

Comparative study on GPS and cellular probe techniques for freeway travel speed estimations

R. R. He¹, T. Z. Qiu² & B. Ran²

¹*Department of Civil Engineering, Santa Clara University,
CA 95053, USA*

²*Department of Civil and Environmental Engineering,
University of Wisconsin – Madison, WI 53706, USA*

Abstract

Traffic data, such as travel speed and travel time, is by nature stochastic. In order to deploy Intelligent Transportation Systems successfully, efficient, accurate, and sufficient collection of traffic data in a timely manner without interrupting the traffic flows is one of the imperative first steps. It is evident that travel speed estimates by gathering traffic data with moving vehicles is more representative of highway performance than the point estimates of speed from fixed detectors. Recently, GPS and cellular probe techniques have been researched and practiced in academic and industrial areas all around the world. The use of GPS and cellular probes for obtaining traffic data appears to be technically feasible. Still, there is little evidence to demonstrate that either system can be practically implemented. The comparative study conducted in this paper has integrated the possible real attributes of GPS and cellular probe systems, such as location accuracy, probe penetration rate and update frequency, into simulation on a real transportation network, and has developed methods for freeway travel speed estimation. The findings from simulation show that either of these two techniques has its own pros and cons. Cellular probe has wider probe penetration rate than GPS, which relieves the dependency on location accuracy, but the low update frequency reduces the coverage and accuracy of link travel speed estimation in the network. The GPS probe has a high update frequency, but the low probe penetration rate and the subsequent strong dependency on location accuracy are big obstacles to obtaining enough samples for link travel speed estimation. Thus, combining these two technologies of wide cellular probe rate and high GPS update frequency should provide better estimation of traffic conditions.

Keywords: data collection, speed estimation, GPS, cellular probe, sample size, confidence level.



1 Introduction

In order to deploy Intelligent Transportation Systems successfully, such as providing drivers with accurate estimated and predicted traffic information, efficient, accurate, and sufficient collection of traffic data in a timely manner without interrupting the traffic flows is one of the imperative first steps. Currently, traffic data collection is most commonly done by fixed collectors, such as electronic counters, loop detectors and video cameras, on the roadway sections of interest. The information collected is isolated, static and not continuous, meaning that it is hard to have a whole picture of the traffic conditions. Recently, GPS (Global Positioning Systems) and cellular probe techniques have been researched and practiced in academic and industrial areas all around the world [4, 7] for traffic data collection. By continuously gathering traffic data of moving vehicles without interrupting the traffic flows, it is expected that traffic condition estimation is more representative of highway performance than the point estimates from those fixed detectors. Due to the limited comparative research on these two technologies, the objectives of this paper are to clarify the similarities and differences of these two probe technologies, and to simulate these two methods under different scenarios on a real-world freeway network with consideration of the real attributes of GPS and cellular probe data, such as location accuracy, probe penetration rate and update frequency.

2 GPS probe and cellular probe

GPS, developed by the US military, uses a network of satellites and ground stations to transmit location information to users almost anywhere on the planet. It requires a direct, unobstructed view of at least four of the satellites to provide reasonable accuracy, and users can expect to be within about 20 meters of the GPS reading 95% of the time [2]. However, position data is noisy, especially before May 1, 2000 when President Clinton signed an executive order to discontinue the use of Selective Availability. Still, noises in GPS position data cannot be completely excluded. Random errors come from various sources, such as tree canopies, tunnels, and urban canyons can blackout or distract signals to a great extent.

With the cost of GPS devices decreasing and the use of wireless technology increasing, it is anticipated that vehicles with relatively accurate positioning systems and wireless communication capabilities will become commonplace within the next decade. Vehicles with GPS are good candidates as probes to continuously sample the roadway network. With wireless communication, the centralized processing stations can collect the probes' position traces in real time and process the information for traffic estimation and forecasting. In a GPS-based system, the GPS receiver can usually provide one to ten-second updates of latitude, longitude and time as long as the antenna can "see" four of the GPS satellites, which guarantees providing relatively accurate, continuous and automated data with low operating costs, and further reduces the ambiguity in



the subsequent processes of map matching and travel time/speed estimations. Gathering travel times/speeds with moving vehicles produces measures over space that are more representative of highway performance than the point estimates of speed from fixed detectors, particularly for capturing the essentially stochastic travel speeds/times [1]. To secure sufficient data to describe highway performance in an accurate and timely manner, a relatively large number of participating vehicles would be required. This was unimaginable until very recently.

On the other hand, the cellular network is one other type of modern wireless network. Any vehicle equipped with a cellular phone can serve as a probe, updating locations periodically to network control centers as long as it stays within the cellular network. There are two types of so-called cellular probes, one is network-based and the one is handset-based. According to the FCC (Federal Communication Commission), network-based cellular probes can achieve accuracy of 100 meters for 67% of calls and 300 meters for 95% of calls; while handset-based cellular probes can achieve accuracy of 50 meters for 67% of calls and 150 meters for 95% of calls [8]. Network-based probes use triangulation, time difference of arrival and angle of arrival to determine probe position, which is not considered in this study due to low accuracy. Handset-based cellular probes with GPS assistant can provide more accurate position information and have been researched in this study.

The number of cellular subscribers is very large. As long as the cellular phones stay within a cellular network, whether they are in use or not, their locations can be retrieved at certain instances, such as when the cellular phone enters another location area, or when the cellular phone is being turned on or off. These cellular probes can be so widely dispersed over the transportation network that it is very likely to have sufficient samples on any roadway segment at any time period. And the most important pro of cellular probing is any driver with a cellular phone can be a probe, so there is no initial cost for cellular probes. As long as the customer pays his/her phone bill, there is no additional operating cost. However, the installation cost for a GPS device ranges from a few hundred to a few thousand dollars. Besides the benefit of real-time route guidance from those GPS-based personal navigation systems, drivers may not see other benefits of being equipped with GPS. The reluctance may also come from trying to protect privacy.

3 Research on probe technologies

While it seems to be technically feasible to implement GPS and cellular probe techniques for traffic data collection, a few field trials have been conducted with ambiguous results. In 2000, Yim and Cayford [6] at the University of California conducted an evaluation of GPS and cellular probe tracking. They used an embedded Geographical Information System for map matching. This study involved a small area in the San Francisco Bay area, including part of I-580 and Broadway, which is a major arterial in the City of Oakland. Cellular systems meeting the E-911 requirement of 100-meter accuracy for two thirds of all calls were capable of producing quantitative travel time information for 87% of the



major street segments and 68% of the freeway segments. Global positioning systems with 20-meter accuracy could produce data for 99.2% of the surface street segments and 98.9% of the freeway segments. A study conducted by the French transportation research organization INRETS [3] in 2000 also examined the impacts of varying levels of probe vehicle penetration on the accuracy of travel time estimates, assuming a geolocation error of 150 meters. The results showed that freeway link travel times could be estimated within 10% of their actual values if there is at least a 5% penetration of wireless devices. In 2001 a study by the University of Maryland used simulation to examine wireless location technology based on geometric positioning. The results indicated that the network-based cellular systems could provide a general characterization of traffic flow on a freeway, but could not estimate accurate speeds.

Some other field tests include the CAPITAL field operational test (1997) and the U.S. Wireless operational test (2001). The CAPITAL test showed that cellular phones could be located within 100 meters of their actual positions and that four positions estimated had to be identified for each phone in order to calculate speed, which only occurred 20% of the time. The U.S. Wireless operational test included the UC Berkeley study by Yim and Cayford [6] which determined that estimates generally had a 60-meter accuracy, although 66% of all probe vehicle tracks had at least one data point that deviated from the caller's actual position by more than 200 meters. The Virginia Department of Transportation (VDOT) and Maryland State Highway Administration (MSHA) study [5] determined that only 4.2% of the intervals demonstrating significantly different mean speeds than the actual means for daytime periods. For nighttime periods, 27.3% of the speed estimates were different on the highway and 20.7% on major arterial streets. In 2005, Globis Data Inc. conducted a series of tests with various numbers of cellular phones being driven around the test route under different traffic conditions. Drivers and observers reported the observed travel speeds and made odometer and time reports that provided calculated speeds. The cellular phones appeared on the map in their expected positions according to driver reports. The two main conclusions from this test were that: (1) the average of all errors at each speed was consistently less than 1% and (2) the range of data points varied from +3.2% to -3.0% at just over 100 km/h, to +7.4% to -10.5% at 45 km/h. The positional accuracy tests revealed that the worst-case positional error for over 99% of the data points was 20 meters, and the mean was 7 meters.

In summary, none of the previous tests have considered all the possible impact factors, such as specific network attributes, location accuracy, probe penetration rate and update frequency. The goal of this study is to simulate two types of probe technologies for freeway speed estimations with consideration of all real attributes of GPS and cellular probe data in order to identify the pros and cons of each technology and help determine the best solution for probe data collection.

4 Simulation

First, in order to derive traffic condition information, such as travel time and travel speed, from probe data, it is necessary to match error-prone probe data to



error-prone digital map in order to determine vehicles' trajectories. The errors in probe data may be categorized in: a. Non-vehicle probes that have cellular phones or GPS devices, such as pedestrians, bicyclists, transit riders, and airplane passengers, etc. b. Limitations of positioning technologies, such as urban canopy or trees that block the "sight" of GPS satellites, the limited coverage of cellular networks, and the deliberate distortions for homeland security and military purposes. And the errors in digital maps can be categorized in: a. Error from old paper maps. b. Lack of updating recent changes. These errors can not be eliminated completely so after applying the commonly-used topological map matching methods to the noise-included probe data and the error-included digital map, only the possible traveling paths of probe vehicles can be shown, which will directly affect the estimation of travel times/speeds in the network. Hence, different levels of location accuracy are integrated into the simulation in this study in order to find out how significant the effects are for cellular phone and GPS probe technologies respectively.

Second, the probe penetration rates of cellular phone and GPS probes are another factor affecting the travel time/speed estimation. Travel speeds/times are in essence random numbers with certain means and standard deviations. For freeway segments with highly variable travel times/speeds, such as merging and diverging segments or during congested peak hours, a relatively large sample size is required to capture the stochasticity of traffic conditions. Currently, the market of GPS devices, compared with that of cellular phones, is very small. This study is going to integrate this penetration rate factor into simulation to find out how significant the different probe rates affect the travel time/speed estimation.

The third important factor integrated into the simulation is update frequency. As aforementioned, GPS devices can usually update position within 10 sec, while cellular probes' updating frequency is different and may even be irregular. This study has integrated the update frequency into simulation in order to find out the effects of different levels of update frequency on the map-matching results and travel time/speed estimations for both probe technologies.

4.1 The study freeway corridor

The freeway segment chosen to study consists of 71 freeway links on the Beltline at Madison, Wisconsin, as shown in Figure 1. The total simulation time is 30 minutes, i.e. 1800sec. And total 2,680,287 simulated location records are obtained.

The following different scenarios are simulated to identify the pros and cons of both probe technologies: 1. GPS probe penetration rate of 6%, 7.5%, or 10%. 2. Cellular probe penetration rate of 30%, 50%, or 70%. 3. The locations of GPS probes are updated every 3, 5, or 10 seconds. 4. The locations of cellular probes are updated every 5, 10, or 15 seconds. 5. Location accuracy of GPS is less than 20 meters 95%, 85% or 75% of the time. 6. Location accuracy for cellular probes is less than 50 meters for 67%, 50% or 33% of the time. 7. Travel time estimation interval is 5 minutes or 300 sec.



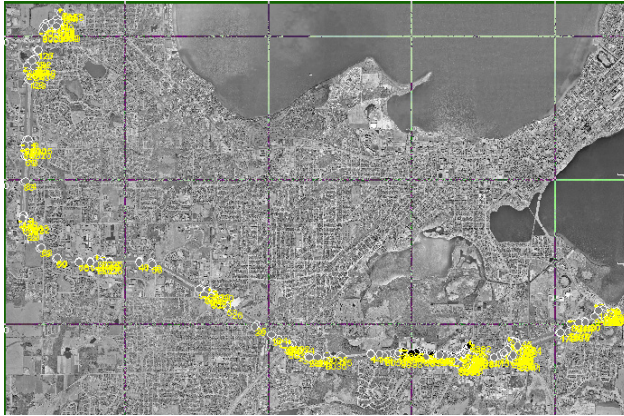


Figure 1: The Beltline in Madison, Wisconsin.

4.2 Data processing

4.2.1 Space mean speed

It is known that spot speeds are not good estimators for travel speed/time, so space speeds for the whole population and probe vehicles are calculated separately from simulation data and compared in this study. There are two ways to get one sample of space speed. First, if there are two updated position records of a vehicle k on link a , then one travel speed sample can be obtained as follows:

$$v_a = \frac{\text{position2} - \text{position1}}{\text{time2} - \text{time1}} * 3600 / 5280 \quad (1)$$

where v_a = travel speed in mph on link a , position1 = the distance in feet of vehicle k from the beginning of link a at time1 in second, position 2 = the distance in feet of vehicle k from the beginning of the link a at time2 in second.

Second, if there are two consecutive records of a vehicle k on two links a and b , the distance between the two records are first computed as follows:

$$\begin{aligned} \text{distance} = & (\text{first_link's length} - \text{position1} \\ & + \text{all_the_links' lengths_inbetween} + \text{position2}) \end{aligned} \quad (2)$$

Then a space speed sample for these two links and all the links in between is calculated as follows:

$$v = \frac{\text{distance}}{\text{time2} - \text{time1}} * 3600 / 5280 \quad (3)$$

The assumption behind this calculation is that the travel speed between two consecutive records is consistent. For one link, if more than one space speed sample comes from one vehicle, they are averaged to get just one sample speed. Then space mean speed for one link is the average of all the travel speed samples, and standard deviations are computed accordingly.

4.2.2 Links with one observation or unreasonable speed estimation

If there is only one observation on a certain link, no population travel speed is estimated. If there is only one probe observation on a certain link, no probe travel

speed is estimated. Note that the speed limit on Beltline at Madison is 65 mph. If, for a certain link, there is unreasonably high or low speed estimation, e.g. greater than 100mph or less than 5mph, those unreasonable estimations are excluded from the study. The reason for those unreasonable speeds may come from mismatching between vehicle position and the map, or the speed consistency assumption between two records.

4.2.3 Sample size

Given that the population mean and standard deviation on all freeway links are derived via the methods described in 3.2.1, the following equation is utilized to estimate minimum probe sample size in order to have $(1-\alpha)100\%$ confidence level that the probe estimations will not exceed a specified amount e (in mph) when sample size is:

$$n = \left(\frac{z_{\alpha/2} \sigma}{e} \right)^2 \quad (4)$$

where n = sample size, $z_{\alpha/2}$ = z-value leaving an area of $\alpha/2$ to the right under the standard normal curve, usually 95% confident level corresponding to $z_{0.025} = 1.96$, σ = population standard deviation in mph. It can be seen that if a freeway link has large variance, large probe sample size is required to capture the attribute. So if probe technologies are going to implement in a network with highly varying speeds, probe penetration rate, location accuracy and location update frequency are key factors that need careful consideration.

4.3 Simulation results

For the six 5-minute intervals in the 30-min simulation, 71 freeway links are observed to have more than 3500 vehicles traveling on them. Population mean, standard deviation of travel speed and required sample size for all freeway links during each 5-min interval are obtained. According to the description of different scenarios in Section 3.1, 27 scenarios have been tested for GPS and cellular probes, respectively. As an example, the results of one scenario that 70% of the total vehicles are handset-based cellular probes, the locations of cars and times were recorded every 5 seconds from cellular probes, and the position accuracy for cellular probes is less than 50 meters for 67% of calls show that handset-based cellular probes have adequate samples on 68 links (95.77% of all links in the network) with a 95% confident level to have sample mean speed within 2mph from true mean speed. Comparing the sample means computed from cellular probe data with true means for all links show that only 77.46% of the links' sample means are within 2mph range of true mean speed. The results of another scenario that 10% of the total vehicles are GPS-based probes, the locations of cars and times were recorded every 3 seconds from GPS probes, and the position accuracy of GPS probes is less than 20 meters 95% of the time show that GPS probes have adequate samples on only 38 links (53.52% of all links in the network) with a 95% confident level to have sample mean speed within 2mph from true mean speed. However, comparing the sample means computed from GPS probe data with true means for all links show that 80.28% of the links' sample means are not more than 2mph away from true mean speed.



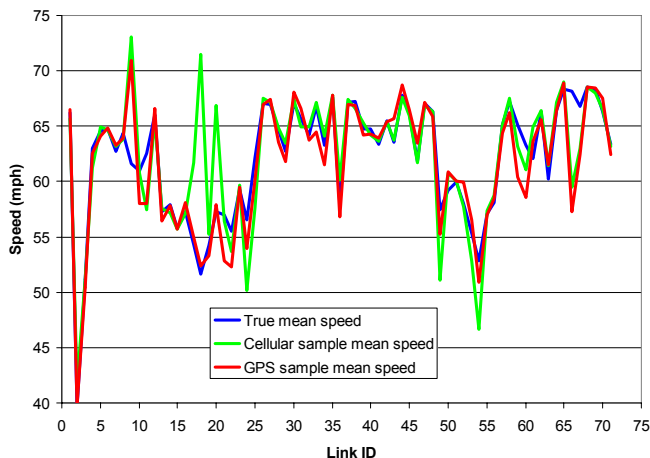


Figure 2: Comparison of true speed, cellular probe speed and GPS probe speed.

It is interesting to find that although GPS probe data seems insufficient due to a low probe penetration rate, it has better mean speed estimation (closer to true mean speed) than cellular probe data, which is also shown in Figure 2.

The summary over the total 54 scenarios show the evident trend: for both probe technologies, lower probe penetration rate, smaller probe coverage of network; longer update interval, smaller probe coverage of network; and lower location accuracy, smaller probe coverage of network. However, the impacts from these three factors are more significant on probe coverage of network than on the mean speed estimation. Each factor's specific impact on GPS and cellular probe data is analyzed and summarized in the following section.

4.4 Comparative analysis

Using the proposed data processing and computation procedures, 54 scenarios have been tested to find out the pros and cons of two probe techniques. In specific, the following attributes are analyzed and discussed in detail:

4.4.1 Location accuracy

As aforementioned, position accuracy of GPS is less than 20 meters 95% of the time, while position accuracy for cellular probes is less than 50 meters for 67% of calls and 150 meters for 95% of calls. With consideration of the error-prone digital map and the limitation of map-matching technology, different levels of location accuracy (95%, 85%, and 75% for GPS probes; 67%, 50%, and 33% for handset-based cellular probes) are tested and summarized in Table 1 and 2.

The 33% decrease in accuracy of cellular probes from 67% to 33% reduced on average 1% ~ 2.5% of links having sample mean speeds within 2mph range of true mean speeds when update rates are 5 ~ 15sec. This finding tells two things: one is that the effect of location accuracy is limited for cellular probe

Table 1: Cellular probe data summary 2 (bold numbers stand for percentage of links having sample mean speed not more than 2mph away from true mean speed).

Probe rate 70%		Location Accuracy		
		67%	50%	33%
Update Frequency	5sec	77.46%	77.46%	74.65%
	10sec	69.01%	63.38%	63.38%
	15sec	59.15%	57.75%	57.75%
Probe rate 50%		Location Accuracy		
		67%	50%	33%
Update Frequency	5sec	77.46%	80.28%	76.06%
	10sec	66.20%	63.38%	66.20%
	15sec	56.34%	54.93%	53.52%
Probe rate 30%		Location Accuracy		
		67%	50%	33%
Update Frequency	5sec	74.65%	76.06%	74.65%
	10sec	61.97%	61.97%	64.79%
	15sec	57.75%	57.75%	54.93%

data; the other is if cellular probe vehicles update their locations longer than every 10 sec, then the good travel speed/time estimation can only be done for approximately half of the whole network, which is not adequate. It is important to guarantee the less-than-10sec update frequency of cellular probes in order to have reasonable link coverage for good travel speed/time estimation. On the other hand, the 20% decrease in accuracy of GPS probes from 95% to 75% reduced on average 20% to 30% of links having sample mean speeds within 2mph range of true mean speeds when update rates are 3 ~ 10sec. This finding tells that estimation from GPS probes is very sensitive to location accuracy. Low location accuracy and long update intervals significantly reduce the validity of GPS probe data for travel speed/time estimation.

4.4.2 Sampling rate (probe rate)

The statistics of using GPS devices and cellular phones are not available, so different levels of penetration rate of these two technologies are tested according to the best knowledge of the authors. Different levels of probe rate (6%, 7.5%, and 10% for GPS probes; 30%, 50%, and 70% for handset-based cellular probes) are tested and summarized in Table 1 and 2.

On average, the 40% decrease in cellular probe rate from 70% to 30% reduced less than 3% of link having sample mean speeds within 2mph range of true mean speeds when accuracy rate are 33% ~ 67%. This finding tells two things: one is, again, that location accuracy has limited impacts on cellular probe

Table 2: GPS probe data summary 2 (bold numbers stand for percentage of links having sample mean speed not more than 2mph away from true mean speed).

Probe rate 10%		Location Accuracy		
		95%	85%	75%
Update Frequency	3sec	80.28%	80.28%	78.87%
	5sec	69.01%	71.83%	71.83%
	10sec	54.93%	52.11%	59.15%
Probe rate 7.5%		Location Accuracy		
		95%	85%	75%
Update Frequency	3sec	85.92%	67.61%	60.56%
	5sec	57.75%	56.34%	60.56%
	10sec	49.30%	57.75%	39.44%
Probe rate 6%		Location Accuracy		
		95%	85%	75%
Update Frequency	3sec	59.15%	26.76%	0.00%
	5sec	56.34%	26.76%	0.00%
	10sec	53.52%	19.72%	0.00%

estimation; the other is that the effect of cellular probe rate is also not significant as long as it is more than 30%. On the other hand, as long as the location accuracy is 95%, the effect of probe rate on GPS probe estimation is limited too. However, when location accuracy is 85% or lower, probe rate's effect is getting strong. A special case is that 6% of GPS probe rate with 75% of location accuracy cannot provide any samples for travel speed/time estimation. This finding tells that if GPS probe rate is 7.5% or less, high location accuracy has to be guaranteed in order to have good time/speed estimation.

4.4.3 Update frequency

Different levels of update frequency are also tested to reflect the real attributes of these two probe technologies. In specific, the locations of cars and times were recorded every 3, 5, or 10 seconds from GPS-based probe vehicles. The locations of cars and times were recorded every 5, 10, or 15 seconds from cellular probes. The results are summarized in Table 1 and 2.

On average, the 10-sec increase in cellular probe update rate from 5sec to 15sec reduced approximately 20% of links having sample mean speeds within 2mph range of true mean speeds when probe rates are 30% ~ 70%. This finding tells that cellular probe update rate is significant for travel time/speed estimation. On the other hand, on average, the 7-sec increase in GPS probe update rate from 3sec to 10sec reduced more than 20% of links having sample mean speeds within 2mph range of true mean speeds when probe rates are 10% and 7.5%, but only

4% link reduction when probe rate is 6%. This finding tells two things: one is that the update rate is significant for travel time/speed estimation from GPS data; the other is that GPS probe rate should be at least 7.5% in order to have good travel time/speed estimation coverage in the network.

5 Conclusions and future work

The use of GPS and cellular probes for obtaining traffic data appears to be technically feasible. Still, there is little evidence to demonstrate that either system can be practically implemented. The few field trials that have been conducted have provided mixed results because insufficient sampling size or inaccurate location referencing methods. Simulation based research has tended to consider unrealistic network conditions without differentiating network attributes, probe rate, location accuracy and update frequency.

The comparative study conducted in the paper has integrated the possible real attributes of GPS and cellular probe systems into simulation on a real transportation network, and has developed methods for freeway travel speed estimation. The findings from simulation show that either of these two techniques has its own pros and cons. Cellular probe has wider probe penetration rate than GPS, which also relaxes the dependency on location accuracy, but lower update frequency reduces the coverage and accuracy of link travel time/speed estimation in the network. GPS probe has higher update frequency, but the low probe penetration rate and related strong dependency on location accuracy are big obstacles to obtain enough samples for link travel time/speed estimation. The possible and more efficient way to collect traffic data is to integrate these two probe techniques. Then, the wide cellular probe rate plus high GPS update frequency should provide the best estimations of travel time/speed for the whole network.

References

- [1] He, R., Kornhauser, A. and Yelinek, M. (2006). "A Study on Temporal and Spatial Variability of Travel Time", Proceedings of The 11th International Conference on Computing and Decision Making in Civil and Building Engineering.
- [2] He, R., Kornhauser, A. and Ran, B. (2003). "Map Alignment Using Probe Vehicle Data", Proceedings of Transportation Research Board Annual Meeting.
- [3] Hellinga, B., Fu, L., and Takada H., Obtaining Traveler Information via Mobile Phone Location Referencing – Challenges and Opportunities, 2003 Annual Conference of the Transportation Association of Canada, 2003
- [4] Qiu, Z. Jin, J., Cheng, P. and Ran, B. (2007). "State of the Art and Practice: Cellular Probe Technology Applied in Advanced Traveler Information System." Proceedings of Transportation Research Board Annual Meeting.
- [5] Smith, B., Zhang, H., Fontaine, M. and Green, M. (2004). "Wireless Location Technology-Based Traffic Monitoring: Critical Assessment and



- Evaluation of an Early-Generation System.” ASCE Journal of Transportation Engineering, 130(5), 576–584.
- [6] Yim, Y. and Cayford, R. (2002). “Positional Accuracy of Global Positioning System and Cellular Phone Tracking for Probe Vehicles.” Proceedings of Transportation Research Board Annual Meeting.
- [7] NCHRP Project 70-01, Private-Sector Provision of Congestion Data: Probe-based Traffic Monitoring State-of-the-Practice Report, November 21, 2005
- [8] Fact Sheet – FCC Wireless 911 Requirement, January 2001
http://www.wirelessdevnet.com/e911/factsheet_requirements_012001.pdf

