A STUDY ON ESTIMATION OF PROBABILISTIC CHANGING TRAVEL TIME BASED ON BAYESIAN STATISTICS

Makoto KASAI, Tokyo University of Science, Japan; kasai@rs.noda.tus.ac.jp
Hisao UCHIYAMA, Tokyo University of Science, Japan; uchiyama@rs.noda.tus.ac.jp

ABSTRACT

Intelligent Transportation Systems (ITS) provide highly accurate time travel predictions to support rational decision making. Most current estimation methods, however, deal with travel time as a stationary event; that is, they do not take it into account non-stationary flow states influenced by irregular cases such as traffic accidents.

This paper suggests the necessity of a method to estimate travel time under non-stationary flow states following three steps: 1) assuming that travel time can treated as a stochastic event, 2) probabilistically extracting future estimated travel time distributions from large amounts of accumulated daily travel data, and 3) updating the distributions with observational results of non-stationary conditional factors.

Keywords: Travel time prediction, non-stationarity, Bayesian statistics, ITS

1. INTRODUCTION

There is broad consensus that information services are essential for Transportation Demand Management (TDM). Real-time information about road conditions assists travelers in making decisions such as choosing routes and departure times. For drivers in particular, “future” predicted travel times are more important than current predictions, because travel times change irregularly. Therefore, various methods for travel time prediction have been developed.

Conventional methods for travel time prediction are categorized into two general groups: model-based and data-oriented approaches. The former use linear regressive models (Huang & Barth (2008), Rice & van Zwet (2004), Zhang & Rice (2003)), time-series analysis as represented by Auto-Regressive Integrated Moving Average (ARIMA) models (Billings & Yangt (2006), Guin (2006)), the Kalman filtering technique (Vanajakshi et al. (2009)), stochastic process (Jou et al. (2003), Yeon et al. (2008)), and artificial neural networks (Wen et al. (2005), van Lint (2008), Yu et al. (2008), van Hinsbergen et al. (2009), Zhu & Wang (2009)), and have been applied with remarkable success in recent years. In spite of such
models’ high accuracy estimations under stationary state flow, however, there are few explicit references to handling non-stationary state traffic flow, for example traffic flow under the effect of traffic accidents or debris. This is the case even though travelers would prefer predicted travel times under non-stationary states.

The ARIMA model is often applied to non-stationary time series with trend or seasonal variations, where first differencing of the original time series is assumed to fulfill weak stationarity (Vlahogianni et al. (2006)). If a nonlinear structure is hidden within the observed time series, however, such handling is inadequate. Unfortunately, it is well known that traffic flow has high nonlinearity (Wang et al. (2008)).

Artificial neural network (ANN) models have also attracted attention. Since the advantage of an ANN is to naturally take account of nonlinearity, it is suitable for illustrating traffic flow phenomena. Some shortcomings, however, can be pointed out. For example, calibration is time consuming in general, and the topology of the network is not uniquely determined. In other words, arbitrariness is difficult to exclude (Zhang & Rice (2003)). Nevertheless, if non-stationary factors related to travel time could be specified, application of an ANN might develop a reliable model under both stationary and non-stationary states. To our knowledge, however, there are no studies focusing on non-stationarity.

Now is a suitable time to begin attempting data-oriented approaches, because data from traffic detectors has accumulated over many years. As potential approaches, the k-nearest neighbors method (Zou et al. (2008)) and the pattern matching method (Bajwa et al. (2004), Bajwa et al. (2005), Warita et al. (2004)) are representative methods. Both are predicated on similar hypotheses, namely that traffic patterns of past days are similar to that of the current day, and patterns recur not only according to the current time but also according to successive times. Such pattern matching methods perform particularly well in spite of their use of naïve algorithms. These methods, however, do not consider the abrupt changes in travel time that can occur due to incidents such as traffic accidents or road debris. Although only macroscopic and aggregated data are recorded in most databases, various traffic conditions under not only stationary but also non-stationary states are contained therein. Therefore, new concepts for utilization of these huge amounts of data are required.

We have previously proposed a prediction method for non-stationarity using Bayesian statistics and a data-oriented approach (Kasai et al. (2008)). Although it is confirmed that the Bayesian approach improved precision under statistical tests (Winkler (1972)), the ‘lower-right problem’, a tendency to under-estimate relatively long travel times, was indicated as an area for future research. The method proposed in this paper improves upon the previous method, refining the application of Bayes’ theorem to the prediction, while retaining the advantage for the variation in travel times in a non-stationary state.

2. **DATA PROFILE**

2.1. **The MEX and a study route**

The Metropolitan Expressway (MEX) covers the Tokyo Metropolitan Area, and the large amount of traffic inbound to the central business district (CBD) creates serious traffic jams almost every day. This paper uses MEX Route 6 (called the Misato Line, and linked with the
Mukoujima Line from the Misato Junction to the Edobashi Junction) as a study route. A recently opened commuter train line (called the Tsukuba Express) runs parallel to the route. Figure 1 shows the study route. The length of the section is around 20 kilometers, and includes a configuration of the MEX network.

Figure 1 – MEX network and study route

2.2. Profile of traffic detector and related data

Traffic data from 33 different traffic detectors installed in series along the study route are recorded automatically in 5-min increments. Figure 2 shows the traffic detector data format. This paper assumes that the coverage area of each traffic detector extends to the midpoint of the adjacent detector. The allocated length corresponds to the coverage area. For the development of a methodology, this study uses annual data from 2004, from which 17 days with aberrant values (e.g., travel time is zero) were excluded.

Monthly traffic accidents and debris data (inhibited flow data) as reported by MEX are available for every 5 min during September 2004. Precipitation data for every hour during the same period are also available from the Japan Meteorological Agency (JMA).

Because the detector does not directly measure travel times, times were calculated and assigned to each section as a quotient of the distance from a counter to its adjacent one, divided by velocity. This study uses two types of travel times: time slice travel time and instantaneous travel time.

2.3 Estimation of time slice travel time

Time slice travel time is an approximated value of the actual travel time, and is calculated based on the estimated trajectory of a car as if it ran in a time-space plane (Matsuba et al. (2004)). The procedure can be summarized as follows (also see Figure 3): 1) it is assumed that at time $t = \tau_0$, a car flows into the $i$th section of length $l_i$, and that $v_{i,j}$ is the spatial
A STUDY ON ESTIMATION OF PROBABILISTIC CHANGING TRAVEL TIME BASED ON BAYESIAN STATISTICS

Makoto KASAI; Hisao UCHIYAMA

12th WCTR, July 11-15, 2010 – Lisbon, Portugal

mean velocity in the \( j \)th time slot corresponding to the interval \( \Delta t \) of data collection; 2) temporary local travel time through the \( i \)th section is calculated as the distance \( l_i \) divided by the velocity \( v_{i,j} \); 3) if \( \tau_0 + \tau_j' < \Delta t \) (case 1), that is, if a car passes through section \( i \) in the same time slot \( \Delta t \), then the local travel time \( \tau_{i,j} \) is fixed as \( \tau_{i,j} = \tau_j' \). Otherwise, if \( \tau_0 + \tau_j' \geq \Delta t \) (case 2), that is, if the car traverses the time slot \( \Delta t \), the local travel time is composed by the summation of \( \tau_{i,j} = \Delta t - \tau_0 \) and \( \tau_{i,j+1} = \frac{(l_i - v_{i,j} \tau_j)}{v_{i,j+1}} \); and 4) in a similar way, the trajectory of cars flowing into the study route at a given time from the 1st to the 33th section is estimated by repeating 2)-3).

<table>
<thead>
<tr>
<th>Traffic detector No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allocated length[m]</td>
<td>460</td>
<td>880</td>
<td>940</td>
<td>...</td>
<td>290</td>
</tr>
<tr>
<td>10:00 - 10:05</td>
<td>21.3</td>
<td>23.6</td>
<td>17.0</td>
<td>...</td>
<td>86.5</td>
</tr>
<tr>
<td>10:05 - 10:10</td>
<td>29.9</td>
<td>34.7</td>
<td>22.0</td>
<td>...</td>
<td>89.3</td>
</tr>
<tr>
<td>10:10 - 10:15</td>
<td>27.7</td>
<td>37.2</td>
<td>33.7</td>
<td>...</td>
<td>90.5</td>
</tr>
<tr>
<td>10:15 - 10:20</td>
<td>25.1</td>
<td>29.7</td>
<td>23.5</td>
<td>...</td>
<td>91.7</td>
</tr>
<tr>
<td>10:20 - 10:25</td>
<td>30.9</td>
<td>27.8</td>
<td>16.5</td>
<td>...</td>
<td>93.6</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Figure 2 – Data format of traffic detector data and actual travel time estimation

<table>
<thead>
<tr>
<th>Traffic accident</th>
<th>Rainfall</th>
<th>Fallen debris</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3 – Estimation method of time slice travel time

(Case1: without stride across boundary of interval of data collection) (Case2: without stride across boundary of interval of data collection)
Figure 4 shows an example estimation of a time slice value of travel time. Assuming that a route is composed of 3 sections and a car flows into the route at $t = 0$ [min], the time slice travel time $T_s$ is computed as follows:

$$
\begin{align*}
\tau_{1,1} &= \frac{460}{1000} \cdot 0.60 = 1.30 \text{ [min]} \\
\tau_{2,1} &= \frac{880}{1000} \cdot 0.60 = 2.24 \text{ [min]} \\
\tau_{3,1} &= \frac{940}{1000} \cdot 0.60 = 3.32 \text{ [min]} \\
\tau_{1,2} + \tau_{2,1} + \tau_{3,1} &= 6.86 > \Delta t = 5 \text{ [min]} \\
\tau_{3,2} &= (\Delta t - \tau_{2,1} - \tau_{3,1}) = 5 - 1.30 - 2.24 = 1.46 \text{ [min]} \\
\tau_{3,2} &= \frac{(l_3 - v_{3,1} \cdot \tau_{3,1})}{v_{3,2}} = \frac{940 - (17.0 \times 60/1000) \times 1.46}{22.0} \cdot 60 = 1.43 \text{ [min]} \\
T_s &= \tau_{1,1} + \tau_{2,1} + \tau_{3,1} + \tau_{3,2} = 6.43 \text{ [min]}
\end{align*}
$$

(1)

2.4 Estimation of Instantaneous Travel Time

Instantaneous travel time is measured as a summation of the number of allocated lengths divided by instantaneous velocity. Generally, instantaneous travel time is expressed as follows:

$$
T_i = \sum_k I_k \frac{l_k}{v_k},
$$

(2)
where \( l_k \) and \( v_k \) are the length and velocity of the \( k \)th section, respectively. The instantaneous travel time is used in the screening process described in Section 4.

For example, the instantaneous travel time at \( t = 0 \) in Figure 4 is estimated as follows:

\[
T_i = \frac{l_1}{v_1} + \frac{l_2}{v_2} + \frac{l_3}{v_3} = \frac{460/1000}{21.3} \cdot 60 + \frac{880/1000}{23.6} \cdot 60 + \frac{940/1000}{17.0} \cdot 60 = 6.85 \text{ [min].} \tag{3}
\]

### 3. PREDICTION UNDER NON-STATIONARY STATES

This section describes dealing with abrupt changes in travel time with the help of Bayes’ theorem (Winkler (1972)).

Suppose that an observer estimates in advance a probability \( p(c_i) \) of future travel time \( c_i \), which in Bayesian statistics is called a prior probability distribution. Later, an event \( x \) such as a traffic accident or road debris causes an abrupt change in travel time. Immediately, the prior distribution is updated by Bayes’ theorem (4):

\[
p(c_i | x) = \frac{p(x | c_i) p(c_i)}{\sum_k p(x | c_k) p(c_k)}, \tag{4}
\]

where \( p(c_i) \): prior probability of travel time \( c_i \); \( p(x | c_i) \): conditional probability of an event \( x \) given \( c_i \); \( p(c_i | x) \): conditional probability of an event \( c_i \) given \( x \).

\( p(c_i | x) \) is also called a posterior probability of \( c_i \) given new information \( x \). \( p(x | c_i) \) can also be prepared in advance for estimation of concrete posterior probabilities.

Hereinafter, preparing \( p(c_i) \) is called the screening process and estimating \( p(c_i | x) \) is called handling a non-stationarity process.

Kasai et al. (2008) implied that the concept above works acceptably as a whole, although the model occasionally fails to address relatively longer travel times in what is called the “lower-right problem”. Trying to identify the cause of the problem, Miyata et al. (2009) suggested two possibilities: 1) preparation of the prior probability (the screening process) or 2) the effect on travel time of those events assumed as factors (handling non-stationarity process). That paper also pointed out that the former is the dominant factor. The following section discusses necessary arrangements for addressing the prior probability.

### 4. ARRANGEMENTS OF PRIOR PROBABILITY

Referring to the pattern matching method of Warita et al. (2004), Kasai et al. (2008) designed a prototype module for the screening process. Figure 5 shows the screening mechanism. During the matching period, from 2 h before up to the current time, the discrepancy between the historically accumulated data in the database and the current data is calculated as the root mean square error (RMSE). The days satisfying the criterion that the RMSE is smaller than 20 min are extracted.

12th WCTR, July 11-15, 2010 – Lisbon, Portugal
The prototype module rejects historical patterns that do not perfectly fulfill its criterion. For instance, if the RMSE between a historical pattern and the current pattern is 21 min, the possibility of recurrence of the historical pattern is calculated as exactly 0, even though the criterion is exceeded by only 1 min. If the traffic condition of a given day is in a non-stationary state, however, wider screening may be preferable.

In the present paper, we propose a second module, in which it is assumed that the recurrence probability \( p(h_i) \) of historical data \( h_i \) is weighted by similarity \( S_i \) between the historical and the current data. Thus,

\[
p(h_i) = \frac{s_i}{\sum_k s_k}, \quad (i, k = 1, 2, \ldots, N)
\]

where \( N \) denotes the number of days in the database.

The second module assumes the similarity \( s_i \) to be a monotonically decreasing function of RMSE in the matching period. A negative exponential function \( s_i = \lambda e^{-r_i} \), which is named as a similarity function here, is suitable for fulfilling the assumption. \( r_i \) and \( \lambda \) denote the RMSE between the current and historical patterns for a matching period, and the coefficient of the exponential function, respectively. In addition, the pattern matching technique in the prototype corresponds to the case that a discrete function \( s_i \) is applied as similarity \( s_i \), where \( r_0 = 20 \text{ [min]} \). Figure 6 shows the similarity functions in the cases of both the prototype and the second module.
A STUDY ON ESTIMATION OF PROBABILISTIC CHANGING TRAVEL TIME BASED ON BAYESIAN STATISTICS
Makoto KASAI; Hisao UCHIYAMA

12th WCTR, July 11-15, 2010 – Lisbon, Portugal

Figure 6 – Similarity function for the screening process

The prototype module: \( s_i = \begin{cases} 1 & (r \leq r_0) \\ 0 & (r > r_0) \end{cases} \)

The second module the paper proposes: \( s_i = \lambda e^{-\lambda r_i} \)

\( r_0: 20[\text{min}] \quad r_i: \text{RMSE between current and historical pattern in matching period} \)

Figure 7 – Concept of screening process with exponential likelihood function

Travel time \( T(t) \) [min] 

Similarity function: \( s_i = \lambda e^{-\lambda r_i} \ (\lambda = 0.1) \)

\[ r_1 = \sqrt{\frac{12^2 + 6^2 + 7^2 + 13^2}{2}} = 10.0 \]

\[ s_1 = 2e^{-0.1 \times 10.0} = 3.68 \times 10^{-2} \]

\[ r_2 = \sqrt{\frac{21^2 + 19^2 + 18^2 + 21^2}{2}} = 20.0 \]

\[ s_2 = 2e^{-0.1 \times 20.0} = 1.35 \times 10^{-2} \]

Prior probability of future travel time:

\( p(h_1) = \frac{s_1}{s_1 + s_2} = \frac{3.68 \times 10^{-2}}{3.68 \times 10^{-2} + 1.35 \times 10^{-2}} = 0.73 \)

\( p(h_2) = \frac{s_2}{s_1 + s_2} = \frac{1.35 \times 10^{-2}}{3.68 \times 10^{-2} + 1.35 \times 10^{-2}} = 0.27 \)

Day (1)
Current day
Day (2)

Matching period
The current time: \( t_0 \)
The future time: \( t_0 + \Delta t \)

Figure 7 shows a simple example of the screening process based on the negative exponential similarity function. Suppose that 1) only two days (Day 1 and Day 2) are recorded in the database of historical travel times, 2) a future travel time \( T(t_0 + \Delta t) \) yielding to a binomial distribution, that is, either \( 70 \leq T(t_0 + \Delta t) < 100 \) or \( 100 \leq T(t_0 + \Delta t) < 130 \), is observed, and 3) \( \lambda \) is assumed as 0.1[\text{min}^{-1}]. If the travel time for the current day is recorded...
as the bold line in the figure, the probability of recurrence for Day 1 and Day 2 of the current day are 0.73 and 0.27, respectively. The future travel time distribution is probabilistically determined as an accumulation of the historical data, weighted by the probability of recurrence. In the example, the prior probability distribution of future travel times is prepared as shown in the figure simply because the historical data is limited to two days.

For matching, the instantaneous travel time is used since time slice travel time for a driver flowing into the study route is not calculated at the current time. Therefore the similarity between historical data and the current day’s data is measured using the instantaneous travel time. On the other hand, the prior distribution of travel time at a future time (the target time) is estimated based on time slice travel time in the historical database.

5. ARRANGEMENTS OF CONDITIONAL PROBABILITY

As mentioned in Section 4, the prior probability distribution of travel times is updated using the handling non-stationarity process, which is the conditional probability of travel times given events on the road. This section shows 1) how to translate historical data to a conditional probability, and 2) which events are considered.

The historical travel time data, the inhibited-flow data from MEX, and precipitation data from JMA are each translated to the frequency by travel time in events on the road. The frequency is regarded as the joint probability of travel time and events. Table I is an example showing the case of the frequency of travel time 30 min after event observation. In addition, only traffic accidents and rainfall are recorded as events for explanation in the table.

Suppose that 1) prior probabilities of travel time are distributed uniformly, that is,
\[ p(c_i) = \frac{1}{16} = 6.25 \times 10^{-2}, (i = 1, 2, \ldots, 16), \]
2) when a traffic accident has just been observed \((x_1)\), it has not been raining on the study route \((x_2)\), and 3) a traveler will flow into the route 30 min later. Referring to Table I, the posterior probability \(p(c_1 | x_1, x_2)\) of travel time \(10 \leq c_1 < 20\) is then estimated as follows:

\[
p(c_1 | x_1, x_2) = \frac{p(x_1, x_2 | c_1) p(c_1)}{\sum_{k=1}^{16} p(x_1, x_2 | c_k) p(c_k)}
\]

\[
= \frac{(3.13 \times 10^{-2} / 43.7 \times 10^{-2}) \cdot 6.25 \times 10^{-2}}{(3.13 \times 10^{-2} / 43.7 \times 10^{-2}) \cdot 6.25 \times 10^{-2} + \cdots + (0.07 \times 10^{-2} / 0.07 \times 10^{-2}) \cdot 6.25 \times 10^{-2}}
\]

\[
= 2.27 \times 10^{-2}.
\]

Similarly, the other posterior probabilities are
\[ p(c_5 | x_1, x_2) = 0.58 \times 10^{-2}, \ldots, p(c_{16} | x_1, x_2) = 0.318. \]

Miyata et al. (2009) considered four representative events related to expanding travel time: traffic accidents, debris, rainfall, and traffic weaving. In addition to such events, the traffic rate from the Misato Junction flowing into the study route is also considered in the present paper.
Table I – Frequency by travel time in rainfall and traffic accident event

<table>
<thead>
<tr>
<th>Travel time [min]</th>
<th>Traffic accident $x_1$ and rainfall $x_2$ [%]</th>
<th>Not-traffic accident $\bar{x}_1$ and rainfall $x_2$ [%]</th>
<th>Traffic accident $x_1$ and not-rainfall $\bar{x}_2$ [%]</th>
<th>No events $\bar{x}_1, \bar{x}_2$ [%]</th>
<th>Total [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$ 10 - 20</td>
<td>0.19</td>
<td>0.57</td>
<td>3.13</td>
<td>3.98</td>
<td>43.7</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$c_5$ 50 - 60</td>
<td>0.05</td>
<td>0.59</td>
<td>0.13</td>
<td>6.23</td>
<td>6.99</td>
</tr>
<tr>
<td>$c_6$ 60 - 70</td>
<td>0.23</td>
<td>0.79</td>
<td>0.16</td>
<td>4.32</td>
<td>5.50</td>
</tr>
<tr>
<td>$c_7$ 70 - 80</td>
<td>0.09</td>
<td>0.96</td>
<td>0.31</td>
<td>3.10</td>
<td>4.47</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$c_{16}$ 160 - 170</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
<td>0</td>
<td>0.07</td>
</tr>
<tr>
<td>Total [%]</td>
<td>1.68</td>
<td>6.37</td>
<td>6.66</td>
<td>85.31</td>
<td>100.0</td>
</tr>
</tbody>
</table>

6. FINDINGS

6.1. Calibration of the second module for the screening process

In this section, we discuss the accuracy of travel time predictions for 1 h in the future. To estimate the prior probability of travel time in the screening process, the unknown parameter $\lambda$ of the exponential function (similarity function) is estimated as shown below. This paper searches for a value of $\lambda$ minimizing the root mean square error between the actual and predicted travel times for all days in September 2004. The calibration uses 8329 samples, data samples taken every 5 min from September 2 to September 30. Since the RMSE is not an analytical function of $\lambda$ we apply the simplex method, an explanatory solving method for nonlinear optimization. Consequently, the estimated value of $\lambda$ is 0.636 [min$^{-1}$], while RMSE concurrently reaches 5.78 min.

The right part of Figure 8 shows the results of the calibration as a scattered plot graph. The left part shows the results before the Bayes’ update. The horizontal axis indicates estimated travel time calculated as the expected value of the probability distribution, and the vertical axis indicates the actual time. The left part shows a trend of slight under-estimation in cases where travel time was over 120 min, while after the update the trend seems to disappear.
A STUDY ON ESTIMATION OF PROBABILISTIC CHANGING TRAVEL TIME BASED ON BAYESIAN STATISTICS
Makoto KASAI; Hisao UCHIYAMA

12th WCTR, July 11-15, 2010 – Lisbon, Portugal

As a comparison of these results with those of the previous paper, Figure 9 shows the results obtained by the prototype module for the screening process. As above, predicted travel times on the left are determined according to the expected value of the prior probability, and those on the right correspond to the posterior probability. Even after the update using Bayes' theorem (the right graph), the lower-right problem, the tendency to underestimate travel time, is unsolved, while Figure 8 indicates that the problem seems to have disappeared.

Figure 8 – Scattered graph of relation between actual and predicted travel time with the third model the paper proposes

Figure 9 – Scattered graph of relation between actual and predicted travel time with the prototype module
6.2. Comparison from variation of travel time on the day

In this section, we discuss the following two topics: 1) comparison of the prototype and the improved module this paper proposes for the screening process and 2) the effect of using Bayes’ theorem for handling non-stationarity processes. Figure 10 and Figure 11 are time-series graphs of travel time. Figure 10 shows results from the prototype module, and Figure 11 shows results from the second module. In both figures, the green line indicates the actual travel time on September 27, 2004. For drivers flowing into the study route at around 11:00, travel time was over 150 min. The blue line is the expected value of the prior probabilities from the prototype module, and the red is that of the posterior probabilities.

In Figure 10, the expected value of the prior probability (blue line) suddenly falls off shortly after 10:00. Although the Bayes’ theorem update changes the blue line to the red one, a large gap between the actual and updated travel times remains. In Figure 11, neither line falls suddenly as in Figure 10. This implies that the proposed second module is more effective than the prototype module for prior probability estimations. Moreover, examining the red line, the expected value of the posterior probability from the second module becomes closer to the actual travel time. The Bayes’ update is thus more effective than the prototype module.

In order to measure the effect of the Bayesian update, the Kullback-Leibler divergence, which indicates the difference of two distributions, is calculated every 5 min and drawn in both figures. Given the probability distributions \( f(x_i) \) and \( g(x_i) \), the divergence \( I(g; f) \) is defined as follows:

\[
I(g; f) = \sum_i g(x_i) \log \frac{g(x_i)}{f(x_i)}.
\]  

We assigned the prior probability and the posterior probability to \( f(x_i) \) and \( g(x_i) \), respectively. In both figures, the Kullback-Leibler divergence becomes large where rainfall or traffic accidents were observed. In particular, Figure 11 indicates large values for the Kullback-Leibler divergence between 9:00 and 12:00. In fact, the expected value of the posterior probability approaches the actual travel time during the time slot, except for around 10:30. It is considered that this outlier results from the calculation of the conditional probability because the posterior estimation is worse than the prior one. This indicates that consideration of the compliment of traffic accidents and debris data is required for future research.
7. CONCLUSIONS

This paper attempted to improve upon previous models for travel time prediction. To that end, the following points were discussed: 1) how to refine the screening process for historical databases to develop prior probability of travel time distributions, and 2) adding events that suddenly change travel times. We confirmed that prediction accuracy is satisfied in terms of RMSE. The results indicated that the improved model has a much higher ability to predict travel times in not only stationary states but also in non-stationary states. Therefore, the pilot model developed here contributes toward such predictions through 1) the modeling of travel time that considers abrupt change, despite using a simple application of Bayes’ theorem, and 2) its indication that the vast amount of accumulated traffic detector data should be more commonly utilized for modeling.
Information related to predicted traffic conditions may push TDM into practice in the near future. The concept of Dynamic Park & Ride, where travelers transfer from the expressway to railway running parallel to, has high potential for mitigating congestion. Travel time prediction methods will play an important role in this context.

ACKNOWLEDGEMENTS

This research was supported by Kosokudoro-kanren-shakaikoken-kyogikai in Japan. The authors thank MEX (Metropolitan Expressway Co., Ltd.) for providing the traffic detector data and related data used in this study. We gratefully acknowledge the valuable comments from Dr. Warita (MEX). We also thank Mr. Konagai (Nomura Research Institute, Ltd), Mr. Rokutan (Pacific Consultants Co., Ltd.) and Mr. Miyata (Bank of Yokohama Co., Ltd.) for cooperation in support of this study.

REFERENCES


A STUDY ON ESTIMATION OF PROBABILISTIC CHANGING TRAVEL TIME BASED ON BAYESIAN STATISTICS

Makoto KASAI; Hisao UCHIYAMA


