Travel Time Prediction on expressways using traffic detectors*

Shamas ul Islam Bajwa**, Edward Chung***, Masao Kuwahara****

1. Introduction

In recent years, ITS related research has been carried out all over the world. In Japan, a lot of research is performed in areas such as VICS and AHS. ITS research aims at increasing efficiency, safety and convenience of traffic networks. In traffic control systems, research on traffic flow prediction and travel time estimation has been carried out to improve efficiency and convenience. Travel time information has been identified as particularly important data. Studies have shown that drivers with knowledge of the travel time to reach the destination, found driving less stressful as they know what to expect ahead of their journey. Travel time information is a good indicator of traffic condition. Provision of accurate travel time information to expressway users not only help in spatial and temporal dispersal of traffic but also encourage users to try other modes of transportation. So far, different methods based on statistical[1)] and neural network[2),3)] models are proposed for travel time prediction on expressways using traffic detector data. Pattern matching technique is used for travel time prediction on intercity highways using magnetic ticket data[4),5)] and on arterial roads using AVI data[6)]. The aim of this research is to develop a model for travel time prediction on expressways. In the following sections, a methodology is proposed for travel time calculation using pattern matching technique on traffic detector data. The prediction time horizon is 0, as travel time prediction is zero minutes ahead of latest data.

2. Methodology

Ultrasonic detectors are used for flow, velocity and occupancy measurements on metropolitan expressways in Tokyo. The data from detectors is recorded in 5 minutes intervals. These traffic detectors are located at approximately 300 m apart on all routes of metropolitan expressways and provide measurements for their discrete locations. For the purpose of simplicity, these measurements are assumed as of certain section of road instead of discrete locations. The section of road extends halfway to next detector stations on upstream as well as downstream sides. Pattern matching technique is based on finding the closest matched traffic pattern in the historical database to the present traffic pattern and then to use this closest matched traffic pattern for travel time prediction. The process is as follows:

We had three types of data available, namely flow, speed and occupancy. Generally, the historical database of the prediction target(e.g. travel time) is used for pattern matching. But as time delay cannot be avoided for departure travel time measurements, so the latest travel time always lag by travel time of the latest completed trip. That is why, in the case of travel time prediction, traffic detector data was used for pattern matching. First of all, the traffic detector data was arranged into a matrix format on spatial and temporal scales. From preliminary studies it has been found that occupancy data is the best indicator of closest match patterns.

Occupancy distributions in one hour time window from prediction target time along the whole road length were used as representative of the present traffic condition. Occupancy patterns of all days in historical database within a time frame of ±1 hour of prediction target time were searched for closest patterns. The ±1 hour time frame was used as probability is low that traffic situations will recur exactly at the same time as they occurred before. At the same time, occupancy data on the road at the prediction target time is more important than occupancy data one hour before. Similarly, traffic entering the roadway is more important than traffic leaving the roadway. So, spatial and temporal weights were applied on the occupancy distributions data using following linear functions.

$$w_s(i) = 2 - \frac{i}{n_i}$$

where, $w_s(i)$ is spatial weight at *ith section*, i is the section number starting from upstream and n_i is total number of road sections along the road length. Similarly,

---

*Key words: Travel time prediction, Pattern matching.
**B.Sc., Graduate Student, Institute of Industrial Science,University of Tokyo. (4-6-1 Komaba, Meguro ku, Tokyo.Tel. 03-5452- 6001 ext. 58175, Fax. 03-5452-6420).
Email: shamas@nishi.iis.u-tokyo.ac.jp
***Ph.D., Visiting Professor, Center for Collaborative Research, University of Tokyo.
****Ph.D., Professor, Institute of Industrial Science, University of Tokyo.

$$w_t(j) = 2 - \frac{j}{n_j}$$

where, $w_t(j)$ is temporal weight at *jth time slot*, j is time slot number starting from latest time of time window (e.g. j=5 means, 25 minutes from start of time window, as data is in 5 minute resolution). After applying these weights, squared difference of prediction target time occupancy distribution to the historical occupancy distributions was used as criteria for finding closest matches from historical patterns.

The objective function for pattern matching is

$$Sq\_Diff(t,p,h,t_s) = \sum_{i=1}^{n_i}\sum_{j=1}^{n_j} w_s(i)\cdot w_t(j)\cdot \{O(i,j,p,o) - O(i,j,h,t_s)\}^2$$

where, $Sq\_Diff(t,p,h,t_s)$ *is* square difference of occupancy at prediction target time t between prediction target day p and occupancy of a time window starting from t_s at historical day h. i, j, n_i and n_j are same as for weights. $O(i,j,p,0)$ represent occupancy distribution for prediction target time on prediction target day. Here, p represent prediction target day and 0 represent relative location of start of data time window from prediction target time which is zero in this case. $O(i,j,h,t_s)$ represent occupancy distribution from historical database. Here, h represent the day in historical database, $h= 1,2,3,4,\ldots,n_h$, where, n_h represent total number of days in historical database and t_s is the relative location of the start of data time window from prediction target time. As historical database is searched in ±1 hour of the target prediction time and data is at 5 minute resolution so $+12 = t_s = -12$.

Since, minimum square difference is regarded as a closest matched i.e. most similar traffic condition. After calculating squared differences, the results are sorted in ascending order and five patterns with smallest squared difference i.e. most similar historical patterns to the prediction target time pattern are selected.

In the next step, travel times corresponding to start time of selected historical data time windows are extracted from database. These values are modified by the following rule.

$$T(t,p,h) = \frac{T(t',p)}{T(t',h)}\cdot T(t,h), \qquad h \in \Omega(t,p), \qquad t = t' + T(t',p).$$

Where, $T(t,h)$ is travel time from one of the best similar patterns. $T(t',p)$ represents the latest completed travel time on prediction target day. $T(t',h)$ represents the travel time on historical day h at departure time lagging from start of historical time window by $T(t',p)$. $h \hat{I} W(t,p)$ represents a set of days h with the best n_k similar patterns for target time t on target day p. Finally, the travel time corresponding to prediction target time is calculated as:

$$T(t,p) = \frac{\sum_{h\in\Omega(t,p)} T(t,p,h)}{n_k}.$$

where, $T(t,p)$ is predicted travel time of a vehicle departing from the entry point at time t on prediction target day p based on the travel time of day h.

The process mentioned above is executed for every prediction target time on prediction target day.

3. Application of Model

(1) Site Description

The site selected for the application of model is inbound section of route no. 3 of Tokyo metropolitan expressway, i.e. from Yoga to Tanimachi. The length of road is approximately 12km. Historical travel time record shows that travel time on this route varies from 9 minutes in free flow condition to 70 minutes in severe congestion.

(2) Data

There are 40 detectors on inbound section of route no. 3 of Tokyo metropolitan expressway, i.e. from Yoga to Tanimachi. For this research, detector data from November 99' to July 2000' was used as historical data. This forms a historical database of 274 days in total. Sometimes at some detector stations, data was found missing due to apparent malfunctioning of the detectors. Missing data was interpolated on spatial as well as on temporal scale. Three days were used as test days. The test days and corresponding test time zones are as follows:

I)	August 1,2000	(Tuesday)	01:00-24:00
II)	August 2,2000	(Wednesday)	01:00-24:00
III)	August 3, 2000	(Thursday)	01:00-24:00

(3) Results

Figure-1 shows the results of our method and for the purpose of comparison actual travel time on test days is also plotted. Figure (1-a),(2-a) and (3-a) show the actual travel time, smoothed travel time, predicted travel time and error between actual and predicted travel time on three test days. Figure (1-b),-1(2-b) and -1(3-b) show the correlation between the actual and predicted travel time.

From Figure-1, we can verify that the trend of the predicted travel time by our proposed method is similar to that of measured travel time and time lag is small. Statistical evaluations about the three predicted days are presented in Table 1. The mean absolute error, mean absolute percentage error and standard deviation of error, measure the distribution of error around the correct value.

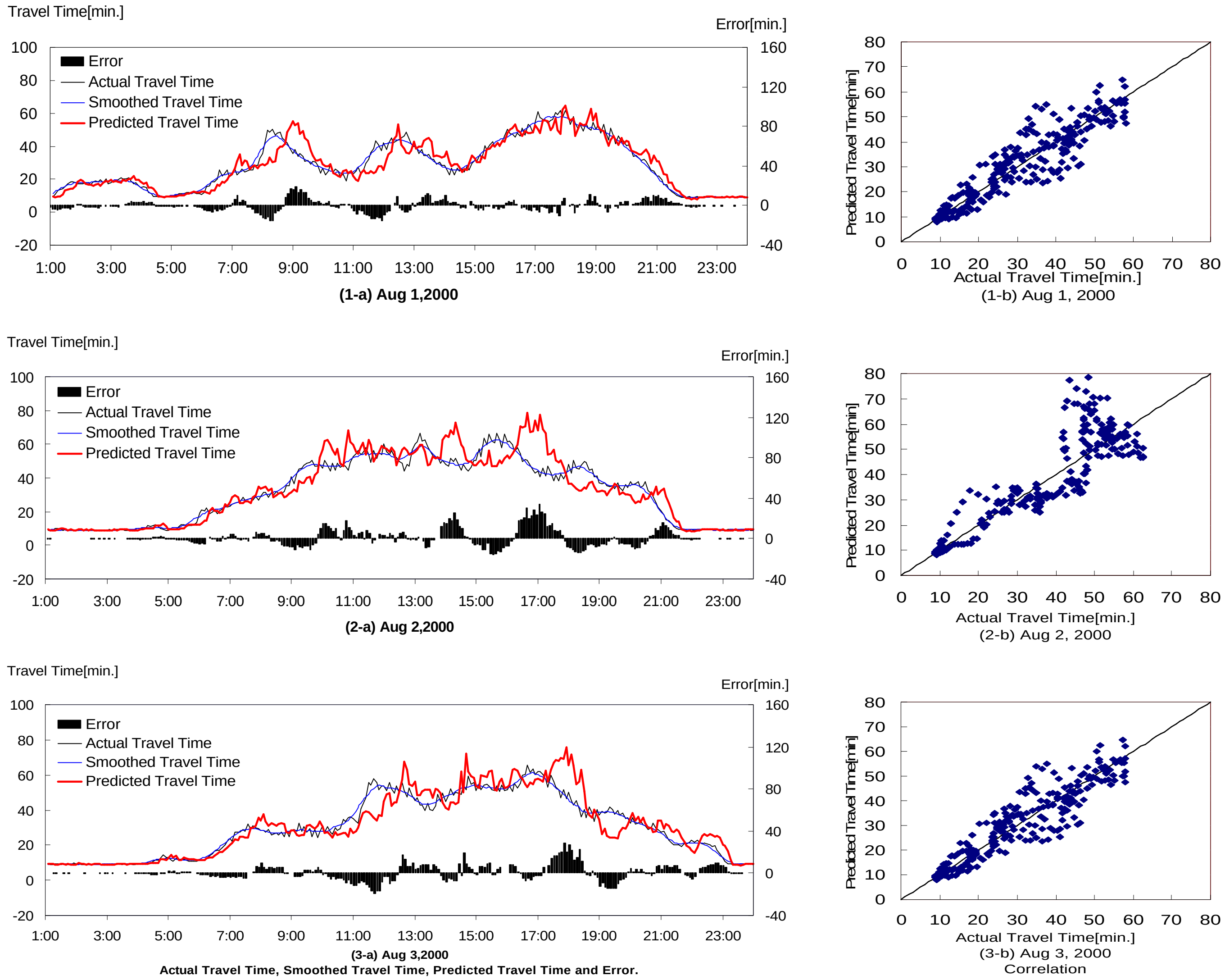

Figure-1 Results of the proposed method

Test Days	Mean Absolute Error[min.]	Mean Absolute Percentage Error[%]	Standard Deviation of Error[min.]	Correlation Coefficient
August 1,2000	3.90	13.69	3.92	0.931
August 2,2000	5.18	14.33	6.15	0.920
August 3,2000	4.79	14.23	5.17	0.923

Table-1 Statistical evaluation of predicted travel time

Test days	Percentage of predicted travel time within ±5 minutes of actual travel time	Percentage of predicted travel time within ±10 minutes of actual travel time	Percentage of predicted travel time greater than or less than 10 minutes of actual travel time
August 1,2000	73.19%	89.86%	10.14%
August 2,2000	62.28%	82.25%	17.75%
August 3,2000	60.14%	87.32%	12.68%

Table-2 Evaluation of predicted travel time from road users' point of view

The correlation coefficient, describe the relationship between the actual and predicted travel time. Along with statistical evaluation of the travel time, we have also analyzed the accuracy with respect to the road users' point of view. According to a previous study[7)] conducted on the Tokyo metropolitan expressway, 68% of users consider that errors of ±5 minutes between the travel time displayed on VMS and actual travel times are acceptable while 16% of users think that an error of ±10 minutes is allowable. Hence, keeping in view of, these findings, the accuracy of prediction results is also analyzed from road users' point of view and results are presented in Table-2. Results show that more than 82% of time, difference between actual and predicted travel time is less than 10 minutes.

4. Conclusions

In this paper, we have proposed a travel time prediction method applying pattern matching to traffic detector data. Moreover, the model was tested against the actual data obtained from field and comparison indicated that the proposed method yielded good results. Still, the results show some lag at the start and end of congestion. We have tried to search historical patterns for most similar pattern in ±1 hour time frame from prediction target time. It may be possible that no similar pattern existed in historical database in searched time frame. So, in the future, we may try to search in a wider time frame. Also, our method needs to search the database for a most similar pattern, which is a time consuming procedure. Hence, in order to apply the method in real time, we will further research on how to optimize pattern searching so that computational time can be minimized. It is obvious that the proposed method can be further refined and we are now studying it for further improvement.

At the same time, it will also be useful to study how such travel time information provided to users affect their behavior as the users may try to opt for different routes or different departure times in response to the information. It also needs to be studied that how this change of user behavior can be incorporated into the travel time prediction models.

5. Acknowledgements

The authors would like to thank Tokyo Metropolitan Expressway Public Corporation for providing traffic detector data.

6. References

1) Chung, E. and Rosalion, N. : Short Term Traffic Flow Prediction, Papers of the 24th Australasian Transport Research Forum(CD-ROM), Hobart, 2001.
2) Dia, H. : An object-oriented neural network approach to short-term traffic forecasting, European Journal of Operations Research, 131, pp. 253-261, 2001.
3) Innamaa, S. : Short term prediction of highway travel time using MLP-Neural Networks, 8th World Congress on Intelligent Transportation Systems (CD-ROM), Sydney, 2001.
4) Ohba, Y., Ueno, H. and Kuwahara, M. : Travel Time Prediction Method for Expressway using Toll Collection System Data, 7th World Congress on Intelligent Transport Systems, Turin, 2000.
5) Ohba, Y., Ueno, H. and Kuwahara, M. : Travel Time Information Calculation Method for Expressway using Toll Collection System Data, 6th World Congress on Intelligent Transport Systems, Toronto, 1999.
6) Otokita, T., Hashiba, K. and Oda, T. : Travel Time Prediction Based on Pattern Extraction from database, 5th World Congress on Intelligent Transport Systems, Seoul, 1998.
7) Warita, H., Okada, T. and Tanaka, A., : Evaluation of Operation for the Travel Time Information on the Metropolitan Expressway, 8th World Congress on Intelligent Transportation Systems (CD-ROM), Sydney, 2001.