Final report of ITS Center project: Cellphone probes as an ATMS tool

A Research Project Report

For the National ITS Implementation Research Center

A U.S. DOT University Transportation Center

Cellphone Probes as an ATMS Tool

Principal Investigators:

Dr. Brian L. Smith
Han Zhang
Mike Fontaine
Matt Green

Technical Report Prepared by:

Smart Travel Laboratory
Charlottesville, Virginia
Center for Transportation Studies
University of Virginia
CTS Website http://cts.virginia.edu
351 McCormick Road, P.O. Box 400742
Charlottesville, VA 22904-4742
434.924.6362

June 2003
Smart Travel Lab Report No. STL-2003-01

ABSTRACT

The foundation of traffic operations and management is the ability to monitor traffic conditions. One approach to traffic monitoring is to sample conditions by "tracking" a limited number of probe vehicles as they traverse a network. An emerging technology known as wireless location technology (WLT) has been developed to allow for the geolocation of mobile wireless devices (the most common of which are cellular telephones). Over the past decade, a number of research studies and operational tests have attempted to develop probe traffic monitoring systems based on WLT. However, there still exists significant confusion and misperceptions concerning WLT-based traffic monitoring. To address this problem, this paper seeks to provide a comprehensive assessment of WLT-based traffic monitoring. To do so, the specific purposes of this paper are to: (a) fully describe the concept of WLT-based traffic monitoring, (b) present a critical assessment of past studies of WLT-based traffic monitoring, (c) document the evaluation of one of the most recent operational tests – the 2001 Virginia Department of Transportation (VDOT), Maryland State Highway Administration (MSHA), and US Wireless Corporation (USWC) effort in the Washington, D.C. region, and (d) discuss the unique challenges that these systems pose to the field of traffic engineering.

INTRODUCTION

Since the intelligent transportation systems (ITS) program began to take shape in the early 1990s, transportation agencies at the federal, state, and local levels have focused significant resources on using information technology to improve surface transportation. While ITS has evolved and now takes a myriad of forms, the foundation of nearly every ITS initiative is the ability to measure traffic conditions on the network – generally referred to as traffic condition monitoring. Without knowledge of traffic conditions, transportation professionals can do little to manage traffic or provide traveler information. Furthermore, traffic condition data collected by ITS are now being used in a wide range of transportation applications, such as planning and infrastructure management.

Transportation professionals generally measure traffic conditions using a network of "point" sensors installed at strategic locations throughout the network. While this approach is functional, it suffers from practical limitations that have resulted in incomplete and erratic traffic condition data. The communications, installation, and maintenance costs of the sensor networks have forced transportation agencies to install them only on the most critical routes – often with widely spaced sensors. This leaves many "holes" in the system that lack traffic condition data.

An alternative approach to traffic condition monitoring is to sample a portion of vehicles as they traverse the network. This approach is generally referred to as probe-based monitoring. An emerging information technology, wireless location technology (WLT), holds great promise to provide the platform for a probe-based traffic condition monitoring system. This technology allows wireless devices to be geolocated while in use. WLT supports "location-based services"

(LBSs) in which wireless subscribers (whether using cellular phones, automobile-based telematic devices, mobile computing products, etc.) are provided with targeted information content based on their specific location (i.e. latitude/longitude). Thus, WLT offers an existing and growing infrastructure that can conceptually be tapped to provide probe-based traffic information, without the need for transportation agencies to install extensive infrastructures solely devoted to traffic monitoring.

Given the appeal of this concept, it is no surprise that WLT-based monitoring has captured the interest of the ITS community over the past decade. In fact, there have been a number of relatively large-scale operational tests undertaken in an effort to accelerate development of WLT-based traffic monitoring systems. While these tests have generally been categorized as unsuccessful, the project participants have gained significant insight into the abilities of WLT-based traffic monitoring. Unfortunately, relatively little of this insight has made its way into formal research papers or reports. This has led to significant confusion in the transportation community, and a lack of information on which to base future decisions concerning this technology.

To address this need, this paper seeks to provide a comprehensive introduction to the concepts, experiences with, and performance of early-generation WLT-based traffic monitoring systems. The specific purposes of this paper are to (a) describe the concept of WLT-based traffic monitoring, (b) present a critical assessment of past studies of WLT-based traffic monitoring, (c) document the evaluation of one of the most recent operational tests – the Virginia Department of Transportation (VDOT), Maryland State Highway Administration (MSHA), and US Wireless

Corporation (USWC) effort in the Washington, D.C. region, and (d) discuss the challenges that these systems pose to the field of traffic engineering.

WLT-BASED TRAFFIC MONITORING

WLT can be classified into two general groups: handset-based systems and network-based systems. Handset-based systems rely on global positioning system (GPS)-enabled wireless phones. The GPS unit in the handset determines the location of a phone, and this information is relayed from the phone to a central processing system maintained by the wireless carrier. Network-based systems utilize signal information from cell phones to derive their location. In some cases, network-based systems require special equipment to be installed throughout a metropolitan area in order to analyze signal characteristics of calls. For example, some network-based systems determine positions by analyzing signal power and angle of arrival at multiple cellular towers. In other cases, network-based systems derive location estimates purely from signaling information already available at cellular towers. Since network-based systems do not require users to have GPS-enabled phones, they generally provide less spatial accuracy than handset-based systems.

There have been several attempts to use WLT data to generate traffic condition information. In order to produce this type of information, a WLT system has to be able to perform three basic tasks:

- **Location Determination.** The location of the probe must be determined. These position estimates are usually inaccurate to some degree, and may not lie directly on the roadway network.

- **Map Matching.** In map matching, the position estimates are matched to a specific road. Techniques for map matching vary from simple geometric methods to more complex statistical approaches that account for link connectivity and past travel history.
- **Determination of Traffic Information.** The probes represent a small sample of the entire traffic stream. A WLT-based system must be able to use a set of these samples to estimate the speed or travel time for all traffic on a link.

Several evaluations of early generation WLT-based traffic condition monitoring systems have occurred, but none of these systems adequately performed all three of these tasks. Past evaluations have taken one of two forms: simulation studies or field operational tests. Major findings of these studies are summarized below.

Simulation Studies

Researchers have used simulation studies to explore the potential accuracy and effectiveness of WLT-based systems. While these studies do not replicate the actual conditions precisely, they do provide some indication of the potential performance of a WLT-based system.

French Simulation Study

A study conducted by the French transportation research organization, INRETS, focused on developing a discrete event simulation of traffic flow in order to determine the sample size requirements and accuracy of a hypothetical WLT system (Ygnace, et. al., 2000). The simulation examined the impact of varying levels of probe vehicle penetration on the accuracy of travel time estimates. A location error of 150 meters was assumed, and researchers examined a series of traffic and geometric conditions. The simulation 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. These promising results are based on relatively simple geometric conditions.

Berkeley Simulation Study

A recent evaluation by the Berkeley Institute for Transportation Studies examined factors that could affect the utility of WLT-based traffic monitoring systems (Cayford and Johnson, 2003). 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 five-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, however. The major findings of this research effort were:

- Assuming a network-based system accurate to within 100 meters, at least one probe speed sample 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 accurate to within 50 meters, a probe speed sample 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 tracked probe 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.

Field Operational Tests

Several field tests of WLT-based systems have been performed in the United States. Network-based systems have been used in all cases since GPS-enabled phones currently account for a relatively small portion of available probes.

CAPITAL Field Operational Test

The first major operational test of a WLT-based system was conducted over a 27-month period in the mid-1990's on several interstates and state routes in Virginia (UMD, 1997). This project was named CAPITAL (Cellular APplied to ITS Tracking And Location), and was the result of a cooperative agreement between FHWA, VDOT, MSHA, and several private sector firms. This evaluation produced 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 cellular 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.
- In order to calculate speed, at least four position estimates had to be identified for each phone, and this occurred only 20 percent of the time. As a result, link speed estimates could not be generated for the network. This was due to the small number of data points, as well as a lack of well-developed algorithms to match vehicles to links.

While the CAPITAL test showed that WLT could provide reasonably accurate positional data, it was unsuccessful in producing traffic information.

U.S. Wireless Operational Tests

In addition to the VDOT/MSHA/USWC evaluation documented later in this paper, the U.S. Wireless Corporation (no longer in business) also participated in an operational test in Oakland, California (Yim and Cayford, 2001). Researchers at the University of California - Berkeley obtained 44 hours of wireless location data. 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.

While the researchers were generally able to identify the location of a vehicle, they were not successful in matching vehicles to roads or in generating speed or travel time information. The researchers noted that the call lengths were generally very short, with a median call length of only 30 seconds. This made it impossible to estimate speeds on links since position estimates were not available for long distances. As a result, the researchers were not able to match 60 percent of vehicles to a roadway link.

The most recently completed large-scale field test of WLT-based traffic monitoring was the VDOT/MSHA/USWC effort that took place in the Washington D.C. region beginning in the year 2000. The next sections of this paper introduce this test, describe the evaluation methodology, and present a summary of the evaluation results.

DEPLOYMENT DESCRIPTION – VDOT/MSHA/USWC OPERATIONAL TEST

Beginning in 2000, VDOT and MSHA participated in a WLT demonstration project with USWC in the southern suburban region of Washington D.C.. This region includes the Capital Beltway, a heavily traveled 8-lane freeway that experiences significant congestion, and many major arterials. The purpose of the demonstration project was to demonstrate the feasibility of WLT-based traffic condition monitoring.

EVALUATION METHODOLOGY

The University of Virginia's Center for Transportation Studies served as the traffic monitoring evaluator in this field test (the University of Maryland also served as an evaluator, focusing on location estimation of individual vehicles). This section details the methodology developed for the WLT-based traffic monitoring evaluation.

Data Collection

A key evaluation challenge was to collect accurate baseline traffic data to serve as the ground truth against which WLT-based system estimates of macroscopic traffic parameters could be compared. It was essential that the baseline data accurately represent the actual conditions on the facilities – and that this could be verified through manual means. Two possible approaches to collecting the baseline data were available to the evaluator: probe vehicles and point sensors.

Conceptually, the use of probe vehicles for baseline data collection is desirable. Given that the WLT-based system collects data on "probes" over the entire length of roadway links, baseline probes would allow for the same measurement. However, without tolled facilities in the project region, it was impractical to collect significant samples of baseline probe data. The use of a point sensor, on the other hand, allows the speed of every vehicle passing a point on a link to be collected. However, it requires the key assumption that conditions are uniform throughout the link in order to compare with the WLT-based system results. Traffic volumes can also be determined using point sensors, something that is impossible with a probe-based approach. This is important in that it allows for the assessment of the temporal adequacy of samples from the WLT-based system by providing true population data for the macroscopic traffic parameters. Given these advantages, as well as data collection constraints, it was determined that point sensors would be used as the baseline measurement approach in this evaluation.

A van-mounted video detection system was used for baseline data collection. The video detection system was used to derive point measures (spot speeds and counts) by processing video from a camera mounted on a 45-foot telescoping mast installed on the van. This system provided the flexibility to position the location of collection in the mid-point of WLT-based system defined links (relatively short, 0.4 miles in length), and allowed for manual verification of baseline data accuracy (using a hand-held Lidar unit for speed, and manual counts for volume).

Given that the focus of this evaluation was on the ability of WLT-based systems to support traffic monitoring applications, it was decided that a relatively short "polling" interval should be used in the effort. Based on information in the *Highway Capacity Manual* (TRB, 2000) and recent research on flow rate stability (Smith and Ulmer, 2003) 10 minutes was selected as the polling interval used.

Finally, it was important to clearly define the term "sample" in the data collection effort. In many cases, a WLT-based system may sample the speed of the same vehicle multiple times as it traverses a single link. While it is possible to argue the validity of treating each sample of the same vehicle as an independent sample of link speed (i.e. assuming a vehicle's speed is solely governed by conditions over the entire length of the link), a conservative approach that most directly corresponds to traditional traffic monitoring practice is to consider the average speed of the multiple samples from the same vehicle as the single sample for that vehicle/link pair. This was the approach used in this research.

Comparative Analysis

The comparative analysis included two key components. First, the baseline link population data was used to compute confidence intervals on mean speed estimates to identify minimum required sample sizes for deriving traffic data of particular levels of "quality." These were then compared with the WLT-based system samples to ascertain if the system is theoretically capable of providing sufficient numbers of samples. Second, link data from the baseline video detection system and WLT-based system were directly compared to determine if the WLT-based system produced link mean speed results that accurately reflect the ground truth.

Sample Size Adequacy Evaluation

Given that a WLT-based traffic monitoring system will produce individual vehicle speed samples on links, or samples of the random variable, *U*, and the system is attempting to estimate the mean link speed, μ, the Central Limit Theorem can be used to estimate the number of samples required to estimate μ to within some level of allowable error at an assumed confidence level. This approach is widely used in experimental design and has been used, for example, to determine the number of probes required for speed estimation in the Houston automatic vehicle identification (AVI) based traffic monitoring system (Turner and Holdener, 1996).

Based on the Central Limit Theorem, one can collect *n* probe samples and compute a confidence interval about the population mean as follows:

$$P\left(\bar{u} - \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \le \mu \le \bar{u} + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right) = \alpha \qquad (1)$$

In other words, there is an α probability that the true mean link speed falls in the interval defined above. Note that this effort's methodology allows for the direct computation of the speed population parameters since the van-based video detection system measured the speed of every vehicle traversing the link over the 10-minute interval.

Working with equation (1), if the "width" of the confidence interval is defined as 2*d* miles/hour (i.e., an error of $\pm d$ miles/hour can be accepted in the estimation of the true link mean speed, μ) an equation can be derived to determine the minimum required sample size, *n*. Note that a confidence level, α, must be assumed.

$$n = \left(\frac{Z_{\alpha/2}\sigma}{d}\right)^2 \qquad (2)$$

Finally, it must be noted that this methodology assumes that a WLT-based traffic monitoring system measures the speed of each vehicle traversing a link without error. Certainly, this will not be the case. As such, the values of *n* presented in the results represent the absolute minimum number of samples required. A fielded system will likely require a somewhat larger number of samples to account for errors in the system's ability to measure individual vehicle speeds.

Link Conditions Estimation Evaluation

For each 10-minute interval, mean speeds produced by the WLT-based system were compared with the baseline data to measure percentage error in the speed estimates of the WLT-based system. In addition, hypothesis tests were conducted to rigorously assess the difference in mean speed values. The Wilcoxon Rank-sum Test, the most widely used nonparametric alternative to the independent samples *t*-test, was selected for this purpose. The mean speed of baseline and WLT-based system samples were compared at the 95 percent confidence level using the following hypothesis test:

$$Ho: \mu = \mu_o$$

$$Ha: \mu \neq \mu_o$$

Where,

μ = WLT-based system mean 10-minute interval link speed

μ_o= Baseline mean 10-minute interval link speed

RESULTS

Three major data collection efforts were conducted to gather the baseline data needed for the evaluation. These took place in the Fall of 2001, and included links that could be classified as freeway, high-speed major arterial, and low-volume/speed urban links. Data collection was conducted primarily during daylight hours, however, data were collected in the evening for several hours on a freeway link. It should first be noted that the WLT-based system was unable to reliably collect sufficient samples to estimate conditions on low-volume/speed urban links – therefore no results are available for this category of facility. Table 1 provides detail on the links analyzed in this research.

TABLE 1,

Link Descriptions

Link ID	Road Name	Direction	Speed Limit (mph)	AADT	# Lanes	Description
201	I-495	East	55	65,000	4	Between Telegraph Rd & US1
202	I-495	West	55	73,000	4	Between Telegraph Rd & US1
103	US-1	North	45	58,000 (Bidirectional)	3	Immediately south of I-495 interchange
104	US-1	South	45	58,000 (Bidirectional)	3	Immediately south of I-495 interchange
105	US-1	North	30	64,000 (Bidirectional)	3	Immediately north of I-495 interchange
121	George Washington Parkway	North	40	27,000 (Bidirectional)	1	Immediately south of I-495 interchange
122	George Washington Parkway	South	40	27,000 (Bidirectional)	1	Immediately south of I-495 interchange
242	Duke Street	West	40	23,000 (Bidirectional)	3	Immediately west of US-1

WLT-Based System Sample Size Adequacy Evaluation

The WLT-based system was only capable of sampling a relatively small portion of total vehicles traversing the links analyzed. Table 2 presents mean results for links classified as freeway (daylight), freeway (evening), and arterial. It should also be noted that on many occasions, the

WLT-based system only sampled 1 or 2 vehicles per link per 10-minute interval. During four 10-minute intervals, the WLT-based system did not sample a single vehicle on a link (this occurred twice on link 201, I-495, and twice on link 104, US-1).

TABLE 2.

WLT-Based System Mean Sample Sizes

Link Classification	**Mean WLT Samples (per 10 minutes)**	**Mean Traffic Count (per 10 minutes)**	**Percentage of Traffic Stream Sampled**
I-495 (daytime)	7.0	636	1.1%
I-495 (evening)	3.7	462	0.8%
Arterials	4.1	216	1.9%

While the above results seem to paint a rather bleak picture, it is important to consider needed sample sizes based on the Central Limit Theorem. The Central Limit Theorem shows that relatively small sample sizes are required when speed variance is low, so the small sample sizes observed in the field test may be acceptable. Table 3 provides the minimum sample sizes required for freeway links given two confidence intervals. The first confidence interval was intended to capture the strict requirements of traffic management and control – with the intervals defined as ±5 mph with a confidence level of 99%. The other interval was designed to represent the less strict requirements of traveler information systems – with the intervals defined as ±10 mph with a confidence level of 95%. In addition to these minimum required sample sizes, the actual numbers of samples collected by the WLT-based system are reported.

TABLE 3.

Freeway Link Results – Sample Size

	Time Interval	Total Vehicles	Population Mean Speed μ	Population Standard Deviation, σ	Sample Required (99% C.I., Error=±5mph)	Sample Required (95% C.I., Error=±10mph)	USWC Sample Size
Sept I-495 East Link 201	9:50-10:00	430	68.7	7.9	16.6	2.4	3
	10:00-10:10	466	51.6	21.6	124.2	18.0	5
	10:10-10:20	592	22.0	6.2	10.3	1.5	4
	10:20-10:30	745	20.0	6.4	10.9	1.6	13
	10:30-10:40	824	23.7	6.5	11.2	1.6	6
	10:40-10:50	729	51.8	21.6	123.4	17.9	3
	10:50-11:00	694	67.5	7.7	15.9	2.3	12
	11:00-11:10	697	68.2	8.8	20.4	3.0	3
	11:10-11:20	405	68.0	8.4	18.9	2.7	7
Oct I-495 East Link 201	11:35-11:45	597	65.7	7.4	14.4	2.1	4
	11:45-11:55	613	66.2	7.0	13.0	1.9	6
	11:55-12:05	567	66.0	7.7	15.6	2.3	4
	12:05-12:15	361	46.0	19.1	96.5	14.0	1
	12:15-12:25	574	19.1	6.0	9.6	1.4	5
	12:25-12:35	741	22.4	7.7	15.7	2.3	8
Oct I-495 West Link 202	13:30-13:40	733	66.4	6.4	10.8	1.6	5
	13:40-13:50	732	66.1	6.0	9.6	1.4	6
	13:50-14:00	828	64.9	6.4	11.0	1.6	2
	14:00-14:10	750	65.9	6.3	10.7	1.5	7
	14:10-14:20	705	65.3	6.6	11.4	1.7	6
Nov I-495 East Link 201	17:40-17:50	850	17.0	5.2	7.1	1.0	22
	17:50-18:00	736	14.2	7.4	14.5	2.1	11
	18:00-18:10	756	15.4	6.8	12.4	1.8	6
	18:10-18:20	873	19.2	6.5	11.2	1.6	16
	18:20-18:30	667	40.9	19.1	100.4	14.5	9
	18:30-18:40	523	64.8	7.9	16.4	2.4	4
	18:40-18:50	530	65.7	7.6	15.2	2.2	5
	18:50-19:00	498	65.3	7.5	14.9	2.2	5
	19:00-19:10	510	66.5	7.0	13.1	1.9	0
	19:10-19:20	318	64.9	7.7	15.9	2.3	3
	19:20-19:30	439	67.2	7.7	15.8	2.3	3
	19:30-19:40	470	65.8	7.3	14.2	2.1	0
	19:40-19:50	483	64.6	7.7	15.6	2.3	2
	19:50-20:00	443	67.3	7.2	13.6	2.0	1
	20:00-20:10	352	66.8	6.5	11.1	1.6	3
	20:10-20:20	392	67.4	7.2	13.8	2.0	4
	20:20-20:30	394	66.2	6.9	12.8	1.9	8
	20:30-20:40	363	65.9	7.0	13.1	1.9	4
	20:40-20:50	328	67.2	7.0	13.1	1.9	2

The results presented in Table 3 indicate that the feasibility of an early-generation WLT-based system providing adequate quantities of samples is dependent on accuracy requirements. For the more stringent traffic management and control requirements, the WLT-based system only provided sufficient samples in 3 of the 39 intervals. On the other hand, when the requirements are relaxed to traveler information standards, the WLT-based system met the sample size requirements in 30 of the 39 intervals.

As evident in Table 3, speed variance plays a major role in the minimum sample size requirements. As seen in equation 2, as the standard deviation of U increases, the sample size requirements increase dramatically. For example, when the speeds dropped significantly during the 12:05-12:15 interval on link 201 in October, a minimum of 97 samples are required to produce an average speed estimate within ±5 mile/hour error at a 99% confidence interval. Thus, it is clear that when any probe-based system is deployed in an area that experiences frequent changes in conditions (such as those that experience frequent incidents), the sample size requirements will increase significantly.

The sample size adequacy analysis was also conducted for arterial links. A total of 35 10-minute intervals were analyzed on the major arterials (US 1, Duke Street, George Washington Parkway). Table 4 presents a summary of the results.

TABLE 4.

Arterial Link Results – Sample Size

	Time Interval	Total Vehicles	Population Mean Speed μ	Population Standard Deviation, σ	Sample Required (99% C.I., Error=±5mph)	Sample Required (95% C.I., Error=±10mph)	USWC Sample Size
Oct US-1 North Link 103	15:55-16:05	240	36.0	6.0	9.4	1.4	7
	16:05-16:15	244	35.3	5.7	8.5	1.2	3
	16:15-16:25	268	29.6	11.2	33.1	4.8	1
	16:25-16:35	261	29.6	8.5	19.2	2.8	4
	16:35-16:45	248	32.3	7.5	15.1	2.2	4
	16:45-16:55	261	22.2	10.2	27.4	4.0	3
Oct US-1 South Link 104	17:35-17:45	116	37.3	8.6	19.6	2.8	2
	17:45-17:55	292	33.4	8.2	17.9	2.6	3
	17:55-18:05	290	33.0	8.6	19.8	2.9	0
	18:05-18:15	298	32.8	8.9	20.9	3.0	3
	18:15-18:25	375	26.6	9.5	24.0	3.5	4
	18:25-18:35	338	28.4	7.7	15.6	2.3	4
Oct GW North Link 121	8:25-8:35	251	43.8	8.7	20.2	2.9	2
	8:35-8:45	329	44.5	5.4	7.8	1.1	1
	8:45-8:55	355	44.4	5.6	8.3	1.2	2
	8:55-9:05	282	46.1	6.0	9.4	1.4	3
	9:05-9:15	232	46.7	5.8	8.9	1.3	1
	9:15-9:25	185	46.7	6.2	10.2	1.5	4
Oct US-1 South Link 104	17:35-17:45	116	37.3	8.6	19.6	2.8	2
	17:45-17:55	292	33.4	8.2	17.9	2.6	3
	17:55-18:05	290	33.0	8.6	19.8	2.9	0
	18:05-18:15	298	32.8	8.9	20.9	3.0	3
	18:15-18:25	375	26.6	9.5	24.0	3.5	4
Nov US-1 North Link 105	8:15-8:25	283	15.0	5.7	8.5	1.2	2
	8:25-8:35	299	16.1	5.3	7.4	1.1	7
	8:35-8:45	282	15.9	6.0	9.5	1.4	3
	8:45-8:55	271	17.6	5.2	7.2	1.0	5
	8:55-9:05	271	12.8	4.1	4.5	0.6	9
	9:05-9:15	243	15.1	4.8	6.1	0.9	7
Nov Duke St. Link 242	10:30-10:40	160	27.2	8.3	18.3	2.7	4
	10:40-10:50	178	27.0	9.0	21.6	3.1	5
	10:50-11:00	180	28.1	8.0	17.1	2.5	12
	11:00-11:10	152	28.4	8.7	20.0	2.9	8
	11:10-11:20	211	26.3	9.3	23.1	3.3	8
	11:20-11:30	178	26.3	10.2	27.5	4.0	4

The results presented in Table 4 are consistent with the results for freeway links presented in Table 3. For the more stringent requirements, the WLT-based system only provided sufficient samples in 2 of the 35 intervals. On the other hand, when the requirements are relaxed to ±10 mile/hour error with a 95% confidence interval, the WLT-based system met the sample size requirements in 26 of the 35 intervals.

In general, the sample size requirements are greater for arterial links than for freeway links when controlling for the level of accuracy. This is to be expected given the larger variations in speeds on arterials due to traffic control devices and greater number of access points. Furthermore, the arterials chosen for this case study are major arterials, and likely to experience less variation than many arterial facilities. Thus, the results presented in Table 4 should be considered as a "best case" scenario.

WLT-based System Link Condition Estimation Evaluation

Before summarizing the results of the link speed evaluation, it is important to note that the WLT-based system was unable to produce a speed estimate due to a lack of any samples in 4 out of 74 ten-minute intervals considered in the evaluation. This illustrates that early-generation WLT-based systems struggle with the ability to consistently produce condition data. Table 5 presents the average 10-minute interval mean error for the link classifications considered in this evaluation. Every interval was considered, regardless if the sample size adequacy analysis conducted in the previous section indicated that the interval did not have sufficient WLT samples.

TABLE 5.

Average 10-minute Interval Mean Absolute Error – WLT-based System

Link Classification	Average Mean Absolute Speed Error (mph)	Minimum Mean Absolute Speed Error (Mph)	Maximum Mean Absolute Speed Error (Mph)
I-495 (daytime)	7.2	0.1	23.9
I-495 (evening)	9.2	0.5	22.8
Arterials	6.8	0.1	23.2

The results presented in Table 5 indicate general agreement with the results of WLT-based system sample size adequacy analysis: the WLT-based system sample size is frequently larger than the sample size requirement on the allowable error of ±10 mph, but frequently smaller than the requirement on the allowable error of ±5 mph. One will note that the maximum absolute mean error reaches 23.9 mph for a 10-minute interval. This level of error is far greater than can be accepted for even the least rigorous traffic management or traveler information application. Finally, Figure 1 presents a histogram of WLT-based system mean errors, considering the “sign” of the error (indicating over- or under-estimation). As seen in the figure, there are more than twice as many 10-minute intervals in which the WLT-based system underestimates the mean speed as opposed to overestimating the speed. This may indicate that the WLT-based system is susceptible to including speed samples of stopped or slow moving vehicles located near the link in question.

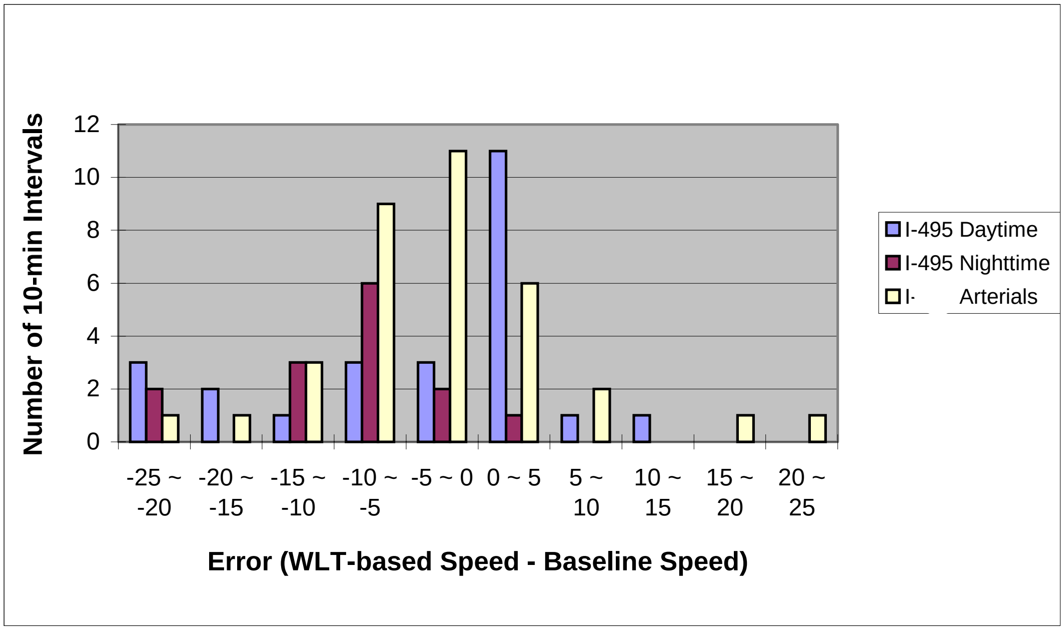

Figure 1. Interval Error Histogram

Finally, hypothesis tests were conducted to investigate the significance of the differences between WLT-based system speed estimates and the baseline data. The WLT- based system proved to be relatively effective for the I-495 freeway links during daytime periods – with only 4.2% of the intervals demonstrating significantly different mean speeds than the population mean. On the other hand, the WLT-based system produced poor performance for freeway links at night and arterial links. In the case of I-495 links at night, 27.3% of the WLT-based system speed estimates were significantly different from the population mean speed. Similarly, for arterials, 20.7% of the WLT-based system speed estimates were significantly different from the population mean speed.

While these results may appear to be fairly positive, one must consider the tests in more detail. As stated earlier, the average WLT-based system sample size is quite small (less than 4) for a number of 10-minute intervals, especially on I-495 nighttime and arterials. On many occasions, the WLT-based system only produced 1 or 2 samples of vehicle speeds on a link in a 10-minute interval. As a result, the small sample size significantly reduces the "reliability" of hypothesis test results. In other words, it is important to recall the underlying premise of a hypothesis test. When one "fails to reject" the null hypothesis, this is not equivalent to stating that that the mean speeds are equal. Essentially, there was just insufficient evidence to reject this claim. In many cases, the evidence is insufficient due to the very small sample sizes.

CHALLENGES POSED TO TRAFFIC ENGINEERING

Considering the results of the case study, one can reach two relatively simple conclusions. First, based on the experience of this field test, early-generation WLT-based traffic monitoring systems are not ready to provide the accuracy and availability needed by modern traffic management systems (and to a large extent, traveler information systems). Second, the ability of an early generation system to generally measure speeds within an error of 10 percent on major routes points to the potential of WLT-based systems.

Readers will certainly note that the evaluation methodology required a number of rather strong assumptions in order to compare the WLT-based system to traditional point detector data. These assumptions were necessary due to the fact that sampling with a WLT-based system (or any probe-based system) is fundamentally different from sampling with point sensors. This section explores these differences and the challenges that WLT-based monitoring systems pose to traffic engineering. First, the limitations that WLT-based monitoring shares with other probe-based systems are reviewed. Next, the issues specific to WLT-based systems are summarized.

Sampling Issues

One of the difficulties of using any probe-based system is determining the relationship between the characteristics of the probe sample and the entire population. Using Equation 2, one can apply the Central Limit Theorem to determine the number of samples needed to estimate the average link speed within some level of allowable error at some confidence level. However, this method assumes that the data within each measurement interval is stationary and that the measured speeds for each vehicle are free of error.

The stationary data assumption may be grounds for concern especially during peak hour periods and incidents when speeds change rapidly with time. Shortening the measurement interval would reduce the variation experienced during the interval. However, shortening the measurement interval would also increase the total number of samples needed over some unit length of time, using more computational resources. Non-stationary data is a problem that is inherent in any sort of travel time or speed estimation system that relies on data measured over some time interval and it must be understood when interpreting results from Equation 2.

The exact data assumption is also troubling in a WLT system. While technology may be improving, there will always be some inaccuracies in measuring travel times and speeds. These inaccuracies are caused by wireless position estimation error, vehicle map-matching error (including errors in distinguishing between phones being used on the road and those being used in office buildings, parking lots, etc.), and underlying error in the road position data on the map. Such errors must be taken into account when using Equation 2 to determine minimum sample size. Future research will need to address these issues before WLT-based systems can become generally accepted.

Role of Speed Variance

Another issue encountered by probe-based systems is how to account for changes in population speed variance. As this evaluation showed, the speed variance of the population plays an important role in the minimum sample size required to determine mean speed. A number of papers have attempted to investigate how to deal with this issue in probe-based monitoring systems.

Chen and Chien (2000) modeled a section of Interstate 80 in New Jersey with CORSIM in an attempt to determine the minimum sample size needed for travel time estimation. They found that, on account of heavy volumes and intense weaving situations, some links may not have normally distributed travel times. For those that did have normally distributed travel times, Chen and Chien determined the number of probes required for five different volume levels, ranging from 50% of the average flow rate to 150% of the average flow rate. Their simulation showed that at 80% and 100% of the average flow rate, only 3 samples were needed per 5-minute interval. However, at very low and very high flow rates (when travel time variance would presumably be greatest), as many as 12 samples were required per 5-minute interval. This shows the importance of variance in determining the minimum required sample size.

Holdener and Turner (1996) used real data from toll tag systems in Houston to determine the minimum sample size needed for vehicle speed estimation. They found that, for 5-minute collection periods, the necessary sample size was between 1 and 4 for a 90th percentile confidence interval and it was between 1 and 6 for a 95th percentile confidence interval.

A number of other papers have attempted to specify minimum required numbers of probe vehicles using simulation and real data. However, the differences seen in their results just go to emphasize the importance of speed variance in determining minimum sample size and the location-specific nature of that variance. While one particular road in a particular area may require a certain number of samples most of the time, another road in another area may have a completely different requirement. Thus, minimum sample size determination must be performed on a location-by-location basis.

Issues Specific to WLT-based Monitoring

While there are similarities, WLT-based systems differ substantially from probe-based traffic monitoring systems that use AVI tags. With AVI tag systems, the location of an individual vehicle can be determined with no ambiguity. WLT-based systems, on the other hand, do not produce data at fixed points, and often have a considerable amount of error in the position estimates that are reported. This positional error can have a direct impact on whether a WLT-based system can match a vehicle to a road, and also on the speed estimates generated for a vehicle. These errors could potential require larger sample sizes for WLT-based systems than probe systems based on AVI readers. Procedures to deal with these issues need to be developed.

There are additional sample size concerns that result from the need to reduce computational overhead in a real-time system. In order to provide travel time estimates based on wireless location data in real time, the WLT requirements must not exceed the capabilities of the computational resources that exist. Furthermore, systems that assess a cost each time a phone is located (which appears to be a likely business model for location-based services) make sample minimization a cost-saving measure as well. Hence it is important to understand the minimum necessary sample size so that the minimum number of vehicle tracks can be sampled to conserve both processing and financial resources.

CONCLUSIONS

WLT-based traffic monitoring is an extremely appealing conceptual approach to collecting traffic information. In fact, the appeal of this concept has led to the relatively large field trials conducted using WLT while the technical aspects of the technology and traffic information derivation were still in their infancy. As seen in the results of the VDOT/MSHA/USWC trial presented in this paper, an early generation WLT-based system produced link speed estimates of moderate quality. This, along with the theoretical capabilities of WLT, give reason for much optimism for use of this technology in future traffic monitoring applications.

In order to make this optimism a reality, however, there is a need for concerted effort to address the significant sampling challenges raised by such systems. AVI tag-based probe vehicle systems provide a starting point, but WLT raises a number of additional challenges. Based on the results of previous field tests and the critical analysis presented in this paper, it is recommended that a basic research program commence addressing the complex sampling and map matching challenges that must be surmounted to make accurate, reliable WLT-based system monitoring a reality.

REFERENCES

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

Chen, Mei and Steven I. J. Chien. "Determining the Number of Probe Vehicles for Freeway Travel Time Estimation by Microscopic Simulation." *Transportation Research Record.* No 1719: 61-68. 2000.

Smith, B. L., Ulmer, J. M. "Traffic Flow Rate Measurement: An Investigation Into the Impact of the Measurement Time Interval." *ASCE Journal of Transportation Engineering.* Vol. 18, No. 3. 2002. Vol. 129, No. 3. 2003.

TRB. *Highway Capacity Manual.* The National Research Council, Washington, D.C. 2000.

Turner, Shawn M. and Douglas J. Holdener. Probe Vehicle Sample Sizes for Real-Time Information: The Houston Experience. *Proceedings of the Sixth International Conference on Vehicle Navigation and Information Systems*, Seattle, Washington. 1995.

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

Ygnace, J-L., J-G. Remy, J-L Bosseboeuf, and V. Da Fonseca. *Travel Time Estimates on Rhone Corridor Network Using Cellular Phones as Probes: Phase 1 Technology Assessment and Preliminary Results.* INRETS Report. Arcueil, France. November, 2000.

Yim, Y.B.Y., and R. Cayford. *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, CA, 2001.