2012

The Use of Cell Phone Network Data in Traffic Data Collection and Long-Haul Truckshed (Geographic Extent) Tracking

Ming-Heng Wang
University of Kansas

Steven D. Schrock
University of Kansas, schrock@ku.edu

Nate Vander Broek
University of Kansas

Thomas Mulinazzi
University of Kansas

MID-AMERICA TRANSPORTATION CENTER

Report # MATC-KU: 361

Final Report

25-1121-0001-361

The Use of Cell Phone Network Data in Traffic Data Collection and Long-Haul Truckshed (Geographic Extent) Tracking

Ming-Heng Wang, Ph.D.
Postdoctoral Research Associate
Transportation Research Institute
University of Kansas

Steven D. Schrock, Ph.D., P.E.
Associate Professor

Nate Vander Broek
Graduate Research Assistant

Thomas Mulinazzi, Ph.D., P.E., L.S.
Professor

2012

A Cooperative Research Project sponsored by the U.S. Department of Transportation Research and Innovative Technology Administration

The Use of Cell Phone Network Data in Traffic Data Collection and Long-Haul Truckshed (Geographic Extent) Tracking

Ming-Heng Wang, Ph.D.
Postdoctoral Research Associate
Transportation Research Institute
University of Kansas

Steven D. Schrock, Ph.D., P.E.
Associate Professor
Civil, Environmental, and Architectural Engineering
University of Kansas

Nate Vander Broek
Graduate Research Assistant
Transportation Research Institute
University of Kansas

Thomas Mulinazzi, Ph.D., P.E., L.S.
Professor
Civil, Environmental, and Architectural Engineering
University of Kansas

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska-Lincoln

December 2012

Technical Report Documentation Page

1. Report No. 25-1121-0001-361	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Use of Cell Phone Network Data in Traffic Data Collection and Long-Haul Truckshed (Geographic Extent) Tracking		5. Report Date December 2012 6. Performing Organization Code
7. Author(s) Ming-Heng Wang, Steven D. Schrock, Nate Vander Broek, and Thomas Mulinazzi		8. Performing Organization Report No. 25-1121-0001-361
9. Performing Organization Name and Address Mid-America Transportation Center 2200 Vine St. PO Box 830851 Lincoln, NE 68583-0851		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Organization Name and Address Research and Innovative Technology Administration 1200 New Jersey Ave., SE Washington, D.C. 20590		13. Type of Report and Period Covered Final Report July 2009 - December 2012 14. Sponsoring Agency Code MATC TRB RiP No. 17135
15. Supplementary Notes Project Title: Improving the Long-Range Tracking Algorithm for use in Tracking Long-Haul Trucks using Cellular Data		
16. Abstract This study analyzed the potential of cell phone positioning techniques in freight truck data collection and long-haul truckshed (geographic extent) tracking. Freight truck identification and tracking algorithms were developed by means of cell phone network data and the established freight truck analysis GIS, to recognize freight trucks and determine their geographic extent (trucksheds). A case study was conducted to illustrate the truckshed tracking process and verify the tracking results from the cell phone network. Cell phones leaving from the test logistics distribution center were tracked and classified based on the developed tracking algorithms. The case study also demonstrated the processes of determining the geographic extent and traffic impact on the transportation network from the test logistics distribution center. The results showed that the proposed tracking algorithms can identify a similar percentage of freight truck data from the test facility compared with manual counts. The analysis of geographic extent indicated that 60 percent of freight traffic stayed within 30 miles of the facility, and approximately 20 percent of the traffic was considered long-haul freight traffic traveling more than 80 miles away. The long-haul tracking results found that most of the long-haul trucks returned to the original test facility during the same day. It is recommended that the tracking algorithms and data analysis process could also be applied to any other freight trucking terminal or intermodal transportation facility, as long as the cell phone network data are available. A complete freight GIS analysis network around the study area is also recommended to understand the likely destinations.		
17. Key Words Cellular Phone, Freight, Intermodal, Truckshed		18. Distribution Statement No restrictions.

19. Security Classification (of this report) Unclassified.	20. Security Classification (of this page) Unclassified.	21. No. of Pages 27	22. Price

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Geographic Information System (GIS)
Identification Number (ID)
Mid -America Transportation Center (MATC)
Origin-Destination (O-D)

Acknowledgements

The authors would like to thank AirSage, Inc. for technical assistance and providing the raw data for this research. This study was funded by Mid-America Transportation Center at the University of Nebraska.

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. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Abstract

This study analyzed the potential of cell phone positioning techniques in freight truck data collection and long-haul truckshed (geographic extent) tracking. Freight truck identification and tracking algorithms were developed by means of cell phone network data and the established freight truck analysis GIS, to recognize freight trucks and determine their geographic extent (trucksheds). A case study was conducted to illustrate the truckshed tracking process and verify the tracking results from the cell phone network. Cell phones leaving from the test logistics distribution center were tracked and classified based on the developed tracking algorithms. The case study also demonstrated the processes of determining the geographic extent and traffic impact on the transportation network from the test logistics distribution center. The results showed that the proposed tracking algorithms can identify a similar percentage of freight truck data from the test facility compared with manual counts. The analysis of geographic extent indicated that 60 percent of freight traffic stayed within 30 miles of the facility, and approximately 20 percent of the traffic was considered long-haul freight traffic traveling more than 80 miles away. The long-haul tracking results found that most of the long-haul trucks returned to the original test facility during the same day. It is recommended that the tracking algorithms and data analysis process could also be applied to any other freight trucking terminal or intermodal transportation facility, as long as the cell phone network data are available. A complete freight GIS analysis network around the study area is also recommended to understand the likely destinations.

Chapter 1 Introduction

The widening supply chain has increased demand for freight truck movement across the country. Freight truck bottlenecks are evident in all modes of transportation. As logistics facilities, distribution centers, and intermodal facilities spread throughout the region and the nation, a major issue is the impact of freight traffic from those facilities on traffic in the local and regional transportation networks. The truckshed—defined as the geographic extent of freight traffic based on the normal use pattern of freight movements within a region (1)—is an important step in understanding traffic impacts and any potentially negative consequences from the intermodal facilities. The truckshed data is also a critical input in the support of freight demand analysis and planning.

There are many methods and techniques which can be used to collect freight data and identify the geographic extent of freight from logistics and intermodal facilities. The Quick Response Freight Manual II (1) provides detailed methods such as vehicle classification counts, roadside intercept surveys, and travel diary surveys, to collect freight truck data. Traditional origin-destination (O-D) methods, such as roadside, facility gateway, and freight truck driver surveys, may be performed but are time consuming. Innovative technologies have become available that can reflect the traffic condition and also catch O-D data by tracking vehicle trajectories and destinations. Among these new technologies, cell phone tracking technology has the potential to track vehicle location and movement, and has received strong interest from the transportation community. It can also be used for updating past traffic data and reducing the frequency of exhaustive data collection efforts. A number of researchers (2 – 4) have evaluated the application of cell phones in travel speed and travel time estimation, as well as the identification of congestion sections. Some researchers have also used cell phone data to derive

the O-D matrix (5,6). However, the application of cell phone data to the collection of freight data and tracking of the geographic extent of logistics or intermodal facilities has not been reported in the literature.

The objective of this study was to demonstrate the process of using cell telephone location data in freight data collection, and to determine the geographic extent of freight traffic from the logistics and intermodal facilities. Freight truck identification and tracking algorithms were developed to recognize freight truck data from anonymous cell phone position data. A nearby logistics distribution center was selected for comparing ground count real freight truck data to the freight truck data estimated by the proposed algorithm and cell data. Traffic impact from the logistics facility was also illustrated based on cell phone trajectory data.

Chapter 2 Cell Phone Network Data

2.1 Cell Phone Positioning Technology

Cell phone positioning technologies can generally be divided into two categories: network-based and handset-based (GPS-enabled). Due to data availability, this study focused on the system that uses network-based technology. Network-based systems utilize signal information from cell phones to derive their locations. In such systems, one or several base stations (signal towers) are involved in locating a cell phone. All required measurements are conducted at the base stations, and the measurement results are sent to a location center, where a position is calculated. There is no requirement to make any changes to the current handsets. However, the cell phone position can be generated only when the cell phone user is actively talking or texting (both sending and receiving text messages) on the phone, or when the phone is connected to the Internet. Thus, only a partial number of cell phone users can be detected during an observation period. A long-term observation could increase the sample size of cell phone location data.

The accuracy of location estimation for network-based cell phone positioning technology is about 50-300 meters (7). Several systems have been tested for accuracy in location tracking and travel speed measurement. The results have shown that most systems can locate cell phones to within 200-330 meters (8,9). However, the accuracy of estimating travel speeds has varied (10–13). Some systems have proved to be feasible in determining travel time only under the normal conditions of free traffic flow, while accuracy decreased rapidly as congestion increased.

2.2 Cell Phone Data Structure

The cell phone data were provided by AirSage Inc., a provider of location, movement, and real-time traffic information based on cell phone signaling data (14). Their data were comprised from one cell phone carrier, constituting about 18 percent of the total cell phone subscribers in the United States. Penetration of the cell phone data may vary by location, however. According to internal and independent testing (*14*), there is an average uncertainty radius of 320 meters, and a median of 220 meters; moreover, at some peak usage periods, additional locational error may be introduced when users are automatically transferred by the network from the nearest cell phone signal tower to one farther away but less heavily-loaded. The cell phone data provided by the vendor for this study covered four Midwestern states: Missouri, Iowa, Nebraska, and Kansas (MINK).

The provided cell phone data included a unique encrypted identification number (ID), location information, and the times when the cell phones were actively being used. The location information was shown as longitude and latitude. The encrypted ID number contained more than 30 digits, and was a unique string for every individual cell phone. The epoch time was recorded in the UNIX timestamp format. The data were provided on a 30-second base. For every 30 seconds, all location data of cell phones actively being used for calls or typing a message within the network were recorded in the database. The travel time and Euclidean distance between two adjacent location data points were also computed based on the provided data.

2.3 Cellular Phone Data Availability

The available cellular phone data included only those individuals who carried a cell phone and made a call, texted, or were browsing the Internet during the observation period. Thus, only a partial number of drivers could be detected in the cellular phone database. Because the vender's cell phone coverage only contained a portion of the total cell phone subscribers in the United States, those people who did not carry a cell phone or whose cell phones were not part of the vender's system were not included in the cell phone data.

Chapter 3 Methodology and Tracking Algorithms

Due to security issues and the need to protect the privacy of cell phone users, all cell phone data were encrypted. Thus, it was infeasible to identify the attributes and classification (freight truck or non-freight truck users) of the cell data through the encrypted cell ID. The proposed methodology for classifying cell phones was to track the trajectories of the cell phones in the transportation network. The classification of cell phones depends on the movement status and the land use of their trajectories and destinations. Thus, the cell phone location data were mapped out in a freight analysis network using a geographic information system (GIS) to plot the trajectories of the cell phones in the transportation network. The freight truck tracking process and algorithms were developed based on the established freight network.

3.1 Freight Transportation Analysis GIS Network

A freight transportation analysis network was established using ArcGIS software on the basis of the National Transportation Atlas Database 2010 (15). The network converted the U.S. state boundaries, freight analysis network, intermodal facility, and railway network in Region 7, covering the states in the MINK region. The freight truck stop location database in the analysis network was developed based on an Internet search using Google™ Maps, American Trucker (16), and Truck Master Fuel Finder (17). The potential freight truck stops included warehouses, gas stations, rest areas, and weigh stations. The truck stop database was used as the criteria to determine and extract the freight traffic from the cell phone network data. A total of 800 freight truck stops were obtained in the MINK region.

3.2 Cell Phone and Truck Tracking Algorithms

The cell phone tracking algorithm modules and the operational process are shown in figure 3.1. The first four modules are data formation modules, which filter and generate the tracking databases including cell phone location, trajectory, movement states, and the origin and destination for every identified trip. Module 6 is the application module developed for the study to identify travel patterns and determine whether tracked traffic are long-haul trucks based on their trajectory and their geographic locations in relation to the trucking facilities.

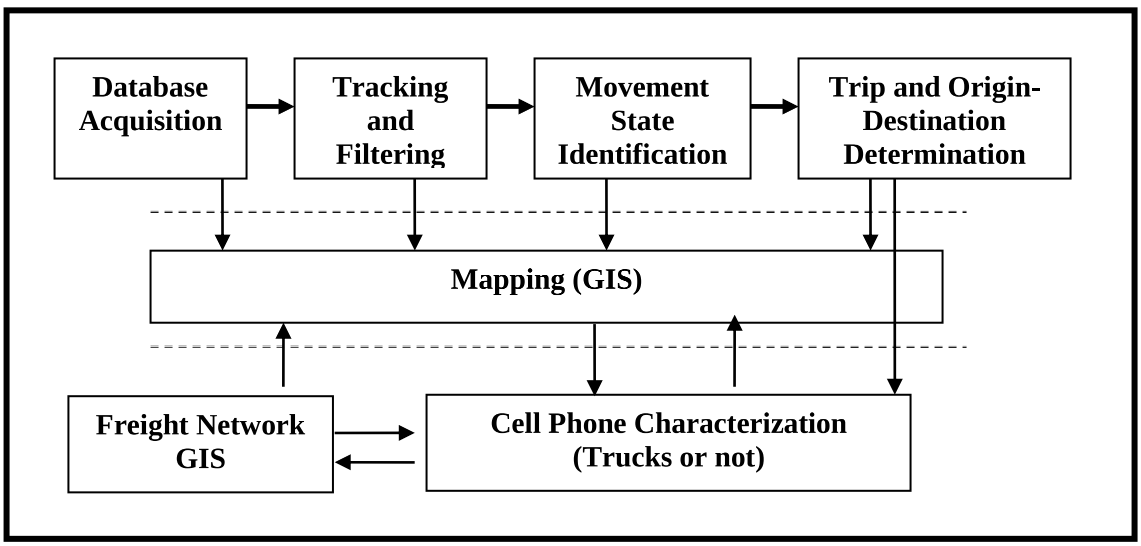

Figure 3.1 Cell phone tracking algorithm modules and the operational process

3.3 Module 1: Data Acquisition and Localization

The purpose of the first module is to establish the initial database for tracking and filtering processes. Cell phone network data provided from the vendor are coded and transferred into the database format. Every individual piece of data represents a geographic position point based on its longitude and latitude. In most cases, a single call could generate multiple data points because the database recorded the cell phone positions on a 30-second interval base. In

this module, a location or region code is assigned to every data point. The region codes vary depending on the demands of the study. The region codes could be ZIP codes, county or city codes, specific codes for facilities or landmarks, or a road section or small street block for travel time estimation. Thus, the parameters for every data point in the database include: an encrypted identification number (*ID*); longitude and latitude, in decimal format; a timestamp, in epoch format; and the given region code. Depending on study demands, the cell phone database size could be reduced by trimming off any unnecessary data points not related to the study from the entire network data.

3.4 Module 2: Data Tracking and Filtering

The purpose of this module is to ascertain the trajectory data and other operational parameters for every cell phone, and also to eliminate any erroneous locational data. The database established in Module 1 is sorted by encrypted ID and timestamp. The time difference, distance, and speed parameters related to the previous data point are added to the database. Note that the distance between two consecutive points is the Euclidean distance. The speed data between consecutive points are not intended to reflect the travel speed but are instead used for data filtering. Due to the cell phone signal transition process, there are some errors in the raw cell phone position data. Those position errors will lead to an irrational distance and speed between consecutive points. In order to filter those erroneous points, a small trajectory data within five minutes are inspected. Any points which have irrational travel distance or travel speed (more than 120 miles per hour) are removed from the database. The trajectory and all other parameters are re-calculated for the remaining data in the database. Other parameters, such as distance or travel time from the given regions, can also be added depending on the needs of the study.

3.5 Module 3: Movement State Identification

The purpose of this module is to identify whether the cell phones are in *moving* or *non-moving* condition, not along the entire trajectories, but merely within the given regions or facilities. This module will determine if the cell phones are the passing-by traffic or non-moving users in the given regions. The judgment criterion is based on dwell time, which is the duration of the cell phone users staying in the given region. If the dwell time at the same given region is longer than the threshold time (T_c), which is approximately equal to the regular travel time between the boundaries of the given region, the cell phones are considered *non-movements*. If the dwell time is less than T_c, the cell phones are in *moving* conditions and are considered passing-by traffic.

Figure 3.2 shows the time and distance distributions from a set of consecutive points for a cell phone related to a given region. In this example, Point 1 is the last point that appears at the previous region before entering the target region, Point 2 is the first point after entering the region, Point 3 is the last point that appears at the given region, and Point 4 is the first point in the next region after leaving the target region.

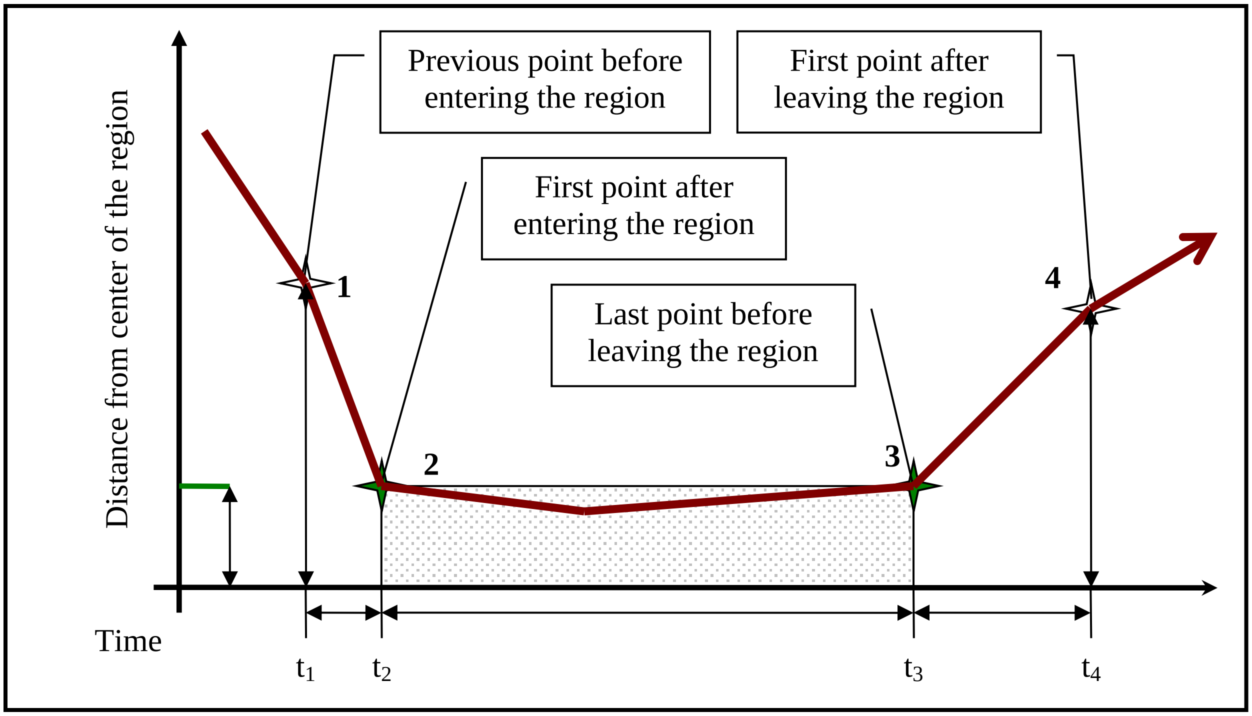

Figure 3.2 Time and distance distributions of cell phone points related to a given region

The dwell time can be obtained simply by estimating the time difference ($\Delta t_{2,3}$) between the first and the last consecutive cell phone points appearing at the given region. However, due to the availability of the cell phone data, the first point appearing in the given region may not be the actual time that the cell phone user entered the region, nor the last point that the user entered the region. Particularly when $\Delta t_{2,3}$ is less than (T_c), the time difference ($\Delta t_{4,1}$) between Point 1 and Point 4 should be taken into account to estimate the dwell time. The regular travel time (T_E) between Point 1 and Point 4 is the other judgment criteria. The dwell time should not be less than (T_c) but less than $\Delta t_{4,1} - T_E + Tc_r$. The expected dwell time (DW) is estimated as:

$$E(DW) = p \cdot T_c + (1 - p) \cdot (\Delta t_{4,1} - T_E + T_c) \qquad (3.1)$$

where,

$$p = \frac{T_c}{\Delta t_{4,1} - T_E + T_c}, \text{ and } \Delta t_{4,1} - T_E \geq 0.$$

If $\Delta t_{4,1} - T_E < 0$, then the dwell time would be less than T_c. The cell phone is included in the passing-by traffic. The T_E can be obtained by simply dividing the distance ($\Delta D_{4,1}$) between point 1 and point 4 by the average speed limit (S_{Limit}) of roadways around the given region.

3.6 Module 4: Trip and Origin-Destination Determination

The purpose of the fourth module is to determine trips and the origin-destination regions based on the movement states of cell phones among all given regions. The other threshold value (T_k) is set to identify whether the non-movement cell phones are temporarily stopping or users staying for a long time at the given region. If the estimated dwell times are greater than T_k, the cell phones are long-time staying users at the region. The origin or destination regions are the regions where the phone user stays for more than T_k. A new trip is generated once the cell phone user heads to other regions after a long dwell time at the identified origin region. As a result, the trip database is created by including the sequence number of trip, origin and destination region codes, last point (longitude and latitude) and time at the origin, and first point and time at the destination. The dwell time at the origin region and the travel time between the last point at the origin and first point at the destination are also calculated and added into the database.

The T_k value area varies with given regions and depends upon the study subjective. For example, if a resident or business area is the given region, the T_k may be set to be more than four

or six hours to assure that cell phone users are residents or workers within the region. In this example, the origin region was the logistics distribution center.

3.7 Module 5: Mapping

The order of the mapping module does not necessarily follow the other modules. It can be processed in any module above, as long as the location data points are available to identify any possible invalid points and demonstrate the spatial distribution of the data points. All the tracked data points, trajectories, trips, and O-D data can be converted into Google Earth or GISs to display their geographic locations. The tracking map can also illustrate the spatial distribution of the cell phone points and determine the geographic extent or service area from certain targets, such as the intermodal transportation facility or terminal. Trajectory data converted into a GIS can further be used to identify characteristics by analyzing the temporal distributions and their related locations or interactions to given objects or facilities for the cell phones in the network.

3.8 Module 6: Cell Phone Characterization

As mentioned earlier, because of privacy rights and security issues, all of the cell data were encrypted. Thus, it was infeasible to obtain the attribute and classification (traveler types) of the cell phone users through the encrypted ID. However, the classification and the traveler characteristics can still possibly be identified by observing cell phone users' trajectory, origin and destination information. In this study, the tracking algorithms were used to determine whether cell phone users were freight or non-freight drivers based on the trajectories of the cell phones leaving from the origin facility. In this module, all the cell phone points were converted into an established freight stops GIS containing possible freight stops or destinations such as truck stops, rest areas, warehouses, or other logistic facilities. Based on the trajectories of the cell phones, if a cell phone point has appeared around the freight stops, the cell phone user could

be considered a freight truck driver. However, Module 3 should be applied to check the movement status after the cell phone reached the stop, i.e., whether or not they are "passing-by" traffic. If the cell phone stays at a stop for a long period of time, the cell user could be considered an employee of or a resident near the stop. Similarly, if the cell phone points appear within the stops in a very short time, the cell phones are not considered as freight traffic, but instead could be passing-by traffic through the truck stops.

In addition, the freight types "long-haul" or "nearby shipment" are also determined based on the trajectory of the cell phone data points. If the cell phone data points routinely appear between some nearby warehouses and the intermodal facilities, the cell phone user is considered as a nearby shipment freight. If the freight travels a long distance (for instance, more than 80 miles from the intermodal facilities), it will be considered long-haul freight traffic. Based on the trajectory of these cell phone data, the geographic extent, service area and the impact on the traffic network by the facility also can be determined.

3.8.1 Tracking Process and Analysis Methods

Cell phone data and the developed tracking and freight identification algorithms were processed by a data management and statistical analysis program, SAS (18). The cell phone data that had appeared within the study facility were located and tracked as they dispatched from the study facility. The tracked cell phone data were integrated into the established freight GIS analysis network, which included major highways and freight truck stops. The Tracking Analyst function included in the ArcGIS program was used to illustrate the trajectory of every tracked individual cell phone. It was also used to analyze the travel patterns for the identified freight trucks, particularly for long-haul freight trucking.

The Tracking Analyst function is an extension of the ArcGIS program that enables users to create time series visualizations to analyze information relative to time and location. From a collection of random data, ArcGIS's Tracking Analyst creates a visible path, or track, showing movement through space and time of the analyzed phenomena. Through the Tracking Analyst, users are able to view complex time series and spatial patterns in an integrated ArcGIS system.

Chapter 4 Case Study and Analysis

A case study was conducted in a selected intermodal facility to illustrate the process of using cell phone network data in freight truck data collection and in determining the geographic extent of freight trucks from the test facility. Ground traffic counts were collected and compared to cell phone tracking results. The spatial distribution and traffic impact analysis of the freight trucks from the facility on the transportation network were also demonstrated.

4.1 Study Site and Area

A logistic distribution center with simple traffic components was selected for the field test. The logistics and distribution center (Netherlands Park) was located at Missouri Route 210 between Interstate 435 and Missouri Route 291, as shown in figure 4.1. The center consists of several logistics companies, such as TNT, CEVA, and Delivery Logistics, a FedEx Less Than Truckload (LTL) distribution center, and a Norfolk Southern Railroad operation center. All the traffic from the logistics companies use only two streets (Kimball Dr. and N. Norfleet Rd.) to connect to Missouri Route 210. The speed limit in the section of Missouri Route 210 is 55 mph (88.5 km/hr). The distance between these two intersections is 1 mile (1.6 km). The area range of the facility is about 1.5 miles (2.4 km) east-west by 2.5 miles (4.2 km) north-south.

Figure 4.1 shows the location of the study facility and area, along with sample points of four cell phones leaving the facility along the major highways. Cell phones which had appeared around the test facility were tracked as long as the data points showed up within the four-state MINK region. The threshold value (T_c) in Module 3, used for judging the passing-by traffic, was set at five minutes, which is approximately the travel time crossing the facility.

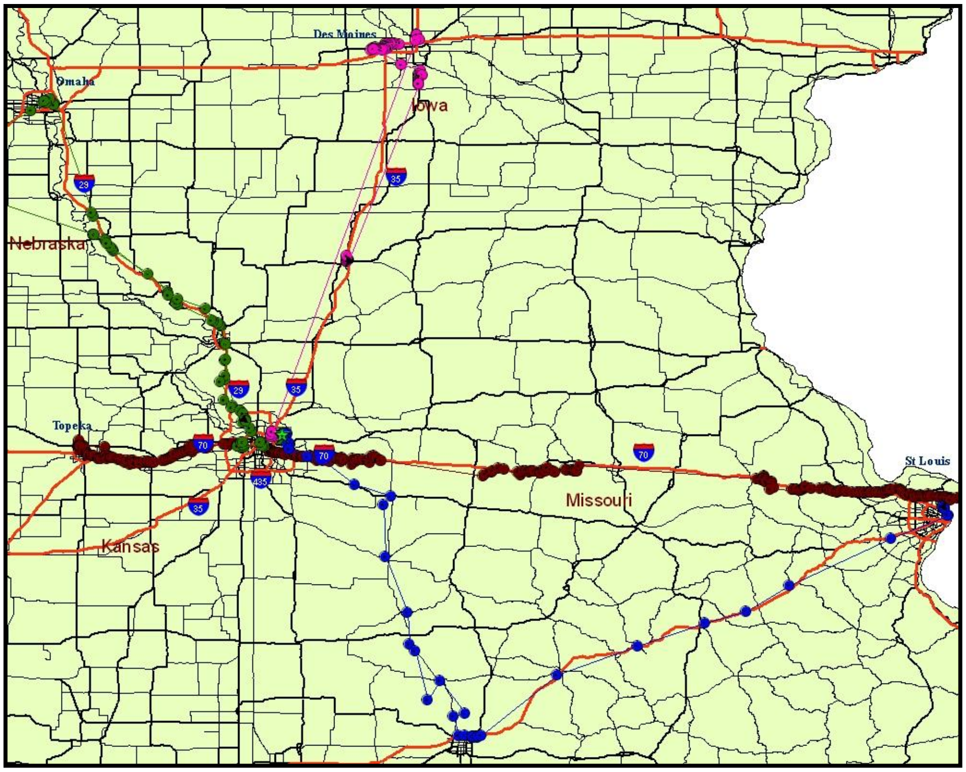

Figure 4.1 Samples of cell phone points from the distribution center

4.2 Data Collection and Comparison

The ground count traffic data from the test facility were collected from 9:00 a.m. to 4:00 p.m. on May 19th, 2010. The cell phone data appearing within the facility on May 19th were tracked after leaving the facility throughout the day until 11:59 p.m. However, only those cell phones which left the test facility from 9:00 a.m. to 4:00 p.m. were used for the comparison analysis. In this study, the time of the last point appearing within the facility (which is t_3 in figure 3.2) was used to represent the time of the cell phone leaving the facility. These data were processed using the SAS program, and were converted into the established freight analysis GIS network to illustrate the corresponding locations of the cell data and geographic information.

The classification of the cell phones was determined by the proposed tracking algorithms. The T_k was set as four hours to determine whether the cell phone users were traffic or non-moving phone users in the facility. The ground count traffic (collected at 30-minute intervals), as well as traffic recognized by cell phones, are shown below in table 4.1.

Table 4.1 Traffic data obtained by ground count and from cell phone data

Time	Ground Count Traffic (From the facility)				CellularPhone Tracked Traffic				% of Traffic Tracked by Cellular Phone
	Trucks	Cars	Total	% of Trucks	Trucks	Cars	Total	% of Trucks	
9:00-9:30	49	29	78	62.8%	7	3	10	70.0%	12.8%
9:30-10:00	37	33	70	52.9%	5	7	12	41.7%	17.1%
10:00-10:30	45	36	81	55.6%	9	10	19	47.4%	23.5%
10:30-11:00	43	42	85	50.6%	6	6	12	50.0%	14.1%
11:00-11:30	26	53	79	32.9%	8	4	12	66.7%	15.2%
11:30-12:00	33	56	89	37.1%	6	8	14	42.9%	15.7%
12:00-12:30	49	76	125	39.2%	7	14	21	33.3%	16.8%
12:30-13:00	46	63	109	42.2%	8	13	21	38.1%	19.3%
13:00-13:30	37	53	90	41.1%	8	6	14	57.1%	15.6%
13:30-14:00	43	37	80	53.8%	5	7	12	41.7%	15.0%
14:00-14:30	42	49	91	46.2%	8	6	14	57.1%	15.4%
14:30-15:00	29	52	81	35.8%	6	3	9	66.7%	11.1%
15:00-15:30	30	45	75	40.0%	4	5	9	44.4%	12.0%
15:30-16:00	53	72	125	42.4%	6	14	20	30.0%	16.0%
Total	562	696	1258	44.7%	93	106	199	46.7%	15.8%

Because of the availability of cell phone network data, the cell phone network data could only reflect part of the traffic data. The results show that, in the study area, about 16 percent of total traffic from the test facility were tracked and recognized using the cell phone location data. This percentage is close to the available cell phone coverage (17 percent) in the study area. Figure 4.2 shows the relationship between total ground traffic counted by a 30-minute interval

and traffic recognized from cell phone data. The linear regression model of the relationship is shown below with R^2 = 0.6731:

$$Traffic_{Ground} = 3.3426 \cdot Traffic_{Cellular} + 42.344 \tag{4.1}$$

The regression model is a calibration equation which can be used to estimate the total number of vehicles leaving from the facility when the number of vehicles is recognized from the cell phone data. It should be noted that the model may be varied and dependent on the coverage of the cell data around the study area.

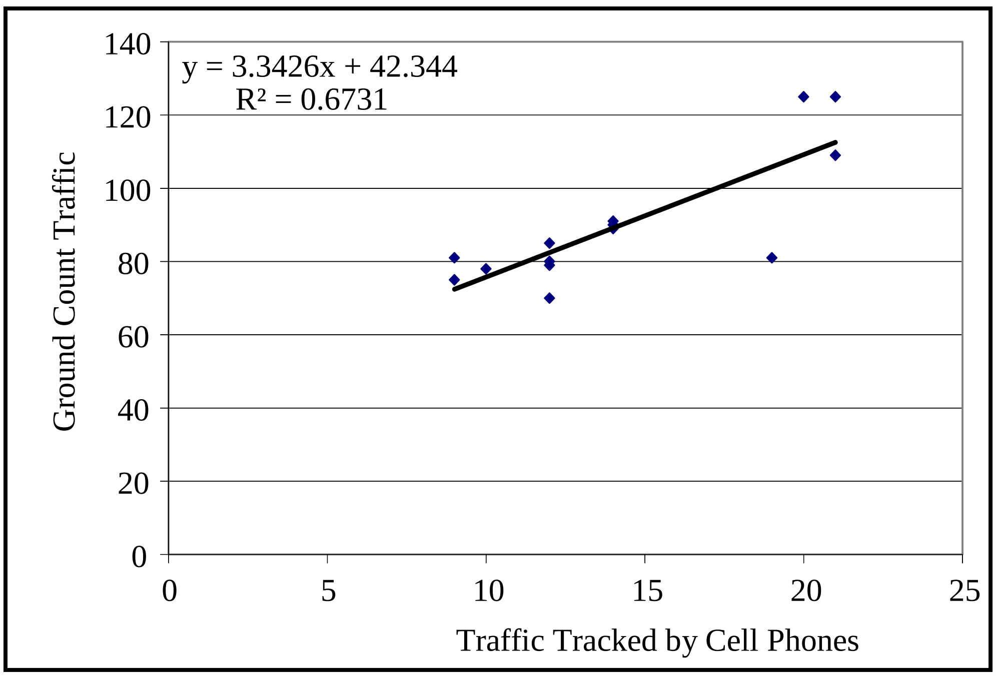

Figure 4.2 Relationships between ground count traffic and traffic tracked from cell phone data

Regarding the freight truck volumes, about 45 percent of total vehicles leaving from the test facility were freight trucks. Using the cell phone location data, about 47 percent of total vehicles were identified as freight trucks. The Chi-square test was used to test whether the proposed cell phone tracking and freight truck identification algorithms could recognize a similar proportion of freight trucks obtained by ground count. The Chi-square test results showed that the difference of the proportions was not significant at the 0.05 level of significance (χ^2 =0.217 $<\chi^2_{0.95,1}$ =3.84), indicating that using the cell phone location data and proposed freight truck tracking algorithms can provide a similar proportion of real freight truck data from the test facility.

4.3 Determining the Freight Truck Geographic Extent

Understanding the geographic extent of the reach of the transportation intermodal facility is the first step in realizing the traffic impact of the facility and freight traffic on the traffic network. Once the freight truck data were identified from the cell phone data, the geographic extent of the freight traffic could be determined by tracking the trajectory of individual freight trucks.

In this case study, the farthest destination of every identified freight truck leaving the facility between 9 a.m. and 4 p.m. was obtained to represent the geographic extent of freight trucks. The data were sorted to derive the spatial distribution of the freight data. Figure 4.3 shows the spatial distribution of freight traffic spreading out from the facility by 25, 50, 75 and 100 percent in brown, blue, purple and green lines, respectively. It also represents the service area of freight from the facility.

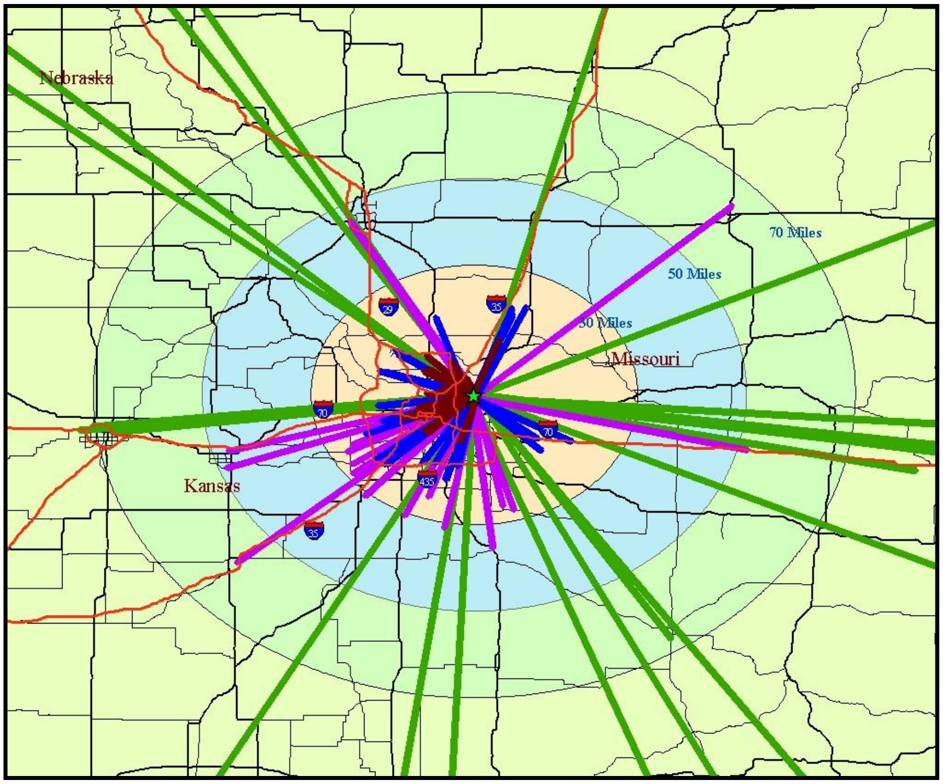

Figure 4.3 Spatial distribution of freight traffic spreading out from the facility

It was observed that about 25 percent of the freight service area was within 16 miles of the test facility, while 50 percent was within 23 miles, and 75 percent was within 65 miles. Regarding the farthest service area, about 40 percent of freight trucks served within 20 miles of the facility. About 60 percent of freight served within 30 miles of the facility around the Kansas City metropolitan area. About 80 percent stayed within 80 miles, dispersing to nearby cites. If traveling more than 80 miles is considered long-haul trucking, it means that only about 20 percent of the freight trucks were long-haul freight trucks.

4.4 Long-Haul Truckshed Tracking

As the results above show, only 20 percent of identified freight trucks were long-haul trucking from the test logistics distribution center. This means that a total of 19 freight trucks identified by cell phone data were long-haul trucking from the test facility between 9:00 a.m. and 4:00 p.m. on May 15, 2010. Based on the location and direction of four study states, the five major destinations of long-haul trucking from the test facility were Des Moines (IA), Omaha and Lincoln (NE), St Louis (MO), Wichita (KS), and Joplin and Springfield (MO). The truck counts identified from cell phone data and the estimated long-haul truck volume are shown in table 4.1. The estimated truck counts were based on equation 3.1 and the percentage of trucks from the test facility. The farthest points and the trajectories of long-haul truckshed are shown in figure 4.4.

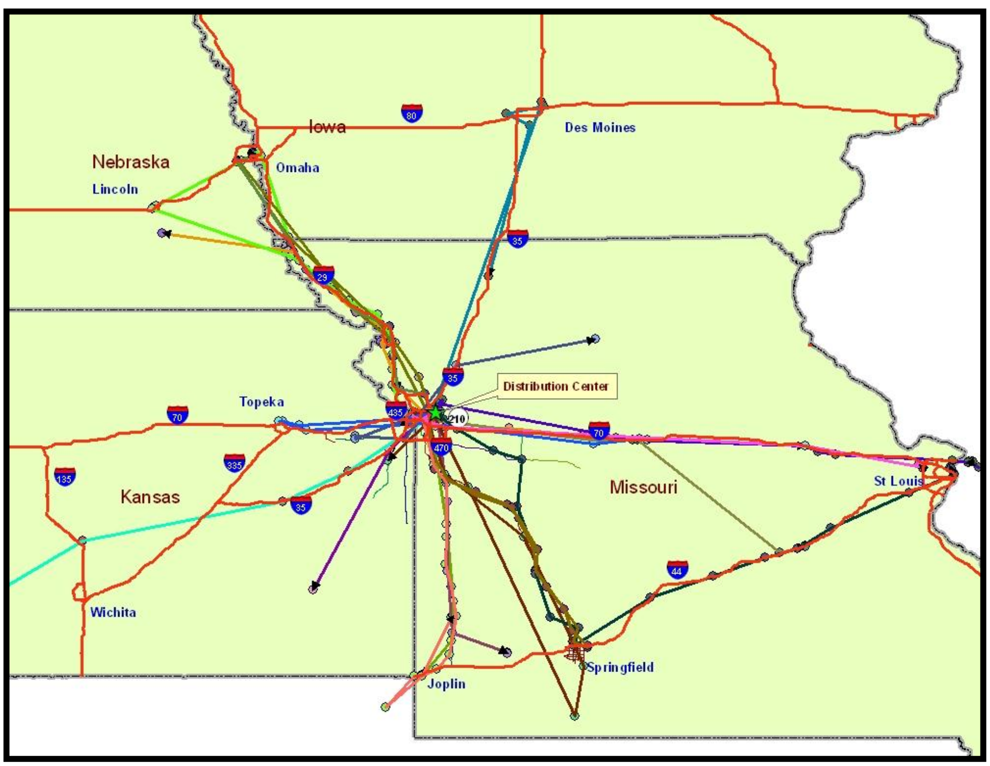

Figure 4.4 Geographic extents and trajectories of long-haul trucks

Among the long-haul trucks, about 26 percent of long-haul trucks headed east to St. Louis, 2 percent south to Springfield and Joplin, 21 percent headed northwest to Omaha and Lincoln, and 5 percent each headed north to Des Moines and southwest to Wichita. The remaining 16 percent of long-haul trucks stopped at Macon and Columbia, Missouri, and southeast of Kansas around Chanute.

It was observed that most of the long-haul trucks that headed north to Des Moines, Omaha and Lincoln (four out of five) turned back to the origin (Kansas City). The same tendency was found in the long-haul trucks heading south to the Missouri cities of Springfield and Joplin. Four out of five of trucks headed back to Kansas City in the latter part of the day. This pattern was not found among those trucks heading to St. Louis or Wichita. However, for trucks heading east to St. Louis, two out of five trucks traveled on Interstate 44 instead of using the more direct route of Interstate 70.

As a result, a significant pattern of long-haul trucking from the test logistics facility was mostly trucks heading to southern or northern cities and turning back to the origin facility on the same day. This pattern was not found in the trucks heading to the St. Louis area.

4.5 Analysis of Traffic Impact

Traffic impact on the transportation network was assessed by estimating the truck volume from the original facility on the transportation network. This can be done by tracking the trajectories and traveling routes of freight trucks from the test logistics facility. Since cell phone data may not be continuous due to the limitation of active usage, the cell phones traveling on the network segments were detected based not only on the position points location, but also through determination of the shortest route between the origin and destination. Figure 4.5 shows the calculated results of truck volumes on the transportation network around the Kansas City area

based on the cell phone data from the test logistics distribution center during the study time period.

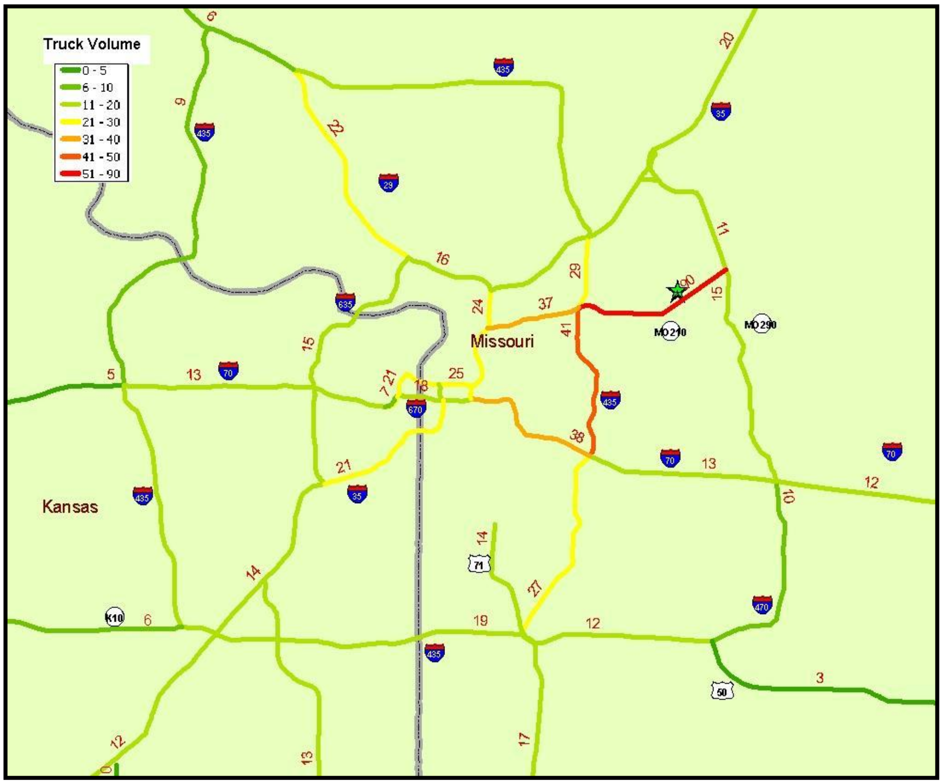

Figure 4.5 Freight truck volumes from origin facility on transportation network around Kansas City

The real traffic impact of freight truck from the origin facility on the network can be estimated by applying the percentage of trucks (47 percent) and equation 3.1 to the truck volumes shown in figure 4.5. For example, the freight truck volume on the segment of Interstate 70 between Interstate 435 and Interstate 470 is 13. The estimated freight truck volume from the origin facility traveling on this segment is approximately 64.

Chapter 5 Conclusions and Recommendations

The use of cell phone tracking technology can assist in updating past traffic data and reduce the frequency of exhaustive data collection efforts. This study analyzed the potential of cell phone positioning techniques in freight trucks data collection and long-haul truckshed (geographic extent) tracking. The freight truck identification and tracking algorithms were developed by means of cell phone network data and the established freight truck analysis GIS to recognize the freight trucks and determine their geographic extent. The case study results found that the available cell phone data covered approximately 16 percent of traffic from the test facility in the Kansas City area. It also found that the proposed tracking algorithms could identify a similar percentage of freight truck data from the facility. The case study developed a relationship model between the ground count traffic and traffic estimated by cell phone data around the area. The analysis of geographic extent indicated that 60 percent of freight traffic stayed within 30 miles of the facility, and approximately 20 percent of the traffic was considered long-haul freight traffic traveling more than 80 miles away. The long-haul tracking results found that most of the long-haul trucks turned back to the original facility during the same day.

It is recommended that the tracking algorithms and data analysis process could also be applied to any other freight trucking terminal or intermodal transportation facility, as long as the cell phone network data are available. However, the calibration factor may vary by facility locations and research areas. The other critical challenge is that the truck stop database in the freight analysis GIS network will influence the truck identification results. A complete freight GIS analysis network around the study area is also needed before processing the study. For a future study, conducting more site tests, accomplishing freight GIS databases, and investigating freight truck data directly from the industries are recommended to verify the proposed algorithms

and the application of using cell phone data in freight truck data collection.

References

1. Federal Highway Administration. 1997. *Intelligent Transportation Intermodal Freight Symposium Workbook*. Washington, D.C.: author. Accessed November 15, 2011. http://ntl.bts.gov/lib/jpodocs/proceedn/5501.pdf.

2. Beagan, D., M. Fischer, and A. Kuppam. 2007. "Quick response freight manual." *Federal Highway Administration Report. FHWA-HOP-08-010*. Washington, D.C., Federal Highway Administration.

3. Lovell, D. J. 2001. "Accuracy of speed measurements from cell phone vehicle location systems." *ITS Journal*, 6, no. 4: 303-325.

4. Smith, B. 2006. "Wireless location technology-based traffic monitoring demonstration and evaluation project - Final evaluation report." University of Virginia Smart Travel Laboratory Center for Transportation Studies, University of Virginia. Charlottesville, VA.

5. Cayford, R., and Y. B. Yim. 2006. "A field operation test using anonymous cell phone tracking for generating traffic information." *Paper no. 2511, Transportation Research Board 2006 Annual Meeting* [CD-ROM].

6. Caceres, N., J. P. Wideberg, and F. G. Benitez. 2007. "Deriving origin–destination data from a mobile phone network." *IET Intelligent Transportation Systems*, 1, no. 1: 15–26.

7. Sohn, K., and D. Kim. 2008. "Dynamic origin–destination flow estimation using cell communication system." *IEEE Transactions on Vehicular Technology*, 57, no. 5: 2703-2713.

8. Openwave Inc. *Overview of location technologies*. Accessed July 30, 2010. http://developer.openwave.com/omdtdocs/location_studio_sdk/pdf/Intro_to_Location_Technologies.pdf.

9. Zhao, Y. 2000. "Mobile phone location determination and its impact on intelligent transportation systems." *IEEE Transactions on Intelligent Transportation Systems*, 1, no. 1: 55-64.

10. Wunnava, S. V., K. Yen, T. Babij, R. Zacaleta, R. Romero, and C. Archilla. 2007. "Travel time estimation using cell phones for highway and roadways, final report. Tallahassee, FL, Florida Department of Transportation.

11. University of Maryland Transportation Studies Center. 1997. *Final evaluation report for the CAPITAL-ITS Operational Test and Demonstration Program*. University of Maryland, College Park, MD.

12. Yim, Y. B., and R. Cayford. 2001. "Investigation of vehicles as probes using global positioning system and cell phone tracking: Field operational test." *Report UCB-ITS-PWP-2001-9*, California PATH Program, Institute of Transportation Studies, University of California, Berkeley, CA.

13. Smith, B. L., H. Zhang, M. Fontaine, and M. Green. 2003. *Cellphone probes as an ATMS tool*. National ITS Implementation Research Center, Center for Transportation Studies, University of Virginia, Charlottesville, VA.

14. University of Virginia National Cooperative Highway Research Program 70-01. *Private-sector provision of congestion data: Probe-based traffic monitoring state-of-the-practice report*. University of Virginia Center for Transportation Studies, Virginia Transportation Research Council, Charlottesville, VA.

15. Airsage Inc. Accessed August 1, 2010. http://www.airsage.com/site/index.cfm.

16. Research and Innovative Technology Administration Bureau of Transportation Statistics, National Transportation Atlas Database. 2010. Accessed October 25, 2010. http://www.bts.gov/publications/national_transportation_atlas_database/2010/.

17. American Trucker, Truck Stops in Iowa/ Kansas/ Missouri/ Nebraska. Accessed January 31, 2011. http://www.trucker.com/TruckStops.aspx.

18. Truck Master Fuel Finder. Accessed January 31, 2011. http://www.findfuelstops.com/.

19. Statistical Analysis Software (SAS). http://www.sas.com/technologies/analytics/statistics/stat/index.html.