COMPARISON OF TIME/SPACE POLLING SCHEMES FOR A PROBE VEHICLE SYSTEM

Kai LIU
COE Researcher, Kobe University, Japan
1-1 Rokkodai, Nada, Kobe, Hyogo, 657-8501, Japan
Tel/Fax : +81-78-803-6346
E-mail: liukai@landscape.kobe-u.ac.jp

Toshiyuki YAMAMOTO
Associate Professor, Nagoya University, Japan
1, Furo-Cho, Chikusa-Ku, Nagoya, 464-8603, Japan
Tel : +81-52-789-4636 Fax : +81-52-789-5728
E-mail: yamamoto@civil.nagoya-u.ac.jp

Taka MORIKAWA
Professor, Nagoya University, Japan
1, Furo-Cho, Chikusa-Ku, Nagoya, 464-8603, Japan
Tel : +81-52-789-3564 Fax : +81-52-789-5728
E-mail: morikawa@nagoya-u.jp

ABSTRACT

Two types of probe-vehicle polling scheme, using time-based probe intervals and space-based probe intervals, are investigated in this paper in order to provide more insight into their cost-effectiveness in real-time traffic data collection. The basic characteristics of the two schemes are examined and compared using simulated data at lower polling frequencies. Subsequently, data performance at various polling frequencies is measured and compared from the perspective of map-matching accuracy and link travel time/speed estimation accuracy. The results for the two polling schemes indicate their comparative advantages with respect to each other in the domain of traffic surveillance.

INTRODUCTION

Over the past decade, probe vehicles have been widely employed around the world for real-time traffic surveillance. Probe vehicles report their location on the road network along with other traffic information to a management center at a certain polling intervals through a wireless communications network. There are two common polling schemes in use for the transmission of this real-time information: at fixed intervals of time and at fixed intervals in space. It is suggested that choice of an appropriate polling frequency can improve the cost-effectiveness of a probe system (1). In order to determine an optimum polling frequency, an insight into the characteristics of these two polling schemes, as well as their respective predominance, is necessary.

Probe technology is considered a useful means of improving the efficiency of real-time traffic monitoring. The prospect of covering the entire network opened up when the Selective Availability (SA) of GPS under control of the U.S. military was removed in the year 2000 and thus GPS satellite location services became free for civilian usage. Subsequently, an efficient communication method became the urgent requirement for full accomplishment of probes for ubiquitous coverage of the entire road network. Travel time/speed estimates made by Quiroga and Bullock (2, 3) revealed the trade-offs between sampling rate and the reliability of section speed estimates when the data is communicated at specific time intervals. Horiguchi (4) pointed out that event-based transmission is an effective data collection scheme, as it reduces data transmission by about 18% compared with communication at a fixed time interval of 30 seconds without a significant loss in the quality of travel time and congestion information.

Although some studies have already investigated the efficiency issue with respect to time-interval data and space-interval data separately, few studies have paid close attention to the differences between these two polling schemes. Which scheme can be more easily implemented? What are the advantages of each scheme? And what about efficiency and performance in an actual urban network with each polling scheme? This paper tries to answer such questions by implementing a comprehensive comparison of the two polling schemes and, consequently, provides some insights into the essential characteristics of both. The results will help in the configuration of more cost-effective probe-based traffic surveillance systems.

Nagoya once had one of the largest worldwide probe systems thanks to the cooperation of all 32 taxi companies operating in the city. The Internet ITS consortium ran this experiment (the Nagoya probe experiment) for the last five years. Fortunately, both time-based and space-based polling schemes were implemented during data collection, so the large amount of probe data obtained gives us a unique opportunity to compare the feasibility and performance of different polling schemes.

In the next section, the data used in this study is described. Next, the basic characteristics of the probe data collected under the two polling schemes are compared from the viewpoint of feasibility and efficiency in data collection. Finally the relative accuracy of probe data obtained under the two schemes at different frequencies is evaluated using two indexes: map matching accuracy and link travel time/speed measurements accuracy. The paper ends with our conclusions and recommendations for further study.

NAGOYA PROBE EXPERIMENT

The Nagoya probe experiment consisted of 1,570 taxis fitted with GPS receivers automatically reporting their location (5). Data collection was conducted in three stages from January 2002 to June 2004. These probe vehicles 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, from October 1st, 2002 to March 31st, 2003, are used in this study.

DATA COLLECTION

The probe taxis can be grouped into three types according to data transmission type and the in-vehicle map matching capability of the vehicle. The critical difference is that the GPS data submitted by type II probe vehicles was already on-line map matched owing to an enhanced GPS receiver integrated with the in-vehicle navigation device; in this data, there is a reduced ratio of missing GPS records, commonly high in downtown areas due to the barrier effect of tall buildings. Moreover, seven patterns of pre-set polling intervals were implemented. A mobile electronic communications network was employed to transmit data, with charging by the amount of packets (DoPa technology by DOCOMO, one of Japan's mobile communications giants).

Type II probe vehicles in the Nagoya probe experiment mainly used two types of polling scheme: on a time interval basis (5s and 10s) and on a space interval basis (50m, 100m and 300m). Only data collected from taxis with passengers are used in this study because it is suggested that occupied taxis provide lower variance and are more realistic in reflecting traffic conditions than empty cars (6). Large numbers of taxis congregate around the downtown area of Nagoya where demand for taxi services is high. All in-service trips (occupied taxis) that begin, end, or pass through the downtown area are selected for use in this research because these trips are found to have a definite aim and hence more accurately reflect real-world traffic.

Although a total of 29 items of data, including time, GPS latitude/longitude, speed, direction, distance traveled (through a gyroscope device) and so on, were submitted in the Nagoya probe experiment, only time information and GPS coordinates are used in our dynamic route guidance system so as to avoid possible great uncertainties in other information, where reporting errors are more common than those in time and location information and are difficult to differentiate.

DATA SIMULATION AND CLEANING

In order to understand the difference between two types of probe polling schemes, it is necessary to explore the performance of probe data at various polling intervals. Probe data at various transmission frequencies are required for this comparison analysis. Simulated probe data can be generated at any desired interval and can be easily compared with data at the highest frequency without any external influence. Lower-frequency data transmissions are simulated by selectively deleting parts of the data record for the highest-frequency data; in this case, the highest frequency data was obtained at 5-second intervals and 50-meter intervals. Since both of these datasets are among the Type II data, a lower ratio of missing GPS data can be guaranteed.

For both polling schemes, data at lower polling frequencies are generated for each trip by every vehicle over the whole collection period. The simulated polling intervals range from 10 to 60 seconds (abbreviated as 10s data, 15s data, and so on hereafter) for time-interval data and from 100 to 600 meters (similarly abbreviated as 100m data, 150m data, and so on hereafter) for space-interval data. Trips of less than 800 meters or with a travel time of less than 180 seconds (about 5% of total trips) are discarded from the original data because the number of data points obtained from such short trips do not reach the minimum number required by our map-matching algorithm (4 records) once lower-frequency data have been simulated.

Meanwhile, trips in the 5s data set with a distance exceeding 150 meters between consecutive GPS reports (about 28% of all trips) are also discarded due to the desire to record at least one GPS point per Digital Road Map link (where the mean link length is 100 meters in Nagoya). Similarly, the greatest distance between reports for space-interval data is 75 meters, given that there is no missing GPS data in the datasets. This data cleaning process eliminates potential external influences to a great extent. After cleaning, a total of 1195 trips of time-interval data and 1427 trips of space-interval data remained for use in examining performance at various polling intervals.

BASIC CHARACTERISTICS OF PROBE DATA UNDER THE TWO POLLING SCHEMES

The polling scheme determines the spatial distribution of GPS observation points. With space-interval data, the points are evenly distributed, while with time-interval data they are sparsely distributed on some links and densely distributed on other links where there is signal delay or congestion. Intuitively, it is clear that space-interval data provide more data points when traffic is light and all vehicles travel at the free-flow speed, while time-interval data might provide more information than space-interval data in conditions of heavy traffic. This makes it difficult to judge which polling scheme is superior for practical traffic data collection. In this section, the two schemes are compared from the viewpoint of data collection. Two issues are considered: feasibility and efficiency.

MISSING DATA RATIO AND GPS ERRORS

Some GPS coordinate reports are inevitably missed in metropolitan areas due to the barrier of urban canyons. It is clear that probe data with a lower missing data ratio and with lower GPS errors would help to produce accurate speed estimates and thus offer better feasibility. In the Nagoya probe experiment, probe vehicles of Type II, which were equipped with a navigation system, were able to overcome this type of missing data problem to a great extent.

One drawback of space-interval data is that the intervals between data points are not always stable. For example, when data is supposed to be transmitted at 50 m intervals, the majority of intervals range from 45m to 75m, but at greatest the interval may be even more than 100m. This means that there are missing records. In 5s data, missed records are very rare, while the rate is greater in 50m data. Furthermore, stable transmission of space-interval data depends on the accuracy and performance of the in-vehicle distance measuring device.

This problem is even more noticeable with a commercial probe system such as a taxi dispatch system, which typically uses wireless radio technology to achieve real-time communications (through an analogue wireless network). A previous study by Yamamoto *et al.* (7) demonstrated that taxi probes may suffer from more missing data points due to communications congestion (Table 1). For time-interval data, it is rare to lose a data point because of congested channels because the communication scheme can be pre-planned on a time interval basis to overcome such problems, while space-interval data always suffers from this limitation. If many taxis attempt to transmit data simultaneously to the base station, some data might be lost.

Table 1 shows the missing data ratios of GPS coordinates by polling schemes for both probe data and taxi dispatch data. The results of taxi dispatch data are based on one month of data for 250 randomly selected taxis with both time interval scheme (50s) and space interval scheme (2km for occupied taxis and 300m for vacant taxis). The results suggest that probe data under both the time-interval polling scheme and the space-interval polling scheme, when transmitted by a packet communications technology, has acceptable missing data ratios when integrated with an in-vehicle navigation device. For probe data, although the different polling schemes do not affect the missing data ratios so much in total, great difference exists for occupied taxis.

Table 1. Ratios of Missing GPS Coordinates

	Probe data		Taxi dispatch data	
	Time interval (5s)	Space interval (50m)	Time interval (50s)	Space interval (2km or 300m)
Occupied taxis	0.47%	2.09%	16.4%	30.5%
Vacant taxis	1.67%	1.73%	20.6%	68.7%
Total	1.38%	1.85%	19.6%	67.0%

Generally speaking, the accuracy of GPS information has little relationship to the type of polling scheme. The GPS location error for both types of data is within 15 meters (with 90% confidence). It should be mentioned that there are still some occasions where the minimum distance from a GPS point to the nearest Digital Road Map link is greater than 50 meters. Besides errors in GPS location, errors in the Digital Road Map links contribute to such issues.

USEFUL RECORDS

From the viewpoint of efficiency, GPS records transmitted from a single location during a particular trip have limited value because no additional information besides the stationary state can be obtained from such records. In fact, there were many such unnecessary records in the time-interval data; that is, many records give no valuable information. On the other hand, almost all records provide information in the space-interval data. In this work, records that lead to acquisition of valuable information are regarded as useful records, taking into consideration small fluctuations in GPS coordinates that might occur even when there is no motion. A record from a location at least 5 meters away from its upstream neighbor is regarded as useful. The results presented below are based on approximately 5,000 trips of space-based data and approximately 1,500 trips of time-based data.

Figure 1 reveals the difference in trip distributions by polling frequency measured by the space density (record numbers per km) between the original probe records and the selected useful records for data at 5s intervals. Figure 2 then shows a pair of distributions of trip percents by polling frequency of useful records for the 5s data and 50m data.

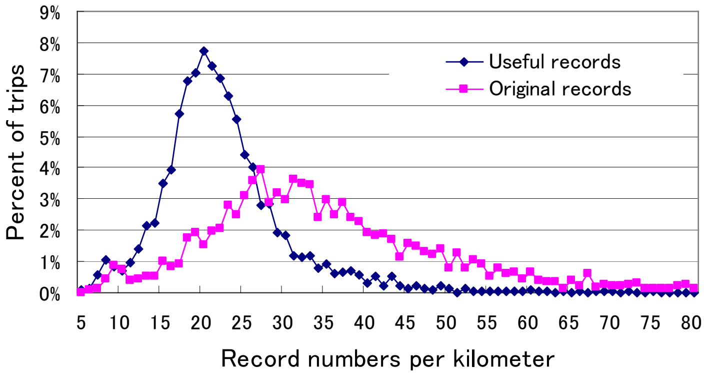


Figure 1. Frequency distribution of original 5s data and useful 5s data

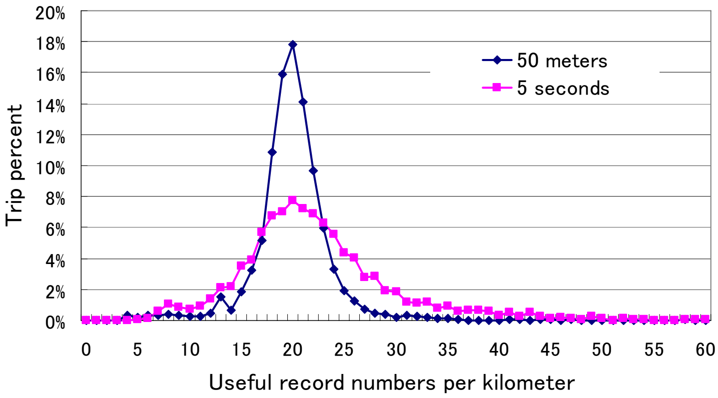


Figure 2. Frequency distribution for 50m data and useful 5s data

Under both polling schemes, trips are symmetrically distributed along the polling frequency axis, with a central peak at around 20 records per km. While the concentration of trips at this central polling frequency in the 50m data is over twice that of the 5s data, trip percentages outside this central frequency are much larger for 5s data than for 50m data. It can be postulated that data obtained under space-based polling scheme are superior to those obtained using the time-based polling scheme in terms of uniformity and stability.

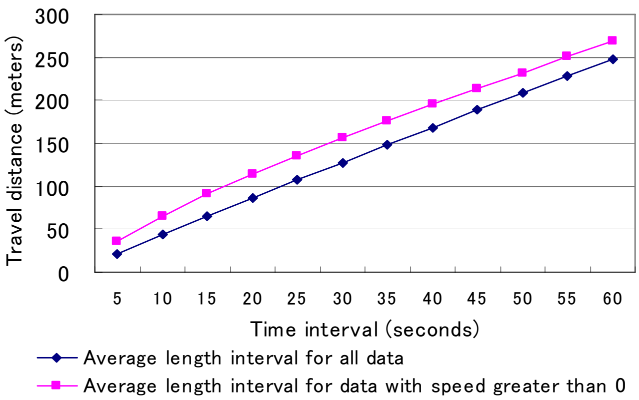


Figure 3. Average travel distance for various polling time intervals

The respective polling frequencies and average travel distance between two successive data points in the study samples are shown in Figure 3 for data obtained at various time intervals (from 5s to 60s). Both useful data and all data are plotted and the two sets of data appear as a pair of parallel lines with a gradient of about 1 in 2. The average polling interval for the 5s data is much less than 50m: 22m for all records and 36m for useful records. This relationship between the two schemes indicates a possible difference in communications costs and therefore efficiency per unit data.

Although probe vehicle data obtained at a high transmission frequency contains much detail, there is also redundant information that might lead to reduced efficiency in information processing and undoubtedly increases costs significantly. With decreasing polling frequency, the percentage of data transmitted while the probe is stationary falls from 40% to about 8%; that is to say, more of the received data becomes useful. The 5s data is poor in terms of cost-efficiency because a 40% reduction in communications costs could be achieved by simply not transmitting these unnecessary data points.

Figure 4 presents the average data density (number of records per km) for all data transmitted at various time intervals and at various space intervals. This plot shows that 30s data has the same transmission frequency as 150m data (about 6.5 records/km), with 40s data similar to 200m data (about 5 records/km) and 55s data similar to 250m data (about 4 records/km). The average record density for all data is directly proportional to the total quantity of data transmitted and therefore is a reflection of data communications costs. The average record density for useful data reflects the actual frequency of useful information.

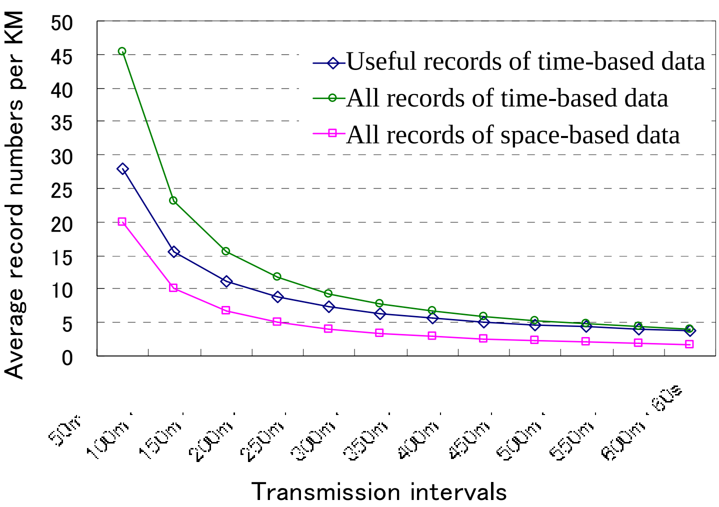


Figure 4. Average record density for both polling schemes at various intervals

PERFORMANCE OF PROBE DATA UNDER TWO POLLING SCHEMES

The results presented above suggest that variation in spatial distribution is the mechanism responsible for the different performance of the two polling methods. This section gives a general overview of the performance of probe data for various polling intervals under both polling schemes. Two aspects of performance are considered: map-matching accuracy and link travel-time measurement accuracy. It is intuitively understood that high-frequency probe data obtained using either polling scheme would easily support traffic surveillance, road management, and so on. Little difference in performance can be expected at such high frequencies. On the other hand, if the polling intervals are too long, probe data could not be used to create a realistic reconstruction of travel routes and travel times. For this reason, the comparison focuses mainly on lower-frequency probe data.

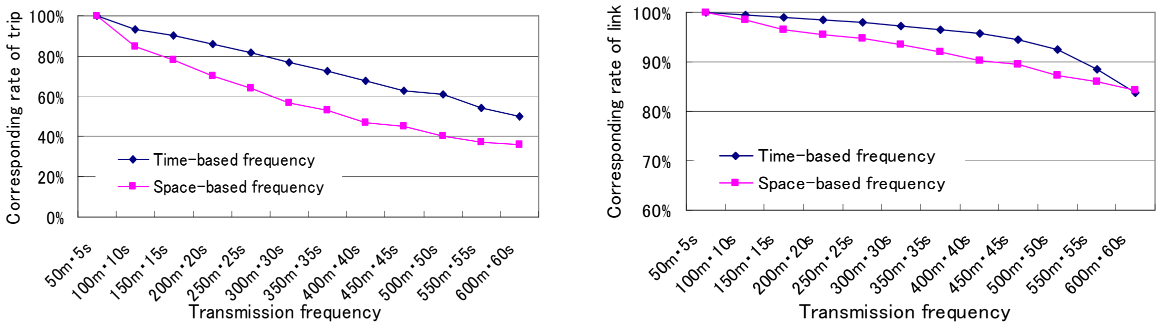
The map-matching algorithm proposed by Miwa (8) is used in this study. This algorithm was developed to overcome the problems of differentiating elevated urban expressways from roads on the surface while coping with long intervals between data points using only location and time information. The advantage of this algorithm is that the map matching result is not simply the shortest route with respect to the pre-set link costs; rather, the route provides the shortest distance as a whole to all GPS points. It should be noted that the algorithm still gives rise to errors in some situations, such as in the case of significant GPS location errors and where GPS observations are sparse and at very large intervals. In a word, polling interval, link length, link density and GPS location errors all have a combined effect on map-matching accuracy.

In practice, two kinds of link travel time/speed measurements (or observations) are widely used: GPS speed observations and measurements through GPS time/location information. Here, the latter is employed because it seems more feasible for probe data at lower polling frequencies. Since it is difficult to determine when a probe vehicle changes speed or stops using only position and time information, uniform motion is assumed between every pair of consecutive GPS reports. Thus, link travel time can be easily calculated by estimating the in and out time for a link under the assumption that a constant velocity is maintained between the two successive points.

MAP-MATCHING ACCURACY

To compare differences in probe data obtained at various time intervals and at various space intervals as regards map-matching accuracy, the first step is to calculate accuracy for probe data at various time/space intervals. The results of map matching for the 5s/50m data are assumed to be correct and the route obtained with this data is treated as the standard route. Subsequently, the accuracy of map-matching results using simulated data at longer time/space intervals can be examined by comparing the results with this standard route.

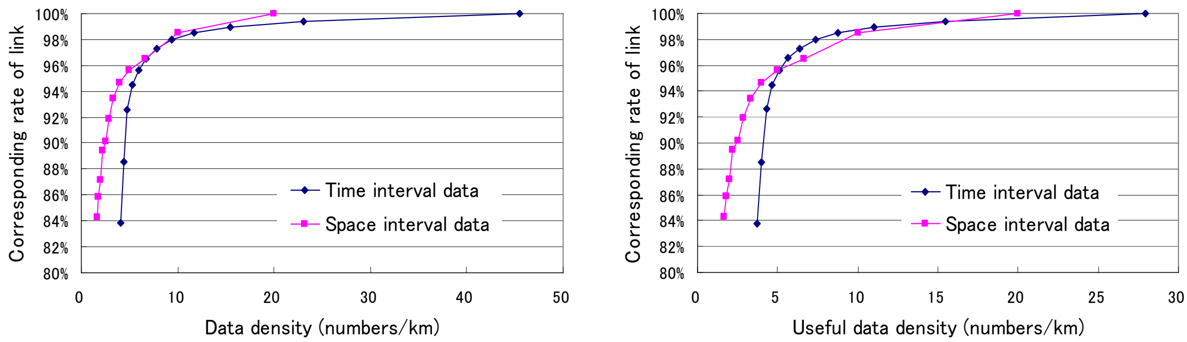
Figure 5 shows the rate of correct map matching for complete trips (Figure 5a) and the rates for traversed links (Figure 5b) for both time-interval data and space-interval data. Clearly, map-matching accuracy reduces as the polling interval increases, regardless of whether polling is on a time-interval basis or a space-interval basis. The correspondence of traversed links falls below 95% if the time intervals are greater than 40 seconds or if the space intervals are greater than 250 meters. Map-matching accuracy for data at intervals over 40 seconds falls abruptly, since errors for trips with travel speeds above 30km/h become significantly large (1). On the other hand, with space-interval data accuracy falls gradually and smoothly. As a result, the difference in accuracy between the two schemes first widens and then gradually closes as polling intervals increase. By the time the intervals have increased to 60s and 600m, both sets of data once again offer a similar accuracy of about 84%.



(a) Correct trips (b) Correct links

Figure 5. Rates of correct map matching by probe data at various frequencies

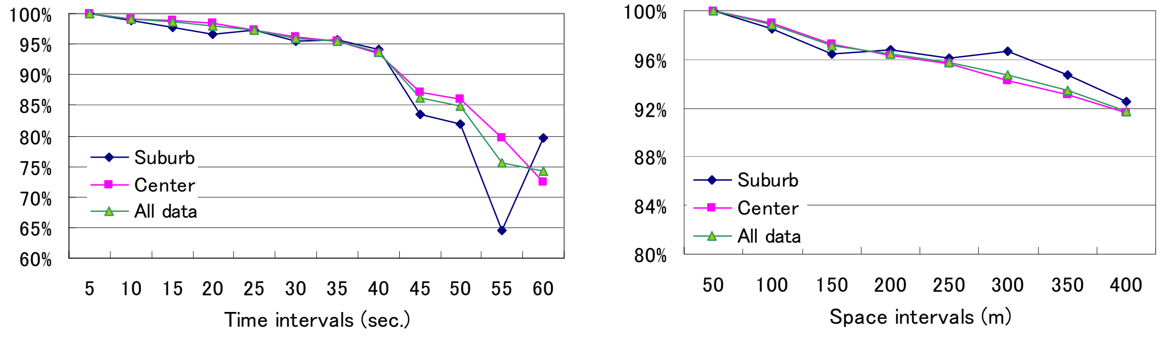
It seems interesting that the variations in map-matching accuracy with increasing intervals for these two polling schemes differ when considered in terms of trips and links. It is easy to understand that the accuracy of data obtained at intervals of space decreases more rapidly than that obtained at interval of time, because the space-interval data includes no redundant information so is sensitive to lower polling rates. Of further interest are the corresponding rates of map-matched links according to data density (number of data points per km), which are presented in Figure 6. For all data, the curve of space-interval data is always above the curve of time-interval data. That is to say, data polled at space intervals is more effective at matching GPS points onto network links than data obtained at time intervals. If the impact of non-useful data is excluded, data transmitted at intervals of less than 40s offers higher efficiency, whereas 200m data provide the same accuracy.



(a) All data (b) Useful data

Figure 6. Difference in map-matching accuracy between time-interval and space-interval data

Further studies found that map-matching accuracy using time-interval data in Nagoya city center is higher than that in the suburbs, whereas the reverse is true for space-interval data; that is, map-matching accuracy in the suburbs is greater than in city center. Figure 7 shows the proportion of correctly map-matched link length by area. The reason for this difference is that there is difference in network density and vehicle speed between the two areas. Time-interval data provide lower map-matching accuracy as vehicle speed increases, while space-interval data provide reduced accuracy when a vehicle is in an area of higher network density.



(a) Time-interval data (b) Space-interval data

Figure 7. Proportion of correctly map-matched link length by area

ACCURACY IN LINK TRAVEL TIME/SPEED MEASUREMENTS

Generally speaking, data obtained at short time intervals provides reliable measurements of link travel time/speed because most traffic conditions are easily reflected in the data. Although somewhat inferior, data obtained at short intervals of space is also sensitive to congestion and signal delays, so it provides relatively high accuracy in link travel time/speed measurements. Intuitively, it is obvious that data obtained at longer time intervals may still offer the ability to detect congestion or signal delays, while data obtained at longer intervals in space loses sensitivity.

To investigate this behavior, the average error in link travel time/speed for each correctly map-matched link is calculated on the assumption that the 5s/50m data provides the true value. Figure 8 presents these average errors in travel speed measurements for various polling intervals. Generally, time-interval data offer slightly better accuracy than space-interval data. It is clear that the average error increases rapidly at first as polling frequency increases, but the rate of increase gradually falls. The error at lower frequencies is very significant for both time-interval data and space-interval data: about 17km/h for 600m data and 14km/h for 60s data (where the average link travel speed is about 28km/h).

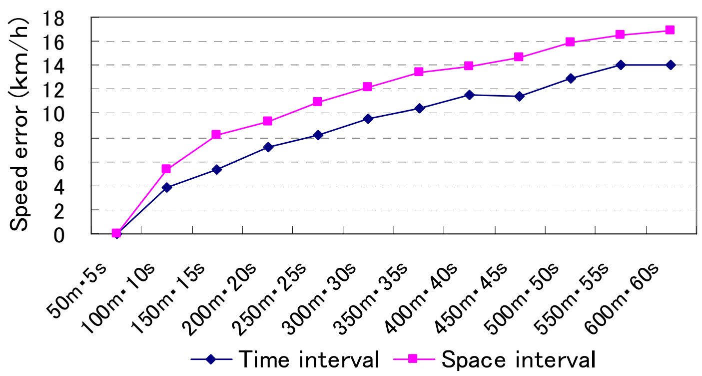


Figure 8. Average errors in link travel speed

It should be noted that the assumption that 5s/50m data reflects the absolute true link travel time/speed value is, in fact, unreasonable because of the uniform motion assumption of the conventional link travel time/speed measurement method. This means that the value of these comparative average speed errors for both schemes relies on the accuracy level acquired with 5s/50m data and the difference between the two values. In practice, it is very difficult to claim superiority of one transmission scheme in link travel time observation over the other, especially given that the data are used for varying applications. One urgent issue at this point is thus to improve the conventional link travel time/speed measurement method to more reflect the actual situation.

SUMMARY

Two basic types of polling scheme are available for a probe vehicle system: time-interval polling and space-interval polling. Despite the availability of these two methods, few previous studies have investigated the differences between the two schemes. This paper attempts to provide some insights into the respective advantages of each and their appropriate implementation environments by comparing differences both from the perspective of data collection as well as data processing. Probe data collected during the Nagoya probe experiment make such a study possible.

From the perspective of data collection, two questions are investigated: feasibility and efficiency. Generally speaking, time-interval polling offers superior feasibility due to the better completeness of the data, while the space-interval scheme is more efficient since almost all GPS reports can be treated as useful data. It should be mentioned that completeness and stability with respect to polling schemes depends on the employed communication method, which might vary according to the technology available in a particular city.

Performance from standpoint of map-matching accuracy and link travel time/speed measurement accuracy are examined for both schemes. The results indicate that space-interval data generally give better map-matching accuracy, while time-interval data are more sensitive at reflecting various traffic conditions. These advantages become weaker when polling intervals exceed a certain level, with the level depending on traffic conditions at different times of day and days of the week.

Probe data transmitted at 5s or 50m intervals seem cost inefficient, although a relatively high accuracy is guaranteed. The cost superiority of one polling scheme over the other relies on the ability to provide traffic surveillance timely and accurately at lower polling frequencies, as well as on the scope of the potential application domain. These issues require more attention in future studies.

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 related to the research. The authors also wish to thank the Internet ITS Project Group for providing the probe data used in this study.

REFERENCES

(1) Liu, K., Yamamoto, T. and Morikawa, T. (2006) An Analysis of the Cost Efficiency of Probe Vehicle Data at Different Transmission Frequencies. *International Journal of ITS Research*, Vol. 4, No.1, pp. 21-28.

(2) Quiroga, C.A. and Bullock, D. (1998) Travel time studies with global positioning and geographic information systems: an integrated methodology, *Transportation Research Part C*, Vol. 6, 101-127.

(3) Quiroga, C.A. and Bullock, D. (1999) Travel time information using GPS and dynamic segmentation techniques, *79th Transportation Research Board Annual Meeting*, Washington, D.C.

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

(5) Website of P-DRGS, http://www.p-drgs.com/english/index.html

(6) Miwa, T. (2005) Travel behavior analysis in road networks based on probe vehicle data, Doctorate Dissertation (In Japanese), Nagoya University, Japan.

(7) Yamamoto, T., Liu, K. and Morikawa, T. (2006) Preliminary Analysis of Taxi Dispatch Data as Potential Probe Data, *Journal of Infrastructure Planning Review* (in Japanese), Vol.23, No.4, pp.863-870.

(8) Miwa, T., Sakai, T. and Morikawa, T. (2004) Route identification and travel time prediction using probe-car data, *International Journal of ITS Research*, Vol. 2, No.1, pp.21-28.