FINAL REPORT

IMPROVING THE EFFECTIVENESS OF TRAFFIC MONITORING BASED ON WIRELESS LOCATION TECHNOLOGY

Michael D. Fontaine, P.E.
Research Scientist

Brian L. Smith, Ph.D.
Faculty Research Scientist
and
Associate Professor of Civil Engineering

Virginia Transportation Research Council
(A Cooperative Organization Sponsored Jointly by the
Virginia Department of Transportation and
the University of Virginia)

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

December 2004
VTRC 05-R17

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

EXECUTIVE SUMMARY

Introduction

Reliable and accurate data on traffic flow characteristics are a fundamental requirement for the effective operation of transportation facilities. Data on traffic flow can be used for multiple purposes, ranging from providing traveler information to creating performance measures for the transportation system. The current state of the practice in Virginia, as well as the rest of the nation, is to rely on a network of inductive loop detectors (ILDs) to gather information on traffic flow at fixed points on the roadway network.

Although ILDs are a proven technology, they have a number of limitations. Since loops are actually installed within the pavement, roadway lane closures must be used to maintain the ILDs. The cost to maintain ILDs and their associated communications infrastructure is also significant. As a result, a dense network of point detectors are usually available only on heavily traveled urban freeways. The Virginia Department of Transportation (VDOT) will likely need to expand the size of the network that it monitors to support better the increased emphasis VDOT is placing on system operations.

Traffic monitoring based on wireless location technology (WLT) offers an opportunity to expand the size of the transportation network being monitored at a lower cost than with loops. The Federal Communications Commission is requiring wireless providers to be able to determine the location (latitude and longitude) of wireless devices to help aid emergency response.[1] Since this location information will become available, it may be possible to use these data to create a probe-based traffic monitoring system that anonymously tracks the locations of individual wireless devices. By tracking a series of positions for devices located in vehicles, it is theoretically possible to derive a speed for the vehicle carrying the device. The speeds of these probe vehicles can then be used to determine the average overall travel speed on a road.

WLT-based monitoring is appealing, in part, because of the widespread availability of potential probe vehicles. Industry data indicate that in 2003 more than 70 percent of individuals over the age of 15 owned a wireless phone.[2] Any phone that is turned on can potentially be tracked, so there is a broad potential base of probe vehicles available to this type of system. In contrast to ILDs, WLT-based monitoring uses infrastructure that has already been installed by wireless service providers. Since WLT-based monitoring does not rely on the physical installation of sensors in the pavement, there is the potential to expand traffic monitoring activities to any road that is covered by wireless communications services. This could potentially allow VDOT to monitor conditions on primary and secondary roads that are not currently examined.

Although the concept of WLT-based monitoring is attractive, results to date have been mixed. VDOT has funded two field deployments of this technology,[3,4] and a third is currently underway in Hampton Roads. Unfortunately, the former two deployments have not been able to generate traffic information of the quantity or quality that would be useful to VDOT. Field tests have been successful at locating vehicle positions but have not been able to generate traffic information of the quality or reliability required for most applications. There is a need to gain a better understanding of the factors affecting the effectiveness of these systems and to develop basic guidelines for when they should be used.

Purpose and Scope

VDOT has already invested in several field deployments of WLT-based traffic monitoring systems, but so far, none has been able to produce usable traffic condition data. There is a need to understand what factors influence the efficacy of these systems. By better understanding how these systems work, VDOT may be able to make a preliminary assessment of a particular vendor's system prior to entering into an agreement to deploy the system in the field.

This research investigated the relative importance of system design and roadway network characteristics on the overall performance of WLT-based monitoring systems. This examination was used to determine guidelines for suitable applications of such systems. The objectives of the research included:

- Enhance and evaluate map-matching algorithms to determine their impact on WLT-based monitoring system performance.

- Investigate the role of system design and the roadway network in system efficacy. System design and roadway characteristics that affect errors in speed estimation and coverage will be identified through a controlled evaluation.

- Investigate and quantify desirable WLT system characteristics and determine their impact on system effectiveness on simulated real-world networks. Identify problematic situations where WLT-based monitoring does not perform well and quantify their effects. Determine roadway characteristics that are amenable to monitoring by a WLT-based system and quantify the impact of different roadway situations.

- Examine sampling requirements for WLT-based systems, with an emphasis on determining how well existing sampling concepts for probe systems perform in a WLT environment.

- Develop general guidelines for the design and application of WLT-based systems.

Methodology

The methodology for this project consisted of six major tasks:

1. *Literature review.* Methods to locate vehicle positions were identified, and past deployments of WLT-based traffic monitoring were reviewed.

2. *Development of map-matching algorithm.* Methods to take inaccurate probe positions and match them to the roadway network were developed.

3. *Development of a test bed.* A test bed was developed to simulate the operation of a WLT-based monitoring system. The test bed consisted of a microscopic traffic simulation model, an application that emulated the location output of a WLT system, and processes to generate traffic conditions data.

4. *Testing on simple networks.* Exploratory testing was performed on simple hypothetical networks to gain a better understanding of how roadway and WLT system design factors affect the effectiveness of WLT-based systems.

5. *Case studies on real-world networks.* The results of the exploratory testing were applied to case studies on three simulated networks from Virginia. This testing was intended to examine conditions that were not explicitly evaluated during exploratory testing.

6. *Generation of guidelines for the application of WLT-based monitoring systems.* The exploratory testing and case study results were used to develop broad guidelines for designing and applying WLT-based traffic monitoring systems.

Results

Map Matching

Three forms of map matching were developed for testing in this research, ranging from relatively simple geometrically based methods to more complex statistically based approaches:

1. *Simple geometric map matching.* A simple point-to-curve method was developed. In this method, probe positions were simply projected onto the closest roadway link. This method served as a baseline against which other alternatives could be compared. Although the past WLT deployments often do not explicitly describe the map matching method used, indications are that simple map matching was used in many of those evaluations.

2. *Geometric map matching with topology.* This method built on the simple geometric method but included consideration of link orientation, link connectivity, and vehicle travel history. Vehicle trajectories were compared to link orientations to determine likely paths, and only paths that were physically possible were candidates for matching. A scoring function was developed to select the path that appeared to be both physically possible and the best match with regard to the travel history of a vehicle. This method was expected to offer improved performance over simple geometric map matching.

3. *Probabilistic technique based on the multiple hypothesis technique (MHT).* The MHT method builds on the topology method but adds consideration of multiple path alternatives to improve matching robustness. This method maintained a series of hypothetical paths that were evaluated based on normalized scores created by the cumulative differences in (1) the distance between the position estimates and projected positions and (2) the estimated vehicle trajectory and link orientation. When a score reached a predefined threshold, the matched positions for the best path were retained and all others were discarded. If no path reached convergence, the path with the highest probability was selected when a vehicle left the network or was lost by the system.

Exploratory Testing

The exploratory testing on simple networks revealed several trends in the performance of WLT-based monitoring. The most significant findings include:

- Any WLT-based monitoring system should use a proven form of map matching, such as the topology or MHT method, to identify vehicle positions onto the roadway network.

- Using a relatively infrequent mean time between samples generally improves speed estimation over frequent sampling intervals. Using longer sampling intervals allows the system to gather information over longer distances, reducing the chance of capturing a non-representative speed.

- Large errors in vehicle positions usually translate into larger errors in speed estimation. Location errors should be minimized to improve system accuracy.

- Speed estimation errors are largest when the mean time between location estimates is short and position errors are large. If the system has large errors in positioning that cannot be reduced, it should be designed to use long sampling intervals to mitigate this problem.

- WLT-based systems will need to have the ability to change system parameters, especially the number of vehicles tracked, based on the characteristics of different parts of the network. More vehicles will be required for complex networks than for simple networks to ensure that adequate probe vehicle penetration is achieved. This indicates that the design of a WLT-based monitoring system is not a "one size fits all" problem and that systems will have to be scalable to accommodate location-specific characteristics.

Case Studies

Several recurring trends emerged from the three case studies. Key findings from the case studies include:

- WLT system design needs to be able to be scaled to approximately a single cell. As the area being monitored increases in size, traffic conditions tend to vary considerably across the facilities in the monitoring area. The WLT system characteristics must be able to be tailored to a relatively small scale for the system to be effective. Attempting to monitor large networks without any type of zone-specific sampling will likely result in large gaps in the monitoring system. Random sampling of the entire network will result in a disproportionately large number of vehicles being sampled on congested roads, which often need the smallest number of samples because of their low speed variance.

- The MHT method significantly increased the number of samples available over the topology method, usually by a statistically significant margin. The MHT method should be used to match vehicles to the roadway network.

- Longer sampling frequencies produce better results when there are long, continuous links on the network. Shorter sampling frequencies performed better when link travel times were short, because of short lengths, high speeds, or a combination of both. Again, this points to the need to tailor system characteristics to specific components of the network.

- Data screening methods need to become more robust to solve the problems posed by wrong-way matching between parallel free-flowing and congested links. This problem was most pronounced for the MHT matching methods, but this was primarily because the MHT method generated so many more samples than did the topology method. This provided a greater opportunity for mismatches to create problems.

- Non-vehicular sources do not affect speeds provided their orientation of travel differs from that of the surrounding roadway. For similar reasons, the matching process had difficulty discerning the difference between high-occupancy vehicle (HOV) lanes and the main lanes because their orientations were similar, if not identical.

Conclusions

Map Matching

The results of the testing showed that having a reliable, accurate map-matching method can create statistically significant improvements in the number and accuracy of speed estimates. Important conclusions include:

- Map-matching techniques originally developed for in-vehicle navigation systems can be successfully adapted and applied to WLT-based systems. The algorithms were adapted and modified to ensure that they provided the real-time matching results needed for these applications. Using well-developed map-matching techniques offers the opportunity to improve the usage of data significantly over the simple geometric or empirically driven methods used in previous evaluations.

- The topology and MHT methods performed well and offered significant improvements over the simple geometric map-matching method. These methods could perform well without the input of vehicle-based information and did not have a significant computational load.

- The modified MHT method consistently produced more speed estimates than the topology method in the case studies and should be the preferred form of map matching for these systems since it did not significantly increase computational time. By increasing the number of speed samples, the MHT method usually estimated speeds better than the topology method.

WLT System Design Parameters

Several system parameters were shown to have a direct impact on system coverage and system accuracy. Major findings included:

- Larger errors in position estimation translate into larger errors in speed estimation. Systems should seek to minimize location error if possible, but there are methods to mitigate large position errors if necessary. If needed, the mean time between samples can be extended to help reduce the impact of large errors in location estimation.

- Speed estimation errors are most pronounced when a low mean sampling frequency and high location error are present. Errors are generally minimized by the use of longer mean sampling frequencies. Links that are either short or have high travel speeds are the exception. In those cases, a short mean time between samples should be used to ensure system coverage. The requirements to balance system accuracy and system coverage must be weighed by the designer based on specific characteristics on individual roads being monitored.

Network Characteristics

Several trends were apparent with regard to the types of networks likely to be successfully monitored by a WLT-based system:

- The algorithms had difficulty when there were large differences in traffic flow speed between parallel facilities with similar orientations. Specifically, vehicles on the low-speed road were often mismatched to the higher-speed road. Improved data screening may correct this problem.

- The map-matching process could not successfully distinguish between concurrent facilities with different characteristics, such as HOV lanes and general-purpose lanes. These distinctions could be made where HOV lanes diverged from main lanes, but not when they are concurrent.

- Non-vehicle sources did not generate significant problems, primarily because their orientation was not parallel with the facilities being monitored. The lack of impact by non-vehicle sources can be seen by the fact that speed estimates in the area near these sources did not drop considerably.

Sampling and Accuracy Requirements

Trends in the number of samples and the accuracy of speed estimates were observed. These included:

- The number of samples generated appears proportional to the vehicle hours of travel per mile on a link. Long, congested links act as "sinks" for the samples. This results in large number of samples on slow-moving links and relatively few on faster-moving links. This can create problems with sample size when congested facilities are over sampled and uncongested roads are under sampled. If networks are small or the density of traffic is approximately evenly distributed over the monitored roads, this is less of a problem. More research in this area is needed to ensure that adequate samples are obtained for all facilities.

- More samples than the minimum predicted by the central limit theorem are required to generate good speed estimates. This is likely attributable to the impacts of errors in map matching and potential over sampling of lower speed vehicles.

- The case studies show that WLT-based systems are reasonably accurate. Accuracy levels should be sufficient for the purposes of highway performance measurement or the determination of whether a road is congested.

- In general, the WLT-based monitoring speed estimates compared favorably to estimates by point detectors. Speed estimates for WLT were much more accurate than those of point detectors for arterial systems, but point detection appeared to perform better when speeds were high and traffic flow was uniform (such as with uncongested freeways). WLT-based monitoring has the potential to improve condition estimation on arterial roads significantly since it can capture control delay.

- For large networks, it would be desirable to gain a finer-grained control over system sampling. On large networks, it would be desirable to be able to set specific sampling requirements for different regions of the network to ensure that minimum probe penetration and sampling is maintained based on location-specific characteristics. Likewise, dynamic sample reallocation can be tested to ensure that heavily congested links do not act as sampling sinks. This level of control would likely need to be at the level of an individual cell.

RECOMMENDATIONS

1. *VDOT's Mobility Management Division (MMD) and Smart Traffic Centers (STCs) should continue to support the deployment of WLT-based systems but should be more discriminating in which systems they fund for field tests.*

2. *The MMD and STCs should apply the guidelines proposed in this research to help screen any future WLT-based monitoring systems that are being considered for deployment in Virginia.*

3. *In particular, the MMD and STCs should apply the results of this research to ensure that any future systems deployed in Virginia will be able to:*

   — Document the form of map matching used in their system, and provide proof that their method can accurately project inaccurate vehicle positions onto the roadway network.

   — Provide an estimate of the mean and distribution of location error for the system.

   — Indicate the expected mean time between position samples and its distribution.

   — Detail how vehicles are sampled over a network. Vendors that approach WLT system design with a "one size fits all" approach should be viewed with caution. Likewise, VDOT should ensure that the vendor can sample enough vehicles to meet requirements for probe penetration. The information provided by those vendors should be examined in light of the findings of this research to determine if it is likely that the proposed system would perform adequately in terms of accuracy and coverage.

FINAL REPORT

IMPROVING THE EFFECTIVENESS OF TRAFFIC MONITORING BASED ON WIRELESS LOCATION TECHNOLOGY

Michael D. Fontaine, P.E.
Research Scientist

Brian L. Smith, Ph.D.
Faculty Research Scientist
and
Associate Professor of Civil Engineering

INTRODUCTION

Departments of transportation (DOTs) are placing an increasingly heavy emphasis on the efficient operation of existing facilities rather than the construction of new capacity. The Virginia Department of Transportation (VDOT) has recently established the position of chief of system operations in recognition of this trend. Accurate and reliable information on traffic conditions is a fundamental requirement for many operational initiatives. Traffic condition data allow DOTs to provide traveler information, identify incidents, and create performance measures. Although this type of data is critical if operational performance is to be maximized, the cost of installing and maintaining a sensor network to collect these data often limits the number of roads on which detailed data can be collected.

Traffic condition information is usually collected through a network of inductive loop detectors (ILDs). Although ILDs are a proven technology, they are often difficult and costly to operate and maintain. Since ILDs are located in the roadway pavement, lane closures must be implemented to replace or repair an ILD. This can create delays for the driving public and may act as an impediment to the timely maintenance of detectors. The cost to install and maintain ILDs and their associated communications network is also significant, particularly for large sensor networks. In 2004, VDOT estimates that the cost to install a single ILD station was $40,000, with $7,500 in annual operating and maintenance costs. As a result, closely spaced ILDs are usually installed only on heavily traveled routes in urban areas, creating significant gaps where there is little or no information on system performance.

Monitoring systems based on probe vehicles have been used in several places in the United States to monitor the road network without the infrastructure requirements of an ILD point detector system. In a probe-based system, the speeds of a small subset of traffic are sampled, and these samples are used to estimate a mean speed for all traffic on a road. Automatic vehicle identification (AVI) systems are one of the methods often used in these systems. In these situations, AVI readers are installed at key locations on the network and the time it takes vehicles with toll tags to travel between readers is used to estimate link travel speeds. Although this can produce better point-to-point estimates of link travel speeds than can point detection, it still requires a significant investment in infrastructure and communications.

Traffic monitoring based on wireless location technology (WLT) offers an opportunity to use the probe-based monitoring concept without the infrastructure requirements of AVI-based systems. Recent advances in technology have allowed the location (latitude and longitude) of wireless devices to be determined to within a reasonable degree of accuracy. Federal requirements stipulate that wireless carriers must be able to provide accurate position information for wireless devices so that emergency response to wireless 911 calls can be improved.[1] As a result, wireless service providers are implementing systems to comply with the regulations.

A variety of commercial uses for this location information is also being developed in the wireless provider community. One potential application of this location information is traffic monitoring. If a series of positions for a wireless device in a vehicle can be monitored, those positions could be used to derive the speed of the vehicle that contains the device. By sampling the locations of devices in multiple vehicles, it should be possible to estimate the overall speed of traffic on a particular road. Wireless device position information could allow traffic monitoring systems to use vehicles equipped with wireless devices as probes in the traffic stream, permitting agencies to detect incidents and determine traffic flow characteristics without installing an expensive point detector network of ILDs. These position data could be used to derive average link speeds, travel time information, and possibly origin and destination information for a network.

WLT-based monitoring is appealing, in part, because of the widespread availability of potential probe vehicles. Industry data indicated that in 2003 more than 70 percent of individuals over age 15 owned a wireless phone.[2] Any phone that is turned on, even if it is not in use, can potentially be tracked using WLT, so there is a broad potential base of probe vehicles available to this type of system. Further, the roadway network monitored could be potentially expanded to any facility that has wireless coverage. This means that conditions on primary and secondary roads could be monitored along with the freeways.

Although the idea of WLT-based monitoring is conceptually attractive, past field deployments have met with mixed results. VDOT was a partner in two operational tests of WLT that were conducted in the Washington, D.C., area in the mid-1990s and in 2001[3,4] and is a current partner in an ongoing test in the Hampton Roads region. Although past tests showed that individual wireless device users could be located to within a reasonable degree of accuracy, the tests were unsuccessful in translating the location information into traffic condition data that would be useful to VDOT or the traveling public. In particular, these deployments have had difficulty matching vehicles to roads and determining accurate link travel speeds.

The failure of these deployments illustrates the need for further research related to the use of WLT for traffic monitoring. Previous deployments have relied on empirically derived methods for sampling and matching vehicles to the roadway network. In addition, most deployments and simulation tests have examined only a very limited set of roadway and system design characteristics, making it impossible to separate the influence of system and roadway factors. There has not been a systematic examination of how system design and roadway characteristics affect the ability of a system to produce realistic results. Likewise, many potentially problematic conditions have not been examined. There is also a fundamental lack of understanding regarding how WLT system design would affect the sampling requirements on a roadway network. As a result, there is a need to develop a better understanding of the operation and usage of WLT-based systems so that DOTs can make better decisions on the deployment of this technology.

Figure 1 shows a simplified schematic of how WLT-based monitoring systems work. The two elements to the left of Figure 1, the wireless network and determination of location information, deal with how wireless signaling information is processed and used to determine position information regarding individual wireless devices. These areas are within the domain of electrical and telecommunications engineers. The position processing and traffic data fields cover how the WLT position data are processed and converted into usable traffic data. Developers of WLT-based monitoring systems have generally spent a lot of effort in generating accurate position information (the two components to the left of Figure 1). Although this is a basic requirement of these systems, little effort has gone into developing position processing techniques or defining WLT system factors that affect the overall effectiveness of the system.

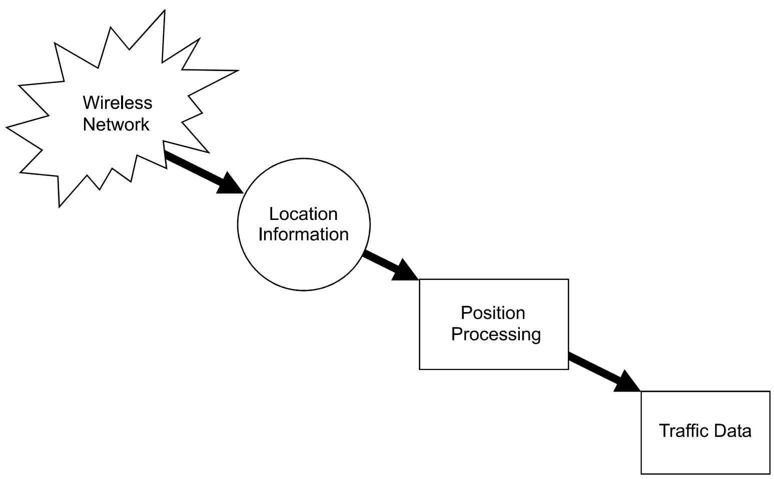

Figure 1. Components of WLT-Based Monitoring System.

PURPOSE AND SCOPE

This study focused on how position data should be processed to generate useful traffic information and investigated the relative importance of system design and roadway network characteristics on the overall performance of WLT-based monitoring systems. The specific objectives of the study included:

- Enhance and evaluate procedures for matching inaccurate vehicle positions onto the roadway network, and determine their impact on WLT-based monitoring system performance.

- Investigate the role of WLT system design and roadway network configuration in system efficacy. System design and roadway characteristics that affect errors in speed estimation and coverage will be identified through a controlled evaluation.

- Investigate and quantify desirable WLT system characteristics, and determine their impact on WLT system effectiveness on simulated real-world networks. Identify problematic situations where WLT-based monitoring does not perform well, and quantify their effects. Determine roadway characteristics that are amenable to monitoring by a WLT-based system, and quantify the impact of different roadway situations.

- Examine sampling requirements for WLT-based systems, with an emphasis on determining how well existing sampling concepts for probe systems perform in a WLT environment.

- Develop general guidelines for system design and application of WLT-based traffic monitoring systems.

The concept of WLT was treated in a technology-independent manner so that the results would not be biased toward any particular vendor. As a result, a simulation-based approach was used to eliminate reliance on any particular technology. The study used only the output of an emulated WLT-based system, which was assumed to include a time stamp of when the location was generated, a unique identification number for a particular wireless device, and estimated latitude and longitude coordinates for a device

These outputs are consistent with what has been produced in past deployments. In the real-world deployments, the identification number is randomly assigned and nothing is reported that could allow a particular device to be identified.

The study did not explicitly cover the privacy and legal implications of using WLT-based traffic monitoring systems. A wide range of commercial applications for wireless location information is being explored beyond traffic monitoring. Given the potential commercial implications of this type of information, a number of wireless providers and potential location-based content providers are investigating privacy and legal issues. In previous systems, wireless providers have stripped all identifying information from a vehicle position prior to sending it to a third party company that produces traffic information. An "opt-in" concept is also being considered whereby users would agree to serve as probes in exchange for free travel information or some other benefit. This is still a developing area, and the final form of the legal framework for these systems has yet to be resolved completely although it appears likely that these issues will be decided.

METHODOLOGY

A large number of factors have the potential to affect the accuracy of the vehicle position and traffic condition estimates generated by a WLT-based monitoring system. Past evaluations have made little attempt to quantify the impact of the interactions of roadway network geometry, WLT system design parameters, and traffic conditions on overall system effectiveness. To investigate these issues, a methodology consisting of six major tasks was developed:

1. literature review
2. development of a map-matching algorithm
3. development of a test bed
4. exploratory testing on simple roadway networks
5. case studies on real-world roadway networks
6. generation of guidelines for the design and application of WLT-based monitoring systems.

Literature Review

The Transportation Research Information Service (TRIS), the University of Virginia library, and the Virginia Transportation Research Council library were consulted for the literature review. The literature review was performed to gather information in the following topic areas:

- operation of cellular networks
- wireless location technology operation
- past deployments and simulation studies of WLT-based traffic monitoring
- map-matching techniques
- probe vehicle sampling strategies

Development of a Map-Matching Algorithm

Before evaluations could be conducted, it was necessary to develop methods to match inaccurate vehicle position estimates to the roadway network. Map-matching procedures that were originally developed for in-vehicle navigation systems were reviewed for possible application to WLT-based traffic monitoring systems. The map-matching techniques from the literature review were evaluated based on the following criteria:

1. ability to be extended to track many vehicles simultaneously
2. likely computational demands
3. ability to match vehicle positions to the roadway correctly.

In some cases, these objectives were at odds, particularly the tradeoffs between computational demands and accuracy.

Once candidate methods were identified, they were modified to make them appropriate for application in a WLT-based monitoring system. In most cases, this involved changes to the algorithms to remove factors that required information that would be generated by the vehicle, such as acceleration or dead-reckoning direction information. Algorithms were adapted to make matches based purely on known roadway characteristics, system design parameters, and WLT position estimates.

Development of Test Bed

A major task in this research was developing a test bed that was capable of accurately simulating a WLT-based monitoring system. A test bed had to be created that could perform the following functions:

1. *Accurately represent traffic flow on a network and report actual vehicle positions and travel times.* This involves being able to accurately represent "true" conditions on a road.

2. *Create a digital network that could be used for map matching based on network characteristics.* This network is similar to a digital roadway centerline map that is readily available to most transportation agencies.

3. *Degrade the accuracy, number, and frequency of vehicle positions based on desired test parameters to simulate the output of a WLT-based monitoring system.* This emulates the raw location information that would be generated by a WLT-based system.

4. *Match the degraded data to the roadway network using the map-matching methods developed earlier.* This involves the processing of the raw position data to determine estimated locations for vehicle probes on the roadway network.

5. *Generate speed estimates based on the new, matched positions and compare them to the actual speeds on the road.* This uses the processed data to generate traffic estimates.

Figure 2 shows the basic operation of the test bed developed. The process consisted of two major parts: the simulation of the roadway network using microscopic simulation and the simulation of the operation of a WLT-based monitoring system. The specific parts of the process related to simulating the operation of a WLT-based system are noted in the figure using dashed boxes. The traffic simulation provides baseline comparison data and inputs to the simulated WLT-based monitoring system.

The microscopic traffic simulation VISSIM was used to simulate the roadway network being evaluated, and it recorded the true position of each vehicle on the roadway network every second. Positions were expressed in X,Y coordinates from a user-defined origin. User-defined parameters were used to degrade the positions to simulate the output of a WLT-based monitoring system. The degradation process involved intentionally introducing errors into the position estimates, increasing the elapsed time between position estimates on a vehicle, and reducing the number of vehicles being monitored at any given time. The values of these factors were based on expected WLT performance characteristics as dictated by technology limitations and potential

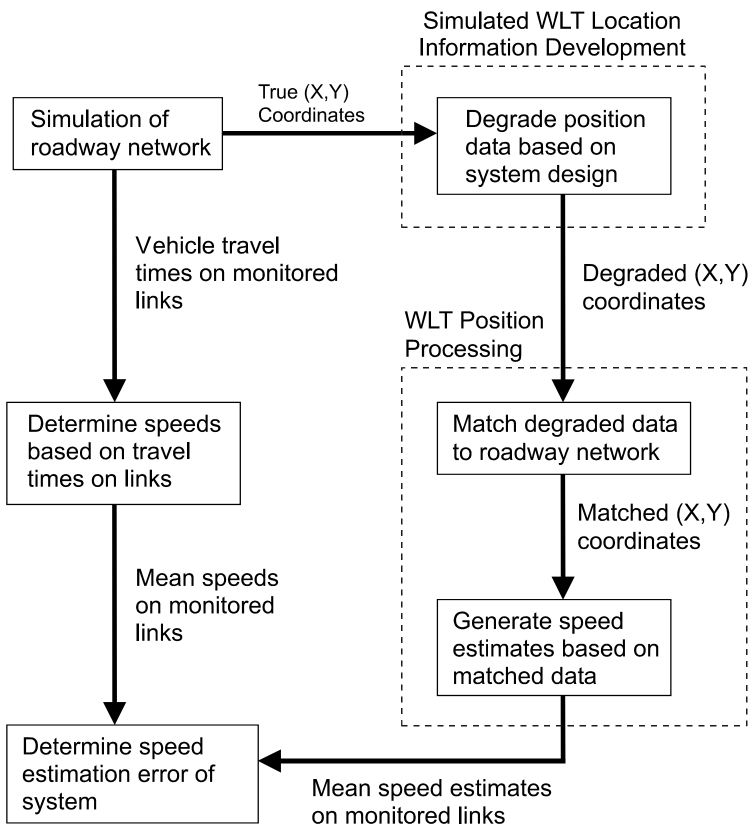

Figure 2. Test Bed Design.

demands on the wireless network. This degraded position data set represents the raw output that would be achieved from an actual WLT-based traffic monitoring system.

Since these positions contain error, there is ambiguity about the actual location of a vehicle on the roadway network. As a result, these positions must be re-matched to the roadway. These matched positions can then be used to create speed estimates, which are compared to the speeds generated by the original VISSIM simulation. This provides a measure of the errors in speed estimation created by using WLT-based monitoring systems.

The operation of the test bed can be divided into four tasks: generation of traffic data, creation of the matching network, degradation of traffic data, and position matching and speed summarization.

Generation of Traffic Data

Roadway networks were coded using the VISSIM model based on the conditions required by the testing. The simulations were used (1) to compare the estimates generated by the simulated WLT-based monitoring system to the “reality” of the simulation and (2) to generate position data for simulated probe vehicles. Multiple replications of each simulation were not required since the goal of the evaluation was to determine whether the monitoring system was capable of determining the traffic flow in the simulation, not to simulate the actual "typical" performance of the road. Data were simulated for 2 hours, but only the data from the second hour was used for analysis to allow the network to reach equilibrium. VISSIM was used to generate two data files: a summary of travel times and a listing of vehicle locations.

VISSIM allows the user to specify links where travel times are to be measured. These travel times represent the actual time it took for a specific vehicle to travel between user-defined start and end points. Travel time information was generated using this function to determine the ground truth for traffic flow on the network. The information was collected on an individual vehicle basis and aggregated into 5- or 15-minute intervals. These average travel times were then translated into average speeds over the link based on the link length. The speeds generated by the system were compared to these speeds to determine the errors in speed estimation.

VISSIM was also used to generate a file consisting of vehicle positions. The location of every vehicle on the network was recorded every second. Each record consisted of a vehicle identification number (VIN), a time stamp, and the X and Y coordinates of the vehicle (in meters from a user-defined origin). This represented a very dense set of position data that often consumed several gigabytes. This position file was used as the basis for the simulation of the WLT-based monitoring system.

Creation of Matching Network

One critical component of the test bed program was the ability to simulate and match to any roadway network. Given the wide variety of possible network configurations, it was necessary to be able to generate a matching network based on the VISSIM input file. The file consists of a series of text inputs, which includes coordinate information for all roads in the network. As a result, this file could be parsed to generate a matching network.

A program was designed to parse the VISSIM text input file to identify link and connectivity information. The parsing process was concerned with VISSIM links and connectors. In VISSIM, links represent roads with consistent cross sections, and connectors define possible turning movements between links. The start and end coordinates for all links are included in the VISSIM input file, and connectors note which links are connected. Direction of travel is implied by the order in which the coordinates are presented.

The program would identify the start and end coordinates for each link on the network and write those to a file. A separate listing of connectors was also created to determine which links were connected to one another. The connector information was used later when the map-matching algorithms had to determine if a path was physically possible between two location estimates. This information was used by the program to trace possible routes between two position estimates. The final product of the parsing process was a complete listing of link coordinates and directionality, coupled with a listing of connectivity information.

Degradation of Traffic Data

A C++ application was developed to translate the dense VISSIM vehicle coordinate file into a file that would simulate the output of a WLT-based monitoring system. Position estimates generated by WLT systems are unlikely to be perfectly accurate, so the quantity and accuracy of the VISSIM positions had to be significantly reduced to emulate a WLT-based system. The data were degraded based on several expected system design characteristics and technology limitations. The application allowed the user to degrade the data based on the following factors:

- *Number of vehicles to track simultaneously.* This factor was used to specify the maximum number of vehicles the system tracked at any given point in time. The program would use the unique VIN to ensure that the proper number of vehicles was maintained. When a vehicle left the network or its position was lost, another vehicle would be randomly picked up somewhere on the network so that the same number of vehicles would continue to be tracked. In the real world, this factor is a function of system design characteristics and practical limitations on the load of the cellular network.

- *Time between vehicle positions.* The user specified the mean and standard deviation of time between position readings on an individual vehicle based on a normal distribution. Using this distribution, the program randomly determined how long it would be between position estimates for a specific vehicle. Intermediate position readings were then removed, and the total number of position estimates was reduced from the 1-second time between readings of the original coordinate file. Since a stochastic distribution was used, the time between readings on an individual vehicle varied between consecutive readings. This factor would be primarily a designed characteristic of the system, although concerns about load on the wireless network might factor into how it is set.

- *Error of vehicle positions.* The X-Y coordinates of each remaining vehicle position were degraded according to a normal distribution of error. The user specified the mean and standard deviation of error. The error of the position estimates is likely to be driven by technology limitations.

- *Probability of instantaneously losing a vehicle.* The user specified a probability that a vehicle was no longer tracked after reading a particular position. This was intended to simulate situations where a call is dropped or a user turns off his or her wireless device. When this occurs, a new vehicle is picked by the system to be tracked to replace the vehicle that was lost.

The levels tested for each of these factors are described in the experimental design section later. Once the data were degraded, a file consisting of the degraded position estimates for each monitored vehicle was created. This file was similar to the one generated by VISSIM in that it consisted of the VIN, time, and degraded X and Y coordinates.

It should be noted that this system was not designed to pick probe vehicles based on their location on the network. Vehicles were randomly sampled somewhere on the network and then tracked until they were no longer on the network or their position was lost. The method evaluated is similar to the manner in which previous deployments and simulated systems have operated. In those cases, the system would retain particular VINs for vehicles being monitored and track only those vehicles.

Position Matching and Speed Summarization

Map Matching

Following data degradation, the program produced a file that consisted of degraded position estimates for a subset of vehicles on the network at relatively infrequent time intervals. This output data file represented the raw position data that would be used as an input into a WLT-based traffic monitoring system. Given the inaccuracies in the vehicle positions, there are usually ambiguities about the true position of the vehicle on the roadway network. The map-matching techniques developed earlier were used to attempt to determine the probable location of a vehicle on the network.

The user of the application was asked to determine the form of map matching that should be applied to the data. Based on the user's input, the appropriate algorithm was applied and a new set of matched positions was created based on the algorithm results. The position readings would either (1) lie somewhere on the matching network created earlier or (2) be noted that they could not be successfully matched to the network.

Speed Summarization

The matched position data were translated into speed estimates. When a vehicle had two location estimates that lay upon links with the same road name, a speed estimate was generated. The speed estimate was created by dividing the distance traveled along the network by the time elapsed between readings. To create the distance traveled, the program summed the distance between the two points by tracing the path along the network.

Speeds were not generated when a vehicle was determined to have turned onto a different road. In this case, it became difficult to determine what portion of time was spent on each facility, so the program did not attempt to assign time traveled to the different roads. The speed data were then summarized for each link monitored. Speed data were then filtered to remove vehicles that were traveling more than 100 mph. This filter was intended to remove those vehicles that were determined to be traveling at speeds that were obviously erroneous because of the impacts of map-matching and location error. Mean speeds for a link were then calculated using the remaining speeds. It should be noted that all speed estimates on a link were weighted equally and a single vehicle could generate multiple speed estimates on a link. The speeds generated by the program could then be compared to those produced by the original VISSIM model. It should be noted that the speeds generated by the simulated WLT system were based on sampling a small subset of vehicles whereas the VISSIM baseline was produced based on the travel speeds on all vehicles on a link.

Exploratory Testing on Simple Networks

The test bed program was used to explore the impact of a variety of system design and roadway factors on the potential accuracy of a WLT-based monitoring system. In this task, simulation was used to examine the impacts of the system design on geometrically simple networks. The test networks were composed of a relatively small number of links with varying numbers of intersections and parallel roads. The purpose of this testing was to determine which factors generally have an impact on the ability of a WLT-based system to generate accurate estimates of traffic conditions. Simple networks were used in this stage to evaluate a large number of scenarios and to help differentiate roadway network impacts from system design impacts. The results of this analysis were used to help define the roadway and system factors that are important to the overall performance of a WLT-based system. This analysis was not intended to define definitive performance functions but rather to provide an indication of general trends in performance related to different factors.

Experimental Design

For the exploratory testing, it was desirable to investigate the main effects and interactions of a wide range of factors to determine how a broad base of WLT system, technology, and roadway network characteristics could impact the effectiveness of WLT-based monitoring systems. As a result, the challenge was to find an experimental design that would allow exploration of a relatively large number of factors but limit the number of trials to a number that was manageable.

The first design issue that had to be examined was homogeneity of the experimental units. Completely randomized designs assume that all experimental units are homogeneous with respect to their effect on the response variable.[5] For this testing, the experimental units are links on the roadway networks being evaluated. There are obvious potential differences in how different network types could impact the accuracy and availability of speed estimates, mainly because of differences in speed variance and network configuration. For example, WLT-based systems are likely to provide different performance on an urban grid system and an isolated freeway. Therefore, it could not be assumed that the effect of the experimental units on the response variables was homogeneous for all network types.

As a result, it was necessary to block the data by roadway network configuration. It was reasonable to assume that the impact of network type on individual links would be relatively uniform for a particular type of network. By using the network type as a blocking variable, homogeneous blocks were created for the experiment. Blocking allows for differences attributable to experimental units to be removed from treatment contrasts and removes variability attributable to heterogeneous groups from the experimental error.[5]

The second design issue that had to be examined was the number of treatments to be examined and the total number of trials required. A 2^k factorial design was initially selected so that a wide range of variables could be examined to determine underlying trends in response. Only two levels for each factor were investigated to minimize the number of trials required while still providing indications of response trends. With a factorial design, the number of combinations of treatments can increase very rapidly as new factors are added. As a result, an experimental design was sought that would maintain experimental power while reducing the number of individual combinations of treatments that had to be examined.

A confounded incomplete block design was eventually selected for this study. This type of experimental design confounds block effects with several interactions so that it is impossible to attribute the degree to which each one impacts the response variable. Determining block effects was not seen as a critical issue, since the specific roadway network characteristics tested would not necessarily be directly extendable to other networks. The factor interactions confounded with the block effects are typically higher order interactions that have less physical meaning than lower order interactions.[5] As a result, the confounded incomplete block design controls for the block effects of roadway network type and reduces the number of runs at the expense of being able to determine the impact of several higher order interactions. These higher order interactions rarely have a practical meaning, so this was not seen as an issue.

The factors examined in the exploratory analysis were broken up into blocking factors and treatments. For the purposes of this experiment, factors relating to network geometric configuration were used as blocking factors. By using network characteristics as blocking factors, the effect of roadway geometry was controlled in the experiment. Factors connected to the WLT system design, potential technology limitations, and how the network is represented digitally were used as treatment factors. These experiments were repeated for each map-matching algorithm evaluated.

Blocking Factors

Four network types were used as blocks in this experiment. Although all cases were geometrically simple, each network represented a different combination of parallel and intersecting facilities. The network types investigated were:

1. *Isolated freeway.* A 10-mile-long isolated freeway with no interchanges or ramps was simulated. This freeway had two lanes in each direction.

2. *Two parallel freeways.* Two 10-mile-long freeways with no interchanges or ramps were simulated. The freeways were placed 1,000 feet apart and had two lanes in each direction.

3. *Arterial road.* A 10-mile-long arterial with signalized intersections every 0.5 mile was simulated. The arterial had two lanes in each direction, with an additional left-turn bay at intersections. Cross streets were a single lane, with an additional left-turn lane at intersections.

4. *Signalized grid system.* This network was similar to the arterial road, but in this case there were two parallel arterial roads placed 1,000 feet apart. Cross streets were still placed every 0.5 mile.

The distances between parallel facilities and cross streets were selected based on situations that could be encountered in the real world. A spacing of 1,000 feet between roads represents a reasonable distance between a freeway and a parallel frontage road, which could represent a challenging situation for the map-matching algorithms. A spacing of 0.5 mile between intersections could also be found in an urban area and could also create problems in map matching.

Several factors remained constant throughout all of the networks. The mean desired speed was 55 mph in all cases, although the actual traffic speeds were dictated by volume and traffic control conditions. Cross streets carried 30 percent of the traffic on the mainline, and signalized intersections were optimized using the Synchro computer program for the specific traffic volumes. At each intersection, 80 percent of traffic went through, and 10 percent turned left and right. Exclusive turn bays were present at all intersections.

Treatments

Six treatments were evaluated for each map-matching technique. Again, each factor was evaluated at two levels. The factors evaluated were:

1. *Traffic volume (Q).* Volumes of 500 and 1,500 vehicles per hour per lane (vphpl) were evaluated on the mainline. These volumes were selected to represent uncongested and near/over capacity conditions on the network.

2. *Number of vehicles that can be tracked simultaneously (N).* Maximum readings of 25 and 100 vehicles were used. This is based on 5 percent and 20 percent of current wireless system capabilities as reported in a study by the University of California at Berkeley.[6]

3. *Frequency of readings (F).* The mean time between readings was set at a mean of 15 and 60 seconds. A standard deviation of 10 seconds was assumed for these readings. These values were based on research that defined these levels as the maximum and preferred time between readings identified by wireless providers.[6]

4. *Probability of losing a vehicle (P).* The probability that a vehicle will be lost and not recovered was 1 percent and 10 percent after each reading. "Lost" means that a vehicle will no longer be tracked. The probability factor is intended to simulate conditions such as a call being dropped or a person turning off his or her phone. This is being tested to determine if the length of time that a vehicle is tracked is an important factor in the overall system accuracy.

5. *Error characteristics (E).* For initial testing, a normal distribution of error was assumed. The error distribution has a mean of 0, and standard deviations of 10 and 150 m. These standard deviations represent global positioning system (GPS) error and E911 requirements for network based WLT systems.[1]

6. *Length of links (L).* Links with an average length of 0.5 mile and 5 miles were examined. These link lengths represent very short link segments typical of urban areas and longer links representative or rural areas. It should be noted that the length of links is partially a function of how the network is represented digitally and could be altered by the user.

Incomplete Block Design

Table 1 shows the design of this evaluation, which was created after several texts on experimental design were consulted.[5,7,8] Since each factor was tested at two levels, a combination of treatments is represented by showing whether a factor is present at its "high" or "low" level. If the factor is present at its "high" level, a letter corresponding to the factor was used. Factors present at the low level are not shown. For example, if variables A, C, and D were present at their high level and variables B and E were present at that their low level in a block, then the treatment combination would be represented as "ACD." The notation (1) represents all variables present at their low level. The following variables are used in Table 1:

- Q = traffic volume
- N = number of vehicles to be tracked
- F = time frequency between readings
- P = probability of losing a vehicle
- E = position error
- L = length of link.

Table 1. Experimental Design of Simple Networks

Block 1: Isolated freeway, no intersections	Block 2: Signalized arterial	Block 3: Parallel freeways, no intersections	Block 4: Signalized grid
• F	• Q	• QF	• (1)
• QNF	• N	• NF	• QN
• QP	• FP	• P	• QFP
• NP	• QNFP	• QNP	• NFP
• QE	• FE	• E	• QFE
• NE	• QNFE	• QNE	• NFE
• FPE	• QPE	• QFPE	• PE
• QNFPE	• NPE	• NFPE	• QNPE
• L	• QFL	• QL	• FL
• QNL	• NFL	• NL	• QNFL
• QFPL	• PL	• FPL	• QPL
• NFPL	• QNPL	• QNFPL	• NPL
• QFEL	• EL	• FEL	• QEL
• NFEL	• QNEL	• QNFEL	• NEL
• PEL	• QFPEL	• QPEL	• FPEL
• QNPEL	• NFPEL	• NPEL	• QNFPEL

Q = traffic volume, N = number of vehicles to be tracked, F = time frequency between readings, P = probability of losing a vehicle, E = position error, L = length of link.

This design confounds three four-factor interactions with the block effects. The following interactions are confounded with the block effects in this design: Q×N×F×L, F×P×E×L, and Q×N×P×E. Since the four-term interactions are unlikely to have much practical significance, this design should identify all effects of interest.

Evaluation Metrics

Two principal areas are of concern with WLT-based monitoring: system coverage and system accuracy. *System coverage* refers to the availability of speed estimates on the monitored roadway network, whereas *system accuracy* denotes how well the WLT-based system estimates the true travel speed on the road. Systems need to perform well in both areas to be useful. The metrics used to measure these areas were (1) percentage of intervals with data, (2) number of speed estimates per link per interval, and (3) accuracy of speed estimates.

The accuracy of the speed estimates variables was examined using the experimental design described earlier. Summary data for system coverage measures are also presented, and additional analyses of system coverage were performed during the case studies of the networks.

Definition of Preferred Parameter Sets

The analysis results were examined to determine the system parameters that produced the best overall performance in terms of system coverage and accuracy. Roadway network situations that created problems were also noted. The error distributions and system coverage generated by candidate parameter sets were also examined. These comparisons were used to define the parameter sets that would be tested and evaluated during the case studies.

Case Studies

The preferred parameter sets defined during exploratory testing were evaluated on three simulated real-world networks. The purpose of this testing was to examine how the system designs identified earlier handled geometric and traffic conditions that were more representative of reality, as well as to determine if they could accommodate potentially problematic conditions that were not explicitly tested earlier. The characteristics of the three case studies are summarized in Table 2.

Table 2. Summary of Case Study Characteristics

Location	Major Roads	Centerline Miles Simulated	No. of Signalized Intersections	No. of Freeway Interchanges	Centerline Miles Monitored
Charlottesville	US 29, US 250	11.33	10	2	3.47
Tyson's Corner	SR 7, SR 123, I-495, Dulles Toll Road	85.45	32	7	9.84
Springfield	I-95, I-395, I-495, Franconia-Springfield Parkway, Franconia Road, Van Dorn Street	80.76	41	5	14.39

Development of Case Studies

The three locations noted in Table 2 were selected because (1) they represented a variety of network and traffic characteristics and (2) the input data to build the models were readily available. In the case of the Charlottesville simulation, a calibrated simulation that had been developed for the Thomas Jefferson Planning District Commission was available for testing. This model already incorporated all appropriate volume and traffic control features for the afternoon peak hour and required no modifications.

For the two Northern Virginia cases, the simulation models had to be created from scratch. The cases were developed using 1-foot resolution aerial photos obtained from VDOT's Northern Virginia District and the relevant Synchro files generated by the district for those areas. The Synchro files contained the signal timings and turning movement counts for all signals in the simulated areas for the afternoon peak hour. The aerial photos were imported into VISSIM, and the roadway network was created by drawing the simulated roads on top of the roads in the aerial photos. The Synchro files were merged with the VISSIM network created from the aerial photo to generate the signal network. Traffic volumes were determined from the Synchro files for the arterial system, and VDOT's Traffic Monitoring System was consulted to determine approximate freeway volumes.

The two Northern Virginia cases were not calibrated and are not perfect representations of actual travel conditions at those locations. Freeway ramp volumes were usually not available, so it was assumed that turning movements on ramps would be proportional to the intersecting mainline volumes. Likewise, freeway counts were sometimes older than the intersection turning movement counts. As a result, the models are not a perfect representation of reality, although the simulation results were visually verified to make sure that the animation appeared reasonable. Since the goal of the project was to test whether the simulated WLT-based monitoring system could reproduce speeds, deviations from what one would actually see in Northern Virginia were not seen as a significant issue so long as the VISSIM results could be replicated by the WLT system and there were no obvious flaws or errors in the simulation.

The Northern Virginia cases were also used to assess conditions that were not explicitly tested earlier that could create problems for the WLT-based system. Examples are:

- ability to deal with non-roadway cellular sources, such as people talking on phones in office buildings or shopping malls
- ability to distinguish between adjacent high-occupancy vehicle (HOV) lanes and general purpose lanes
- influence of heavy rail public transportation
- ability to separate free flowing roads from slower parallel facilities.

All of these factors could be explicitly simulated within VISSIM. For example, pedestrians could be simulated walking around the network and a Washington Metro line was present in the Springfield case. The locations of these elements were recorded by VISSIM along with passenger vehicles, so those elements might also be sampled and tracked by the simulated WLT system.

Travel time and position data were generated for each network as in the earlier exploratory testing. Interstates and major primaries (numbered U.S. and state routes) were selected for monitoring in all cases. For the Charlottesville case, Hydraulic Road was also selected for monitoring. These roads were selected because it was considered reasonable that VDOT would be principally interested in monitoring the primary system.

Data Degradation and Matching

The position data were degraded in accordance with the preferred parameter set defined during exploratory testing. Although this parameter set provided some guidance, there was still ambiguity in some of the factors, particularly in the number of vehicles to be tracked.

Given that some of the case studies were on very large networks, it was necessary to estimate the number of vehicles to be tracked to ensure adequate probe coverage on the network. The number of vehicles tracked was initially based on the central limit theorem (CLT) requirements for the link with the largest standard deviation on the network, inflated to cover the entire network. The number of vehicles to be tracked was determined as follows:

$$N_{i,t} = \frac{\sigma_{i,t}^2 z_{\alpha/2}^2}{d^2} \times \frac{L_{network}}{L_i} \times \frac{F}{Int}$$

where N = number of vehicles to be tracked
i = link number
t = time (seconds)
d = desired accuracy level, assumed to be ±10 mph
σ = standard deviation of speeds for a particular link and interval (mph)
$L_{network}$ = centerline miles of roadway in entire network
L_i = length of link i (miles)
F = mean time between position samples (15 or 60 seconds)
Int = analysis interval length (300 seconds).

This equation inflates the CLT-based sample estimate based on the ratio of the monitored link length to the total network length and the ratio of mean time between readings to the analysis interval. This method assumes a uniform distribution of vehicle-hours of travel over the network, which represents a simplification of reality. The number of vehicles to be tracked was calculated in this manner for each interval for each link. The interval that created the maximum number of samples for each analysis interval size and mean time between samples was selected for testing on the networks. As a result, an N value was determined for two different cases: F = 15 and F = 60.

It was expected that this modification of the CLT-based sampling requirements might not produce adequate system coverage on the network because of errors created by WLT position estimation and errors in map matching. As a result, sensitivity analysis was performed for N. Each N value was then tested at four levels: 100, 125, 150, and 200 percent of what was calculated using the equation.

For the purposes of this research, a desired accuracy of ±10 mph was used to determine the CLT-based sample size. This accuracy level was selected because many traffic management and performance measure applications are more concerned with whether traffic flow is congested or uncongested, rather than an absolute estimation of travel speed. An accuracy of ±10 mph will give an indication of the general traffic flow conditions of the site.

Data Analysis

Once again, both system coverage and system accuracy were examined in the case studies. The following measures were examined for all potential optimal parameter sets on each network:

1. number of speed estimates on monitored links
2. percentage of intervals where the CLT-required sample size was met or exceeded
3. accuracy of speed estimates
4. processing time.

These measures were intended to illustrate the relative performance of the WLT-based monitoring systems. The percentage of intervals where the sample size exceeded the CLT-based requirement for a road was examined to determine whether previously developed sampling guidelines applied to WLT-based systems.

The processing time it took to match vehicles to the network and generate speed estimates was also examined. All matching was performed on a desktop PC with a 1.4 GHz processor. This was done to compare the processing times of different matching alternatives and to ensure that the matching could be performed in less than real time.

These variables were summarized for each set of parameters tested. The percentage of intervals where the CLT-required sample size was met or exceeded and the processing time information are presented in summary form, and the number of speed estimates and speed accuracy were analyzed using Fisher's least significant difference (LSD) test. The LSD test was used to perform pairwise comparisons of the means created by different parameter sets to determine which were significantly different.[5] This made the analysis fine-grained and helped to distinguish potential parameter sets that would be appropriate for different conditions. The LSD test was selected because it was not as conservative as other pairwise tests, making it more likely to detect significant differences. This test was performed using an $\alpha = 0.05$.

Speed estimates generated by simulated point detectors were also compared to WLT-based estimates and actual travel speeds. This was not a rigorous analysis but rather was intended to illustrate some of the differences between WLT-based monitoring and point detection. Areas of application where each method could be used were noted.

Generation of Application Guidelines

The exploratory testing and case study results were examined to determine factors that influence the successful operation of a WLT-based system, as well as conditions that could cause problems for these systems. These results were used to develop some guidelines for the design and application of WLT-based systems. It was hoped that these guidelines could be used by departments of transportation (DOTs) to help screen potential vendors of this technology to ensure that only systems that are likely to perform adequately will be deployed.

RESULTS

Literature Review

The literature review was conducted to gather information on several areas that are important for WLT-based traffic monitoring. The results are grouped into five general topics:

1. operation and design of cellular networks
2. wireless location technology
3. evaluations of WLT for traffic monitoring
4. map-matching methods
5. probe vehicle sampling issues.

Cellular Network Operation and Design

The concept of channel and frequency reuse forms the basis of cellular communications. In cellular systems, there are two channel types: voice transmission channels and control channels. Control channels transmit information relating to call initiation, requests for service, and handoffs between base stations, whereas voice channels carry actual voice communications.[9] Cellular systems are able to cover large areas through reuse of channels in a relatively small geographic area. This is accomplished by assigning a subset of available channels to small geographic divisions within an area, known as cells.[9] Different groups of channels are assigned to adjacent cells so that there will be no overlap in channel usage between nearby cells. Base stations within a cell are designed to provide cellular coverage only within a single cell, so channels can be reused by other base stations provided there is a large enough distance between stations that there will be little or no interference between calls on the same channel in different cells.

Figure 3 provides a simplified example of this concept. In Figure 3, each hexagon represents an individual cell, and each letter represents a different group of channels that are used by that cell. It can be seen that cells with the same channel grouping are never adjacent. In reality, the shapes of the cells are not so regular and there is often overlap between adjacent cells. Likewise, cellular coverage areas are dynamic and can vary depending on network load, site-specific topographic characteristics, and other external factors.[10]

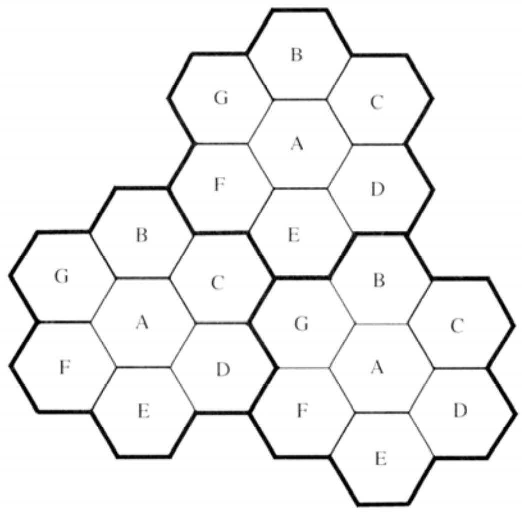

Figure 3. Simple Schematic of Cells and Frequency Reuse.

When a cellular phone is turned on, it scans for the base station with the strongest signal. The phone locks onto this channel and uses it until the signal strength drops below a pre-defined minimum. At that point, the phone begins scanning again to find another base station that has a stronger signal. A process called a handoff allows cellular phone users to travel between cells without losing a call. When signal strength falls below the pre-defined threshold, the base station "instructs" the phone to switch channels to another cell.[10] This can be done without disrupting a phone call and is a critical component of a cellular system. When a call is ended, the phone alerts the base station and the voice channel is opened up for other potential users. The phone continues to scan the control channels to identify the base station that has the strongest signal while the phone is not in use.[10] This allows the phone to receive and make calls efficiently since the phone always "knows" the closest base station. This feature also allows WLT to operate without a call actually being made as long as the phone is on.

Wireless Location Technology

As mentioned earlier, WLT involves determining the spatial location of a wireless device. Wireless location technology can be generally classified into two groups: handset-based systems and network-based systems.[1] Handset-based systems rely on GPS-enabled wireless phones. The GPS unit in the handset determines the location of a phone, and this information is relayed from the phone to a central processing system maintained by the wireless carrier. Network-based systems use signal information from wireless phones to determine locations. In this case, special equipment is installed at cellular towers throughout a metropolitan area to analyze signal characteristics of calls. Network-based systems often determine positions by analyzing signal power and angle of arrival as seen from multiple cellular towers. Although network-based systems do not rely on users having GPS-enabled phones, they are usually less accurate than handset-based systems.

Beginning in 1996, the Federal Communication Commission (FCC) issued a series of orders to improve the quality of 911 service provided by wireless carriers.[1] The purpose of these orders, commonly called E911, was to improve emergency vehicle response by providing specific location information about where a wireless telephone call originated. The rollout of this program is ongoing and is scheduled to conclude at the end of 2005. E911 has been the driving force in the deployment of WLT to date. The FCC allowed each wireless carrier to determine whether it would support a handset-based or a network-based WLT system but established minimum accuracy standards for each system type:

- *For handset-based systems:* Calls must be located to within 50 meters for 67 percent of calls and 150 meters for 95 percent of calls. Handset-based systems are preferred by most cellular providers.

- *For network-based systems:* Calls must be located to within 100 meters for 67 percent of calls and 300 meters for 95 percent of calls.

A secondary factor influencing the deployment of WLT is the development of location-based services (LBSs). The LBS concept has been embraced by wireless communications companies, who foresee that the public will pay for information tied to their current geographic location. An example of an LBS would be a driver paying a fee to find the nearest hotel or "fast food" restaurant to their current position. The LBS market is continuing to grow worldwide, with the strongest growth being seen in Asia.[11] Worldwide revenues for LBSs are estimated at $500 million today and are expected to exceed $3.6 billion by the end of the decade.[11] Since the success of the LBS market is dependent on knowing the location of the user of a device, this is also factoring into the development of WLT around the United States. Given the large potential market for these services, many companies continue to explore new applications and methods to use these data commercially.

Since both the government and the private sector have an interest in the development of WLT, it seems probable that these systems will be widely available in the near future. The private sector is still attempting to determine which commercial applications are viable, and the precise levels of error that WLT will provide are unclear. One of the many proposed possible applications of WLT is traffic condition monitoring.

Evaluations of WLT for Traffic Monitoring

Several evaluations of WLT-based traffic monitoring systems have been conducted around the world, but there has not been an entirely successful evaluation to date. Past evaluations have taken one of two forms: simulation studies or field operational tests. The simulation studies have generally dealt with simplified conditions. The field operational tests have been successful in generating reasonably accurate location information but have failed to produce the fidelity of traffic data that would be useful to DOTs or to the public.

Simulation Studies

Researchers have used simulation studies to determine the potential accuracy and effectiveness of WLT systems. Although these studies do not replicate the actual conditions precisely, they may provide some indication of the potential performance of a WLT system. The literature review identified three major simulation-based studies. They provide some useful information but generally do not conclusively show that WLT-based monitoring is feasible in a real-world environment. Further, the measures of effectiveness were sometimes inappropriate, and the system design and roadway characteristics were not examined to a great degree of detail.

French researchers conducted a project to evaluate the effectiveness of wireless positioning systems.[12] Their initial efforts focused on developing a discrete event simulation of traffic flow to determine the sample size requirements and accuracy of a WLT system. The simulation examined the impact of varying levels of probe vehicle penetration on the accuracy of travel time estimates. An error of 150 meters was assumed, and researchers examined several simple geometric conditions. The simulations results showed that freeway link travel times could be estimated to within 10 percent of their actual value if there is at least 5 percent penetration of wireless devices in the traffic stream. Although these results appear promising, they are based on simple geometric conditions on small networks. Further, sampling frequency was not examined. The researchers also do not explicitly describe how they matched vehicles to the network, although the report implies that it was an extremely simple form of geometric map matching where positions were simply matched to the nearest link.

Researchers from the Helsinki Institute for Information Technology developed a simulation of a network-based WLT.[13] They created a hypothetical road network and overlaid simulated wireless towers on the network. Their model used signal strength, combined with a wireless signal propagation model, to try to estimate vehicle locations. As a comparison, the researchers compared their model results to a very simple process whereby locations were assigned to the wireless tower that received the largest signal strength. The researchers found that their algorithms produced more accurate location estimates than simply assigning positions to the nearest cellular tower. Errors in positioning were still significant in the model, however. It should be noted that the researchers were concerned only with determining position estimates, not with developing speed or travel time estimates.

A recent evaluation by the Berkeley Institute for Transportation Studies also examined factors that could affect the use of WLT.[6] The researchers examined three variables in their simulation: location accuracy, frequency of locations of a single wireless device, and the total number of locations that could be determined per square mile per second. The variation in the number of roads that could be traversed by at least one vehicle within a 5-minute period was used as the measure of effectiveness to compare different alternatives. The researchers did not attempt to address whether the observed sample sizes were sufficient to produce accurate estimates of speeds or travel times for the entire traffic stream, and they did they attempt to produce any estimates of speeds. Further, they do not explicitly talk about how they matched vehicles to roads or describe their test network or simulation method to any degree of detail. The major findings of this research effort were:

- Assuming a network-based system with an accuracy of 100 meters, at least one measurement can be generated on 85 percent of the roads every 5 minutes. This assumes that positions are updated every 30 seconds and a maximum of 40 locations are determined every second per square mile.

- Assuming a handset-based system with an accuracy of 50 meters, a measurement can be generated on 90 percent of the roads every 5 minutes. This assumes that positions are updated every 30 seconds and a maximum of 40 locations are determined every second per square mile.

Again, these results only show whether a vehicle traveled on a particular road at least once during a 5-minute period. The researchers did not state whether this would be sufficient to determine actual speeds or travel times on a road.

Field Operational Tests

Several field tests of WLT have been performed in the United States. Network-based systems have been used in all cases since GPS-enabled phones composed a relatively small portion of available probes at the times of the studies. As noted earlier, the field tests to date have not reliably produced traffic condition information on specific roads. Further, the tests have been conducted on a small number of roadway networks, so the impacts of system design and roadway conditions cannot be separated in these studies. An additional test of WLT is currently underway in the Hampton Roads area of Virginia, but no data from that evaluation are currently available. This section discusses the results of these previous field tests.

The first major operational test of WLT was conducted over a 27-month period on I-66, I-495, and various state routes in Virginia.[3] This project was dubbed CAPITAL (Cellular APplied to ITS Tracking And Location) and was the result of a cooperative agreement between the Federal Highway Administration, VDOT, the Maryland State Highway Administration, and several private sector firms. This evaluation yielded the following major findings:

- By the end of testing, wireless telephones could be located within 100 meters of their actual position. The accuracy of the position estimates improved considerably as the number of towers providing directional information increased. The evaluators noted that accuracies on the order of 5 to 25 meters might be needed to perform accurate speed estimation for a network.

- Speeds could be determined for only 20 percent of all wireless phones that were located. To calculate speed, at least four position estimates had to be identified for each phone, and this occurred only 20 percent of the time.

- Link speed estimates could not be estimated for the network. This was attributable to the lack of well-developed algorithms that matched vehicles to links.

Although the CAPITAL test showed that wireless phones could provide reasonably accurate positional data, it was unsuccessful in producing traffic information that would be useful to DOTs or motorists. The position data were not used to generate link or system information that would be useful to users.

The U.S. Wireless Corporation tested their proprietary RadioCamera technology in the Washington, D.C., area in 2001.[14] The RadioCamera system examined radio frequency and multi-path characteristics of a signal and attempted to match the received signal to similar patterns stored in a central database. The average positional error of the RadioCamera system was approximately 30 meters. The system was scaled to track 160 phone calls every 2 seconds, generating 4,800 data points every minute. This system also failed to produce traffic condition data for the network. Although a reasonable amount of data was produced, speed estimation errors were often significant. This could be partially traced to small sample sizes on some roads, but errors varied by statistically significant margins, with errors of up to 25 mph.[4] It was unclear if these errors were attributable to network characteristics, system design issues, or a combination of the two.

Researchers at Berkeley obtained 44 hours of wireless location data from US Wireless Corporation for roads around Oakland, California.[15] The researchers found that the position estimates generally had a 60-meter accuracy, although 66 percent 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 researchers noted that the call lengths were generally very short, with a median call length of only 30 seconds. This made it very difficult to estimate speeds on links since position estimates were not available for long distances. The researchers were also not able to match 60 percent of vehicles to a roadway link.

Map-matching Methods

Operational tests of WLT systems have shown that it can be very difficult to match inherently inaccurate vehicle position estimates to the roadway network. Past evaluations either used simple methods to match vehicles to roads or did not explicitly describe how this matching was done, however. Well-developed map matching provides a possible way to take these inaccurate estimates and determine an estimated position for a vehicle that lies on the roadway network.

Although map matching can improve the accuracy of the position estimates generated by WLT, it is still subject to errors. Figure 4 shows an example of a simple map-matching process where the measured positions of a vehicle are simply projected onto the nearest road with no consideration of network topology or past vehicle travel history. Figure 4 shows that there are still errors in the "matched" positions, some of which are considerable (to the extent of being matched to the incorrect road).

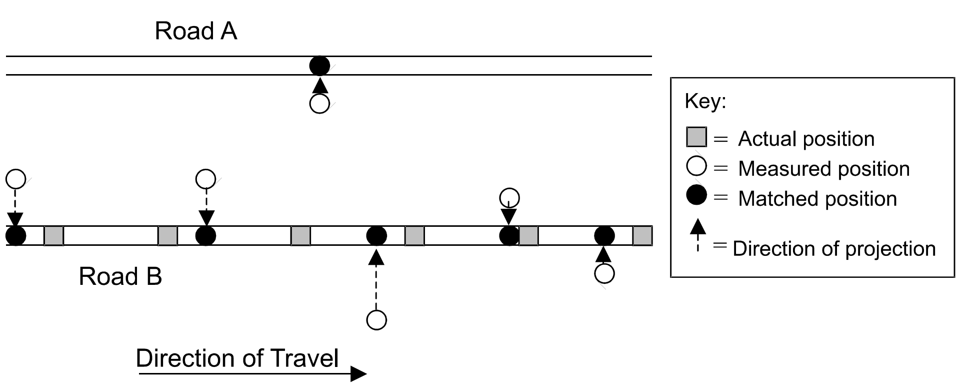

Figure 4. Simple Geometric Map Matching.

A great deal of research has been done on how to best match vehicle paths to networks, with most of it being focused on developing algorithms for in-vehicle navigation systems.[16,17,18] In-vehicle navigation systems are concerned with map matching for only a single vehicle, rather than for the potentially hundreds or thousands of vehicles to a network in real time as with a WLT-based monitoring system. Past map-matching research should provide insight into potential methods that could be extended to a WLT system, however.

Several map-matching techniques are documented in the literature. These techniques can be generally classified as follows:

- simple geometric methods
- geometric methods accounting for topology
- statistically based methods.

Although many researchers have developed map-matching techniques, numerical performance data are often absent from the reports documenting the work. In other words, researchers will present the theoretical framework for their technique but will not provide any quantitative data on how well it works. This makes direct comparisons between the various methods a problem.

Simple Geometric Map-Matching Techniques

Simple geometric map matching is the most basic type of map matching. It relies on purely geometric criteria to project a vehicle's estimated position to the closest node or link on the network. This form of map matching uses only the geometric shape of the links and does not consider issues such as link connectivity or link directionality. This type of map matching is the simplest both conceptually and computationally but could create large errors and matches that are not physically possible based on potential vehicle trajectories or network characteristics. Two forms of geometric map matching are point-to-point and point-to-curve matching.

The simplest form of geometric map matching is point-to-point matching. In this method, vehicle locations are snapped to the nearest network node or shape point. This is usually determined by calculating the Euclidean distance between the vehicle position and all network nodes or shape points and determining the minimum distance.[17,19] The advantages of this method are the relative ease of creating an algorithm to perform the matching and the computational simplicity of the procedure. This method can be quickly and easily implemented, and does not require significant computational time.

This method is likely to be the least accurate form of map matching and is prone to instability and errors. The ability of the method to provide good results is dependent in large part on the characteristics of the network being examined. Links with more shape points will, on average, have more vehicle positions matched to them than links with few nodes or shape points.[17] This can create statistically significant impacts on the accuracy of this method.

A second geometric technique is point-to-curve matching. This technique involves projecting the current vehicle position onto the nearest link, rather than onto the nearest node or shape point. This method should theoretically produce more accurate results than point-to-point matching since the number of nodes or shape points should not influence the matching process. Again, the Euclidean distance between a point and a link is used to determine the closest link to a vehicle's estimated position.[17,19] The perpendicular distance to a line is the minimum distance only if the projection of the vehicle position actually falls on the line. If this is not the case, the distance to the link endpoints may provide a shorter distance. Both the perpendicular distance and the distance to link endpoints must be calculated to determine the overall minimum distance.

Point-to-link matching has the potential to improve accuracy over point-to-point matching. Only one additional calculation must be performed for each link, so it should not increase computation time. A problem with point-to-curve matching is that it can be very unstable, particularly if there are parallel, closely spaced links. Since this method does not consider past travel history or network connectivity, it could potentially assign positions between two parallel links in an oscillating manner even when there is no physical connection between them. An example of this was shown in Figure 4. This represents a potentially serious shortcoming of this method.

Geometric Map Matching with Topology

Geometric methods can be improved by considering the topology of the network along with geometric information. Inclusion of topology allows past vehicle history, link directionality, and network connectivity to be incorporated into map-matching processes. Attempts to incorporate topological considerations have usually added topology to point-to-curve geometric matching methods.[17,19] Topology was often incorporated by adding a performance function to the geometric map-matching procedures along with heuristic procedures to determine whether a path was possible. The performance function usually includes:

- factors for the distance between the measured location of the vehicle and a network link
- a measure of the difference in orientation between the vehicle trajectory and the link direction of travel
- the number of links that had to be traversed between the previous position and currently matched position.

These performance functions are evaluated for all candidate links, and the one with the highest outcome is selected as being the correct match.

There are two sources of documented results for topological map matching. Greenfeld[19] tested his procedure using GPS observations on a route between New Jersey and Manhattan. Like many other researchers in the map-matching field, Greenfeld does not provide a detailed assessment of the performance of the algorithm beyond saying that "a correct match was made virtually along the entire length of the tested route. The very few isolated mismatches that were computed were immediately corrected by consequently computed matches." No quantitative data were presented on the efficacy of the method.

White, Bernstein, and Kornhauser provide a more detailed comparison of several topological methods to simple geometric methods.[16] They evaluated their algorithms by attempting to match field GPS data to digital maps for a county in New Jersey. Position estimates were collected every second, and data were collected on four distinct routes. The routes were in an urban area, and the speed limit was nearly always 25 mph. The average length of the links route ranged from 171 to 608 meters.

The algorithms did not consistently match a vehicle to the proper route, with the best algorithm for a particular route matching only between 66 and 86 percent of GPS readings to the proper link. No single algorithm emerged as the best for all networks. The only major finding from this evaluation was that the simple geometric point-to-link matching performed worse than all other algorithms. In general, it as found that all algorithms tended to perform better when links were long and vehicles traveled at high speeds.

Statistically Based Methods

A variety of more advanced statistical methods have been proposed to perform the map-matching task. They attempt to improve the quality of map matching by better accounting for some of the uncertainty in position estimates, but they also typically require more computational effort than the simpler geometric or topological methods. Two of the methods that have been evaluated for map matching are simple probabilistic methods and the multiple hypothesis technique (MHT).

With simple probabilistic methods, known location error characteristics are used to enhance the map-matching process. If the error characteristics of a location estimate are known, they can be used to restrict the candidate set of links for map matching.[20] Error characteristics can be used to determine an elliptical or rectangular area that defines the limits of the likely location of the vehicle. If only a single link is included in this area, the position of the vehicle can be easily matched to the network. If more than one link, or an intersection, is located on the road, then connectivity checks must be performed to determine likely links. Essentially, consideration of known error characteristics can be used to restrict the search area for a map-matching function and can provide an additional filter on top of geometric or topological methods. Although this method has been proposed in the literature, no quantitative data on its performance have been presented.

The MHT approach has been proposed as another way to perform map matching. This technique involves first identifying roads that lie within a confidence region based on the known location estimate error characteristics, similar to the basic probabilistic method. The probability of a vehicle being matched to a link is determined through a series of recursive equations that incorporates the network topology along with differences in position and heading. These equations are used to develop a series of hypothetical paths for a particular vehicle. As new position estimates are generated, the hypothetical paths are updated and a likelihood score for each path is developed. Hypotheses for paths that have likelihoods that fall below a particular threshold are eliminated to improve computational efficiency.

A study evaluated the effectiveness of MHT techniques in map matching.[21] The MHT algorithms were formulated as a single target tracking problem, and the assumed true position of vehicles was on the roadway network. The measured position and heading of a vehicle using GPS was defined as a function of the measured position and heading of the vehicle, the standard deviation of position and heading estimates, and the number of visible GPS satellites. The measured position was then matched to all feasible candidate links, and a likelihood that the match was correct was calculated for each link. The calculated likelihood that a vehicle is on a specific link is a function of the following factors:

- The probability of a certain type of road facility, such as an underpass, as a function of the number of visible GPS satellites.

- The probability of a link direction. This probability is assigned using the differences in headings between a link and a vehicle's track.

- The probability of a specific position and heading difference.

- The probability of connectivity. This probability was assigned based on the number of links bypassed between the previous measurement and the current candidate map matched position. If large numbers of links had to be bypassed between subsequent readings, this probability was assumed to be low.

The number of hypotheses grew exponentially as time increased, so it became necessary to eliminate hypotheses whose probability became very small to maintain a manageable set of hypotheses. Threshold values were set to eliminate hypotheses when a probability became very small either in absolute terms or relative to the hypothesis with the largest probability. Thresholds were also set for accepting a hypothesis as being correct and eliminating all other hypotheses.

Several field tests of these algorithms were performed to assess their abilities. The algorithm was evaluated in areas with low and high street densities and used a combination of GPS and dead reckoning to provide position information. The algorithm matched to the correct link between 83 and 96 percent of the time, but this algorithm still relied heavily on internal vehicle information (such as dead reckoning data) to produce position estimates.

Issues with Monitoring Based on Probe Vehicles

Traffic monitoring systems using probe vehicles are being used in a number of locations around the United States. These systems primarily rely on AVI-based systems to generate position data. WLT-based monitoring systems represent a new technology that could be applied to the concept of probe-based traffic monitoring, and some of the previous research into how vehicles should be sampled in probe-vehicle based systems may be able to be extended to WLT-based systems.

Associated Advantages and Disadvantages

Probe vehicles have been used in several jurisdictions to generate traffic condition information. To date, most probe-based systems have used AVI transponders that motorists use for electronic toll collection. AVI readers are installed at fixed locations off a toll facility to determine when a vehicle travels by these points. Speed and travel time information can be derived based on the time that the probe vehicles pass by a series of AVI readers. AVI-based probe vehicle systems have been deployed in Houston, New York, and the Puget Sound region.[22] These probe-based systems have generally been successful in generating traffic condition information.

Although AVI-based probe vehicle systems do have lower infrastructure costs than ILDs, there are still a number of drawbacks to their deployment, including:

- Infrastructure costs for AVI readers and communications lines are still significant. Capital costs for a single detector site on a six-lane highway range from $18,000 to $38,000, with annual operating costs of $4,000 to $6,000 per site.[23,24] If sites are to be spaced every 1 to 2 miles, this can be a significant cost.

- Probe vehicles must "opt-in" by either using an AVI tag or transmitting GPS positions back to the traffic management center. If there are no toll roads using AVI transponders in the area, it may be difficult to recruit probe vehicles.

- AVI-based probe systems lack the flexibility of WLT-based systems since permanent installations are required at fixed locations. Although the infrastructure requirements are lower than with ILDs, AVI probe-based systems do not have the ability to cover as wide an area as WLT-based systems potentially do.

Sampling Requirements

Two general methods have been used to generate minimum requirements for probe sample size: static methods and dynamic methods. Static methods rely on a single, fixed sampling requirement based on historic data, whereas dynamic methods vary sampling requirements based on existing conditions on a road.

Static sampling requirements for probe-based systems have been derived both theoretically and empirically. The most common theoretical approach involves creating sample size requirements based on the CLT. In this case, a desired accuracy level and a known or assumed standard deviation of speeds are used to determine the minimum number of probe vehicles required to generate an appropriate speed estimate on a link. Research in Houston used the following equation to estimate the number of probe vehicles required to estimate the speed along a roadway link:[25]

$$n = \left(\frac{t_{n-1} s}{E} \right)^2$$

where n = minimum number of probe vehicles required
t_{n-1} = value of t-distribution with n-1 degrees of freedom at desired confidence level
s = standard deviation of sample speeds (mph)
E = permitted absolute error (mph).

If the number of samples approaches or exceeds 30, the t-distribution can be replaced with a normal distribution. This equation shows that the number of probe vehicles required to estimate speeds on a link will be a function of the confidence level, standard deviation of speeds, and allowable error. In Houston, researchers found that this equation produced minimum sample sizes of 1 to 4 vehicles every 5 minutes at a 95 percent confidence level and 10 percent allowable error.[25] This equation has also been adapted by other researchers to incorporate other measures of reliability and to handle traffic in low volume corridors.[26, 27] Although these sample size requirements are classified as static, it is only because a static standard deviation of speeds is usually assumed. If standard deviations are dynamically changed based on traffic conditions, these are no longer static guidelines.

In other cases, researchers have attempted to define truly static probe penetration requirements based on empirically derived data from simulations. In these cases, the researchers typically sought to achieve "good coverage" of a system. The definition of *good coverage* often varies from study to study, but it almost always involves ensuring that some minimum number of probe vehicles traversed a roadway link in a given amount of time. Table 3 summarizes some of the guidelines proposed in previous studies.

Table 3. Previous Empirically Derived Static Guidelines

Researchers	Date	Guideline
Boyce, Hicks, Sen[28]	1991	400 probes to cover 200 square mile area
Sanwal, Walrand[29]	1995	4% probe penetration
Srinivasan, Jovanis[26]	1996	5% probe penetration
Ygnace, Drane, Yim, Lacvivier[30]	2000	5% penetration for monitoring based on cellular phones

There are several drawbacks to using these static guidelines. First, they do not account for differences in traffic conditions. If the standard deviation of speeds is low, use of static guidelines could result in significant over sampling; if the standard deviation is high, under sampling and errors in estimates of mean speed could occur. Using an overall probe penetration percentage also does not reflect differences in flow levels on different roads. If one road carries a high volume, there will be a large number of probe vehicles on the road, which should produce more accurate speed estimates; lower volume roads will have very few speed estimates.

A major limitation of static sampling guidelines is that certain time periods are likely to be over sampled and others under sampled depending on conditions at the time. Under sampling traffic will result in less accurate speed estimates, and over sampling will produce inefficiencies in terms of costs and system requirements. Recent work by Green et al. compared static sampling guidelines to dynamic CLT-based sampling requirements generated using actual variance information for 5-minute intervals.[31] That research found significant savings in sample sizes by using dynamic variance information, even when small numbers of samples were used with data that were not normally distributed. In a WLT-based system, where there may be a per-location charge, sampling efficiency is a very important issue.

Although the initial results of dynamic sampling requirements were promising, it is often difficult to generate real-time estimates of speed variance to implement the initial CLT-based algorithms. Further work by Green evaluated several dynamic sampling strategies based on the use of historic speed data on a facility.[32] In these methods, data were collected over a series of days on a facility to determine the typical distribution of speeds and speed variance on a road. The typical speed variance was then used to define the CLT-based sample size for a particular time interval. If the mean speeds determined from the probes were not within a pre-defined limit of the historic mean speed, then additional samples were taken. This method showed that sampling requirements could be changed dynamically without any reliance on real-time variance information and still produce accurate speed estimation results.

Development of Map-matching Algorithms

The literature review examined map-matching algorithms that could potentially be extended to WLT-based monitoring systems. Map-matching techniques found in the literature were all focused on the requirements of in-vehicle navigation systems. Although some basic map-matching principles could be applied to WLT-based monitoring, they have a number of limitations for this application, namely:

- In WLT systems, vehicle positions are observed from an external source, so no data generated internally by the vehicle could be used. This is in sharp contrast to in-vehicle navigation systems, which often use internally measured vehicle trajectory and speed as primary data sources.

- Computational power is often not an issue with in-vehicle systems since they track only one vehicle. If multiple vehicles are being tracked, computational power becomes more of a concern.

- Location estimates and updates of in-vehicle systems are likely to occur more frequently than in WLT systems. WLT-based systems are likely to have relatively infrequent position samples because of the load that repeated positions samples would place on the wireless network.

Three general approaches for map matching were explored in this project. They were adapted to account for these differences between in-vehicle navigation systems and the requirements of WLT-based monitoring. The techniques are briefly described here, with a more detailed explanation provided in Appendix A. More information, including flowcharts of program logic and algorithm pseudocode, can be found in the work by Fontaine.[33]

- *Simple geometric map matching.* A simple point-to-curve method was developed where positions estimates were simply projected onto the nearest link. This method served as a baseline against which other alternatives could be compared. Although the WLT deployments and simulation studies reviewed were often unclear, indications were that simple map matching was used in many of the evaluations. This method was also expected to require the least computational power.

- *Geometric map matching with topology.* This method included consideration of travel history and network topology in determining the current position of a vehicle. Inclusion of topology allows the algorithm to take link connectivity, network relationships, and vehicle history into account. Vehicle trajectories were compared to link orientations to determine likely paths, and only paths that were physically possible were candidates for matching. This method was expected to offer improved performance over simple geometric map matching.

- *Probabilistic technique based on the MHT.* The MHT method builds upon the topology method but considers of multiple path alternatives to improve its robustness. This method tracks multiple hypothetical paths until the estimated locations allow the algorithm to be relatively certain of a vehicle's true path. Potential paths for a vehicle were identified based on normalized scores created by the cumulative differences in (1) the distance between a position estimate and its matched location and (2) differences in orientation between a link's direction and a vehicle's estimated past trajectory. When a score reached a predefined threshold, the matched positions for that hypothetical path were retained and all other hypotheses were discarded. If no path reached convergence when a vehicle left the network or was lost by the system, the path with the highest likelihood score was selected.

Exploratory Testing Results

This section summarizes the results of the exploratory testing using simple networks. First, the main effects of each factor on the coverage and accuracy measures of effectiveness are presented to show the magnitudes of differences created by the levels of each factor. Second, significant main effects and interactions are identified, and factors that are meaningful within the context of the design and application of a WLT-based monitoring system are discussed for each form of map matching. Third, these results were used to define parameter sets that generate good accuracy and coverage for WLT-based monitoring systems.

Summary of Main Effects

The analysis of the simple network results focused on determining the impact of each factor on system coverage and system accuracy. The main effects of each factor on the following measures are discussed:

1. percentage of intervals with data
2. number of speed estimates per interval
3. accuracy of speed estimates.

The first two factors measure the availability of speed estimates and system coverage, and the third concerns system accuracy.

Percentage of Intervals with Speed Estimates

The first measure of system coverage examined was the proportion of time intervals where the system generated at least one valid speed estimate for a particular link. This was intended to measure the likelihood that no data would be generated for a link on the network during a given monitoring period. Table 4 summarizes the percentage of time intervals where at least one speed estimate was generated for a link. Table 4 also shows that a large total number of potential intervals were monitored in the simulations. Variations in the number of intervals monitored are attributable to test network characteristics. For example, since the parallel roads and signalized grid cases monitored two mainline roads, those situations had many more observations that the isolated freeway and signalized arterial cases that monitored only one road.

Table 4. Proportion of 15-Minute Intervals with at Least One Speed Estimate

Factor	Level	No. of Intervals Monitored	Proportion of Intervals with At Least 1 Speed Estimate		
			Simple	Topology	MHT
Network type	Isolated Freeway	1408	96.5	98.0	97.9
	Signalized Arterial	1408	93.1	96.9	93.8
	Parallel Freeways	2816	87.1	94.0	94.0
	Signalized Grid	2816	73.3	84.6	85.6
Volume (Q)	500 vphpl	4224	88.8	95.7	95.1
	1500 vphpl	4224	81.4	88.3	88.5
No. of vehicles tracked (N)	25	4224	75.0	85.6	85.5
	100	4224	95.1	98.4	98.1
Frequency of readings (F)	μ =15 sec	4224	93.2	97.8	98.4
	μ = 60 sec	4224	77.0	86.2	85.2
Probability of losing vehicle (P)	1%	4224	88.3	94.7	95.1
	10%	4224	81.8	89.3	88.5
Error (E)	σ = 10 m	4224	90.3	91.9	90.4
	σ =150 m	4224	79.9	92.1	93.2
Link length (L)	0.5 mi	768	83.7	91.2	91.1
	5.0 mi	7680	98.8	100	99.9

Several interesting trends are visible in Table 4. First, availability of data is often lowest with the simple map matching. The topology and MHT methods often have comparable results in terms of the percentage of intervals with data. With the exception of simple map matching, results comparable to or better than those documented in the University of California-Berkeley study were achieved for percentage of links traversed by at least one vehicle in an interval.[6] There are also trends in terms of the impact of factors on the availability of speed estimates:

- *Network type.* Generally speaking, the availability of data declined as network complexity increased. With the exception of simple map matching, all methods performed comparably for the non-signalized cases. The topology method performed the best for the signalized arterial, and the MHT method worked better for the grid network. This finding is related both to the relative sizes of these networks and the increasing complexity of map matching on roads with several parallel and intersecting facilities.

- *Traffic volume*. The lower volume cases had better network coverage than the higher volume cases. When volumes are high, monitored vehicles may spend an extended period of time on a relatively small number of links, especially in signalized systems. In low-volume situations, vehicles traverse the network under much higher speeds than the high volume cases. As a result, an individual vehicle will traverse a larger number of links in the same amount of time with the low-volume cases than the high-volume cases. This means that data are not as evenly distributed throughout the network for the high-volume cases, creating gaps in coverage.

- *Number of vehicles tracked.* Increasing the number of vehicles tracked improved the system coverage for all cases. This finding is intuitive since tracking more vehicles will increase probe penetration on the network, making it more likely that a vehicle will traverse a link in any given interval.

- *Frequency of readings.* In all cases, the shorter mean time between readings produced more intervals with at least one speed estimate.

- *Probability of losing a vehicle.* The 1 percent probability produced more intervals with speed data than did the 10 percent interval.

- *Error of position estimates.* Interestingly, the higher error of position estimates generated more intervals, with at least one data point for all cases except for simple map matching. The reason for this is unclear, but it may be related to the increased ambiguity of a vehicle's position when error estimates are high. This may result in more cases where there is an estimate on a link, even though the estimates are not accurately located.

- *Link length.* Longer links tend to have better coverage, which seems intuitive given that there is more opportunity for a speed estimate on a road.

Number of Speed Estimates

Although having at least one speed estimate available for a link is important, it does not ensure that sufficient data are available to generate an accurate speed estimate. The CLT-based sampling method discussed in the literature review showed that the potential accuracy of a probe-based system depends on the number of samples collected and the underlying variance of the speed distribution. As a result, increasing the number of speed estimates should decrease the errors in estimation of the mean speed. Table 5 shows the main effects of the number of speed estimates that are generated per link per 15 minutes.

Table 5. Mean Number of Speed Estimates per 15 Minutes per Link

Factor	Level	Mean No. of Speed Estimates per 15 Minutes per Link		
		Simple	Topology	MHT
Network type	Isolated Freeway	100.5	155.0	155.2
	Signalized Arterial	51.2	78.9	75.4
	Parallel Freeways	39.0	73.9	74.4
	Signalized Grid	21.1	34.3	37.1
Volume (Q)	500 vphpl	51.6	88.7	85.3
	1500 vphpl	54.3	82.3	85.8
No. of vehicles tracked (N)	25	21.4	34.1	35.3
	100	84.5	136.9	135.8
Frequency of readings (F)	μ = 15 sec	73.6	111.7	117.6
	μ = 60 sec	32.3	59.4	53.4
Probability of losing vehicle (P)	1%	62.9	98.5	100.9
	10%	43.0	72.5	70.2
Error (E)	σ = 10 m	72.2	100.2	95.3
	σ = 150 m	33.7	70.8	75.8
Link length (L)	0.5 mi	9.6	15.6	15.7
	5.0 mi	96.3	155.4	155.4

Mean Absolute Speed Estimation Error

Table 6 summarizes the main effects of each factor on the mean absolute speed estimation error. The signed mean was not used to describe central tendency because extreme values sometimes exerted undue influence. In general, the topology method and the MHT method performed the best in terms of speed accuracy. All methods usually performed reasonably well on the isolated freeway and parallel freeway cases. This was to be expected because there was little ambiguity as to which road a vehicle was on. All four methods had mean absolute errors within 4 mph of reality for these situations.

Table 6. Mean Absolute Speed Estimation Error (mph)

Factor	Level	Mean Absolute Speed Estimation Error per 15 Minutes per Link (mph)		
		Simple	Topology	MHT
Network type	Isolated Freeway	2.40	2.06	2.11
	Signalized Arterial	8.73	7.80	7.78
	Parallel Freeways	3.58	2.58	2.60
	Signalized Grid	11.52	11.30	10.54
Volume (Q)	500 vphpl	6.79	6.06	6.08
	1500 vphpl	6.34	6.09	5.63
No. of vehicles tracked (N)	25	6.95	6.42	6.39
	100	6.28	5.77	5.40
Frequency of readings (F)	μ = 15 sec	7.49	7.31	6.77
	μ = 60 sec	5.47	4.67	4.82
Probability of losing vehicle (P)	1%	6.79	5.68	5.53
	10%	6.35	6.49	6.22
Error (E)	σ = 10 m	5.14	4.74	4.70
	σ = 150 m	8.20	7.40	6.99
Link length (L)	0.5 mi	6.96	6.33	6.19
	5.0 mi	3.29	3.75	2.88

Across all factors, the topology method and MHT method were the most consistent at estimating the true speed on a road. Several underlying trends were visible across all forms of map matching, however:

- *Network type.* As network complexity, size, and speed variance increased, the speed estimation errors increased. Errors were low for simple networks with no cross streets and increased for large networks with interrupted flow facilities. These differences can probably be attributed to the lower numbers of speed observations available on the larger networks and the larger speed variances for the interrupted flow facilities.

- *Volume (Q).* Although differences in errors were generally low, errors were usually smaller when higher volumes of traffic were on the network. This reduction in error is likely attributable to lower speed variances and low overall magnitudes of speeds.

- *Number of vehicles tracked (N).* Speed estimation errors decreased as the number of vehicles tracked increased. This is likely attributable to the increased number of speed estimates with the higher value of N shown in Table 5.

- *Frequency of readings (F).* Speed estimation errors declined as the time between readings increased. By using a longer time between readings, the contribution of the location error to the total overall distance traveled was minimized. Shorter mean times between readings were also influenced by interactions with high standard deviations for position error.

- *Probability of losing a vehicle (P).* With the exception of simple geometric map matching, the accuracy of speed estimates improved when there was a low probability of losing a vehicle. This implies that it is better to track a single vehicle for a prolonged period of time than to track multiple vehicles for a short period of time.

- *Error of position estimates (E).* A lower standard deviation of error improved speed estimation for all four types of map matching. This is likely attributable to the interaction of sampling frequency and location error, discussed in the next section.

- *Link length (L).* For all four methods, longer links produced more accurate speed estimates. The improved accuracy is probably attributable to the larger number of speed estimates available on longer links.

Analysis of Main Effects and Interactions

Next, the main effects and interactions were analyzed separately for each form of map matching to determine what factors were significant in terms of speed estimation error. Once again, network type was confounded with three four-factor interactions. Although network type appears to play a role in accuracy, it could not be explicitly analyzed because of the experimental design. Given the wide variability in field geometric conditions, the extent to which conclusions about network type could be extended to "real-world" situations is unclear in any case.

The results of the experimental analyses are summarized here. In the interest of brevity, only the main effects and interactions that were significant to at least $\alpha = 0.05$ are discussed. All forms of map matching were successfully run in less than real time on a desktop PC.

Simple Geometric Map Matching

Table 7 shows the results of the analysis of the speed estimation error for simple map matching. Of the sources shown to be significant, the F, E, L, and F×E sources are the most interesting in terms of their system design implications:

- *Frequency of readings (F).* The results indicate that the errors are significantly different between the two sampling frequencies tested, with a longer sampling frequency producing significantly less speed estimation error. The longer time between readings allows for speeds to be measured over a larger proportion of the link, resulting in more representative speeds. It also reduces the contribution of position error on a percentage basis to the overall distance traveled.

- *Error of location estimates (E).* When the standard deviation of error was large, the speed estimation error increased. This increased the ambiguity of position locations and resulted in increased speed estimation errors.

- *Link length (L).* Longer links had lower errors, likely because more speed estimates were available.

- *F × E Interaction.* The results showed that errors were highest when short sampling frequencies were used in conjunction with large error standard deviations. In these cases, there were large errors in positions but two position readings were taken relatively close to each other in time. In these cases, speeds were often estimated to be much larger than reality or were matched to the wrong direction of travel. When the error standard deviation was large and the frequency between readings was also large, these errors were mitigated somewhat.

Table 7. Significant Main Effects and Interactions for Simple Geometric Map-matching

Source	F	Significance
F	40.6	0.000
E	9.2	0.002
L	4.8	0.029
F x E	15.6	0.000
P x L	21.6	0.000
F x P x E	3.9	0.047
Q x N x L	7.2	0.007
F x P x L	19.1	0.000
P x E x L	41.2	0.000
Q x N x F x P x E	7.5	0.006

Geometric Map Matching with Topology

Table 8 shows the results of the analysis of the speed estimation error for geometric map matching with topology. Of the sources shown to be significant, the F, P, E, L, F×E, Q×E, and P×L sources are the most interesting in terms of their system design and roadway network implications. The other interactions either had no obvious physical meaning or were extensions of the main effects. The practical meaning of these results is as follows:

- *Frequency of Readings (F).* A longer sampling frequency produced significantly less speed estimation error. Again, the longer sampling interval allows for a more representative link-wide speed to be estimated.

- *Probability of losing a vehicle (P).* Errors were minimized when the probability of losing a vehicle was low.

- *Error of location estimates (E).* When the standard deviation of error was large, the speed estimation error increased. This could also be partially attributed to the role of the F × E interaction.

- *Link Length (L).* Longer links minimized errors, likely attributable to the increased number of speed samples that was available on those links.

- *F × E Interaction.* Again, errors peaked when short sample frequencies were combined with large standard deviations of error.

Table 8. Significant Main Effects and Interactions for Geometric Map Matching with Topology

Source	F	Significance
Q	5.6	0.018
F	198.4	0.000
P	5.6	0.018
E	40.8	0.000
L	4.8	0.029
Q × E	6.2	0.013
F × E	40.0	0.000
N × L	7.0	0.008
P × L	20.1	0.000
Q × F × E	6.1	0.014
Q × N × L	7.9	0.005
F × P × L	26.5	0.000
F × E × L	10.4	0.001
P × E × L	131.4	0.000
N × F × P × E	9.5	0.002
N × P × E × L	4.0	0.046
Q × N × F × P × E	6.0	0.014
Q × F × P × E × L	4.3	0.038

- *Q × E Interaction.* When the standard deviation of error was high, speed estimation errors were lower at low volume than at high volumes. This was likely attributable to wrong-way matching during congested conditions. When the standard deviation of error was low, errors were lower at high traffic volumes.

- *P × L Interaction.* Overall speed estimation errors were minimized when P = 10 percent and long links were used. This seems to indicate that many vehicles should be sampled on longer links. When the link length was 0.5 mile, the use of P = 1 percent provided the best results. This may indicate that shorter links should retain vehicles they track whereas longer links should try to sample as many vehicles as possible.

MHT Map Matching

Table 9 shows the results of the analysis of the speed estimation error for MHT map matching. Of the sources shown to be significant, the F, E, F × E, and P × L sources are the most interesting in terms of their system design implications. The other interactions either had no obvious physical meaning or were extensions of the main effects.

Table 9. Significant Main Effects and Interactions for MHT Map-matching

Source	F	Significance
F	76.3	0.000
E	4.8	0.028
N × F	4.6	0.031
F × E	10.1	0.001
N × L	6.1	0.013
P × L	14.6	0.000
Q × F × E	5.9	0.015
Q × N × L	6.3	0.012
Q × P × L	4.9	0.027
F × P × L	11.7	0.001
P × E × L	76.3	0.000
N × F × P × E	6.5	0.011
N × P × E × L	4.2	0.040
Q × N × F × P × E	7.0	0.008

These results can be interpreted as follows:

- *Frequency of readings (F).* Much like the topology method, a longer sampling frequency produced a statistically significant reduction in speed estimation error.

- *Error of Location estimates (E).* When the standard deviation of error was large, the speed estimation error increased. Again, this was influenced by the F × E interaction.

- *F × E Interaction.* The same trends as with the topology method were seen with regard to the interaction of F and E. The interaction of short sampling frequencies and large standard deviations of error continued to be significant.

- *P × L Interaction.* The same trends as with the topology method were seen with MHT. Shorter links performed better with the use of P = 1 percent, and longer links performed better with the use of P = 10 percent.

Definition of Preferred Parameter Sets

Based on the exploratory testing, it was possible to identify preferred parameter sets that minimized the errors in speed estimation while providing adequate coverage. The topology and the MHT methods produced reasonably accurate speed estimates with no obvious problematic matching conditions. Based on the analysis of the main effects and interactions, it was possible to define a desired parameter set for these two forms of map matching. Using the values of the parameters tested, errors in speed estimation are minimized under the following conditions:

- The mean frequency between readings (F) is 60 seconds.
- The standard deviation of error of the location estimates (E) is 10 meters
- The probability (P) of losing a vehicle is 1 percent.

Although the link length was found to play an important role in reducing error, it is primarily a function of network geometry and is often beyond the scope of the system designer to affect. Likewise, the number of vehicles being tracked did not emerge as a significant main effect on accuracy, but when N = 100 in the system, coverage was usually improved for the MHT and topology methods. The N variable was also significant in several interactions, which indicated that the larger value of N produced better speed estimates. Thus, although the value of N was not found to be a significant determinant of the accuracy of the system, it does play an important role in determining the availability of data over the network being monitored. The effect of the level of N in conjunction with other factors (most notably the mean time between readings) will have to be considered because of its potential impact on the load on the wireless system, however.

The impact of N on system coverage is shown in Table 10 for each type of map matching and each combination of variables. In general, the MHT and topology methods produced similar results in terms of system coverage. When N = 100, between 95 and 100 percent of intervals had a speed observation for both methods.

Table 10. Percentage of Intervals with at Least One Data Point for Selected Parameter Set

Map Matching	No. of Vehicles Tracked	Network Type: Isolated Freeway	Isolated Arterial	Parallel Freeways	Signalized Grid
Topology	25	96.0	91.5	90.1	87.2
	100	99.5	99.5	98.0	95.2
MHT	25	96.0	90.3	90.0	85.5
	100	98.3	99.5	98.0	94.9

Table 10 seems to indicate that using an N = 25 does not produce coverage much worse than with an N = 100. These results should be viewed with caution, however, given the relative complexity and homogeneity of the networks evaluated. It is expected that N will play a more significant role as networks increase in size and complexity. As a result, the N variable was explored further in the case studies.

Figures 5 and 6 show the cumulative distribution of errors for uninterrupted and interrupted flow facilities, respectively. Figure 5 includes the isolated freeway and the parallel freeways. Figure 6 includes the signalized arterial and the grid network. These distributions were created by aggregating the accuracy results for all traffic flow levels and link lengths for applicable network types. Since N = 100 was shown to provide good system coverage, both figures were created using N = 100.

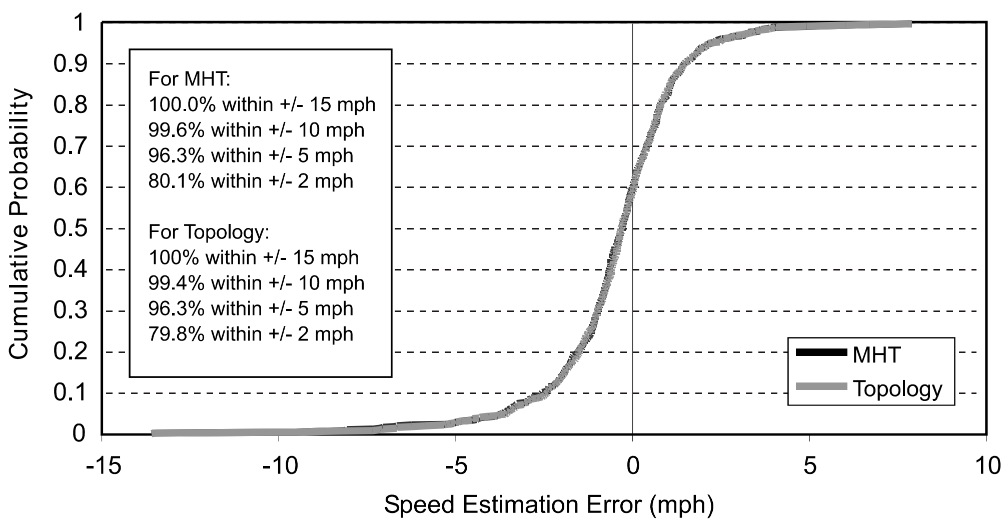

Figure 5. Distribution of Speed Estimation Error for Uninterrupted Flow Facilities, Topology Method and MHT Without Volume Factor.

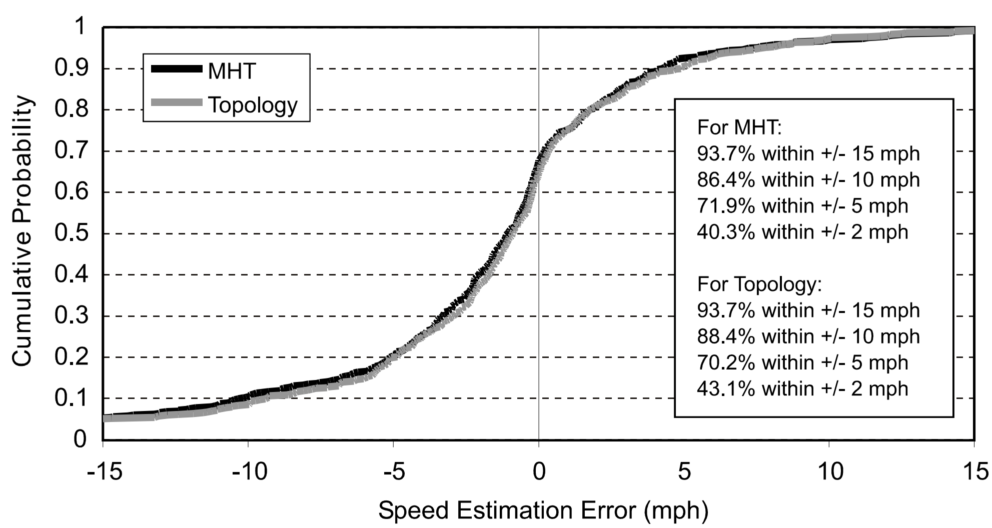

Figure 6. Distribution of Speed Estimation Error for Interrupted Flow Facilities, Topology Method and MHT Without Volume Factor.

The topology and MHT methods produced similar error distributions. In both cases, speed estimates were more accurate for uninterrupted flow facilities. There is also a slight bias in both methods toward underreporting the true speed on a facility, probably attributable to over sampling of lower speed vehicles. Figure 5 shows that both map-matching methods produce speed estimates that are within 15 mph of the true speed for more than 93 percent of time intervals. Likewise, about 70 percent of intervals are within 5 mph of the true link travel speed. The uninterrupted flow error distribution shown in Figure 6 is substantially more accurate, with more than 95 percent of intervals within 5 mph of the true speed and around 80 percent within 2 mph of the true link mean speed.

The differences in errors between the uninterrupted and interrupted flow facilities are probably not surprising. The uninterrupted flow facilities are relatively simple geometrically and have a relatively low speed variance. As a result, any individual sample on a road is likely to be fairly representative of the true mean speed on the road. For interrupted flow facilities, speed variance is larger because of the influence of traffic signals and the more complex network geometry, making matching more difficult. As a result, a single speed estimate may be mismatched to intersecting roads or a vehicle may be sampled at a speed that is not representative of the speeds on a link. These results may also be influenced by the relatively poor performance on the grid network. The performance could probably be improved by increasing the number of vehicles tracked so that the density of speed samples is greater on the networks.

Lessons Learned from Simple Networks

Several general conclusions could be drawn from this analysis that could be useful to DOTs considering using this technology:

- Any WLT-based monitoring system should use a proven form of map matching that includes consideration of network geometry, such as the topology or MHT method, to identify vehicle positions onto the roadway network.

- Using a relatively infrequent mean time between samples generally improves speed estimation over frequent sampling intervals. Using longer sampling intervals allows the system to gather information over longer distances, reducing the overall contribution of location error to the estimated distance being traveled.

- Large errors in vehicle positions usually translate into larger errors in speed estimation. Location errors should be minimized to improve system accuracy.

- Speed estimation errors are largest when the mean time between location estimates is short and position errors are large. If the system has large errors in positioning that cannot be reduced, it should be designed to use long intervals between location estimates.

- WLT-based systems will need to have the ability to change system parameters, especially the number of vehicles tracked, based on the characteristics of different parts of the network. More vehicles will be required for complex networks than for simple networks to ensure that adequate probe vehicle penetration is achieved. This indicates that the design of a WLT-based monitoring system is not a "one size fits all" and that systems will have to be scalable to accommodate location-specific characteristics.

Although the testing on simple networks provided some insight into how system design and roadway factors impact the coverage and accuracy of a WLT-based monitoring system, these evaluations were conducted for a relatively constrained set of conditions. Further testing and evaluation are necessary to consider more fully the range of situations that might be encountered in the real world.

Case Study Results

The case studies were used to explore and assess the application of the selected parameter sets in real-world situations. A variety of geometric and traffic conditions that were not explicitly tested on the simple networks were evaluated and potential areas of application and limitations of WLT-based monitoring were identified.

Network Preparation

This section describes the preparation of the case study networks.

Parameters to Be Tested

Before the data could be analyzed, the exact parameter sets to be tested for each network had to be determined based on the results of the exploratory testing. The exploratory testing showed that the following levels produced good speed estimation and system coverage results:

1. either topology or MHT map matching
2. error standard deviation of 10 meters
3. 1 percent chance of losing a vehicle.

The data from the simple networks indicated that although the 60-second mean sampling frequency generally produced the best speed estimation results, the 15-second mean frequency improved system coverage. Given that trade-offs between coverage and accuracy are likely, both options were examined in the case studies to gain further insight.

The results also showed that probe penetration was important, especially as network size increased. Since the sizes of the networks used in the case studies were different, the number of vehicles tracked was also examined. The CLT-based approach was used to determine the number of vehicles tracked for each network. The number was then tested at 100, 125, 150, and 200 percent of this requirement. Table 11 shows the number tested for each network for each combination of variables. The method used to inflate the number of vehicles to be monitored also did not explicitly look at system penetration but rather examined the issue from the perspective of the overall number of speed estimates.

Table 11. Number of Vehicles Tracked by Network and System Design

Network	Mean No. of Vehicles in Network	Monitored Time Interval (min)	Maximum CLT-based Sample	Mean Sampling Frequency (sec)	Inflated No. of Vehicles to be Monitored			
					100%	125%	150%	200%
Charlottesville	759	5	8.78	15	25	32	38	50
				60	95	119	143	190
Springfield	7,507	5	11.74	15	125	157	188	250
				60	490	613	735	980
Tyson's Corner	18,537	5	5.14	15	60	75	90	120
				60	235	294	353	470

Table 11 also shows the average number of vehicles in the network at any given time. This information gives an indication of the percentage of traffic being sampled. As can be seen, that there is a relatively large variation in the percentage in each situation. For the Charlottesville case, approximately 25 percent of traffic on the network is being sampled under the 200 percent, F = 60 case for a 5-minute interval. This is the maximum percentage of traffic sampled for all case studies. This is in contrast to the Tyson's Corner case, where a maximum of only 2.5 percent of all traffic is sampled. The percentage of traffic is always far below the current more than 70 percent penetration of wireless phones in the United States. Table 11 shows that although all cases sample only a portion of the traffic on the network, differences in speed variance, sampling frequency, and network size can create large differences in the proportion of vehicles being sampled. As WLT-based systems are implemented, the corresponding load the monitoring system will place on the wireless communications network must also be considered.

Comparison of Case Studies to Typical Cell Sizes

Proprietary cellular coverage information for typical urban, suburban, and rural areas was obtained from a cellular provider, and these data were examined to determine how the case studies compared to typical individual cell sizes. The test bed system applies the same WLT system design characteristics across the entire simulated roadway network. It appears that past systems deployed also used this approach, so it was continued for this evaluation. Monitoring large areas introduces concerns about how a system can be designed to handle the wide variety of traffic conditions in the network, however.

If a large area was being monitored, there might be the potential to subdivide it based on the size of individual cells. This means that a large area could potentially be divided up in a manner similar to that shown in Figure 3, with WLT system parameters set at the individual cell level. The coverage data are interesting in that they provide a comparison between the area being monitored by the simulated WLT-system and the potential to develop a finer-grained system design. Table 12 summarizes the data. Given that the cell size data are proprietary, it is presented in only broad categories.

Table 12 shows that the Charlottesville case study was the only case where the roadway network size was comparable to that of a typical urban or suburban cell. Depending on whether Springfield is categorized as urban or suburban, the data suggest it would be composed of 6 to 18 individual cells, on average.

Table 12. Comparison of Networks to Typical Cell Sizes

Area	Mean Cell Size (mi^2)	Minimum Cell Size (mi^2)	Maximum Cell Size (mi^2)
Typical urban cell	0.9	0.02	3.52
Charlottesville case study	2.1	N/A	N/A
Typical suburban cell	2.8	0.07	35.7
Tyson's Corner case study	9.5	N/A	N/A
Springfield case study	16.0	N/A	N/A
Typical rural cell	28.1	3.93	120.4

As a result, it can be concluded that the two Northern Virginia cases studies represent monitored areas that are much larger than a single typical cell size. Although the simulated system uses area-wide WLT design parameters, there may be an opportunity to subdivide the monitored areas based on individual cells if area-wide performance is lacking. This would provide an opportunity to set system design parameters by cell based on the conditions at more localized locations.

Charlottesville Case Study

The Charlottesville case study represented the least complex network and was the only one whose area was comparable to the size of a single cell. In many ways, this case study represented an ideal application of a WLT-based monitoring system because of its relatively small size and simple network. Data tables relevant to this case study are given in Appendix B.

Network Description

The Charlottesville network tested was produced from a calibrated model of afternoon peak hour conditions developed for the Thomas Jefferson Planning District Commission. Major facilities simulated included U.S. 29, U.S. 250, and Hydraulic Road. There are 10 traffic signals on this network, and two interchanges with freeways. The network is shown in Figure 7. This case study was based on conditions in fall 2003, so it does not take into account several recent changes to the actual network. Ten roadway links were monitored on the network, and information on these links is shown in Table 13. The maximum required number of samples was created from the CLT-based estimate of the most samples required to determine mean speed in a given 5-minute interval during the simulation.

Number of Samples

Tables B1 and B2 in Appendix B summarize the number of samples obtained on the Charlottesville network. Sample sizes varied dramatically. As noted earlier, the probability of acquiring a vehicle in the system is a function of the vehicle-hours of travel on a monitored link. Thus, links with high traffic volumes, long lengths, and/or low speeds will have more samples than links with low volumes, short lengths, and/or high speeds.

Table B1 shows the median number of samples per 5-minute interval for the network. Table B2 shows whether each link achieved a sufficient number of speed observations to meet

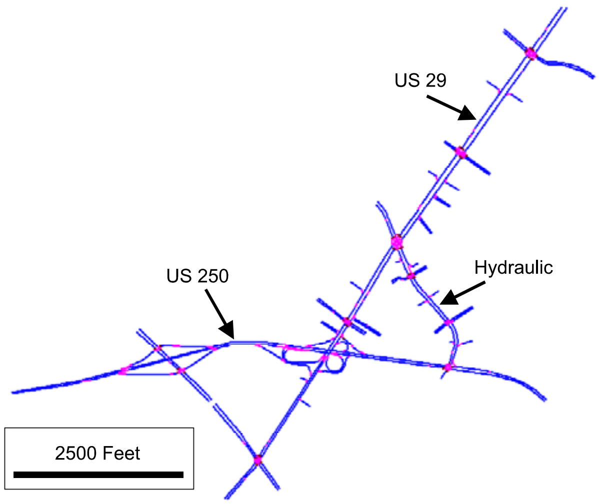

Figure 7. Charlottesville Case Study Network

Table 13. Charlottesville Case Study: Characteristics of the Monitored Links

Link	Length (ft)	Maximum Required No. of Samples	Mean Speed (mph)
US 29 NB–Barracks to US 250	1456.1	0.41	35.7
US 29 NB–US 250 to Hydraulic	1427.8	1.12	14.1
US 29 NB–Hydraulic to Greenbrier	3222.4	2.49	25.5
US 29 SB–Greenbrier to Hydraulic	3225.4	1.17	22.3
US 29 SB–Hydraulic to US 250	1529.6	1.57	11.2
US 29 SB–US 250 to Barracks	1370.9	3.76	23.7
US 250 EB–US 29 to Hydraulic	1127.5	8.00	29.2
US 250 WB–Hydraulic to US 29	1126.2	1.21	42.1
Hydraulic E–US 29 to US 250	1920.9	2.03	17.1
Hydraulic WB–US 250 to US 29	1915.7	1.84	17.0

the CLT-defined requirements for a 10 mph accuracy. The "required number of samples" column is based on the maximum number of samples required according to the CLT for a specific 5-minute interval. The four sub-columns represent a number of vehicles equal to 100, 125, 150, and 200 percent of the inflated number required theoretically to satisfy the CLT. Cases where at least 90 percent of intervals comply with the CLT requirements are shown in bold in Table B2.

As shown in Table B1, there was considerable variability in the median sample size on the different links. The U.S. 29 links have a significantly higher sample size than did the U.S. 250 or Hydraulic Road links. This is caused by the higher density of traffic on U.S. 29. It is also influenced by the fact that vehicles on Hydraulic and on U.S. 250 do not stay on the same road for as long a period as do vehicles on U.S. 29. In other words, vehicles on these streets often turn onto other roads, which interrupts the ability to trace vehicle speeds on shorter roads. Not surprisingly, however, increasing the number of vehicles tracked does increase the number of speed estimates that are usually available for a link.

The MHT method almost uniformly increased the number of samples per time interval over the topology method. The median number of speed estimates per time interval and the proportion of intervals that met the CLT-defined minimum sampling requirements increased when the MHT method was used.

The number of samples obtained was examined using the LSD test. The purpose of this analysis was to identify underlying trends in how system and link factors interacted to determine the number of samples on a link. These data provide insight into the relative importance of these factors in determining the number of valid data points. The LSD test was used to evaluate differences between combinations of map matching and mean sampling frequency for all network types. Table 14 summarizes the results of the LSD test. The results show that the combination of MHT and a 60-second sampling interval produced a statistically significant increase in the number of data points over all other alternatives.

Table 14. Charlottesville Case Study: LSD Results by Map-matching Type and Mean Sampling Frequency

Map-matching Type	Frequency	Mean	Standard Error	Significantly Different From
Topology	15	12.22	0.52	MHT, 15 MHT, 60
	60	12.22	0.62	MHT, 15 MHT, 60
MHT	15	22.81	0.89	All others
	60	32.11	1.36	All others

Accuracy of Speed Estimates

In general, the system was able to produce good speed estimates of traffic along U.S. 29, but speed estimates on side streets were less accurate. The percentage of speed estimates within 10 mph for a particular 5-minute interval is shown in Table B3. Cases where at least 90 percent of intervals were within 10 mph are shown in bold. These data are summarized to give an indication of the distribution of error and to show if the accuracy requirement used in creating the number of vehicles to be tracked was met.

The table shows that the MHT and topology methods were generally successful in producing speed estimates for the U.S. 29 portion of the corridor. The lack of accuracy on other segments was primarily attributable to a lack of speed samples because of higher speeds, short lengths, or low volumes.

The accuracy of the speed estimates was examined using the LSD test. The absolute value of error was examined to help eliminate the impact of extreme values on the mean errors being tested. The test was used to evaluate differences in the accuracy of speed estimates between combinations of map matching and mean sampling frequency for all network types. Table 15 summarizes the results.

Table 15. Charlottesville Case Study: LSD Results, Speed Estimation Accuracy, MHT vs. Topology

Map-matching Type	Frequency	Mean Accuracy	Standard Error	Significantly Different From
Topology	15	6.14	0.26	All others
	60	6.91	0.19	All others
MHT	15	5.16	0.15	All others
	60	8.01	0.17	All others

The results indicate that the MHT method and a 15-second mean sampling frequency produced the smallest mean absolute error for the Charlottesville case, and the MHT and 60-second mean sampling frequency produced the largest. These results show that short sampling frequencies work better than longer sampling intervals for a network of this size. Thee results are interesting given that the MHT, F = 60 case actually generated the largest number of speed estimates. It appears that this case, despite generating the most speed estimates, cannot adequately monitor the network because of its relatively small size and the relatively short time vehicles spend on the roads other than U.S. 29.

Processing Time

Map-matching time was negligible for the Charlottesville network, with an entire hour of position data matched in less than 7 seconds. The MHT method required slightly more time than the topology method, but times were within 2 seconds in all cases. Given that the matching time was so small, the map matching should not serve as a significant impediment to implementing a WLT-based monitoring system in a network like Charlottesville.

Springfield Case Study

The Springfield case study represents a much larger, more complex network than does the Charlottesville case study. The network consists of several major freeways and arterial roads, as well as some local roadways. Features that could create ambiguity within the map-matching process are also present. In many ways, the Springfield case study represents a worst-case situation for a WLT-based monitoring system. Data tables relevant to this case study are provided in Appendix B.

Network Description

The Springfield case study uses a simulated network consisting of 80.76 centerline miles of roadway. Some of the features include:

- 41 traffic signals
- five freeway interchanges
- a heavy rail line
- an HOV facility
- four locations that simulate the potential impact of pedestrians with handheld phones
- numerous parallel roads where speeds between facilities will be very different, such as residential roads next to interstate highways.

The netwc 'as simulated for afternoon peak hc ent in 1999, prior to the current ongoir on activities are not included in the m(

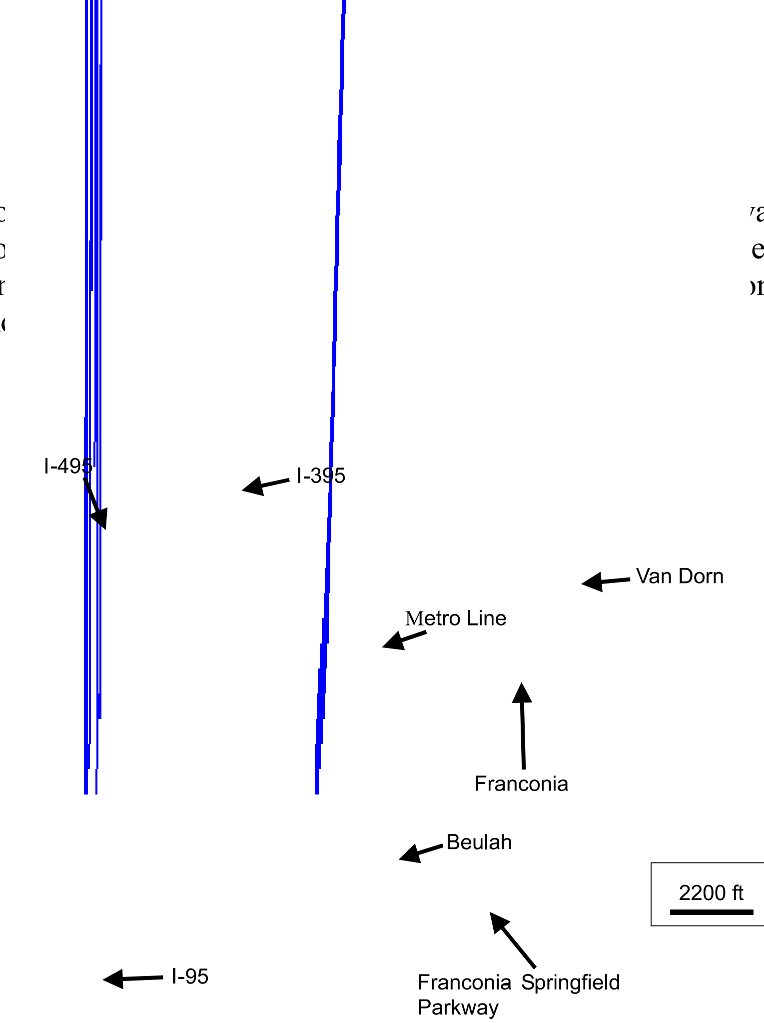

Figure 8. Springfield Case Study Network

Several potentially problematic features were included in this model. Four locations where there are shopping malls or schools were simulated. In each of these locations, several hundred pedestrians were put into the network going an average of 2 mph to evaluate whether non-roadway sources would impact speed estimates on nearby roads. Likewise, a Metro line was simulated, with trains traveling at 60-second headways. An HOV lane was also present on I-95 within the confines of this model.

Eighteen roadway links were monitored on the network. Summary information on the links is shown in Table 16. The maximum required number of samples were based on the CLT-based estimate of the most samples required to determine mean speed in a given 5-minute interval during the simulation. In this case study, links were longer than in the Charlottesville case study and there was greater variability in the speeds across the network.

Severe congestion was noted on I-95 NB between the Springfield Interchange and Van Dorn Street, on Franconia Road WB between Beulah and Commerce, and on Van Dorn Street SB between I-95 and Franconia Road. Intermittent congestion was observed on I-95 SB between the Springfield Interchange and Franconia Road.

Table 16. Springfield Case Study: Characteristics of Monitored Links

Link	Length (ft)	Maximum Required No. of Samples	Mean Speed (mph)
Franconia EB–Commerce to Beulah	8062.5	3.91	28.2
Franconia EB–Beulah to Van Dorn	2437.8	6.35	28.0
Franconia WB–Van Dorn to Beulah	2466.7	3.15	40.4
Franconia W–Beulah to Commerce	8056.1	0.34	5.2
Van Dorn NB–Franconia-Springfield Parkway to Franconia	3588.1	4.03	33.5
Van Dorn NB–Franconia to I-95	3058.2	5.03	13.9
Van Dorn SB–I-95 to Franconia	3062.1	0.04	2.1
Van Dorn SB–Franconia to Franconia-Springfield Parkway	3660.4	5.23	30.8
Franconia-Springfield Parkway EB–Henry to Frontier	2380.3	0.88	62.0
Franconia-Springfield Parkway EB–Frontier to Beulah	3049.1	1.80	50.7
Franconia-Springfield Parkway EB–Beulah to Van Dorn	7101.2	3.93	27.6
Franconia–Springfield Parkway WB–Van Dorn to Beulah	7035.9	4.50	27.1
Franconia-Springfield Parkway WB–Beulah to Frontier	3034.0	0.22	48.6
Franconia-Springfield Parkway WB–Frontier to Henry	2374.5	0.55	64.5
I-95 NB–Franconia to Springfield Interchange	2056.5	0.74	63.2
I-95 NB–Springfield Interchange to Van Dorn	6572.9	0.06	3.1
I-95 SB–Van Dorn to Springfield Interchange	5916.9	0.60	64.4
I-95 SB–Springfield Interchange to Franconia	2088.8	11.74	27.4

Number of Samples

Tables B4 through B7 show the number of speed samples obtained on the Springfield network. The MHT method almost uniformly increased the number of samples per time interval over the topology method. The median number of samples per time interval and the proportion of intervals that met the CLT-defined minimum sampling requirements increased when the MHT method was used. Tables B4 and B5 show the median sample size per 5-minute interval. Tables B6 and B7 show the percentage of intervals that met the CLT-based sampling requirements based on the same conditions. Cells shown in bold in Tables B6 and B7 met the CLT-based sample requirements at least 90 percent of the time.

Much like the Charlottesville case study, sample sizes varied dramatically, although the variation was more extreme for Springfield. Again, increasing the number of vehicles tracked usually results in a larger number of speed estimates on a link. The congested roads acted like "sinks" for the samples, however, which caused higher speed and/or lower volume roads to have fewer samples. As a result, although the congested high-volume roads were successfully monitored, lower volume roads had few samples. As shown in Tables B4 and B5, there was considerable variability in the median sample size on the different links. The congested links, such as I-95 NB between the Springfield Interchange and Van Dorn Street, had a significantly higher sample size than did the uncongested links, such as the Franconia-Springfield Parkway.

The number of samples obtained was examined using the LSD test. The purpose of this analysis was to identify underlying trends in how system and link factors interacted to determine the number of samples on a link. The test was used to evaluate differences between combinations of map matching and mean sampling frequency. Given the different characteristics of the freeways and arterial segments, the analysis was further segregated by roadway type. Table 17 summarizes the results. The results show that the combination of MHT and a 60-second sampling interval produced significantly more data points than did all other alternatives for each type of facility.

Table 17. Springfield Case Study: LSD Results by Map-matching Type, Mean Sampling Frequency, and Link Type

Road	Map-matching Type	Frequency	Mean	Standard Error	Significantly Different From
Arterial	Topology	15	3.92	0.19	All Others
		60	7.51	0.60	All Others
	MHT	15	9.74	0.46	All Others
		60	17.76	1.15	All Others
Freeway	Topology	15	13.52	1.33	All Others
		60	48.60	5.96	All Others
	MHT	15	30.97	2.91	All Others
		60	79.91	8.90	All Others

Accuracy of Speed Estimates

The percentage of speed estimates within 10 mph for a particular 5-minute interval are shown in Tables B8 and B9. Cases where at least 90 percent of intervals were within 10 mph are shown in bold. These results indicate that the accuracy of the speed estimates generated by the simulated WLT-based monitoring system for the Springfield network was much worse than for the Charlottesville case study.

The tables show that the system produced accurate estimates on the congested facilities, but other facilities were not monitored as precisely. To a large degree, the monitoring problems were attributable to a lack of speed samples at those site. Although this could be corrected by increasing the number of vehicles to be monitored, this would be an extremely inefficient way of approaching the problem. The 200 percent case is monitoring 980 vehicles simultaneously, and increasing that value on a network-wide basis is probably not an ideal solution to the problem. The WLT-based monitoring system seemed to produce more accurate speed estimates on higher speed facilities using a shorter sampling frequency, although the congested facilities still acted as a sink for the samples acquired.

Although a lack of data points was certainly part of the problem, another source of error in the speed estimates was errors in map matching. The most pronounced problems occurred when one direction on a road was congested while the other was moving at high speeds, such as the stretch of I-95 between the Springfield Interchange and Van Dorn Street. In these cases, the slow-moving traffic moves a relatively small distance, Δs, in a relatively long time period, Δt. Errors in position estimation can actually cause the apparent orientation of a vehicle to be the reverse of its actual direction of travel, resulting in a match to the wrong direction. When there is a large disparity in the speeds by direction, this can create significant problems.

To illustrate this problem, consider the I-95 case between Springfield and Van Dorn Street. I-95 NB has a median number of samples per 5 minutes of 300.0 for the MHT, 100 percent case. For the same conditions, I-95 SB has a median number of samples per 5 minutes of 6.0. If five of six individual vehicle speed samples matched to I-95 SB were 65 mph and a single speed of 3 mph was incorrectly matched to I-95 SB, then the overall average would be reduced to 54.7 mph, more than 10 mph off the true average speed. Although improved data screening would help alleviate this problem, it would not address the issues of the small numbers of samples on the other facilities.

There were no obvious problems with speed estimation in the areas around the simulated non-vehicular wireless sources. In these cases, the simulated pedestrians were often traveling in a slightly different direction from the surrounding roadways, so the algorithms would not match them to the roadway network. Problems occurred, however, with the HOV lanes. Since the HOV lanes had basically the same orientation as the southbound lanes of I-95, vehicles were often matched incorrectly from the main lanes to the HOV lanes and vice versa. Even with the relatively accurate standard deviation of location error of 10 m, the system could not differentiate between two facilities that were basically concurrent, except in the areas where the orientations changed significantly from one another. As a result, speed estimation errors for I-95 between the Springfield Interchange and Franconia are relatively high.

Table 18 summarizes the results of the LSD test. The results show no significant difference in the speed estimation errors for any of the arterial cases. This is likely a result of the limited number of data points available for many arterial roads. For freeways, the topology method performed significantly better than the MHT method, with the 15-second sampling frequency performing the best. This is again attributable to the influence of errors in map-matching in the MHT cases and the limited sample size on uncongested high-speed freeway links.

Table 18. Springfield Case Study: LSD Results, Speed Estimation Error by Map Matching Type, and Mean Sampling Frequency

Road	Map-matching Type	Frequency	Mean	Standard Error	Significantly Different From
Arterial	Topology	15	9.50	0.32	None
		60	9.77	0.31	None
	MHT	15	9.65	0.27	None
		60	9.36	0.25	None
Freeway	Topology	15	8.31	0.58	MHT, 15 MHT, 60
		60	9.22	0.50	MHT, 15 MHT, 60
	MHT	15	14.83	0.83	Topo, 15 Topo, 60
		60	13.31	0.58	Topo, 15 Topo, 60

Processing Time

The processing time for the Springfield case is shown in Figure 9. Again, the map-matching times for the topology and the MHT methods are very close, so there is not a significant computational savings when using the topology method for this network. Once again, the overall time needed to match the positions to the network is not a significant impediment to implementation of the WLT-based monitoring system. Even when a relatively large number of vehicles is being monitored, the processing time is under 5 minutes to process 1 hour of data.

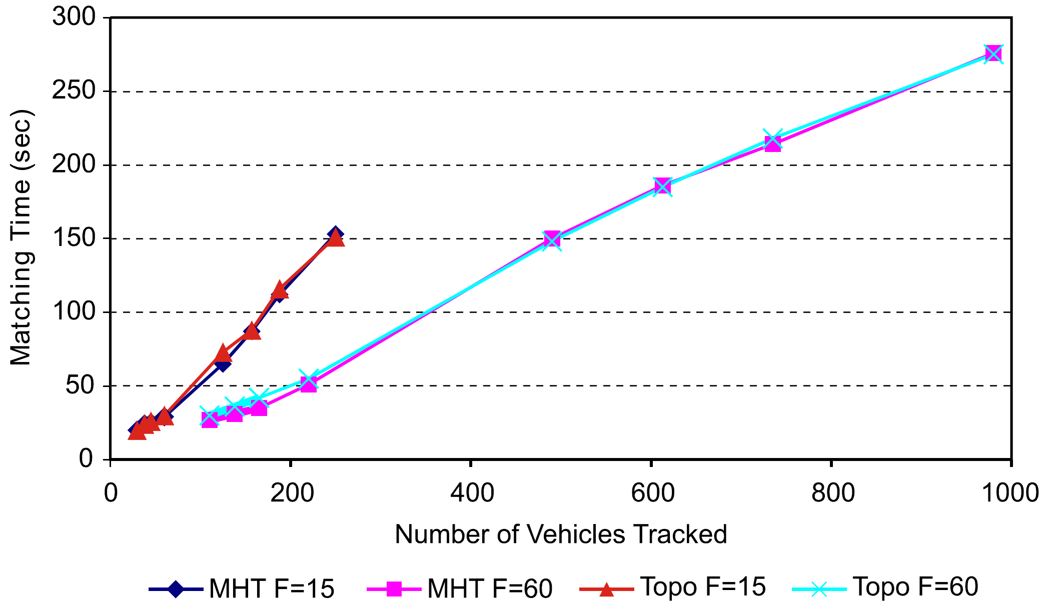

Figure 9. Springfield Case Study: Matching Time vs. Number of Vehicles Tracked.

Tyson's Corner Case Study

The Tyson's Corner case study represents an intermediate case between the Charlottesville and Springfield scenarios. The area of this case study was approximately halfway between the two, and traffic was distributed relatively evenly on the major roads in the network. Data tables relevant to this case study are provided in Appendix B.

Network Description

The Tyson's Corner case study used a simulated network consisting of 85.45 centerline miles of roadway during the afternoon peak hour. This simulation was intended to provide an intermediate test of the performance of a WLT-based monitoring system with some features that could create problems in the matching process. Some of the features included:

- 32 traffic signals
- seven freeway interchanges
- two locations that simulate the potential impact of pedestrians with handheld phones on system estimates
- slow-speed toll facilities paralleling free-flow freeway roads.

The network is shown in Figure 10. Fourteen links were monitored. Summary information on the monitored links is provided in Table 19. The maximum required number of samples was based on a CLT-based estimate of the most samples required to determine mean speed in a given 5-minute interval during the simulation. Although International Drive had also been selected for monitoring, no vehicles speeds were generated in 1 hour of simulation. The other facilities being monitored had a much higher number vehicle-hours of travel, so lower volume roads such as International Drive could not be monitored.

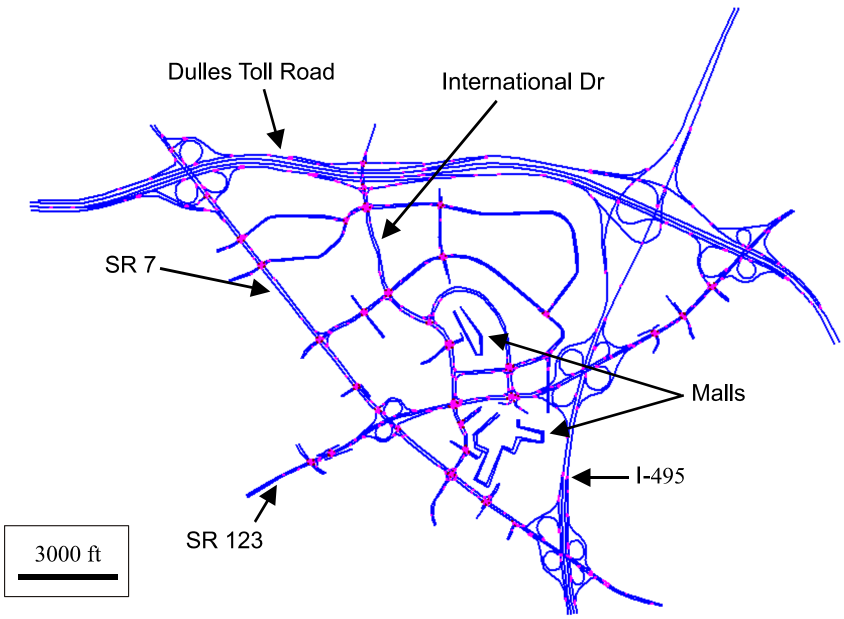

Figure 10. Tyson's Corner Case Study Network

Table 19. Tyson's Corner Case Study: Characteristics of Monitored Links

Link	Length (ft)	Maximum Required No. of Samples	Mean Speed (mph)
Dulles Toll Road E–SR 7 to International Drive	1880.2	1.02	58.2
Dulles Toll Road EB– International Drive to I-495	4386.3	1.55	61.3
Dulles Toll Road WB–I-495 to International Drive	4557.4	1.20	59.5
Dulles Toll Road WB–International Drive to SR 7	2207.8	0.80	57.2
I-495 WB–SR 7 to SR 123	1927.4	5.14	49.1
I-495 EB–SR 123 to SR 7	1024.9	11.48	28.3
SR 7 SB–Dulles Toll Road to SR 123	7143.5	0.50	17.5
SR 7 SB–SR 123 to I-495	4577.0	3.33	17.0
SR 7 NB–I-495 to SR 123	4487.6	2.23	19.6
SR 7 NB–SR 123 to Dulles Toll Road	7662.2	0.20	12.7
SR 123 EB–SR 7 to I-495	3016.4	1.07	9.5
SR 123 EB–I-495 to Dulles Toll Road	2731.2	0.76	8.1
SR 123 WB–Dulles Toll Road to I-495	2729.7	4.24	22.7
SR 123 WB–I-495 to SR 7	3634.6	3.86	17.4

The simulation was not intended to be a perfect replication of the Tyson's Corner area. Numerous unsignalized access points and business driveways in this network could not be simulated because of a lack of data on turning movement counts. It is likely that the conditions as simulated are not a perfect reflection of true conditions on the system since these factors were not included in the simulation.

Number of Samples

Tables B10 and B11 summarize the number of samples for this case study. As with the other case studies, the MHT method almost uniformly increased the number of samples per time interval over the topology method. The median number of samples per time interval and the proportion of intervals that met the CLT-defined minimum sampling requirements increased when the MHT method was used. Tables B10 and B11 show (1) the median sample size per 5-minute interval and (2) the percentage of intervals that met the CLT-based sampling requirements based on the same conditions. Entries in Table B11 shown in bold represent cases where at least 90 percent of the intervals met the CLT-based sampling requirements.

Although there was still disparity in the number of samples on each link, the distribution of speed samples throughout the network was more uniform than with the Springfield case study. In this case study, four relatively high-traffic and/or congested routes were monitored. As a result, there was no single link that acted as a "sink" for the samples. Even though samples were better distributed, lower volume links still did not have as many speed estimates available because of the influence of routes such as the Dulles Toll Road, S.R. 123, and S.R. 7 NB. As a result, routes such as S.R. 7 SB had a relatively low number of data points, and many of the side streets such as International Drive had no data points at all. This is reflected in Table B11, which shows that many of the monitored links met the CLT-based sampling requirements.

The number of samples obtained was examined using the LSD test. The test was used to evaluate differences between combinations of map matching and mean sampling frequency for all network types. The data analysis was segregated by whether a link type was an arterial or freeway segment. Table 20 summarizes the results. The results show that the combination of MHT and a 60-second sampling interval produced significantly more data points that did all other alternatives for both arterials and freeways.

Table 20. Tyson's Corner Case Study: LSD Results for Number of Samples by Map-Matching Type and Mean Sampling Frequency

Road	Map-matching Type	Frequency	Mean	Standard Error	Significantly Different From
Arterial	Topology	15	7.37	0.410	MHT, 15 MHT, 60
		60	9.08	0.578	MHT, 15 MHT, 60
	MHT	15	14.00	0.680	All others
		60	20.87	1.051	All others
Freeway	Topology	15	4.59	0.238	MHT, 15 MHT, 60
		60	4.11	0.223	MHT, 15 MHT, 60
	MHT	15	7.48	0.378	All others
		60	13.98	0.628	All others

Accuracy of Speed Estimates

The percentage of speed estimates within 10 mph for a particular 5-minute interval is shown in Table B12. Cases where at least 90 percent of intervals were within 10 mph are shown in bold. In general, there is a lot of agreement between the roads that met the CLT-based sampling requirements and the roads that achieved a high proportion of intervals within 10 mph.

Table B12 shows that the system did a relatively good job of providing system coverage for the network, especially when the MHT matching was used. The exceptions to this were (1) when links had a high-speed variance (and therefore a high CLT-based sampling requirement) such as the I-495 links and (2) when the traffic volume on a route was lower than on the other major routes (such as S.R. 7 SB). Accuracy of speed estimates for links with limited traffic was improved by using the 15-second mean sampling frequency, although the majority of the network produced more accurate speed estimates with the 60-second mean frequency.

Many of the potentially problematic conditions in this case study did not generate significant errors. In general, wrong-way matching did not cause as much of a problem with this network as with Springfield. Parallel routes had relatively similar speeds, so these problems were mitigated somewhat and did not create large adverse impacts. In the case of the Dulles Toll Road, the low-speed toll lanes were far enough removed from the through route that speed estimates were not matched to the wrong road. Likewise, the pedestrian traffic at the simulated malls did not bias any results for those facilities, even though one mall did come close to S.R. 7. Differences in the heading of the pedestrians versus the nearby roadways likely account for the lack of effect.

The system did have difficulty in the areas around freeway interchanges. In these cases, very few positions were matched to the networks. In most cases, any positions in these areas were identified as errors by the program and discarded. In this case study, this phenomenon limited the areas that could be monitored. As a result, no link between S.R. 123 and the Dulles Toll Road on I-495 was monitored.

The accuracy of the speed estimates was examined using the LSD test. The absolute value of error was examined to help eliminate the impact of extreme values on the mean errors being tested. Table 21 summarizes the results. The trends in the mean absolute error are less dramatic than trends in the number of data points. For the arterial roads, the 60-second mean sampling interval provided the best results. Although the MHT method appeared to provide slightly better results than the topology method, differences between the two forms of map matching were not statistically significant. For freeways, the topology method with a 15-second sampling interval was the only case that was statistically different from the others, although the MHT, 60-second case did appear to be slightly better than the others.

Table 21. Tyson's Corner Case Study: LSD Results for Speed Estimation Accuracy, Map-matching Type by Mean Sampling Frequency

Road	Map-matching Type	Frequency	Mean	Standard Error	Significantly Different From
Arterial	Topology	15	7.11	0.212	All others
		60	5.27	0.253	Topo, 15 MHT, 15
	MHT	15	6.19	0.184	All Others
		60	4.85	0.163	Topo, 15 MHT, 15
Freeway	Topology	15	8.19	0.377	MHT, 15 MHT, 60
		60	7.33	0.289	None
	MHT	15	7.17	0.306	Topo, 15
		60	6.92	0.244	Topo, 15

Processing Time

The processing time for the Tyson's Corner case is shown in Figure 11. The map-matching time for the topology and the MHT methods were very close, so there was not a significant computational savings when using the topology method for this network. The overall time needed to match the positions to the network is not a significant impediment to implementation of the WLT-based monitoring system. Even when a relatively large number of vehicles is being monitored, the processing time is under 2 minutes to process 1 hour of data.

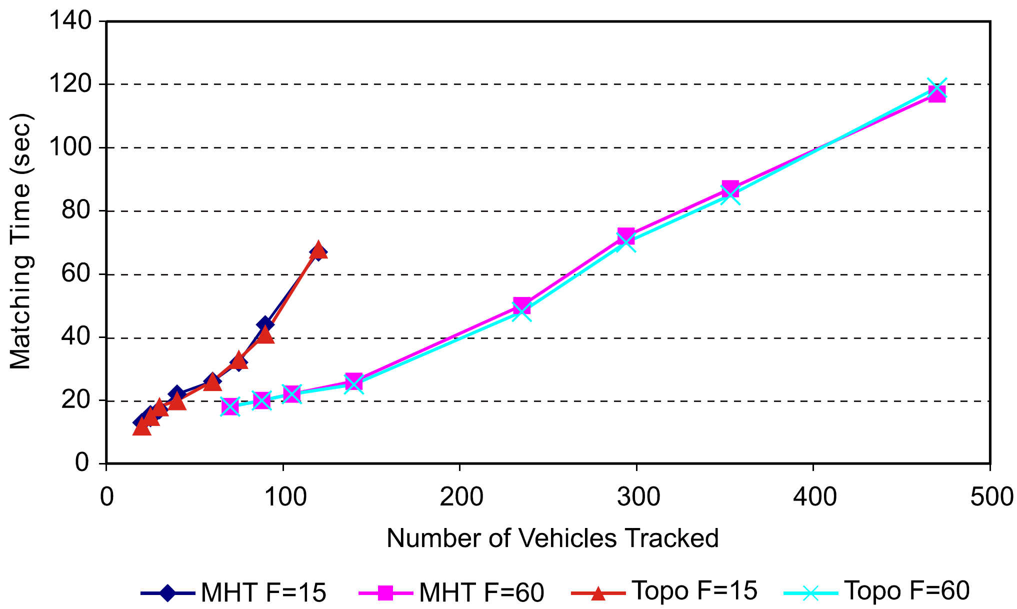

Figure 11. Tyson's Corner Case Study: Matching Time by Number of Vehicles Tracked.

Comparison to Point Detection

The analyses for the case studies focused on defining the performance of the WLT-based monitoring system relative to simulated "reality." Although this is certainly an important measure of performance, it is also worth investigating how WLT-based systems would compare to simulated point detectors on the network. This section compares errors in WLT-based estimates to errors in simulated point detectors for several examples in each case study. This was not a rigorous analysis but rather was intended to show some potential comparisons between these two forms of detection and identify potential areas of application for each. Differences in speed estimation performance between the two types of detection are attributable to (1) the errors created by WLT position estimation and (2) the fact that point detectors gather speed data for all vehicles at a discrete location whereas WLT samples a subset of vehicles over a prolonged distance.

Charlottesville Comparison

For this comparison, point detectors were placed approximately halfway between Angus Road and Hydraulic Road on U.S. 29. The actual speeds on this link were compared to the point detector speeds and the WLT-based estimates. Figures 12 and 13 show some of the speed estimation errors using 5-minute intervals. Positive numbers indicate situations where the travel speed has been overestimated, and negative numbers show cases where the speed has been underestimated.

In both cases, point detection overestimated speeds on the link and WLT provided relatively accurate speed estimates. Point detection cannot easily capture the control delay at a traffic signal, whereas WLT-based systems can determine this time. As a result, point detection may generate arterial speed estimates that are much larger than a true average speed over a link. This indicates that WLT-based systems could offer significant improvements in speed estimation on arterial networks.

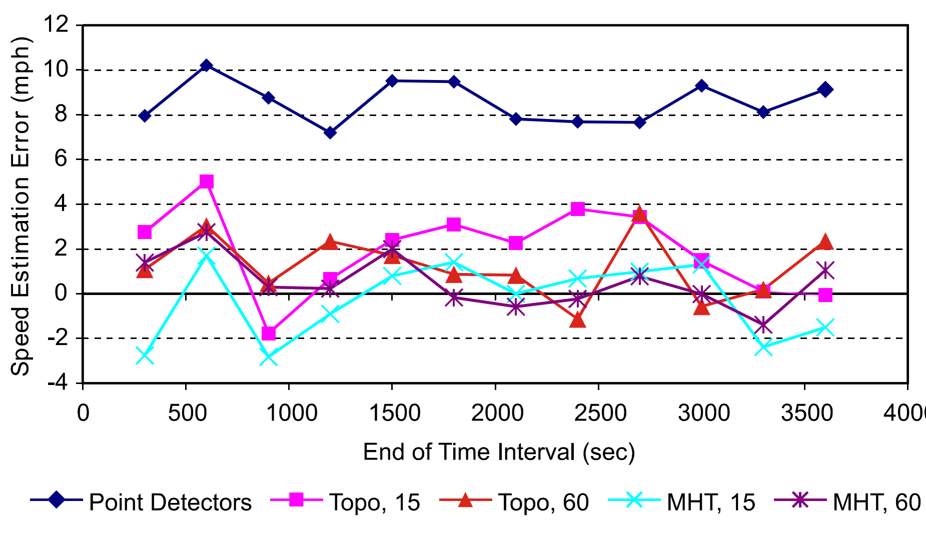

Figure 12. Charlottesville Case Study: US 29 S Speed Estimation Errors: WLT vs. Point Detection

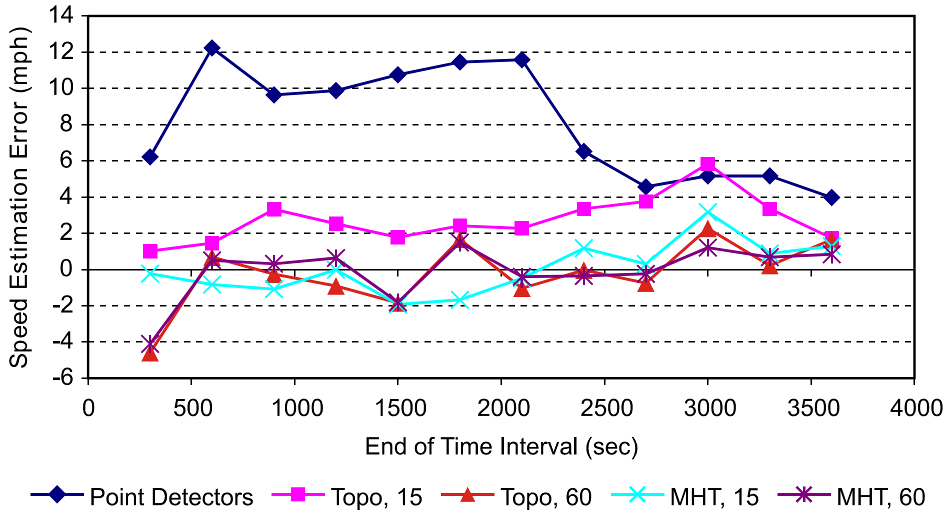

Figure 13. Charlottesville Case Study: US 29 N Speed Estimation Errors: WLT vs. Point Detection.

Springfield Comparison

Point detectors were compared to WLT-based results on I-95 between Van Dorn and I-395 in the Springfield case study. Unlike the Charlottesville example, this is a freeway facility and was one of the situations where the WLT-based accuracy was poor in one direction. This example was chosen to try to identify a situation where point detectors might offer superior performance. Figures 14 and 15 show the speed estimation accuracy for I-95 N and S, respectively.

These two figures highlight some of the problems with speed estimation that occurred in the WLT-based monitoring system. For I-95 N, the system offered improved speed estimation accuracy over point detection. This was a case where traffic speeds were very low, so many samples were available on the road with the WLT-based system. For I-95 S, point detection produced far superior results to those with the WLT-based system. In this case, speeds were

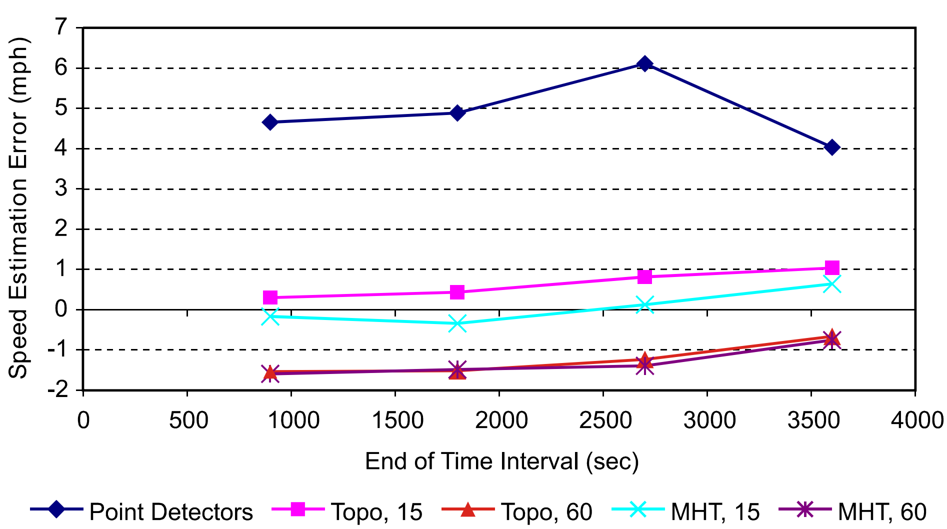

Figure 14. Springfield Case Study: I-95 N Speed Estimation Error: WLT vs. Point Detection.

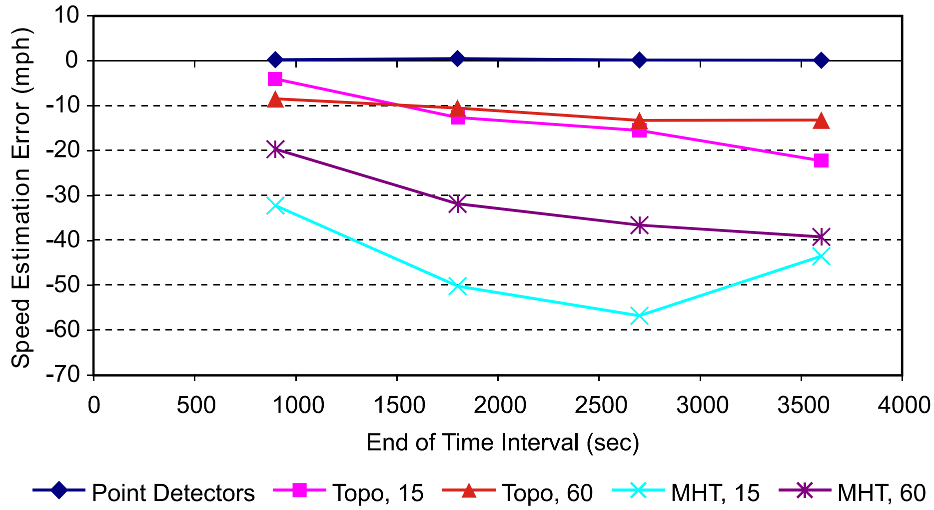

Figure 15. Springfield Case Study: I-95 S Speed Estimation Error: WLT vs. Point Detection.

relatively uniform along I-95 S, which represents an ideal application for point detection. In addition, there were relatively few samples on this link for the WLT-based systems and wrong-way matching problems introduced significant errors into the speed estimation process. Improved data screening and larger sample sizes would be needed to generate better WLT-based speed estimates for I-95 S.

Case Study Findings

Several recurring trends emerged from the three case studies. The key findings were as follows:

- *As the area being monitored increased in size, traffic conditions tended to vary considerably across the facilities in the monitoring area.* The WLT system characteristics must be able to be tailored to a relatively small scale for the system to be effective. Attempting to monitor large networks without any type of zone-specific sampling will likely result in large gaps in the monitoring system. Random sampling of the entire network will result in a disproportionately large number of vehicles being sampled on congested roads, which ironically often need the smallest number of samples because of their low speed variance.

- *The MHT method significantly increased the number of samples available over the topology method, almost always by a statistically significant margin.* The MHT method should be used to match vehicles to the roadway network.

- *Longer sampling frequencies produced better speed estimation results when there were long, continuous links on the network. Shorter sampling frequencies produced better speed estimation results when link travel times were short, because of short lengths, high speeds, or a combination of both.* This points to the need to tailor system characteristics to specific components of the network.

- *Data screening methods need to become more robust to deal with the problems posed by wrong-way matching between free-flowing and congested links.* This problem was most pronounced for the MHT matching methods, but this was primarily because the MHT method generated so many more samples than did the topology method. This provided a greater opportunity for mismatches to create problems.

- *Non-vehicular sources did not impact speeds, provided their orientation of travel was different from that of the surrounding roadway. For similar reasons, the matching process had difficulty discerning the difference between HOV lanes and the main lanes because their orientations were similar, if not identical.*

- *Given the large number of possible paths at freeway interchanges, the matching process could not accurately determine vehicle locations within the confines of the interchange.*

CONCLUSIONS

General

- *Although WLT monitoring systems have some advantages over traditional point detector systems, the WLT systems tested to date were not well suited to the traffic monitoring application.* Further refinement of these systems is necessary before they can provide useful operational data.

- *System design and location post processing plays an important role in the effectiveness of a WLT-based monitoring system.* Monitoring system design can create statistically significant impacts on the accuracy and availability of speed estimates. Systems must be designed to handle localized conditions and use well-developed post process techniques to ensure that data quality is sufficient for traffic management purposes.

- *Some roadways are more likely to be successfully monitored using a WLT system than are others.* As roadway network geometry becomes more complex, WLT-based monitoring systems are likely to have more difficulty in distinguishing the true location of vehicles. The high speed variance present on urban arterials can also create difficulties for WLT-based monitoring systems. Isolated freeways or roads with relatively simple geometric conditions and uniform traffic flow characteristics are likely to be able to be monitored successfully.

- *These findings should be weighed when WLT-based monitoring systems are being considered for deployment.*

Map Matching

- *A reliable, accurate map-matching method can produce statistically significant improvements in the number and accuracy of speed estimates.*

- *Map-matching techniques originally developed for in-vehicle navigation systems can be successfully adapted and applied to WLT-based systems.* These algorithms were adapted and modified to ensure that they provided the real-time matching results needed for these applications. Using well-developed map-matching techniques offers the opportunity to improve the usage of data significantly over the simple geometric methods used in previous evaluations.

- *On simple networks, the modified topology and MHT methods performed well without the input of vehicle-based information and did not have a significant computational load.*

- *The MHT method generally produced better coverage and accuracy than the topology method in the simulated real-world case studies.* This method appeared to be the best for handling more realistic, complex geometric conditions.

WLT System Design Parameters

- *Larger errors in position estimation translate into larger errors in speed estimation.* Systems should seek to minimize location error if possible, but the mean time between samples can be extended to help reduce the impact of large errors in location estimation if needed.

- *Speed estimation errors are most pronounced when there is a short mean time between position readings and a high location error.* Errors are generally minimized by using longer times between position samples. Links that are either short or have high travel speeds are the exception. In those cases, a short mean time between samples should be used to ensure system coverage. The requirements to balance system accuracy and system coverage must be weighed by the designer based on the specific characteristics of the individual road being monitored.

Network Characteristics

- *The processing approach as tested had difficulty generating accurate speed estimates in the areas around freeway interchanges.* Only mainline freeways, away from interchanges, could be successfully monitored.

- *The algorithms had difficulty when there were large differences in traffic flow speed between parallel facilities with similar orientations.* Specifically, vehicles on the low-speed road were occasionally mismatched to the higher speed road. Improved data screening may correct this problem, however.

- *The map-matching process could not successfully distinguish between concurrent* facilities with different characteristics, such as HOV lanes and general-purpose lanes. These distinctions could be made where HOV lanes diverged from main lanes.

- *Non-vehicle wireless sources did not generate significant problems, primarily because their orientation was not parallel with that of the facilities being monitored.* The lack of impact by non-vehicular sources can be seen by the fact that speed estimates in the area near these sources did not drop considerably.

Sampling and Accuracy Requirements

- *Long, congested links act as "sinks" for the samples.* This results in a larger number of samples on slow moving links and relatively few on faster moving links. This can create problems with sample size in cases such as the Springfield case study where several congested facilities were over sampled and uncongested roads were under sampled. If networks are small (as with Charlottesville) or the density of traffic is approximately evenly distributed over the monitored roads (as with Tyson's Corner), this is less of a problem.

- *When sample sizes are reasonably large, the relationship between the percentage of intervals within 10 mph and the number of speed observations appears relatively strong.*

- *More samples than the minimum predicted by the CLT are required to generate good speed estimates.* This is likely attributable to the impacts of errors in location estimation and potential over sampling of lower speed vehicles.

- *The Charlottesville and Tyson's Corner case studies show that WLT-based systems can produce reasonably accurate speed estimates.* Accuracy levels should be sufficient for the purposes of highway performance measurement or the determination of whether a road is congested or uncongested.

- *In general, the WLT-based monitoring speed estimates compared favorably to those for the point detectors.* Speed estimates for WLT were much closer to truth than those for point detectors for arterial systems, but point detection appeared to perform better when speeds were high and traffic flow was uniform (such as uncongested freeways). WLT-based monitoring has the potential to improve condition estimation on arterial roads since it can capture control delay.

- *For large networks, gaining a finer-grained control over system sampling would be desirable.* On large networks, being able to set specific sampling requirements for different regions of the network could ensure that minimum probe penetration and sampling is maintained. Likewise, dynamic sample re-allocation could be tested to ensure that heavily congested links do not act as sampling sinks. This level of control would likely need to be at the level of an individual cell.

GUIDELINES FOR IMPLEMENTATION OF A WLT-BASED SYSTEM

Network Type

WLT-based systems can monitor all facility types within certain bounds. The following are defining guidelines for the types of networks that should be monitored using these systems:

- *WLT-based monitoring systems should not be used to monitor networks that contain long, congested facilities in close proximity to lower volume or higher speed streets. The high-volume, congested facilities act as sinks for sampling.* Dynamically dropping vehicles if too many are found to be assigned to a particular link might overcome this problem. These dropped samples could then be re-assigned to other links. Although this would improve system coverage, it would likely increase the load on the cellular network. Alternatively, system design could be specified on a smaller scale, allowing minimum sampling requirements to be maintained for lower volume roads.

- *WLT-based monitoring systems should not be used to monitor roads with concurrent facilities, such as HOV lanes or rail in the median of a road.* The system cannot distinguish these facilities from the general purpose lanes even when location errors are relatively small. These facilities can be distinguished if their orientations differ, but completely concurrent roads cannot be distinguished from parallel roads.

- *On freeways, only the mainline should be monitored with the system.* The system could not perform matches around freeway interchanges because of the high number of possible locations for a match.

- *The system as tested should not be used to monitor roads where one direction of a road was moving at much different speeds than the other direction.* This is especially important if the MHT method is used. If these conditions are likely to exist on the network, improved data screening should be investigated to deal with the problem of wrong-way matching.

- *Until data screening is improved and sampling requirements can be more focused spatially, WLT-based monitoring should not be used for high-speed facilities in close proximity to lower speed facilities.* Point detection will be a better option at those locations.

WLT System Design

System design plays a critical role in the accuracy of a WLT-based monitoring system. Features that are critical to system performance include:

- *A WLT-based monitoring system should be able to set system parameters to handle localized traffic conditions.* Factors such as the number of vehicles monitored and the mean time between position readings need to be set at local levels to ensure adequate coverage and accuracy. The practical minimum possible with WLT is the size of an individual cell, approximately 3 square miles. This area is comparable to the size of the Charlottesville network. There are often significant variations in speed in larger areas, which suggests that zonal systems that can alter sampling characteristics by cellular coverage area are needed to accommodate localized conditions.

- *The system should use a form of map matching that relies on the geometric characteristics of the network.* The MHT and topology methods produced reasonable estimates of system speeds in conjunction with other characteristics, with the MHT method performing better in nearly all situations. Improved data screening should be applied to improve the performance of the MHT method in situations where high-speed and low-speed facilities parallel each other, however.

- *The system should generate approximately 2 to 3 times the minimum number of samples required based on the central limit theorem to ensure system accuracy.* This is required to overcome over sampling of slower vehicles and errors created by position estimation and map matching.

- *Probe penetration must be sufficient to generate speed estimates on links.* As a result, agencies should ensure that potential vendors have agreements with enough wireless providers to ensure adequate penetration of wireless probes.

- *The time between vehicle position samples should be set to balance the requirements of speed estimation accuracy and data availability.* Longer mean times between position estimates generate more accurate speed estimates than shorter times between estimates. However, shorter sampling times may be required on short or fast-moving links to ensure adequate system coverage.

- *Errors in location estimation should be minimized to the maximum degree possible.* Lower errors in location estimation translate into lower errors in speed estimation. If high location errors cannot be avoided, then the mean time between readings should be extended to mitigate the influence of errors on speed estimates.

RECOMMENDATIONS

1. *VDOT's Mobility Management Division (MMD) and Smart Traffic Centers (STCs) should continue to support the deployment of WLT-based systems but should be more discriminating in which systems they fund for field tests.*

2. *The MMD and STCs should apply the guidelines proposed in this research to help screen any future WLT-based monitoring systems that are being considered for deployment in Virginia.*

3. *In particular, the MMD and STCs should apply the results of this research to ensure that any future systems deployed in Virginia will be able to:*

— Document the form of map matching used in their system, and provide proof that their method can accurately project inaccurate vehicle positions onto the roadway network.

— Provide an estimate of the mean and distribution of location error for the system.

— Indicate the expected mean time between position samples and its distribution.

— Detail how vehicles are sampled over a network. Vendors that approach WLT system design with a "one size fits all" approach should be viewed with caution. Likewise, VDOT should ensure that the vendor can sample enough vehicles to meet requirements for probe penetration. The information provided by those vendors should be examined in light of the findings of this research to determine if it is likely that the proposed system would perform adequately in terms of accuracy and coverage.

SUGGESTIONS FOR FURTHER RESEARCH

Although this research illustrated the potential feasibility of a WLT-based system, a number of potential future research avenues were also identified. Many of these areas could further refine and improve the potential for using WLT-based traffic monitoring systems. Specific areas where future work is recommended include:

- *Investigation of zone-based WLT-based systems.* In this research, vehicles were randomly sampled on a region-wide basis. As a result, sampling requirements and system designs sometimes had to be generated for large areas, which have varying traffic conditions and sampling requirements. There is the potential to improve system accuracy and sampling efficiency by using a zonal system, where sampling requirements can vary by cell. This would allow sampling parameters to be set based on the localized conditions observed in a cell.

- *Measures to avoid over sampling.* As noted in the case studies, severely congested links act as sinks for samples. The system could be redesigned so that once a link had achieved a minimum number of samples, further vehicles on that link would not be used. Those additional tracked vehicles could be assigned elsewhere on the network.

- *Implementation of dynamic sampling strategies in the WLT-based network.* Green[32] showed the merits of using a dynamic sampling strategy to determine speeds on a link. This type of system should be tested to determine its efficacy within the context of a WLT-based system.

ACKNOWLEDGMENTS

The authors acknowledge Steven Henn for his excellent programming work on this project. Matthew Green also provided input on some of the sampling issues related to probe-based monitoring systems.

REFERENCES

1. Federal Communications Commission. *Fact Sheet: FCC Wireless 911 Requirements.* http://www.fcc.gov/911/enhanced/releases/factsheet_requirements_012001.pdf. Accessed March 3, 2003.

2. Cellular Telephone and Internet Association. *Statistics.* http://www.ctia.org. Accessed August 25, 2004.

3. University of Maryland Transportation Studies Center. *Final Evaluation Report for the CAPITAL-ITS Operational Test and Demonstration Program.* University of Maryland, College Park, 1997.

4. Smith, B.L., Zhang, H., Fontaine, M.D., and Green, M.W. Wireless Location Technology-Based Traffic Monitoring: Critical Assessment and Evaluation of an Early Generation System. *ASCE Journal of Transportation Engineering*, Vol. 130, No 5, September/October 2004, pp. 576-584.

5. Lentner, M., and Bishop, T. *Experimental Design and Analysis.* Valley Book Company, Blacksburg, Va., 1993.

6. Cayford, R., and Johnson, T. Operational Parameters Affecting the Use of Anonymous Cell Phone Tracking for Generating Traffic Information. On *Transportation Research Board 82nd Annual Meeting CD-ROM.* Transportation Research Board, Washington, D.C., 2003.

7. Montgomery, D.C. *Design and Analysis of Experiments.* John Wiley and Sons, New York, 1997.

8. Kempthorne, O. *The Design and Analysis of Experiments.* John Wiley and Sons, New York, 1952.

9. Rappaport, T.S. *Wireless Communications Principles and Practice.* Prentice Hall, Upper Saddle River, N.J., 2002.

10. Lee, W.C.Y. *Mobile Cellular Telecommunications Systems.* McGraw-Hill, New York, 2001.

11. Gonsalves, A. *Location Based Services Market Poised for Growth,* April 27, 2004. http://www.mobilpipeline.com/data_voice/19202084. Accessed July 12, 2004.

12. Ygnace, J-L, Remy, J-G, Bosseboeuf, J-L, and Da Fonseca, V. *Travel Time Estimates on Rhone Corridor Network Using Cellular Phones as Probes: Phase 1 Technology Assessment and Preliminary Results.* French Department of Transportation, Arceuil, France, 2000.

13. Roos, T., Myllymaki, P., and Tirri, H. A Statistical Modeling Approach to Location Estimation. *IEEE Transactions on Mobile Computing,* Vol. 1, No. 1, January-March 2002, pp. 59-69.

14. Smith, B.L., Pack, M.L., Lovell, D.J., and Sermons, M.W. Transportation Management Applications of Anonymous Mobile Call Sampling. On *80th Annual Meeting Preprint CD-ROM*, Transportation Research Board, Washington, D.C., 2001.

15. Yim, Y.B.Y., and Cayford, R. *Investigation of Vehicles as Probes Using Global Positioning System and Cellular Phone Tracking: Field Operational Test.* Report UCB-ITS-PWP-2001-9. California PATH Program, Institute of Transportation Studies, University of California, Berkeley, 2001.

16. White, C.E., Bernstein, D., and Kornhauser, A.L. Some Map Matching Algorithms for Personal Navigation Assistants. *Transportation Research Part C*, Vol. 8, 2000, pp. 91-108.

17. Bernstein, D., and Kornhauser, A. *An Introduction to Map Matching for Personal Navigation Assistants.* New Jersey TIDE Center, Newark, NJ, 1996.

18. Kim, W., Jee, G-I, and Lee, J.G. Efficient Use of Digital Road Map in Various Positioning for ITS. *IEEE Position Location and Navigation Symposium Proceedings*. San Diego, Calif., 2000, pp. 170-176.

19. Greenfeld, J.S. *Matching GPS Observations to Locations on a Digital Map.* Department of Civil Engineering, New Jersey Institute of Technology, Newark, N.J., 2001.

20. Zhao, Y. *Vehicle Location and Navigation Systems.* Artech House, Boston, 1997.

21. Pyo, J.S., Shin, D.H., and Sung, T.K. Development of a Map Matching Method Using the Multiple Hypothesis Technique. In *IEEE Intelligent Transportation Systems Conference Proceedings.* Oakland, Calif., 2001, pp. 23-27.

22. Turner, S.M., Eisele, W.L., Benz, R.J., and Holdener, D.J. *Travel Time Data Collection Handbook.* Research Report 7470-1F. Texas Transportation Institute, Texas A&M University, College Station, 1998.

23. Wright, J. and Dahlgren, J. *Using Vehicles Equipped with Toll Tags as Probes for Providing Travel Times.* Report UCB-ITS-PWP-2001-13. California PATH Program, Institute of Transportation Studies, University of California, Berkeley, 2001.

24. Ullman, G.L, Balke, K.N., McCasland, W.R., and Dudek, C.L. Benefits of Real-Time Travel Time Information in Houston. In *Intelligent Transportation: Realizing the Benefits.* Proceedings of the 1996 Annual Meeting of ITS America, Vol. 1, Houston, Tex., 1996, pp. 189-197.

25. Holdener, D.J., and Turner, S.M. Probe Vehicle Sample Size for Real-Time Information: the Houston Experience. In *Intelligent Transportation: Realizing the Benefits.* Proceedings of the 1996 Annual Meeting of ITS America, Vol. 1, Houston, Tex., 1996, pp. 287-295.

26. Srinivasan, K.K., and Jovanis, P.P. Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network. In *Transportation Research Record 1537.* Transportation Research Board, Washington, D.C., 1996.

27. Chen, M., and Chien, S.I. Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation. In *Transportation Research Record 1719.* Transportation Research Board, Washington, D.C., 2000.

28. Boyce, D., Hicks, J., and Sen, A. In-Vehicle Navigation System Requirements for Monitoring Link Travel Time in a Dynamic Route Guidance System. In *70th Annual Transportation Research Board Pre-print.* Transportation Research Board, Washington, D.C., 1991.

29. Sanwal, K.K., and Walrand, J. *Vehicles as Probes*. California PATH Working Paper UCB-ITS-PWP-95-11. University of California, Berkeley, 1995.

30. Ygnace, J-L., Drane, C., Yim, Y.B., Lacvivier, R. *Travel Time Estimation on the San Francisco Bay Area Network Using Cellular Phones as Probes.* California PATH Working Paper UCB-ITS-PWP-2000-18. University of California, Berkeley, 2000.

31. Green, M.W., Fontaine, M.D., and Smith, B.L. Investigation of Dynamic Probe Sample Requirements for Traffic Condition Monitoring. In *Transportation Research Board 83rd Annual Meeting CD-ROM,* Washington, D.C., January 2004.

32. Green, M.W. *Investigation of Sampling Issues in Probe-Based Traffic Monitoring*. University of Virginia Master of Science Thesis, Charlottesville, 2004.

33. Fontaine, M.D. *Improving the Effectiveness of Wireless Location Technology-Based Traffic Monitoring: An Exploration of System Design Issues.* University of Virginia Ph.D. Dissertation, Charlottesville, 2004.

APPENDIX A

MAP-MATCHING ALGORITHMS

This appendix provides further detail on the specific map-matching algorithms used in this project. The algorithms are described, and relevant equations are presented.

Simple Geometric Map Matching

Overview

The simple geometric map-matching method was tested to provide a baseline against which to compare all other map-matching techniques. In the simple geometric method, each position estimate was simply projected onto the nearest link without any regard for past travel history or network topology. This represents the most basic form of map matching and was not expected to be computationally demanding.

Algorithm Description

In this algorithm, each vehicle position estimate was read sequentially and point-to-curve matching was used to project vehicle positions onto the roadway network. The simple geometric algorithm was adapted from work by Greenfeld, Bernstein, and Kornhauser.[1,2] The perpendicular distance between each vehicle position and each link in the network was calculated using the following equation:

$$\text{distance} = \sqrt{\frac{\left[\left(y_a - y_b\right)x_c + \left(x_b - x_a\right)y_c + \left(x_a y_b - x_b y_a\right)\right]^2}{\left(y_a - y_b\right)^2 + \left(x_b - x_a\right)^2}}$$

where: (x_a, y_a), (x_b, y_b) = Start and end coordinates of a link
(x_c, y_c) = Coordinates of vehicle position estimate

Once distances were determined between the location estimate and all links on the network, the position was projected onto the link that minimized the distance. A simple screening method was implemented to remove potential off-network data points. Any position that was greater than 500 m from a link was discarded. This threshold was set based on WLT network-based system error characteristics to ensure that only positions that were far off network were discarded.

The process of projecting positions onto links that minimized the distance was repeated for each position estimate processed by the program. There was no consideration of whether a path between two consecutive vehicle position estimates was possible. As a result, it was expected that this method would generate some inaccuracies and instabilities in assigning positions to the network.

Map Matching with Topology

Overview

The simple geometric map-matching technique was improved by including consideration of network topology. Topology refers to the spatial relationships between connecting or adjacent features. By adding topology, the map-matching process can ensure that matches are possible given (1) the shape and connectivity of the roadway network and (2) past vehicle trajectories. This method ensures that positions are only matched to roadway links where (1) there is a physically possible path between consecutive position estimates and (2) the vehicle trajectory matches the orientation of the roadway link being considered. It was expected that this method would offer and improvement in accuracy over the simple geometric map matching. The algorithm tested modifies the methods proposed in several previous evaluations.[1,2,3] Several modifications have been made to existing algorithms:

- A preliminary screening technique based on known error characteristics has been incorporated into the algorithm. This helps reduce computation time by eliminating potential unfeasible matches early in the process. This represents an application of the simple probabilistic method.

- The scoring system used to determine which link the vehicles should be assigned to has been modified to eliminate vehicle-centric information. As a result, only network information is used to match to the network.

- The scoring system has been further modified to place equal weights on distance and orientation. Earlier methods were more heavily weighted towards distance. It was expected that orientation would play an important role in the matching process, especially as network complexity increases.

Algorithm Description

Map matching with topology builds upon the simple geometric map-matching method. An additional attribute that has been added to the matching process is link orientation. The orientation of a link describes the direction of travel on the link in polar coordinates. For example, a link traveling from west to east would have a orientation of "0," a link traveling from south to north would have an orientation of "π/2," and link traveling from west to east would have an orientation of "π," and a link going from east to west would have an orientation of "2π." These orientation codes are generated automatically by the test bed program based on the start and end coordinates of a link.

When a probe position estimate is initially read into the system, all roadway links within 3 standard deviations of error of the position estimate are identified. This preliminary screening was preformed to eliminate any obviously unlikely links. This criterion ensures that there is a 99.9 percent confidence (based on the normal distribution) that the true location of the vehicle will be included in the set of selected links. This screening process improves computational efficiency by not allowing the algorithm to attempt to match to any unlikely alternatives.

The initial path match for a particular vehicle is based off the first two probe position readings. Two readings are required for an initial match so that a vehicle trajectory can be established. As new position estimates become available, they are added to the vehicle trajectory that was created based on the two initial matches.

An azimuth of travel was then generated for two successive vehicle position estimates using the following equation:

$$Azimuth = \tan^{-1}\left(\frac{x_2 - x_1}{y_2 - y_1}\right)$$

If $y_2 - y_1 < 0$, π should be added. If $x_2 - x_1 < 0$, 2π should be added

Comparison of the vehicle azimuth and the link orientation will help determine whether a vehicle is matched to the proper link. A performance function was then developed to help determine which link a reading should be matched to. This performance function was evaluated for every link that was determined to be within 3 standard deviations of the most recent position estimate. The performance function was:

$$Score = \frac{D_j}{Min(D)} + \frac{\Delta Azimuth_{ij}}{Min(\Delta Azimuth)}$$

where D = the perpendicular distance between the most recent position estimate, j, and a link

Min(D) = the minimum perpendicular distance between the position estimate and the matched position of all candidate links

$\Delta Azimuth_{ij}$ = the difference between the azimuth formed by the most recent two position estimates (i and j) and a candidate link's orientation

Min(ΔAzimuth) = the minimum difference in orientation between the vehicle trajectory and all candidate link orientations.

This function was calculated for every possible link in the candidate set. The most recent position estimate was assigned to the link that minimized the score variable. This represents a deviation from previous studies by providing equal weight to distance and heading. In previous studies, heading was not weighted as heavily as distance. It was believed that vehicle azimuth will play an important role in the matching process, and should be weighted at least equally with the distance from a link. Likewise, the score eliminates information generated by the vehicle (such as velocity or dead reckoning data), which was used in some previous studies to ensure better matching.

Link connectivity was also considered. Connectivity information was provided to indicate whether links were connected to one another. The user defines a "link look ahead" distance that defines the maximum number of links that can be between two matches for a single vehicle. This was done to ensure that links that cross but do not connect, such as freeway overpasses, are not seen as viable paths by the algorithm. Likewise, this will eliminate cases where positions skip between two parallel roads with similar orientations that do not connect. Given that link lengths will vary throughout a network, the link look ahead distance had to be set to a large number. As a result, these numbers were typically on the order of 15 to 20 links based on the speeds and average link lengths seen on the networks.

MHT Method

Overview

The MHT method developed for this research represented an extension of what was proposed in the work by Pyo, Shin, and Sung.[421] This method had to be adapted considerably to handle the particular data limitations of a WLT-based system. Specifically, a WLT-based monitoring system is unlikely to have some of the information that would be available to an in-vehicle navigation system, such as the number of GPS satellites visible. As a result, the map-matching system had to be altered to handle map matching from a network perspective, rather than a vehicle-centered perspective. This meant that the MHT method was adapted to consider the measured position and trajectory of the vehicle only. In addition, the Kalman-filtering aspect of the earlier research was removed to improve the computational efficiency of the algorithm. This meant that vehicle trajectories were not smoothed since no Kalman filter was used.

MHT map matching relies on maintaining a series of possible paths for each vehicle tracked by the system. A probability that a particular hypothetical path is correct is developed based on (1) the distance of the position estimates from the potential matched location on the network and (2) the difference between the link headings and the estimated vehicle direction of travel. Like the topology method, the MHT method relies on geometric characteristics of the network but maintains multiple potential paths until convergence criteria are met or a vehicle leaves the network. At that point, the vehicle positions are assigned to the path with the highest likelihood of being correct.

Algorithm Description

The MHT process can be generally broken down into three phases: initial matching, hypothesis development, and hypothesis maintenance.

Initial Matching

When the first position estimate for a vehicle is determined, all links that are within 3 standard deviations of error of the position are identified. The vehicle is tentatively matched to each of these links, and the distance between the position estimate and each matched position is calculated. The likelihood of the position is then determined based on the normal distribution of error input by the user. That is, a likelihood is assigned based on the probability that the error is greater than or equal to the distance between the position estimate and its matched location (based on the normal distribution). Thus, when there are large differences between the position estimate and the matched position the likelihood is less than when distances are smaller. This process may produce a large number initial candidate positions on the network.

The candidate positions are then examined to determine if any can be eliminated. If the maximum likelihood that was generated for a particular position is more than 10 times the second largest likelihood, then all positions except for the maximum are removed. Likewise, if any likelihood is less than 10 percent of the maximum likelihood, it is removed from consideration. This will result in the production of fewer initial potential starting positions for a vehicle path.

The thresholds for retaining or discarding potential matches were developed through preliminary testing of the algorithm to determine their impact on matching accuracy and computational time. The literature did not provide any guidance on how to appropriately set these thresholds. Initial testing focused on using thresholds of 5, 10, and 30 percent for removal, and 5×, 10×, and 30× for convergence. Testing revealed that the 5 percent case for removal and the 30× cases for convergence could not be run in real time, although they produced the most accurate results. The testing also revealed that the less restrictive criteria (30% case for removal and 5× case for convergence) reduced computation time but increased the number of matching errors on the network. For the less restrictive cases, matching errors were 39 percent more common than the 10 percent and 10× case. As a result, the 10 percent case for convergence and the 10× case for acceptance were set as the criteria for the MHT method because they balanced matching accuracy with computational time.

Hypothesis Development

Next, an additional position for a vehicle was read into the system. This position was used along with the initial position estimate to develop a series of hypothetical paths through the network. As the next position is read into the system, the program again identifies all links within 3 standard deviations of error from the new location estimate. This will result in a new set of candidate positions for the second position estimate. The likelihood of the second position estimate is again calculated based on the normal distribution.

Next, position pairs are generated for each possible combination of initial candidates and second position candidates. A vehicle azimuth is calculated for each position pair and compared to link orientations, much like the topology method. Connectivity is also checked to ensure whether there is a physical path from the initial match to the second match. Assuming that there is a physical path that is possible between the two points, the path is added to the candidate list for paths for the two points.

Hypothesis Maintenance

In the hypothesis development phase, a list of potential paths was developed based on sets of position pairs. Since the number of hypotheses is likely to increase exponentially as each new position is added to a vehicle track, it is necessary to prune hypotheses that are unlikely to be true to reduce computational demands. A score for the likelihood of a particular path was developed using the following equation:

$$Score = \frac{Initial_j}{Max(Initial)} \times \frac{Final_k}{Max(Final)} \times \frac{Min(\Delta Azimuth)}{\Delta Azimuth_{j,k}}$$

where Initial = Likelihood of previous matched position, j, based on distance and normal distribution

Final = Likelihood of newest matched position, k, based on distance and normal distribution

ΔAzimuth = difference between vehicle azimuth and link heading in radians.

This score was calculated for each potential path. Each element of the score is normalized, and a score that approaches 1.0 represents a perfect vehicle track along a link. Once scores have been calculated for each possible position pair, then the scores are examined to determine if any hypothetical tracks can be eliminated. First, the scores are checked to determine if one path is far superior to all of the others. If the largest score is more than 10 times as large as the next largest score, then all other hypothetical tracks are eliminated and the matched positions are retained as the only estimates of vehicle positions. If any scores are less than 10 percent of the maximum score, then those scores are also discarded.

If multiple hypothetical paths remain following pruning, the process continues by adding additional position readings. The procedures in hypothesis development are repeated, and a new score is calculated by:

$$Score = OldScore \times \frac{Final_l}{Max(Final)} \times \frac{Min(\Delta Azimuth)}{\Delta Azimuth_{k,l}}$$

Where OldScore is the likelihood score calculated after the previous position reading was read into the system. The new score is based on the previous score for the hypothesis, the distance probability of most recent reading, and the azimuth between the two most recent readings. This process is repeated until the pruning criteria are reached or no more position estimates are available for a particular vehicle. If no more estimates are available for a vehicle, then positions are assigned to the path with the highest probability.

Example of MHT

An example of the physical application of the MHT method is shown in Figure A1. In this case, the initial position reading “A” can only be matched to one candidate link. This matched position is determined to be the only possible option, so it is retained. The next position reading, “B,” is within 3 standard deviations of two locations, which results in the formation of two hypothetical paths: Path $(AB)_1$ and Path$(AB)_2$. When the scores are calculated for these two, the results for Path$(AB)_2$ are slightly larger since it is closer to a link, even though the heading deviations are slightly off because of the intersection. As a result, both paths are determined to be viable at this point. When the third position estimate, “C,” is generated, the algorithm determines that there is only one possible location where position “C” can be matched. Since there is no viable path from Path$(AB)_2$ to C, that path is eliminated. As a result, the route from Path $(AB)_1$ to C is retained as Path $(ABC)_{Final}$.

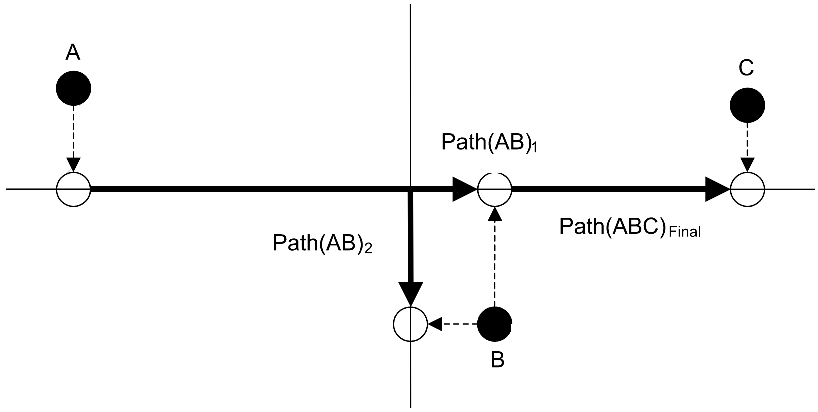

Figure A1. MHT Path Tracking and Finalization.

REFERENCES

1. Bernstein, D., and Kornhauser, A. *An Introduction to Map Matching for Personal Navigation Assistants.* New Jersey TIDE Center, Newark, 1996.

2. Greenfeld, J.S. *Matching GPS Observations to Locations on a Digital Map.* Department of Civil Engineering, New Jersey Institute of Technology, Newark, 2001.

3. White, C.E., Bernstein, D., and Kornhauser, A.L. Some Map Matching Algorithms for Personal Navigation Assistants. *Transportation Research Part C*, Vol. 8, 2000, pp. 91-108.

4. Pyo, J.S., Shin, D.H., and Sung, T.K. Development of a Map Matching Method Using the Multiple Hypothesis Technique. In *IEEE Intelligent Transportation Systems Conference Proceedings.* Oakland, Calif., 2001, pp. 23-27.

APPENDIX B

DATA TABLES FROM CASE STUDIES

Table B1. Charlottesville Case Study: Median Number of Data Points, 5-Minute Intervals

Link	Length (ft)	Map Matching	Required Number of Samples	Median Number of Data Points – F=15				Median Number of Data Points – F=60			
				100	125	150	200	100	125	150	200
US 29 NB – Barracks to US 250	1456.1	Topology	0.41	**3.0**	**3.0**	**4.0**	**5.0**	**1.0**	**1.5**	**1.0**	**2.0**
		MHT		**5.0**	**6.0**	**8.0**	**12.0**	**10.5**	**12.0**	**14.5**	**18.0**
US 29 NB – US 250 to Hydraulic	1427.8	Topology	1.12	**10.5**	**18.0**	**16.5**	**25.5**	**16.5**	**21.5**	**25.5**	**31.0**
		MHT		**24.5**	**38.5**	**37.5**	**51.0**	**51.0**	**63.0**	**77.5**	**93.5**
US 29 NB – Hydraulic to Greenbrier	3222.4	Topology	2.49	**16.5**	**31.0**	**34.0**	**38.0**	**16.0**	**27.5**	**26.5**	**38.0**
		MHT		**31.0**	**50.5**	**54.0**	**71.5**	**43.0**	**61.0**	**63.5**	**84.5**
US 29 SB – Greenbrier to Hydraulic	3225.4	Topology	1.17	**15.0**	**19.0**	**27.0**	**35.0**	**12.0**	**16.0**	**14.5**	**26.0**
		MHT		**26.0**	**34.0**	**48.0**	**62.0**	**43.0**	**51.0**	**60.5**	**91.0**
US 29 SB – Hydraulic to US 250	1529.6	Topology	1.57	**13.5**	**17.0**	**25.0**	**30.0**	**24.0**	**35.0**	**39.5**	**57.0**
		MHT		**28.0**	**30.5**	**45.0**	**51.5**	**47.0**	**66.5**	**78.5**	**106.0**
US 29 SB – US 250 to Barracks	1370.9	Topology	3.76	**4.0**	**5.0**	**6.5**	**10.0**	**5.5**	**9.5**	**14.0**	**19.0**
		MHT		**5.5**	**9.0**	**10.5**	**16.0**	**11.5**	**18.0**	**22.5**	**26.5**
US 250 EB – US 29 to Hydraulic	1127.5	Topology	8.00	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
		MHT		3.5	4.5	4.0	**10.5**	3.5	4.5	8.0	**10.0**
US 250 WB – Hydraulic to US 29	1126.2	Topology	1.21	**2.5**	**3.0**	**3.0**	**3.0**	**2.0**	**2.0**	**3.0**	**5.0**
		MHT		**4.5**	**5.5**	**7.0**	**6.0**	**8.5**	**10.5**	**14.5**	**20.5**
Hydraulic EB – US 29 to US 250	1920.9	Topology	2.03	0.5	0.5	1.0	1.0	0.0	0.0	0.0	0.0
		MHT		**3.0**	2.0	**3.0**	**4.5**	2.0	2.0	**3.0**	**3.5**
Hydraulic WB – US 250 to US 29	1915.7	Topology	1.84	**2.0**	**3.0**	**4.5**	**5.5**	**2.0**	**2.0**	**4.5**	**5.5**
		MHT		**6.5**	**6.5**	**11.0**	**12.0**	**2.5**	**3.0**	**5.0**	**3.5**

Bold text indicates that the median number of samples exceeds the minimum required number of samples.

Table B2. Charlottesville Case Study: Percentage of Intervals with Sufficient Data, 5-Minute Intervals

Link	Length (ft)	Map Matching	Percent of Intervals with Sufficient Data – F=15				Percent of Intervals with Sufficient Data – F=60			
			100	125	150	200	100	125	150	200
US 29 NB – Barracks to US 250	1456.1	Topology	83.3	66.6	83.3	**91.7**	58.3	75.0	58.3	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 NB – US 250 to Hydraulic	1427.8	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 NB – Hydraulic to Greenbrier	3222.4	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 SB – Greenbrier to Hydraulic	3225.4	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 SB – Hydraulic to US 250	1529.6	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 SB – US 250 to Barracks	1370.9	Topology	66.6	66.6	83.3	**91.7**	83.3	**91.7**	**91.7**	**100.0**
		MHT	75.0	**91.7**	**91.7**	**100.0**	**100.0**	**91.7**	**91.7**	**100.0**
US 250 EB – US 29 to Hydraulic	1127.5	Topology	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		MHT	8.3	0.0	8.3	66.6	16.7	8.3	41.7	83.3
US 250 WB – Hydraulic to US 29	1126.2	Topology	66.6	75.0	66.6	**91.7**	75.0	66.6	75.0	**91.7**
		MHT	83.3	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Hydraulic EB – US 29 to US 250	1920.9	Topology	25.0	0.0	16.7	8.3	0.0	0.0	0.0	0.0
		MHT	50.0	41.7	58.3	66.6	25.0	41.7	66.6	75.0
Hydraulic WB – US 250 to US 29	1915.7	Topology	66.6	75.0	**91.7**	**100.0**	50.0	75.0	**91.7**	**91.7**
		MHT	**91.7**	**91.7**	**100.0**	**100.0**	58.3	66.6	75.0	75.0

Bold text indicates that at least 90 percent of intervals exceed the minimum required number of samples.

Table B3. Charlottesville Case Study: Percentage of Intervals Within 10 mph, 5-Minute Intervals

Link	Length (ft)	Map Matching	% Within 10 mph - F=15				% Within 10 mph – F=60			
			100	**125**	**150**	**200**	**100**	**125**	**150**	**200**
US 29 NB – Barracks to US 250	1456.1	Topology	41.7	58.3	75.0	83.3	16.7	0.0	8.3	0.0
		MHT	50.0	58.3	41.7	83.3	0.0	0.0	0.0	0.0
US 29 NB – US 250 to Hydraulic	1427.8	Topology	83.3	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 NB – Hydraulic to Greenbrier	3222.4	Topology	**100.0**	**100.0**	**100.0**	**100.0**	75.0	**91.7**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	75.0	**91.7**	**91.7**	83.3
US 29 SB – Greenbrier to Hydraulic	3225.4	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	75.0	66.7	75.0	83.3
US 29 SB – Hydraulic to US 250	1529.6	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
US 29 SB – US 250 to Barracks	1370.9	Topology	66.7	**91.7**	83.3	**91.7**	**91.7**	**91.7**	**100.0**	**100.0**
		MHT	83.3	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**91.7**	**91.7**
US 250 EB – US 29 to Hydraulic	1127.5	Topology	8.3	25.0	25.0	16.7	16.7	25.0	25.0	41.7
		MHT	41.7	66.7	50.0	**91.7**	**91.7**	83.3	83.3	**91.7**
US 250 WB – Hydraulic to US 29	1126.2	Topology	66.7	83.3	75.0	66.7	33.3	41.7	50.0	41.7
		MHT	75.0	50.0	66.7	75.0	8.3	16.7	83.3	8.3
Hydraulic EB – US 29 to US 250	1920.9	Topology	8.3	8.3	25.0	25.0	16.7	0.0	0.0	0.0
		MHT	41.7	25.0	50.0	58.3	75.0	58.3	50.0	58.3
Hydraulic WB – US 250 to US 29	1915.7	Topology	33.3	50.0	50.0	58.3	50.0	75.0	**91.7**	**91.7**
		MHT	58.3	75.0	83.3	83.3	8.3	8.3	8.3	0.0

Bold text indicates that at least 90 percent of intervals are within 10 mph of the actual speed of then link.

Table B4. Springfield Case Study: Median Number of Data Points, Arterials, 5-Minute Intervals

Link	Length (ft)	Map Matching	Required Number of Samples	Median Number of Data Points – F=15				Median Number of Data Points – F=60			
				100	125	150	200	100	125	150	200
Franconia EB - Commerce to Beulah	8062.5	Topology	3.91	**9.5**	**7.5**	**9.5**	**9.0**	3.0	4.0	3.5	**7.0**
		MHT		**11.5**	**12.5**	**20.5**	**19.5**	**40.0**	**39.5**	**50.0**	**62.5**
Franconia EB – Beulah to Van Dorn	2437.8	Topology	6.35	2.5	2.5	3.0	3.5	3.0	3.5	2.0	5.0
		MHT		4.5	6.0	5.0	**7.0**	6.0	**6.5**	**10.0**	**12.0**
Franconia WB – Van Dorn to Beulah	2466.7	Topology	3.15	1.0	0.0	2.0	2.5	1.0	1.5	3.0	**6.0**
		MHT		2.0	2.0	**3.5**	**5.5**	**8.0**	**6.5**	**9.0**	**13.5**
Franconia WB – Beulah to Commerce	8056.1	Topology	0.34	**9.5**	**13.0**	**15.0**	**14.5**	**35.0**	**46.0**	**65.5**	**88.5**
		MHT		**29.0**	**34.0**	**41.0**	**41.5**	**73.0**	**88.0**	**12.5**	**159.0**
Van Dorn NB – Franconia-Springfield Parkway to Franconia	3588.1	Topology	4.03	0.5	0.5	1.0	2.0	0.0	0.0	0.0	0.0
		MHT		1.5	0.5	3.5	4.0	1.0	2.0	1.5	**4.5**
Van Dorn NB – Franconia to I-95	3058.2	Topology	5.03	4.5	5.0	**6.0**	**6.0**	**6.0**	**11.0**	**9.0**	**16.0**
		MHT		**14.0**	**17.5**	**21.0**	**25.5**	**20.5**	**33.5**	**33.0**	**44.5**
Van Dorn SB – I-95 to Franconia	3062.1	Topology	0.04	**5.5**	**6.0**	**9.0**	**13.0**	**6.0**	**9.5**	**11.5**	**14.5**
		MHT		**13.5**	**19.0**	**19.5**	**31.0**	**46.0**	**51.5**	**68.5**	**87.0**
Van Dorn SB – Franconia to Franconia-Springfield Parkway	3660.4	Topology	5.23	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		MHT		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Bold text indicates that the median number of samples exceeds the minimum required number of samples.

Table B5. Springfield Case Study: Median Number of Data Points, Freeways, 5-Minute Intervals

Link	Length (ft)	Map Matching	Required Number of Samples	Median Number of Data Points – F=15				Median Number of Data Points – F=60			
				100	**125**	**150**	**200**	**100**	**125**	**150**	**200**
Franconia-Springfield Parkway EB – Henry to Frontier	2380.3	Topology	0.88	**1.5**	0.5	**3.0**	**1.5**	0.5	0.0	**1.0**	**1.0**
		MHT		**3.0**	**2.0**	**3.5**	**3.0**	**4.0**	**4.5**	**5.5**	**5.5**
Franconia-Springfield Parkway EB – Frontier to Beulah	3049.1	Topology	1.80	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		MHT		0.0	1.0	0.5	1.0	1.5	1.0	1.0	**2.0**
Franconia-Springfield Parkway EB – Beulah to Van Dorn	7101.2	Topology	3.93	1.0	1.0	1.0	1.0	1.0	1.0	1.5	2.5
		MHT		2.5	3.5	2.0	2.5	**4.5**	**4.5**	**5.5**	**6.5**
Franconia–Springfield Parkway WB - Van Dorn to Beulah	7035.9	Topology	4.50	0.0	2.0	0.5	2.0	2.5	2.0	3.0	**5.0**
		MHT		1.0	2.0	2.0	3.5	4.0	3.0	4.5	**8.0**
Franconia-Springfield Parkway WB – Beulah to Frontier	3034.0	Topology	0.22	0.0	**1.0**	**1.0**	**2.5**	**1.0**	**1.0**	**1.5**	**3.5**
		MHT		**1.5**	**3.0**	**2.5**	**4.0**	**2.0**	**3.5**	**3.5**	**6.0**
Franconia-Springfield Parkway WB – Frontier to Henry	2374.5	Topology	0.55	0.0	**1.0**	**1.5**	**1.0**	**1.0**	0.0	**1.0**	**1.5**
		MHT		0.5	**1.5**	**2.0**	**2.0**	**2.0**	**2.0**	**2.0**	**4.0**
I-95 NB – Franconia to Springfield Interchange	2056.5	Topology	0.74	**1.0**	**1.0**	**2.0**	**3.5**	0.0	0.0	0.5	0.0
		MHT		**2.5**	**1.0**	**3.5**	**5.0**	**1.0**	**2.0**	**1.5**	**2.5**
I-95 NB – Springfield Interchange to Van Dorn	6572.9	Topology	0.06	**48.0**	**52.5**	**61.5**	**67.0**	**183.5**	**289.5**	**330.5**	**492.5**
		MHT		**112.5**	**123.5**	**128.5**	**159.0**	**300.0**	**430.5**	**479.5**	**679.5**
I-95 SB – Van Dorn to Springfield Interchange	5916.9	Topology	0.6	**1.0**	**1.0**	**1.0**	**3.5**	**2.5**	**3.0**	**4.0**	**3.5**
		MHT		**9.0**	**11.0**	**11.5**	**15.5**	**6.0**	**6.5**	**9.0**	**10.5**
I-95 SB – Springfield Interchange to Franconia	2088.8	Topology	11.74	1.0	4.0	5.5	4.0	5.0	7.5	6.0	11.0
		MHT		3.5	8.0	8.0	6.0	11.0	**17.5**	**17.0**	**26.0**

Bold text indicates that the median number of samples exceeds the minimum required number of samples.

Table B6. Springfield Case Study: Percentage of Intervals with Sufficient Data, Arterials, 5-Minute Intervals

Link	Length (ft)	Map Matching	Percent of Intervals with Sufficient Data – F=15				Percent of Intervals with Sufficient Data – F=60			
			100	125	150	200	100	125	150	200
Franconia EB – Commerce to Beulah	8062.5	Topology	75.0	83.3	75.0	**91.7**	41.7	58.3	50.0	66.7
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**91.7**	**91.7**	**100.0**	**100.0**
Franconia EB – Beulah to Van Dorn	2437.8	Topology	25.0	8.3	16.7	16.7	0.0	25.0	16.7	41.7
		MHT	41.7	41.7	33.3	50.0	41.7	50.0	58.3	**100.0**
Franconia WB – Van Dorn to Beulah	2466.7	Topology	0.0	8.3	33.3	16.7	25.0	16.7	33.3	83.3
		MHT	16.7	41.7	50.0	58.3	83.3	**91.7**	**100.0**	**100.0**
Franconia WB – Beulah to Commerce	8056.1	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Van Dorn NB – Franconia-Springfield Parkway to Franconia	3588.1	Topology	0.0	8.3	16.7	0.0	0.0	0.0	0.0	8.3
		MHT	16.7	16.7	33.3	41.7	0.0	16.7	8.3	50.0
Van Dorn NB – Franconia to I-95	3058.2	Topology	33.3	41.7	50.0	75.0	50.0	58.3	58.3	58.3
		MHT	**91.7**	**91.7**	**100.0**	**100.0**	75.0	83.3	**91.7**	**100.0**
Van Dorn SB – I-95 to Franconia	3062.1	Topology	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Van Dorn SB – Franconia to Franconia-Springfield Parkway	3660.4	Topology	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		MHT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Bold text indicates that at least 90 percent of intervals exceed the minimum required number of samples.

Table B7. Springfield Case Study: Percentage of Intervals with Sufficient Data, Freeways, 5-Minute Intervals

Link	Length (ft)	Map Matching	Percent of Intervals with Sufficient Data – F=15				Percent of Intervals with Sufficient Data – F=60			
			100	125	150	200	100	125	150	200
Franconia-Springfield Parkway EB – Henry to Frontier	2380.3	Topology	83.3	50.0	**91.7**	66.6	50.0	41.7	75.0	75.0
		MHT	**91.7**	83.3	**91.7**	**91.7**	**91.7**	**100.0**	**100.0**	**100.0**
Franconia-Springfield Parkway EB – Frontier to Beulah	3049.1	Topology	8.3	0.0	0.0	0.0	16.7	0.0	0.0	8.3
		MHT	0.0	33.3	8.3	25.0	50.0	33.3	58.3	58.3
Franconia-Springfield Parkway EB – Beulah to Van Dorn	7101.2	Topology	25.0	33.3	8.3	41.7	0.0	0.0	16.7	16.7
		MHT	33.3	50.0	25.0	41.7	58.3	58.3	75.0	75.0
Franconia–Springfield Parkway WB – Van Dorn to Beulah	7035.9	Topology	8.3	25.0	16.7	25.0	16.7	16.7	25.0	58.3
		MHT	16.7	33.3	33.3	41.7	41.7	25.0	50.0	66.6
Franconia-Springfield Parkway WB – Beulah to Frontier	3034.0	Topology	33.3	66.0	58.3	**91.7**	66.6	83.3	75.0	83.3
		MHT	58.3	75.0	58.3	**91.7**	**91.7**	**100.0**	**100.0**	**100.0**
Franconia-Springfield Parkway WB – Frontier to Henry	2374.5	Topology	33.3	58.3	66.6	58.3	58.3	33.3	75.0	**91.7**
		MHT	50.0	75.0	75.0	66.6	**91.7**	**91.7**	**91.7**	**100.0**
I-95 NB – Franconia to Springfield Interchange	2056.5	Topology	75.0	58.3	**100.0**	75.0	33.3	16.7	50.0	41.7
		MHT	83.3	83.3	**100.0**	**91.7**	66.6	83.3	**91.7**	**91.7**
I-95 NB – Springfield Interchange to Van Dorn	6572.9	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
I-95 SB – Van Dorn to Springfield Interchange	5916.9	Topology	75.0	58.3	66.6	83.3	**91.7**	**100.0**	**91.7**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
I-95 SB – Springfield Interchange to Franconia	2088.8	Topology	16.7	33.3	41.7	41.7	33.3	50.0	41.7	50.0
		MHT	41.7	50.0	50.0	41.7	50.0	50.0	50.0	50.0

Bold text indicates that at least 90 percent of intervals exceed the minimum required number of samples.

Table B8. Springfield Case Study: Percentage of Intervals Within 10 mph, Arterial Roads, 5-Minute Intervals

Link	Length (ft)	Map Matching	% Within 10 mph - F=15				% Within 10 mph – F=60			
			100	125	150	200	100	125	150	200
Franconia EB – Commerce to Beulah	8062.5	Topology	50.0	16.7	33.3	75.0	50.0	25.0	50.0	83.3
		MHT	41.7	41.7	33.3	50.0	58.3	58.3	83.3	66.7
Franconia EB – Beulah to Van Dorn	437.8	Topology	41.7	33.3	58.3	33.3	66.7	58.3	66.7	75.0
		MHT	58.3	66.7	58.3	66.7	**91.7**	83.3	75.0	83.3
Franconia WB – Van Dorn to Beulah	2466.7	Topology	25.0	41.7	50.0	41.7	16.7	41.7	8.3	8.3
		MHT	16.7	58.3	58.3	16.7	0.0	0.0	8.3	0.0
Franconia WB – Beulah to Commerce	8056.1	Topology	58.3	66.7	66.7	33.3	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Van Dorn NB – Franconia-Springfield Parkway to Franconia	3588.1	Topology	16.7	16.7	50.0	58.3	0.0	25.0	0.0	8.3
		MHT	58.3	33.3	66.7	66.7	25.0	16.7	25.0	16.7
Van Dorn NB – Franconia to I-95	3058.2	Topology	58.3	41.7	75.0	75.0	75.0	66.7	66.7	75.0
		MHT	75.0	58.3	66.7	66.7	75.0	83.3	66.7	75.0
Van Dorn SB – I-95 to Franconia	3062.1	Topology	83.3	83.3	**100.0**	**91.7**	58.3	75.0	75.0	50.0
		MHT	**100.0**	**100.0**	83.3	**100.0**	**91.7**	**91.7**	83.3	83.3
Van Dorn SB – Franconia to Franconia-Springfield Parkway	3660.4	Topology	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		MHT	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Bold text indicates that at least 90 percent of intervals are within 10 mph of the actual speed of then link.

Table B9. Springfield Case Study: Percentage of Vehicles Within 10 mph, Freeways, 5-Minute Intervals

Link	**Length (ft)**	**Map Matching**	**% Within 10 mph - F=15**				**% Within 10 mph – F=60**			
			100	**125**	**150**	**200**	**100**	**125**	**150**	**200**
Franconia-Springfield Parkway EB – Henry to Frontier	2380.3	Topology	58.3	41.7	83.3	25.0	33.3	0.0	25.0	8.3
		MHT	**91.7**	58.3	83.3	58.3	8.3	0.0	0.0	16.7
Franconia-Springfield Parkway EB – Frontier to Beulah	3049.1	Topology	0.0	16.7	0.0	0.0	16.7	8.3	25.0	25.0
		MHT	0.0	16.7	8.3	50.0	41.7	33.3	41.7	50.0
Franconia-Springfield Parkway EB – Beulah to Van Dorn	7101.2	Topology	33.3	58.3	16.7	41.7	50.0	33.3	50.0	58.3
		MHT	66.7	58.3	41.7	58.3	58.3	75.0	75.0	83.3
Franconia–Springfield Parkway WB – Van Dorn to Beulah	7035.9	Topology	16.7	41.7	33.3	16.7	66.7	75.0	66.7	**91.7**
		MHT	25.0	33.3	58.3	41.7	83.3	66.7	58.3	66.7
Franconia-Springfield Parkway WB – Beulah to Frontier	3034.0	Topology	33.3	58.3	25.0	75.0	33.3	41.7	58.3	75.0
		MHT	41.7	75.0	41.7	**91.7**	58.3	83.3	58.3	75.0
Franconia-Springfield Parkway WB – Frontier to Henry	2374.5	Topology	33.3	33.3	58.3	41.7	41.7	25.0	50.0	58.3
		MHT	50.0	66.7	75.0	66.7	83.3	58.3	75.0	50.0
I-95 NB – Franconia to Springfield Interchange	2056.5	Topology	50.0	25.0	58.3	66.7	25.0	8.3	41.7	16.7
		MHT	75.0	50.0	83.3	75.0	58.3	66.7	**91.7**	75.0
I-95 NB – Springfield Interchange to Van Dorn	6572.9	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
I-95 SB – Van Dorn to Springfield Interchange	5916.9	Topology	33.3	25.0	41.7	41.7	41.7	25.0	41.7	58.3
		MHT	0.0	0.0	0.0	0.0	8.3	8.3	16.7	8.3
I-95 SB – Springfield Interchange to Franconia	2088.8	Topology	41.7	58.3	58.3	66.7	58.3	41.7	50.0	58.3
		MHT	25.0	58.3	58.3	66.7	50.0	58.3	50.0	58.3

Bold text indicates that at least 90 percent of intervals are within 10 mph of the actual speed of then link.

Table B10. Tyson's Corner Case Study: Median Number of Data Points, 5-Minute Intervals

Link	Length (ft)	Map Matching	Required Number of Samples	Median Number of Data Points – F=15				Median Number of Data Points – F=60			
				100	125	150	200	100	125	150	200
Dulles Toll Road EB – SR 7 to International Drive	1880.2	Topology	1.02	1.5	1.0	2.5	2.0	1.0	3.0	2.0	3.0
		MHT		2.5	2.0	3.0	4.0	21.0	9.5	7.5	11.5
Dulles Toll Road EB – International Drive to I-495	4386.3	Topology	1.55	1.0	1.0	1.0	2.5	0.5	1.0	1.0	2.5
		MHT		2.5	3.0	3.5	5.0	4.5	7.5	7.5	13.0
Dulles Toll Road WB – I-495 to International Drive	4557.4	Topology	1.20	4.0	5.5	4.0	8.0	4.0	4.0	4.0	7.0
		MHT		5.5	7.5	6.5	14.0	12.0	11.5	14.0	19.0
Dulles Toll Road WB – International Drive to SR 7	2207.8	Topology	0.80	3.0	3.5	3.5	9.5	5.0	6.0	8.0	11.0
		MHT		3.5	4.0	4.0	12.0	12.5	12.5	16.5	25.0
I-495 WB – SR 7 to SR 123	1927.4	Topology	5.14	1.0	3.0	2.0	3.0	1.0	2.0	2.0	2.0
		MHT		2.5	4.5	3.0	5.5	8.0	12.0	12.0	21.0
I-495 EB – SR 123 to SR 7	1024.9	Topology	11.48	4.0	6.0	6.5	10.0	2.0	2.5	3.0	6.5
		MHT		7.0	10.0	12.5	14.0	9.0	10.0	11.0	15.5
SR 7 SB – Dulles Toll Road to SR 123	7143.5	Topology	0.50	0	0	0	0	0	0	0	0
		MHT		1.0	0.5	1.0	2.5	0.5	0	1.0	1.0
SR 7 SB – SR 123 to I-495	4577.0	Topology	3.33	0	1.0	0.5	0.5	0	0	0	0
		MHT		3.5	5.0	5.5	6.5	2.0	3.0	5.0	7.0
SR 7 NB – I-495 to SR 123	4487.6	Topology	2.23	6.0	6.5	5.5	12.0	8.0	7.5	11.5	15.0
		MHT		10.0	12.0	12.5	25.0	19.0	16.0	23.0	31.0
SR 7 NB – SR 123 to Dulles Toll Road	7662.2	Topology	0.20	18.5	16.5	28.5	33.5	26.0	32.5	46.5	51.5
		MHT		31.5	27.0	43.5	61.0	49.0	64.0	79.5	90.0
SR 123 EB – SR 7 to I-495	3016.4	Topology	1.07	3.5	5.0	7.0	7.0	9.0	7.0	9.0	12.0
		MHT		7.0	12.0	19.0	15.5	17.0	19.5	25.0	32.5
SR 123 EB – I-495 to Dulles Toll Road	2731.2	Topology	0.76	2.0	6.0	4.5	8.0	8.5	6.0	10.5	17.5
		MHT		5.5	9.5	11.0	18.0	21.0	19.5	25.5	35.0
SR 123 WB – Dulles Toll Road to I-495	2729.7	Topology	4.24	1.0	1.5	3.5	3.0	1.0	1.0	1.5	1.5
		MHT		3.0	3.0	6.5	6.0	4.0	5.0	4.5	9.5
SR 123 WB – I-495 to SR 7	3634.6	Topology	3.86	4.0	4.5	6.0	10.0	3.0	4.5	4.0	7.5
		MHT		5.5	9.5	11.0	18.0	21.0	19.5	25.5	35.0

Bold text indicates that the median number of samples exceeds the minimum required number of samples.

Table B11. Tyson's Corner Case Study: Percentage of Intervals with Sufficient Data, 5-Minute Intervals

Link	Length (ft)	Map Matching	Percent of Intervals with Sufficient Data – F=15				Percent of Intervals with Sufficient Data – F=60			
			100	125	150	200	100	125	150	200
Dulles Toll Road EB – SR 7 to International Drive	1880.2	Topology	50.0	41.7	58.3	50.0	25.0	75.0	66.7	66.7
		MHT	75.0	58.3	83.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Dulles Toll Road EB – International Drive to I-495	4386.3	Topology	33.3	41.7	41.7	75.0	16.7	41.7	50.0	75.0
		MHT	75.0	83.3	58.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Dulles Toll Road WB – I-495 to International Drive	4557.4	Topology	83.3	**91.7**	66.7	**91.7**	83.3	83.3	**91.7**	**100.0**
		MHT	**91.7**	**91.7**	83.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Dulles Toll Road WB – International Drive to SR 7	2207.8	Topology	75.0	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**91.7**	**100.0**
		MHT	83.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
I-495 WB – SR 7 to SR 123	1927.4	Topology	8.3	25.0	33.3	16.7	16.7	16.7	16.7	33.3
		MHT	25.0	25.0	41.7	50.0	75.0	**91.7**	**100.0**	**100.0**
I-495 EB – SR 123 to SR 7	1024.9	Topology	25.0	25.0	25.0	33.3	0	8.3	0	25.0
		MHT	25.0	50.0	58.3	66.7	50.0	41.7	41.7	75.0
SR 7 SB – Dulles Toll Road to SR 123	7143.5	Topology	33.3	33.3	25.0	25.0	0	0	0	0
		MHT	75.0	50.0	**91.7**	**100.0**	50.0	33.3	66.7	58.3
SR 7 SB – SR 123 to I-495	4577.0	Topology	0	0	0	8.3	0	0	0	0
		MHT	50.0	58.3	58.3	83.3	25.0	41.7	83.3	83.3
SR 7 NB – I-495 to SR 123	4487.6	Topology	75.0	75.0	**91.7**	**100.0**	83.3	**91.7**	**100.0**	**100.0**
		MHT	**91.7**	**91.7**	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 7 NB – SR 123 to Dulles Toll Road	7662.2	Topology	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 EB – SR 7 to I-495	3016.4	Topology	75.0	83.3	**100.0**	**91.7**	**100.0**	**91.7**	**91.7**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 EB – I-495 to Dulles Toll Road	2731.2	Topology	83.3	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 WB – Dulles Toll Road to I-495	2729.7	Topology	8.3	0	16.7	33.3	0	8.3	8.3	0
		MHT	33.3	41.7	75.0	66.7	41.7	50.0	50.0	83.3
SR 123 WB – I-495 to SR 7	3634.6	Topology	66.7	66.7	58.3	83.3	41.7	75.0	58.3	83.3
		MHT	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Bold text indicates that at least 90 percent of intervals exceed the minimum required number of samples.

Table B12. Tyson's Corner Case Study: Percentage of Intervals Within 10 mph, 5-Minute Intervals

Link	Length (ft)	Map Matching	% Within 10 mph - F=15				% Within 10 mph – F=60			
			100	125	150	200	100	125	150	200
Dulles Toll Road EB – SR 7 to International Drive	1880.2	Topology	58.3	58.3	58.3	58.3	50.0	**91.7**	83.3	**91.7**
		MHT	83.3	58.3	58.3	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**
Dulles Toll Road EB – International Drive to I-495	4386.3	Topology	58.3	50.0	58.3	**100.0**	41.7	41.7	50.0	58.3
		MHT	75.0	83.3	75.0	**100.0**	**100.0**	**91.7**	**100.0**	**100.0**
Dulles Toll Road WB – I-495 to International Drive	4557.4	Topology	66.7	**91.7**	50.0	83.3	75.0	66.7	83.3	83.3
		MHT	75.0	**91.7**	66.7	66.7	**91.7**	83.3	**91.7**	75.0
Dulles Toll Road WB – International Drive to SR 7	2207.8	Topology	75.0	**100.0**	66.7	**91.7**	**91.7**	**100.0**	**91.7**	**100.0**
		MHT	83.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
I-495 WB – SR 7 to SR 123	1927.4	Topology	25.0	41.7	41.7	50.0	16.7	41.7	58.3	50.0
		MHT	41.7	58.3	66.7	41.7	83.3	50.0	75.0	66.7
I-495 EB – SR 123 to SR 7	1024.9	Topology	50.0	83.3	66.7	58.3	41.7	41.7	66.6	33.3
		MHT	66.7	**100.0**	83.3	**91.7**	41.7	66.7	33.3	50.0
SR 7 SB – Dulles Toll Road to SR 123	7143.5	Topology	0	0	8.3	0	0	0	0	0
		MHT	25.0	8.3	16.7	33.3	0	0	0	8.3
SR 7 SB – SR 123 to I-495	4577.0	Topology	0	8.3	16.7	16.7	0	0	0	0
		MHT	83.3	50.0	50.0	**91.7**	75.0	66.7	83.3	66.7
SR 7 NB – I-495 to SR 123	4487.6	Topology	75.0	83.3	83.3	83.3	50.0	41.7	75.0	50.0
		MHT	**91.7**	83.3	83.3	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 7 NB – SR 123 to Dulles Toll Road	7662.2	Topology	**91.7**	75.0	**100.0**	**100.0**	**100.0**	**100.0**	**91.7**	**100.0**
		MHT	**100**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 EB – SR 7 to I-495	3016.4	Topology	66.7	83.3	75.0	50.0	**100.0**	**91.7**	**100.0**	**100.0**
		MHT	83.3	**100.0**	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 EB – I-495 to Dulles Toll Road	2731.2	Topology	66.7	**91.7**	83.3	66.7	**100.0**	**100.0**	**100.0**	**100.0**
		MHT	**100.0**	**100.0**	**100.0**	**91.7**	**100.0**	**100.0**	**100.0**	**100.0**
SR 123 WB – Dulles Toll Road to I-495	2729.7	Topology	33.3	41.7	75.0	50.0	50.0	41.7	50.0	66.7
		MHT	58.3	50.0	58.3	50.0	75.0	75.0	**91.7**	75.0
SR 123 WB – I-495 to SR 7	3634.6	Topology	**91.7**	66.7	75.0	83.3	75.0	**91.7**	**100.0**	**91.7**
		MHT	**100.0**	**91.7**	**91.7**	**91.7**	**91.7**	**91.7**	**100.0**	**91.7**

Bold text indicates that at least 90 percent of intervals are within 10 mph of the actual speed of then link.