# An Analysis of the Cost Efficiency of Probe Vehicle Data at Different Transmission Frequencies

Kai Liu[*1] Toshiyuki Yamamoto[*2] Taka Morikawa[*3]
*Department of Civil Engineering, Nagoya University*[*1]
*(Furo-cho, Chikusa-ku, Nagoya City, Japan, 464-8603, +81-52-789-3730,*
*liukai@trans.civil.nagoya-u.ac.jp)*
*Department of Civil Engineering, Nagoya University*[*2]
*(Furo-cho, Chikusa-ku, Nagoya City, Japan, 464-8603, +81-52-789-4636,*
*yamamoto@civil.nagoya-u.ac.jp)*
*Department of Environmental Engineering and Architecture, Nagoya University*[*3]
*(Furo-cho, Chikusa-ku, Nagoya City, Japan, 464-8603, +81-52-789-3564,*
*morikawa@nagoya-u.jp)*

Probe vehicles are promising tools for collecting travel time data because of their ability to provide large amounts of real-time information. However, there are many obstacles to the practical deployment of probe vehicles, one of which is the enormous expense of data transmission. This study attempts to assess the effects of differing transmission intervals (and therefore costs) on the quality of travel time data collected. Two accuracy/error values are examined: map-matching accuracy and the mean absolute percent error (MAPE) of link travel information. In order to determine the most cost-effective probe transmission frequency, the accuracy elasticity of cost is compared for six different ratios of data transmission charge to total communication cost.

***Keywords:*** *Cost efficiency, Probe vehicle data, Transmission frequencies*

## 1. Introduction

Travel time data collection is one of the fundamental components of an Intelligent Transportation System (ITS). Several data collection techniques can be used to measure travel times, among which probe vehicles equipped with GPS rank highly because of advantages that include flexibility and relatively low operating cost per unit of data [1].

One of the problems preventing large-scale data collection using probe vehicles is the tradeoff between reducing operation costs and improving the accuracy of individual reports. The cost effectiveness of travel data collection is an important design issue that should not be ignored. Most previous studies, however, have focused on methodologies for improving the accuracy of data collection and estimation without considering system costs [2] [3] [4], while others have investigated minimum values of probe vehicle density [5] and probe size [6] that meet the basic requirements of coverage area, update cycle and sample size. Most of the ITS projects investigated in these studies have selected relatively high transmission frequencies (usually second-by-second) for probe vehicles so as to easily trace travel behavior and obtain realistic real-time network information. It seems that all have neglected transmission frequency even though it is a promising variable for improving the cost effectiveness of probe vehicle data. High-frequency data obtained at high transmission cost can generate more useful traffic information, but generally speaking it is unaffordable and unnecessary. The continuing collection of real-time traffic information on a large scale depends on reducing transmission costs to a reasonable level.

Event-based transmission is regarded as an effective data collection technique as it reduces the amount of data by about 18% as compared with transmission at a fixed time interval of 30 seconds, without a significant loss in the quality of travel time and congestion information [7]. Event-based transmission schemes have been implemented in some field studies, including the Nagoya P-DRGS (Probe-vehicle-based Dynamic Route Guidance System) project. One drawback of event-based transmission, however, is that a corresponding in-vehicle device is required, which are not usually installed beforehand. Reducing data transmission frequency in a time-based transmission scheme is, therefore, an attractive alternative. Unlike event-based transmission, time-based transmission is generally in use already by taxi dispatch centers, and taxi dispatch information can be also used as probe information. The value of such data, which is costless, less frequent and simple in content (including only position and time information), depends largely on the accuracy of the measurement on real time traffic information.

Quiroga and Bullock (1998) examined minimum sampling intervals for probe vehicles (i.e. the data transmission interval) on freeways and highways from the viewpoint of sampling theory [8]. They suggest that the maximum time interval between consecutive GPS points is at most half the roadway segment travel time. Theoretical results suggesting a maximum 6 seconds data transmission interval for freeway travel negate our idea of improving cost-effectiveness by increasing transmission interval. However, the optimal transmission interval should be selected not only in consideration of the accuracy of information about traffic conditions but also on actual operation costs. Probe vehicle transmission frequency as a factor influencing operational costs has not yet attracted enough attention. This study attempts to assess the effect of transmission interval on the quality of travel time data collection. Focus is concentrated on the accuracy and cost-efficiency of probe data obtained at different frequencies.

In the section that follows, the data for this empirical study will be described. Then the relative accuracy of probe data obtained at different frequencies is calculated from two points of view: map matching and link travel time/speed. Finally the concept of elasticity is employed to assess the responsiveness of performance accuracy to transmission cost associated with transmission frequency on various hypothetical cost structures. The paper ends with our conclusions and recommendations for further study.

## 2. Data Description

Nagoya once had one of the largest probe systems in the world as a result of cooperation with 32 companies of the Nagoya Taxi Association. For the past three years, it has consisted of 1,570 taxis with GPS receivers automatically reporting their location [9]. These probe vehicles have provided a unique opportunity for the collection of travel time information and travel time prediction in a complicated urban network.

Data obtained during the second collection period, implemented from October 1st, 2002 to March 31st, 2003, are used in this study. The probe taxis can be grouped into eight types according to the pre-set transmission interval used to send real-time information back to the information center. The ten probe taxis that transmit their location at the highest frequency of 12 times per minute (5 second intervals) are selected for study in our empirical analysis. A total of 29 items of data, including time, GPS latitude/longitude, speed, direction, and distance traveled, are submitted by the GPS receiver and a gyroscope device. The GPS location error for these data is within 15 meters (95% of total data).

Lower frequency data transmissions are simulated by selectively deleting parts of the records. Twelve groups of data (one original and eleven simulated) are obtained for each trip by each vehicle over the whole collection period, with simulated transmission intervals ranging from 5 to 60 seconds (abbreviated as 5s data, 10s data, and so on in this paper).

Large numbers of taxis congregate around the downtown area of Nagoya where demand for service is high. All in-service trips that begin, end or pass through the downtown area are selected. 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 of Nagoya and old Nagoya airport. 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.

Trips less than 800 meters in length or with a travel time of less than 180 seconds (about 5% of total trips) are removed from the data because the records obtained from such short trips do not reach the minimum number required by our map-matching algorithm (at least 5 records) after simulating the lower frequency data. Meanwhile, no trip in the 5s data set with a distance between GPS reports exceeding 150 m (about 28% of all trips) is included due to the desire to record at least one GPS point per DRM link (taking into account the mean link length of 100 meters). After this data cleaning, a total number of 3,411 trips remain. As already noted, 11 groups of data for lower frequency transmissions were simulated from 5s data. The respective transmission frequencies and average travel distance between two successive data points of study samples are shown in Figure 1.

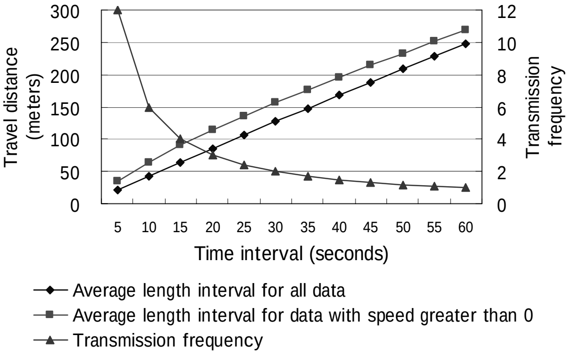


**Figure 1. Basic characteristics of probe data with different time intervals**

Although detailed, probe vehicle data with high transmission frequency include much redundant information that might lead to reduced efficiency in information processing and doubtlessly increases costs significantly. For example, about 40% records in the 5s data set consisted of readings obtained when vehicle speed was zero. As a result, many of the data points were at similar locations near intersections. Simply not transmitting such data points could lead to savings of at least 40% of communication costs. This data, therefore, is poor in terms of cost-efficiency. Figure 2 gives a comparison of cumulative distribution of travel distance for the twelve groups of probe data as described above. With decreasing transmission frequency, the percentage of data transmitted at zero vehicle speed falls from 40% to about 8%; that is to say, much more of the data is useful. However the much greater travel distances in this case prevent these data from being used to create a realistic reconstruction of the travel route and travel time.

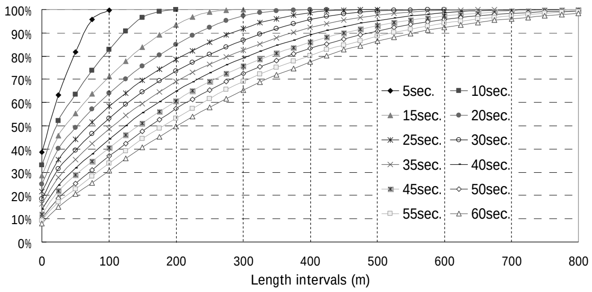


**Figure 2. Cumulative distribution of travel distance of 12 groups of probe data**

It is difficult to measure the absolute accuracy of probe vehicle data for all transmission frequencies, while it is simple and still valuable to calculate the accuracy/error ratio relative to the most accurate dataset in practice. This is, of course, the dataset with the highest frequency – the 5s data in this study.

## 3. Map-matching accuracy

Location and time information from the probe vehicles must be map-matched on DRM links for transformation into valuable traffic information. Sophisticated map-matching algorithms are required to overcome inaccuracies in the positioning system and in the digital road map. It is easy to understand that transmission frequency is also a factor that affects map-matching accuracy. In most cases, there are many routes within the DRM network that can match a set of GPS records if the time interval (or travel distance) between two adjacent records is too long (in the case of low-frequency data). Miwa *et al.* [10] developed a map-matching algorithm for the case of Nagoya city that overcomes problems of differentiating elevated urban expressways from surface roads while coping with long intervals between data transmissions using only location and time information, because other information is not always included in low-cost data such as taxi dispatch data. Miwa's algorithm was an improvement of Asakura *et al.* [11] and PROLIMAS (Probe car Link Matching System) developed by the Japan's Ministry of Land, Infrastructure and Transportation [12]. This algorithm utilizes a weighted link cost, Dijkstra algorithm, and a screening method to maintain a list of all potentially feasible routes for one trip, and then re-evaluates the likelihood of each potential route as a candidate of the correct route, which helps to ascertain the most likely location of the probe vehicle on the network. The link cost depends on the length of link, the distance from the link to the nearest GPS point, and a weighted parameter based on both vehicular velocity and road category of the link. Through a validation study, Miwa showed that his algorithm enhances accuracy effectively in Nagoya. The algorithm is used in our study on map-matching accuracy.

To what extent the Miwa algorithm can overcome the multiple route problem thrown up by lower transmission frequencies is beyond the scope of this empirical study. No map-matching algorithm can, in practice, overcome this problem completely, though the actual route can sometimes be determined by the human eyes with considerable accuracy. Map-matching algorithms that incorporate some reasonable behavioral assumption (e.g. the least turns assumption) may help to ensure correct matching, but sometimes these induce other errors. Consequently, it is inevitable that low-frequency data will results in incorrect route selection, while map-matching using higher frequency data will obviously determine routes more precisely under the same environment. Given constant external conditions, meaning that there is a constant level of deviation when using the same equipment with the same algorithm, map-matching errors with low-frequency data can be thought of as simply the result of multiple possible routes.

Almost all the records in the 5s data set (about 99.8%) have a travel distance of less than 100 m, which makes us confident of the limited links without GPS reports and affirms that map-matching of routes is possible with high accuracy. The accuracy of our map-matching algorithms on 5s data is examined by checking 100 randomly selected trips (about 3% of the total) for which correct paths were constructed manually. The results show that 99.4% of links, comprising 99.3% of total trip length, is correctly map-matched.

One and only one map-matching result can be obtained for one dataset with a given transmission frequency. Twelve routes for every trip were obtained by map-matching the 12 sets of data at different frequencies, among which the route derived from the original data (5s data) is considered the real one, with the other 11 data sets yielding an identical route or a partly different one. To evaluate the difference in accuracy of the map-matching results, we calculated the percentage of correct matches (including trips, links and link length) for all trips. A match was judged to be correct if the route/link was exactly the same as the actual one at that time, and it was judged to be incorrect otherwise.

Figure 3 may help provide a draft of route identification accuracy for different frequency data. The three curves in this figure show that all three indices rise with decreasing frequency and the slope is steep for errors in the number of links and link length when the time interval is over 40 seconds, while the slopes is relatively gentle from 5 seconds to 40 seconds. The 40-seconds point seems to be a good choice of inflection point, as it yields about 95.6% correct links; that is, this appears to be a good choice offering relatively higher accuracy and lower cost. The lowest map-matching accuracy is with the 60s data, where the relative accuracy values for trip, link and length are 50.0%, 83.8% and 83.1% respectively. That is to say, a mean value of 83.5% correct route identification can be obtained by utilizing the 60s data while the amount of transmitted information is sharply lower at about 1/12.

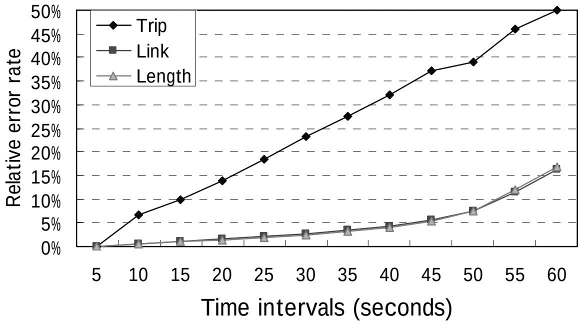
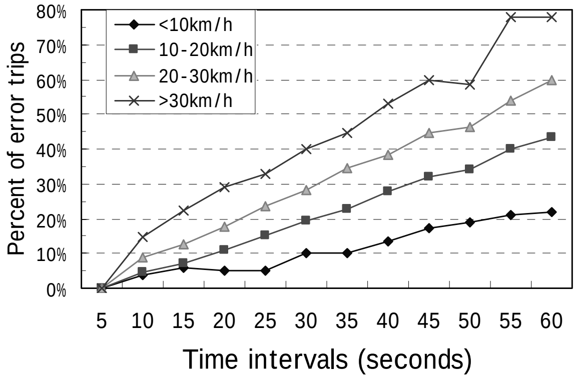


**Figure 3. Map-matching accuracy for data at different frequencies**

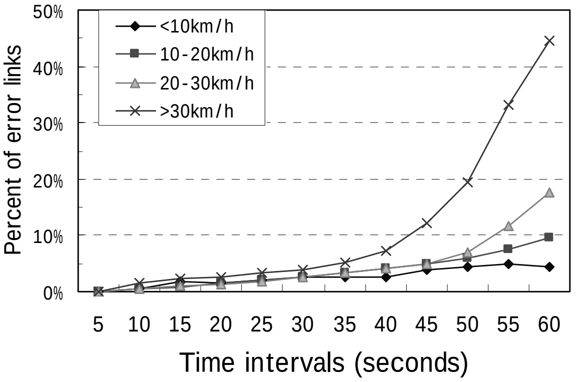
A comparison of the variability of map-matching accuracy is also carried out among sub-group trips, classified by travel speed and trip length. Firstly, the factor of travel speed is examined. All trips were divided into four categories according to average travel speed: below 10km/h, 10-20km/h, 20-30km/h and over 30km/h. The results are given in Figure 4a, 4b and 4c for accuracy of trip, link and length, respectively. The wide divergence in trip accuracy curves for the four sub-categories implies that travel speed has a considerable effect on map-matching accuracy.

Link accuracy and length accuracy of all 11 groups of simulated lower-frequency data for trips with speeds less than 10km/h show a relatively high accuracy in general. The map-matching errors become much larger for those trips with speeds over 30km/h when the time interval is greater than 40 seconds. Those trips with speeds from 10-20km/h and from 20-30km/h have larger errors in terms of the number of links and link length than those categories as below 10km/h when the time interval is greater than 50 seconds. It is noteworthy that the higher accuracy at lower speeds makes low-frequency data more valuable if the purpose of the travel time collection is to find the traffic congestions.

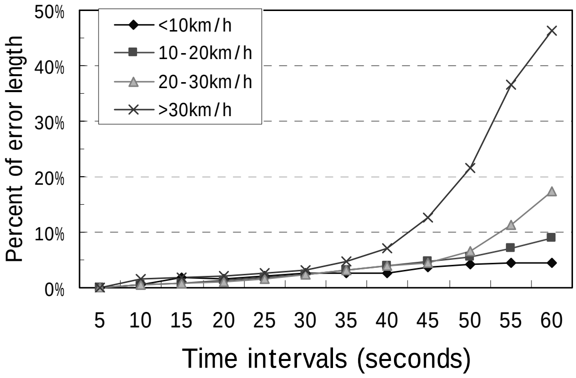
If 95% is defined as an acceptable level of accuracy (as is usually considered a significant value in statistics) the 30s data is able to satisfy all sub-categories of trips, the 45s data can meet the requirements for sub-categories of trips with travel speeds less than 30km/h, while 60s data can only satisfy the slowest trips. An interesting phenomenon is that data with a time interval shorter than 30 seconds offers similar accuracy for all classes of trips. That is to say, in Nagoya city data with a time interval below 30 seconds allow for relaxation of length interval requirement.



(a) Trip error rate



(b) Link error rate



(c) Length error rate

**Figure 4. Map-matching accuracy of sub-categories with different travel speeds**

As compared with travel speed, trip length seems to have little effect on map-matching accuracy. The length interval at a given frequency shows little difference among the four length-based sub-categories: below 2km, 2-4km, 4-7km, and over 7km, so almost the same accuracy values are obtained. For example, the percentage of correct links using 40s data are 95.6%, 96.0%, 95.9% and 94.7%, respectively, for the four sub-categories. The difference in accuracy among sub-categories widens for transmission time-intervals longer than 40 seconds. One main cause of this is that such longer trips always have a relatively higher travel speed, resulting in lower map-matching accuracy.

Actually, acceptable accuracy from the map-matching point of view depends on the purpose for which the data is being collected, so it is difficult to determine optimal transmission interval based only on relative accuracy as examined above. As GPS accuracy, DRM accuracy and map-matching algorithms all improve, restraints on data transmission frequency can be relaxed to some extant (but not completely), and data transmitted at slightly longer intervals might have acceptable potential.

## 4. Accuracy of link travel time and speed

One useful type of information that can be derived from probe vehicles is link travel time, the average value of which can, in general, be seen as an estimate of travel cost. This information is used for route guidance and travel time forecasting [2]. There are two possible methods of calculating individual link travel time from GPS point data, based on assuming uniform motion or uniform acceleration. The simpler uniform motion assumption is employed in this study, because it is difficult to judge when probe cars change speed or where they stop using only position and time information, even for the highest frequency data used in this study (5s data). Thus link travel time can be easily calculated by estimating the inflow and outflow time under the assumption that a constant velocity is maintained between two adjacent GPS records; details of the method can be found in the literature [10].

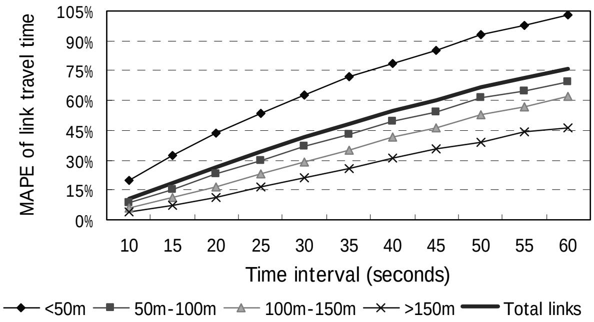
The estimation error when high-frequency probe data is used (e.g., with a time interval of 1 second) under this assumption can be neglected, while for the data with a longer transmission time interval it should not be underestimated. This methodology of evenly distributing time over distance smoothes out velocity variations during a travel sequence, especially variations due to stopping at red traffic signals. Due to the lack of one-second data, the data with 5 second intervals is assumed to be correct here; that is to say, the 5s data is assumed capable of giving the true time and true speed at any location along the whole route. The Mean Absolute Percent Error (MAPE) is used to represent these link-based attribute errors, with the percentage error meaning the average absolute percentage difference between the correct value and the observation [13]. The MAPE of link travel time and link travel speed are described by Equation (1) and (2), respectively.

$$\mathrm{MAPE}_{\mathrm{Time}} = \frac{1}{n} \times \sum_{i=1}^{n} \left| \frac{T_i - T_f}{T_f} \right| \times 100\% \tag{1}$$

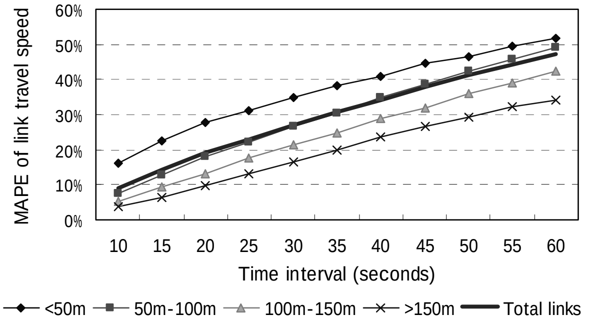
$$\mathrm{MAPE}_{\mathrm{Speed}} = \frac{1}{n} \times \sum_{i=1}^{n} \left| \frac{L/T_i - L/T_f}{L/T_f} \right| \times 100\% = \frac{1}{n} \times \sum_{i=1}^{n} \left| \frac{T_f - T_i}{T_i} \right| \times 100\% \tag{2}$$

where $T_i$ is the $i^{\mathrm{th}}$ observed link travel time, $T_f$ is the true value based on 5s data, $L$ is the link length, and $n$ is the number of sample cases.

The values are shown in Figure 5 for different link lengths. The MAPE values of link travel time and speed are much greater than the authors expected, especially for link travel time. All these MAPE values increase rapidly as the transmission frequency decreases. The MAPE value of travel time changes much more rapidly than that of speed for all links.



(a) Link travel time



(b) Link travel speed

**Figure 5. Mean Absolute Percent Error (MAPE, %) of link-based attributes by link length**

It is easy to imagine that when traffic is free flowing or traffic is extremely heavy, the value of $T_i$ is similar to that of $T_f$ for all groups of data with different frequencies, which results in the similar value of MAPE for both link travel time and speed. However, all vehicles actually have to face red signals at intersections at unpredictable times. By making the assumption that time is evenly distributed over distance between adjacent records, the observed $T_i$ for lower frequency data would be longer than $T_f$ in links without stops and shorter than $T_f$ in links with stops; generally speaking the number of links that a vehicle passes through without stopping is greater than in which it stops. Consequently $T_i$ is greater than $T_f$ in most situations, which results, ultimately, in the much more rapidly varying MAPE value of travel time.

Unlike other link attributes such as road category and number of lanes, link length is found to have a great effect on MAPE values. Figure 5 shows detailed MAPE results for four categories of link lengths: below 50 meters, 50-100 meters, 100-150 meters and over 150 meters. Shorter links seem to be more sensitive at lower transmission frequencies than longer links owing to the relative larger MAPE values and the rapidly changing rate with frequency. For a given frequency, longer links obviously have a higher probability of including more than one GPS record, and this leads to more accurate link travel time and relatively lower MAPE values.

It is very clear that link travel time estimates made using low-frequency data will inevitably lead to significant errors. Such inaccuracy of link travel time and speed estimates is in a sense beyond the acceptable level of performance measurement and efforts to identify a suitable transmission frequency from the point of view of link travel time seem to have little meaning. The widely used methodology of calculating individual link travel time requires improvement for actual application of a low-cost probe data system.

It is worth noting that in order to exclude the effect of map-matching error, only 1,195 trips that have absolutely correct map-matching results for all 12 groups of data are included in this section. The MAPE values would be worse if map-matching errors are also considered.

## 5. Cost efficiency

Elasticity analysis is employed to assess the cost efficiency of different transmission frequencies in order to find the most suitable frequency for a probe vehicle. Elasticity is an economics concept that measures the responsiveness of one variable to changes in another variable. The accuracy elasticity of cost measures the responsiveness of accuracy to changes in transmission cost. The best measure of this responsiveness is the proportional or percent change in the variables. This gives the most practical results for any type or range of data. Thus elasticity is the proportional (or percent) change in accuracy relative to the proportional change in cost. The formula for this accuracy elasticity of cost is:

$$E = \frac{\text{percent decrease in accuracy}}{\text{percent decrease in transmission cost}} \qquad (3)$$

The elasticity of accuracy in response to cost can be used to judge the cost efficiency of one data set; in other words, the data is cost-effective if its elasticity is 1, otherwise the transmission cost can be decreased in some percents with decreasing accuracy in less percents if less than 1, and the accuracy can be increased in some percents by increasing the transmission cost in less percents if greater than 1. The accuracy per cost becomes the highest when the elasticity is 1.

Generally, the transmission cost imposed by local wireless transmission companies consists of two parts: a basic charge (per month) and a charge that is directly proportional to the amount of information transmitted. The ratio of basic charge to data transmission charge can vary by country/city and company. A hypothetical cost system is established to examine different ratios, consisting of six levels of data transmission charge to total charge ratio (D/T ratio) for 5s data, from 10% to 100%. The data transmission charge is assumed to be proportional to the transmission frequency. The case of 0% D/T ratio does not require attention, because this implies no transmission charges based on the amount of data, so the highest possible transmission frequency is clearly the optimal choice, say, second-by-second. Table 1 shows details of the basic charge for 5s data at six levels of D/T ratio, taking the data transmission charge for 60s data as the unit cost (1C). The total cost for different transmission frequencies can be calculated by adding the respective data transmission charge to the basic charge.

**Table 1. Basic charge under different D/T ratios (unit: C)**

| D/T ratio for 5s data | 10% | 30% | 50% | 70% | 90% | 100% |
|---|---|---|---|---|---|---|
| Basic charge | 108C | 28C | 12C | 5.14C | 1.33C | 0C |

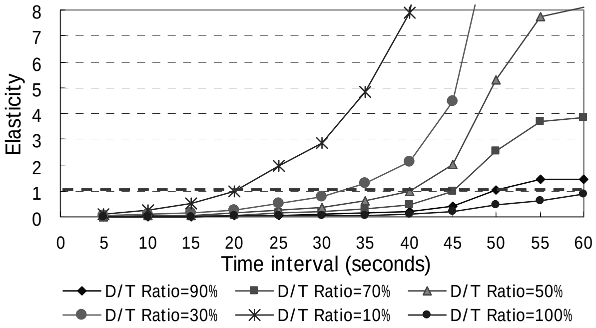


**Figure 6. Cost elasticity of map-matching accuracy**

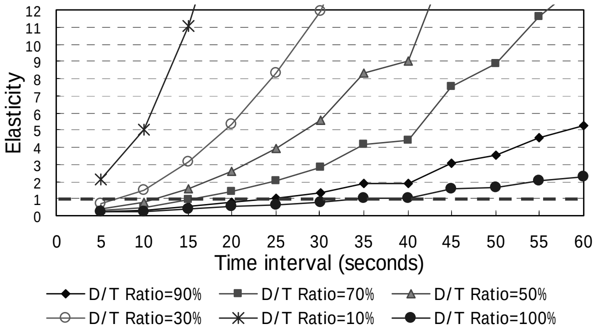


**Figure 7. Cost elasticity of link travel time accuracy**

Probe data with a 5 second interval seems to be poor in cost efficiency, although a relatively high accuracy can be guaranteed. Two indicators of accuracy/error are selected as numerators: link accuracy of map matching and the MAPE of link travel time. The accuracy of link travel time can be simplified as one minus the MAPE value. Figures 6 and 7 represent the elasticity for these two indicators, respectively, and Table 2 shows the resulting most cost-effective transmission intervals for one probe vehicle with different D/T ratios. For Nagoya city, where the D/T ratio is about 50% for 5s data, the best data in terms of cost efficiency is 40s data on a map-matching basis and 10s data on a link travel time basis.

**Table 2. Most cost-effective transmission interval at different D/T ratios**

| D/T ratio | 10% | 30% | 50% | 70% | 90% | 100% |
|---|---|---|---|---|---|---|
| On map-matching basis | 25s | 35s | 40s | 45s | 50s | 60s |
| On link travel time basis | –* | 5s | 10s | 15s | 25s | 35s |

* Most cost-effective transmission interval is less than 5 seconds.

The optimal transmission frequency for a real probe-based navigation project needs to be analyzed in future work by considering both transmission frequency and probe size. The former has an effect on the choice of most cost-effective frequency while the latter is associated with the choice of a minimum number of probe vehicles consistent with estimating the travel cost well for each vehicle. Estimation accuracy under different frequencies and different probe sizes should be further analyzed by considering these two factors jointly before making any statement on optimal transmission frequency.

## 6. Conclusion and future research

This study has focused on finding the most cost-efficient transmission frequency for a single probe vehicle by measuring accuracy and calculating the accuracy elasticity for twelve groups of data at different frequencies ranging from 5 seconds intervals to 60 seconds intervals. Accuracy was estimated from two points of view: map-matching accuracy and link travel time/speed accuracy.

Data transmitted at time intervals of 30 seconds seems to be the best choice in light of map-matching being capable of meeting the 95% accuracy requirement for all trips with different travel speeds and different lengths. However, 40s data was already able to satisfy the 95% level under average traffic conditions. Further, the accuracy of lower frequency data in terms of link travel time/speed is fairly poor compared to the map-matching accuracy, and this might influence route guidance and travel time forecasting using the data. This lower accuracy, to a great extent, results from the methodology by which time is assumed to be evenly distributed over distance between two adjacent GPS records. This implies that the method of link travel time estimation using low-frequency probe vehicle data should be improved in a future study.

The cost efficiency of probe data at different transmission frequencies is based on elasticity analysis. Results show that the most cost-effective frequency depends on the ratio of data communication charge to total cost. In Nagoya city, where the ratio of data transmission charges to total cost is about 50% for 5s data, it corresponds to 40s data in the sense of map-matching accuracy and 10s data in the sense of link travel time accuracy.

Results of this study provide a valuable reference for navigation projects that are burdened with excessive communication costs and, as a result, are forced to rely on only a few probe vehicles for the collection of real-time data. Although based on Nagoya city, the recommendations on appropriate transmission frequency for a single probe vehicle are also suitable for other cities.

However, accuracy and cost efficiency are far from sufficient basis on which to make any decision on optimal transmission frequency, since this usually depends on the purpose for which the traffic data are being collected. Any future study of the optimal transmission frequency for a navigation system should take into consideration the combination of transmission frequency with probe size.

## 7. Acknowledgments

This study is part of 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 for use in this study.

## 8. References

[1] Turner, S.M., Eisele, W.L., Benz, R.J. and Holdener, D.J.: “Travel time data collection handbook”, *Report FHWA-PL-98-035*. U.S. Department of Transportation, Federal Highway Administration, March 1998.

[2] Sen, A., Thakuriah, P.V., Zhu, X.Q. and Karr, A.: “Frequency of probe reports and variance of travel time estimates”, *Journal of Transportation Engineering*, Vol.123, No.4, pp.290-297, July/August 1997.

[3] Torday, A. and Dumont, A.-G.: “Link travel time estimation with probe vehicles in signalized networks”, *3rd Swiss Transport Research Conference*, Monte Verita/Ascona, March 19-21, 2003.

[4] Hellinga, B.R. and Fu L.: “Assessing expected accuracy of probe vehicle travel time reports”, *Journal of Transportation Engineering*, Vol.125, No.6, pp.524-530, 1999.

[5] Fushiki, T., Yokota, T., Kimita, K. and Kumagai, M.:

"Study on density of probe cars sufficient for both level of area coverage and traffic information update cycle", *11th World Congress on ITS*, CD-ROM, Nagoya, Japan, 2004.

[6] Cheu R.L., Xie C. and Lee D.-H.: "Probe vehicle population and sample size for arterial speed estimation", *Computer-aided Civil and Infrastructure Engineering*, Vol.17 (1) , pp.53-60, Jan. 2002.

[7] Horiguchi, R.: "The advantage of event-periodic data recording for probe vehicle system", *Proceedings of Infrastructure Planning*, Vol. 26, CD-ROM, 2002 (in Japanese).

[8] 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.

[9] http://www.p-drgs.com/english/index.html

[10] Miwa, T., Sakai, T. and Morikawa, T.: "Route identification and travel time estimation using probe-car data", *International Journal of ITS Research*, Vol. 2, No.1, pp.21-28, October 2004.

[11] Asakura, Y., Hato, E., Daito, T. and Tanabe, J.: "Monitoring travel behavior using PHS based location data", *Journal of Infrastructure Planning and Management,* No.653/IV-48, pp.95-104, 2002. (in Japanese)

[12] Makimura, K., Kikuchi, H., Tada, S., Nakajima, Y., Ishida, H. and Hyodo, T.: "Performance indicator measurement using car navigation systems", *81th Transportation Research Board*, CD-ROM, 2002.

[13] Toppen, A. and Wunderlich, K.: "Travel time data collection for measurement of advanced traveler information systems accuracy", *Report DTFH61-00-C-00001.* U.S. Department of Transportation, Federal Highway Administration, June 2003.

**K. Liu** Master of Eng., (Tongji University, China, 2003) Candidate of Doctor of Civil Engineering, Nagoya University. Student Member of JSCE.

**T. Yamamoto** Doctor of Eng., (Kyoto University, 2000). Associate professor of Department of Civil Engineering, Nagoya University. Member of JSCE and JSTE.

**T. Morikawa** Ph.D. (MIT, 1989). Professor of Graduate School of Environmental Studies, Nagoya University. Member of JSCE, CPIJ, JSTE, JIMS, WCTRS and EASTS.

- *Received date: 13 December 2005*
- *Received in revised form: 15 March 2006*
- *Accepted date: 19 April 2006*

- *Editor: Hiroyuki Oneyama*