Implementation of Short-term Travel Time Prediction Model on Urban Expressway

Chong Wei[*1] Takamasa Iryo[*1] Yasuo Asakura[*1]

Graduate School of Engineering, Kobe University [*1]
(1-1, Rokkodai-cho, Nada-ku, Kobe-shi, Kobe 657-8501, Japan, TEL: +81-78-803-6360,
E-mail: 058t781t@stu.kobe-u.ac.jp, iryo@kobe-u.ac.jp, asakura@kobe-u.ac.jp)

Prediction of travel times is a vital function of many advanced traveler information systems. In this study, an experimental analysis of the relationship between instantaneous travel time and real travel time is carried out and a method to predict freeway travel times using a linear model is proposed. The proposed method is applied to a data set of the Hanshin Expressway. The implementation process of the travel time prediction model is described in detail in this article.

Keywords: *travel time prediction, implementation, linear model.*

1. Introduction

Many expressway corporations have been providing travel time information on some major road segments by using VMS. These travel time information are essential “instantaneous travel time” (ITT) which is a simple accumulation of the individual link travel times to the present time. Link travel time can be calculated as the length of the link divided by the velocity of the link, which is measured by a detector. However, in actual situations, the link travel time may change according to traffic conditions. Therefore, the instantaneous travel time is not equal to the “real travel time” (RTT). In this article, we propose a method for predicting the real travel time and focus on techniques to implement the method.

Many studies have been conducted on travel time prediction. Methodologies and tools used for such prediction in previous studies include the artificial neural networks [1], Kalman filter [2], regression model [3][4], pattern matching [5], and microscopic simulation [6]. Data collection methods including loop detectors, automatic vehicle identification systems, supersonic detectors, and probe vehicle systems have been used in previous researches. The artificial neural networks are usually employed when there is a nonlinear relationship between the input and the output, whereas the Kalman filter methods are usually adopted when there is a linear relationship between the input and the output. In both artificial neural networks and regression models, part of available data is required for modeling the input and the output parameters; however, there are some differences between these two tools. An artificial neural network can be termed a “black box” because the relationship between the input data and the output data is not explicit. On the contrary, a regression model is devised according to some predefined relationships between the input and the output data. The predefined relationships are generally based on experiences, theoretical analysis, or intuitions. The matching method directly searches for a pattern in the database that matches the input data. On the other hand, the regression method first analyzes the trend of the data (so as to obtain the regression line) and then calculates the output according to the trend.

Problems of these methods can be described as follows. In the artificial neural network model, a long time is required for training the parameters, and trained parameters cannot be transferred to another location. Generally, the structure of the artificial neural network model also cannot be conveniently transferred to another location. Kalman filter requires a primary error matrix as an input; however, a general method to obtain an error matrix has not yet been devised. Thus far, a regression model has been designed on the basis of intuition. Theoretical and logical analyses of the linear model for predicting the travel time are insufficient.

In this study, a varying coefficient linear model was proposed for predicting the RTT. This model provides travel time information that is more precise and reliable than instantaneous travel time. The proposed model can be easily understood and run by expressway managers. This model can also be easily transferred to other locations without any complications.

2. Relationship between RTT and ITT

The relationship between the RTT and the ITT is generally determined by the congestion process. For example, in Figure 1, the ITT is derived from a vehicle track depicted by a dotted line (Figure 1 (b)). This figure is plotted assuming that the traffic condition after 8:00 is the same as that at 8:00. On the other hand, the RTT is derived from a vehicle track under the real traffic condition, which is depicted by a solid line (Figure 1 (a)).

Obviously, the traffic congestion causes a difference between the RTT and the ITT. From Figure 1, we can observe that the RTT is greater than the ITT because of the congestion caused at 8:00.

Figure 2 (a) shows the relationship between the RTT and the ITT in the case of decreasing congestion. In this case, the RTT is less than the ITT. Figure 2 (b) shows the relationship between the RTT and the ITT when the traffic is free flowing. In this case, the RTT is same as ITT.

However, the estimated value of the RTT (see section 4.1) and the measured value of the ITT always include some error. Hence, data obtained from traffic flow detectors may fail to show the relationships mentioned above. Practically, we expect that the ratio of the RTT to the ITT under increasing congestion should be greater than that under decreasing congestion.

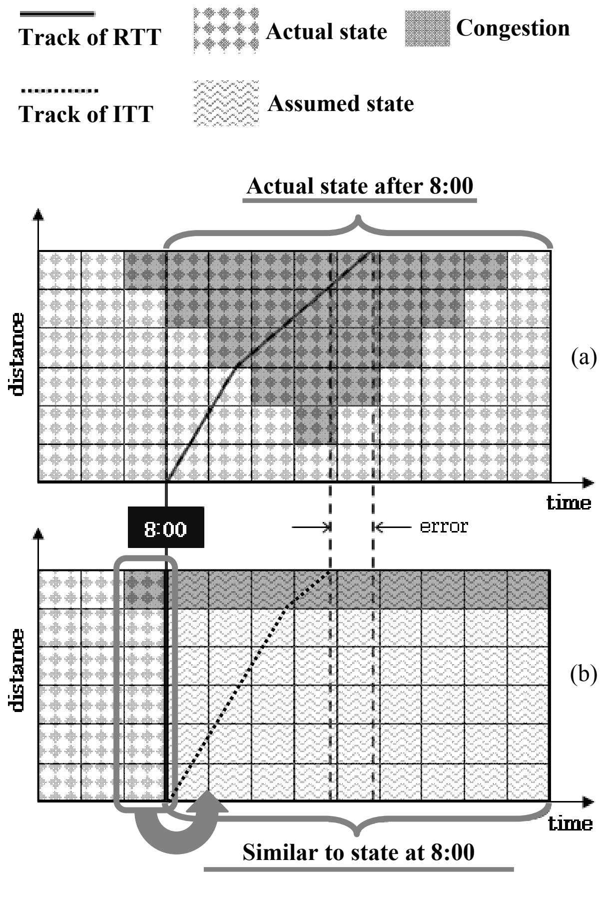


Figure 1. RTT vs ITT (developing congestion)

The relationship between the RTT and the ITT was examined by using real-world data of the Hanshin Expressway. A Visual Basic application (VBA) program was developed for illustrating the relationship between the RTT and the ITT. Figure 3 displays the congestion processes (start to end) from 10:00 to 11:30 on a particular day (45 degree lines are represented by the dotted line). From this figure, we can deduce that the ratio of the RTT to the ITT is influenced by the increase or decrease in congestion (The increase or decrease in congestion can be identified by the value of the ITT).

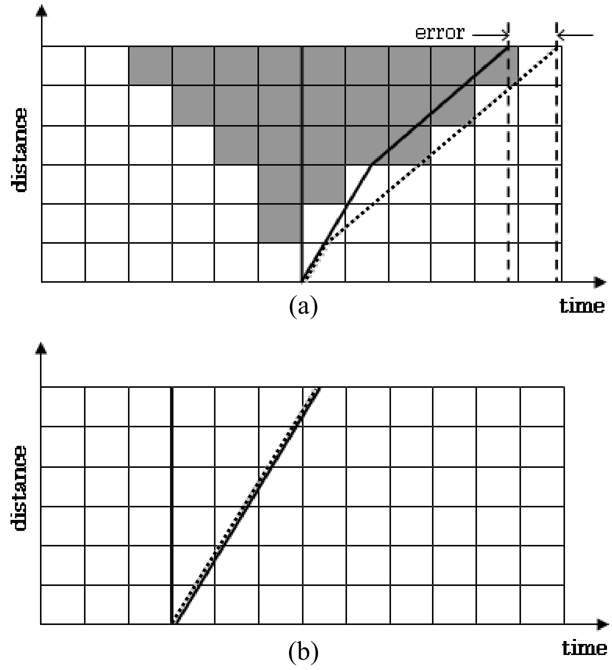


Figure 2. RTT vs. ITT ((a) decreasing congestion and (b) free-flowing traffic)

3. Proposed Model

As explained in section 2, we can use a varying coefficient linear model to describe the relationship between the RTT and the ITT. First, the RTT of an expressway segment can be transformed into the ITT as follows (see Figure 4):

$$RTT(t) = ITT(t + gap\_t) \tag{1}$$

where t denotes the instant at which the vehicle enters the expressway segment and $t + gap\_t$ denotes the instant at which the vehicle encounters congestion. Here, we assume that the bottleneck capacity of the expressway segment is stable (or say, changes slowly) and that there exists only one bottleneck across the segment. The difference between $ITT(t + gap\_t)$ and $ITT(t)$ arises due to two reasons. First, there exists a time interval (gap_t) between $ITT(t + gap\_t)$ and $ITT(t)$; hence, they may be not equal to each other.

Further, if the traffic condition does not vary with time, $ITT(t + gap\_t)$ will be equal to $ITT(t)$; hence, the second reason is the change in the traffic states at instant t and $t + gap\_t$ (Figure 4). A natural consideration is to describe the change by using the expanded/dismissed speed of the congestion; however, it is difficult to detect the speed by using the existing detectors. In this study, we use the term $C(t)$ to describe

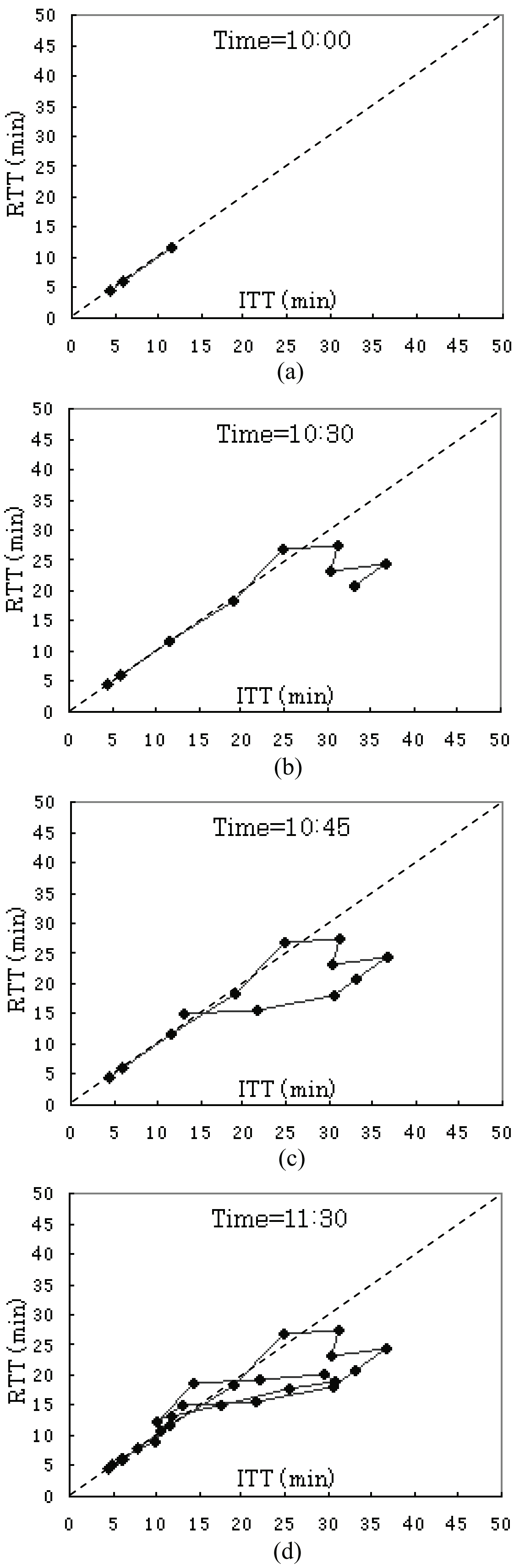


Figure 3. Relationship between RTT and ITT

the change in the traffic state:

$$C\left(t\right) = ITT_t \,/\, ITT_{t-5\min} \tag{2}$$

$C\left(t\right)$ is a measurement that describes the change in the ITT at t. That is, $C > 1$ when congestion is increasing and $C < 1$ when it is decreasing. Consequently, according to the above analysis, the relationship between $ITT(t + gap\_t)$ and $ITT(t)$ can be described as:

$$ITT(t + gap\_t) = f\left\{C(t), gap\_t\right\} \cdot ITT(t) \tag{3}$$

where $f(.)$ is an unknown function. Using equations (1) and (3), we have:

$$RTT\left(t\right) = f\left\{C(t), gap\_t\right\} \cdot ITT(t) \tag{4}$$

From Figure 4, gap_t can be expressed as:

$$gap\_t = x \,/\left\{V_{free} + \overline{V}_{extent}\left[C(t)\right]\right\} \tag{5}$$

where x is the distance between the starting point of the expressway segment and the tail of the congestion corresponding to t. V_{free} is the velocity of the vehicle at free flow and can be considered as a constant. $\overline{V}_{extent}$ is the mean propagation speed of the congestion between t and $t + gap\_t$. Here, we assume that $\overline{V}_{extent}$ is dependent on $C(t)$. gap_t also can be considered as a function of x and $C(t)$, i.e.,

$$gap\_t = gap\_t\left\{x, C\left(t\right)\right\} \tag{6}$$

Function $f\left\{C(t), gap\_t\right\}$ can be rewritten in a more compact form as follows:

$$f\left\{C(t), gap\_t\right\} = F\left(C, x\right) \tag{7}$$

where function $F(.)$ has a similar structure as function $f(.)$ but the independent variables are different. Using the above analysis, we can deduce the relationship between the RTT and the ITT as follows:

$$RTT(C,x) = F(C,x) \cdot ITT(C,x) \tag{8}$$

Zhang and Rice (2003) proposed a varying coefficient linear model for predicting travel time; they assumed that the parameters of linear regression vary with the time of the day. However, according to our analysis, the regression line should vary with the value of C and x. We consider the following linear model based on the above analysis:

$$RTT(C,x) = \beta_0(C,x) + \beta_1(C,x) \cdot ITT(C,x) + \varepsilon \tag{9}$$

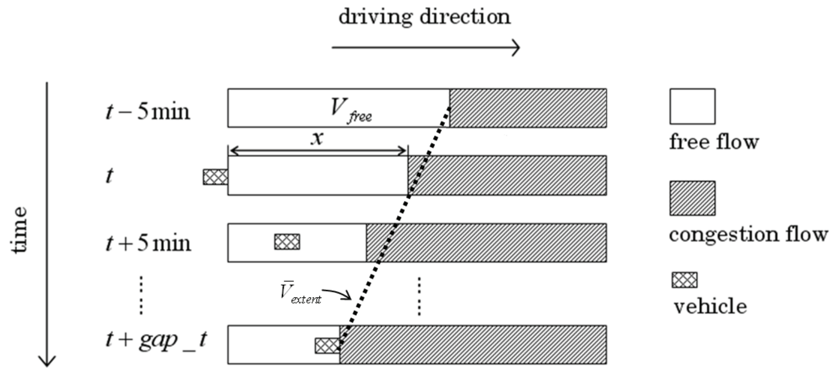


Figure 4. Relationship between ITT and RTT

where $\beta_0(C,x)$ and $\beta_1(C,x)$ are the parameters of the linear model that will be estimated from historical data of a study section. These parameters can vary with the values of C and x. This kind of linear model can be termed a varying coefficient linear model. The weighted least squares (WLS) method is used to estimate the parameters β_0 and β_1. From the present condition at time *t*, if we can assume that the propagation speed of a queue does not change until $t+gap\_t$ (i.e. the mean speed $\bar{V}_{extent}$ between t and $t+gap\_t$ can be exactly determined at time t), RTT can be calculated by using it and speeds of free flow and congested regions. However, this assumption may not be applicable when the propagation speed of a queue changes rapidly (i.e. $d^2x/dt^2 \neq 0$). For example, when the queue length gets longest and going to decrease, d^2x/dt^2 should be less than zero significantly, meaning that the assumption of the constant propagation speed will cause overestimation of RTT. In contrast, the proposed methodology estimates change of future travel time by using historical data, which can incorporate such changes of dx/dt into the estimation. Further, it must be noted that measuring the propagation speed (and free flow/congestion speed sometimes) would be costly or inaccurate. The proposed model incorporates errors included in the measurements as the error term ε in equation (9) and can handle them by WLS method.

Figure 5 shows graphs of RTT versus ITT; the data used to plot these graphs was derived from the data pool of the Hanshin Expressway, under certain conditions. These conditions and the number of samples are mentioned at the top left corner of Figure 5. Generally, if C is greater than 1, $\beta_1(C,x)$ should be greater than 1 and $\beta_0(C,x)$ should be approximately equal to 0. However, in this study, because the traffic state is captured by the detectors, the values of the RTT and the ITT are estimated values rather than true values. This measurement error may lead to biases (e.g. in Figure 5, $\beta_1(C,x) < 1$ although $C > 1$ and $\beta_0(C,x) \cong 0$). On the other hand, although some observed errors and biases exist, the proposed model can successfully

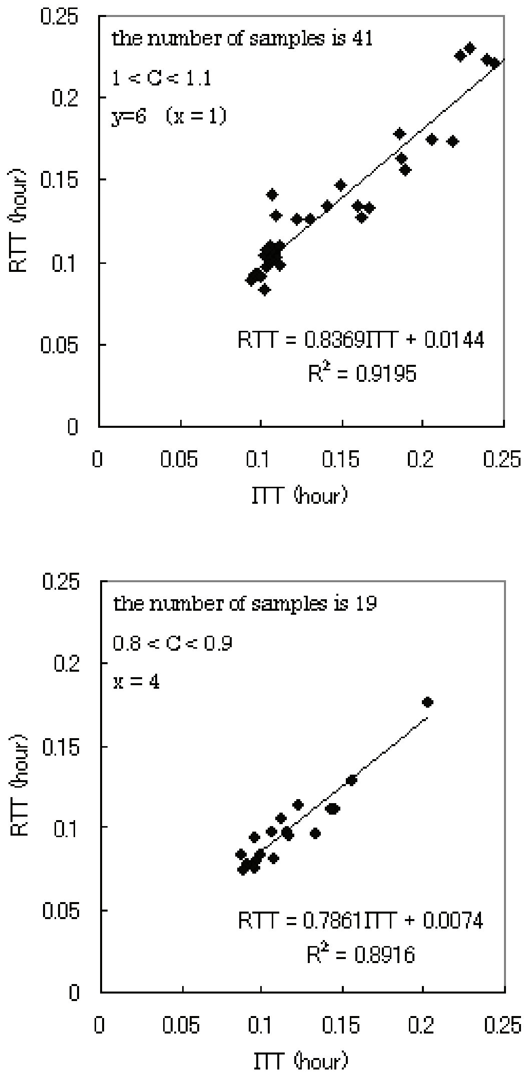


Figure 5. Linear relationship

describe the changing trend of $\beta_1(C,x)$. Generally, the value of $\beta_1(C,x)$ for $C>1$ should be greater than that for $C<1$. We can observe this trend in Figure 5. Congestion may occur at the starting point of the expressway segment at time t. In this case, y is used instead of x, where y is the distance between the starting point of the expressway segment and the head of congestion corresponding to t. Here, x is the ID (ID = 1, 2, …, n) of the first detector counting from the starting point of the expressway segment, and the detected speed is lower than 50 km/h. y is the ID of the last detector counting from the starting point of the expressway segment, and the detected speed is lower than 50km/h.

4. Data processing

4.1 RTT

The value of the RTT cannot be directly obtained by using supersonic detectors or loop detectors and must be estimated. Figure 6 shows an example of estimation of RTT. At the iteration step i, time t_{i+1} and distance d_{i+1} of the next iteration $i+1$ are calculated according to the values of t_i, d_i, and the speed of the cell (the values of both t_0 and d_0 are set to 0). The iteration is terminated when the value of d_i becomes equal to the length of the expressway segment, and the value of the associated t_i is the estimated value of the RTT.

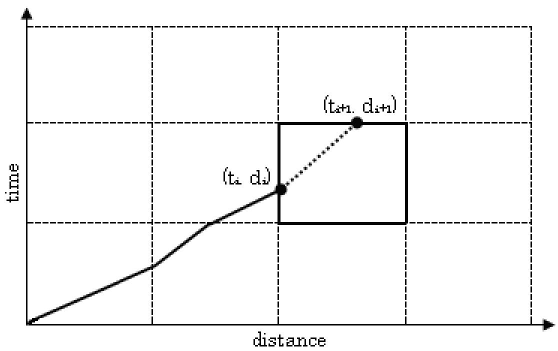


Figure 6. Estimation of RTT

4.2 Values of C, x and y

The values of C, x and y were calculated according to the procedure described in section 3. The calculated values were saved in a data table of the relational database. The variables in the data table include ITT, RTT, C, x and y. The variables corresponding to an instant t are on the same line in the data table.

4.3 Prediction of RTT

The prediction procedure was implemented using a C language program and a relational database. C, x and y correspond to the prediction time that is to be input into the program. x and y are the IDs of the first and last detectors, respectively, when the detected speed is lower than 50 km/h. The prediction procedure is shown in Algorithm 1. The total calculation procedure requires approximately 2 s. The calculation time increases with an increase in the size of the database, but a 5-min interval is permitted between each prediction, so that the proposed model is capable of online prediction of the RTT.

Algorithm 1 shows that only two parameters β_0 and β_1 are necessary to be estimated, and the parameters are independent of the expressway location; hence, the proposed model can be transferred to another location. However, the parameters should be re-estimated if the road or traffic conditions are drastically different from those of the road segment that was originally studied.

Algorithm 1. Prediction procedure

Database = { RTT, ITT, C, x, y }
Input = $\tilde{C}, \tilde{x}, \tilde{y}, I\tilde{T}T$; (values of C, x, y, ITT at the current instant)
Output = $R\tilde{T}T$; (predicted value of RTT at the current instant)

1. If $x = 1$, select RTT and ITT from database where $\tilde{C} - \Delta C < C < \tilde{C} + \Delta C$ and $y = \tilde{y}$;
else, select RTT and ITT from database where $\tilde{C} - \Delta C < C < \tilde{C} + \Delta C$ and $x = \tilde{x}$;
Let { RTT, ITT } denote the selected sample. (In the present study, we set $\Delta C = 0.1$. We have introduced x, y in section 3).
2. Estimate β_0 and β_1 in equation (9) using the selected samples { RTT, ITT }. Let $\hat{\beta}_0$ and $\hat{\beta}_1$ denote the estimation result.
3. Calculate $R\tilde{T}T$ (the predicted value of RTT at the current instant) as
$$R\tilde{T}T = \hat{\beta}_0 + \hat{\beta}_1 \cdot I\tilde{T}T$$

5. Case study

Archived operational data obtained for a period of 3 years (from 03-02-2003 to 05-31-2006) from the Hanshin Expressway were considered in the present study. The expressway segment between the Kashima and Fukushima ramps, which has a length of 6 km, was selected for this study. The data were collected from 9:00 to 22:00 every day, throughout the week. We considered the pool data during the entire three-year period with the exclusion of data from 8 days, which could not be obtained due to technical problems with the detectors. All data were saved in a relational database. The data structure used in this study has already been described in section 4.2.

Supersonic detectors were used to collect traffic condition information from the Hanshin Expressway.

Figure 7 shows the process of detecting vehicle speed. A vehicle is detected by supersonic detectors 1 and 2 at the t2 and t4 instants, respectively. The detected speed can be calculated as Δd / Δt where Δd is the distance between supersonic detectors 1 and 2, and Δt is the time interval between t2 and t4 (Δt = t4 – t2).

We calculated ITT, RTT, C, x and y using the data obtained from the supersonic detectors; here, the RTT was estimated by using the method described in section 4.1.

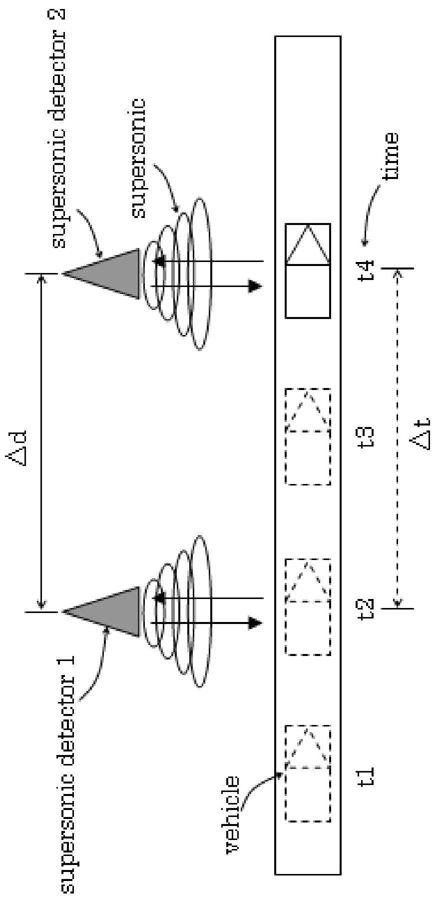


Figure 7. Supersonic detector

The value of C was calculated using equation (2); x and y were derived from the data obtained from the supersonic detectors, as described in section 4.3. All this information was saved in a data table for predicting the RTT.

First, we focused on the analysis in the period from 11:55 to 12:40 on Feb 11, 2006, because heavy congestion had occurred during this time segment. Figure 8 shows the primary validation result for this period. We can observe that the performance of our model is better than that obtained using the ITT .

A more rigorous validation was also carried out. Predicting the travel time under congested traffic conditions is very difficult; this problem has always challenged traffic managers. This section mainly focuses on predicting the travel time under traffic congestion conditions (ITT > 0.2 h). We assume that the traffic conditions are considered abnormal if ITT > 0.5 h, e.g., when an accident delays traffic. Prediction of the travel time under abnormal conditions is out of the scope of this study. Hence, a test data set was drawn from the data pool of the year 2005 having 0.2 h < ITT < 0.5 h. A total of 182 test data points were randomly selected for evaluation purposes.

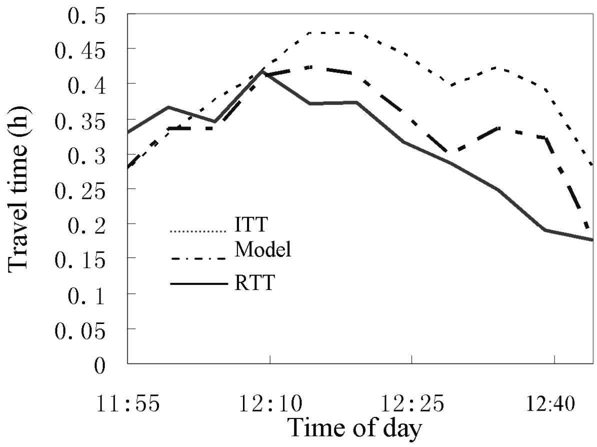


Figure 8. Primary validation result

The mean absolute percentage error (MAPE) was used to evaluate the proposed prediction model. This MAPE was defined as follows:

$$MAPE_{R\hat{T}T} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|R\hat{T}T_i - RTT_i\right|}{RTT_i}\cdot 100\% \tag{10}$$

Let MAPE of the ITT be:

$$MAPE_{ITT} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|ITT_i - RTT_i\right|}{RTT_i}\cdot 100\% \tag{11}$$

Figure 9 shows a comparison of the MAPE of the ITT and that of the proposed model for different values of the ITT . From this figure, it is observed that the MAPE of the ITT obviously increases with an increase in the ITT . This implies that the error will obviously increase where the traffic is under heavy congestion. On the other hand, the MAPE of the proposed model is not influenced by the degree of traffic congestion and remains stable irrespective of the value of the ITT (We use the ITT as the criterion to represent the degree of congestion.). Figure 9 also shows that the proposed model can provide more useful information than the ITT under heavy congestion. For example, when 0.4 h <

ITT < 0.5 h, the MAPE of the proposed model and the ITT are 27% and 40%, respectively. This implies that the error is reduced by about 4 min by using the proposed model.

6. Analysis

The chart shown in Figure 10 was prepared in order to

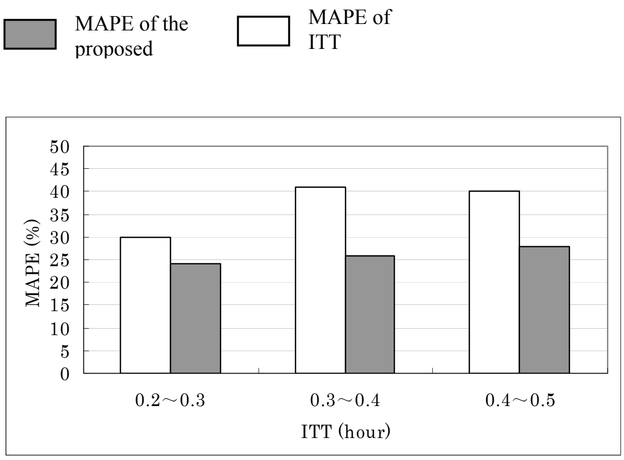


Figure 9 Validation result grouped by ITT

analyze the relationship between the slope of the linear model, the value of C, and the value of x. We can observe that the slope of the linear model is high if the tail of the congestion is located downstream of the segment and the value of C is small (denoted by a solid line). However, no obvious trend can be observed if the tail of the congestion is located downstream of the segment and the value of C is high (denoted by a broken line). A high value of C indicates that the traffic condition is unstable; hence, it is difficult to find an obvious trend under these conditions.

7. Conclusions and future work

The proposed model was tested for traffic on the Hanshin Expressway. From the results, it was found that by using the proposed model, a stable and acceptable travel time prediction can be obtained under heavy traffic congestion. Compared to the ITT model, which currently provides information to the users on this expressway, the proposed model is not only more accurate but also more consistent in the prediction of errors. Moreover, this model is independent of the degree of traffic congestion.

In this study, the influence of an untoward incident was not considered. We intend to combine the incident factor with the travel time prediction model in a future study. For example, an incident could be quantified by a measurement, and then the relationship between the measurement and the parameters of the travel time prediction model could be analyzed. The congestion pattern is not always as described in Figure 1 (a). We can assume that the other congestion patterns are equivalent to Figure 1 (a) if the expressway segment is not very long, but we intend to discuss the effects of variations in the length of the expressway segment in a future study.

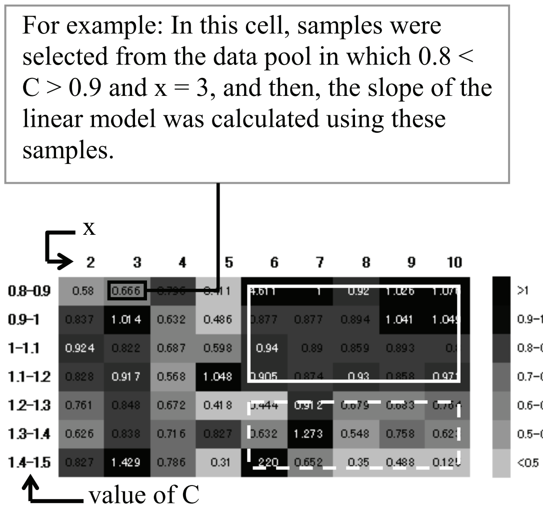


Figure 10. Relationship among C, x, and slope

8. Acknowledgements

The authors would like to thank the Hanshin Expressway Corporation for providing the data used in this research. We also would like to thank the anonymous reviewers for valuable comments.

9. References

[1] Park, D., Rilett, L. (1999) Forecasting motorway link travel times with a multilayer feedforward neural network. Computer-Aided Civil and Infrastructure Engineering 14: 357–367.

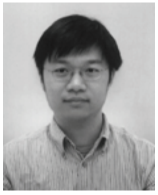
[2] Chien, S., Kuchipudi, C. (2002) Dynamic travel time prediction with real-time and historical data. Transportation Research Board, Washington DC. Remarks: Paper 02–2548.

[3] Zhang, X., Rice, J.A. (2003) Short-term travel time prediction. Transportation Research Part 11 C (3), 187–210.

[4] Kwon, J., Coifman, B., Bickel, P.J. (2000) Day-to-day travel time trends and travel time prediction from loop detector data. In: Transportation Research Record 1717, 120–129.

[5] Otaka, H., Horiguchi, R., Akahane, H. Short Term Travel Time Prediction Using Cumulative Traffic Counts And Uplink Information Collected on Expressways, Proceedings of 11th World Congress on ITS, Nagoya, October, 2004.

[6] Kaumann, O., Froese, K., Chrobok, R., Wahle, J., Neubert, L., Schreckenberg, M. (2000) Online simulation of the freeway network of NRW. Traffic and Granular Flow ’99, 351–356, eds. Helbing, D., Hermann, H.J., Schreckenberg, M., Wolf, D.E. (Springer, 2000)



Chong Wei is a doctoral course student at Kobe University. He received his Masters degree from Kobe University in 2007. His research interests include travel behavior modeling and traffic flow data collection and analysis.

Takamasa Iryo is an assistant professor at Kobe University. He received his Ph.D degree from the University of Tokyo. He is interested in network analysis and traffic assignment.

Yasuo Asakura is a professor at the Graduate School of Engineering, Kobe University. He has been studying travel and traffic data collection and modeling. In particular, he is interested in the analysis of travel behavior in a network. He was awarded the outstanding paper award during the interactive session at the 11th ITS World Congress in Nagoya in 2004.

Received date: March 21, 2008
Received in revised forms: July 15, 2008
December 11, 2008,
April 27, 2009
Accepted date: June 14, 2009
Editor: Hirokazu Akahane