SHORT - TERM TRAVEL TIME PREDICTION USING LINEAR MODEL

Chong WEI  
Graduate Student  
Graduate School of Engineering  
Kobe University  
1-1, Rokkodai-cho, Nada,  
Kobe, 657-8501, Japan  
Fax: +81-78-803-6360  
E-mail: 0581781t@stu.kobe-u.ac.jp

Takamasa IRYO  
Assistant Professor  
Graduate School of Engineering  
Kobe University  
1-1, Rokkodai-cho, Nada,  
Kobe, 657-8501, Japan  
Fax: +81-78-803-6360  
E-mail: iryo@kobe-u.ac.jp

Yasuo ASAKURA  
Professor  
Graduate School of Engineering  
Kobe University  
1-1, Rokkodai-cho, Nada,  
Kobe, 657-8501, Japan  
Fax: +81-78-803-6360  
E-mail: asakura@kobe-u.ac.jp

Abstract: Prediction of travel times is a vital part of many advanced traveler information systems. Linear model is used in this research as it is effective, computationally efficient and reliable to the available freeway detector data. In this paper an explicit analysis of the linear model is shown and a method to predict freeway travel times using a linear model is proposed. The proposed method is applied to a data set of Hanshin Expressway, Japan.

Key Words: Travel time prediction, Linear model, Weighted least squares

1. INTRODUCTION

Travel time information was considered as a vital component of ATIS (Advanced traveler Information Systems). In recent year, the different media have been in use for providing travel time information; they include internet, in-vehicle information, and roadside based information. VMS (Variable Message Signboard) is one kind of widely used equipment for providing travel time information either in developed country or developing country. In Japan, Hanshin Expressway Corporation is providing travel time information of some major segments using VMS. However these travel times information are only so-called instantaneous travel time of the segment as shown in Figure 1. Instantaneous travel time is a simple accumulation of link travel times which is calculated as the length of a link divided by the current velocity of the link include in the segment.
If traffic flow is stable and the link travel time is constant, the instantaneous travel time is equal to the real travel time. However, the link travel times may change due to the traffic conditions, therefore the instantaneous travel time is not equal to the real travel time as shown in Figure 1. (Isii 2006)

There have been studies on the topic of travel time prediction. The methodologies used in the previous work include artificial neural networks (Park and Rilett, 1999), Kalman filter (Chien and Kuchipudi, 2002), regression model (Zhang and Rice, 2003), pattern matching (Horiguchi et al., 2003) and microscope simulation (Kaumann et al., 2000). In this study, first, an explicit definition of travel time is presented: the concern is focused on a highway system without signal control; the travel time is defined as the passing time interval between the origin and the destination, where a traveler moving from the origin to the destination along a highway segment. We propose a method for predicting the travel time. The method will provide the precise and reliable travel time information than the method based on the instantaneous travel time. The proposed method not only can be operated but also easily understood by the highway managers, and the method can be easily transplanted to the other places without complicated localization processing.

2. METHODOLOGY DESCRIPTION

2.1 Analysis
In this section, the relationship between the instantaneous travel time and the real travel time has been analyzed. The focus is on the general phenomena of traffic flow and transforms the real-world to an abstract model. The capability limitation of the detector is considered in the analysis.
Figure 2 describes the relationship between instantaneous travel time and real travel time.

In Figure 2, $V_{\text{free}}$ is the velocity of the vehicle at free flow and $V_{\text{extent}}$ is the speed of the increase or decrease of the congestion extent. $t$ is the moment of the vehicle entering the highway segment. $x$ is the distance between the entrance of the highway segment and the tail of the congestion correspond to $t$. At $t + \text{gap\_t}$ moment the vehicle meet with the congestion. Traffic condition information is updated every 5 min interval.

An assumption made here that there are only two conditions on the highway: free flow and congestion flow. Let $RTT$ and $ITT$ represent real travel time and instantaneous travel time respectively. The aim is to analyze the relationship of $RTT$ and $ITT$ at $t$ moment. However $RTT$ and $ITT$ are two different conceptions that trouble us to directly analyze the relationship of $RTT$ and $ITT$. For solving this problem, $RTT$ is transformed into $ITT$ first. As shown in Figure 2, the equation follows:

$$RTT(t) = ITT(t + \text{gap\_t})$$  \hspace{1cm} (1)

where $RTT(t)$ is the $RTT$ at $t$ moment, and $ITT(t + \text{gap\_t})$ is the $ITT$ at $t + \text{gap\_t}$ moment. Now the aim can be transformed into the analysis of the relationship between $ITT(t + \text{gap\_t})$ and $ITT(t)$. The relationship can be described as:

$$ITT(t + \text{gap\_t}) = f(\cdot) \cdot ITT(t)$$  \hspace{1cm} (2)

where $f(\cdot)$ is an unknown function. One naturally consideration is that the value of the function $f(\cdot)$ is decided by $\text{gap\_t}$ and one measurement which can describe the change of $ITT$. Here $C(t)$ is used as the measurement.

$$C(t) = \frac{ITT(t)}{ITT(t - 5\text{min})}$$  \hspace{1cm} (3)

Now the equation (2) can be rewritten as

$$ITT(t + \text{gap\_t}) = f[C(t), \text{gap\_t}] \cdot ITT(t)$$  \hspace{1cm} (4)
\( \text{gap}_t \) is the time gap between the vehicle entering the highway segment and meeting with the congestion. \( \text{gap}_t \) can be calculated as:

\[
\text{gap}_t = \frac{x}{V_{\text{free}} + V_{\text{extent}}}
\]  

(5)

Where \( V_{\text{extent}} \) can be decided by \( C(t) \). \( V_{\text{free}} \) is a constant. So that equation (4) also can be written as:

\[
\text{ITT}(t + \text{gap}_t) = f[C(t), \frac{x}{V_{\text{free}} + V_{\text{extent}}(C(t))}] \cdot \text{ITT}(t) = F[C(t), x] \cdot \text{ITT}(t)
\]  

(6)

Where function \( F[C(t), x] \) has a similar structure as function \( f(.) \) but the independent variables are different. As the analysis above, we can find the relationship between \( \text{RTT} \) and \( \text{ITT} \) as following form:

\[
\text{RTT}(t) = F[C(t), x] \cdot \text{ITT}(t)
\]  

(7)

The value of function \( F(C, x) \) is determined by the value of \( C \) and \( x \) correspond to \( t \). Different \( t \) may have the same value of \( C \) and \( x \). If the couples of \( \text{RTT} \) and \( \text{ITT} \) which have the same value of \( C \) and \( x \) are selected, the equation (8) can be used to describe the relationship of the selected sub data set:

\[
\text{RTT}(C, x) = F(C, x) \cdot \text{ITT}(C, x)
\]  

(8)

Now consider that \( \text{RTT} \) and \( \text{ITT} \) have the linear relationship which have the same value of \( C \) and \( x \). The value of \( F(C, x) \) can be calculated by the linear regression method. Figure 3 shows the example of \( \text{RTT} \) versus \( \text{ITT} \) where the data was selected from data pool of Hanshin Expressway according to the given condition. The number of samples and the given conditions are shown at the top left corner of Figure 3. Compared the proposed theoretical model and observed there is a constant term in \( \text{RTT} \) (Figure 3), the constant term can be considered as the model error, and we can identify that the constant term is very small; this can be observed from Figure 3.
Congestion is not always as described in Figure 2, it may happen at the entrance of the highway segment at $t$ moment as shown in Figure 4. At this time, $y$ will be used instead of $x$, where $y$ is the distance between the entrance of highway segment and the head of congestion correspond to $t$ (See section 3.2 for more details).

2.2 Varying-coefficient Linear Model

The varying coefficient linear model has been proposed by Hastie and Tibshirani (1993), in the model both the slope and the intercept of the regression line vary with the appropriate value. Varying coefficient linear model had been applied for predicting travel time (Zhang and Rice, 2003) that “time of day” was chosen as varying coefficient. However, the regression line should vary with the value of $C$ and $x$ according to the analysis in the section 2.1. Consider a linear model based on the analysis in section 2.1:

$$RTT(C, x) = \beta_0(C, x) + \beta_1(C, x) \cdot ITT(C, x) + \varepsilon$$

(9)

Where $\beta_0(C, x)$ and $\beta_1(C, x)$ are the parameters of the linear model and the parameters of linear model vary with the value of $C$ and $x$. This kind of linear model is so-called varying-coefficient linear model. Because of the variance of the linear model is not equal, the weighted least squares (WLS) to estimate the parameters $\beta_0$, $\beta_1$ is used:

$$\text{Minimizing } Z = \sum_{i=1}^{n} \left[ RTT_i - \beta_0(C, x_i) - \beta_1(C, x_i) \cdot ITT_i \right]^2 \cdot w_i$$

(10)

Where $i$ is the identification number of the data saved in the database, and $n$ is the number of samples saved in the database. $RTT_i$ and $ITT_i$ is the value of $RTT$ and $ITT$ correspond to the same identification number $i$ respectively. $C$ and $x$ are the value of $C$ and $x$ correspond to the present moment respectively. $w_i$ is a weight function. The weight function is defined as:

$$w_i = \begin{cases} 
0 & \text{if } C_i \not\in [C - \Delta C, C + \Delta C] \text{ or } x_i \neq x; \\
\frac{1}{\sigma_i^2} & \text{if } C_i \in [C - \Delta C, C + \Delta C] \text{ and } x_i = x; 
\end{cases}$$

(11)

Where $C_i$ and $x_i$ are the values of $C$ and $x$ of sample $i$ respectively. $\sigma_i^2$ is the variance of the $\varepsilon$ (equation 9). In practice the values of $\sigma_i^2$ are not known and must be estimated. For estimating the $\sigma_i^2$, the regression of $RTT$ on the $ITT_i$ with the general least
squares and the residuals \( e_i \) is obtained, then \( e_i^2 \) is an estimator of \( \sigma_i^2 \). Then calculate \( w_i \) by equation (11), \( \beta_0 \), \( \beta_i \) can be estimated by equation (10).

3. IMPLEMENTATION

3.1 Data Structure
The historical data was managed by a relational database system. The data structure in the database is \{time, ITT, RTT, C, x, y\}. time is the moment of traffic condition detected by the detectors; ITT and RTT are the instantaneous travel time and the real travel time correspond to time respectively, ITT can be directly calculated (Figure 1) and RTT can be calculated by the trajectory described in Figure 1. The value of C can be calculated by the equation (3). As shown in Figure 5, x can be denoted by an ID of the detector which is the first detector from the entrance of highway segment and the detected speed is lower than 50km/h. Similarly, y can be denoted by an ID of the detector which is the last detector that the detected speed is lower than 50km/h from the entrance of highway segment.

![Figure 5 The value of x and y](image)

3.2 Processes
The method was implemented in a C program. C, x and y correspond to prediction time be inputted into the program, where x and y are the ID of the first and last detector which the detected speed is lower than 50km/h respectively. At the first step, the appropriate data are selected from the database according to the given condition: \( C_i \in [C - \Delta C, C + \Delta C] \) and \( x_i = x \). Here \( \Delta C \) is set to 0.05. If \( x = 1 \), select the data by \( C_i \in [C - \Delta C, C + \Delta C] \) and \( y_i = y \). Second, parameters of linear model \( \beta_0(C, x) \) and \( \beta_i(C, x) \) be estimated using the weighted least squares for the selected data. At last, the prediction of RTT be calculated. The process time is about 2s. Of course calculate time would increase with the size of data in database, but there is 5 min interval between each calculation loop, so that the proposed method is capable of predicting RTT on line.

4. CASE STUDY

4.1 Study Site
The method was carried out on Hanshin expressway which between Osaka city and Kobe city in western Japan. The study segment was 6 km long. The study segment was equipped with
12 pairs supersonic detector stations spaced approximately 0.5 kilometers. The information of volume, occupancy and speed at every 5 min interval from this monitoring system is collected by Hansin Expressway control center (Figure 6).

![Diagram of Hansin Expressway monitoring system]

**Figure 6 Date collection and the information display of Hansin Expressway**

### 4.2 Data Pool
The study was based on the data collected during the year 2003 year to 2006. The data was from 9:00 to 22:00 on every day, throughout the week. We use data from period 2003-3-2 to 2006-05-31 as the study data pool with the exclusion of 8 days’ data due to technical problem of the detectors. All the data be saved at a relational database and the data structure is same as the description of section 3.1. The outline of the study data pool is shown in Table 1. The general travel time of data pool is 0.08 hour. The presented study segment does not exist obvious congestion period, however we identified the traffic congestion between 13:00 to 15:00 hours is serious than other period of the test data of section 4.3, presented by Table 3.

<table>
<thead>
<tr>
<th>Time</th>
<th>Average of ( RTT )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00-11:00</td>
<td>0.10 hour</td>
</tr>
<tr>
<td>11:00-13:00</td>
<td>0.08 hour</td>
</tr>
<tr>
<td>13:00-15:00</td>
<td>0.08 hour</td>
</tr>
<tr>
<td>15:00-17:00</td>
<td>0.08 hour</td>
</tr>
<tr>
<td>17:00-19:00</td>
<td>0.08 hour</td>
</tr>
<tr>
<td>&gt;19:00</td>
<td>0.08 hour</td>
</tr>
</tbody>
</table>

### 4.3 Result
Predicting travel time under congestion traffic condition is very difficult; this problem always troubles traffic system managers. In this article, the case study just focuses on the key problem: predicting travel time under congestion traffic condition. In this study, it is not considered to predict travel time under extremely abnormal traffic condition (for bad weather or accident). A test data set (Table 2, Table 3) come from study data pool had been selected for evaluation and then the test data was deleted from the data pool. The traffic condition is considered under extremely abnormal condition if \( ITT > 0.5 \).

<table>
<thead>
<tr>
<th>Year</th>
<th>Selected condition</th>
<th>Number of the test data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>0.2 &lt; ( ITT ) &lt; 0.5</td>
<td>182</td>
</tr>
</tbody>
</table>
We use Mean Absolute Percentage Error (MAPE) and Error Decrease Rate (EDR) as the quantify standards to evaluate our prediction model. Let define MAPE of the proposed method as:

\[
MAPE_{\hat{RTT}} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\hat{RTT} - RTT}{RTT} \right) \cdot 100\%
\]  

(12)

Where \( \hat{RTT} \) is the prediction result of model and \( n \) is the number of the test data. Let MAPE of \( ITT \) as:

\[
MAPE_{ITT} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{ITT - RTT}{RTT} \right) \cdot 100\%
\]  

(13)

And let Error Decrease Rate (EDR) as:

\[
EDR = \frac{MAPE_{ITT} - MAPE_{\hat{RTT}}}{MAPE_{ITT}} \cdot 100\%
\]  

(14)

The evaluation result is shown in Table 3:

<table>
<thead>
<tr>
<th>MAPE_{\hat{RTT}}</th>
<th>MAPE_{ITT}</th>
<th>EDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.4%</td>
<td>25.2%</td>
<td>26.7%</td>
</tr>
</tbody>
</table>

The evaluation result is divided into the different groups according to the time of day and the value of \( ITT \) for making more detail evaluation of the proposed method.
Figure 7 exhibits the MAPE of \( I'TT \) and the MAPE of the proposed method with different time segment. Generally, the MAPE of the proposed method is lower and stable (about 25%), which does not increase with the increment of the MAPE of \( I'TT \). Sometimes, the performance of \( I'TT \) is better when the congestion is not heavy (for example the MAPE of \( I'TT \) is lower than the MAPE of the proposed method of time segment 11:00 to 13:00). However, we cannot say that the performance of \( I'TT \) is usually better than the proposed method when the congestion is not heavy, for example, the congestion of the time segment 9:00 to 11:00 is not heavy (see Table 3) and also we identified that the MAPE of \( I'TT \) is obviously higher than the MAPE of the proposed method.

Figure 8 depicts the MAPE of \( I'TT \) and the MAPE of the proposed method with the different value of \( I'TT \). The MAPE of \( I'TT \) obviously increase with the increment of \( I'TT \), or say that the error obviously increase where the traffic is under very heavy congestion. On the other hand, the MAPE of the proposed method will not be influenced by the degree of traffic congestion and keeps stable whether high or low of \( I'TT \).
As mentioned in section 3.2, for fitting the linear model, some sample data are selected from the database. The sample size is a major factor which influences the accuracy of prediction, so that it is important to understand the behavior of the sample size. Figure 9 shows the number of the sample of the linear model. There are about 150 tests which the number of samples is more than 25 (The total number of test is 182).

The parameters of linear model are shown in Figure 10. Here the focus is on the slope of the regression line (the intercepts are very small). When the value of $C$ is less than 1.2, the slope of regression line will reduce with the decrease in the value of $C$. However, the obvious trend of the slope can not be observed when the value of $C$ is bigger than 1.2.

5. CONCLUSIONS

The main purpose of this article was to describe a method of travel time prediction. We tested the proposed method using a data set which is under heavy congestion situation. The result shows that the method can provide the stable and acceptable prediction under heavy congestion condition. Comparing with ITT (Now, which is provided to the users of highway
as the information of travel time.), the proposed method is not only more accurate but also more stable that the prediction error of the method does not increase with the increment of the degree of traffic congestion.

The purposed method is not only simple in pursuing the accuracy but also consider the implementation issue. The method can be easily applied to the other place also. The relationship of input data (independent variable) and output data (dependent variable) is explicit, all the variables and parameters have the clear physical meaning, so that, the highway managers can well communicate with the proposed method, they can improve the method according to their operation experiment.

ACKNOWLEDGEMENTS

The authors would like to thank Hanshin Expressway Corporation for providing the data used in this research.

REFERENCES


