EN-ROUTE UPDATING METHODOLOGY OF TRAVEL TIME PREDICTION USING ACCUMULATED PROBE-CAR DATA

Tomio MIWA
Ph.D. Candidate, Graduate School of Environmental Studies, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
Tel: +81-52-789-3730, Fax: +81-52-789-3738
E-mail: miwa@trans.civil.nagoya-u.ac.jp
Yusuke TAWADA
Water and Sewer Commission, City of Nagoya
3-1-1 Sannomaru, Naka-ku, Nagoya 460-8508, Japan
Tel: +81-52-961-1111
E-mail: tawada@trans.civil.nagoya-u.ac.jp
Toshiyuki YAMAMOTO
Associate Professor, Department of Civil Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
Tel: +81-52-789-4636, Fax: +81-52-789-3738
E-mail: yamamoto@civil.nagoya-u.ac.jp
Takayuki MORIKAWA
Professor, Graduate School of Environmental Studies, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
Tel: +81-52-789-3564, Fax: +81-52-789-3738
E-mail: morikawa@civil.nagoya-u.ac.jp

ABSTRACT

Probe-car data include the information about the trajectory of speeds driver have experienced while passing through arbitrary selected sections. This paper discusses the methods of en-route updating travel time prediction, making effective use of probe-car data. Several methods suggested in this paper, which are en-route updating methods using both accumulated probe-car data and driver-experienced data, can predict travel time more accurately than methods that do not use both data in most cases.

INTRODUCTION

Travel time is critically important information for drivers in both pre-trip and en-route. That is, travel time influences decisions about departure time, route choice and route switching. Travel times, however, can vary significantly depending on traffic conditions, so travel time prediction is an important subject of research. Thanks to recent developments in IT (information technology), a variety of observation equipments (automated vehicle identification (AVI) systems, loop detectors, probe-car systems, etc.), are available for gathering traffic data. Currently, AVI systems and loop detectors are used to predict travel times. These equipments require considerable costs for its implementation, however, and provide data only for those sections of road in which it is installed. In the probe car system, vehicles travel through the traffic network gathering a wide range of data from the entire network. Currently, the probe car system gathers less data than conventional systems, but it is much cheaper and covers a broader area.

Most existing researches into travel time predictions, for example, using Regression models (1-3), Pattern-matching algorithms (4-7) and Kalman filtering theory (8-9), use data —real time or accumulated— that is obtained from specific sections of road. Travel times for the same sections are then predicted for the near future (a subsequent short period of time). Therefore, only drivers that are about to enter that particular section can receive predicted travel times. The length of the predictive section is expected to be longer, however, since drivers cannot accurately calculate the time needed to travel the long sections. If the section is long, traffic conditions may change while traveling through the section. Therefore, drivers would benefit from receiving updated travel time predictions while en-route. In other words, travel time information for long sections should be provided not only to drivers that are about to enter the section but also to drivers that have entered it. Previous studies used Kalman filtering theory to handle dynamically changing conditions of the section, but that method uses only the aggregate traffic data gathered from stationary equipment, which does not include the trajectories of speeds that drivers have experienced. The probe car system provides information about the trajectories of speeds that drivers have experienced while traveling through the section. In other words, probe car data include not only the time needed to travel through the entire section but also how it is composed of the times needed to travel shorter segments of the section. We think that en-route predictions of travel times can be updated using accumulated probe car data. When the travel time would be updated, the accuracy of the predictions can be improved by using the relationship between accumulated data and the trajectory of the speeds (travel times in shorter segments of the sections) that a driver has just experienced.

This paper describes methods for updating travel time predictions using accumulated probe car data and empirically compares the predictive accuracy of several updating methods. The following section discusses en-route updating of travel time predictions. The third section describes the data used in this paper, the area of study, and compares the predictive accuracy of the suggested methods. The final section discusses future directions.

METHODS OF UPDATING TRAVEL TIME PREDICTIONS

In this section, we discuss various methods of updating travel time predictions while en-route using accumulated probe car data. The term 'en-route updating' includes revisions of the initial travel time prediction while moving towards the destination. Updating is done using the relationship between the trajectory of the speeds that a driver has just experienced and the trajectories of speeds in data accumulated during past trips (see Figure 1).

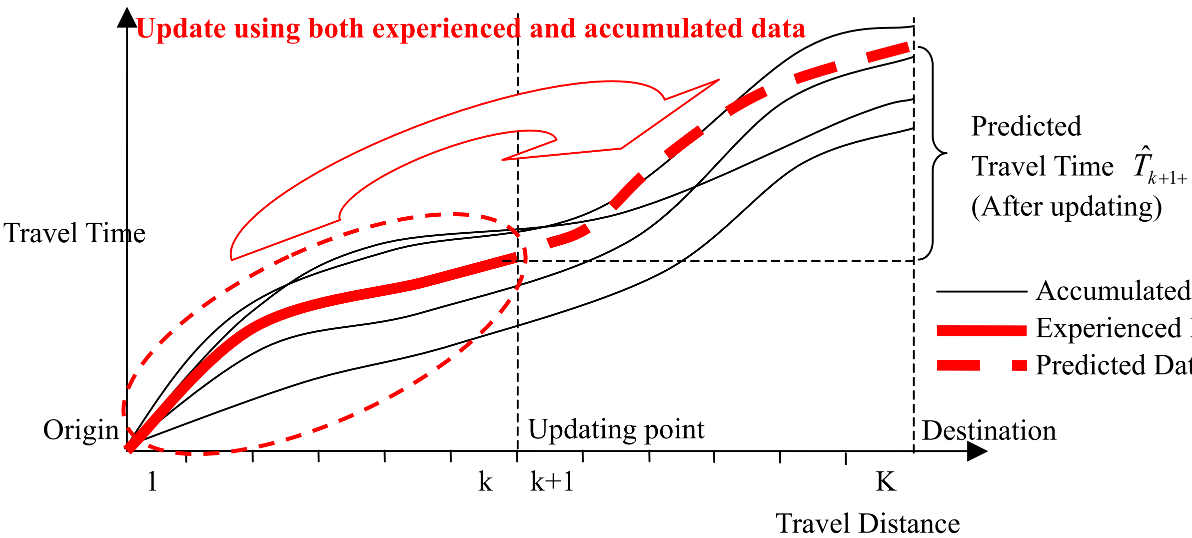


Figure 1. Concept of Updating Travel Times

Assuming there are N accumulated data for a particular section of road, the predicted travel time for a driver who is about to enter the section can be expressed by

$$\hat{T} = \frac{1}{N}\sum_{n=1}^{N}\tilde{T}^{n} \tag{1}$$

where $\hat{T}$ is the predicted travel time and $\tilde{T}^{n}$ is the travel time of nth accumulated data. If the section consists of K sub-sections of equal length, then the travel times for these sub-sections can be expressed by $\{\tilde{t}_1^{\,n} \cdots \tilde{t}_k^{\,n} \cdots \tilde{t}_K^{\,n}\}$ for nth accumulated data. The travel time for the entire section can be expressed by

$$\tilde{T}^{n} = \sum_{k=1}^{K}\tilde{t}_k^{\,n}\ . \tag{2}$$

After a driver has passed from the first sub-section to the end of the kth sub-section, the following four methods can be used to predict the travel time from the k+1th sub-section to the end of the Kth sub-section.

A. Method using mean squared error

Let $\{v_1, \cdots, v_k\}$ be speeds that a driver has experienced while passing through sub-sections $\{1, \cdots, k\}$, and $\{\tilde{v}_1^n, \cdots, \tilde{v}_k^n\}$ as speeds of the nth accumulated probe-car data. Then the difference between the experienced data and accumulated data can be written as

$$Err^n = \frac{1}{k}\sum_{i=1}^{k}\left(v_i - \tilde{v}_i^n\right)^2 . \tag{3}$$

Then predicted travel time is expressed by equation (4) using the N' ($\le N$) accumulated data with the minimum erroneous value Err^n.

$$\hat{T}_{k+1+} = \frac{1}{N'}\sum_{n=1}^{N}\tilde{T}_{k+1+}^n \delta_{N'}^n \tag{4}$$

where $\hat{T}_{k+1+}$ is the predicted travel time to pass from the k+1th sub-section to the Kth sub-section. $\delta_{N'}^n$ is 1 if nth data is in the N' minimum erroneous data and 0 in all other cases. In the empirical analysis, N' is set to 30.

B. Method using weight function

Let $f_w(x_1, x_2)$ be the weight function that calculates the similarity between x_1 and x_2. The value calculated by $f_w(\ )$ is assumed positive. The similarity in the traveling conditions in the sub-sections can be described as

$$W^n = \sum_{i=1}^{k} f_w\left(v_i, \tilde{v}_i^n\right). \tag{5}$$

The predicted travel time is then expressed by equation (6) using the N'' ($\le N$) accumulated data with the maximum similarity value W^n.

$$\hat{T}_{k+1+} = \frac{\sum_{n=1}^{N}\tilde{T}_{k+1+}^n W^n \delta_{N''}^n}{\sum_{n=1}^{N} W^n \delta_{N''}^n} \tag{6}$$

where $\delta_{N''}^n$ is 1 if nth data is in N'' maximum similarity data and 0 in all other cases. In the empirical analysis, N'' is set to 30 and the weight function given by equation (7) is used.

$$f_w\left(v_i, \tilde{v}_i^n\right) = \exp\left(-\gamma\left|v_i - \tilde{v}_i^n\right|\right) \tag{7}$$

where we set $\gamma = 1.0$.

C. Method using bivariate statistical inference

When the target section is divided into two groups of sub-sections that are expressed by $\{1, \cdots, k\}$ and $\{k+1, \cdots, K\}$, the travel times for the two groups are expressed by $\tilde{T}_{-k}^n$ and $\tilde{T}_{k+1+}^n$. When the travel times for the two groups are distributed following the normal distribution,

($\tilde{T}_{-k}^{n}$, $\tilde{T}_{k+1+}^{n}$) is distributed according to bivariate normal distribution. When a driver has passed the first group $\{1,\cdots,k\}$ in a time of T_{-k}, the predicted travel time is calculated as

$$\hat{T}_{k+1+} = \overline{T}_{k+1+} + \tilde{\rho}_{-k,k+1+} \times \frac{\tilde{\sigma}_{k+1+}}{\tilde{\sigma}_{-k}} \times \left(T_{-k} - \overline{T}_{-k}\right) \tag{8}$$

where $\overline{T}_{-k}, \overline{T}_{k+1}$ are the mean travel times of groups $\{1,\cdots,k\}$ and $\{k+1,\cdots,K\}$ as calculated from the accumulated data, $\tilde{\rho}_{-k,k+1+}$ is the correlation between $\tilde{T}_{-k}^{n}$ and $\tilde{T}_{k+1+}^{n}$, and $\tilde{\sigma}_{-k}, \tilde{\sigma}_{k+1+}$ are the standard deviations for each group.

D. Method using multidimensional statistical inference

This method considers the travel times of each sub-section as probability variables using *K*-dimensional normal distribution. Every time the probe-car passes a sub-section, the travel time prediction is updated. In this case, the joint probability density function is expressed by

$$f_{x_1,\cdots,x_K}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^K |\mathbf{V}|}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T\right\} \tag{9}$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_K)$ is the probability variable following *K*-dimensional normal distribution, $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_K)$ is the mean vector and $\mathbf{V}$ is the $K \times K$ variance-covariance matrix. In our context, the conditional mean vector $\overline{\boldsymbol{\mu}}\big|_{x_1=t_1} = (\hat{t}_2, \cdots, \hat{t}_K)$ under the condition of $x_1 = t_1$ is expressed by

$$\overline{\boldsymbol{\mu}}\big|_{x_1=t_1} = (\mu_2, \cdots, \mu_K) - (t_1 - \mu_1)(w_{12}, \cdots, w_{1K})\mathbf{U}^{-1} \tag{10}$$

where $\mathbf{U}^{-1}$ is the $(K-1)\times(K-1)$ variance-covariance matrix under the condition of $x_1 = t_1$ and *w* is the element of matrix $\mathbf{V}^{-1}$. Additional information can be found in the appendix.

Two travel-time prediction methods are also used as base cases for comparing the four suggested methods.

E. Method using only accumulated data

$$\hat{T}_{k+1+} = \frac{1}{N}\sum_{n=1}^{N} \tilde{T}_{k+1+}^{n} \tag{11}$$

F. Method using only driver-experienced data

$$\hat{T}_{k+1+} = \frac{K-k}{k}\sum_{i=1}^{k} t_i \tag{12}$$

Method E does not use driver-experienced data, while method F, which is the one used in conventional car-navigation systems, does not use accumulated data.

EMPIRICAL ANALYSIS OF UPDATING TRAVEL TIMES

DATA SET

The probe car data used in this study was collected as part of the Nagoya Probe Taxi Project (2002.1 - 3, 2002.10 – 2003.3) established by the Internet ITS Consortium (10). This project was set up to develop information technology system infrastructure. In the Nagoya pilot program, Internet ITS infrastructure was established in Nagoya City and experiments using more than 1,500 taxis were conducted in cooperation with 32 member companies of the Nagoya Taxi Association. The objectives of the project include technical verification of the Internet ITS infrastructure, the effectiveness and potential business application of the information collected from the vehicles, the effectiveness and economics of the Internet ITS infrastructure in the taxi business, and the business opportunities in information content delivery business for taxi passengers.

PROBE CAR DATA

The probe car data used in this study were collected and transmitted to the operation center whenever one of several pre-defined events occurred. Table 1 describes these events and the frequency of their occurrence. The "Distance," "Short stop" and "Short trip" events comprised about 30-35% of the total events.

Event	Definition	Percent
Distance	Vehicle travels 300 m without an event	31.3%
Short start	Vehicle starts	35.1%
Short stop	Vehicle stops	29.8%
Passenger boarding/departing Time Engine start/stop Hazardous movements	Passenger enters or exits the vehicle 550 seconds pass without an event Engine starts or stops Vehicle exceeds speed limit or accelerates or decelerates too quickly	3.8%

Table 1. Events that trigger transmission of information to the operation center

Since taxis are used as probe cars, each in-service taxi trip, that is, from the time a passenger enters the taxi to the time the passenger exits the taxi, can be treated as a trip. This study used only in-service taxi data. Since taxis are more time-sensitive while carrying passengers, this data is more appropriate for analyzing travel time predictions.

AREA OF STUDY

Figure 2 shows the area of study and the target routes for which travel times are predicted. These are the most often used routes for taxis traveling from Nagoya Airport to Nagoya

Station. One route is an expressway, and the other is an arterial road. The former is about 12.0 km and latter about 7.0 km in length. Figure 3 shows the number of trips through each route and the average weekday travel time by time of day. This figure shows that between 19:00 and 23:00, the number of trips on the expressway route is high and the average travel time is relatively constant. This is also true for the arterial road route between 18:00 and 22:00. Therefore, the data for these hours was used for the following analysis. Figure 4 shows the relationship between the travel distance and the travel time for a section of each route. This figure shows that for the trajectory of speeds for the arterial road route, there is a strong correlation between the first half of the trip and the latter half. In other words, if the first half of the trip takes considerable time, so will the latter half of the trip. The same figure also shows a moderate correlation for the expressway route.

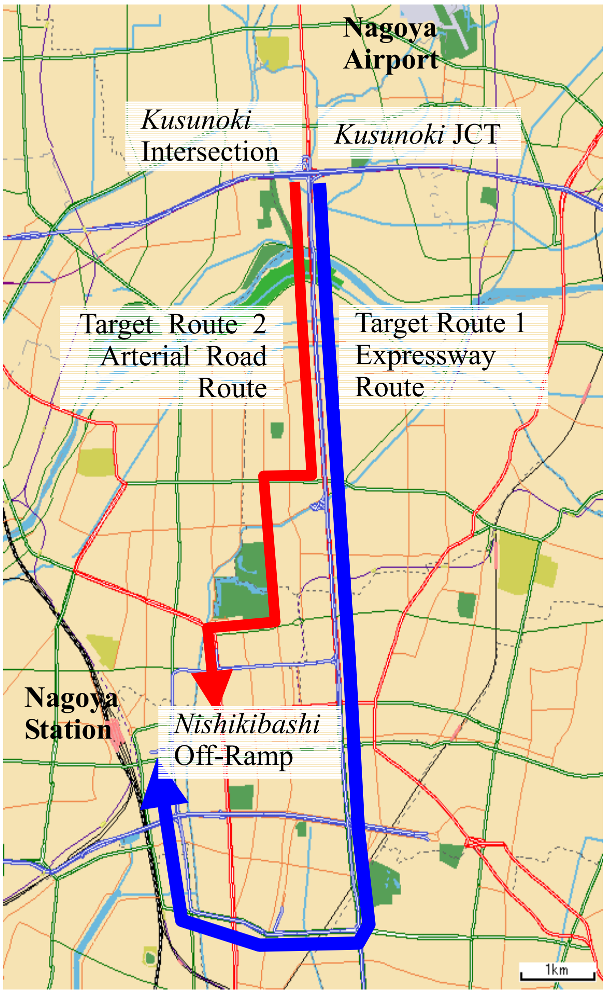


Figure 2. Study Area

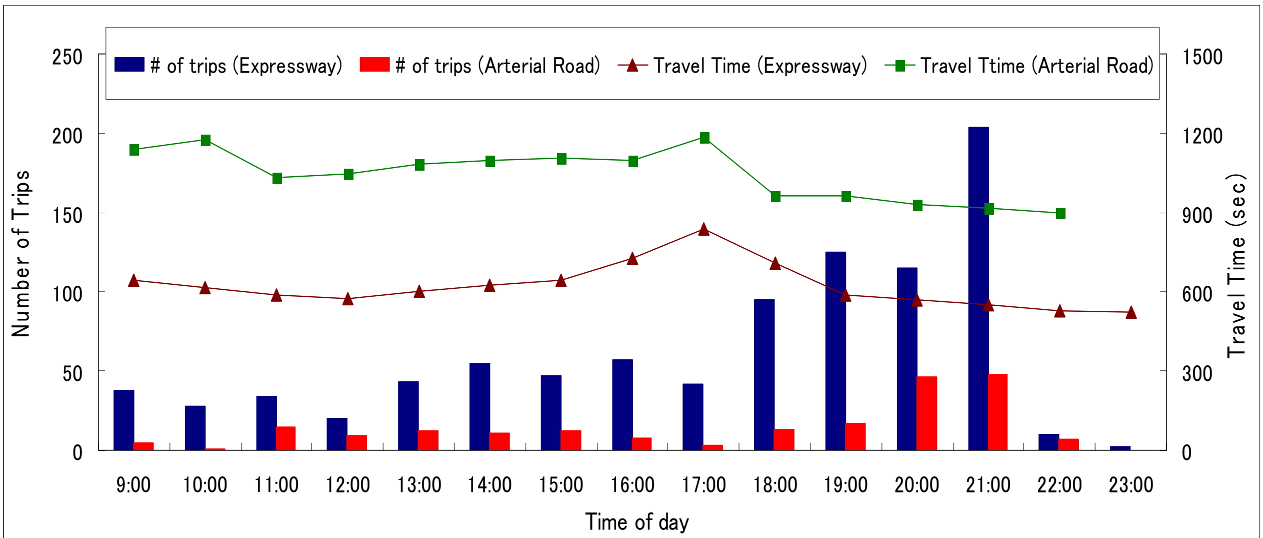
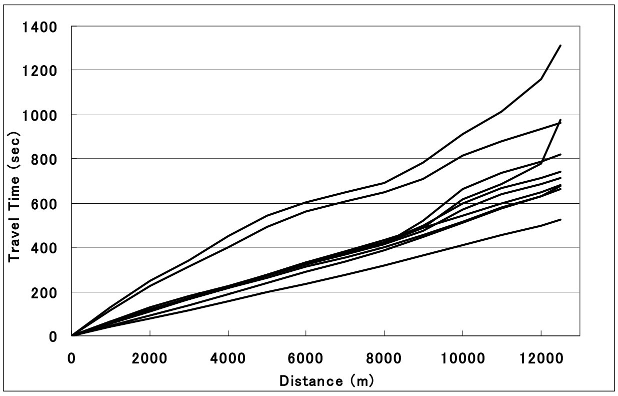
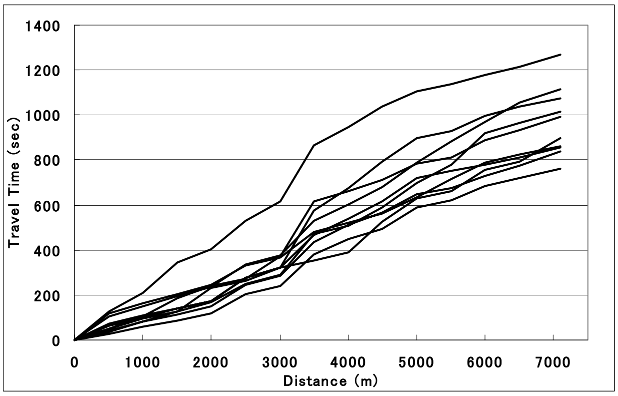


Figure 3. Number of trips and their associated travel times





Expressway Section   Arterial Road Section

Figure 4. Relationship between travel distance and travel time

COMPARISON OF PREDICTIVE ACCURACY

In this section, we compare the accuracy of the predictions provided by the six methods (methods A through F) described above. To compare the accuracy of these methods, we use a prediction error index, the mean absolute relative error (*MARE*). This is expressed by

$$MARE_k = \frac{1}{N}\sum_{n=1}^{N}\frac{\left|T_{k+1+} - \hat{T}_{k+1+}\right|}{T_{k+1+}} \tag{13}$$

where N is the number of samples, subscript "k" of $MARE_k$ is the travel distance (prediction point), $MARE_k$ is the prediction error at the end point of the kth sub-section. The expressway and arterial road section samples are 451 and 131 trips, respectively.

The expressway route has twelve sub-sections each of which is 1 km in length, while the arterial road route has thirteen sub-sections, each of which is 0.5 km in length. To simulate a travel time prediction, we first classified all cases as either driver-experienced data or accumulated data. In this study, one case was considered experienced data and all other cases were considered accumulated data; the simulation was repeated as many times as there were cases.

Table 2 and Table 3 show the values of $MARE_k$ for the expressway section and the arterial road section, respectively. For the expressway section, method A provided the most precise predictions of travel time at all prediction points, while method B provided the most precise predictions at most points for the arterial road section. In most cases, the travel times predicted by the four methods using both driver-experienced data and accumulated data (A, B C and D) were more precise than those predicted by methods E and F. In the expressway section, $MARE_1$ is significantly lower than $MARE_0$ (at the origin). Therefore, we can say that the high accuracy of the predictions derives from the appropriate use of driver-experienced data. The accuracy of the predictions for the arterial road section is lower than that for the expressway section; the shorter the remaining sub-sections, the lower the accuracy. This can be attributed to the high degree of variability in the speeds of vehicles traveling the arterial road section. The accuracy of the methods that use statistical inferences (C and D) is lower than that of the non-statistical methods (A and B). One explanation for this is the low correlation between one sub-section and the next sub-section. Another, more reasonable, explanation is that the statistical methods use accumulated data that are aggregated. Therefore, if adequate data have not been collected, extemporary conditions (i.e. traffic accidents and traffic jams) cannot be accommodated. The accumulated data used in this study consist of only a few hundred cases, so there are only a few cases with extemporary conditions. On the other hand, non-statistical methods use only N' or N'' accumulated data that are not aggregated and include appropriate data for making predictions. With data such as these, non-statistical methods provide more accurate travel time updates.

	Travel Distance k (prediction distance: 12.0km-travel distance)						
Method	Origin	1km	3km	5km	7km	9km	11km
A	0.0731	**0.0412**	**0.0347**	**0.0342**	**0.0378**	**0.0406**	**0.0530**
B	0.0731	0.0419	0.0367	0.0343	0.0391	0.0436	0.0544
C	0.0731	0.0426	0.0406	0.0397	0.0398	0.0450	0.0560
D	0.0731	0.1200	0.0616	0.0563	0.0548	0.0423	0.1012
E	0.0731	0.0729	0.0706	0.0662	0.0650	0.0696	0.0732
F	0.0731	0.1164	0.0631	0.0515	0.0548	0.0533	0.0619

Table 2. $MARE_k$ of each prediction method (Expressway Section)

	Travel Distance k (prediction distance: 7.0km-travel distance)						
Method	Origin	1km	2km	3km	4km	5km	6km
A	0.0710	0.0713	0.0776	0.0934	**0.1252**	**0.1887**	0.2192
B	0.0710	**0.0685**	**0.0769**	**0.0923**	0.1273	0.1926	**0.2187**
C	0.0710	0.0747	0.0790	0.0938	0.1270	0.1948	0.2210
D	0.0710	0.0729	0.0856	0.1008	0.1273	0.2100	0.2706
E	0.0710	0.0743	0.0787	0.0931	0.1273	0.1932	0.2193
F	0.0710	0.2192	0.2556	0.2328	0.1677	0.2463	0.3697

Table 3. *MAREk* of each prediction method (Arterial Road Section)

CONCLUSION AND TOPICS FOR FUTURE STUDY

Probe-car data provide not only travel times for specific sections but also the trajectory of speeds that a driver experienced while traveling those sections. In this study, we analyzed several methods for updating predicted travel times using both accumulated probe car data and the trajectory of speeds that a driver has experienced. The results suggest that using both accumulated and driver-experienced data provides a more precise prediction in most cases. The squared error and weight function methods are more precise than other methods. In reality, the travel times for long sections should be updated and provided to those drivers that have entered these sections more than once. Topics for future study include an analysis of the relationship between the lengths of the sub-sections and the accuracy of the predictions. The amount of accumulated data used in this study is insufficient for precisely predicting and updating travel times. Therefore, we will seek other methods that can be used to predict travel times when the accumulated data are insufficient. Finally, the time lag between the transmission of data from the probe-car and the provision of the prediction is an important issue that must be investigated in the near future.

ACKNOWLEDGEMENTS

We would like to express our thanks to the *Internet ITS Project Group* (Japanese Ministry of Economy, Trade and Industry, Keio Research Institute at SFC, Toyota Motor Corporation, DENSO Corporation, and NEC Corporation) for generously providing the data used in this study.

REFERENCES

(1) Zhang, X. and J. A. Rice (2003) Short-term travel time prediction, Transportation Research Part C, Vol.11, pp.187-210

(2) Wu, C.H., C.C. Wei, D.C. Su, M.H. Chang and J.M. Ho, (2003) Travel Time Prediction with Support Vector Regression, Proceedings of IEEE 6th International Conference on Intelligent Transportation Systems

(3) Sun, H., H.X. Liu, H. Xiao and B. Ran, (2003), Short Term Traffic Forecasting Using the Local Liner Regression Model, Transportation Research Board 82nd Annual Meeting, CD-ROM

(4) Bajwa, S., E. Chung, M. Kuwahara (2003) Sensitivity Analysis of short-term travel time prediction model's parameters, Proceedings of 10th World Congress on Intelligent Transport Systems, CD-ROM

(5) Bajwa, S., E. Chung, M. Kuwahara (2002) Travel Time Prediction on Expressways using Traffic Detectors, Proceedings of Infrastructure Planning ,Vol.26, CD-ROM

(6) Ueno, H., Y. Ohba and M. Kuwahara (2002) The Comparison of Two Type Travel Time Prediction Methods Using Toll Collection System Data, Proceedings of Symposium on ITS 2002, pp.515-520 (in Japanese)

(7) Warita, H., H. Morita, M. Kuwahara and A. Tanaka (2003) A Study of a Statistical Traffic

Condition Forecast Method in Consideration of the Latest Traffic Conditions Corresponded to the Advancement of Road Traffic Information, Proceedings of Symposium on ITS 2003, pp.215-220 (in Japanese)

(8) Chen, M. and S. Chien (2001) Dynamic Freeway Travel Time Prediction Using Probe Vehicle Data: Link-based vs. Path-based, Transportation Research Board 80th Annual Meeting, CD-ROM

(9) Yasui, K., K. Ikenoue and Y. Fukai (1993) Prediction of Journey Time on Streets, Infrastructure Planning Review, No.11, pp.137-144 (in Japanese)

(10) Keio University Internet ITS Research Group (2002) Internet ITS Research Report, http://www.internetits.org/ja/report.html

APPENDIX

Supposing $\mathbf{x} = (x_1, x_2, \cdots, x_K)$ is the probability variable following K-dimensional normal distribution, then $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_K)$ is the mean vector of $\mathbf{x}$, and $\mathbf{V}$ is the $K \times K$ variance-covariance matrix that is expressed by

$$\mathbf{V} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1K} \\ v_{21} & v_{22} & \cdots & v_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ v_{K1} & v_{K2} & \cdots & v_{KK} \end{pmatrix} \tag{a.1}$$

$$v_{ij} = E(x_i - \mu_i)(x_j - \mu_j) \tag{a.2}$$

If the number of samples of x_i is N, then μ_i and v_{ij} are expressed by

$$\mu_i = \frac{1}{N}\sum x_i \tag{a.3}$$

$$v_{ij} = \frac{1}{N-1}\sum (x_i - \mu_i)(x_j - \mu_j) \tag{a.4}$$

Then the joint probability density function is expressed by

$$f_{x_1,\cdots,x_K}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^K \det \mathbf{V}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})^T \right) \tag{a.5}$$

where subscript "T" means transposed matrix. The distributed form of the probability density function expressed by Equation (a.5) depends on the quadratic form expressed by following equation.

$$(\mathbf{x} - \boldsymbol{\mu})\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})^T \tag{a.6}$$

For simplicity, we set $\mathbf{W} = \mathbf{V}^{-1}$. Under the condition of $x_1 = t_1$, equation (a.6) becomes equation (a.7).

$$\sum_{i \neq 1, j \neq 1} (x_i - \mu_i) w_{ij} (x_j - \mu_j) + 2\sum_{j \neq 1} (t_1 - \mu_1) w_{1j} (x_j - \mu_j) + (t_1 - \mu_1) w_{11} (t_1 - \mu_1) \tag{a.7}$$

The third term is negligible because it is constant.

Here, matrix $\mathbf{U}$ is separated from $\mathbf{W}$ and expressed by the following equation.

$$\mathbf{U} = \begin{pmatrix} w_{22} & \cdots & w_{2K} \\ \vdots & \ddots & \vdots \\ w_{K2} & \cdots & w_{KK} \end{pmatrix} \tag{a.8}$$

The other variables $\mathbf{x}'$, $\boldsymbol{\mu}'$ and $\overline{\mathbf{b}}$ are

$$\mathbf{x}' = (x_2, \cdots, x_K) \tag{a.9}$$

$$\boldsymbol{\mu}' = (\mu_2, \cdots, \mu_K) \tag{a.10}$$

and

$$\overline{\mathbf{b}} = (t_1 - \mu_1)(w_{12}, w_{13}, \cdots, w_{1K}) \tag{a.11}$$

Using the above equation (a.8)-(a.11), we can transform equation (a.7) into (a.12).

$$(\mathbf{x}' - \boldsymbol{\mu}')\mathbf{U}(\mathbf{x}' - \boldsymbol{\mu}')^T + 2\overline{\mathbf{b}}(\mathbf{x}' - \boldsymbol{\mu}')^T \tag{a.12}$$

Completing the square form of equation (a.11) produces the following equation.

$$(\mathbf{x}' - \boldsymbol{\mu}' + \overline{\mathbf{b}}\mathbf{U}^{-1})\mathbf{U}(\mathbf{x}' - \boldsymbol{\mu}' + \overline{\mathbf{b}}\mathbf{U}^{-1})^T - \overline{\mathbf{b}}\mathbf{U}^{-1}\mathbf{U}(\overline{\mathbf{b}}\mathbf{U}^{-1})^T \tag{a.13}$$

Therefore conditional expectation $\overline{\boldsymbol{\mu}}$ can be obtained from the following equation.

$$\overline{\boldsymbol{\mu}} = \boldsymbol{\mu}' - \overline{\mathbf{b}}\mathbf{U}^{-1} = (\mu_2, \cdots, \mu_K) - (t_1 - \mu_1)(w_{12}, \cdots, w_{1K})\mathbf{U}^{-1} \tag{a.14}$$

and conditional variance $\overline{\mathbf{V}}$ is obtained from the following equation.

$$\overline{\mathbf{V}} = \mathbf{U}^{-1} \tag{a.15}$$