DATA FUSION ALGORITHM
FOR IMPROVING TRAVEL TIME FORECAST

Sunghan LIM, Chungwon LEE
Advanced Transport Research Division, Department of Transportation Engineering
Korea Institute of Construction Technology, University of Seoul
aday@kict.re.kr, chungwon@uos.ac.kr

ABSTRACT

Travel time is considered to be more useful to users than other travel related information such as speed. It can be mainly estimated by point detection systems or spatial detection systems. This paper investigated the deficiency of such detection systems when they are used solely in estimating the travel time. To resolve it, a fusion algorithm, which simultaneously utilizes the data from both of point detection and spatial detection systems, is suggested. The study precisely defined so called the time-lag issue originating from the latter. To overcome it, the traffic states variation between the present time period and the next is analyzed using the online data from point detection systems. The traffic states are classified into several types by fuzzy logic and represented by the states diagram with the flow and the density dimensions. Considering the traffic states variation, the travel time which was estimated using the data from spatial detection system is renewed.

1. INTRODUCTION

Travel time, among traffic information, which is regarded the most typical and preferred by the user is generally obtained using point detection systems or spatial detection systems. Point detection systems are the indirect estimate system designed to convert the speed collected at the certain point into spatial mean speed to estimate the travel time, while spatial detection systems are the direct estimate system designed to estimate the travel time from the probe car based on AVI(Automatic Vehicle Identification) device, etc.

The speed from point detection systems represents the traffic state of the road at a point of time and the travel time calculated based on such system has a specific meaning on assumption that the state is continuously sustained, which is however almost impractical, thereby causing discrepancy between actual travel time and estimated travel time by point detection systems.

In such aspect, the travel time obtained from spatial detection systems is what reflected the temporal and spatial variation of the traffic state, which is considered to be more accurate travel time. However the travel time is generated at the time when the vehicle passes the end point of the section, it's the late information as much as the time spent for travel. Thus, a prediction module is needed to estimate the travel time information required by the user using spatial detection systems.
Viewing from the time standpoint, traffic information contains the limitation having the truth only at a point of time (Korea ITS Association, 2008). The travel time information which the user wants is the travel time with the present time as departure time, which incorporates the temporal and spatial variation of the traffic state in future. That is, a prediction is needed as a precondition in order to grant the travel time to be traffic information.

This study, in an attempt to cope with the shortcomings of point detection systems and spatial detection systems, was intended to develop the fusion algorithm, making use of point detection systems and spatial detection systems. To deal with the time-lag issue of the spatial detection systems, the travel time which was estimated using the data from spatial detection systems is renewed considering the traffic state variation which was predicted using a real-time traffic data from point detection systems.

2. REVIEW OF PREVIOUS STUDIES

The study on travel time prediction has been carried out at home and abroad in diversified way in line with development of detection technologies. Categorizing them from detection system-wide, the studies using the data from point detection systems such as loop detector, video Image detector, etc were mostly carried out at early stage, and recently, the studies using the data from spatial detection systems such as AVI, GPS, etc have been carried out. To deal with the shortcoming of point detection systems and spatial detection systems, a fusion algorithm that predicts the travel time combining the two systems has been under study.

A travel time prediction model using traffic data from the point detection systems is designed to analyze the correlations between traffic data (volume, speed and occupancy rate) and travel time and then predict the travel time, making use of real-time data. Many ITS operators have adopted the system converting the point speed to spatial mean speed. It's used because of the benefit of convenience in using, but the accuracy tends to fall during congested state and the error often occurs in the course of converting it to spatial data. The existing travel time forecasting models also include the time-series model (Anderson et al, 1994 ; Van Arem et al, 1997 ; Al-Deek, 1998), regression model (Clark, 2003 ; Nikovski, et al, 2005) and traffic flow model (Kim Seung-il, 2007 ; Kim Ji-hong, 2006).

The study using the traffic data from the spatial detection systems started in earnest since Advance project began in early 1990s. San Antonio’s TransGuide System is one of the system representing it at abroad, which estimates the travel time using travel time prediction algorithm developed by SwRI(1998). This algorithm was intended to cope with the errors of AVI which occurred due to insufficient number of samples, but it still has a limit as it estimates the arrival time-based travel time (Korea ITS Association, 2008). Domestically, Lee Chung-won(2002) developed the path travel time forecasting model using Kalman filter to provide the reliable traffic information. This model was intended to forecast the travel time by link using the data obtained from probe vehicle and applying the Kalman filter and then using the expected value, it estimates the experienced path travel time. Kim Jae-jin(2006) defined the on-line departure time-based travel time and using bayesian theory, he
developed on-line departure time-based travel time forecasting algorithm. Kim Jae-jin and et al. (2007), in an attempt to determine the appropriateness of spatial design to provide the on-line departure time-based travel time information, employed AVI travel time data. To evaluate the optimal link length, he defined the concept of error in providing the on-line departure time-based travel time information and suggested the improved method to provide the real-time travel time data based on departure time.

Linn and Hall (1991) classified the current data fusion algorithm into 5 categories; data association, positional estimation, identity fusion, pattern recognition and artificial intelligence, which were further broken down to three stages, depending on the characteristic of information by Dailey and et al. (1996).

Level 1 data fusion methods include figure of merit (FOM), gating techniques, and Kalman filters. The previous studies on travel time forecast are mostly the studies applying the variance ratio of travel time (standard deviation, variance) or reliability ratio as fusion weight (Berka, et al., 1995; Yamane, et al., 2000; Kim Sung-hyun, 2005; Lee Hyun-jai, 2005; Shin Chi-hyun, 1998). Yoo Jung-hoon (2008) developed a hybrid model designed to estimate the spatial travel time using VDS (Vehicle Detection System) information and estimate the time passing the intersection using AVI information and then combine them. That is, it used the model comprising the travel time by combining the travel time between the intersection and the delay at the intersection. Chu and et al. (2004), based on typical traffic flow theory and Kalman filters, studies the algorithm (AKF Fusion Algorithm) forecasting the travel time by combining the point detector data (traffic data from a single loop) and probe car data (travel time from sample car). A point detector data was used to determine the state equation comprising the Kalman filters and a probe car data was used to comprise the measurement equation. Takahashi (1996) forecasted the travel time based on travel time data from AVI and the point detector data, using Kalman filter method.

Level II data fusion methods include bayesian decision theory, dempster-schafer evidential reasoning, adaptive neural networks, and cluster methods. Sisiopiku, et al. (1993) developed data fusion algorithm with the loop detector data, probe car data, and historical data, using regression method and bayesian method. There are also studies adopting neural network (Lee Eui-eun, et al., 2002; Chi Xie, et al., 2004; Wen, et al., 2005; Byun Sang-chul, 2006).

Level III data fusion methods include expert systems, blackboard architecture, and fuzzy logic. Kim Young-chan, et al. (2001) developed data fusion algorithm for interrupted and uninterrupted facility. For interrupted road, a fusion algorithm using fuzzy theory was developed. Three models which are data fusion model employing a simple regression method, hybrid neuro-fuzzy model, and hybrid fuzzy-genetic model are developed. Choi Gui-joo, et al. (1998) carried out the study on data fusion algorithm to create the single traffic information by combining the different collection systems with different reliability. Through the initial conversion of the data collected using fuzzy linear regression model, reliability rates were given by collection system and then a final travel time was estimated by D-S (Dempster-Schafer) rule and bayesian theory.

3. TIME-LAG ISSUE

Travel time is usually obtained from point detection systems or spatial detection systems, and as
seen in Table 1, they have their own advantage and disadvantage. The traffic data collected from point detection systems well represented the online traffic state, but it requires additional process to estimate the section travel time, resulting in some errors. On the contrary, a spatial detection systems designed to directly measure the travel time with two detectors at both ends relatively provides more accurate time travel incorporating a dynamic traffic state at the section. But whereas the travel time required by the user is the future travel time information with the present time as departure time, the travel time provided by the spatial detection systems is the past travel time information with the present time as arrival time (arrival time-based travel time).

Table 1. Comparison between the point detection and spatial detection systems

<table>
<thead>
<tr>
<th></th>
<th>Point detection</th>
<th>Spatial detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantage</td>
<td>- A real-time traffic state</td>
<td>- Accurate travel time</td>
</tr>
<tr>
<td>Disadvantage</td>
<td>- A complex process to covert it to spatial data and with errors.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Low accuracy during congestion</td>
<td>- Time-lag due to arrival time-based calculation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- A past travel time information that lack the reality</td>
</tr>
</tbody>
</table>

The issue of arrival time-based travel time results from the information generation principle of the spatial detection systems. Fig 1 indicates the information generation principle of the spatial detection systems with 5-minute time aggregation interval. Irrespective of departure time of individual cars, the travel time collected from spatial detection systems at 09:00 is the representative travel time of the cars arrived at 08:55~09:00.

![Figure 1. The travel time information generation principle of the spatial detection systems](attachment:image)

As seen in Fig 2, the travel time information required by the driver is the travel time in future with t
as departure time, $T_r(t)$. But the travel time from spatial detection systems is the travel time in past with $t$ as arrival time, $T_{sds}(t)$. Where, $\Delta=T_{sds}(t)-T_r(t)$ which is the difference between $t$ and $t-\Delta$ is called time-lag and the error caused by the time-lag is $T_{sds}(t)-T_r(t)$. A time-lag is attributable to time difference between the cars estimating the travel time and the cars receiving the travel time, and the error in travel time is caused by the traffic state variation resulting from the time difference.

In general, a travel time varies depending on driver’s characteristics even under the same traffic state. Thus, the car started earlier doesn’t necessarily arrive earlier, supporting the relationship of $T_{sds}(t) \not= T_r(t-\Delta)$. But in the process of eliminating outliers, the travel time which is significantly different from the mean travel speed is eliminated, making it possible the relationship, $T_{sds}(t) \cong T_r(t-\Delta)$

![Figure 2. Time-lag and travel time error](image)

As seen in Fig 3, the error of travel time resulting from time-lag is in proportion to the changes of traffic state, and thus should the traffic remain unchanged, the error be not occurred due to time-lag though time-lag still exists. A real travel time at time $t$, $T_r(t)$ can be represented with the function of the travel time by spatial detection systems at time $t$, $T_{sds}(t)$. And should the traffic state $S(t, t-\Delta)$ between the two times ($t, t-\Delta$) which is the base of $T_{sds}(t)$ and the traffic state $S(t, t+\delta)$ between the two times ($t, t+\delta$) which is the base of $T_r(t)$ be similar with each other, $T_r(t)$ and $T_{sds}(t)$ also have a similar value. In contrast, should $S(t, t-\Delta)$ and $S(t, t+\delta)$ be different, $T_r(t)$ and $T_{sds}(t)$ are different too. Such theory can be represented by Equation (1). Thus to deal with the error resulting from time-lag, it’s necessary to analyze and forecast the traffic state variation by time interval.

$$T_r(t) = f(T_{sds}(t)), \begin{cases} \text{if } S(t-\Delta, t) \cong S(t, t+\delta), & T_r(t) \cong T_{sds}(t) \\ \text{if } S(t-\Delta, t) \not= S(t, t+\delta), & T_r(t) \not= T_{sds}(t) \end{cases}$$

Where,

$T_r(t) =$ Real travel time at time $t$

$T_{sds}(t) =$ Travel time by spatial detection systems at time $t$

$S(t-\Delta, t) =$ Traffic state between two times ($t-\Delta, t$) which is the base of $T_{sds}(t)$

$S(t, t+\delta) =$ Traffic state between two times ($t, t+\delta$) which is the base of $T_r(t)$
4. TRAVEL TIME FORECASTING ALGORITHM

While point detection systems have the advantage of incorporating the traffic state at the location on a real-time base, spatial detection systems provide accurate travel time incorporating a dynamic traffic state. This study was intended to develop the reliable travel time forecasting algorithm combining a real-time data from point detection systems and accuracy from spatial detection systems.

As reviewed previously, a real travel time, $T_r(t)$ could be represented by the function of travel time from spatial detection systems, $T_{sds}(t)$ and the difference in travel time due to traffic state variation needs to be incorporated. Hence, a real travel time, $T_r(t)$ may be represented as follows.

$$T_r(t) = T_{sds}(t) + \Delta T$$

$$\Delta T = T_{S}(\delta t) - T_{S}(\Delta t)$$

Where,

$T_r(t)$ = Real travel time at time $t$

$T_{sds}(t)$ = Travel time by spatial detection systems at time $t$

$\Delta T$ = Difference in travel time due to traffic state variation

$T_{S}(\delta t)$ = Travel time in traffic state between two times ($t-\delta$, $t$) which is the base of $T_{sds}(t)$

$T_{S}(\Delta t)$ = Travel time in traffic state between two times ($t$, $t+\delta$) which is the base of $T_r(t)$

To analyze and forecast the traffic state variation, the traffic state needs to be defined first and then traffic state by time interval needs to be determined using a real-time data. Youngho Kim (2002) analyzed the changes of traffic flow-density relation over the time and expressed it in states diagram. Then he analogized the dynamic flow-density relation by applying fuzzy logic to states diagram, and categorized the traffic flow at homogeneous motorway sections into 6 traffic states. In this study, a traffic states diagram was developed from a dynamic flow-density relation based on point detection.
systems by adopting Youngho Kim (2002)’s approach among the previous studies on classification of traffic states. The traffic state by time interval was determined by applying fuzzy-logic. Figure 4 is the traffic states diagram developed, making use of traffic data from video image detector installed at Pangyo → Ilsan Jangsoo IC on Seoul Outer Beltway for 5 days (Mon, Mar 16 ~ Fri, Mar 20) and membership functions and fuzzy if-then rules developed for applying fuzzy-logic are illustrated in Figure 5 and Table 3.

![Traffic states diagram](image)

**Figure 4. Traffic states diagram of traffic flow (Pangyo → Ilsan Jangsoo IC on Seoul Outer Beltway)**

![Fuzzy membership functions](image)

**Figure 5. Fuzzy membership functions of the input and out variables**
Table 2. Fuzzy if-then rules

<table>
<thead>
<tr>
<th>Rule</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speed</td>
<td>Flow</td>
</tr>
<tr>
<td>1</td>
<td>Very High</td>
<td>Low</td>
</tr>
<tr>
<td>2</td>
<td>High</td>
<td>Very Low</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>4</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>5</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>6</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>7</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>8</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>9</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>10</td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>11</td>
<td>Very Low</td>
<td>Medium</td>
</tr>
<tr>
<td>12</td>
<td>Very Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

In formula (2), Tsd(t) is directly collected by spatial detection systems, thus the performance of algorithm depends on accuracy of prediction for difference in travel time, $\Delta T$ affected by the traffic states variation. The online data from point detection systems is more useful in analyzing and forecasting the traffic states variation, compared to the data from spatial detection systems. So it’s necessary to estimate $\Delta T$ using the data from point detection systems.

$\Delta T$ is estimated as the difference between $T_S(\Delta t)$ and $T_S(\delta t)$. $T_S(\Delta t)$ could be estimated using the past data (traffic volume, speed and occupancy rate) from point detection systems between the two times $(t, t-\Delta)$, but $T_S(\delta t)$ could be estimated after forecasting the traffic data (traffic volume, speed and occupancy rate) between the two times $(t, t+\delta t)$ in future. In this study, a traffic data by time interval was forecasted using Kalman filter model, which dynamically applies the filter satisfying the minimal requirements of mean square error in such a way of comparing the value collected by time interval with the value forecast based on previous time. It’s preferred on on-line system as it could accommodate the new value collected repeatedly. It, in most of cases, tends to rapidly converge to true value, irrespective of initial value (Lee Chung-won, 2002). Kalman filter has been used in various areas handling the time series data and in traffic sector, it was used for the study on traffic volume forecast (Okutani et al, 1984 ; Vythoulkas, 1993) and the study on travel time forecast (Takahashi, 1996 ; Lee and Choi, 1998 ; Lee Chung-won, 2002 ; Chu et al, 2004).

A travel time forecast algorithm discussed in this study was based on premise that the travel time variation caused by traffic states variation could be modeled. Thus it’s intended to model the relations between the traffic state index from the point detection systems and the travel time from the spatial detection systems and then incorporate the difference $\Delta T$ resulting from the variation of traffic state to forecast the travel time.
A data fusion algorithm for forecasting the travel time is summarized as follows.

**Definition of Traffic States**
- Analysis of traffic flow–density relation
- Development of traffic states diagram

**Estimate of TSI**
- Application of Fuzzy logic
- Input variables: volume(q), density(k), speed(v)
- Output variable: shockwave speed ⇒ TSI

**Modeling traffic state–travel time relation**
- Modeling TSI – Tr relations

**Forecast traffic variable(q, k, v)**
- Application of Kalman Filter model

**Forecast traffic state & travel time variation(ΔT)**
- Applying Fuzzy logic ⇒ TSI
- Applying TSI – travel time(Tr) model

**Travel time by spatial detection systems**
- Historical data

**Travel time by spatial detection systems**
- A real-time data(Tsds)

Figure 6. Data fusion algorithm for improving travel time prediction

**Stage 1** Definition of traffic states
- Develop the traffic states diagram by analyzing the changes of traffic flow-density relation over the time from point detection systems

**Stage 2**: Development of traffic state index(TSI)
- Employment of TSI to quantify the traffic states
- Set the online traffic data(traffic volume, density and speed) from point detection systems as input variable and set the shockwave speed as output variable, by applying fuzzy logic to traffic state diagram.
- Since a dynamic shockwave speed (Δq/Δk) is possibly be estimated from the online traffic flow – density relation, set the shockwave speed as TSI.

**Stage 3** Modeling traffic state – travel time relations
- Set the TSI and Tr relation to incorporate the travel time variation depending on the traffic state variation.

**Stage 4** Forecast the traffic variable
- Forecast the traffic variable(traffic volume, density and speed) at the time interval (t+n) by applying
Kalman filter model to the online data from point detection systems

[Stage 5] Estimate of TSI and travel time variation (ΔT)
- Estimate the TSI variation by applying fuzzy-logic to traffic variable forecasted at [Stage 4]
- Estimate the travel time variation (ΔT) by applying traffic state – travel time relation model developed at [Stage 3]

[Stage 6] Estimate of forecast travel time (Tp)
- Estimate the forecast travel time (Tp = Tds + ΔT) by incorporating travel time variation (ΔT) to travel time by spatial detection systems (Tds)

5. CONCLUSION

The study, based on analysis of the characteristics of point detection systems and spatial detection systems, is intended to develop the data fusion algorithm to enhance the reliability of travel time forecast. It made use of the advantage of spatial detection systems which are able to estimate the accurate travel time incorporating the dynamic traffic states and in an effort to deal with the shortcoming of the spatial detection systems, time-lag, it also adopted the point detection systems which have the advantage of real-time data.

Based on premise that the variation of travel time resulting from transfer of traffic state would possibly be modeled, it attempted to model the relations between TSI from the point detection systems and travel time from the spatial detection systems, and based on result, algorithm was developed to forecast the travel time by incorporating travel time difference resulting from variation of traffic state.

A spatial detection systems-based algorithm designed to estimate the travel time from the probe car based on AVI device, etc had the problem of travel time error caused by insufficient samples. But AVI system which was installed on a limited lane on national highway (interrupted flow facility) has been installed on entire lanes recently, and when it comes to expressway (uninterrupted flow facility), such problems are expected to be resolved thanks to High-Pass and increasing number of cars equipped with the terminal. And considering the increasingly expanding point and spatial detection systems-based ITS facilities, the algorithm recommended in this study is expected to be utilized for its applicability on a real-time base.

Kalman Filter model which will be applied to forecast the traffic variable by time interval in future has a limit in accommodating the rapidly-changed traffic state due to the characteristic of time series model. Given the characteristics of traffic flow which is affected by the upstream and downstream traffic conditions, more reliable traffic variable forecast and travel time forecast would be possible when utilizing the data from upstream and downstream point detection systems.

The study was not able to deal with the performance of algorithm and field verification, and thus further study on field applicability needs to continue. Verification of algorithm using field data remains as the open issue for further study.
REFERENCES

17. Yamane, Kenichiro and Takumi Fushiki.(2000), Development of Travel time estimation system combining license plate recognition AVI and Ultrasonic vehicle detectors, 6th ITS World Congress.
19. Sunghyun Kim(2005), A Link Travel Time Estimation Algorithm Based on Point and Interval Detection Data over the Homogeneous National Highway Section, Ph.D Thesis at Seoul National University, Seoul, Korea.