**Proceedings of the 8th International
IEEE Conference on Intelligent Transportation Systems
Vienna, Austria, September 13-16, 2005**

Performance Evaluation of An Adaptive Travel Time Prediction Model

Shamas ul Islam Bajwa, Edward Chung, Masao Kuwahara

***Abstract*—This paper presents a travel time prediction model and evaluates its performance and transferability. Advanced Travelers Information Systems (ATIS) are gaining more and more importance, increasing the need for accurate, timely and useful information to the travelers. Travel time information quantifies the traffic condition in an easy to understand way for the users. The proposed travel time prediction model is based on an efficient use of nearest neighbor search. The model is calibrated for optimal performance using Genetic Algorithms. Results indicate better performance by using the proposed model than the presently used naïve model.**

***Index terms*—Intelligent Transportation Systems (ITS), Travel Time Prediction, Pattern Matching Technique, Adaptive Parameters**

I. INTRODUCTION

THIS paper presents a travel time prediction model. Travel time information helps the users of the transport system to make better travel decisions resulting in the efficient utilization of the available transport resources. Travel time information can be provided to the users either pre-trip or en-route. Pre-trip travel time information helps users to decide the optimal route as well as the departure time to reach their destination. En-route information can only help the users to decide a better route if alternatives are available. Even if the alternatives are not available and user cannot choose another less congested route, travel time information helps users by reducing the stress especially during the congestion as the users know before-hand how much time they will have to spend on the road.

Travelers' Information Systems can be either reactive or proactive. Reactive systems provide the users either the historical information or at the most information about the current traffic state, which may be outdated by the time travelers come on the road. On the other hand, the proactive systems provide the traffic state which the users will experience when they will use the system. A proactive system is necessary in order to build the confidence of the travelers in the provided information. A proactive system requires the prediction of the future traffic state using the present and past traffic state information.

Manuscript received March 1, 2005.

Shamas ul Islam Bajwa is with the Institute of Industrial Science, University of Tokyo, Tokyo, JAPAN. (phone: +81-3-5452-6419; fax: +81-3-5452-6420; e-mail: shamas@ iis.u-tokyo.ac.jp).

Edward Chung is Adjunct Professor, Institute of Industrial Science, University of Tokyo, Tokyo, JAPAN. (e-mail:edward@iis.u-tokyo.ac.jp).

Masao Kuwahara is Professor in the Institute of Industrial Science and Center for Collaborative Research, University of Tokyo, Tokyo, JAPAN(e-mail: kuwahara@ iis.u-tokyo.ac.jp).

Different types of data collection instruments can be used to assess the traffic state. These include fixed detectors which provide the traffic information at the specified locations. These detectors include loop detectors, ultrasonic detectors and video cameras. Another method is use of probe cars, which provide the link-based information. The proposed method uses detector data for the prediction of the travel time.

Many researchers have developed different travel time prediction methods based on statistical time series methods [1],[2],[3], Kalman Filtering [4],[5] and Artificial Neural Networks [6],[7],[8],[9],[10],[11] and [12]. Most of these methods have been tuned for the local traffic conditions for which these models have been developed and transferability of such models to other expressways is not evaluated. To test the transferability of the proposed model is one of the main objectives of this research.

II. TRAVEL TIME MODEL

The basic hypothesis of the proposed model is that traffic conditions are recurrent in nature and if present traffic state is defined in a way which properly relates to the travel time, similar traffic states can be searched in the history and can be used to extrapolate the present traffic state to predict the travel time. Fig. 1 shows the outline of the proposed model

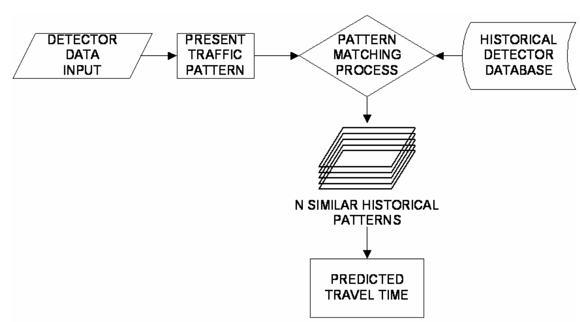

Fig. 1. Outline of the travel time prediction model

in which present traffic detector data is used to define present traffic pattern which is searched for similar instances in the historical database and n most similar patterns are selected from historical data to extrapolate the present traffic state and predict travel time.

A. Definition of Traffic Pattern

Traffic detectors can collect speed, flow and occupancy information. It is already established that inverse speed, best represents the relationship between traffic condition on the road and travel time [13]. It is assumed that point measurements from the detector stations represent the average traffic conditions on the segment of road from half of the distance to upstream detector to half of the distance to downstream detector. But, the detectors may not be equi-distant thus one detector may be representing traffic conditions for a longer length of road than other detector. This causes an unjustifiably equal weight to traffic conditions of different sections. To overcome this problem, inverse speed matrix is weighted by the distance which each corresponding detector represents.

The traffic pattern is defined as a matrix on spatial as well as temporal scale. On the spatial scale, traffic pattern includes the whole section of the road for which travel time needs to be predicted. On the temporal scale, it includes sufficient length of time to define the image of traffic reflecting the effect on travel time. Congested regions in the traffic considerably affect the travel time. To capture this effect, weights are applied to different values in the matrix based on the congestion level at given time and location. Speed is used as an indicator of congestion level at a given time and location and hence of weights.

B. Pattern Matching Process

The basic aim of the pattern matching procedure is to find the most similar historical pattern(s). Hence, the first task is to create some historical days' database. One way of searching the patterns is to search the whole historical database for the most similar pattern, but this makes the search process computationally intensive. Therefore, by making use of the assumption that traffic patterns are recurrent in nature and adding that these recur in a tight time frame such as similar traffic state for 9AM traffic can be searched from 8AM to 10AM. This restricted search window can quickly find a similar traffic pattern. Traffic patterns of all days in the historical database within a time frame of $\pm x$ minutes of prediction time were searched for closest patterns.

Sum of the squared difference between the prediction time traffic pattern and the historical traffic patterns is used as a criterion for finding similarities between the traffic patterns. The historical traffic pattern having minimum sum of squared difference is regarded as the most similar pattern.

The road section under study is assumed to consist of small links i's each representing a section of road equipped with one traffic detector , where $i=0,1,2,3,......n_i$ and similarly time period is divided into slots , $j=0, 5, 10,......n_j$ as data is collected in 5 minute intervals, here n_j is the length of the time window. In this way, the detector data is represented on a temporal and spatial scale. If t is prediction time on prediction day p, then $v\ (i,\ t\text{-}j,\ p)$ represents speed on prediction day p at link i at time t-j.

where,

$i=0$ refers to the most upstream link;

$i=n_i$ refers to the most downstream link;

$j=0$ refers to the time slot corresponding to start of pattern on temporal scale; and

$j=n_j$ refers to the time slot at the end of pattern on temporal scale.

Similarly, if $h=1,2,......n_h$ represents the number of days in historical database and t_s represents the start time of the traffic pattern on historical days then $v(\ i,\ t_s\text{-}j,\ h)$ represents speed on historical day h at link i at time t_s-j. As the search is performed in $\pm x$ minutes of prediction time t on historical days so $t+x \geq t_s \geq t\text{-}x$ and the final form of the objective function for pattern matching is,

$$\min.\ of \quad \Delta^2(p,t,h,t_s) = \sum_{i=0}^{n_i}\sum_{j=0}^{n_j} w(i,j).\frac{L(i)}{L}\left[\frac{1}{v(i,t-j,p)} - \frac{1}{v(i,t_s-j,h)}\right]^2 \quad (1)$$

where,

$w(i,j)$ represents the weight of the traffic in *cell* (i,j)

$L(i)$ represents the length of section i.

L represents the length of the stretch of road for which prediction is required

$\Delta^2(p,t,h,t_s)$ represents the squared difference between the present and historical pattern.

C. Travel Time Extraction and Cleansing

In this step, travel times, corresponding to selected historical traffic patterns, are extracted from database. It has been found that despite all the care exercised in the selection of most similar patterns; sometimes a few selected patterns have larger travel time differences from rest of the selected patterns. To overcome this problem, Box Plot technique is employed. By using Box Plot technique, outlier travel time values are excluded.

In Box Plot technique, the upper bound, *UB*, and lower bound, *LB*, of a data set is calculated using the following formulae.

$$UB = Upper\ quartile + 1.5 * Interquartile\ range$$

$$LB = Lower\ quartile - 1.5 * Interquartile\ range$$

where,

Lower and upper quartiles are the 25th and 75th percentiles, and

Interquartile range is the difference between upper and lower quartiles.

Any data point lying above *UB* or below *LB* is regarded as an outlier and is discarded. This technique helps to improve the prediction, especially when the travel time values are uncharacteristically different from other selected values.

Finally, after exclusion of the outliers, there are n_k patterns out of the historical database for every prediction time, t, on prediction day, p. The final prediction is calculated as:

$$\hat{T}(t,p) = \frac{\sum_{h \in \Omega(t,p)} T(t_s, h)}{n_k} \qquad (2)$$

Where, Ω *(t, p)* represents the set of patterns out of the *n* best matched patterns that are not the outliers; $T(t_s, h)$ is the travel time values extracted from these historical patterns; and $\hat{T}(t,p)$ represents the final prediction.

D. Model Parameters

Above described model has four parameters. These parameters are temporal size of traffic pattern, weight, size of search window and number of selected patterns from historical data.

Traffic conditions on the road changes during the day. In free flow condition the traffic state can be represented properly even with a small size of the traffic pattern and during congested traffic conditions, a bigger pattern size is needed for better representation and same is true for other parameters also. Hence a fixed set of parameters may not provide the optimum prediction and parameters should change with the change in traffic condition to better reflect the present traffic state while saving the extra computations where it can be saved. This can be achieved by using adaptive parameters which are a function of present traffic state. These parameters will self-adjust to the changes in traffic condition and provide optimal performance. Average speed on the whole road section is used to represent the present traffic condition and the parameters are calculated as a function of it.

1) Temporal Size of the Pattern

On the temporal scale, sufficient detector data in immediate history should be used. The temporal size of pattern should be long enough to replicate the evolution of the traffic state but should not be so long as to include the unnecessary details which can mislead the similar patterns search in historical database. Temporal size is assumed as a function of current average speed and the following relationship is proposed:

$$Pattern\_Size = nearest\,\mathrm{int}\left(\frac{A}{V_{av}}\right) * D.I \geq 10 \qquad (3)$$

where, A is a constant, V_{av} denotes the average speed on the road and $D.I$ represent the detectors' data collection interval which is 5 minutes in this case. A minimum value of 10 minutes is used.

2) Weights

Travel time on a road is mainly affected by the congestion present on the road. The congestion may occur either due to permanent bottlenecks, such as a sag section or sections with reduction in number of lanes etc., or temporary bottlenecks which can be further of two types, stationary for example if some incident has occurred or moving such as a slow moving vehicle forming a platoon. Weights are applied to account for the congestion produced due to these type of bottlenecks, whenever and wherever it occurs. These weights are based on the instantaneous speed of each section and should be higher for the sections with lower instantaneous speeds indicating congestion. Following functional form is used:

$$w(i,j) = \frac{1}{\left[v(i,t-j,p)\right]^B} \qquad (4)$$

where, B is a constant.

3) Search Window

The optimal size of the search window needs to be investigated in order to reduce the computational effort while maintaining a sufficient level of accuracy. Following functional form is proposed for this purpose:

$$Search\_Window\_Size = nearest\,\mathrm{int}\left(\frac{C}{V_{av}}\right) * D.I \geq 15 \qquad (5)$$

where, C is a constant and a minimum value of 15 minutes for the search window is fixed.

4) Number of Best Matched Patterns

Using only one best matched pattern for prediction can result in an abrupt change in travel time. To solve this problem, average of a larger number of patterns is used which helps to smooth the change in travel time from one instant to next. Following relationship is used for this purpose, which requires higher number of similar patterns searched for congested conditions than free flow conditions.

$$Number\_of\_selected\_patterns = \mathrm{int}\left(\frac{D}{V_{av}}\right) \qquad (6)$$

where, D is a constant.

The model was calibrated for optimal performance using Genetic Algorithms in a previous study [14]. The calibrated optimal values for the constants are as following:

TABLE I

OPTIMAL VALUES OF PARAMETRIC CONSTANTS

Constant	Optimal Value
A	40
B	0.25
C	180
D	200

This model was calibrated using the detector data from inbound direction of the Shibuya Line of Tokyo Metropolitan Expressway, which has a length of about 12 km and is equipped with 40 detectors. This section has high variability in travel time ranging from 10 minutes in free flow condition to 70 minutes or more during severe congestion.

III. MODEL APPLICATION

The basic objective of this study was to extensively test the proposed model for the network-wide applications. Two main issues are thoroughly investigated. These issues are:

1) Transferability
2) Effect of the length of route

Transferability of the proposed model while maintaining sufficient level of accuracy is a highly desirable property as it will provide a one-for-all solution eliminating the need of calibrating the model for each different site. The effect of the length of the route on accuracy of travel time prediction is also investigated in this study.

A. Site Description

The site selected for the application of the model includes different routes of Tokyo Metropolitan Expressway. Tokyo Metropolitan Expressway has a total length of 283.3 km and carries an average daily traffic volume of 1.12 million vehicles[15]. Fig. 2. shows the Tokyo Metropolitan Expressway Network. Data from following routes is used for the evaluation of the proposed travel time prediction model:

1) Shibuya Line(Route 3) 11.94 km
2) Shinjuku Line(Route 4) 13.50 km
3) Ikebukuro Line(Route 5) 21.44 km
4) Misato Line(Route 6) 09.52 km
5) Wangan Line(Bayshore route) 24.20 km

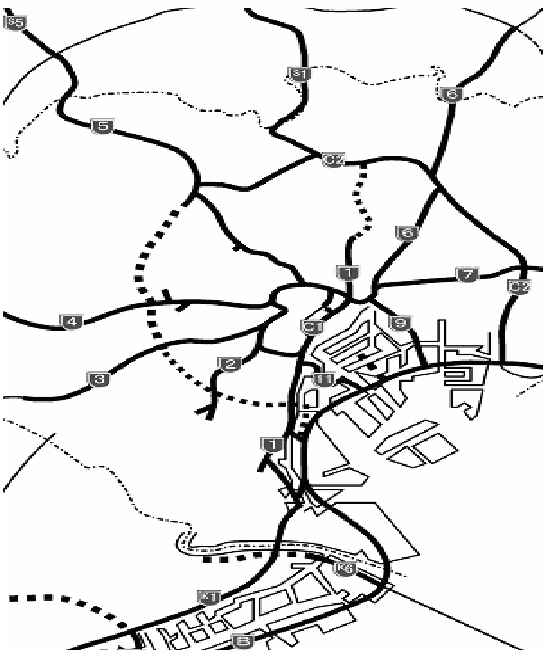

Fig. 2. Tokyo Metropolitan Expressway Network

B. Data

Following days are selected as test days for all the above-mentioned routes:

1) January 21, 2004(Wednesday)
2) January 22, 2004(Thursday)
3) January 23, 2004(Friday)

Data collected during the year 2003 is used as the historical database. Generally, the quality of data from the detectors is quite good but sometimes at some detector stations, data was found missing due to apparent malfunctioning of the detectors. In this case, missing data was interpolated on spatial as well as on temporal scale.

C. Travel Time Prediction Methods

Following travel time prediction methods are tested with the above data and their performance is compared.

1) Proposed Model

The model proposed in earlier sections is applied to predict the travel time for the given road sections on the above-mentioned days.

2) Instantaneous Travel Time Model

This model calculates the travel time based on the present information. This model provides good travel time prediction when traffic conditions are stationary, which is seldom true when the travel time information is most needed. This method assumes traffic conditions at present will prevail until the vehicle entering the road section now will reach the destination. Suppose a road section composed of links $i=0, 1, 2, 3, ...n_i$ where, each link is equipped with a traffic detector and speed data for each link is available. For a car that enters the road section at time T_o, total travel time is the summation of travel times of all the links based on the present speed. This can be represented by this simple formula.

$$TT(T_o) = \sum_{i=0}^{n_i} \frac{L(i)}{v(i, T_o)} \tag{7}$$

where, $TT(T_o)$ represents travel time of the vehicle entering at time T_o; $L(i)$ represents length of each link; and $v(i, T_o)$ represents speed in link i at time T_o. This method is currently used for travel time prediction on the test sites.

D. Performance Measures

To evaluate and compare the performance of the proposed travel time prediction method and currently employed methodology, correlation coefficient (R), root-mean-square-error(RMSE) and hit ratios E_p i.e. percentage of predictions within $\pm p$ percent of actual travel time are used as performance indices.

$$R = \frac{\sum_{l=1}^{n} (\hat{T}_l - \hat{T}_{mean})(T_l - T_{mean})}{n\sigma\hat{\sigma}}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{l=1}^{n}\left(\frac{T_l - \hat{T}_l}{T_l}\right)^2}$$

$$E_5 = \frac{\sum_{l=1}^{n} j_l}{n} \times 100 \quad where \quad j_l = 1 \ if \ \left[\frac{\left|\hat{T}_l - T_l\right|}{T_l}\right] \le 5, else \quad j_l = 0$$

$$E_{10} = \frac{\sum_{l=1}^{n} k_l}{n} \times 100 \quad where \quad k_l = 1 \ if \ \left[\frac{\left|\hat{T}_l - T_l\right|}{T_l}\right] \le 10, else \quad k_l = 0$$

where, T_l and $\hat{T}_l$ represent the actual travel time and predicted travel time at lth instant respectively. σ and $\hat{\sigma}$ represent the standard deviation for actual and predicted travel time respectively.

IV. Results

Travel time prediction results by using the proposed model are shown in Fig. 3. to Fig. 7 for the five different routes of the Tokyo Metropolitan Expressway. Results clearly indicate a good correlation between the predicted and actual travel time. The proposed model shows a better performance especially at the onset of congestion when traffic is still in transition state and travel time information is most needed. One shortcoming of this model is, it cannot predict travel time accurately during the non-recurrent congestion because of its dependence on historical traffic patterns. At some instants, especially during the congestion, the prediction is not good. The possible reason may be non-availability of a similar pattern from the historical database indicating non-recurrent congestion. This emphasizes the need to develop models which can be used to predict travel time during the non-recurring congestion such as during incidents.

The results in Table II show that for first four routes, the proposed travel time prediction model performs better than the instantaneous travel time model which is presently used for the travel time prediction, while on the last route the performance of proposed model is comparable. This is due to relatively low congestion levels for this route. The speed limit on first four routes in 60 km/hr and these routes have two lanes while the Wangan Line has a speed limit of 80km/hr and have three lanes. The lower level of congestion is also evident from the Fig. 7. which shows fewer variations in travel time as compared to the other routes. Lesser variations in travel time indicate a lower congestion level representing stationary traffic conditions, which is the basic assumption of the instantaneous travel time model and hence of its good performance in this case. Overall, it is evident from the results that on all those routes which experience severe congestion levels, proposed travel time prediction model performs better underlining the success of this model.

To investigate the effect of the length of the route on the travel time prediction using proposed model, we plotted the RMSE versus length of route, see Fig. 8. The results indicate that accuracy of model decreases with increase in length of road except for the Wangan Line. Though the increase in RMSE is not significant compared with an increase in length of the route but it needs to be investigated whether adjusting the parameters of model for change in length of route can improve the performance of the model.

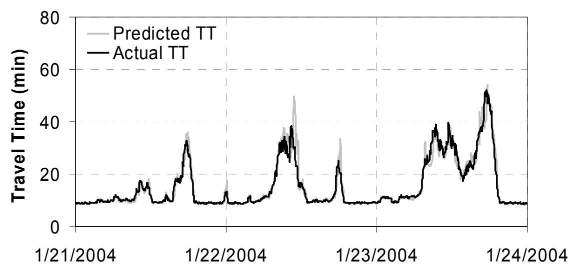

Fig. 3. Actual and predicted travel time profile for Shibuya Line

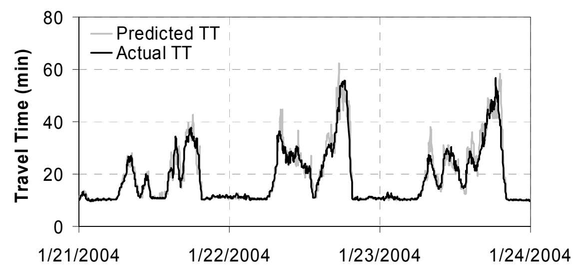

Fig. 4. Actual and predicted travel time profile for Shinjuku Line

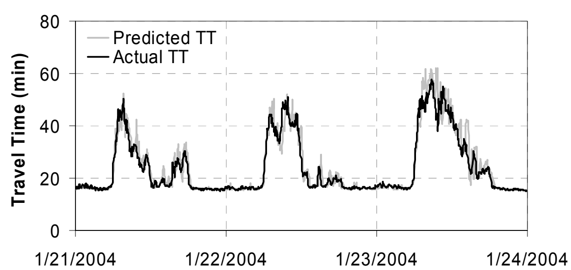

Fig. 5. Actual and predicted travel time profile for Ikebukuro Line

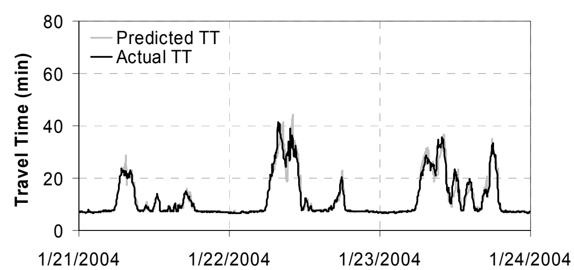

Fig. 6. Actual and predicted travel time profile for Misato Line

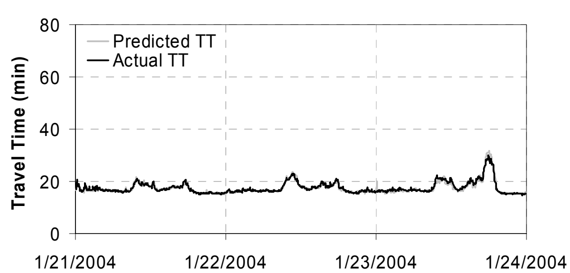

Fig. 7. Actual and predicted travel time profile for Wangan Line

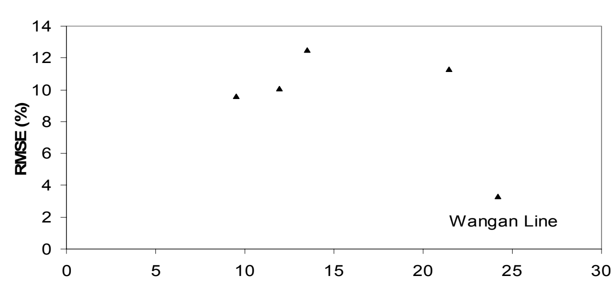

Fig. 8. Error of travel time prediction versus length of route

TABLE II
COMPARISON OF PERFORMANCE INDINCES FOR DIFFERENT ROUTES

Route	Prediction Method	Correlation Coefficient	RMSE	E_5	E_{10}
Shibuya Line	Proposed	0.965	10.1%	69%	84%
	Instantaneous	0.941	14.4%	62%	77%
Shinjuku Line	Proposed	0.954	12.5%	55%	73%
	Instantaneous	0.936	18.1%	52%	69%
Ikebukuro Line	Proposed	0.956	11.3%	59%	76%
	Instantaneous	0.920	14.6%	56%	73%
Misato Line	Proposed	0.967	9.6%	68%	81%
	Instantaneous	0.942	13.9%	61%	74%
Wangan Line	Proposed	0.963	3.3%	89%	99%
	Instantaneous	0.963	3.2%	90%	99%

V. CONCLUSION

Extensive testing of the proposed travel time prediction model clearly indicates its success in predicting travel time especially for the congested expressways. The high accuracy of the model on routes other than for which it is calibrated proves that the model is fully transferable to other routes without any significant loss of generality and accuracy. This gives this model a significant edge as it provides an off-the-shelf approach provided the required data is available.

ACKNOWLEDGMENT

The authors gratefully acknowledge the valuable discussions and insights contributed by Prof. H. Morita and other members of the Travel Time Research Group at Kuwahara Laboratory. The authors would also like to thank Tokyo Metropolitan Expressway Public Corporation for providing the data used in this research.

The valuable comments by three anonymous referees helped the authors in improving the paper.

REFERENCES

[1] K. F. Petty, et al. "Accurate estimation of travel times from single loop detectors," *Transportation Research Part A: Policy and Practice*, vol. 32(1), pp. 1-17, Jan. 1998.

[2] J. Rice, and E. V. Zwet, "A simple and effective method for predicting travel time on freeways." *IEEE Transactions on Intelligent Transpotation Systems,* vol. 5, pp. 200-207, Sep. 2004.

[3] J. Kwon, B. Coifman, and P. Bickel, "Day-to-Day Travel Time Trends and Travel Time Prediction from Loop Detector Data," *Transportation Research Record No. 1717*, pp. 120-129, 2002.

[4] S.I. Chien, X. Liu, and K. Ozbay, . "Predicting travel times for the South Jersey real-time motorist information system," in *Proc. of 82nd Annual TRB meeting*, Washington D.C., 2003, CD-ROM.

[5] M. Chen, and S. Chein, "Dynamic freeway travel time prediction using probe car data: link-based vs. path-based," in *Proc. of 80th Annual TRB meeting*, Washington D.C., 2001, CD-ROM.

[6] D. Park, et al. "Optimal aggregation internal size of probe-based travel time information size for ATIS," in *Proc. of 80th Annual TRB meeting*, Washington D.C., 2001, CD-ROM.

[7] D. Park, L. Rilett, and G. Han, "Spectral basis neural networks for real-time travel time forecasting," *Journal of Transportation Engineering*, vol. 125(6), pp. 515-523, Nov. 1999.

[8] D. Park, and L. R. Rilett, "Forecasting multiple-period freeway link travel times using modular neural networks," in *Proc. of 77th Annual TRB meeting,* Washington D.C., 1998, CD-ROM.

[9] G. Huisken, and E. C. Berkum, "A comparative analysis of short-range travel time prediction methods," in *Proc. of 82nd Annual TRB meeting*, Washington D.C., 2003, CD-ROM.

[10] H. Dia, "An object oriented neural network approach to short term traffic forecasting," *European Journal of Operations Research*, vol. 131(2), pp. 253-261, Jun. 2001

[11] J. W. C. Van Lint, S. P. Hoogendorn, and H. J. Zuelyn, "State space neural network for freeway travel time prediction," in *Proc. of the International Conference on Artificial Neural Networks*, Madrid, Spain. 2002

[12] S. Innamaa, "Short-term prediction of highway travel time using MLP-neural networks," in *Proc. of 8th World Congress on Intelligent Transportation System,* Sydney, 2001, (CD-ROM).

[13] S. I. Bajwa, E. Chung, and M. Kuwahara, "A travel time prediction method based on pattern matching technique," in *Proc. of 21st ARRB and 11th REAAA Conference,* Cairns, 2003 , (CD-ROM).

[14] S. I. Bajwa, E. Chung, and M. Kuwahara, "An adaptive travel time prediction model based on pattern matching," *in Proc. of 11th Intelligent Transport Systems World Congress,* Nagoya, 2004, (CD-ROM).

[15] Metropolitan Expressway Public Corporation(2004), "Metropolitan Expressway Public Corporation Guide" [online], Available: http://www2.mex.go.jp/profile/guide_english/index.html