Estimating Delay Time at Signalized Intersections by Probe Vehicles

Kai Liu[1], Toshiyuki Yamamoto[2], and Taka Morikawa[3]

ABSTRACT

Delay times at intersections are one of the most important types of information for ATIS (advanced traveler information systems). Even though probe vehicles have already been used to directly measure delay times, the ongoing collection of real-time delay information on a large scale depends on reducing transmission costs to a reasonable level. Reducing the GPS transmission frequency is one effective means of lowering costs. This paper attempts to assess the sensitivity of delay measurements to differing data transmission intervals. Two delay measurement algorithms are developed corresponding to high-frequency probe data (at 5 second intervals in this paper) and lower-frequency data (from 10 to 60 seconds). Results show that delay detection and measurement sensitivity decrease rapidly with increasing transmission interval. Delays measured from data transmitted at 10-second intervals match the 5-second values in about 74% of cases. The relative accuracy drops to 37% for data transmitted at 60-second intervals.

KEYWORDS delay time measurements, sensitivity, probe transmission interval

INTRODUCTION

Delay times at signalized intersections are one of the most important performance measurements for traffic control systems. As an essential indicator of LOS (level of service) in a signal control system, intersection delay times are also widely employed to construct ATIS (advanced traveler information systems) and estimate vehicle

[1] Doctoral Candidate, Dept. of Civil Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan. Tel: 0081-52-789-3730;
e-mail: liukai@trans.civil.nagoya-u.ac.jp
[2] Associate Professor, Dept. of Civil Engineering, Nagoya University
e-mail: yamamoto@civil.nagoya-u.ac.jp
[3] Professor, Dept. of Environmental Engineering and Architecture, Nagoya University
e-mail: morikawa@nagoya-u.jp

emissions at intersections. Therefore, accurate measurement and better understanding of measurement accuracy are necessary for ITS-related researchers.

Intersection delay time is defined in this study as the difference in travel time between vehicles affected by the signal system and those not affected (and known as total/control delay in some of the literature). Position information received from probe vehicles equipped with GPS offers the ability to trace vehicle trajectories and thus supports the measurement of such delay times directly for individual vehicles. The study by Quiroga and Bullock (1999) is an example of intersection delay time measurement using second-by-second position information from probe vehicles. In this work, an automated procedure was developed to detect critical GPS points where vehicle stops occur. Colyar and Rouphail (2003) improved on Quiroga and Bullock's procedure by taking traffic conditions into greater consideration. The high accuracy of delay times measured in this way is alluring and attractive. However, the enormous cost of transmitting the huge amount of data associated with large coverage areas cannot in general be afforded by most ITS projects. Kühne and Thiessenhusen (2004) integrated the relationship between stopped time and travel time by employing taxi dispatch data from the field that included only timestamp and position information at a longer transmission interval (30 seconds) as a way to reduce operational costs. The accuracy of such data is not yet known, although the authors mentioned the limitations of longer transmission intervals and hence inaccuracies at places where several intersections are located in close proximity.

Generally speaking, the accuracy of delay measurements obtained from probe data decreases inevitably with increasing transmission interval because the reported information is far from continuous. It is hypothesized that only data rates within a threshold time/distance interval can satisfy the requirements for measuring delays with different signal schemes in different cities. It is necessary to prove this hypothesis by measuring the sensitivity of delays determined from probe vehicle data to different transmission frequencies. In this empirical study, an attempt is made to determine such sensitivity of measured delays at arterial intersections.

The different frequencies of probe data will be explained in the next section. Then the test network and the method of sensitivity measurement are introduced. Clearly, the sensitivity of delay measurements to probe data at different frequencies is affected by the magnitude of the delay time itself. Consequently, the method of intersection delay time measurement using probe vehicle data obtained at 5-second intervals is explained in detail, as well as the method used for lower-frequency probe data. Subsequently, the total delay time in a trip is calculated and measurement sensitivity is investigated. Finally, the findings and conclusions of the study are given.

DATA COLLECTION AND CLEANING

Nagoya once had one of the largest probe systems in the world as a result of cooperation between P-DRGS (Probe vehicle based – Dynamic Route Guidance System) consortium and the 32 companies in the Nagoya Taxi Association. Among the total of 1,570 taxis fitted with GPS receivers, ten probe taxis transmitting location at the highest frequency of 12 times per minute (5-second intervals) were selected for this empirical analysis. Data obtained from October 1st, 2002 to March 31st, 2003 are used in the study. A total of 29 items of information, including time, GPS latitude/longitude, speed, acceleration, direction, and distance traveled, are submitted by the GPS receiver in conjunction with a gyroscope device. Of this information, only timestamps and locations are used for intersection delay measurements, because this information would also be available from taxi companies that receive data for taxi dispatch purposes, a possible costless data source that might reduce the enormous operational cost of probe systems. The GPS location accuracy for the data is 15 meters (95% of total data).

Data transmissions at lower frequencies are simulated by selectively deleting parts of the data records. In total, six groups of data are simulated for each trip by each vehicle over the whole collection period. The simulated transmission intervals range from 10 to 60 seconds (abbreviated as 10s data, 20s data, and so on).

The DRM (digital road map) network includes all roads except those less than 5.5 meters in width. The mean link length is about 100 meters. A large number of taxis abound the downtown area of Nagoya, where demand for taxi services is high. All in-service trips that begin, end or pass through the downtown area are selected for study. The road density in this downtown area is a little higher than 25 km/km^2. The selected trips are scattered around the city center and the old Nagoya airport. Then trips where the distance traveled between any two adjacent GPS records was over 150 meters were deleted. Trips including GPS errors or with missing data were also deleted. After data cleaning, a total of 1,090 trips remained.

COMPUTING DELAY TIME

The method utilized to measure the total delay time accurately from the 5s data is crucial to this study. Previous investigations of intersection delay time measurement using probe vehicles equipped with GPS receivers employed similar means for achieving this: a predetermined threshold value of speed or deceleration/acceleration that locates critical points when a vehicle begins to decelerate, stop, and accelerate up to speed. However, the exploration of such delay control points with our 5s data would be difficult and fruitless, because the boundaries at critical points are smoothed out during the process of calculating speed/acceleration so that the sensitivity with which travel conditions are tracked decreases significantly compared with second-by-second GPS data. It is, however, still possible to detect those data intervals in which vehicles begin to decelerate or stop accelerating due to intersection signal control. Intersection delays can be measured by assuming that vehicles go through an intersection at cruise speed if there is no interruption from the signal system. As will be described in more detail in subsequent sections, the maximum value of speed during the previous data interval and at the previous data point is adopted in this research as the normal cruise speed through an intersection.

Quiroga and Bullock (1999) employed a relatively small threshold (between 2.0 and 3.5 km/h/s) to detect the critical boundaries using a forward/backward average acceleration algorithm. Such a scheme is difficult to implement in urban networks where traffic conditions are complicated and the speed value or the change rate of speed can vary significantly. It is evident that low-speed driving and stoppages due to over-capacity or light congestion should be distinguished from intersection delays. Colyar (2003) developed a decision tree to distinguish the two under complicated traffic conditions. However, this algorithm still mistakenly categorizes queue procession with an average speed over the previous 20 seconds of less than 3 mph (about 5 km/h) as a delay resulting from traffic signals. This results in an overestimate of intersection delay. Furthermore most previous studies neglect the fact that the real cruise speed, if unaffected by signals, is not necessarily the free-flow speed or any predefined speed limit.

To overcome such problems, accurate cruise speed values are required. The widely used central difference scheme is employed in this study to calculate the cruise speed. This method has been proved to be effective and simple to use (Mousa, 2002; Quiroga and Bullock, 1999). Figure 1 describes the relationships among record i, time interval T_i, and distance L_i, while the following equations give the speed and acceleration value for record i and distance i respectively.

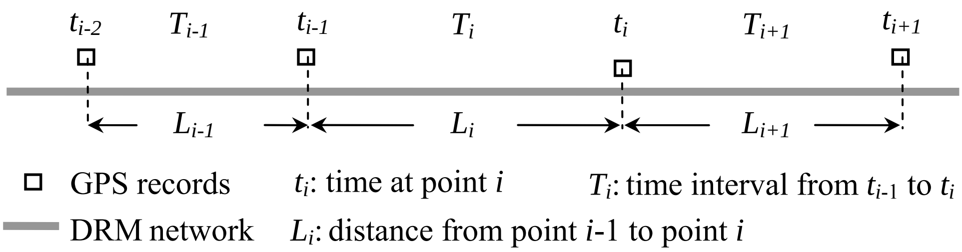


Figure1. Some important variables and their relationships

$$v_i = (L_i + L_{i+1})/(T_i + T_{i+1}) \quad (1)$$

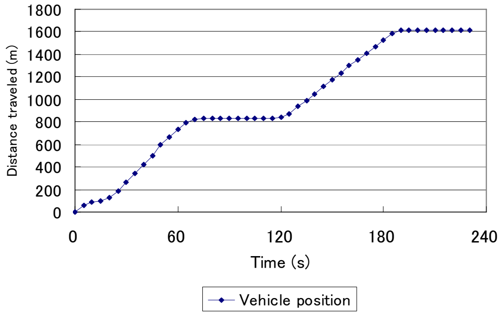
$$V_i = L_i / T_i \quad (2)$$

$$a_i = 2 * (V_{i+1} - V_i)/(T_i + T_{i+1}) \quad (3)$$

$$A_i = (v_i - v_{i-1})/T_i \quad (4)$$

where v_i is the speed at record i; V_i *is* the average speed over time interval T_i; a_i is the acceleration associated with GPS record i; and A_i is the acceleration over time interval T_i. For a time-based transmission scheme, T_i is constant and here it is notated as period i, too.

Figure 2 shows the time−distance diagram for a typical trip. It was made westbound on Sakura-Dori, Nagoya, on January 8, 2003, at about 8:00 p.m. Figure 3 delineates this trip data into a speed-time diagram and an acceleration-time diagram. Both speed and acceleration are the values at the GPS timestamp (v_i and a_i). Ghulam and Rahim (2002) evaluated and compared 12 acceleration models using speed profile data from field data. They found that the acceleration rate increased quickly from a nonzero value to a maximum value and then fell to zero at maximum speed. A scatter plot of speed and acceleration for 100 trips shown in Figure 4 suggests the same relationship. Thus, the constant acceleration threshold does not fit the data at all.



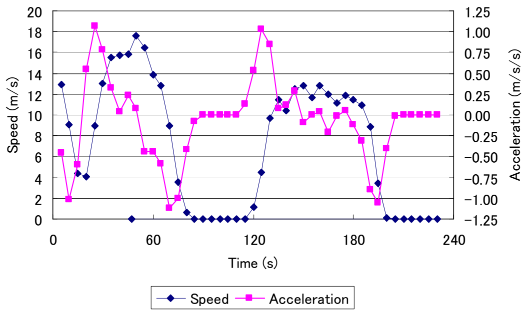


Figure 2. Time-distance diagram

Figure 3. Speed and Acceleration diagram

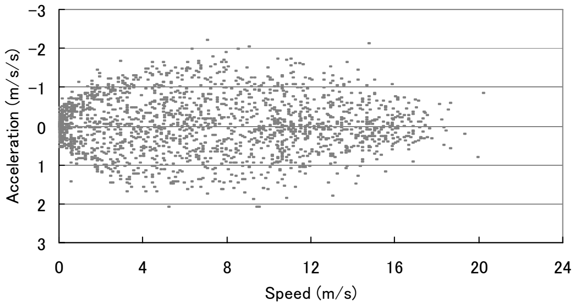


Figure 4. Scatter plot of speed and acceleration for 100 field trips

A first-order differential judgment algorithm corresponding to different acceleration stages is developed in this study. With this algorithm, each GPS point in the time sequence is tested as a delay candidate according to whether the preceding five-second interval includes a delay-related (deceleration, stop or acceleration) event.

A GPS point is first classified according to its point speed. Traffic conditions during the GPS interval are then judged from both the acceleration in the preceding five-second period and the instantaneous acceleration at the GPS point (A_i and a_i). Cruise speed, if unaffected by signals, is then determined after confirming a delay event. The traffic condition differentiation process and cruise speed selection are defined in detail in Table 1. It should be noted that the greatly fluctuating values of speed and acceleration observed in practice are smoothed based on our scheme and thus the threshold value should be a little smaller than usually used.

In practice, when the above algorithm is used, delay times owing to congestion are still included in intersection delays to some degree if more than one stop occurs in a single DRM link, a common phenomenon in the stop-go conditions experienced during queuing. One way of simplifying the problem is to assume that only the final delay in a consecutive series of intervals with delays is regarded as the intersection delay. This can be reserved as the signal delay time while other delays are treated as congestion time. Hence, deceleration/acceleration delay events that happen before or after congestion delays are also treated as normal running conditions.

Table 1. Traffic condition differentiation and real-time cruise speed

Speed	Acceleration	Traffic condition	Cruise speed
Greater than 15m/s	——	Flowing	V_i
10-15m/s	If both A_i and a_i < -0.1m/s/s and Min(A_i, a_i) ≤ -0.6m/s/s	Deceleration	Max(v_j, V_{j+1})*
	If both A_i and a_i > 0.1m/s/s and Max(A_i, a_i) ≥ 0.4m/s/s	Acceleration	Max(v_k, V_k)**
	Otherwise	Flowing	V_i
5-10m/s	If both A_i and a_i < -0.1m/s/s and Min(A_i, a_i) ≤ -0.5m/s/s	Deceleration	Max(v_j, V_{j+1})*
	If both A_i and a_i > 0.1m/s/s and Max(A_i, a_i) ≥ 0.5m/s/s	Acceleration	Max(v_k, V_k)**
	Otherwise	Flowing	V_i
2-5m/s	If both A_i and a_i < 0m/s/s and Min(A_i, a_i) ≤ -0.3m/s/s	Deceleration	Max(v_j, V_{j+1})*
	If both A_i and a_i > 0m/s/s and Max(A_i, a_i) ≥ 0.4m/s/s	Acceleration	Max(v_k, V_k)**
	Else if both v_i *and* V_i ≤5m/s	Congestion	V_i
	Otherwise	Flowing	V_i
Less than 2m/s	——	Stopped	Max(v_j, V_{j+1})*

* *j* is the first preceding five-second period with no delay.
** *k* is the first forward five-second period with no delay.

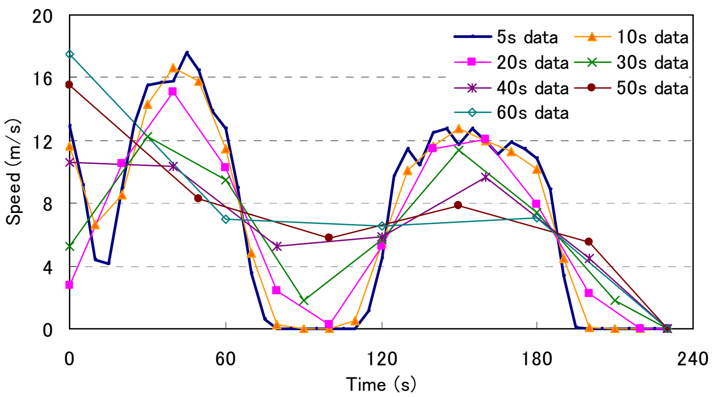


Figure 5.
Speed-time diagram

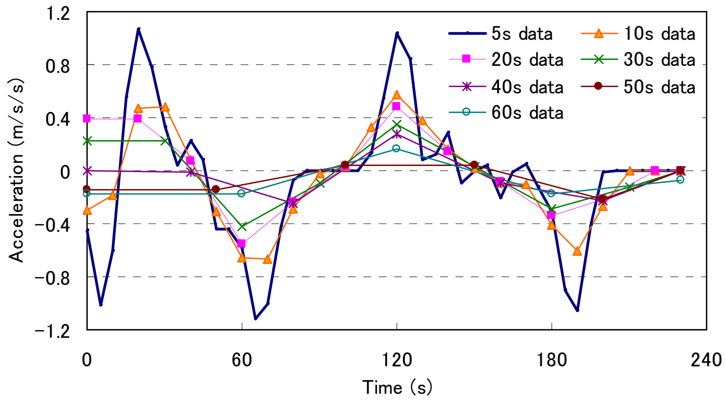


Figure 6. Acceleration-time diagram

Using the lower-frequency data, it is somewhat difficult to differentiate travel conditions. Figures 5 and 6 compare the speed-time and acceleration-time curves for the different frequencies of probe data. Obviously, neither the speed-time diagram nor the acceleration-time diagram appears to efficiently detect where and when delays happen, especially for data with transmission intervals of 30 seconds or more. We propose a simple three-step algorithm to measure delay time in this relatively lower-frequency data. The algorithm begins by searching for the periods during which the vehicle is stopped (judged as a speed of less than 3m/s). The algorithm then searches for cruise periods (that include no delay events), upstream and downstream of the period with the stop, to determine the cruise speed without affected by signals. A simple indicator of speed change rate between the period in question and the previous one was used to judge whether the vehicle is cruising; that is, the vehicle is cruising if the speed change rate is more than 0.8 (backward) and less than 1.2 (forward). Finally, the delay time can be calculated using the method described for 5s data.

RESULTS

Intersection delay times are aggregated on a trip basis. Table 2 presents the average trip delay characteristic based on the 5s data. The table shows that the average delay per stop is approximately 30 seconds. Actually, only 8.7% of all intersection delays range from 20s to 30s, while about 41% of delays are less than 10s. The majority of these short delays include only deceleration/acceleration events and no stops.

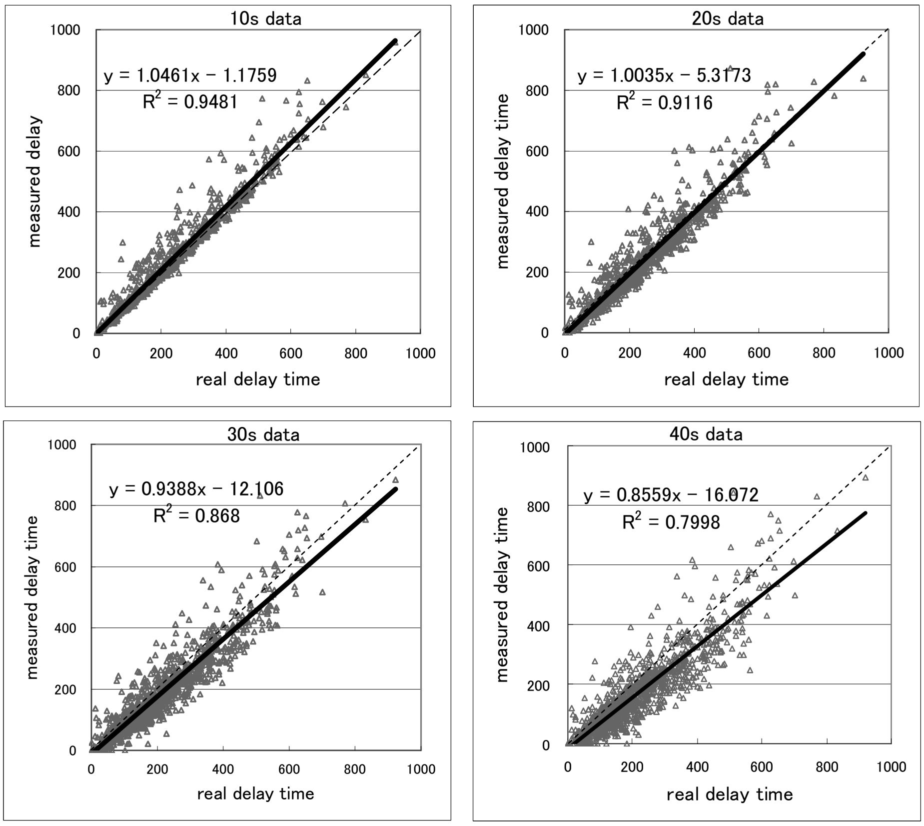
Table 2. Average intersection delays for all selected trips

Avg. trip length (m)	Avg. number delay events per km	Avg. delay per km (s)	Avg. delay/travel time ratio
2,038	3.9	120.8	50.3%

Real cruise speeds are found to range from 3m/s to 20m/s in a normal distribution although the distribution is not shown here due to limitations of space. These speeds are far from free-flow speed or speed limits. The average delay/travel time ratio of

50.3% seems a little higher than would be expected for an efficient traffic signal control system. One possible reason is that one of the trip selection criteria was that the taxi passed through Nagoya city center, where large numbers of pedestrians use the sidewalks. The relatively long signal cycle times (ranging from 120s to 140s) might also help to explain this high delay/travel time ratio.

It is clear that use of the lower-frequency data makes it difficult to differentiate low-speed cruising from short intersection delays, and this would prevent traffic managers from making accurate judgments of traffic congestion. This in turn might result in inappropriate guidance with respect to route choice and departure time. This problem consists of two types of simultaneous judgment error: intersection delays being treated as low-speed cruising and vice-versa. The former always results in an underestimate of intersection delay, or missing it altogether, while the latter leads to an overestimate of intersection delay.



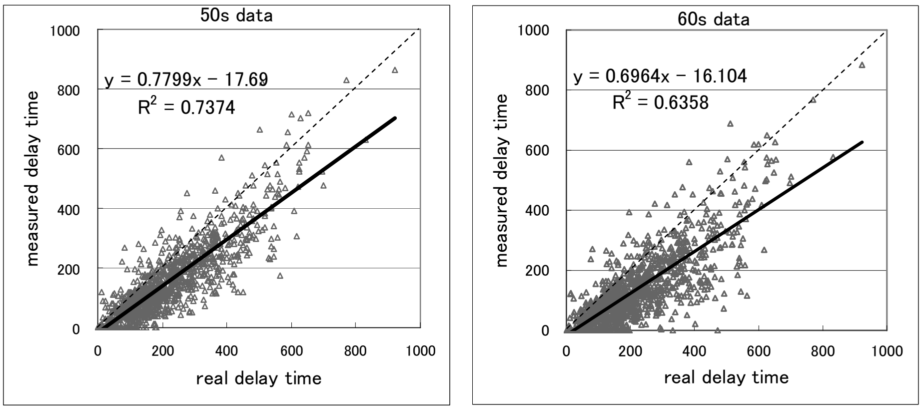


Figure 7. Scatter plots of measured/real delay times for six simulated probe data sets

Figure 7 depicts scatter plots of measured/real delay time for all sample trips corresponding to the six lower-frequency sets of probe data. The 10s data and 20s data give much better measure of delay times; the plotted points are concentrated more on the diagonal than in other cases. The 10s data set is quite distinct in that it gives an overestimate of delay, although not by much, for the majority of trips. As transmission interval increases, more and more intersection delays are underestimated or missed completely.

Sensitivity analysis can be used to examine the ability of lower-frequency probe data to accurately provide delay detection and measurement. Sensitivity is explored by first assuming that high-frequency data (5s data in this paper) is able to provide correct delay measurements and then examining the relative differences between 5s data and other lower-frequency data. Two indicators of sensitivity are investigated: the average ratio of correct judgments and the average error in measured delay time.

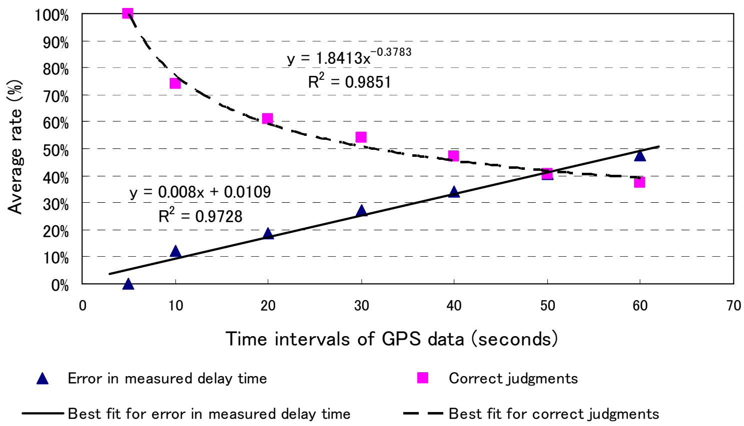


Figure 8. Relative sensitivity by frequency

Figure 8 shows the relative sensitivity for all six sets of lower-frequency probe data. The 10s probe data provides correct judgments in about 74% of cases and has an average error of 12% in delay time, while the indicators for 60s data are 37% and 47%, respectively. A pair of regression lines was estimated (and shown in Figure 8) to obtain best-fit curves for both indicators; these can be used to calculate the relative sensitivity of probe data with longer transmission intervals. The extreme time interval for probe data would be 120s under normal traffic conditions, since the average error ratio of measured delay time would reach 100% in that case; that is, data with intervals greater than 120s would have no ability to detect intersection delays. It should be noted that this extreme value is smaller for a city with shorter signal cycle times and is found to be at most equal to the cycle time, although the detailed discussion is omitted due to limitations of space.

CONCLUSIONS

Delay measurements based on the probe technique offer the advantage of providing data on all intersections in the network directly, while conventional delay estimation methods based on vehicle counts are limited to particular intersections fitted with detector systems. The financial constraints on acquiring and maintaining large-scale probe systems, however, limit the utility of the technique in most applications. Reducing the probe data transmission frequency is one effective means of saving operational costs. The sensitivity of delay measurements to reduction in probe data frequency is explored in this paper with the motivation of understanding the degree of delay measurement reliability that can be achieved with a given probe transmission frequency.

Two delay measurement algorithms, corresponding to high-frequency probe data (5s data in this paper) and lower-frequency probe data (from 10s data to 60s data) were developed in consideration of vehicle behavior at intersections. The 5s data have the advantage of judging travel conditions accurately and thus giving much better delay measurements. The lower-frequency data are simulated from 5s data by deleting parts of the records. The difference in measured delay times between these simulated data sets and the original 5s data provides an indication of their sensitivity.

Results show that with increasing transmission interval, more and more intersection delays are underestimated or missed completely. The 10s data set achieves approximately 74% correct judgments and has an average error of 12% in measured delay time, while these indicators for the 60s data set are 37% and 47%, respectively. This finding indicates that any decisions based on lower-frequency probe data should be circumspect. Future studies should aim to improve the accuracy of lower-frequency probe data applications by developing better algorithms.

ACKNOWLEDGEMENT

This study is part of the Probe-based Dynamic Route Guidance System (P-DRGS) project. The authors are grateful to the consortium members for providing valuable suggestions. The authors also wish to thank the Internet ITS Project Group for providing the probe data used in this study.

REFERENCES

Colyar, J.D., and Rouphail, N.M. (2003): Measured distributions of control delay on signalized arterials, *Transportation Research Record*, No.1852, Pp1-9.

Ghulam H.B., and Rahim F.B. (2002): Development, evaluation, and comparison of acceleration models, TRB, 81st Annual Conference, Washington, D.C.

Kühne, R.D., and Thiessenhusen, K.-U. (2004): Traffic at urban intersections: stop-time travel-time relation gained from vehicle probe data, *Proceedings of ICTTS 2004*, Pp11-17.

Miwa, T., Sakai, T., and Morikawa, T. (2004): Route identification and travel time estimation using probe-car data, *International Journal of ITS Research*, Vol.2, No. 1, Pp21-28.

Mousa, R.M. (2002): Analysis and modeling of measured delays at isolated signalized intersections, *Journal of Transportation Engineering*, Vol.128, No.4, Pp347-354.

Quiroga, C.A. and Bullock, D. (1999): Measuring control delay at signalized intersections, *Journal of Transportation Engineering*, Vol.125, No. 4, Pp271-81.