IMPROVING THE METHOD OF LINK TRAVEL TIME MEASUREMENT FROM INFREQUENT PROBE DATA*

by Kai Liu**, Toshiyuki Yamamoto*** and Taka Morikawa****

1. Introduction

Currently, the lack of effectively obtaining adequate traffic condition data is recognized as one of the most serious obstacles to the widespread application of probe vehicles. It is suggested that reducing the transmission frequency of probe data would improve the cost-effectiveness of a probe vehicle system and the saved budget can be used to employ more probes to increase coverage area and market penetration. After a comprehensive comparison between probe data in Nagoya probe experiment and taxi dispatch data[1)], we found that the challenge of using such infrequent data arises from creating a realistic reconstruction of the travel route and travel time due to sparse data (long interval between successive position information of the vehicle).

As compared with problems of reduced map matching accuracy, large errors in the measurement of link travel time (LTT) seem prominent, which result from the conventional method under the assumption that a constant velocity is maintained between two adjacent GPS record. Accordingly, the improvement of LTT measurement method is required and urgent for the application of general commercial vehicles as probe vehicles. A new LTT measurement method that attempt to locate intersection delays is developed in this paper. The focuses are concentrated on the performance of proposed method and some influence factors.

Probe data from Nagoya probe experiment are used, which were collected from October, 2002 to March, 2003 and transmitted at 5s interval. Lower frequency data transmissions are simulated by selectively deleting parts of the records. Total twelve groups of data are simulated for each trip by each vehicle over the whole collection period, with simulated transmission intervals ranging from 10s to 120s (abbreviated as 10s data, 20s data, and so on).

2. Proposed Method

Delays at signalized intersections and stoppages due to over-capacity or light congestion are showed to be the main part responsible for increasing uncertainty in LTT reports. Conventional methods of LTT measurements, with the assumption of constant speed between any two adjacent GPS records, generate significant errors inevitably for urban arterials when using infrequent position information. Because the constant speed assumption simplifies the time-speed curve to a group of horizontal lines, as a result, delay time happened on a certain DRM link is distributed to other links inevitably.

An improved method of LTT estimation is developed in this study under the consideration that re-location of signal delays to the corresponding network link might improve the measurement accuracy. The proposed methodology includes two algorithms: a traffic detection algorithm to separate the intersection delay time from the

---

* Keywords: link travel time, probe vehicle data, infrequent
** Student Member of JSCE, M.Eng., Department of Civil Engineering, Nagoya University,
(Furo-cho, Chikusa-ku, Nagoya 464-8603 JAPAN, Tel: +81-52-789-3730, Fax: +81-52-789-3565, E-mail: liukai@trans.civil.nagoya-u.ac.jp)
*** Member of JSCE, Dr. Eng., Department of Civil Engineering, Nagoya University,
(Furo-cho, Chikusa-ku, Nagoya 464-8603 JAPAN, Tel: +81-52-789-4636, Fax: +81-52-789-3565, E-mail: yamamoto@civil.nagoya-u.ac.jp)
**** Member of JSCE, Ph.D., Department of Environmental Engineering and Architecture, Nagoya University,
(Furo-cho, Chikusa-ku, Nagoya 464-8603 JAPAN, Tel: +81-52-789-3564, Fax: +81-52-789-3565, E-mail: morikawa@nagoya-u.ac.jp)

running time, and a delay location algorithm to assign delays to corresponding links. Then LTT can be improved by summing the running time and corresponding delay time.

Our previous study examined the sensitivity of infrequent probe vehicle data on the intersection delay detection[2). Delays can be measured by assuming that vehicles go through intersections at a cruise speed if without interruption by signal system. Although missing detection rates increase significantly as transmission intervals increase, the relative long-time delay still can be detected from infrequent data. The detailed delay detection algorithm can be reviewed in Liu *et al.* (2006); here a simple explanation is introduced. The algorithm begins by searching for the periods during which the vehicle is stopped (judged by a threshold speed of 3m/s), subsequently traffic condition during each transmission interval is judged by weighting the speed change rate (deceleration if less than 0.7 acceleration if more than 1.3, and running otherwise), lastly the cruising speed is estimated and used to calculate delay time. Note that because lack of moment speed reports, the widely used central difference scheme[3)-4)] was employed for this study. Although has been proved to be effective and simple to use, such scheme still smoothes the time-speed relationship and makes traffic detections difficult and leads to uncertainties.

Since the results of traffic condition detection algorithm are associated with each transmission intervals rather than DRM links, the traversed distance increases as the interval increases and might cover several links even if delay events happened during the interval. Therefore, a delay location algorithm is required to assign delay to corresponding links. Intuitively the link where the latter GPS record of one interval is located has the highest probability of including delay events, especially for a stop delay. Furthermore, consecutive delay events located at adjacent links should be re-assigned to the link where longer delay time is detected. Uncertainties arise from the difficulty of where the acceleration/deceleration delay events should be located.

If considering that an integrated delay event should comprise all three kinds of delay events, namely deceleration, stop and acceleration, it is better to locate a consecutive of delay events (including deceleration, stop and acceleration) on the link where stops are recorded. However, it is actually not always the case. Long intervals make situations very complicated. For example, one short delay after another long delay may be thought incorrectly as an acceleration event since little information between these two delay events can be obtained. In such a situation, traffic condition at the adjacent interval (forward for acceleration and backward for deceleration) is referred to make the judgment. It should be noted that such uncertainties can not be absolutely avoided by the proposed delay location algorithm. There are still some occasions on which delay events are assigned to false link and lead to a worse LTT environment. The longer the interval is, the higher the probability of making a mistake will be.

3. Some Influence Factors of the Proposed Method

There are three indicators with respect to three uncertainties that influence the accuracy of LTT measurement: the rate of missing detected delays for all links, the incorrect delay location rate for links with detected delay events, and relative errors of estimated cruise speeds for links with delay events. The rate of missing detected delays depicts the sensitivity of data on delay detector algorithm; the incorrect delay location reveals that how many delay events are located incorrectly, including incorrect detection and incorrect location; and relative speed error indicates that to what extent the detected delays can be quantified. These indicators are measured and the results in Figure 2 show that missing detects rate increase as intervals increase, while calculated relative error in cruise speed estimates (assuming that 5s data obtained the correct cruise speed and calculate the MAPE) present an increase and a subsequent fluctuation around 39%, and a constant decrease except 5s data indicates the variation of incorrect delays location. The calculated cruise speeds are found to approach to the reality as compared with any free flow speed assumption, although errors still seem large. The main cause of such large errors is the lack of instantaneous speed information and calculated speeds using central difference scheme are smoothed to a great extent. Errors for data at intervals longer than 40s tend to be steady because a minimum cruise speed of 5m/s is used in this study.

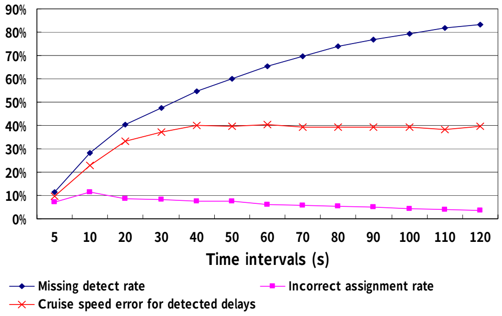

Figure 1: Three influence factors on LTT reports

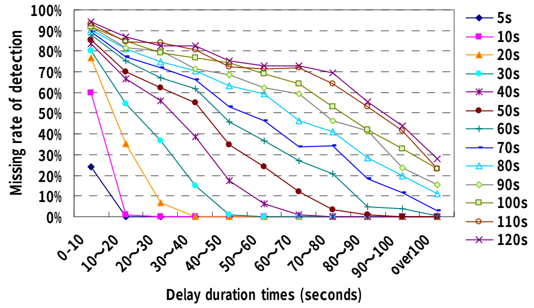

Figure 2: Missing detection rate by delay duration

Data with shorter intervals tend to include higher percentage of errors in delay locations. The incorrectly located delays are found to be always shorter than data transmission intervals. Considering the distribution of delay duration times (larger amount for shorter delays), it becomes easy to understand why 10s data include more delay location errors. In general the amount of delay location errors and accordingly reduced accuracy in LTT reports are limited and can be tolerated. Figure 2 shows the missing rate of delay detection by delay duration, which indicates the sensitivity of all thirteen levels of probe data on various delay duration times. The sensitivity of probe data intervals on delay duration decreases rapidly. If data intervals are more than 80s, the missing rates are always more than 10% for all data duration levels.

4. Performance of Proposed Method

To understand the performance of this proposed method, it is important to compare it with the conventional method. If the proposed method does not show significant improvements over the measurements from conventional method, there is no merit in adopting the method. In this paper the mean average percentage error (MAPE) is used to quantify the performance of the proposed method, with the percentage error meaning the average absolute percentage difference between the true value and the observation. Since no second-by-second data are available for our study, the LTT calculated by using 5s data is assumed to be absolutely correct. Figure 3 illustrated the MAPE difference between the conventional method and the proposed one.

Data at all intervals except 10s show a decrease in MAPE although being not as significant as expected, which means an improvement in LTT measurements is achieved. For data at an interval of less than 50 seconds such improvement seems to be inappreciable, while the improvement becomes prominent as intervals increased. For example, by using proposed method 120s data can achieve the same accuracy level as 90s data under the conventional method. Similarly, data at a 90s interval are optimized to 70s; 60s data are optimized to 50s.

The improvement by proposed method is examined by separating the trips in peak hours from those in off-peak hours. Results show that the MAPE of LTT in peak hours is less than that in off-peak hours from 30s data to 90s data, whereas the case is different for data with longer intervals. Both curves are below the curve results from conventional method. Figure 4 shows the percentage of links with LTT reports catching the same accuracy, better accuracy or worse accuracy as compared with those results by using the conventional method for all datasets. The amount of links with the same accuracy decreases rapidly until 20 seconds interval, accompanied with a slight increase thereafter.

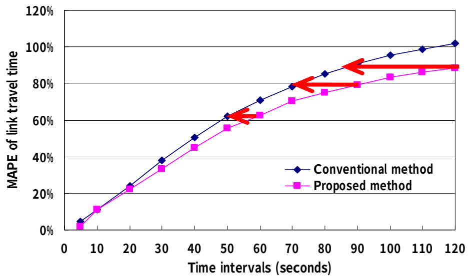

Figure 3: Comparison of MAPE of LTT

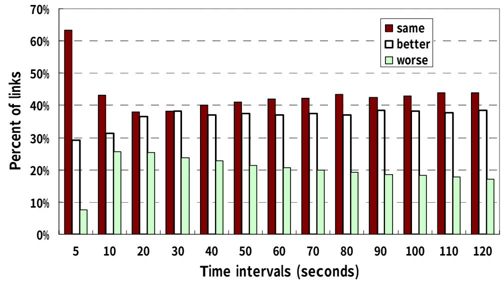

Figure 4: Percent of links by changed accuracy

Links with better reports experience a constant increase when transmission intervals are less than 30s and tend to be steady when longer than 30s, while links with worse measurements for all datasets except 5s data decrease constantly as intervals increase. This result is a little far away from our expectations because 10s data should have included few links with worse LTT reports considering that 10s data have better capacity of detecting delays than other low frequency data, and therefore hold sufficient information to assign delay correctly. One of the possible reasons that the 10s data lead to more reports with worse accuracy is that the rate of incorrect assignment of delay for 10s data is a little higher than those for other data.

5. Conclusion

Signal delays are found to be the main obstacle that affects the accuracy of LTT measurement from infrequent vehicle tracking data, and the conventional method fails to overcome such obstacle due to its inherent limitation of constant speed assumption. Better considerations of traffic condition would help to make more accurate LTT measurement.

An evaluation of the proposed methodology found that errors of LTT reports decrease as compared with the conventional method, which means an increased accuracy is achieved. However, such improvement for data at an interval of less than 50 seconds seems to be insignificant. It should be noted that only one speed criterion and a pair of speed change ratio criteria are employed in the proposed method to deal with all data at various intervals ranging from 5s to 120s, therefore it seems a little coarse. A particular criterion for one dataset at a given interval might provide significant improvements, that is to say, detecting more delay events correctly, locating delays more accurately and estimating more reality cruise speed.

References

1) Yamamoto, T., Liu, K. and Morikawa, T.: Preliminary Analysis on Taxi Dispatching Data as Potential Probe Data, Journal of Infrastructure Planning Review. (In Japanese)
2) Liu, K., Yamamoto, T. and Morikawa, T.: Estimating Delay Time at Signalized Intersections by Probe Vehicles, Proceedings of the Fifth International Conference on Traffic and Transportation Studies, August, 2006.
3) Mousa, R.M.: Analysis and modeling of measured delays at isolated signalized intersections, Journal of Transportation Engineering, Vol.128, No.4, pp. 347-354, July, 2002.
4) Quiroga, C.A. and Bullock, D.: Travel time studies with global positioning and geographic information systems: an integrated methodology. Transportation Research Part C, Vol.6, pp. 101-127, 1998.