Operational Parameters Affecting the Use of Anonymous Cell Phone Tracking for Generating Traffic Information

By

Randall Cayford
Institute of Transportation Studies
University of California, Berkeley
111 McLaughlin Hall
Berkeley, CA 94720-1720
510-642-1008
Fax: 413-653-2533
Cayford@its.berkeley.edu

And

Tigran Johnson
TTech Systems, Inc.
25A Crescent Drive, #181
Pleasant Hill, CA 94523
925-256-8384
Tigran@ttechsys.com

Words: 6,245 + 4 figures

ABSTRACT
This paper is an investigation of the use of anonymous tracking of cellular phones to generate traffic information on a road network. The E911 initiative mandated by the FCC presents the possibility that a traffic surveillance system based on the movements of cell phones could generate travel information for all roads in an urban street network. The effectiveness of a cell phone based traffic monitoring system depends on the interaction of several parameters including: the accuracy of the locations, the frequency with which location measurements are taken, and the number of locations available to support traffic monitoring in a given area. This research examined the effects of these parameters on the operation of a possible traffic monitoring system. Ranges of likely values were determined for the operating parameters based on the technologies being installed by the cellular telecommunications companies and combined into sets representing likely operating modes. For each operating mode, the effectiveness of a cell phone based traffic monitoring system was examined.

SUMMARY

A variety of technologies are being deployed by the cellular telecommunications companies to more accurately locate callers. These technologies have differences in the kind of location data which they can provide and differences in how they can be operated. This research project compared the capabilities and characteristics of the location technologies currently being deployed by the wireless carriers against the software requirements for an effective traffic monitoring system based on analyzing cell phone locations.

Three variables were examined to determine their effect on the behavior of a traffic monitoring application. The first was location accuracy. This is a characteristic of the particular location technology deployed and affects the ability of a traffic monitoring system to correctly determine which road a particular vehicle is traveling along. The second variable examined was the time between successive location measurements for a particular phone. This is a characteristic of both the location technology used and of the mode of operation of the location measurement system. Generating locations uses carrier resources and carriers may not wish to allow locations to be measured as frequently as is possible technically. Finally, the third characteristic was the number of locations per second available for traffic monitoring in a particular area. Using more locations causes greater load on the carrier's network while using fewer locations decreases the probability that traffic conditions on a particular street will be measured.

Values or ranges of likely values for each of the three characteristics were determined and combined into sets representing realistic operating conditions. Software from an earlier UCB study(*1*)was used to evaluate the performance of a traffic monitoring application based on these parameter sets. The evaluation was done using simulated vehicle traffic over the road networks of two California counties.

The first task was to evaluate the effect of location accuracy on the ability of the UCB software to place the phone on the correct road. Earlier research had found that using location technologies with 100 meter accuracy were capable of unambiguously determining the correct road for 87% of surface roads in two counties in California but for only 68% of the freeways in those counties(*1*). A major emphasis of this project was to improve those results. Two enhancements were developed which dramatically improved the correct road identification: a driver behavioral model and a spatial shift model.

A behavioral model was developed which used driver behavior and the physics of vehicle movement to compare the characteristics of alternate paths. This allowed the discard of many of the ambiguous paths. A second model, a spatial shift model, was developed which compared the characteristics of paths with respect to whether a driver was likely to shift from one to the other under normal conditions. Paths to which drivers were unlikely to shift were discarded.

The behavioral model and the spatial shift model were applied to simulated traffic on the road networks of the same two California counties used in the earlier study. The results showed a large improvement for all classes of roads and particularly for freeways. Using the improved algorithms, a location technology with 100 meter accuracy can determine the correct road for

98.4% of all surface streets and 98.9% of all freeways. A technology with 50 meter accuracy can determine the correct road for 99.5% of all road segments.

Values for two parameters affecting the providing of cell phone locations were established through interviews with wireless carriers and with the providers of the location technologies being implemented by the wireless carriers. The technologies examined included those currently being deployed by 6 of the top 7 carriers and which will cover at least 68 million cell phones in the United States. The two key parameters established were frequency of position update, expressed in number of seconds between successive locations of a particular phone, and load on the carrier network, expressed as number of simultaneous locations per second per square mile.

Examination of the impact of different frequencies of position update showed a decline in the percentage of road segments which could be correctly identified as the time between updates increased. For a location technology accurate to 50 meters, the percentages fell from 99.5% of the roads at 1 second intervals to 98% at 45 second intervals. At 30 second intervals—the frequency preferred by most of the carriers—98.8% of the roads segments could be identified.

Constraints on the number of locations available at a time impact the likelihood that a particular road segment will have a probe vehicle traveling on it. The impact of the number of locations on the spatial distribution of roads that could be covered in a 5 minute time interval was examined. As the number of locations increases from 1 per second per square mile to 10 per second per square mile, the percentage of roads covered increases very rapidly. Above 20 per second, the increase in the percentage of roads covered declines. There is little benefit to using more than approximately 40 locations per second per square mile.

The results of all the algorithm improvements were incorporated into an evaluation model which combined the requirements of the analysis software and the carrier location technology constraints. The operation of the software was then evaluated using three sets of optimal values for the key parameters identified.

Matching the characteristics of the location data stream available from the wireless carriers with the requirements of the analysis software showed the effectiveness of a cell phone based probe system. With a network based location technology, measurements on 85% of the roads can be generated using approximately 5% of the location capacity of a single carrier, in every 5 minute interval. Operating continuously, traffic information coverage generated over time will approach 97.7% of the road segments, the maximum possible with a location technology accurate to 100 meters using a 30 second update frequency. With handset based location technologies, using 5% of the location capacity of a single carrier can generate measurements for over 90% of the road segments in every 5 minute interval. Over multiple time intervals, the 50 meter accuracy of handset based technologies is capable of generating traffic information for 98.8% of the roads using a 30 second update frequency. With either the handset or the network based E911 location systems currently being installed by the wireless carriers, anonymous cell phone locations can generate traffic information for almost all the roads in a metropolitan area almost all the time.

BACKGROUND

In 1996, the Federal Communications Commission, issued a mandate, generally referred to as E911, which requires cell phone operators to be able to locate cell phones used to make 911 emergency calls. The mandate can be met with either network based systems or handset based systems or a combination of the two. For network based solutions, which are implemented in the cell towers, the FCC requires that all calls be locatable to within 100 meters for 66% of the calls and 300 meters for 95% of the calls. For handset based solutions, requiring new handsets, the requirement is 50 meters for 66% of the calls and 120 meters for 95% of the calls. While the October, 2001 deployment required by the original FCC mandate has been slightly delayed, initial deployments of enhanced location technologies have been made in many counties and several models of enhanced handsets are currently being sold. Both network and handset based location solutions are theoretically capable of generating the tracking data needed to generate traffic information.

Current traffic data gathering

Traffic information has traditionally been gathered in two ways. Qualitative data is gathered through reporting via traffic helicopters or traveler call-in. Quantitative data is usually gathered by public agencies (state Departments of Transportation) via surveillance devices installed in the roadway. Alternatively, quantitative data can be collected for specific roads by sending test vehicles, called probes, along a route and measuring their travel.

The most common installed systems generating quantitative data use inductive loops embedded in roadbeds. Costs are approximately $100,000 per installation with approximately $2000 per year in communication and maintenance costs. Loops are usually installed at 1/3 mile to one mile intervals. Other surveillance systems have been proposed or are in development, using radar guns, microwaves, video surveillance or electronic toll tags. Though cheaper than loops, all these systems suffer from a high cost of installation, typically $12,000 to $20,000 per unit, since they require equipment installed along the roads, and from high operating costs, since they require a communications link back to a central office and roadside maintenance. Additionally, they require installation on public property, limiting the ability of private companies to install and operate surveillance systems. The cost of installation limits use of all these technologies to freeway or highway surveillance.

An alternative to fixed installation systems uses traveling vehicles as probes to gather information about traffic conditions. Probe systems have been proposed using several technologies. The most common technique uses GPS receivers with wireless communications to report back to a base station. This setup is found in Automatic Vehicle Location (AVL) systems used for fleet tracking. The primary limitation of probe systems is lack of sufficient probes. AVL systems require installation of moderately expensive equipment in each vehicle as well as wireless communication capability for each. A large number of vehicles would have to be so equipped and an insufficient number of probes limits the ability of the system to generate information for large numbers of streets and limits the accuracy of the results.

Cellular phone based traffic data gathering

The implementation of E911 location systems presents an opportunity to develop a probe system with much greater capabilities than previously possible. Tracking cell phones overcomes the two primary hurdles to a successful probe system: insufficient probes and the high cost of in-vehicle equipment. The primary cost of a cell phone probe system is the installation of locating equipment in the cell towers. Since this is being required by the E911 mandate, a traffic surveillance system can be implemented with small additional cost.

A cellular phone based system provides several major advantages over competing techniques. It can provide traffic information for all streets in an area, not just freeways, unlike fixed installation systems. Unlike AVL systems, it can provide enough probes to generate accurate times for all roads (as many as 25,000 probes may be required to provide optimal coverage of the San Francisco Bay Area(*2*)). It can provide quantitative data about all the roads continuously, unlike traditional qualitative traffic reporting systems. This allows the building of extremely sophisticated user information systems which have never before been possible.

Four years ago, the California Department of Transportation (Caltrans) funded a study at the Institute of Transportation Studies at the University of California, Berkeley (UCB) to investigate the feasibility of using emerging location technologies to build a traffic surveillance system(*1*). The project developed software which took a series of locations from an individual probe, calculated the path followed and the travel time for each of the road segments traversed. A key characteristic of the software developed by UCB was that the travel times could be accurately calculated using quite inaccurate locations. Previous proposed probe systems required knowing the location of the vehicle to within 5-10 meters. The UCB software could calculate travel information for 85% of all roads in an urban area using locations which are only accurate to 100 meters, the required accuracy of network based E911 location systems.

RESEARCH FINDINGS

The FCC mandate (E911) specifies only the minimum accuracy with which emergency calls must be located. An evaluation of current cell phone location technologies was conducted to identify any additional parameters or constraints affecting a traffic monitoring application. Carrier discussions identified two important characteristics of their location technology—frequency of position update and number of simultaneous locations available.

In each of these areas, accuracy, frequency, and number of positions, algorithms were developed and parameters affecting the performance of a cell phone based probe system were examined. Operational values were determined for each parameter and sensitivity of the probe system to changes in the parameter value was studied. Optimal values or ranges of values of key parameters were combined to simulate likely operating conditions. For each likely set of parameters, the performance of a traffic monitoring system based on those parameters was estimated.

Current Cell Phone Location Technologies

The E911 mandate specifies two accuracy ranges which must be met by an acceptable location technology. Other parameters which may impact use of the location infrastructure for other purposes, such as traffic monitoring, are not relevant to the base task the FCC requires and are not specified by the FCC. As part of this research, extensive investigation into the current and planned technologies used by the wireless carriers to meet the FCC mandate was conducted.

The provision of E911 services requires the involvement of multiple companies. The wireless carriers, the manufacturers of network equipment, the manufacturers of handsets, the manufacturers of the location technologies, and the manufacturers of position determining equipment. Interviews were conducted with representatives of several of these players: 6 major wireless carriers, 4 location technology providers, and 3 position determining equipment/location platform providers.

The FCC allows two approaches to meeting the E911 mandate, network-based solutions and handset-based solutions. All the major carriers have chosen handset-based solutions to provide E911 services as their preferred approach. There are three standard technologies used for cellular phones in the US. Analog phones use TDMA while newer digital phones use either CDMA or GSM. Different location technologies have been developed for each of these cellular phone standard technologies. Two carriers, AT&T and Cingular, are in the process of switching from one wireless technology, TDMA, to another, GSM. They are implementing network based solutions for their legacy TDMA handsets while implementing handset based solutions for their new GSM subscribers.

Handset based solutions for E911 are required to locate cell phones to within 50 meters, 66.7% of the time and to within 120 meters, 95% of the time. For CDMA networks, the preferred location system uses assisted GPS positioning. This technology has been certified to meet the E911 requirements by the FCC in field deployments by Sprint and by Verizon. For GSM networks, the preferred location system is E-OTD (enhanced observed time differential). Cambridge Position Systems is the leading provider of this location technology. This technology has not yet been certified as commercial handsets are not yet available. Testing on prototype handsets has demonstrated accuracies within 70 meters, 66.7% of the time. For legacy handsets on TDMA networks, both TruePosition and Grayson Wireless have been certified to meet the FCC requirements of 100 meters, 66.7% of the time.

The particular choice of location technology introduces constraints on any application using its locations beyond just accuracy. The discussions with carriers and with location technology providers identified two characteristics which are key determinants of the feasibility of a traffic monitoring system, the frequency of updates from a particular phone, and the total number of locations available from a particular cell.

The first factor, frequency of position updates, is affected by a number of factors, including latency of the carrier network, time required for the handset to perform internal calculations, transmission time for transferring handset information to the carrier processing center, and latency in the position determining equipment. While different technologies had different limits, it was possible to establish a range within which a commercial application of locations will likely operate. Limits on frequency of updates ranged from 3 seconds to 30 seconds. For most technologies, the minimum time between updates is decreasing as improvements in the system are being made. Most commercial applications are not expected to require very high update frequencies, however, so there is little interest within the carriers in expending significant resources in deploying systems capable of very frequent updates. Times of 15 seconds were the expected lower limit for most carriers and 30 seconds was the preferred operating limit.

The second important factor— for an application such as traffic monitoring which depends on processing large numbers of locations—is constraints on the ability to locate a large enough sample. The number of locations which can be determined at any one time limits the sample size possible. To determine the range of likely limits, information was collected from the interviewees about the capacity of their systems. There were no hard limits in any of the technologies. The number of locations that can be determined is primarily a function of cost not of capability. If more locations are needed, additional hardware can be added at an additional cost for implementation. Implementations currently being deployed are being sized based on the load required by expected applications. A commonly cited application is “buddy finding”. This is an application where a user checks whether anyone on their buddy list is within a target distance of their current location. A national rollout of such a service is planned by several of the carriers and the location platforms underlying these services are being sized to accommodate several million simultaneous location measurements.

Traffic monitoring is a localized application, however. Locations of an adequate number of probe vehicles are needed from a relatively small geographic area. The most appropriate metric for measuring capacity of the location system for a traffic monitoring application is the number of locations per unit area. An implementation of an A-GPS system tested in August, 2000 was capable of delivering 62 locations per second per square mile. A enhanced version of the same A-GPS system tested in March, 2002 can produce approximately 480 locations per second per square mile.

Based on interviews with carriers, location technology providers and location platform providers a set of parameter ranges were determined and used throughout the feasibility study as constraints on the system. The accuracy was assumed to be one of 3 values, 50 meters, 70 meters or 100 meters. These values cover all the technologies currently in deployment. Locations were assumed to be updated with a frequency from 3 seconds to 30 seconds with greater analysis done on the 15 second and 30 second time intervals as these were determined to be the intervals preferred by the carriers. Finally the number of locations per second available was required to be less than 120 per square mile, 25% of estimated capacity, to show technical feasibility. Since the cost to the carriers of providing locations increases with the number of locations required, emphasis was placed on the development of new algorithms and improvements to existing algorithms to lower the number of locations required to achieve the same level of road coverage and accuracy.

The Effect of Location Accuracy

The earlier UC Berkeley study used a simple path following model for matching the phones to particular road segments. The weakness in that model was that phones on many road segments could not be distinguished from phones on nearby road segments. As the inaccuracy in the location measurement increased, the percentage of unambiguous road segments decreased rapidly, particularly for freeway segments. When using a technology generating locations with 100 meter accuracy, the algorithm could only determine unambiguously the correct road segment for 68% of the freeway segments. With technology accurate to 50 meters, 96% of all roads could be unambiguously determined but only 80% of freeways.

A major goal of this research was to improve the basic algorithm's ability to correctly place phones on road segments in difficult cases, such as when a frontage road runs close beside a freeway. Three areas of the basic algorithm were improved: the underlying network topology was improved, a driver behavioral model was added, and a spatial shift model was developed.

The underlying network topology is an important component in determining possible paths taken by a phone probe. The UC Berkeley model used just basic connectivity and road classification information. An enhanced network topology was developed using more information available from the commercial map data supplied by Etak. Turn restrictions between links was added, determination of connectivity at freeway junctions was improved, and enhancements to the speed classifications were made based on the physical shapes of the roadways.

The second area of improvement was the development of a driver behavioral model. The UCB study used a best case travel time estimate when determining whether a phone might be on one road or another. An improved movement model was developed taking into account the physics of vehicle movement and driver comfort levels(*3*). The result is a much more realistic set of constraints on possible phone locations.

The third area of investigation was the addition of a spatial shift model. Several microscopic traffic simulation models use algorithms which simulate the movement of traffic onto and off of freeways based on differences in travel times between the freeway and parallel arterials(*4*). The characteristics of ambiguous road segments and sequences of road segments were examined and matched against these models. Heuristics were developed to weight one path in favor of another based upon differences in the characteristics of the two paths.

Figure 1 shows the results of these improvements in the basic road assignment algorithm. It shows the percentage of unambiguous road segments identified at each location accuracy. As in the UC Berkeley study, simulated probes were generated for each location accuracy threshold traveling along every road segment. The probes were analyzed and those road segments where the phone could be correctly assigned were identified.

Incorporating driver behavior dramatically improves the basic algorithm. Under the original algorithm phones on only 68% of the freeway segments could be correctly located using locations accurate to 100 meters. Under the new algorithm, this has increased to 98.9%. At 100 meter accuracy, the minimum requirement of the E911 mandate, 98.4% of all roads can be unambiguously identified. At 50 meters, the requirement for handset solutions to the E911 mandate, 99.3% of freeways and 99.5% of all road segments can be correctly identified.

The previous UCB study concluded that a cell phone based traffic monitoring system would need to be augmented with an additional technology, such as magnetic loops, in order to provide a network wide traffic information system that included freeways. The improved algorithm makes cell phone based probe systems a feasible stand-alone technology for a network wide traffic information system.

Frequency of position update

Most research in probe systems has focused on GPS equipped test vehicles. The hardware used in these tests was generally capable of generating locations once per second, a standard rate for GPS receivers. The location systems being deployed by the wireless carriers for E911 are more limited. Even though some of the technologies use a variant of GPS, the rate at which locations can be generated is lower. The effects of the frequency of update on the ability of the basic road assignment algorithm to place phones was examined.

The testing was done using a very large number of simulated vehicle tracks. Vehicle tracks with locations at one second intervals were generated for every road segment and stored. Then, the tracks were analyzed with varying numbers of intermediate points removed. The percentage of road segments correctly identified was calculated for each of the different frequency values.

There are two effects from increasing the number of seconds between successive locations. First, as the time increases, the possible paths a particular vehicle took to reach a given point increase as well. This leads to more situations where the phone can not be assigned to the correct road segment. Figure 2 shows the percentage of all roads identified as a function of update frequency. As the time between locations increases, the percentage of roads which can be correctly identified decreases from 99.5% of the roads at 1 second intervals to 98.0% of the roads at 45 second intervals. The second effect of longer times between locations is an increase in the distance over which a particular phone must be followed in order to determine its path. At 1 second intervals, a vehicle only needs to travel an average of 70 meters before the road segment is identified. At 45 second intervals, the vehicle will move an average of 1027 meters. The increase in path length is not a problem for a traffic monitoring system but it does increase the number of phones that need to be tracked simultaneously for a given level of coverage.

Number of locations and spatial distribution of phones

A key parameter for the carriers is the number of locations required to generate traffic information over the complete road network. The technologies deployed have constraints on the total number of locations which can be produced from each cell tower in a period of time. This parameter has changed over the course of the last two year. In Fall 2000, the maximum number of locations per second was approximately 62 per square mile for one carrier. That number has increased to 480 locations per square mile in urban areas under the current implementation for that carrier and further improvements are expected.

The basis of a probe system is that only a sample of all the vehicles on the road is being measured. The more vehicles sampled, the larger the likelihood that a particular road segment will be traversed by a probe vehicle. Using more vehicles as probes requires more locations, however. To measure the sensitivity of a cell phone based probe system to the number of locations, a system with varying numbers of locations per second was simulated.

The simulation used two parameters as input, the number of locations per second per square mile, and the frequency with which the location of a particular probe was updated in seconds. The total number of probe vehicles per square mile that can be monitored simultaneously is the product of these two parameters as locations are collected every second for 1/frequency of the set of probes. Running the simulation determined what percentage of the road network would be covered by a probe vehicle given the number of probe vehicles available.

The road network was divided into one square mile sections. Within each section, the specified number of probe vehicles per square mile were generated and randomly assigned to links within that one square mile. For each vehicle a randomly determined path was generated approximately 1/2 mile in length using points accurate to 50 meters. The simulated probes were processed for a 5 minute simulation interval. When the end of a vehicle's path was reached, a new vehicle was created, randomly assigned to a starting link and processed. After the end of the simulation period, the total number of links for which a speed and travel time had been calculated was calculated. The results for varying numbers of locations per second per mile are shown in Figure 3.

With vehicles located every 15 seconds, the coverage in a 5 minute interval rises quickly until approximately 40 locations per second per square mile. At 40 locations per second, 87.3% of the roads were traversed by a probe vehicle. A greater number of locations in the 5 minute interval yield little improvement with 100 locations per second generating probes on only 87.9% of the road segments. A similar curve occurs with vehicles located every 30 seconds. The curve rises much more steeply, however, flattening out at only 5 locations per second per mile. The upper bound is slightly higher as well with the maximum tested value of 100 locations per second producing data for 91.9% of the road segments. The higher percentage of coverage for the 30 second interval is due to the fact that the number of locations for a particular probe at 30 second intervals will span a longer distance than that number of locations at 15 second intervals.

Based on the current known carrier constraints of approximately 480 locations per second per square mile in urban areas for a single major carrier, the curve in Figure 3 shows that a high degree of coverage can be achieved well within the constraints of the system. The curve for 30 second updates starts to flatten at 5 locations per second. This represents approximately 1% of the location capacity of a location system. The curve has become close to flat by 20 locations per second, representing 4% of the location capacity of the system. Only minor improvements can be expected using any number of locations larger than 50 per second representing 10% of the system capacity. The coverage curves show that a single major carrier is sufficient to provide enough locations for a cell phone based probe system generating traffic information over almost all roads in an area. Combining locations from multiple carriers decreases the load on each of the participating carriers still further allowing an effective system with very small loads on the networks of each carrier.

Operational parameters

Each research task in this project attempted to define the appropriate value or range of values for parameters affecting the operation and effectiveness of a probe system based on cell phone location data. While there are many factors that affect such a probe system, certain key parameters were investigated. Based on the analysis of the response of the system to the parameter variations, a set of nine combinations of parameters was developed. For each of these parameter combinations, the overall effectiveness of a traffic monitoring system using cell phone probes was evaluated.

Based on FCC certification of the location technologies being deployed, the accuracy of location measurements for network based solutions has reached 100 meters for 66% of the measurements. For handset solutions based on assisted GPS, the accuracy has reached the required 50 meters for 66% of the measurements. The only exception for the accuracy requirements is for GSM based handset solutions. These have not yet been certified by the FCC and have only demonstrated 70 meter accuracy. While the technology vendors claim that they will reach 50 meters, it can not be assured that this level of accuracy will ever be deployed. While the FCC only requires that locations be accurate to the greatest precision 66% of the time, the investigation of error correction algorithms suggest that occasional lower accuracy points will not have an adverse effect on system operation. For the overall system evaluation, the accuracy of all locations was assumed to be either 50, 70 or 100 meters.

The second key parameter is the frequency with which each probe is located. While the maximum coverage that can be achieved decreases as the time interval between locations increases, the effect is fairly minor. Given the carrier constraints and preferences, a value of 30 seconds was chosen. This is the current maximum frequency for several of the carriers and yielded greater road coverage over a 5 minute interval than more frequent updates. While the total percentage of roads which can be unambiguously identified with 30 second updates is slightly smaller than the percentage identifiable with 15 second updates, the difference is less than 1%. The largest impact of the frequency of location updates is on the length of path a vehicle must follow before it can be correctly placed on its road segment. For the overall system evaluation, a path length of twice the average path length required for locating the vehicle was chosen.

The last key parameter was the number of locations per second per mile. For this key parameter, the percentage of roads covered in any time interval will increase as the number of locations per second rises. But the cost of the system both in terms of processing and in terms of carrier cost to produce the locations increases with the number of locations as well. So any choice of value represents a tradeoff between the likelihood that travel on a particular road will be measured in each time period and the cost of producing that measurement. Given the shape of the curve shown in Figure 7.1, three values for the number of locations per second were chosen.

For the minimum system, a value of 5 locations per second was chosen as the improvement in coverage starts to fall off at approximately this point. For a high coverage system a value of 25 locations per second was chosen. The coverage curve has flattened dramatically by this point and values higher than 25 yield relatively small improvements. A third value of 10 locations per second was chosen as well, representing a compromise between lower numbers of locations and reasonable coverage.

Three location accuracies and three values for locations per second were combined with the single valued parameters to produce nine sets of operational parameters. For each of the nine sets, the performance of a probe system based on those parameters was evaluated. A large number of simulated probe vehicles was generated on randomly assigned road segments and paths with a minimum path length of 1,348 meters with locations within the test accuracy threshold. As many of the simulated vehicles as possible were processed given the number of locations per second available over a 5 minute period.

The results are shown in Figure 4. The performance ranged from a low of 75% coverage every 5 minutes for a system based on 100 meter accurate technology with only 5 locations per second to a high of 90% coverage every 5 minutes for a system based on 50 meter accuracy and 25 locations per second. These percentages are for road segments covered in each five minute interval. Over multiple intervals, a different set of road segments will be measured and the coverage will approach the limits for unambiguous road segments, established earlier as 98.8% for a technology with 50 meter accuracy and 30 second update frequency.

Conclusion

The feasibility of a traffic monitoring system based on cell phone locations depends on two issues. First, the wireless carriers must be able to produce locations with the appropriate characteristics. Second, the software which analyzes those positions must be able to produce traffic information over a broad area, including correctly locating the phones and adequately selecting sample phones distributed over the whole road network.

The characteristics of the location data suitable for a cell phone based probe system were determined and compared against the capabilities of the current set of technologies being deployed by the wireless carriers. Key characteristics include: the accuracy of locations, the frequency with which the position is updated, and the total number of locations available.

Software to analyze the locations was developed and tested using simulated vehicles traveling over the road network. A driver behavioral model and a heuristic based spatial shift model were developed and incorporated into the location algorithm which dramatically improved the software's ability to assign phones to the correct road segment. The improvement was particularly great for the hardest case, freeway segments, where the percentage of unambiguous freeway segments rose from 68% in the earlier UCB study to 98.9% with the improved algorithm.

Matching the characteristics of the location data stream available from the wireless carriers with the requirements of the analysis software showed the effectiveness of a cell phone based probe system. With a network based location technology, measurements on 85% of the roads can be generated using approximately 5% of the location capacity of a single carrier, in every 5 minute interval. Operating continuously, traffic information coverage generated over time will approach 97.7% of the road segments, the maximum possible with a location technology accurate to 100 meters using a 30 second update frequency. With handset based location technologies, using 5% of the location capacity of a single carrier can generate measurements for over 90% of the road segments in every 5 minute interval. Over multiple time intervals, the 50 meter accuracy of handset based technologies is capable of generating traffic information for 98.8% of the roads using a 30 second update frequency.

Acknowledgement
This material is based upon work supported by the National Science Foundation under Grant No. 0128649. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Yim, Y.B. Youngbin, Randall Cayford. *Investigation of Vehicles as Probes Using Global Positioning System and Cellular Phone Tracking: Field Operational Test*. University of California, Berkeley, 2001.

Ygnace, Jean-Luc; Chris Drane; Y.B. Yim; Renaud de Lacvivier. *Travel Time Estimation on the San Francisco Bay Area Network Using Cellular Phones as Probes*. University of California, Berkeley, 2000.

Guldner, Jurgen; Han-shue Tan; Satyajit Patwardhan. *Analysis Of Automatic Steering Control For Highway Vehicles With Look-down Lateral Reference Systems*. University of California, Berkeley, 1996.

Tsutomu, Imada; Adolf May. *FREQ8PE: A Freeway Corridor Simulation and Ramp Metering Optimization Model*. University of California, Berkeley, 1985.

LIST OF FIGURES

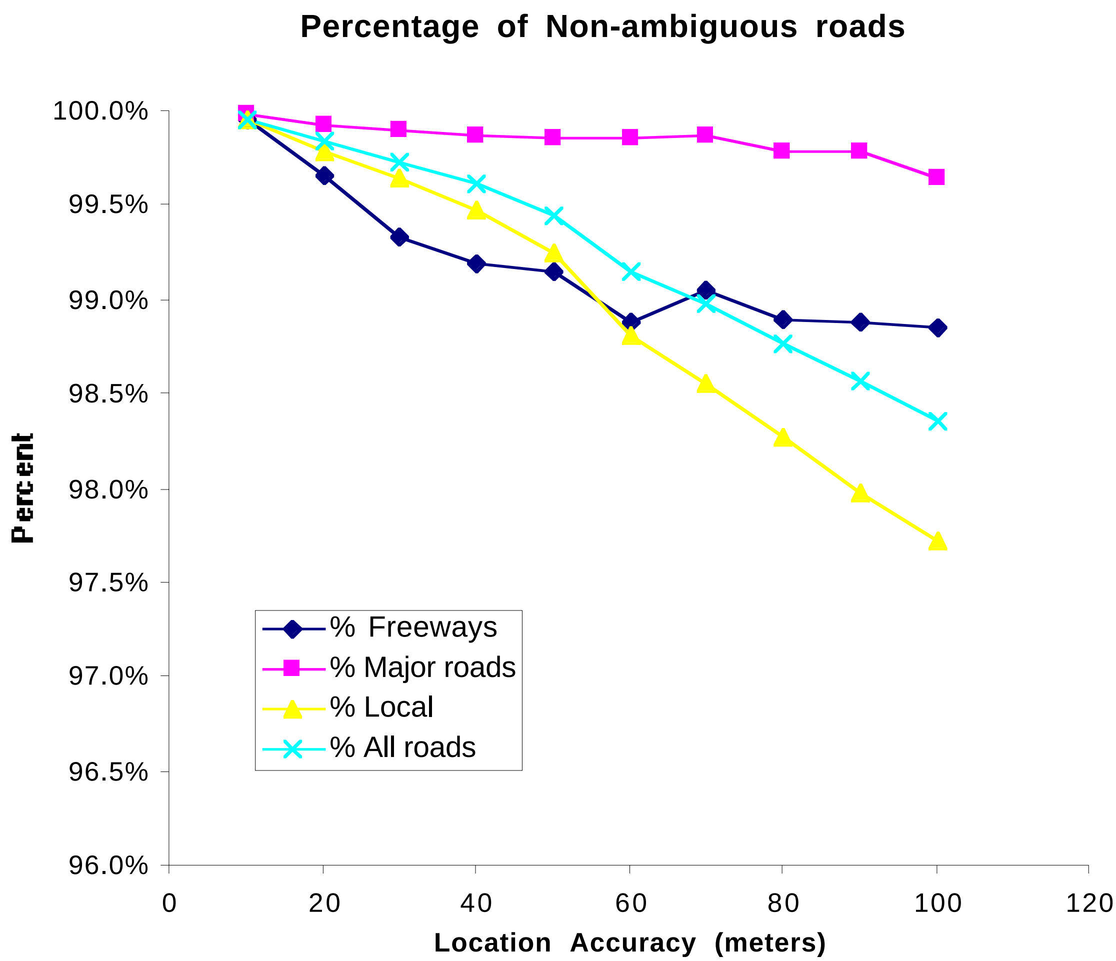

FIGURE 1 – Location Accuracy vs. Road Ambiguity.

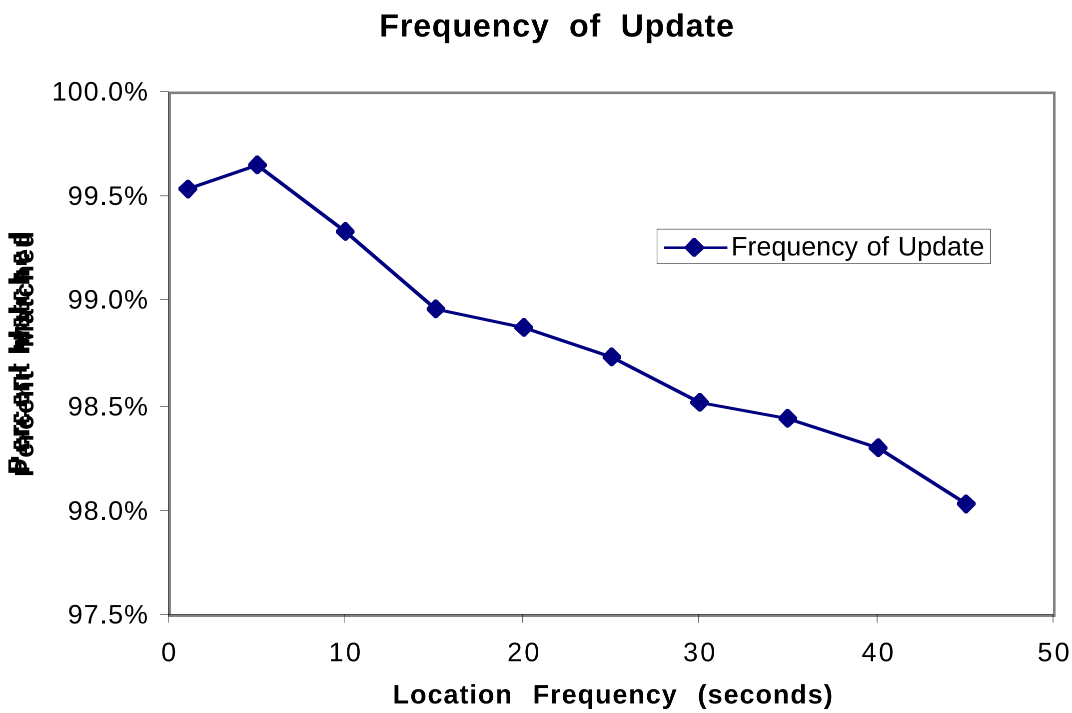

FIGURE 2 – Percentage of Road Coverage vs. Frequency of Update.

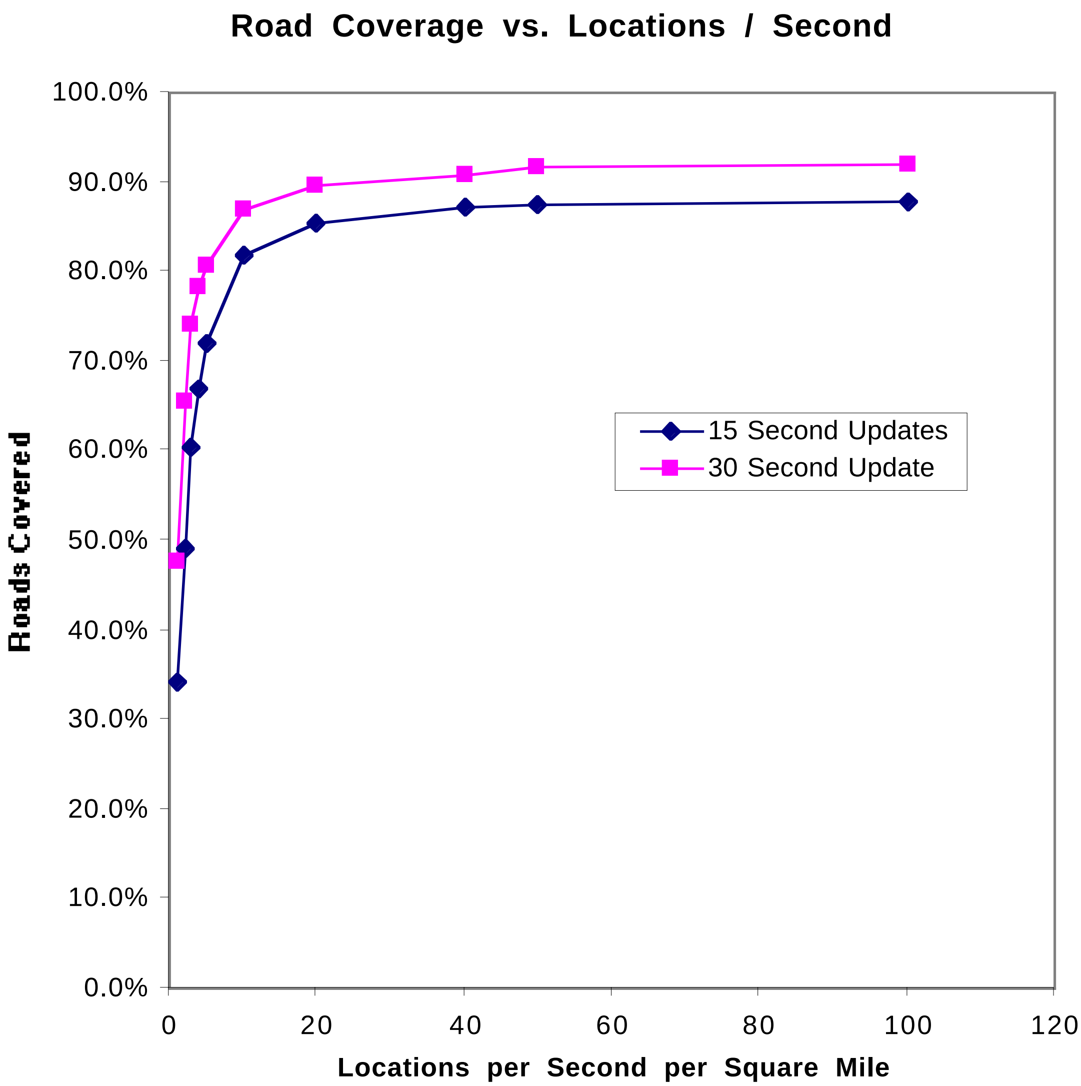

FIGURE 3 – Percentage of Road Coverage vs. Locations / Second / Mile^2.

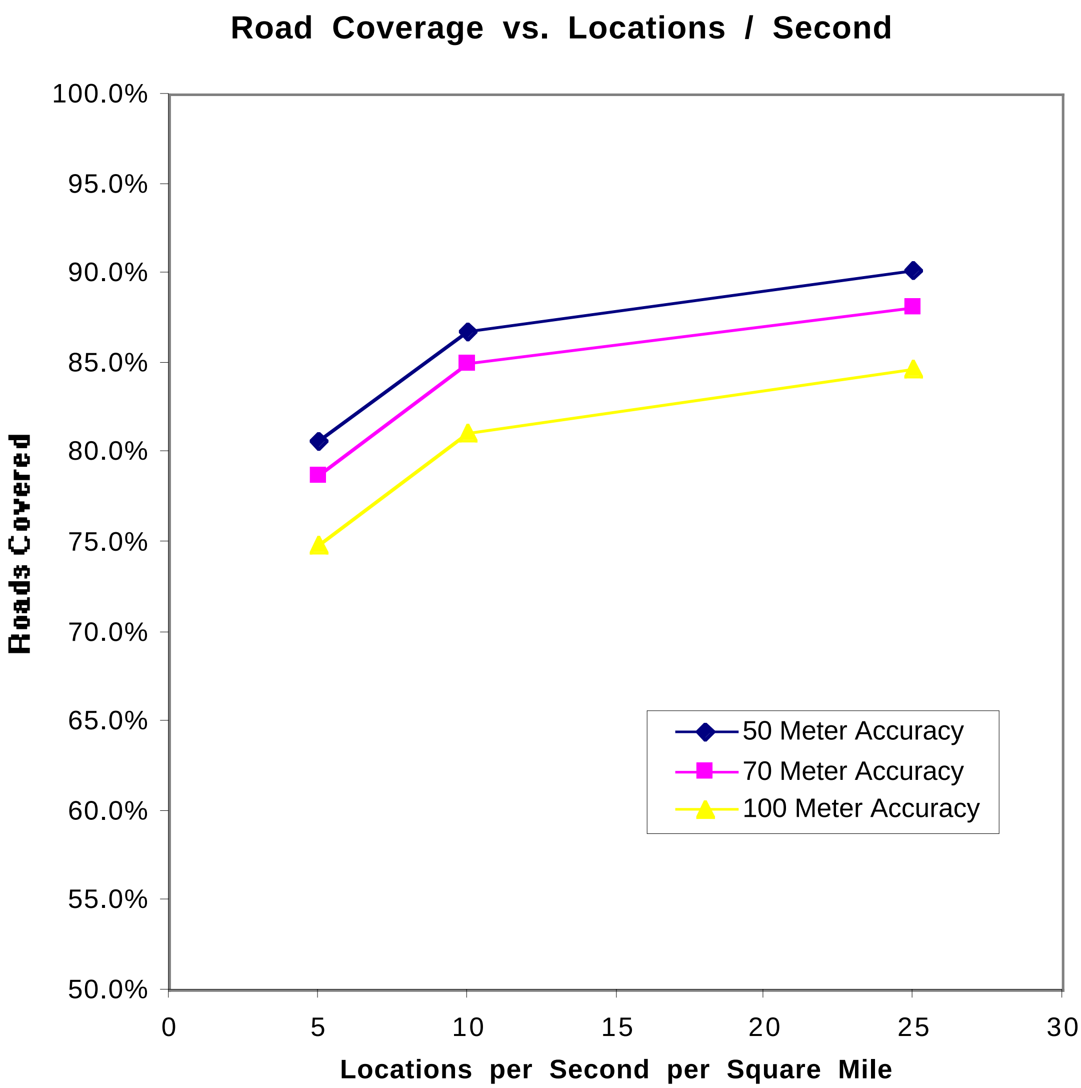

FIGURE 4 – Coverage in 5 Minute Intervals for E911 Accuracies.