EFFECTIVE PROBE DATA TRANSMITTAL WITH DETECTION OF CONGESTION PATTERN

Ryota Horiguchi
Representative Manager, ***i***-Transport Lab. Co., Ltd.
2-12-404 Ageba-cho, Shinjuku-ku, Tokyo 162-0824, Japan
Phone/Fax: +81-3-5261-3077, E-mail: horiguchi@i-transportlab.jp
Koji Wada
Research Manager, Japan Automobile Research Institute
1-1-30 Shiba-Daimon, Minato-ku, Tokyo 105-0012, Japan
Phone:03-3519-4396, Fax:03-3501-5685, E-mail: kwada@jari.or.jp

ABSTRACT

This research proposes a methodology to effectively transmit probe data in terms of congestion detection. By using "short-trip (ST)" and "short-stop (SS)", which represents a vehicle trajectory as piecewise linear shape in time-distance space, it is possible to detect congestion status of traffic with quick response. This allows us to transmit probe data en bloc within a transaction containing ST+SS sequences only recorded in congested traffic conditions, and to reduce the cost with less number of data transmittal. We have validated the proposed methodology through IPCar probe experiment in Yokohama-city.

INTRODUCTION

This research proposes a methodology to effectively transmit probe data in terms of congestion detection. From the sake of business aspects of traffic informative service, a probe vehicle will send its trajectory data to the probe information center by some electric communication in real time. Since the capacity of communication network is limited and the communication is often costly, probe data collection system is strongly required to reduce data transmittal cost as much as possible.

We have so far carried out an experiment for probe data collection in Yokohama-city (IPCar) since 2001. The expression of the probe data in IPCar experiment is based on the idea of "short-trip (ST)" and "short-stop (SS)", which divides a vehicle trajectory in time-space diagram into "running" or "stopping" sections (Figure 1). Through the experiment, ST/SS expression was validated that it can provide a good quality of data to measure section travel times as much as 1 second interval data, and can compress the volume of the data less than 30 seconds interval data. Combining the event records such as blinker switching or hand break on/off with ST/SS, we can cleanse the probe data of non "en-trip" data sections [1].

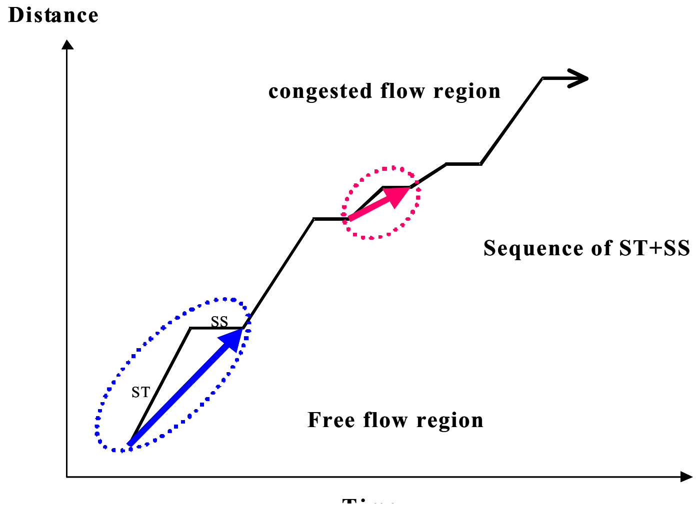

Figure 1: Piecewise linear trajectory with ST and SS.

Another utility of ST/SS expression is to help the interpretation of traffic conditions surrounding a probe vehicle. Suppose a probe vehicle running in a congestion section, the surrounding vehicles with high traffic density rob the probe of movable space, and the shockwaves propagating in the congestion may repeatedly enforce small stop-and-go upon the probe (Figure 2). Such situation can be detected by watching the sequence of ST and SS records [2].

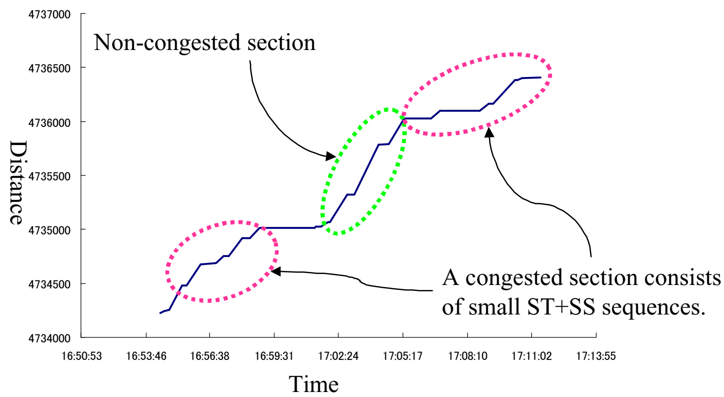

Figure 2: Appearance of ST+SS in congested sections.

When a probe detects a border of congestion and non-congestion sections, it can change the policy of data transmittal. For instance, it might be enough to send data only in congested traffic for travel time information services. Therefore, how much we can reduce the transmittal cost depends on how we can properly identify the congestion sections.

In the following chapters, the outline of the procedure to detect congestion section using pattern classification of ST+SS sequences, at first. Secondly, couple of strategies to delimit "transactions", which is to be considered as evenly congested or non-congested traffic condition, are proposed. Then, those transaction strategies are evaluated their performances in terms of the reduction of data transmittal opportunities with real world probe data. Finally, the paper concludes the best strategy for the probe data collection for travel time information service.

CONGESTION DETECTION WITH ST+SS SEQUENCES

PATTERN CLASSIFICATION OF ST+SS SEQUENCES

The basic idea of congestion detection uses pattern classification of ST+SS sequences. The reason why we do not simply use the average speed within a certain period, e.g. every 30 seconds, is that there is no guarantee the low speed means congestion nor the high speed means non-congestion. On arterial roads, a vehicle may be stopped for a minute or so at a signalized intersection. Especially for the roads in city center, the lengths of links are relatively short and it is difficult for a vehicle to run long distance even in non-congested situation. Indeed, the average speed of the recent period could be very low once the probe vehicle is stopped for a while. Similarly, a vehicle in a congested section may run at some speed when it is enforced stop-and-go by the shockwave propagation. If the periodic survey misses the 'stop', the average speed of the vehicle will be recorded at somehow higher speed.

A sequence of ST+SS provides us not only the average speed but also the additional information "how the probe stops", so that let us classify the sequences with two features, the duration of SS and the average speed of the sequence as shown in Figure 3. Figure 4 illustrates the contour of the appearance of ST+SS sequences with those two features and four clusters, Pattern 'A'~'D', are found in the figure. If we transform a trajectory expressed with ST and SS to a string using the symbols of those patterns, such as "BCCAACDD...", we can now approach the statistical character of the appearance of each pattern. Although the detail of the analysis is left out from the sake of pages, each pattern can be interpreted as follows.

'A' : is a congested pattern that repeatedly appears in a congested section.

'B' : is an ambiguous pattern that independently appears either in a congested or a non-congested section.

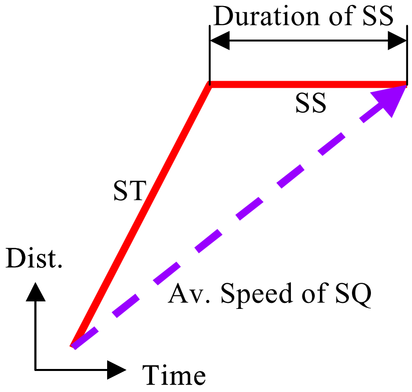

Figure 3: A sequence of ST+SS

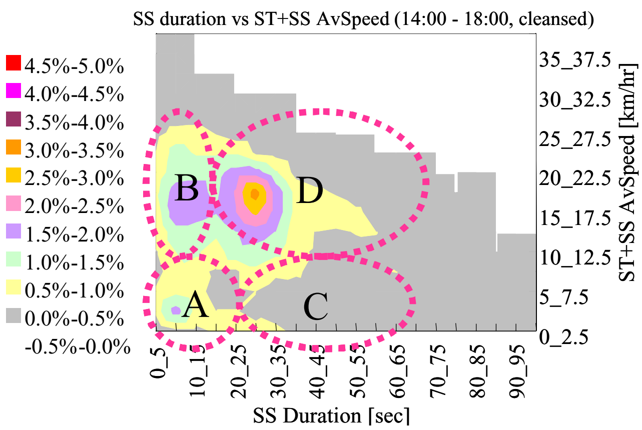

Figure 4: Pattern classification of ST+SS sequences.

'C' : is a congested pattern that independently appears in a congested section, but less frequent than 'A'.

'D' : is a non-congested pattern that repeatedly appears in a non-congested section.

STRATEGIES TO EXTRACT TRANSACTION FOR DATA TRANSMITTAL

Those characteristics of ST+SS sequences will be helpful to delimit probe trajectory data at the boundaries between congested and non-congested sections. Now, let us introduce a transaction, a unit of data transmittal, which consists of a series of continuing ST+SS sequences of which the traffic condition can be regarded as homogeneous, ie. congested or non-congested. If we are interested only in traffic congestion and do not take notice for non-congested sections, we may change the policy of data transmittal from a vehicle to a center whether the vehicle is running in congested sections or not. For example, the most excessive way is to discard the data in non-congested transactions and to send the data only in congested transactions, in order to reduce the transmittal cost.

According to the interpretation of 'A'~'D' in the previous section, the basic strategy to delimit a trajectory is to know the change of traffic status by the appearance of the pattern 'A' or 'D', because they have clear meanings. However, as there are ambiguous patterns, 'B' and 'C', the following two strategies are to be considered.

Strategy-1 : Ignoring 'B' and 'C', let the section where 'A' is continuing be a congested transaction and where 'D' is continuing be a non-congested transaction.

Strategy-2 : Ignoring 'B', let the section where 'A' and 'C' are continuing be a congested transaction and where 'D' is continuing be a non-congested transaction.

The performance of each strategy for the reduction of data transmittal cost is validated with real data collected within IPCar2003 project, as described in the next chapter.

EVALUATION OF DATA TRANSMITTAL STRATEGIES

OUTLINE OF FIELD EXPERIMENT

For a couple of recent years, Japan Automobile Research Institute (JARI) has been working for the experiments of probe data collection system via Internet on Mobile IPv6. The latest experiment in 2003 was carried out in Yokohama-city and Nagoya-city, with the purpose to advance the probe applications for road and traffic administration works [3]. Using the probe data collected in Yokohama-city (IPCar2003-Y), the performances of the proposed strategies were evaluated.

The outline of the data and the subjective field are shown in Table 1 and Figure 5. The raw data is recorded every 1 second and stored on HDD in 'Integrated Probe Equipment' and transferred to the data server in bulk at the begging of a day via Internet. In this research, the raw data is used for the validation of transactions, but specially, 'Integrated Probe Equipment' can transmit the processed data immediately when it is generated from the raw data.

Table 1: Outline of collected data (IPCar2004-Y)

Period of data collection	Sept. 2003 - Jan. 2004
Number of probes	5 buses serving for Line #13 of Yokohama-city Bus.
Number of probe trips	about 2,000 for each up and down of Line #13.
Raw data (1 Hz)	Location (LBH), speed, odometer, blinker status, etc.
Processed data	ST, SS by traffic signals, SS at bus stops, etc.

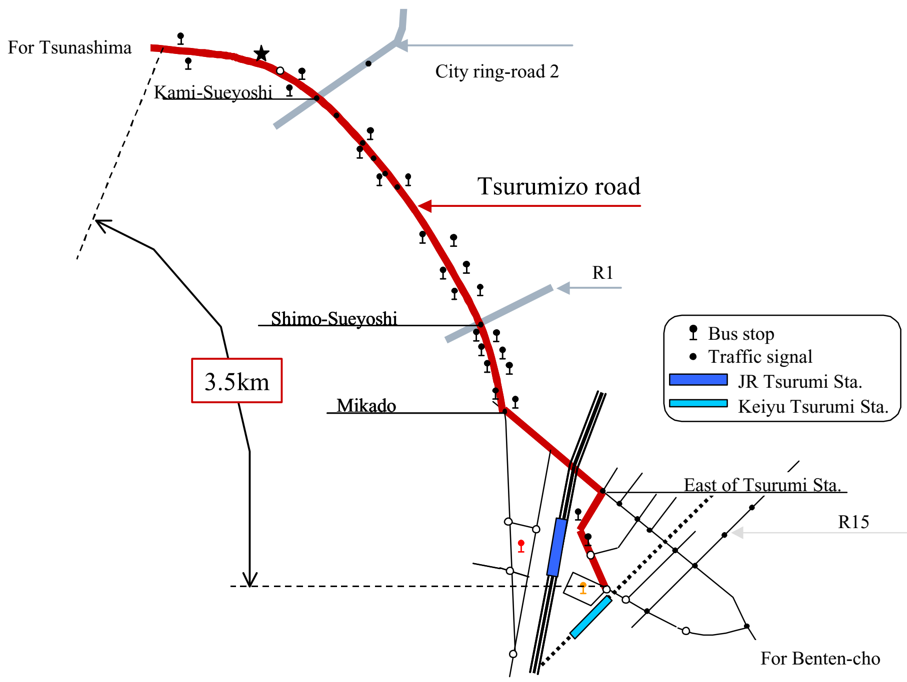

Figure 5: Subjective field of IPCar2003-Y.

COMPARISON OF TRANSACTION STRATEGIES

The performances of two transaction strategies explained in the previous chapter are evaluated regarding how they reduce the opportunities of data transmittal. Table 2 shows the number of appearances for both congested and non-congested transactions. The result says 'Strategy-1' takes the advantage to reduce the number of opportunities to transmit data.

Table 2: Number of extracted transactions with each delimitation strategy.

	# of congested tranx.	# of non-congested tranx.
Strategy-1	1596	2070
Strategy-2	3128	4083

Let us see more detail in terms of the duration and the average speed of transactions. Figure 6 and Figure 7 compares the cumulative percentage of the appearances for each strategy. For 'Strategy-1', though we find some curious "non-congested" transactions at low speed range, the durations of non-congested transactions are mostly longer, and the durations of congested transactions are mostly shorter than the ones of 'Strategy-2'. This implies 'Strategy-1' would be preferable to 'Strategy-2' for traffic informative services, because we may expect quick and dense information in congested sections in where we are interested and vice-versa in non-congested sections.

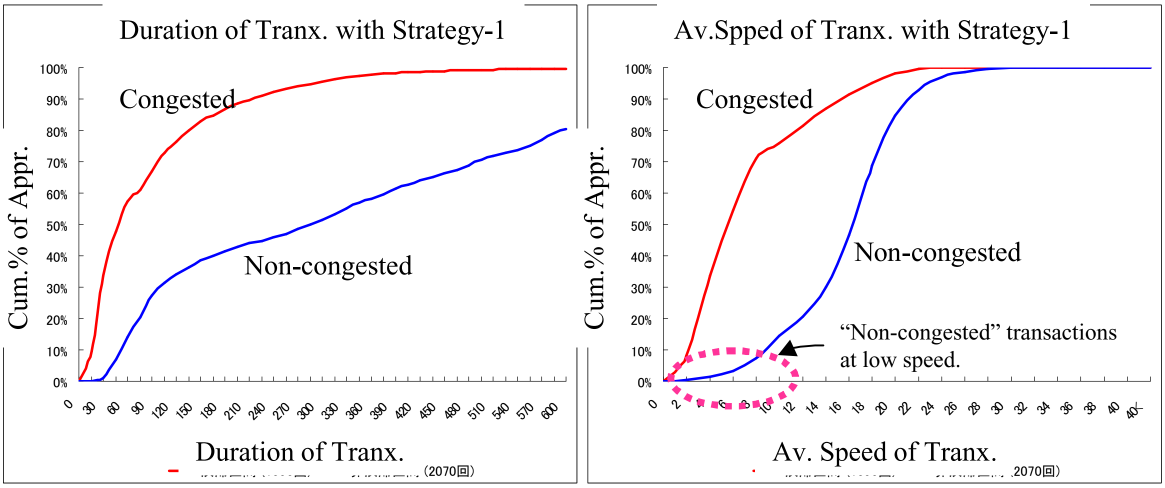

Figure 6: Distribution of the duration and the average speed of transactions with 'Strategy-1'.

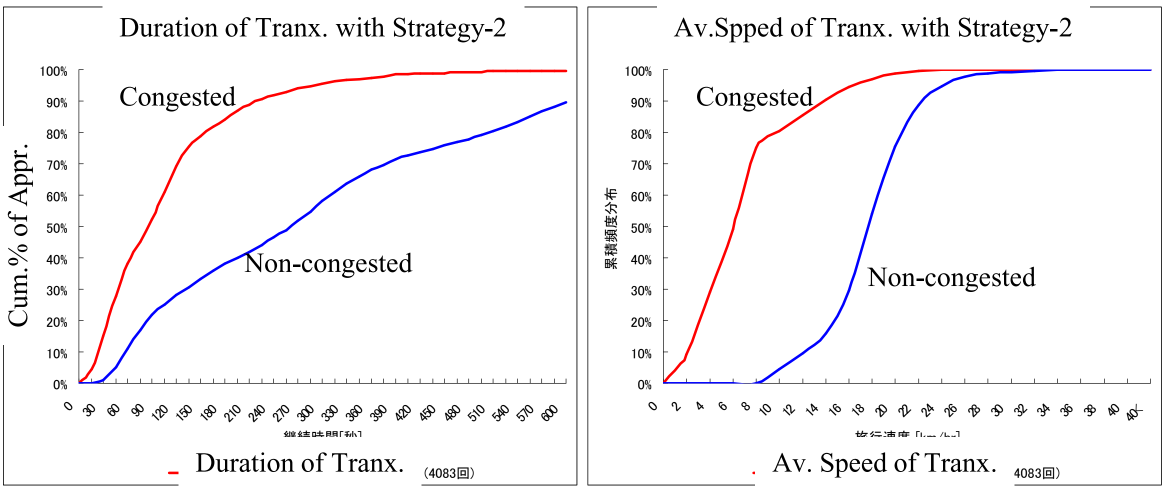

Figure 7: Distribution of the duration and the average speed of transactions with 'Strategy-2'.

UTILITY OF TRANSACTIONS TO REDUCE THE COST OF DATA TRANSMITTAL

In order to highlight the utility of transactions, let us compare the number of data transmittal and the amount of data volume with different way of data transmittal strategies. Table 3 explains two ways, "ST+SS sequence" and "transaction", of probe data approximation to raw data. The items in a record are chosen so as to contain as equivalent information to be valuable for link travel time estimation as possible.

Table 3: The way of probe data approximation

Data approximation	Items in a record	Rec. size	Update frequency in average
Raw data	*time and position.*	*16 byte*	*1 second*
ST+SS sequence	time, position of SS, duration of SS, duration of ST, and distance of ST.	28 byte	Apprx. 50 seconds
Transaction	time, head position, tail position, duration, distance, number of SS and duration of SS.	40 byte	Apprx. 80 seconds (congested) Apprx. 270 seconds (non-congested)

Here, let us consider the following data transmittal strategies.

(a) sending raw data in bulk at every 30 seconds.
(b) sending every ST+SS sequences.
(c) sending every transactions extracted by 'Strategy-1'.

Table 4 shows the result of the performance of each strategy. Let the strategy '(a)' be the base of comparison, '(b)' reduces the number of transmittal almost a half and retrench 97% of data volume. Furthermore, the strategy '(c)' saves 97% of transmittal opportunities and sends equivalent data with only 0.25% volume of raw data.

Table 4: Comparison of data transmittal strategies.

	Data approximation	Transmittal unit	# of transmittal		Amount of data	
(a)	*Raw data*	every 30 seconds	135,038	100%	61.8 MB	100%
(b)	ST+SS sequence	every sequence	77,534	57%	2.07 MB	3%
(c)	Transaction	every transaction.	3,666	2.7%	150 KB	0.25%

Regarding to the time lag to detect congestion traffic, '(c)' takes about 80 seconds in average, while '(a)' takes 30 seconds and '(b)' takes about 50 seconds. This much of the time lag of either '(b)' or '(c)' would be acceptable if we attach importance to the reduction of transmittal cost.

CONCLUSION

In this paper, the basic concept of congestion detection using ST+SS sequence was explained. Delimiting a trajectory into congested and non-congested 'transactions', we may change the policy of data transmittal and reduce the transmittal cost. Through the experiment of IPCar2003 in Yokohama-city, we have validated the performance of data reduction for a couple of transaction strategies. We have concluded that a good strategy with transactions can reduce 97% of the transmittal opportunities and 99.7% of the amount of data volume comparing to raw data collected every 1 second, even though a transaction contains almost equivalent information to the raw data. The time lag to detect congested traffic is about 80 seconds in average, and we may accept this much of delay.

For the future scope, the transaction strategy will be implemented onto 'integrated probe equipments' and validated its performance in the next IPCar2004 experiment.

REFERENCES

[1] E. Chung, M. Sarvi, Y. Murakami, R. Horituchi and M. Kuwahara: “Cleansing of Probe Car Data to Determine Trip OD”, Presented at 21st ARRB Conference in Cairns, Austraria, 2003.
[2] M. Sarvi, R. Horiguchi, M. Kuwahara, Y. Shimizu, A. Sato and Y. Sugisaki: “A Methodology to Identiry Traffic Congestion Using Intelligent Probe Vehicles”, Proceedings of 10th World Congress on Intelligent Transport Systems, Madrid, 2003.
[3] K. Wada: "Research Development and Field Testing of the Probe Car Information System (IV), will be included in Proceedings of 11th World Congress on Intelligent Transport Systems, Nagoya, 2004.