Feasibility of Using Cellular Telephone Data to Determine the Truckshed of Intermodal Facilities

Ming-Heng Wang
Transportation Research Institute
University of Kansas
1530 W. 15th Street, 2160 Learned Hall
Lawrence, KS 66045
mwang4@ku.edu

Steven D. Schrock
Department of Civil, Environmental, and Architectural Engineering
University of Kansas
1530 W. 15th Street, 2150 Learned Hall
Lawrence, KS 66045
schrock@ku.edu

Thomas Mulinazzi
Department of Civil, Environmental, and Architectural Engineering
University of Kansas
1530 W. 15th Street, 2150 Learned Hall
Lawrence, KS 66045
tomm@ku.edu

ABSTRACT

This paper analyzes the feasibility of using cellular telephone location data to determine the geographic extent of trucks from intermodal facilities. The feasibility analysis includes three aspects: location technology, cell phone penetration, and truck tracking methodology. As the results shows, the cell phone location technology is able to provide accurate location data. It can provide updated location information with no time limitation. Cell phones could be located within an average of 100 meters or less of their actual position. Although only partial cell phone data are available for the truck tracking system, a larger sample size is expected during a longer time period observation when more cell phone data are integrated. The principle and process of truck filtering and tracking from the cell phone database was proposed in this study. To extract trucks from the cell phone data, the analysis network should be categorized into several types of regions based on the land use and truck movement characters. In principle, those tracked cell phones frequently located around truck stops and traveling along the Intestate routes are considered as trucks. Further field study is needed to verify the accuracy of the truck identification and derive the calibration factor for generating truck trips using cell phone location data.

Key words: cellular telephone location—geographic extent—truckshed—intermodal facility
INTRODUCTION

In today’s world of global supply chains, the manufacturing of goods is increasingly spread out throughout the world. The widening supply chain has increased the demand for freight movement across the country. It has also created new opportunities for the railroad industry to compete with the trucking industry for long-haul operations through the development of rail-truck intermodal facilities. As intermodal facilities spread through the region and the nation, one of the major questions is the geographic extent (truckshed) that will impact the local and regional transportation network. There is a need to better understand the impacts of this truckshed on the transportation network. Thus, understanding the extent of the reach of the facility is an important first step to mitigate any negative consequences.

There are many methodologies that can be used to catch the geographic extent of trucks from intermodal facilities. Traditional origin-destination (O-D) methods, such as roadside and truck driver surveys, may be performed but are time consuming. These methods are increasingly unpopular due to the disruption they cause to traffic. 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, the cellular phone tracking technology is one of the potential methods in tracking vehicle location and movement. It has received strong interest from the transportation community. A number of researchers (Ygnace et al. 2000; Lovell 2001; Smith 2006; Cayford 2006; Liu 2008) have evaluated the application of cell phones in travel speed and travel time estimation as well as identifying congestion sections. Some researchers also used cell phone data to derive the O-D matrix (Caceres et al. 2007; Sohn and Kim 2008). However, the application of a cell phone tracking system for determining the truckshed from an intermodal facility has not been addressed.

The objective of this paper is to analyze the feasibility of using cellular telephone location data to determine the geographic extent of trucks related to intermodal facilities. The feasibility analysis includes three aspects: technology, penetration analysis, and a subject (truck) filtering and tracking methodology. First, the cell phone location tracking technologies and related research into its reliability were reviewed and analyzed. Cell phone penetrations, such as market penetration and cell phone signal coverage from cell phone carriers, were also investigated. Then, truck tracking, filtering rules, and procedures were developed to identify the trucks and track their movement from intermodal facilities. The remaining challenge and perspective of using cell phone data in determining the truckshed are discussed. Finally, conclusions and recommendations are listed based on the feasibility analysis in the end.

RESEARCH SCOPE

The purpose of this study is to determine the feasibility of using cellular phone tracking technology in determining the geographic extent of truckshed data from intermodal facilities and understanding the impact of intermodal facilities on the traffic network. The feasibility analysis is going to pursue the following questions:

- Can cell phone tracking technology enable tracking accurate locations to realize a long-haul truckshed?
- Does the sample size of cell phone signals provide enough data to reflect the truck volume and impact on the network?
- How is it possible to determine if the cell phone is a truck or not?
- How is it possible to track the truck movement based on cell phone location data?

The feasibility study answered those questions and is presented and discussed in three aspects:
1. Technique analysis: Available technologies, which were used to obtain the location information using the cell phone signal, were briefly introduced. Literature on the accuracy of cell phone location tracking system tests and studies were also reviewed.

2. Data coverage and penetration analysis: The number of cell phone subscriptions and signal coverage were investigated. The cluster of possible types of cell phone data and portion of cell phones in use were also analyzed.

3. Truck identification and tracking process: The process, which can extract data from thousands of anonymous cell phone calls to determine if the cell phones are trucks or not, was proposed. Cell phone locations can be tracked by time and space. Criteria used to filter the data were based on the land use and truck movement and stop characteristics.

The structure and operation process of cell phone location data provided by Airsage, Inc. are briefly introduced. The challenge and perspective of using cell phone location data are also discussed.

TECHNOLOGIES ANALYSIS

Cell phones are now used as the basic and regular communication tool around the world. Cellular phone location technology began to be one of the new technologies for achieving traffic information after the E911 technology was implemented. In 1996, the Federal Communications Commission (FCC) mandated E911 requirements that cellular location should be provided when 911 emergency calls come to emergency management authorities (FCC 2001).

Cell phone location tracking technologies can generally be divided into two categories: network-based and handset-based. A network-based system utilizes signal information from cell phones to derive their location. In network-based 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 the position is calculated. There is no requirement to make any changes to the current handsets. However, the cell phone must be in active mode (i.e., in “talk” mode or sending a signal through the control channel) to enable location measurement. Handset-based systems rely on global positioning system (GPS) enabled wireless phones. The GPS unit in the handset determines the location of a phone, and this information is relayed from the cell phone to a central processing system maintained by the wireless carrier.

Cell Phone Location Calculation and Limitation

According to the requirement by the FCC (2001), location accuracy and reliability should be 100 meters for 67% of calls and 300 meters for 95% of calls for network-based solutions; 50 meters for 67% of calls and 150 meters for 95% of calls for handset-based (GPS-enabled) solutions. Depending on the technology, calculation methodology, and signal path, the accuracy of cell phone location estimation are varied.

For a network-based system, two major methods (as shown in Figure 1) are used to calculate the location of the cell phone. The first one is a triangulation method. In ideal conditions, the cell phone location can be calculated exactly using the triangulation method with computed distances from three nearby stations. However, in reality, the computed distances are dependent on the reflections, diffraction, and multipath occurrences of the phone signal. The triangulation method result is an area instead of a point. The accuracy of triangulation method is about 50–200 meters (Openwave, Inc. 2002)
The other method is an angle of arrival method. In this method, special antenna arrays are installed at the base stations to calculate the direction the signal. Thus, two stations are enough to calculate in what direction the cell phone signal is coming from. Considering the effects of multi-propagation, this method has an accuracy of about 50–300 meters (Openwave, Inc. 2002).

For handset-based systems using the GPS satellite system to calculate the position of the cell phone, the accuracy is between 5 and 30 meters (Openwave, Inc. 2002). The accuracy is affected by factors such as the ionosphere, troposphere, noise, clock drift, ephemeris data, multipath, etc. The use of GPS in cell phones as a location device also suffers from three main disadvantages (Zhao 2000):

1. GPS signals are too weak to detect indoors and in urban canyons, especially with small cellular-sized antennas.
2. The time required to obtain a GPS position is relatively long, ranging from 60 seconds to a few minutes due to the long acquisition of the satellite navigation message.
3. Due to long signal acquisition time, GPS power dissipation is very high. The computations drain the battery of the phone.

The assisted GPS method, where the wireless network uses a server to perform the calculations and to transmit to the phones, can solve the delay and power consumption issues. In addition, the wireless network can use the differential GPS method to reduce the errors. In this method, a tower with a known position is equipped with a GPS receiver to estimate the total error. The errors are roughly the same in nearby areas; the estimated error can be transmitted to the phone for compensation. Thus, the corrected location in a nearby area is the non-corrected GPS location minus the estimated error. The accuracy is improved to 15 m or less (Wunnava et al. 2007).

Cell Phone Location Technology Application and Test

Several researchers have tested the accuracy of location tracking and the estimation of travel speed using cell phone tracking technology. The results showed that most of cell phone location systems can provide reasonably accurate position data. They were unsuccessful in producing traffic flow information.

CAPITAL, deployed in the Washington, D.C. area in the mid-1990s, was the first major deployment of wireless location technology (WLT) in the United States. The system was able to locate cellular phones within 100 meters of their actual position. The accuracy of the position estimates improved considerably as the number of cellular towers providing directional information increased. Speed information could not
be calculated because at least four positions are needed to calculate speed. Less than four position estimates were collected 80% of the time (UMD 1997).

The other deployment using U.S. Wireless was conducted in the San Francisco Bay region in California (Yim and Cayford 2001). Researchers at the University of California, Berkeley obtained 44 hours of wireless location data. They were generally able to determine the location of the cellular phone on the roadway network. They found that the location estimates of cellular phones were regularly accurate within 60 meters, although 66% of cellular devices tracked had at least one outlier with an error of more than 200 meters (Smith et al. 2003).

A research team of the Virginia Transportation Research Council at the University of Virginia (2005) investigated and reviewed over 16 deployments of WLT-based monitoring systems both in the United States and abroad. They concluded that most systems did not produce data of sufficient quality or quantity to provide reliable traffic condition estimates. A similar task was done by a research team at Florida International University (Wunnava et al. 2007). They investigated the maturity of cell phone technologies for application as real-time traffic probes for travel time estimations along the highways and roadways. They found that the cell phone technology is feasible to determine travel time estimations under the normal conditions of free traffic flow, but it is not accurate in congested traffic conditions. The accuracy decreases rapidly as congestion increases.

**Feasibility of Cell Phone Location Technology in Truckshed**

For a truckshed tracking system, the location data needed are trucks traveling from an intermodal facility to an area in the network or long-haul destination. During the truckshed tracking process, the network will be divided into several areas depending on the land use, network distribution, and state boundaries. The cell phone location data are tracked from area to area. A small range of error in location estimation is accepted for the truckshed tracking system.

According to previous research and tests, location estimation using cell phones was able to provide reasonably accurate location data. Cell phones could be located within an average of 100 meters or less of their actual position. Therefore, existing cell phone location technology is feasible to be used in long-haul truckshed tracking from an intermodal facility.

**DATA COVERAGE AND PENETRATION ANALYSIS**

Cell phone data coverage and penetration analysis includes the number of cell phone subscribers, the cell phone signal service area, and the portion of cell phones in use that can be tracked.

**Cell Phone Penetration**

By the end of 2008, there were more than 270 million cell phone subscriptions in the United States, which is about 87% of the total U.S. population (CTIA 2009). About 17.5% of U.S. households are wireless-only. The cell phone market penetration covers nearly all ranges of vehicle users, especially for truck drivers. Thus, using cell phones to develop a truckshed would cover all truck locations, if all the cell phone signals are available.

However, there is no integrated system that could provide all cell phone signals available up to now. The United States is the most competitive in the cell phone market in the world. The top four carriers represent only 86% of the market (CTIA 2009). The cell phone signal data belong to different cell phone carriers.
The top four cell phone carriers covered 28.5%, 26.7%, 18.2%, and 12.1%, respectively, of the cell phone market at the end of 2008 (CTIA 2009). Atlanta-based Airsage, Inc. is one of the providers of location, movement, and real-time traffic information based on cellular signaling data. (Airsage 2009)

Cluster of Cell Phone Data

Since the location data are available only when the cell phone is in use, only a partial number of drivers using the intermodal facility can be detected using cell phone location data. Considering the truckshed tracking process using cell phone location data, the cluster and relationship of the number of truck drivers and available cell phone location data in the facility area are illustrated in Figure 2.

Figure 2. The cluster and relationship of truck drivers and available cell phone data

Total people within the intermodal facility = \{NT_{ns}, NT_{nc}, NT_c, T_{ns}, T_{nc}, T_c\}, where
- \(NTD_{ns}\) = non-truck drivers who do not have a cell phone in provider database,
- \(NTD_{nc}\) = non-truck drivers who did not use cell phone during data collection period,
- \(NTD_c\) = non-truck drivers who used cell phone during data collection period,
- \(TD_{ns}\) = truck drivers who do not have a cell phone in provider database,
- \(TD_{nc}\) = truck drivers who did not use cell phone during data collection period, and
- \(TD_c\) = truck drivers who used cell phone during data collection period.

“Non-truck drivers” include employees, workers, customers, train drivers, and possible residential passing by drivers. “No Service” data consist of the drivers who do not have a cell phone or the cell phone data are not available in the provider’s database. “Cell phones not in use” represent those drivers who do not use their cell phone when they are in the facility during the data collection period. The portion of this cluster would decrease over a long time period of observation. “Available cell phone data for track” are the location data when the cell phones carriers make a call or send a message. These data include truck and non-truck drivers’ cell phone data and are the only known data for the cell phone tracking system during the data collection period.

Based on the cluster distribution above, only partial cell phone data are available for the truckshed tracking system. One of difficulties of using cell data to track a truckshed is determining the relationship between the characteristics of sample size and the entire volume of trucks from the intermodal facility. Fortunately, as the cell phone becomes the basic communication tool and more cell phone signal data will be integrated, more cell phone data will be available in the provider database. For a long time period of observation, the available cell phone data for tracking will increase since the probability of cell phone carriers making cell phones will increase. Therefore, the sample size of cell phone data would be enough for long-haul truck tracking from the intermodal facility.
TRUCK IDENTIFICATION AND TRACK PROCESS

One of the major challenges of tracking a truckshed using cell phone data is how to subtract the truck volume from other available cell phone data. All available cell phone data include truck and non-truck carriers’ location data in the intermodal facility. The first step of tracking trucks is to catch all available cell phone data within and from the targeted intermodal facility. This step can be done directly by matching the cell phone location data to the facility. Every cell phone signal located within the facility area during the data collection period is an initial data point and used for tracking.

The second step is to identify which cell phone signal point is a truck. Since the cell phone data are anonymous, the vehicle type can only be identified by tracking its moving path and destination. Thus, a database with possible truck stops, such as gas stations, rest areas, and logistic warehouses in the analysis network, is needed. A set of filtering truck rules also needs to be established according to the truck movement and truck stop database.

Truck Stop Database

The characteristics of truck movement and traveling pattern and destination are different from passenger cars. The big difference is that most trucks travel for long distances and frequently stop at popular truck stops, such as gas stations, rest areas, and logistic warehouses. Thus, those popular and possible truck stop locations should be established during the truckshed process. The Interstate highway network and city land use information could be referenced to create the database. Truck stop information available on the Internet also can be embedded into the database. During the tracking process, the observed network is divided into several regions based on the truck stops and land use information. In principle, the analysis network can be categorized into several types of regions based on the land use and truck movement characteristics. Possible types of regions include intermodal facilities, truck stops, rest areas, supermarket and logistic warehouses, business regions, and residential regions.

Truck Filter Rules and Tracking Process

Conversely, passenger car trips from the intermodal facility most likely are local and short distance. Those passenger cars can be filtered based on their destination and filter rule. For example, if the cell phone data from the intermodal is found in the residential or business area most of time, this point will be considered as a non-truck driver.

Depending on the size and the divided regions of the network, the truck filtering rules are varied. However, the principle of the filtering rules should be based on the results of a long tracking period. Figure 3 illustrates the procedure of filtering truck and tracking a truckshed. During the analysis process, only and all those cell phone signals that have ever appeared within the facility area are filtered and tracked along the entire network data. Those tracked cell phones frequently located around truck stops with a certain traveling pattern and following the filtering rules could be considered as trucks. All the filtered truck cell phones are embedded into the network map to derive the truckshed and present the impact of intermodal facility on the entire network.
Every cell phone signal point includes a unique identification number, location data in the X and Y axes, and operation time. An electronic map is also needed to embed all the cell phone locations into the designed region of the network. Data are sorted and tracked in the format shown in Table 1.

### Table 1. Sorted cell phone signal data

<table>
<thead>
<tr>
<th>Point ID</th>
<th>t1</th>
<th>t2</th>
<th>T2</th>
<th>t2</th>
<th>Filter result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n1(0,0)</td>
<td>N1(x,y)</td>
<td>n1(x,y)</td>
<td>n1(x,y)</td>
<td>T or NT</td>
</tr>
<tr>
<td>2</td>
<td>n2(0,0)</td>
<td>N2(x,y)</td>
<td>n2(x,y)</td>
<td>n2(x,y)</td>
<td>T or NT</td>
</tr>
<tr>
<td>3</td>
<td>***</td>
<td>N3(0,0)</td>
<td>n3(x,y)</td>
<td>n3(x,y)</td>
<td>T or NT</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

where

ID = cell phone initial ID as the new point appears in the intermodal facility,

T = time interval; system is scanned based on a set time interval (30 sec, 5 min, etc.), and

N(x,y) = cell phone location; the region where the cell phone locates is determined.

Vehicle types are determined based on the cell phone point tracking data and proposed truck filter rules in Table 2 for the network. Note the filtering rules are supposed to be changed based on the analysis system.

Figure 3. Procedure of filtering truck and tracking the truckshed
Table 2. Truck filter rule for the sample network

<table>
<thead>
<tr>
<th>Filtering rule</th>
<th>IF the cell phone point</th>
<th>THEN the point is</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stays in the same region more than 4 hrs</td>
<td>Non-truck driver</td>
</tr>
<tr>
<td>2</td>
<td>Heads to a business or residential area</td>
<td>Non-truck driver</td>
</tr>
<tr>
<td>3</td>
<td>Moves forward and backward between intermodal and logistic warehouse and supermarket</td>
<td>Truck driver</td>
</tr>
<tr>
<td>5</td>
<td>Is long-haul traveling along the Intestate</td>
<td>Truck driver</td>
</tr>
</tbody>
</table>

If the cell phone signal doesn’t appear after leaving the intermodal facility during the data analysis time period, the data point will be considered as invalid data. However, for long-haul tracking, this kind of data would be limited.

LOCATION PROVIDER AND CHALLENGE

Cell Phone Location Provider

Cell phone location data will be provided by Airsage, Inc., a major cell phone location provider in the United States. AirSage has long-term cooperation with Sprint wireless company to provide traffic condition and traffic speed. Airsage system has been tested in several metropolises (Smith 2006; Liu 2008). Recently, Airsage has contracted with Verizon Wireless to generate more cell phone data, which will cover nearly 50% of cell phone data in the United States.

Figure 4 shows the coverage of cell phone location and traffic condition data provided by the Airsage system. The system provides real-time, historical and predictive traffic information for 127 U.S. cities. Most of the data concentrate on the major metropolis areas. The area is expected to expand as more cell phone data are included.

![Figure 4. Airsage live traffic service area (Airsage, Inc.)](image-url)
Location data provided by Airsage system are quite simply a coordinate along an X and a Y axis. In addition, every cell phone has a unique identity number (not cell phone) and the time when the cell phone is in use condition. The location data are provided in a 30 second basis. Every 30 seconds, all the locations of all cell phones in use in the network is recorded and presented. Relative to movement data, AirSage claimed on their website that the Airsage system has the ability to anonymously track the approximate origin and destination of nearly every cell phone signal down to a few hundred feet.

**Challenge and Perspective**

Although the cell phone provider can provide a wide range of cell phone location, several challenges still remain for using cell phone data to generate the truckshed from an intermodal facility.

1. Simple size issues: The cell phone location can only be detected when the cell phone is in use. If the truck drivers did not used a cell phone as the facility, the data would not be tracked. Fortunately, long-term observations may conquer this issue if the trucks routinely return to the facility.
2. Database storage: Since the cell phone location data for the entire network are recorded every 30 seconds, especially for long-term and wide-range network tracking, the storage needed to record these data is expected to be huge.
3. Truck stops and land use categorization: Although truck stop information is available, to embed this information into a truck tracking system is still a big challenge. Since the region designed for filtering trucks will affect the accuracy of truck identification, dividing the network into appropriate regions would be a difficult task.
4. Truck data calibration and verification: Since only partial truck location can be tracked, field trip data collection is needed to check the portion of trucks that have been tracked by using cell phones. The calibration factor is also needed for generating the truck trips from the intermodal facility.

Cell phone location technology is a convenient and effective methodology to track the geographic extent (truckshed) from an intermodal facility. It can provide an updated truckshed with no time limited. As more cell phone data are integrated in the provider’s database, more precise location data will be available to use in tracking the truckshed from an intermodal facilities.

**CONCLUSION AND RECOMMENDATION**

Understanding the extent of the reach of the intermodal facility is an important first step to mitigate any negative consequences. In order to determine the feasibility of using cellular telephone location data in deriving the geographic extent (truckshed) from intermodal facilities, this paper conducts the feasibility analysis in three aspects: technology, penetration analysis, and truck tracking methodology.

According to previous researches and tests, location estimates using cell phones were able to provide reasonably accurate location data. Cell phones could be located within an average of 100 meters or less of their actual position. The cell phone location technology is feasible to use in long-haul truckshed tracking from the intermodal facilities.

The cell phone data coverage and penetration analysis showed that only partial cell phone data are available for the truckshed tracking system. Prospectively, more cell phone signal data will be integrated in the provider database, which will increase the sample size and coverage of cell phone location data. For
a long time period of observations, the available cell phone data for tracking will also increase since the probability of cell phone carriers making cell will increase.

Since the cell phone data are anonymous, the vehicle type can only be identified by tracking its moving path and destination. Thus, the truck identification and tracking process should including a database covering truck stops and land use information as well a set of truck filtering rules. The analysis network can be generally categorized into several types of regions based on the land use and truck movement characteristics. Depending on the size and the divided regions of the network, the truck filtering rules are varied. In principle, those tracked cell phones frequently located around truck stops and traveling along the interstates for long trip are considered as trucks.

Some challenges for using cell phone data in tracking trucks still remain, such as sample size, data storage, and land use categorization for truck filtering. Long-term observations are needed to increase the sample size obtained in the intermodal facility. Field data collection for verifying the accuracy of the truck identification and deriving the calibration factor of generating truck trips are areas for further study.
ACKNOWLEDGEMENTS

This study was funded by the Mid-America Transportation Center in Lincoln, NE.

REFERENCES

Airsage New Release.


Openwave Inc (2002), Overview of Location Technologies.


