Travel Time Prediction for Pre-trip Information Using Latest Traffic Conditions on Expressway

Hiroshi WARITA, Hirohisa MORITA, Masao KUWAHARA, Edward CHUNG, Atushi TANAKA

Metropolitan Expressway Public Corporation

(43-5, Nihonbashi-Hakozakicho, Chuo-ku, Tokyo 103-0015, JAPAN, +81-3-5640-4857, warita@mex.go.jp)

University Research Center, Nihon University

(7-24-1, Narashinodai, Funabashi City, Chiba 274-0063, JAPAN, +81-47-469-5504, hi-morita@irpt.cst.nihon-u.ac.jp)

Center for Collaborative Research, University of Tokyo

(4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, JAPAN, +81-3-5452-6418, kuwahara@iis.u-tokyo.ac.jp)

Center for Collaborative Research, University of Tokyo

(4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, JAPAN, +81-3-5452-6529, edward@iis.u-tokyo.ac.jp)

Oriental Consultants Co.,Ltd

(3-5-7 Hisamoto, Takatsu-ku, Kawasaki City, Kanagawa, 213-0011, JAPAN, +81-44-812-8813, tanaka-at@oriconsul.co.jp)

This study evaluates a method of providing pre-trip information that predicts travel time by matching past accumulated traffic data with traffic data on the day of the trip while taking into account the time required to reach the Metropolitan Expressway. Since pre-trip information is available at the time of access to the information media, it is highly likely that the congestion conditions have changed by the time people actually drive on the Metropolitan Expressway. Predicted future travel time is also very useful for drivers to decide their trip departure times. Therefore, it is necessary to provide pre-trip information while predicting traffic fluctuations. The proposed method is verified using observed traffic data on two major radial routes and it is concluded that estimated travel time is accurate enough for the practical use.

Keywords: Traffic information, Travel time prediction, Forecast method, pre-trip

1. Introduction

Due to the increasing use of the Intelligent Transport Systems (ITS), a wide range of road traffic information, including information on congestion and travel time, is provided on the Metropolitan Expressway through diverse media, such as traffic information boards, Internet and mobile communication terminals. These types of information can be divided into three major categories based on how people use such information: long-term information used several days prior to departure when travel plans are drawn up; medium-term information used just before leaving and short-term information used while driving. Various studies are conducted based on these categories. For people who want to plan trips, the Metropolitan Expressway Public Corporation (MEPC) and the Japan Highway Public Corporation provide statistical data via the Internet [1]-[3]. People can access pre-trip information via the Internet and other media [4]. While people are driving, traffic information boards and other media provide current information on congestion and travel time, but the accuracy of the information on the propagation of the congestion is a problem [5]. There have been several researches on traffic jam prediction methods [6]-[16].

Among the types of information mentioned above, pre-trip information on congestion and travel time is provided via the Internet, mobile phones and other media, but since the information provided is that available at the time of access, it is highly likely that the congestion conditions have changed by the time people actually drive on the Metropolitan Expressway. Therefore, it is necessary to provide pre-trip information that accounts for traffic fluctuations. However, although a broader range of information is being provided and the information providers and tools are becoming more diverse, the content of information remains unchanged. There is a need to improve the quality of information. In a survey conducted by MEPC [17], over 90% participants replied that they needed pre-trip information, demonstrating the high demand for such information.

This study evaluates a method of providing pre-trip information that predicts congestion and travel time under ordinary conditions by matching past accumulated
traffic data with traffic data on the day of the trip while
taking into account of the time required to reach the
Metropolitan Expressway.

2. Outline of the Prediction Method

2.1. Basic Concept

In order to predict travel time, the authors decided to
use time-series fluctuations in travel time (hereafter
referred to as “time-series prediction”). Various time-
series prediction methods have been proposed, including
Kalman filtering and other self learning techniques [9];
method for searching similar patterns based on data on
cumulative traffic volume and occupancy obtained from
vehicle sensors, ETC data, uplink data and others [10]-
[13]; and method of estimating travel time using probe
data and other kinds of data.

In this study, the method of searching for similar
patterns was chosen because it is relatively easy, there is
a lot of high-precision data for the Metropolitan
Expressway, and the Expressway is chronically
congested every day, offering a large volume of similar
patterns. The authors thus studied a prediction method
that can be used one to two hours before departure rather
than short-term predictions that is more suitable when
driving.

In the past, many versions of similar pattern method
have been proposed. The differences between those
versions and the method proposed in this article are as
follows:

a) The method proposed in this article is designed to
provide pre-trip information. For this reason, it takes
into consideration use of the Metropolitan
Expressway two hours ahead of time, which means
that it provides predictions for a longer period of
time than other methods. Previous methods provided
predictions for the period of time required to reach
the Expressway.

b) Our method covers the Metropolitan Expressway,
where accidents, construction work and other
irregularities occur frequently. (For this reason, the
authors came up with better prediction methods,
including referring to the top ten patterns.)

c) Our method uses data accumulated over the past
everal years.


d) Our method analyzes multiple routes while most of
the previous methods analyzed only one route.

e) Our method considers the possibility of providing
practical information on travel time for the
Metropolitan Expressway.

Figure 1 show the basic concept of the method, which
takes into account fluctuations in travel time, for example.
As the figure shows, the method takes into consideration
the fact that fluctuations in traffic conditions in the
immediate past affect traffic conditions in the near future,
and predicts traffic conditions by matching past traffic
data with traffic data of the day for which the prediction
is made.

![Figure 1: Basic Concept of the Prediction Method](image)

2.2. Timing for Providing Predictive Values.

In the monitor survey [17], nearly half the
respondents stated that they would like to have pre-trip
information one hour before they used the Metropolitan
Expressway. The Origin Destination Surveys [18] of the
Metropolitan Expressway show, meanwhile, that the
average distance covered to reach the Metropolitan
Expressway (distance between the nearest ramp and the
center of a given zone) is 21 km. If it is assumed that half
of this represents the distance to the ramp and users drive
on surface streets at an average speed of 20 km/h, the
time required for driving on surface streets is about 30
minutes. Based on these findings, it is assumed that
drivers can easily reach the nearest entrance to the
Metropolitan Expressway if they have one hour. The
authors decided to add an additional hour, to account for
a possible change in departure time, and provide pre-trip
information from two hours before.

2.3. Location Studied and Data Used

In examining the prediction method, the inbound
Yoga-Tanimachi section of Route No. 3 Shibuya Line
was used (see Figure 2). The following data was used:
- Period: April 1, 2000 to March 31, 2002 (two years)
- Content of data: Five-minute data for each 500- to
1,000-meter section (survey data: traffic volume,
vehicle speed, occupancy, calculated data: travel time)
2.4. Evaluation of prediction methods

In order to determine a prediction method, it is necessary to consider the data used for matching, matching methods and the method of calculating predictive values using the matching results. The following section shows the results of such evaluation.

2.4.1. Data used for matching. In order to determine the most suitable data when matching past accumulated data with data for the day of the prediction, the authors compared the travel time for the Yoga-Tanimachi section, as well as the travel time, occupancy and vehicle speed for each section. As a result, since it is assumed that information on both congestion and travel time will be provided, matching using section data is more appropriate. Occupancy takes into consideration too many detailed fluctuations, such as the density or sparsity of traffic as well as the existence of large vehicles in traffic, and travel time enables more appropriate matching than vehicle speed because the former involves the concept of distance. For these reasons, the authors decided to use section travel time.

2.4.2. Matching. Matching involves calculating cumulative squared errors for the time required for traveling the same section at the same time of each day using past accumulated data and data for the day of the prediction. Days showing smaller errors are considered to show a similar tendency.

Accumulated data for a long period make it possible to perform searches using a large volume of data, but such data tend to respond slowly to changes in traffic conditions due to new routes and other factors. For this reason, two-year data are used in this study. There is, however, room for further research on the relationship between accumulated data periods and data accuracy levels.

With respect to the spatial range of matching, since Route No. 3 is affected by congestion ahead on the Inner Circular Route, the effects of such congestion were taken into account in comparing the range of the adjacent Inner Circular Route to be covered in the study (plus one or two junctions) in addition to the 17 sections of the Yoga-Tanimachi section. As a result, if the two junctions on the Inner Circular Route are included, the Yoga-Tanimachi section data are overly affected by traffic fluctuation data on the loop, resulting in extraction of days that do not show a similar tendency to Route No. 3, the prediction target in this study. Therefore, the authors decided to cover only one junction on the Inner Circular Route: the Ichinohashi junction for the inner lanes (two sections) and the Miyakezaka junction for the outer lanes (two sections) (see Figure 3).

2.4.3. Method for Calculating Predictive Values. Based on the results obtained in section 2.4.2, the authors extracted the section travel times for the top 10 days that showed similar tendencies. The main reasons we use the median for the top 10 days rather than for the most similar day are: the data for the top 10 days may include accidents and other unique data; data for the top
20 days would be leveled too much; and there is variation even in the actually extracted data for the top 10 days. Figure 5 compares the accuracy levels for predicted travel time due to differences in the number of days used. As the figure shows, the data for the top 10 days indicate a stable accuracy level.

Then, for the top 10 days obtained after outlier were eliminated, the median for five-minute travel times for each section and time period was calculated as predicted section travel time, and the total predicted travel time for the Yoga-Tanimachi section was calculated by applying and adding up predicted section travel times according to the passage of time (time slice). Five-minute travel time is used here because that is the minimum unit for preparing information on travel time on the Metropolitan Expressway.

In addition, predicted section travel times and the length of the sections were used to calculate section vehicle speed, which was assumed to be the predictive value for congestion maps.

![Figure 5 Differences in Prediction Accuracy Levels Due to Differences in Number of Days Used](image)

2.5. Prediction Procedures

Figure 6 gives an overview of prediction procedures. This analysis is an extension of previous travel time study by Shamas [11] [12] in which the provision of travel time was mainly for drivers already en-route. This study however considers travelers who are about to depart from their origin.

1) The squared error is calculated from the section travel times for the same time and same time period, which are obtained from accumulated traffic data and traffic data for a given day. The $E_k$ employed here has been proposed by Shama [11] [12].

$$E_k = \sum_{i=1}^{n} \sum_{j=1}^{m} (TP_{ikj} - TT_{ij})^2$$

$TP_{ikj}$: Section travel time for section $i$, time $j$ and date $k$, which is obtained from accumulated traffic data

$TT_{ij}$: Section travel time for section $i$ and time $j$, which is obtained from traffic data for a given day

$n$: Total number of sections for which data matching is performed

$m$: Number of times (24) (five-minute pitch; two hours)

$k$: Data accumulation period (730) (two years)

2) Squared errors ($E_k$) are sorted and arranged on a daily basis with the smallest coming first.

3) Section travel times ($S_{ikj}$) for the time periods (two hours), which are used for prediction for the top 10 days, are extracted.

4) The median of the section travel time for each section and each five-minute period is used to calculate the predicted section travel time ($YS_{ikj}$).

5) The section travel time ($YS_{ikj}$) is predicted for the next 2 hours at five minute intervals. By estimating the piecewise linear trajectory based on the section travel time, the total travel time is predicted.

6) The predicted section speed ($YV_{ikj}$) for each five-minute period is calculated by dividing the section length ($L_{ik}$) by the predicted section travel time ($YS_{ikj}$). Predictive values for the congestion chart are created by color-coding speed levels as follows:

- 0-20 km/h - Red (Congestion)
- 20-40 km/h - Blue (Slight congestion)
- 40 km/h or higher - White (Free flow)

![Figure 6: Prediction Procedures](image)
3. Verification of the Prediction Method

3.1. Verification Method

The prediction method was verified using two routes: the Yoga-Tanimachi section of Route No. 3 (see Figure 2) and the Takaido-Miyakezaka section of Route No. 4 (see Figure 7). Verification was performed by randomly choosing three time periods (morning, afternoon and evening) on three weekdays in the months that have a high, average and low traffic volume. On these days, Route No. 3 was not particularly affected by accidents and other irregularities, but Route No. 4 was affected by accidents on October 16 and July 24, and therefore, the time periods when the accidents occurred were avoided when the verification was performed.

Figure 7: Route on Which Verification Was Performed (Route No. 4)

Travel time was used to quantitatively evaluate the verification results, and since the results of a survey[17] indicate that the largest percentage of users, over 40%, demand an accuracy level within ±10 minutes, irrespective of the time when the traffic information is provided, this was considered the standard value. The accuracy of errors within the range of ±5 minutes, and the accuracy of errors within the range of ±33% were used as reference values. The reason for using the latter is that the permissible level of errors for the travel time of 30 minutes is ±10 minutes.

Table 1: Days and Time Periods When Verification Was Performed

<table>
<thead>
<tr>
<th>Route</th>
<th>Date</th>
<th>Time period</th>
<th>Accidents and other irregularities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Morning</td>
<td>Afternoon</td>
</tr>
<tr>
<td>Route No. 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001/7/24 (Tue)</td>
<td>8:00</td>
<td>12:00</td>
<td>16:00</td>
</tr>
<tr>
<td>2001/10/16 (Tue)</td>
<td>8:00</td>
<td>12:00</td>
<td>16:00</td>
</tr>
<tr>
<td>2002/2/12 (Tue)</td>
<td>8:00</td>
<td>12:00</td>
<td>16:00</td>
</tr>
<tr>
<td>Route No. 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001/7/24 (Tue)</td>
<td>8:00</td>
<td>12:00</td>
<td>18:00</td>
</tr>
<tr>
<td>2001/10/16 (Tue)</td>
<td>8:00</td>
<td>10:00</td>
<td>18:00</td>
</tr>
<tr>
<td>2002/2/12 (Tue)</td>
<td>8:00</td>
<td>12:00</td>
<td>16:00</td>
</tr>
</tbody>
</table>

3.2. Results of Verification

Figures 8 and 9 compare the prediction results with actual measurements of congestion and fluctuations in travel time for Route No. 3 Shibuya Line on Tuesday, October 16 and for Route No. 4 Shinjuku Line on Tuesday, July 24. These figures indicate that except for 8 a.m. in Figure 9, actual measurements are almost consistent with predictions in terms of propagation of congestion and increased travel time. The accuracy level for Route No. 4 Shinjuku Line at 8 a.m. is low because, although precise matching was achieved for the period up to 8 a.m., congestion started to quickly spread from 8:30 a.m. due to the congestion on the Inner Circular Route ahead. This type of congestion rarely occurs on days when traffic is heavy and therefore what probably cannot be extracted. In the following sections, this result is excluded.

Table 2 shows the congestion and accuracy for predicted travel time in all cases. About 90% of predicted travel times achieved an accuracy of errors being within the range of ±10 minutes, and about 85% within the range of ±33%, suggesting that the predictions are almost consistent with actual measurements.

By time period, there is no large difference, and almost all time periods achieved an accuracy level of 80% or higher in terms of errors being within ±10 minutes and ±33% for travel time, indicating that for all time periods, the predictions are accurate. Particularly, the accuracy levels for the morning peak time on Route No. 3 Shibuya Line and for the afternoon on Route No. 4 Shinjuku Line are high, at 93% for travel time errors within ±5 minutes and 86% for travel time errors within ±5 minutes, respectively. The accuracy level for the morning peak time is high, probably because many of the users leave home at certain fixed times on weekday mornings. The high accuracy level for the afternoon on Route No. 4 Shinjuku Line is probably because the average travel time is about 20 minutes, indicating that the section is not so heavily congested.

By day, the accuracy level for Tuesday, July 24, when the average travel time was long, is lower than that for other days. It is inferred that this is because the other two days had an average traffic volume, enabling extraction of days that had a traffic volume similar to that for the time period in the immediate past, while July 24 had heavy traffic volume and there are few days that had similar traffic conditions. Under these circumstances, it is likely that the calculation of the median for the top 10 days results in average values, predicting less heavy congestion. This led the authors to predict congestion for the afternoon of July 24, when the largest error was observed for Route No. 3 Shibuya Line, using the median for the top five days, which resulted in accuracy level being raised from 63% to 88% in terms of travel time errors within the range of ±10 minutes.
Figure 8: Comparison of Congestion and Fluctuations in Travel Time for Route No. 3
Shibuya Line on Tuesday, October 16, a Day with Average Traffic Volume

Figure 9: Comparison of Congestion and Fluctuations in Travel Time for Route No. 4
Shinjuku Line on Tuesday, July 24, a Day with Heavy Traffic Volume
This method is effective for situations where traffic volume is high, and in the future, it is necessary to analyze a method of setting a threshold for squared sums and using those below the threshold. Based on the above findings, it can be said that although some problems remain to be addressed in the future, the overall accuracy levels are high and the applicability of the prediction method to inbound Route No. 3 Shibuya Line and Route No. 4 Shinjuku Line to Tokyo has been confirmed.

### Table 2: Accuracy Level for Predicted Travel Time

<table>
<thead>
<tr>
<th>Item</th>
<th>Day</th>
<th>Morning</th>
<th>Afternoon</th>
<th>Evening</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within ±10 minutes</td>
<td>2001/7/24(Tue)</td>
<td>88%</td>
<td>63%</td>
<td>75%</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td>2001/10/16(Tue)</td>
<td>100%</td>
<td>88%</td>
<td>100%</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>2002/2/12(Tue)</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Average</td>
<td>96%</td>
<td>83%</td>
<td>92%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>(Reference)</td>
<td>Within ±5 minutes</td>
<td>79%</td>
<td>29%</td>
<td>25%</td>
<td>44%</td>
</tr>
<tr>
<td></td>
<td>2001/10/16(Tue)</td>
<td>100%</td>
<td>54%</td>
<td>100%</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>2002/2/12(Tue)</td>
<td>100%</td>
<td>33%</td>
<td>83%</td>
<td>72%</td>
</tr>
<tr>
<td>Average</td>
<td>93%</td>
<td>39%</td>
<td>69%</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>(Reference)</td>
<td>±33%</td>
<td>89%</td>
<td>100%</td>
<td>92%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td>2001/10/16(Tue)</td>
<td>100%</td>
<td>58%</td>
<td>100%</td>
<td>86%</td>
</tr>
<tr>
<td></td>
<td>2002/2/12(Tue)</td>
<td>100%</td>
<td>50%</td>
<td>100%</td>
<td>83%</td>
</tr>
<tr>
<td>Average</td>
<td>96%</td>
<td>69%</td>
<td>97%</td>
<td>88%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3: Results of Prediction Including the Tomei Expressway

<table>
<thead>
<tr>
<th>Item</th>
<th>Upstream sections</th>
<th>6:00</th>
<th>8:00</th>
<th>10:00</th>
<th>12:00</th>
<th>16:00</th>
<th>18:00</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within ±10 minutes</td>
<td>Excluding Including</td>
<td>100%</td>
<td>100%</td>
<td>92%</td>
<td>100%</td>
<td>96%</td>
<td>100%</td>
<td>98%</td>
</tr>
<tr>
<td>(Reference)</td>
<td>Excluding Including</td>
<td>100%</td>
<td>100%</td>
<td>79%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>97%</td>
</tr>
<tr>
<td>Within ±5 minutes</td>
<td>Excluding Including</td>
<td>100%</td>
<td>83%</td>
<td>39%</td>
<td>100%</td>
<td>88%</td>
<td>83%</td>
<td>81%</td>
</tr>
<tr>
<td>(Reference)</td>
<td>Excluding Including</td>
<td>100%</td>
<td>92%</td>
<td>39%</td>
<td>67%</td>
<td>75%</td>
<td>58%</td>
<td>70%</td>
</tr>
</tbody>
</table>

4. Conclusion

This study proposes a prediction method of travel time for pre-trip information provision on the Metropolitan expressway using the pattern matching technique. The method is verified using observed data on Routes 3 and 4, and it is confirmed that the accuracy of the predicted congestion and travel time is adequate to provide useful pre-trip information. The predicted travel time is also quite useful for departure time choices and other trip arrangements.

In the future, in order to put the prediction method explained in this article to practical use and improve prediction accuracy for days when congestion occurs downstream or traffic is heavy or for other conditions, the authors plan to expand the scope of application, including weekends and other routes, expand the scope of data matching to include upstream sections, consider the most suitable data accumulation periods and classifications of data into patterns by taking into account new routes and the characteristics of the day of the week.
Acknowledgment
The authors would like to express their gratitude to the Japan Highway Public Corporation and Mr. Yoriyuki Sunaga of the Metropolitan Expressway Public Corporation, who provided valuable data for the analyses. Finally, our study group would like to express sincere thanks to Mr. Shamas ul Islam Bajwa, University of Tokyo for his valuable suggestions.

Bibliography
[4] Japan Road Traffic Information Center’s Website at http://www.jartic.or.jp
Hiroshi WARITA has been with the Metropolitan Expressway Public Corporation since 1991 and is involved in the construction and management of the Metropolitan Expressway. He graduated with a degree in civil engineering from National GUNMA Industry Technical College.

Hiroisa MORITA serves as managing director of Japan Transportation Consultants, Inc. and is also professor of traffic engineering at the University Research Center, Nihon University. After graduating from the University of Tokyo with a bachelor's degree in civil engineering in 1967, he joined the Metropolitan Expressway Public Corporation, where he remained until 1999.

Masao KUWAHARA is Professor of Traffic Engineering at the Institute of Industrial Science, University of Tokyo. After obtaining a Ph.D from University of California, Berkeley, he started teaching and researching at the Department of Civil Engineering, University of Tokyo. He transferred to the institute as Associate Professor in 1987 and became Professor in 2000.

Edward CHUNG is a visiting Professor at the Centre for Collaborative Research, University of Tokyo. He received his bachelor's degree in civil engineering and Ph.D in traffic engineering from Monash University.

Atsushi TANAKA works at the Tokyo headquarters of Oriental Consultants Co., Ltd. and is engaged mainly in transport-related projects. He joined the company upon graduating from Nihon University College of Science and Technology in 1994.

Received: 7 May 2004
Revised: 20 August 2004
Accepted: 3 September 2004
Editor: Fumitaka Kurauchi