data appear to be best suited for the measurement of simple traffic

measures like traffic volume and travel time, although advances are currently being made in the area of OD matrix estimation. Other possible applications of mobile phone positioning data include route choice modeling (Schlaich et al. 2010) and transit passenger volume measurement (Aguiléra et al. 2012). Abdulazim et al. (2013) explored the use of mobile phone location data along with phone sensor data for travel survey data collection, but the reported location errors of 100 to 600 meters, and more than 4000 meters in rural areas, indicate that the data are not yet reliable enough for this technique to be practical.

## Considerations

Despite the obvious benefits of mobile phone positioning data, including availability and sample size, when used in travel demand modeling, a number of issues arise that must be carefully considered:

- False displacements occur when a phone is positioned near a cell boundary, allowing it to communicate with two different base stations and giving the appearance of oscillation between cells. One approach to dealing with false displacements is to discard unrealistic records using a specified time threshold. Wang et al. (2014) and Iqbal et al. (2014) both removed displacements within 10 minutes of each other, while Iqbal et al. additionally removed trajectories with time gaps of over one hour.
- Redundant records can occur as a result of individuals having more than one phone or vehicles holding multiple passengers with phones. Fang et al. (2014) took steps to eliminate redundancy in their data set before generating complete vehicle trajectories, though the specific procedure they employed was not detailed. Horn et al. (2014) evaluated filtering techniques to smooth vehicle trajectories using an average velocity threshold of 250 km/h to identify outliers. The methods significantly improved speed and position estimates.
- Because there are no set standards for pre-processing mobile phone positioning data, a number of different methods have been employed in the literature. More research into this area would be beneficial to determine a generalized approach that could be applied across different data sets and study areas. Whatever approach is taken, due to the typically vast size of these data sets (often in the millions of records), pre-processing procedures will likely need to be automated.
- Cellular network coverage is not necessarily widespread, and for studies in rural areas, gaps in coverage maps are more likely to become an issue. Urban studies suffer from their own unique concerns; despite having the benefit of smaller cell sizes and comprehensive cellular coverage, the mixture of modes including vehicles, transit, walking, and cycling can muddle the data.
- Travel demand model TAZs may not be aligned with BTS cell geographies, which is problematic when generating OD matrices from mobile phone data. Zone assignment procedures need to be implemented, resulting in a loss of precision when the data are re-aggregated.
- Data quality standards cannot be guaranteed since mobile phone positioning data are provided externally by mobile phone carriers.
- There may be sampling bias associated with the failure to capture non-cell-phone users. Palm and Knudsen (2014) compared demographic and travel behavior between cell-

phone-only households and households with landlines using the 2009–2011 Oregon Household Activity Survey. Those in cell-phone-only households were more than twice as likely to participate in a car-sharing program, bike commute, or use public transit, and live in denser areas. These results suggest that cell phone users may be more predisposed to using alternative travel options, although having a landline does not necessarily correspond to the lack of a mobile phone. Identifying the sociodemographic differences between cell phone users and non-users would be even more valuable in determining whether there are significant differences in travel behavior between the groups. Another potential area of research is determining differences in the demographics and travel behavior of subscribers of different cellular providers. Because mobile phone positioning data are commonly acquired from a single source, it is worth investigating whether bias exists due to differences between travelers using different carriers.

## GPS

GPS data can be superior to mobile phone positioning data in that they do not suffer from the same accuracy concerns due to higher precision. GPS data are also typically collected more frequently, often at a sampling rate of between 0.5 to 60 times per minute (Patile et al. 2014), allowing for more precise disaggregate trip data collection. With a finer spatial and temporal resolution, GPS offers high quality route, travel time, and OD data for use in travel modeling, including transit and freight applications. Particularly in the context of smartphones, research has been directed toward the enhancement of content traditionally derived through direct contact with survey participants, such as trip purpose and mode detection, real-time travel prediction, and route choice modeling.

### GPS-Based Surveys

The first major research into GPS-based travel survey data in the United States began in 1996, with many more studies being conducted following the 2000 Census (Wolf 2004). The results of these comparisons between vehicle-based GPS data and traditional computer-assisted telephone interview survey data found that the number of trips recorded by traditional surveys was underreported by 11–81 percent (Wolf 2004). Stopher et al. (2007) discovered a missing trip rate of 7.4 percent when using GPS to assess the accuracy of face-to-face interviews for travel data. It is now accepted that trip frequencies are underreported on travel diaries, though it appears that the opposite phenomenon may occur with trip lengths. Comparing self-reported trip lengths versus GPS-based trip length measurements, Kelly et al. (2013) discovered that self-reported trip lengths were overestimated by between 2.2 and 13.5 minutes per trip. Along with simple errors of perception, it was suggested that trip length overestimation occurred as a result of participants factoring non-travel activities into trip times (e.g., vehicle loading) and sometimes grouped trip chains into a single trip.

These studies reveal the shortcomings of traditionally-collected travel data. The incorporation of GPS with travel diaries can now allow modelers to adjust for these shortcomings and improve model estimation. Despite the reliability of GPS, Safi et al. (2013) suggested prompted recall (PR) surveys may be necessary for user verification to account for errors in the GPS data themselves. PR surveys increase the burden on survey participants compared to passive GPS data collection, but they would allow respondents to correct erroneous trip data and provide valuable trip and activity information. Greaves et al. (2010) demonstrated this technique over an eight-week period with high rates of survey completion (29 out of 30 completed), despite it being a more onerous form of data collection for participants.

Smartphone apps can be used to collect GPS travel survey data while providing an integrated interface for PR user verification. Carrion et al. (2014) had participants install an app that used GPS data to infer activity and mode, which were later user-verified. Results indicated that the smartphone-based survey captured more trips and more precise travel times when compared to the traditional survey. Similarly, Maruyama et al. (2014) reported higher trip rates captured by a smartphone-based survey compared to a traditional travel diary. These surveys can be more comprehensive than traditional travel surveys while being less intrusive to study participants. Through the app interface, users can supply additional trip or demographic data and validate automatically generated trip records to remove or correct erroneous data, though difficulties

regarding participant recruitment and sampling bias remain. Although more time- and labor-intensive for modelers, there is great potential to improve travel forecasting accuracy when traditional travel diaries are integrated with GPS data and PR surveys.

## **Mode and Trip Purpose Imputation**

Research efforts to expand the use of GPS data in travel modeling have included work on the detection of traveler mode and trip purpose. Different approaches employed in the literature include machine learning algorithms, rule-based algorithms, multinomial logit (MNL) models, and fuzzy logic models. Lin et al. (2013) developed an unsupervised technique to detect travel mode from GPS data. Their method relied on the difference in high-speed distributions between modes and resulted in a relatively high classification accuracy of 87 percent and 80 percent for car and bike modes, respectively, but just 55 percent and 62 percent for walk and bike modes, respectively. Another approach is to initially infer trip and activity events based on dwell time and then subsequently determine mode type (Schuessler and Axhausen 2009), although this procedure can lead to a compounding of errors should there be misclassifications in the original event detection (Feng and Timmermans 2014).

Feng and Timmermans (2014) employed an alternative approach by concurrently detecting travel mode and activity episode to improve classification accuracy. Several machine learning algorithms were compared because ad-hoc rules-based approaches are less capable of representing complex data sets and cannot be generalized to other studies. The authors reported extremely high accuracies (99.7 percent and 99.8 percent, respectively), noting that similar studies reported accuracies of 70 percent to 85 percent. Despite the promising results of the study, expectation should be tempered given the inclusion of non-GPS survey data and a limited sample size of just eight participants. Montini et al. (2014) applied another machine learning approach to identify trip purpose from 156 participants outfitted with GPS devices. An overall accuracy of 82 percent was achieved, though non-home- and non-work-based activities including shopping and recreational activities were less accurately categorized. The authors noted that the inclusion of land use data would likely have improved accuracy and should be considered for future studies.

Combining GPS, accelerometer, and transit network data, Broach et al. (2014) compared the effectiveness of these data sources in travel mode detection. They developed three MNL models based on different sets of data, finding that the GPS-only model had difficulty distinguishing between transit and auto modes and poorly identified bicycle trips (68.5 percent). The second model integrated geographic information system (GIS) transit network data, allowing for greater distinction between transit and auto modes. The final model added accelerometer data, which moderately improved bicycle classification accuracy from 68.5 percent to 74.2 percent. Jahangiri and Rakha (2014) also employed GPS and accelerometer data, along with gyroscope and rotation vector data, from smartphone sensors to detect transportation mode using a machine learning algorithm. Results were promising, with high accuracies when combining multiple sensors both with and without GPS data (98.86 percent and 97.89 percent, respectively). Although the study's sample consisted of just one day of travel data from three people, it demonstrated a new approach to smartphone-based mode detection without the aid of GPS data. The study claims to be the first to implement gyroscope and rotation vector sensors for transportation mode detection.

## **Other Applications**

Applications of GPS in transit and freight modeling have also been performed. Kuppam et al. (2014) acquired more than 3 million GPS event records from the American Transportation Research Institute (ATRI) to develop a land-use-based travel demand model for heavy trucks. Similar to mobile phone positioning data, the GPS data were event based and could be initiated by hard braking, vehicle speeds, or pings from a central source. The model predicted the number of tours, the number of stops, and whether the tour returned to its starting location, though the model could have been improved through GIS integration to calculate network distances between points.

At a smaller scale, Fan and Gurmu (2014) acquired external GPS stop data for a single bus line in Macae, Brazil, for use in dynamic bus travel time prediction. Typically requiring several inputs including flow, speed, distance, and weather, the machine learning approach (artificial neural network [ANN]) predicted bus travel time using only GPS stop data as an input. Lu and Li (2014) similarly employed an ANN model for real-time traffic prediction using externally acquired taxi GPS data. Other potential applications of GPS data include automatic vehicle classification (Sun and Ban 2013), bicycle and vehicle route choice analysis (Casello and Usyukov 2014; Spissu et al. 2011), the integration of GPS with loop detector data for improved real-time travel estimation (Patire et al. 2014), and the use of GPS data as ground-truth data in conjunction with Bluetooth or mobile phone data (Omrani et al. 2013).

GPS data have been used in a variety of modeling applications, but have been limited by practical constraints associated with the devices themselves. With the recent growth of GPS-enabled mobile phones (or smartphones), a promising area of research is in the use of smartphone apps to harness the GPS capabilities now prevalent in most phones. Once installed on participants' phones, apps can monitor and record GPS tracks, sometimes in real time. Participants must be smartphone users with a base level of technological savvy and are required to carry their phone with them at all times, but by relying on individuals' personal phones, there is no need for additional equipment and participant burden is lessened.

Jackson et al. (2014) developed a free app for cyclists in Montreal to capture GPS-based cyclist route data. Successful adoption of the app by the cycling community was a result of advertising from the City of Montreal, word-of-mouth marketing, and beneficial features for cyclists including speed, distance, and caloric expenditure data. At 23 days after release, the app had been installed on over 500 phones, with more than 2,300 trips reported. This example demonstrates the great potential for smartphone apps to significantly simplify participant recruitment. It can be challenging to find participants willing to complete detailed travel diaries, but the advent of passive data and smartphone apps may provide an easier avenue for modelers to acquire reliable travel data. Others have used smartphone apps to collect GPS and sensor data (e.g., accelerometer, orientation, or gyroscope sensors) to identify transportation mode through machine learning algorithms. Using such an approach, Nitsche et al. (2012) achieved high classification accuracies of 92 percent and 98 percent for walk and bike modes, respectively, but lower accuracies when identifying car, rail, and motorcycle modes.

## Considerations

The main obstacles for the use of data collected from standalone GPS devices in travel modeling are cost and data acquisition. Each vehicle or participant will need to be outfitted with a dedicated GPS device, which can be costly to procure and burdensome for participants to track. Stopher and Greaves (2007) estimated that more than 400 GPS devices would be required to conduct week-long surveys over the course of a few months for a sample size typical of travel diaries (i.e., 3500–5000 households). These devices may also require recharging depending on the length of the study period and may be considered an inconvenience by participants. The more recent advent of smartphone-based GPS technology takes advantage of the personal phones of participants, mitigating many of these concerns. Study recruitment can remain a challenge because participants must typically install an app allowing researchers to access and record their GPS tracks.

Despite high confidence in GPS's positional accuracy, GPS signal strength can be adversely affected by obstructions, particularly in dense urban environments where the prevalence of tall buildings, trees, and other barriers can lead to data gaps or incomplete trajectories. Small differences in sampling frequencies (e.g., once every second vs. once every three seconds) can also affect the accuracy of GPS-derived traffic parameters. Modelers should be aware of these limitations when attempting to implement GPS-based data. Piccoli et al. (2014) determined that accurate estimation of acceleration and fuel consumption required more frequent GPS sampling, although velocity, density, and travel time appeared to be relatively invariant to changes in sampling frequency.

As with mobile phone positioning data, there are no set standards for the substantial pre-processing often required of raw GPS data, and approaches varied in the literature. Depending on the data set and study objectives, the typically large GPS data sets may need to be aggregated, filtered to clean up outliers and noise, or projected onto a GIS road network. For their GPS-based travel survey, Carrion et al. (2014) removed any data demonstrating gaps or unrealistic speeds. Jackson et al. (2014) and Lin et al. (2013) used Kalman filtering—which recursively estimates discrete GPS points—to smooth noisy GPS tracks. Broach et al. (2014) and Greaves et al. (2010) took a simpler approach by manually inspecting and discarding outliers using their best judgment. Similarly, Kuppam et al. (2014) performed a visual inspection of data acquired from ATRI for a small sample of truck trips, though they did not clean the entire data set. Before they are able to be used in travel modeling applications, GPS data will likely need to be cleaned or pre-processed. Because a number of approaches have been applied in the literature, this is an area modelers should consider carefully.

## BLUETOOTH RE-IDENTIFICATION

Bluetooth re-identification is used to derive location information from devices enabled with Bluetooth wireless technology acting as probes. Bluetooth devices are distinguished by their unique media access control (MAC) address and identified by detectors to develop vehicle trajectories (Chitturi et al. 2014). Because the technique requires the deployment of a series of readers able to detect Bluetooth-enabled phones, most studies are constrained to single stretches of road or small networks. Bluetooth re-identification has been used for OD matrix estimation, route choice modeling, and mode detection, and can be used to accurately and directly measure travel time over long periods. Because of low sampling rates, volume measures need to be inferred and sampling bias will be a concern. Despite growing Bluetooth penetration, sampling rates for recent studies in the United States have typically ranged between 5 and 7 percent (Carpenter et al. 2012; Malinovskiy et al. 2012; Patire et al. 2014; Reiff 2012).

The spatial accuracy of Bluetooth-generated data is much higher than it is for mobile phone positioning data, but being a mostly mobile-phone-based technique, it still suffers from some of the same difficulties including detections from other modes and detections from multiple phones within the same vehicle. Bluetooth sensors can also detect devices up to 300 feet away and so are better suited for isolated highway stretches rather than densely packed arterials. Oversampling is the greatest concern when using Bluetooth data because devices can be detected multiple times by the same sensor, especially in slow or congested traffic, depending on how long they remain in the detection zone. Typically, a heuristic rule, such as selecting only the first detection, is used to manage multiple detections. Detecting outliers in real time is more challenging given the time lag that occurs when measuring travel time between detectors. Moghaddam and Hellinga (2014) developed an adaptive algorithm to tackle this problem, using recent travel time data and historical patterns to establish a dynamic window to detect outliers, although rapid changes in travel time still resulted in false positives.

Confidence in travel time estimations using Bluetooth re-identification is generally strong, having been successfully used as ground-truth data to verify GPS-probe travel times (Aliari and Haghani 2012) and for real-time traffic forecasting (Barcelo et al. 2010). Bluetooth travel times have proven more accurate than in-vehicle navigation and mobile-phone-based GPS data (Omrani et al. 2013), while Haghani et al. (2010) found that measurements did not differ significantly with GPS-based travel times at lower speeds (0–45 mph), although disparities were seen at higher speeds.

To improve traffic speed measurement accuracy, Bachmann et al. (2013) devised algorithms to fuse Bluetooth data with loop detector data, noting that fusion could also be conducted with multiple loop detectors to improve accuracy. Another recent method for improving the accuracy of Bluetooth travel times was developed by Saeedi et al. (2013). They selected MAC addresses based on received signal strength indicator (RSSI) data from their detectors, which are correlated with distance to devices. Their RSSI-based method improved travel time errors by more than 26 seconds on highway segments in Tigard, Oregon. Bluetooth travel times have also been estimated for pedestrians along short walking corridors, but Malinovskiy et al. (2012) cautioned that low Bluetooth sampling rates necessitate that analysis be conducted in areas with heavy pedestrian traffic.

Although most Bluetooth data are limited to travel time estimations due to limited sensor coverage, some research has been conducted in small-scale OD estimation. Chitturi et al. (2014) demonstrated the use of Bluetooth data to generate an OD matrix for a Madison, Wisconsin, highway interchange. The researchers scaled and validated their data against ground-truthed traffic counts using biproportional matrix balancing to account for variable Bluetooth capture rates. The OD matrix generated through time lapse aerial photography (TLAP) outperformed the Bluetooth model, though satisfactory accuracy was obtained when Bluetooth data were aggregated over multiple days. Given the high costs and logistical difficulties associated with obtaining TLAP data, Bluetooth data proved to be a reasonable alternative, though the procedure is only feasible for small, closed networks.

At the corridor level, Carpenter et al. (2012) deployed 14 Bluetooth detectors along 15 miles of highway in Jacksonville, Florida, for OD estimation. Half of the sensors were placed along the highway and half on major cross streets, though not enough were available for comprehensive system coverage, meaning vehicles could enter or leave the system without being tracked. The results of the study were not checked against ground-truth data to determine accuracy. Similarly, the South Carolina Department of Transportation used Bluetooth data to estimate origins and destinations along the congested I-526 corridor in Charleston County in South Carolina by deploying 13 sensors near highway interchanges (Reiff 2012). Despite lower detection rates, the agency selected Bluetooth over automatic license plate matching because it enabled longer data collection and was more cost-effective, at less than one-fourth the price.

In addition, over the past several years, numerous studies by researchers at the Texas A&M Transportation Institute (TTI) have assessed the utility and quality of Bluetooth OD data using various new methods and technologies (Farnsworth et al. 2011; Farnsworth and Hard 2013a, b). The researchers developed and evaluated Bluetooth OD data for corridor studies and external station surveys using mobile Bluetooth readers and TTI's anonymous wireless addressing matching technology. Study outcomes demonstrated the ability of Bluetooth to generate sound estimates of OD movements along a corridor or external-external movements through a modeling study area. Currently the researchers are conducting a comparison of Bluetooth, cellular, and secondary GPS data for the development of external OD matrices. The comparison study, whose results will be available in late 2014, focuses on the case study area of Tyler, Texas (Hard et al. 2014).

Less common applications of Bluetooth re-identification data include single-corridor route choice modeling (Hainen et al. 2011) and mode detection. Araghi et al. (2012) used statistical clustering methods to distinguish between vehicles and bicycles based on Bluetooth travel times with a mean absolute percentage error of 14 percent. The researchers also noted that class of device numbers could be used to identify Bluetooth device type (e.g., phone, laptop, or tablet) in order to improve mode detection algorithms. Continued research into Bluetooth mode detection may make Bluetooth re-identification more viable for data collection on mixed-use arterials.

Bluetooth re-identification allows for reliable, long-term collection of vehicle trajectories and travel time data, which are valuable inputs in travel demand models. Larger-scale data collection and more advanced modeling will depend greatly on the arrangement and number of sensors available. Limited research into mode detection and OD estimation has been conducted with questionable accuracy, though only for small networks and single stretches of highway.

## SOCIAL NETWORKING

Location-based social networking (LBSN) data rely on mobile phone GPS functionality, allowing users to check in at venues and resulting in a database of user-verified trip data. Because check-ins are confirmed by users, the spatial resolution and data quality of origins and destinations are extremely high, making it a logical choice for trip distribution modeling. Although LBSN venue data are too coarse to approximate individual user trajectory data, acquisition of data is simplified because users voluntarily provide location data, which can sometimes be tied to trip activity or demographic data.

Jin et al. (2014) developed a doubly constrained gravity model to generate an OD matrix based on Foursquare® check-ins in Austin, Texas. The authors found significant errors among intra-zonal trips and venue bias because users were more likely to check in at certain venues (e.g., nightlife or recreational venues versus workplace venues). The study also suffered from temporal alignment issues because a 2005 OD matrix was used as a reference data set for the 2012 LBSN data. Refinement of the model could result in a method to provide more frequent updates to an existing OD matrix. Yang et al. (2014) also used Foursquare check-ins to generate an OD matrix in Chicago, Illinois. The researchers began with cluster and regression models to convert check-in data to productions and attractions, which were then used as inputs in a gravity model. The LBSN-derived OD model compared reasonably well to the survey-based OD matrix, although results may be overly optimistic given that the original OD matrix was used to calibrate the model. Discrepancies also existed, partially due to the lack of complete spatial coverage from the venue data among the TAZs. LBSN data may have potential for use in OD modeling, though their use as a standalone data set has not yet been clearly demonstrated. The requirement of historical OD data to calibrate the LBSN-derived OD matrices also means that any errors in the original data set will propagate through the new model.

Research efforts directed toward the mitigation of sample bias in social networking data as part of use in modeling were identified. These included efforts to determine what, if any, differences existed in travel behavior between social media users and non-users. Kamargianni and Polydoropoulou (2014) used a latent class model to segment teenage survey participants in Greece into three social networking classes: indifferent users (lowest social media usage), rational users, and addicted users (highest social media usage). The authors determined that rational and addicted users conducted more social trips than indifferent users, though no other trip activities were explored. Though more work in this area is required, there appear to be differences in travel behavior between heavy users of social media and non-users, which will be an important consideration for any LBSN-based study.

## SMART CARD

For transit-oriented research, smart card technology is now providing a new source of extensive and reliable long-term trip data, which has been used in a variety of applications including the enhancement of travel survey data, trip purpose imputation, and travel time estimation. Transit agencies can track smart card validations, resulting in massive longitudinal databases that are typically absent of socio-economic data, much like other passive data technologies. Similar to GPS or mobile phone data sets, raw smart card data need to be processed to remove transaction errors and to identify transfers and probable alighting points (Chu and Chapleau 2008).

Spurr et al. (2014) used smart card data to validate a Montreal telephone travel survey in order to reveal underlying bias, likely as a result of underreporting, non-response, and entry errors. Results of the comparison indicated that although total daily volume totals were comparable, the traditional household survey overestimated subway ridership during peak periods and underestimated it during off-peak periods. These results were then used to reweight the original survey data to compensate for the discrepancies. Riegel and Attanucci (2014) attempted to merge smart card data with the London Travel Demand Survey in order to take advantage of the demographic and trip purpose data reported in traditional travel surveys. The authors were only able to link approximately half of the trips with smart card data, but the results demonstrated the error inherent in user-reported travel surveys, with the average start time between smart card and survey data differing by 61.2 minutes. These studies indicate that smart card data can be a reliable and beneficial tool for modelers to improve estimation from traditional transit travel data.

In the absence of activity data, Devillaine et al. (2012) attempted to estimate trip purpose using a Santiago, Chile, smart card database and a comprehensive nine-year database of smart card transactions from Gastineau, Quebec. The authors distinguished between four activity types based on activity duration and land use—work, school, home, and other—and used a 30-minute threshold to differentiate activities from transfers. However, without survey data for validation, the authors had no way to assess the accuracy of their approach. The database of Seoul smart card transactions explored by Jang (2010) required significantly less data processing than a typical smart card data set because it already included full trip itineraries. The data were used to analyze transfer behavior and to calculate travel time between stops.

Smart card data have also been used to classify transit passengers based on travel regularity. Kieu et al. (2014) used data from more than one million cardholders in Brisbane, Australia, to mine historical travel itineraries and identify travel patterns. The analysis resulted in four passenger categories: regular OD users, habitual time users, routine OD and time users, and irregular users. This type of segmentation could allow transit operators to deliver more targeted services and information to its customers.

## OTHER CONSIDERATIONS REGARDING THE USE OF PASSIVE DATA

When conducting or evaluating studies researching or implementing passive data for travel modeling, careful consideration must be given to model validation to ensure predictive confidence. The holdout method was the most common validation technique found in the literature and was used for mode detection, trip distribution, and route choice modeling (Cassello and Usyukov 2014; Feng and Timmermans 2014; Jahangiri and Rakha 2014; Lu and Li 2014; Montini et al. 2014; Xia et al. 2014). This validation method sets aside a portion of the data set for model training, while the other independent sample is used for model evaluation. Typically 20–30 percent of the data set was set aside as test data, although Xia et al. (2014) split their data set in half for training and testing in their dynamic OD estimation model. A number of other external data sources were also used in model validation including user-validated GPS data (Feng and Timmermans 2014) and travel mode survey data (Broach et al. 2014) for mode detection studies, while Bluetooth re-identification (Bhaskar et al. 2014) and historical bus travel data (Fan and Gurmu 2014) were used for travel time prediction models. For OD trip distribution modeling, Huntsinger and Donnelly (2014) used traditional survey-based OD data, while Iqbal et al. (2014) employed video-generated traffic counts in their small-scale study. Yang et al. (2014) validated their LBSN-based OD matrix against a survey-based OD matrix, although the survey-based matrix had been used as an input in their model, likely resulting in optimistic performance estimates. Other reviewed studies failed to perform any model validation (Huang and Levinson 2014; Wang et al. 2014), perhaps due to practical constraints.

Another important concern regarding passive data is that of privacy, particularly because individuals are often unaware of their movements being recorded. Although privacy was rarely mentioned in the reviewed literature and typically not a concern for aggregated travel data, there may be ethical implications when developing individual vehicle tracks or disaggregate models. From an online survey of 120 mobile app users, Cottrill (2014) noted that many were concerned about privacy but were willing to give up some personal information for reasonable benefits. Antoniou and Polydoropoulou (2014) attempted to quantitatively determine how much privacy was valued by travelers based on hypothetical GPS-based navigation subscription costs. Results of the survey indicated that participants were willing to give up one level of privacy for 2.2 euros/month, and women were less willing to provide personal information than men. Even when user consent is given for the use of location-based smartphone data, users are often unaware of how these data are being employed. Privacy policies, for example, are infrequently read and written in such a way as to be ineffective at communicating privacy practices to consumers (Cottrill 2014). The explosion of passively collected location data has been a great boon for transportation planners and researchers, but care must be taken to ensure that the anonymity and privacy of individuals are maintained.

## SUMMARY

The current state of new methods of data collection—including mobile phone positioning, GPS tracking, Bluetooth re-identification, social networking data, and smart card data—was evaluated in the context of travel demand modeling. Emerging in recent years with the promise of supplementing or replacing traditional data collection methods, these techniques are providing new sources of travel data for a variety of potential applications including travel survey enhancement, estimation of travel demand modeling inputs, and dynamic traffic management systems. These new data sources are still being studied and evaluated for feasibility and reliability, and each method comes with its own set of considerations.

With these new sources of data, special consideration should be given to sampling bias because they will be drawing from a different population than that of traditionally collected travel data. When possible, integrating new data collection technologies with traditionally collected data to either validate or supplement models could lead to more reliable results. Hybrid approaches appear to be especially attractive for their ability to blend the qualitative information required of activity-based and behavioral modeling with extensive trip data. The data collection techniques evaluated in this report are not comprehensive; other potential areas of research include big data (e.g., credit card purchase records) analysis and vehicle-to-infrastructure data. Although little to no research has yet been conducted in these areas, they may represent an important source of travel data as their use becomes more viable in the future.

## REFERENCES

- Abdulazim, T., H. Abdelgawad, K. M. N. Habib, and B. Abdulhai (2013). “Using Smartphones and Sensor Technologies to Automate Collection of Travel Data.” *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2383, pp. 44–52.
- Aguiléra, V., S. Allio, V. Benezech, F. Combes, and C. Milion (2012). “Estimating the Quality of Service of Underground Transit Systems with Cellular Network Data.” *Procedia-Social and Behavioral Sciences*, Vol. 48, pp. 2262–2271.
- Aliari, Y., and A. Haghani (2012). “Bluetooth Sensor Data and Ground Truth Testing of Reported Travel Times.” *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2308, pp. 167–172.
- Antoniou, C. and A. Polydoropoulou (2014). “How Likely Are Travelers to Give Up Information in Exchange for Better User Information Services?” *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Araghi, B. N., L. T. Christensen, R. Krishnan, and H. Lahrmann (2012). “Application of Bluetooth Technology for Mode-Specific Travel Time Estimation on Arterial Roads: Potentials and Challenges.” *Proceedings from the Annual Transport Conference at Aalborg University, Denmark*.
- Bachmann, C., M. J. Roorda, B. Abdulhai, and B. Moshiri (2013). “Fusing a Bluetooth Traffic Monitoring System with Loop Detector Data for Improved Freeway Traffic Speed Estimation.” *Journal of Intelligent Transportation Systems*, Vol. 17, No. 2, pp. 152–164.
- Barcelo, 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*, Vol. 2175, pp. 19–27.
- Bhaskar, A., M. Qu, and E. Chung (2014). “A Hybrid Model for Motorway Travel Time Estimation: Considering Increased Detector Spacing.” *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Broach, J., N. W. McNeil, and J. Dill (2014). “Travel Mode Imputation Using GPS and Accelerometer Data from a Multi-day Travel Survey.” *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Caceres, N., J. P. Wideberg, and F. G. Benitez (2008). “Review of Traffic Data Estimations Extracted from Cellular Networks.” *IET Intelligent Transport Systems*, Vol. 2, No. 3, pp. 179–192.
- Calabrese, F., G. Di Lorenzo, L. Liu, and C. Ratti (2011). “Estimating Origin-Destination Flows Using Mobile Phone Location Data.” *IEEE Pervasive Computing*, Vol. 10, No. 4, pp. 36–44.

- Carpenter, C., M. Fowler, and T. J. Adler (2012). "Generating Route-Specific Origin-Destination Tables Using Bluetooth Technology." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2308, pp. 96–102.
- Carrión, C., F. Pereira, R. Ball, F. Zhao, Y. Kim, K. Nawarathne, and N. Zheng (2014). "Evaluating FMS: A Preliminary Comparison with a Traditional Travel Survey." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Casello, J. M. and V. Usyukov (2014). "Modeling Cyclists' Route Choice Based on GPS Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Chitturi, M. V., J. W. Shaw, J. R. Campbell IV, and D. A. Noyce (2014). "Validation of Origin-Destination Data from Bluetooth Re-identification and Aerial Observation." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Chu, K. K. A., and R. Chapleau (2008). "Enriching Archived Smart Card Transaction Data for Transit Demand Modeling." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2063, pp. 63–72.
- Cottrill, C. D. (2014). "Considering Smartphones: User Attitudes towards Privacy and Trust in Location-Aware Applications." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Devillaine, F., M. Munizaga, and M. Trépanier (2012). "Detection of Activities of Public Transport Users by Analyzing Smart Card Data." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2276, pp. 48–55.
- Fan, W. and Z. K. Gurmu (2014). "Dynamic Travel Time Prediction Models for Buses Using Only GPS Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Fang, J., M. Xue, and T. Z. Qiu (2014). "Anonymous Cellphone-Based Large Scale Origin-Destination Data Collection: Case Studies in China." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Farnsworth, S. and E. Hard (2013a). "Metropolitan Area Planning Agency External Travel Survey—Summary Report." Technical Paper, Texas A&M Transportation Institute.
- Farnsworth, S., and E. Hard (2013b). *Alice Transportation Study*. Texas A&M Transportation Institute.
- Farnsworth, S., T. Voigt, and D. Borchardt (2011). *Summary Findings: Conduct Field Test Using Bluetooth Technology to Collect External Travel Movements*. Texas A&M Transportation Institute.
- Feng, T. and H. J. P. Timmermans (2014). "Comparison of Advanced GPS Data Imputation Algorithms for Detection of Transportation Mode and Activity Episode." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.

- Freund, S. and H. Bar-Gera (2013). "An Optimization Framework for Travel Pattern Interpretation of Cellular Data." *International Journal of Transportation Science and Technology*, Vol. 2, No. 3, pp. 167–182.
- Greaves, S., S. Fifer, R. Ellison, and G. Germanos (2010). "Development of a Global Positioning System Web-Based Prompted Recall Solution for Longitudinal Travel Surveys." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2183, pp. 69–77.
- Gur, Y. J., S. Bekhor, C. Solomon, and L. Kheifits (2009). "Intercity Person Trip Tables for Nationwide Transportation Planning in Israel Obtained from Massive Cell Phone Data." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2121, pp. 145–151.
- Haghani, A., M. Hamed, K. F. Sadabadi, S. Young, and P. Tarnoff (2010). "Data Collection of Freeway Travel Time Ground Truth with Bluetooth Sensors." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2160, pp. 60–68.
- Hainen, A. M., J. S. Wasson, S. M. Hubbard, S. M. Remias, G. D. Farnsworth, and D. M. Bullock (2011). "Estimating Route Choice and Travel Time Reliability with Field Observations of Bluetooth Probe Vehicles." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2256, pp. 43–50.
- Hard, E., L. Larsen, and I. N. Sener (2014). *Methods for Collecting External Survey Data*. Texas A&M Transportation Institute.
- Horn, C., S. Klampfl, M. Cik, and T. Reiter (2014). "Detecting Outliers in Cell Phone Data: Correcting Trajectories to Improve Traffic Modeling." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Huang, A. and D. Levinson (2014). "A Survival Analysis-Based Choice Set Formation Approach for Single-Destination Choice Using GPS Travel Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Huntsinger, L. F. and R. Donnelly (2014). "Reconciliation of Regional Travel Model and Passive Device Tracking Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Iqbal, S., C. F. Choudhury, P. Wang, and M. C. Gonzalez (2014). "Development of Origin-Destination Matrices Using Mobile Phone Call Data: A Simulation Based Approach." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Jackson, S., L. F. Miranda-Moreno, C. Rothfels, and Y. Roy (2014). "Adaptation and Implementation of a System for Collecting and Analyzing Cyclist Route Data Using Smartphones." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.

- Jahangiri, A. and H. Rakha (2014). "Developing a Support Vector Machine (SVM) Classifier for Transportation Mode Identification Using Mobile Phone Sensor Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Jang, W. (2010). "Travel Time and Transfer Analysis Using Transit Smart Card Data." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2144, pp. 142–149.
- Jin, P. J., M. Cebelak, F. Yang, B. Ran, and C. M. Walton (2014). "Location-Based Social Networking Data: An Exploration into the Use of a Doubly-Constrained Gravity Model for Origin-Destination Estimation." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Kamargianni, M. and A. Polydoropoulou (2014). "Social Networking Effect on Next Generation's Trip Making Behavior. Findings from a Latent Class Model." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Kelly, P., P. Krenn, S. Titze, P. Stopher, and C. Foster (2013). "Quantifying the Difference between Self-Reported and Global Positioning Systems-Measured Journey Durations: A Systematic Review." *Transport Reviews*, Vol. 33, No. 4, pp. 443–459.
- Kieu, L., A. Bhaskar, and E. Chung (2014). "Transit Passenger Segmentation Using Travel Regularity Mined from Smart Card Transactions Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Kuppam, A., J. Lemp, D. Beagan, V. Livshits, L. Vallabhaneni, and S. Nippani (2014). "Development of a Tour-Based Truck Travel Demand Model Using Truck GPS Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Lin, M., W. J. Hsu, and Z. Q. Lee (2013). "Detecting Modes of Transport from Unlabelled Positioning Sensor Data." *Journal of Location Based Services*, Vol. 7, No. 4, pp. 272–290.
- Liu, H. X., A. Danczyk, R. Brewer, and R. Starr (2008). "Evaluation of Cell Phone Traffic Data in Minnesota." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2086, pp. 1–7.
- Lu, Y., and Li, S. (2014). "An Empirical Study of With-in Day OD Prediction Using Taxi GPS Data in Singapore." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Malinovskiy, Y., N. Saunier, and Y. Wang (2012). "Analysis of Pedestrian Travel with Static Bluetooth Sensors." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2299, pp. 137–149.
- Maruyama, T., S. Mizokami, and E. Hato (2014). "A Smartphone-Based Travel Survey Trial Conducted in Kumamoto, Japan: An Examination of Voluntary Participants' Attributes." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.

- Moghaddam, S. S. and B. Hellinga (2014). "Real-Time Prediction of Arterial Roadway Travel Times Using Data Collected by Bluetooth Detectors." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Montini, L., N. Rieser-Schüssler, A. Horni, and K. W. Axhausen (2014). "Trip Purpose Identification from GPS Tracks." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Nitsche, P., P. Widhalm, S. Breuss, and P. Maurer (2012). "A Strategy on How to Utilize Smartphones for Automatically Reconstructing Trips in Travel Surveys." *Procedia-Social and Behavioral Sciences*, Vol. 48, pp. 1033–1046.
- Omrani, R., P. Izadpanah, G. Nikolic, B. Hellinga, A. Hadayeghi, and H. Abdelgawad (2013). "Evaluation of Wide-Area Traffic Monitoring Technologies for Travel Time Studies." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2380, pp. 108–119.
- Palm, M. and B. Knudson (2014). "Differences in Cell Phone Only Households in Large versus Small Metropolitan Areas: Evidence from Oregon." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Patire, A. D., M. Wright, B. Prodhomme, and A. M. Bayen (2014). "How Much GPS Data Do We Need?" *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Phithakkitnukoon, S., T. Horanont, G. Di Lorenzo, R. Shibasaki, and C. Ratti (2010). "Activity-Aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data." *Human Behavior Understanding*, Springer Berlin Heidelberg, pp. 14–25.
- Piccoli, B., K. Han, T. L. Friesz, and T. Yao (2014). "Second Order Models and Traffic Data from Mobile Sensors." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Reiff, R. M. (2012). "Determination of Origin-Destination Using Bluetooth Technology." Paper presented at the ITE 2012 Annual Meeting and Exhibit, Atlanta, GA.
- Riegel, L. and J. Attanucci (2014). "Utilizing Automatically Collected Smart Card Data to Enhance Travel Demand Surveys." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Saeedi, A., S. Park, D. S. Kim, and J. D. Porter (2013). "Improving Accuracy and Precision of Travel Time Samples Collected at Signalized Arterial Roads with Bluetooth Sensors." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2380, pp. 90–98.
- Safi, H., M. Mesbah, and L. Ferreira (2013). "ATLAS Project-Developing a Mobile-Based Travel Survey." *36th Australasian Transport Research Forum (2013 ATRF)*, Australian Government, Department of Infrastructure and Regional Development, pp. 1–15.

- Schlaich, J., T. Otterstätter, and M. Friedrich, M. (2010). "Generating Trajectories from Mobile Phone Data." *Proceedings of the 89th Annual Meeting Compendium of Papers*, Transportation Research Board of the National Academies.
- Schuessler, N. and K. W. Axhausen (2009). "Processing Raw Data from Global Positioning Systems without Additional Information." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2105, pp. 28–36.
- Spissu, E., I. Meloni, and B. Sanjust (2011). "Behavioral Analysis of Choice of Daily Route with Data from Global Positioning System." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2230, pp. 96–103.
- Spurr, T., R. Chapleau, and D. Piché (2014). "Discovery and Partial Correction of Travel Survey Bias Using Subway Smart Card Transactions." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Steenbruggen, J., M. T. Borzacchiello, P. Nijkamp, and H. Scholten (2013). "Mobile Phone Data from GSM Networks for Traffic Parameter and Urban Spatial Pattern Assessment: A Review of Applications and Opportunities." *GeoJournal*, Vol. 78, No. 2, pp. 223–243.
- Stopher, P., C. FitzGerald, and M. Xu (2007). "Assessing the Accuracy of the Sydney Household Travel Survey with GPS." *Transportation*, Vol. 34, No. 6, pp. 723–741.
- Stopher, P. R., and S. P. Greaves (2007). "Household Travel Surveys: Where Are We Going?" *Transportation Research Part A: Policy and Practice*, Vol. 41, No. 5, pp. 367–381.
- Sun, Z. and X. J. Ban (2013). "Vehicle Classification Using GPS Data." *Transportation Research Part C: Emerging Technologies*, Vol. 37, pp. 102–117.
- University of Maryland. 1997. Final Evaluation Report for the CAPITAL-ITS Operational Test and Demonstration Program, College Park, Maryland, University of Maryland Transportation Studies Center.
- Wang, T., C. Chen, and J. Ma. (2014). "Mobile Phone Data as an Alternative Data Source for Travel Behavior Studies." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Wolf, J. (2004). "Applications of New Technologies in Travel Surveys." Presented at the 7th International Conference on Travel Survey Methods, Costa Rica.
- Xia, J., W. Rao, Y.-J. Wu, and L. Guo (2014). "Dynamic OD Estimation for Urban Road Network Using Multi-sensor Traffic Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.
- Yang, F., P. J. Jin, X. Wan, R. Li, and B. Ran (2014). "Dynamic Origin-Destination Travel Demand Estimation Using Location Based Social Networking Data." *2014 TRB Annual Meeting Compendium of Papers*, Washington, D.C.