Short-Term Travel Time Prediction Using the Kalman Filter Combined with a Variable Aggregation Interval Scheme

Jinhwan JANG

Advanced Transport Research Division, Korea Institute of Construction Technology, Goyang, 411-712, South Korea
E-mail: jhjang@kict.re.kr

Abstract: Data aggregation interval is important for reliable travel time predictions in probe-based systems. Where sufficient probes exist, a short interval can be used to minimize the time delay. However, in the opposite case, a short interval can cause unreliable travel time predictions due to small probes. Thus, the optimal aggregation interval may vary according to traffic flow conditions. This study suggests a methodology for selecting the optimal aggregation interval which varies according to a characteristic of probe travel time. The superiority of the proposed methodology compared to a conventional fixed interval is verified using DSRC probe data collected on a multilane highway near Seoul, Korea. The Kalman filter is adopted for a travel time prediction technique. As a consequence, the prediction accuracy is enhanced by approximately 40% compared to a fixed aggregation interval under free flow conditions.

Keywords: ITS, travel time prediction, variable aggregation interval, Kalman filter, probe, DSRC

1. INTRODUCTION

To maximize the efficiency of already-constructed highway facilities, Advanced Traveler Information System (ATIS) has become popular in many areas (FHWA and SwRI, 1998; Houston TranStar, 2003). Recently, 5.8 GHz Dedicated Short-Range Communication (DSRC) traffic information system, which produces section travel times using probes equipped with an On-Board Unit (OBU), has gained interest in Korea (ITS Korea *et al*., 2008). As of June 2012, the market penetration of this OBU is around 25%. Originally, the DSRC OBUs were introduced for an Electronic Toll Collection System (ETCS). Figure 1 illustrates the schematic outline of the DSRC traffic information system. Compared to conventional detector-based systems, the DSRC system has an advantage in that data collection and information provision is simultaneously possible through the same OBU.

However, a probe-based traffic information system inevitably produces delayed travel time (TT) information due to the widely recognized probe characteristic. Hence, TT prediction has been considered a paramount issue for probe-based systems. Chen and Chien (2001), and Chien *et al*. (2003) applied the Kalman filter to produce predicted TTs in a probe-based system, and these TTs were evaluated against the CORSIM simulation TTs as baseline values. Kuchipudi *et al*. and Chien (2003) also used the Kalman filter for reliable TT predictions with link- and path-based schemes for the New York State Thruway traffic information system. Detector data from a freeway in Rotterdam was used by Huisken *et al*. (2003) to predict TTs. They applied a neural network algorithm and verified the superiority of their methodology over existing techniques. Wei et al. (2007) developed a simple linear model to predict travel times gathered on Hanshin Expressway in Japan. The prediction performance, even though its simplicity, were proven to be satisfactory even in congested situations. Myung *et al*. (2011) predicted TTs on the basis of the *k* nearest neighbor (*k*-NN) method and used data provided by the vehicle detector system and the automatic toll collection system of a freeway in Korea.

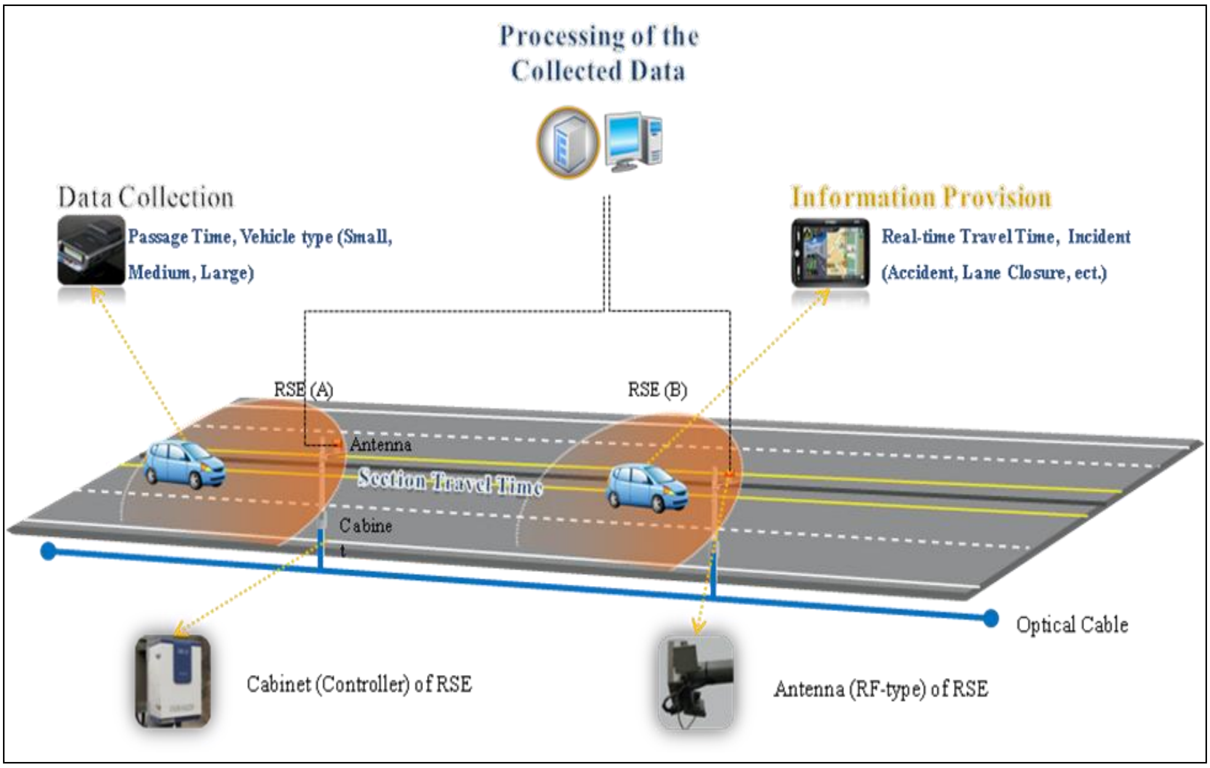


Figure 1. Schematic of DSRC Traffic Information System

Another important issue in probe-based systems is the optimal aggregation interval of probe data. That is, shorter aggregation intervals can diminish the time delay phenomenon, but the TT reliability can be deteriorated due to small sample size. On the contrary, longer aggregation intervals may induce the opposite effects. Therefore, the optimal aggregation interval for a specific traffic flow condition is of importance for reliable real-life traffic information. Some researches into the optimal aggregation interval for reliable TT information have been conducted. Gajewski *et al*. (2000) chose optimal aggregation widths to estimate speed data from loop detectors using a cross-validated mean square error (MSE) approach. Qiao *et al*. (2003, 2004) and Oh *et al*. (2005) applied various statistical methods for choosing optimal aggregation intervals, and Park *et al*. (2009) proposed a methodology based on MSE for identifying the optimal aggregation interval for estimating and forecasting TT. However, because these studies did not provide an algorithm for selecting optimal aggregation intervals on a real-time basis, their findings cannot be applied to a real-world TT prediction system that needs to dynamically select optimal aggregation intervals.

In this paper, the author proposes an algorithm to select the optimal aggregation interval on a real-time basis in a DSRC TT information system that applies the Kalman filter as a TT prediction technique. The optimal aggregation intervals are determined by the variance characteristic of the probe TTs. The predicted TTs and those based on conventional static aggregation intervals are compared with the actual TTs from vehicles that traveled the target link immediately after receiving the predicted TTs.

2. REVIEW OF METHODOLOGIES

2.1 The Kalman filter

The Kalman filter constantly updates its parameters to predict the required state variables (e.g., TT) as new state variables are obtained (Grewal and Andrews, 1993); therefore, it is applied to predict TT in this research. The Kalman filter operates in the following manner: let *x(t)* denote the TT to be predicted at time interval *t*. Let the parameter *ϕ(t)* denote the transition parameter at time interval *t*, which is calculated from current and historical TTs, and let *w(t)* represent a noise term that is normally distributed with a mean of zero and variance of *Q(t)*. The state equation (or system model) can be written as follows:

$$x(t) = \varphi(t-1) + x(t-1) + w(t-1) \tag{1}$$

Let *z(t)* denote the observed TT and *v(t)* denote the measurement error at time interval *t*; *v(t)* is normally distributed with a mean of zero and variance of *R(t)*. As no parameter except TT is considered, the observation equation related to the state variable *x(t)* can be written as follows:

$$z(t) = x(t) + v(t) \tag{2}$$

In this study, *z(t)*, the average TT between two successive Road Side Units (RSEs) at time interval *t*, is collected from the probes by the DSRC OBU. The data in the previous time interval is used to calculate the transition parameter *ϕ(t)*; this parameter indicates the relationship between the state variables (TT in this study) over successive time intervals. Let us suppose that for every *i* and *j*, *E[w(i)v(j)]* = 0, and let *P(t)* represent the covariance of estimation errors at time step *t*; consequently, the Kalman filter can be applied using the procedure given below. Generally, in a linear system, the value of *I* is set to 1 (Kuchipudi and Chien, 2003); we follow this convention here. The Kalman filtering algorithm is applied as follows:

Step 1 Initialization
Let $t = 0$ and let $E[x(0)] = \hat{x}(0)$ and $E[x(0) - \hat{x}(0)^2] = P(0)$
Here, $\hat{x}(0)$ is the predicted TT at $t = 0$

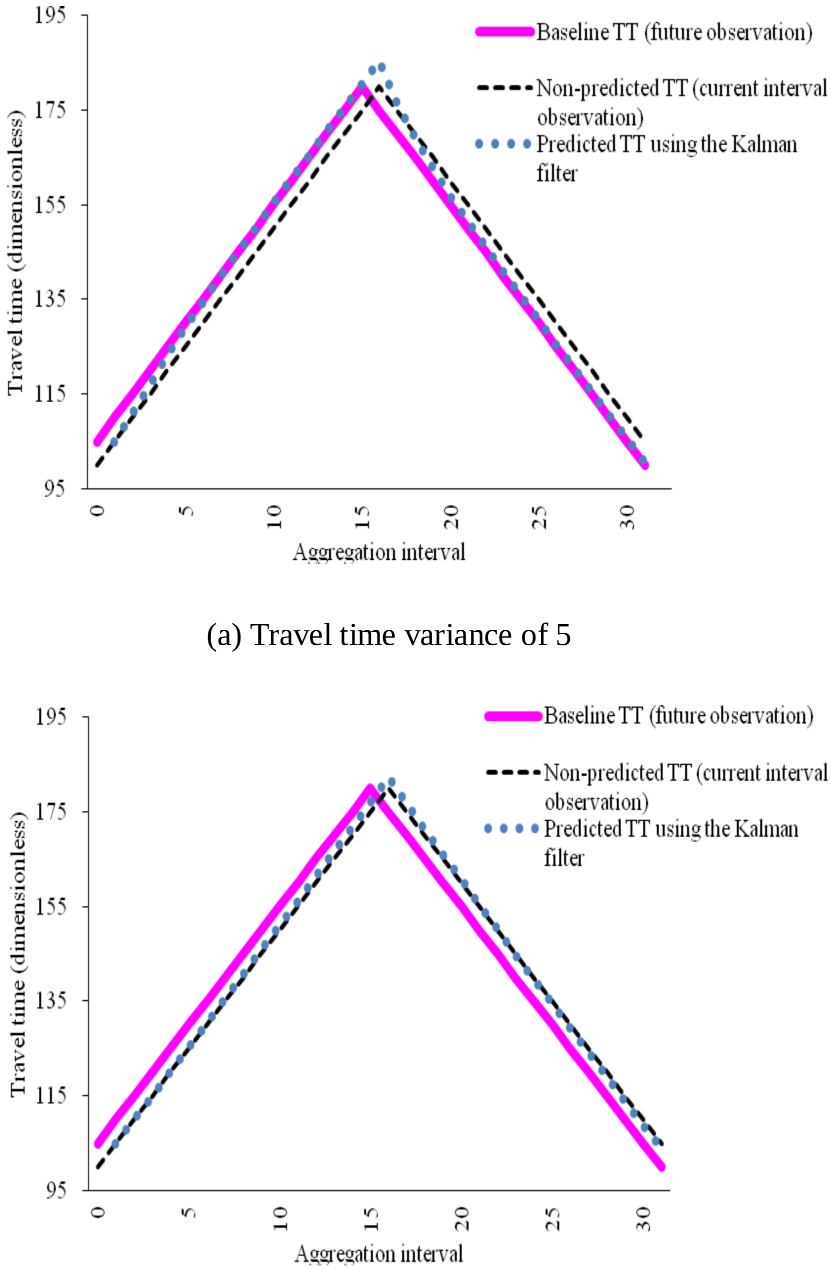
Step 2 Extrapolation
Extrapolate state estimate: $\hat{x}(t)_- = \varphi(t-1)\hat{x}(t-1)_+$
Extrapolate error covariance: $P(t)_- = \varphi(t-1)P(t-1)_+ + \varphi(t-1) + Q(t-1)$

Step 3 Calculation of Kalman gain: $K(t) = P(t)_-[P(t)_- + R(t)]^{-1}$

Step 4 Parameter update
Update state estimate: $\hat{x}(t)_+ = \hat{x}(t)_- + K(t)[z(t) - \hat{x}(t)_-]$
Update error covariance: $P(t)_+ = [I - K(t)]P(t)$

Step 5 Next iteration
Let $t = t + 1$. The next iteration begins from step 2

Analogous to other prediction methods, such as neural network algorithms and time series analyses, the TT prediction performance of the Kalman filter largely depends on the consistency between historical and current TT patterns. In addition, the variance of TT in the present aggregation interval plays a crucial role in prediction performance. Figure 2 shows the influence of variance on TT predictions, comparing the Kalman filter performance with TT variances of 5 and 500. As shown in Figure 2(a), a lower TT variance allows the Kalman filter to predict the baseline TTs (future observations the prediction targeted) extremely well, but with the larger variance in Figure 2(b), there is no notable difference between the predicted TTs and non-predicted TTs (current interval observations).



(a) Travel time variance of 5

(b) Travel time variance of 500

Figure 2. Travel time prediction simulations using the Kalman filter

2.2 MSE-based aggregation interval size identification

Park et al. (2009) suggested a novel methodology for determining the optimal aggregation interval for TT estimation. They used mean square error (MSE), shown in Equation (3), as a performance measure in determining the optimal aggregation interval.

$$
\begin{aligned}
MSE(h) &= E\left[\frac{\sum_{i=1}^{v(h)}\left(x^i(h)-\mu_{\bar{X}(h)}\right)^2}{v(h)}\right] = E\left[\frac{\sum_{i=1}^{v(h)}\left(x^i(h)-\bar{X}(h)+\bar{X}(h)-\mu_{\bar{X}(h)}\right)^2}{v(h)}\right] \\
&= E\left[\frac{\sum_{i=1}^{v(h)}\left(x^i(h)-\bar{X}(h)\right)^2+2\sum_{i=1}^{v(h)}\left(x^i(h)-\bar{X}(h)\right)\left(\bar{X}(h)-\mu_{\bar{X}(h)}\right)+\sum_{i=1}^{v(h)}\left(\bar{X}(h)-\mu_{\bar{X}(h)}\right)^2}{v(h)}\right] \\
&\approx E\left[\frac{\sum_{i=1}^{v(h)}\left(x^i(h)-\bar{X}(h)\right)^2}{v(h)}+\left(\bar{X}(h)-\mu_{\bar{X}(h)}\right)^2 \ (cross\ term\ is\ zero)\right]
\end{aligned}
\tag{3}
$$

where $x^i(h)$ = Observed link TT of i-th vehicle at time period h

$\mu_{\bar{X}(h)}$ = Expected sample mean link TT on a link at time period h

$v(h)$ = Observed number of vehicles at time period h

$\bar{X}(h)$ = Sample mean link TT on a link at time period h

In Equation (3), the cross term is assumed to be zero because the two terms are independent and their expectations become zeros. In other words, the first term represents the difference between an individual probe TT during time period h and the observed mean speed in the same time period. The second term represents the difference between the observed mean TT during time period h and the expected mean TT in the same time period. The mean square error (MSE) is composed of two components. Park et al. referred to the first component as the estimated mean square error of prediction (MSEE) and the second component as the variance of the predictor. The optimal aggregation interval is given by the point at which the MSE is minimized.

It was argued that this minimum MSE results from the trade-off between MSEE and the variance of the predictor. That is, in congested traffic conditions, a smaller aggregation interval induces a smaller MSEE but a larger variance of predictor, and vice-versa for a larger aggregation interval. However, in their research, aggregation intervals above 5 min were found to have a relatively constant variance of the predictor. This implies that the variance of probe TTs for an aggregation interval has a great effect on TT estimation for the aggregation intervals above 5 min. In reality, in a probe-based traffic information system, an aggregation interval less than 5 min is not recommended because of the small probe sample size.

Based on the findings of the Kalman filter feature and literature review, this research assumes that the optimal aggregation interval can be determined at which the probe TT variance is minimized when adopting the Kalman filter as a prediction method. To verify this assumption, we collect and analyze probe data. Details are given in the following sections.

3. DATA COLLECTION AND EVALUATION INDEX

Data for this research were obtained on a multilane highway near Seoul, Korea. The roadway section (3.1 *km* long) with four directional lanes lies between two access points, namely Jayuro service area and Moonbal IC (see Figure 3). Due to unavailability of detector data, the traffic volume on the day this study performed could not be identified exactly. However, the annual traffic volume statistics tells that 15,000 average daily traffic (ADT) traveled on the road, implying around 50 vehicles traveled at five minutes interval on average. There is no intermediate access point between the two points, implying that no sophisticated outlier filtering algorithm is needed. The data were obtained from two RSEs installed at the access points for 12 h (including congested and non-congested periods). Severe congestion occurred between 16:00 and 20:00 due to a surge in traffic during this period.

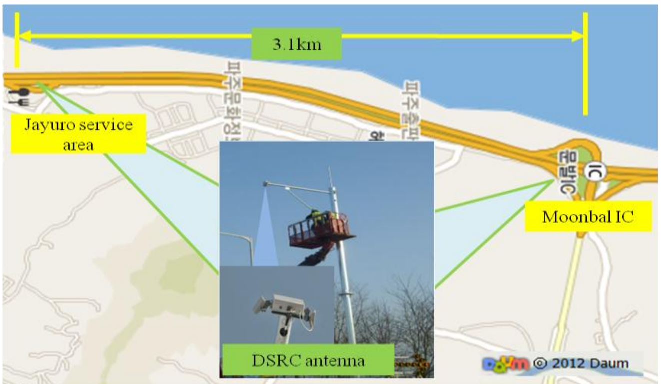


Figure 3. Data collection area and DSRC antenna

In a probe-based traffic information system, outliers might occur due to abnormal maneuvers (i.e., excessively fast/slow vehicles). If such outliers are included in the aggregated data, the mean TT can be biased. To address this issue, the core 90% of probe TTs at an aggregation interval was averaged to obtain an estimate of the section TT. The 90% was arbitrarily selected according to rule-of-thumb knowledge of the roadway section, and this proved to be an effective method. However, for more complicated roadway sections containing multiple access points, a more sophisticated outlier filtering method is recommended or, if necessary, need to be developed.

The reference TTs, against which the short-term TT prediction scheme suggested in this research was evaluated, were also gathered from the RSEs. The difference is that the reference (or baseline) TTs were those experienced by vehicles that passed the start point (Jayuro service area) during a given time period. That is, given that the predicted TTs were derived from probes that arrived at the end point (Moonbal IC) during an aggregation interval, the corresponding baseline was the average TT experienced by probes that passed the start point in the same aggregation interval. This is regarded to be reasonable because the predicted TTs are intended for disseminating to drivers in the vehicles passing the start point in that time interval. The mean absolute percent error (MAPE), which is used as the official evaluation index for traffic detectors in Korea, was selected as the evaluation index for the predicted TT in this research. MAPE is calculated as:

$$MAPE\,(\%) = \frac{\sum \left| TT_{ref} - TT_{pre} \right| \div TT_{ref} \times 100}{n} \tag{4}$$

where TT_{ref} is reference travel time and TT_{pre} is predicted travel time.

4. TRAVEL TIME PREDICTION AND ANALYSIS

4.1 Travel time prediction with fixed aggregation interval scheme

TT predictions were obtained using the Kalman filter combined with fixed aggregation intervals (5, 10, 15, 20, 25, and 30 min), and the predicted TTs were compared to the aforementioned baseline TT. Due to unavailability of historical TTs, analogous to the TTs this research is based on, current TT was exploited as a substitute for the historical TTs. As stated earlier, the objective of this research is to seek the more appropriate prediction scheme between static and variable aggregation intervals using the Kalman filter, this may not be a critical concern.

The results show that 5 min aggregation interval posed the lowest overall error (Table 1 and Figure 4). However, under non-congested conditions, the lowest prediction error was obtained with 15 min aggregation interval, implying that the optimal aggregation interval may vary according to traffic flow status. To distinguish traffic flow conditions (congestion or non-congestion), we used the measure of effectiveness (MOE) from the Korean Highway Capacity Manual. That is, if observed TTs were lower than the threshold speed corresponding to a level of service F, it was assumed to be congested.

Table 1. TT prediction errors by fixed aggregation intervals

Traffic flow conditions	Prediction errors (%) by aggregation intervals					
	5 min	10 min	15 min	20 min	25 min	30 min
Overall	6.2	7.7	9.6	12.7	15.1	18.6
Under congestion	8.6	15.5	22.7	32.3	37.5	46.4
Under non-congestion	5.0	3.8	3.3	3.4	3.9	5.6

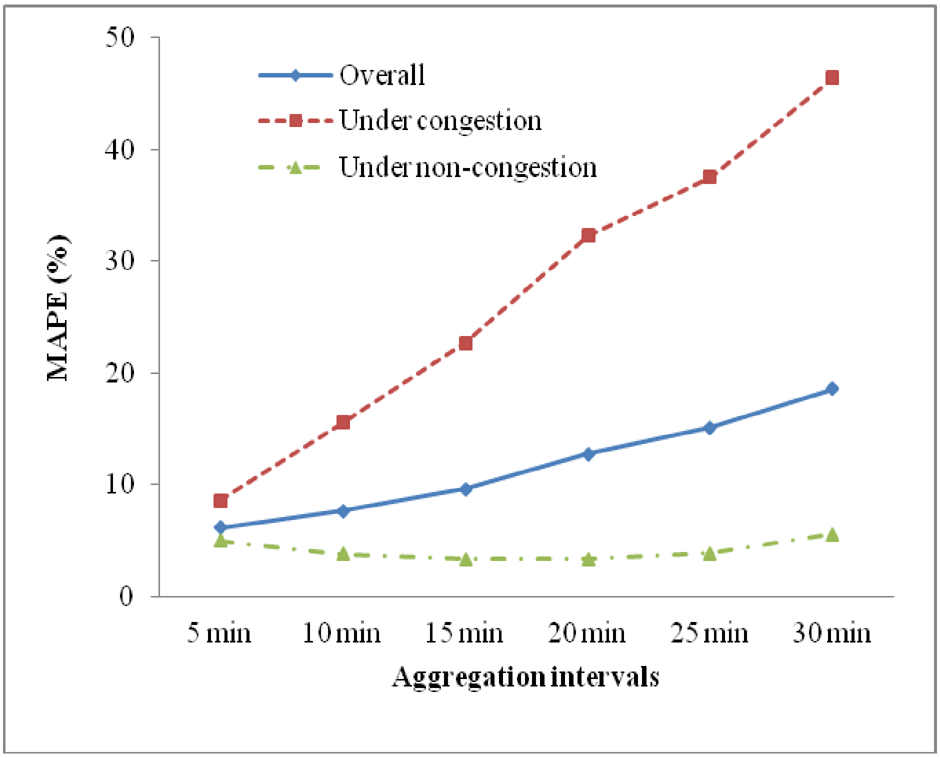


Figure 4. TT prediction errors for each aggregation interval

4.2 Analysis of prediction error by aggregation interval

To verify the assumption of a correlation between TT prediction error and TT variance, a correlation analysis on the two parameters was performed. The consequence, as shown in Table 2, is that the two parameters are highly correlated, with correlation coefficients of 0.98, 0.99, and 0.85 for the overall, congested, and non-congested traffic conditions, respectively. This result presumably indicates that TT prediction errors tend to diminish as TT variance decreases; hence, TT predictions based on aggregated TTs with the lowest variance are the best strategy for accurate real-life TT information. Generally, a lower aggregation interval ensures a lower time delay, resulting in lower TT estimation or prediction errors. However, in a probe-based TT information system, too short aggregation interval might induce unexpectedly high TT variance due to the small number of probes. As observed in our review on the Kalman filter characteristic, as well as in the previously mentioned literature and others (Sen *et al.*, 1997; Chen and Chien, 2001), a higher TT variance may increase the TT prediction error.

Table 2. Correlation analysis of the relationship between TT prediction error and variance

Agg. Int. (min)	Overall		Under congestion		Under non-congestion	
	TT Prediction error (MAPE)	TT variance (s)	TT Prediction error (MAPE)	TT variance (s)	TT Prediction error (MAPE)	TT variance (s)
5	6.2	417	8.6	1361	5.0	183
10	7.7	478	15.5	1631	3.8	159
15	9.6	586	22.7	2278	3.3	165
20	12.7	667	32.3	2838	3.4	161
25	15.1	780	37.5	3161	3.9	190
30	18.6	825	46.4	3604	5.6	260
r	0.98		0.99		0.85	

4.3 Travel time prediction with variable aggregation interval scheme

To utilize the high degree of correlation between prediction error and variance for accurate, real-life TT prediction, we developed an algorithm (see Figure 5) to select the optimal aggregation intervals for TT prediction where the Kalman filter is used for the TT prediction method. The algorithm uses the aggregated (or block) data with the lowest TT variance. From Table 1, the aggregation intervals with the minimum TT variances were 5 and 15 min for congested and non-congested conditions, respectively. Thus, the algorithm only considers the 5 to 15 min aggregation intervals in 5 min increments. However, the maximum and minimum aggregation intervals, as well as the increment, can be adjusted for a specific TT information system.

The predicted TTs with the variable aggregation interval scheme were evaluated using the aforementioned baseline TTs. The results were compared to those of the 5 min aggregation interval. It was observed that the prediction error decreased considerably under non-congested conditions (by around 40%, see Table 3 and Figure 6(a)), though there was no notable difference under congested conditions. This improvement was shown to be significant (p value of 0.01) using a t-test at 5% significance level. To investigate the probable cause of error with the fixed interval, we plotted the prediction errors and the number of probes in each 5 min aggregation interval against the time of day. As illustrated in Figure 6(b), the prediction errors of the fixed 5 min aggregation interval surge after 23:00. Interestingly, the probe sample size in this time period was considerably smaller (fewer than 5 probes) than those in other time periods. The lower sample size caused a high TT variance, resulting in the higher prediction error. The improvement in TT prediction accuracy using the variable aggregation interval is highly regarded because it means that accurate real-life TT information can be secured regardless of traffic conditions. Generally, real-life TT information during non-congested (e.g., nighttime) conditions might be regarded as unimportant; however, in the event of an incident, which could occur in a non-peak period and cause traffic congestion, the accuracy and drivers' reliability of real-life TT information should be emphasized. Figure 7 shows a graphical comparison of TT predicted by the variable aggregation interval scheme to the baseline TT.

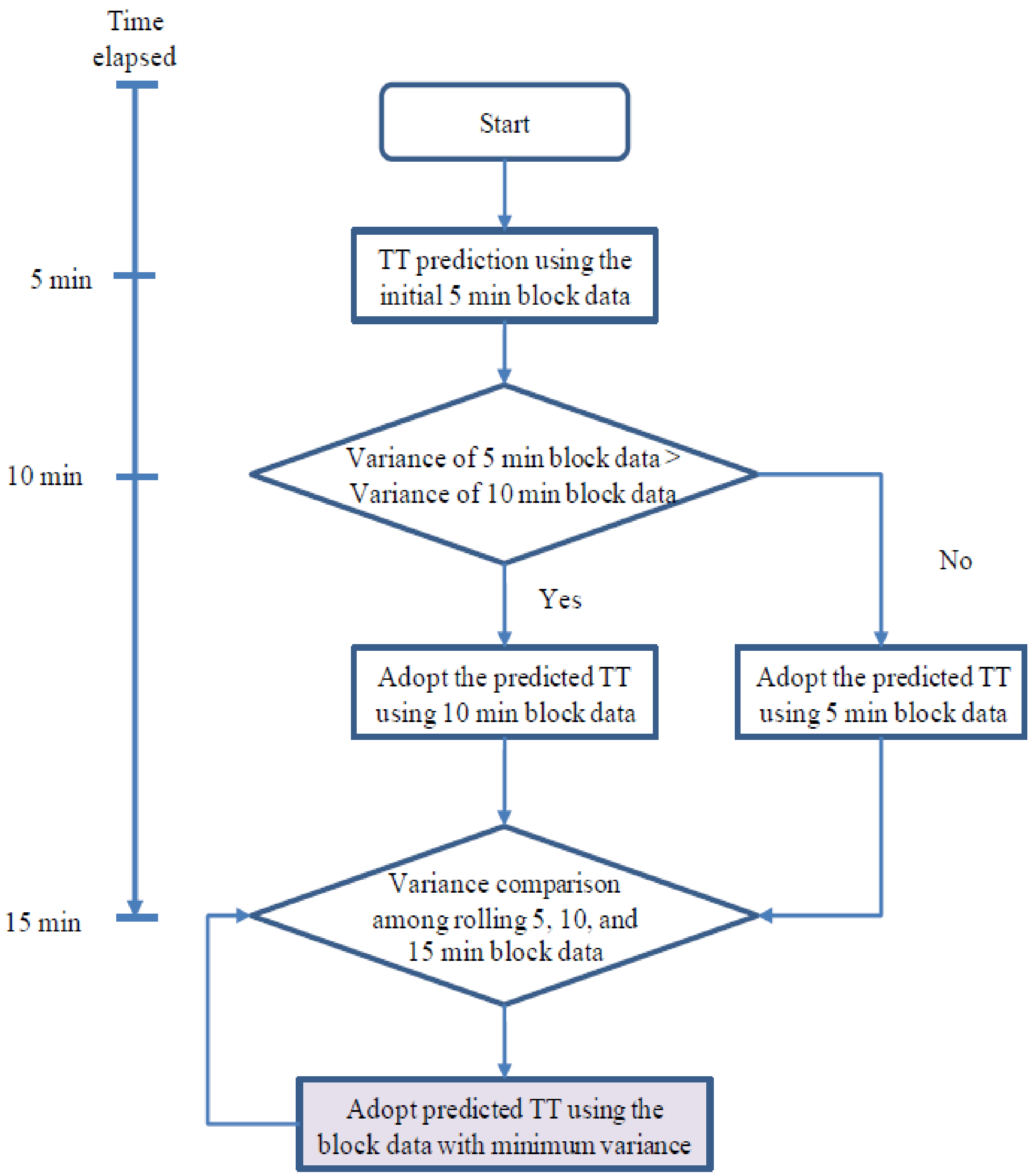
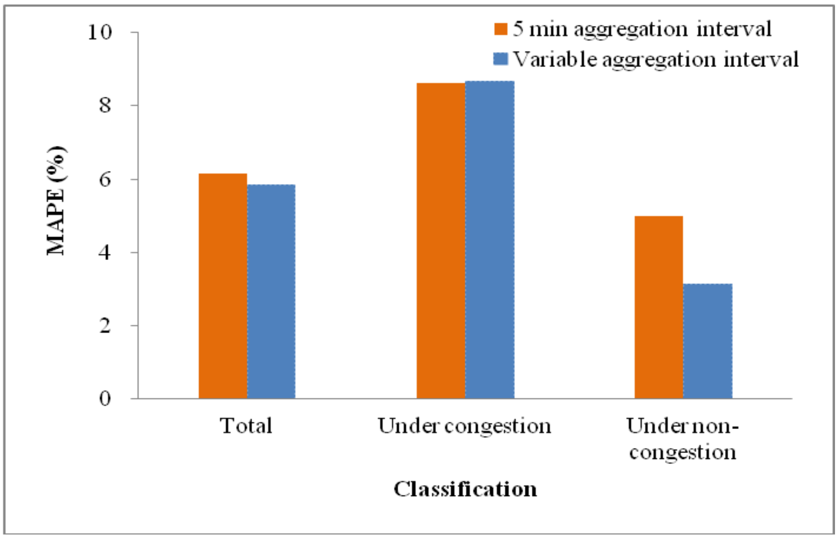


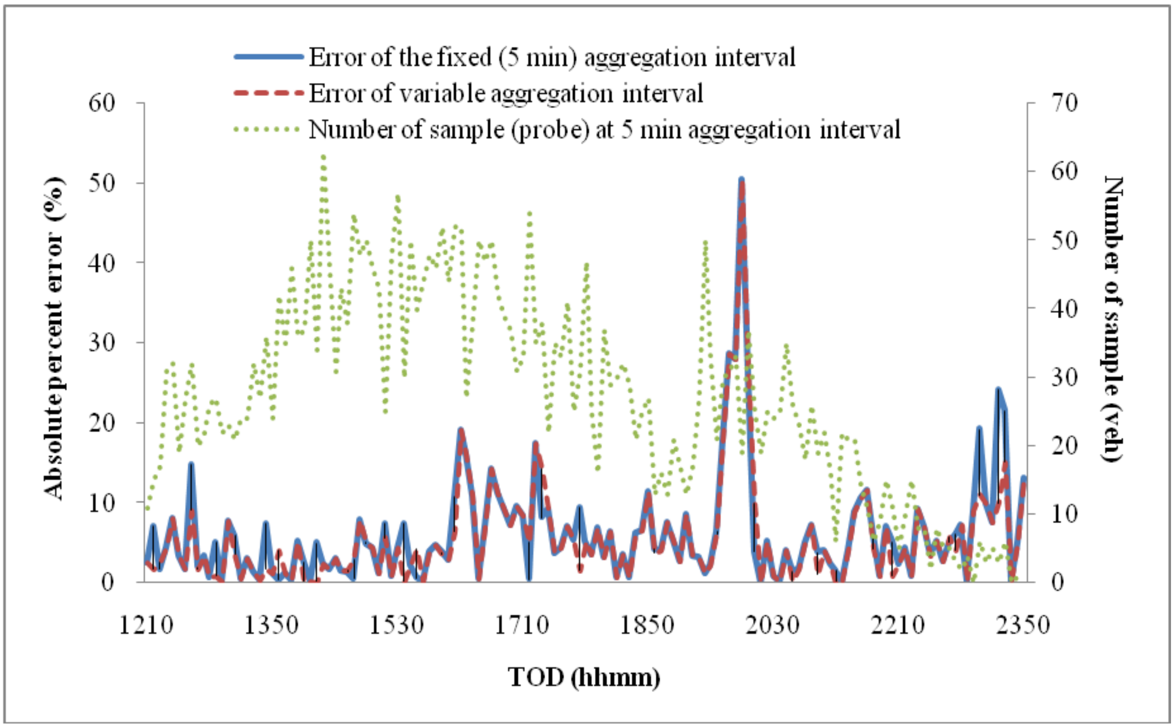
Figure 5. Algorithm for selection of optimal aggregation interval for TT prediction using the Kalman filter

Table 3. Analysis of TT prediction errors by variable and fixed (5 min) aggregation interval

Traffic flow condition	Absolute Percent Error (APE) by aggregation interval				Number of sample	*t*-test (α = 0.05)	
	Fixed (5 min)		Variable			*t*-static	p-value
	Mean	Variance	Mean	Variance			
Overall	6.2	46	5.8	46	141	0.40	0.69
Congested	8.6	79	8.7	80	45	-0.03	0.98
Non-congested	5.0	27	3.1	10	96	2.67	0.01



(a) Aggregated errors by traffic flow condition



(b) Detailed errors by time of day (TOD) combined with number of samples

Figure 6. TT prediction errors by aggregation scheme

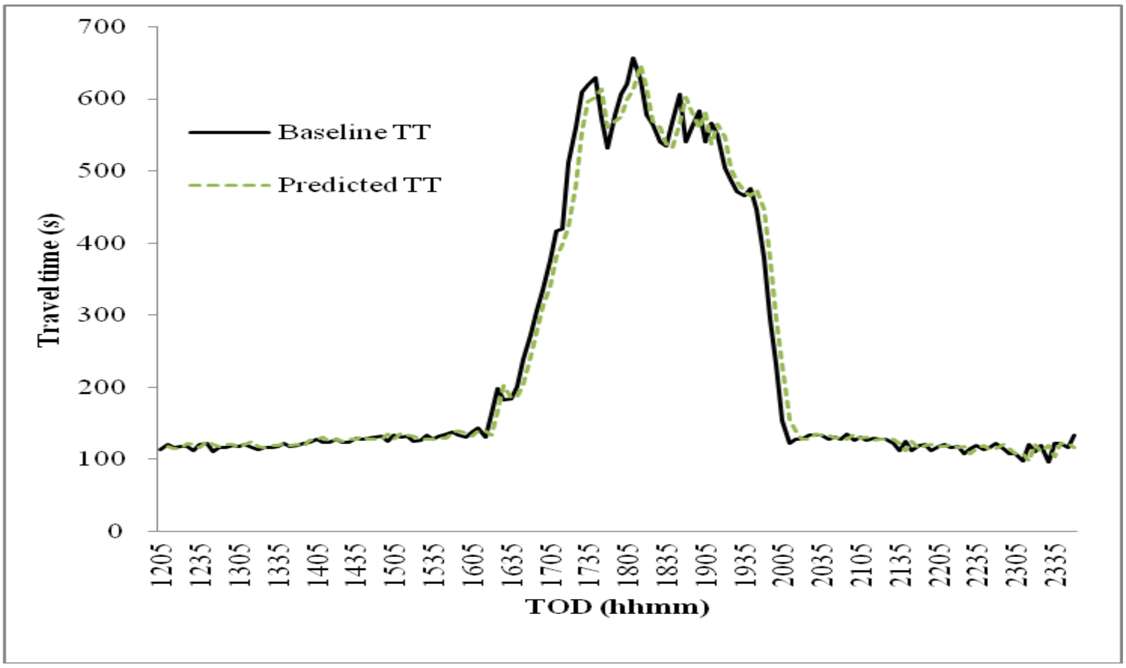


Figure 7. Baseline TT vs. TT predicted using the variable aggregation interval scheme

5. CONCLUSIONS AND FUTURE RESEARCH

This study proposed a variable aggregation interval scheme for predicting short-term TT with the Kalman filter. Such a technique could be utilized with the DSRC traffic information system that is being actively deployed in Korea as a result of the increasing market penetration of OBUs, which were originally introduced for ETCS. In the proposed scheme, the optimal aggregation interval was determined according to the minimum probe TT variance. Predicted TTs were evaluated using the baseline TTs of vehicles that passed the start point of the road section during the period that the predicted TT was posted on information provision media such as variable message signs, the Internet, and smartphones. We used MAPE, which is used for official detector evaluation in Korea, as an evaluation index.

Under congested conditions, little difference between fixed and variable aggregation intervals was observed. However, under non-congested conditions, predicted TT error decreased by approximately 40% in the variable aggregation interval scheme compared to the error of the 5 min aggregation interval; the significance of this improvement was verified with a *t*-test. As stated above, because accurate real-life TT information can be even emphasized during non-peak periods, the improvement could be highly regarded. Consequently, TT estimation/prediction error tends to increase with TT variance for aggregation intervals above 5 min, showing similar results to the Kalman filter characteristic and a previous research (Park *et al*., 2009).

We expect the algorithm developed in this research to be practically applied to probe-based traffic information, particularly where there is limited probe data, for the dissemination of reliable TT information. However, this research verified the effectiveness of the suggested algorithm using only one set of data from one location, though sufficient samples (over 12 h) were secured for statistical analysis. Therefore, the spatial transferability of the algorithm may need to be further verified with more data from other sites.

REFERENCES

Chien, S. I. J., Liu, X., Ozbay, K. (2003) Predicting Travel Times for the South Jersey Real-Time Motorist Information System, TRB 82nd Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Chien, S. I. J., Kuchipudi, C. M. (2003) Dynamic Travel Time Prediction with Real-Time and Historic Data, *Journal of Transportation Engineering*, Vol. 129, No. 6, American Society Civil Engineering.

Gajewski, B., Turner, S., Eisele, B., Spiegelman, C. (2000) ITS data archiving: statistical techniques for determining optimal aggregation widths for inductance loop detectors, *Transportation Research Record 1719*, Transportation Research Board.

Chen, M., Chien, S. I. J. (2001) Dynamic Freeway Travel-Time Prediction with Probe Vehicle Data (Link Based versus Path Based), *Transportation Research Record 1768*, Transportation Research Board.

FHWA. (1998) *ITS Field Operational Test Summary - TRANSCOM System for Monitoring Incidents and Traffic*.

Grewal, M. S., Andrews, A. P. (1993) *Kalman Filtering-Theory and Practice*, Prentice Hall.

Houston TranStar (2003) Houston TranStar Fact Sheet.

Huisken, G., Berkum, E. C. (2003) A Comparative Analysis of Short-Range Travel Time Prediction Methods, TRB 82nd Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

ITS Korea, HITECOM Co. Ltd., Ajou University. (2008) *Development of Utilization Technology for Road Traffic Information System using Dedicated Short Range Communication*, Final Report, Korea Highway Corporation. (in Korean)

Myung, J., Kim, D.-K., Kho, S.-Y., and Park, C.-H. (2011) Travel Time Prediction Using *k* Nearest Neighbor Method with Combined Data from Vehicle Detector System and Automatic Toll Collection System, *Transportation Research Record 2256*, Transportation Research Board.

Kuchipudi, C. M., Chien, S. I. J. (2003) Development of a Hybrid Model for Dynamic Travel Time Prediction, TRB 82nd Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Oh, C., Ritchie, S.G., Oh, J. (2005) Exploring the relationship between data aggregation and predictability toward providing better predictive traffic information, TRB 84th Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Park, D., Rilett, L. R., Gajewski, B. J., Spiegelman, C. H., Choi, C. (2009) Identifying optimal data aggregation interval sizes for link and corridor travel time estimation and forecasting, *Transportation*, 36 77-95, Springer.

Qiao, F., Wang, X., Yu, L. (2004) Double-sided determination of aggregation level for ITS data, TRB 83rd Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Qiao, F., Wang, X., Yu, L. (2003) Optimizing aggregation level for ITS data based on wavelet decomposition, TRB 82nd Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Sen, A., Thakuriah, P., Zhu, X. Q., Karr, A. (1997) Frequency of Probe Reports and Variance of Travel Time Estimates, *Journal of Transportation Engineering*, Vol. 123, No. 4, American Society Civil Engineering.

SwRI (1998) *Automatic Vehicle Identification Model Deployment Initiative-System Design Document*.