Data Fusion Prototype for Dynamic Signal Coordination on Urban Corridors *

by Farhana NAZNIN**, Babak MEHRAN *** and Masao KUWAHARA****

1. Introduction

Availability of traffic information from different sources provides more opportunities for dynamic evaluation of traffic conditions and real time traffic management. Considering different accuracies and limitations of traffic data from various sources, data fusion techniques are applied to extend the spatial and temporal coverage of data.

So far, data fusion techniques have been used mostly for travel time estimation on urban streets using fixed and probe sensor data. Majority of fusion techniques to estimate travel time are merely based on statistical methods without considerations for traffic engineering concepts. In addition, existing methods do not use rich probe trajectory data which represent traffic conditions in time and space. Instead probe trajectories are used simply to extract travel times. Travel times estimated from probe vehicles (e.g. taxis or buses) might vary according to vehicle types, though location and waiting time of probe vehicles stopping in a queue does not depend on vehicle types. Such valuable information of probe trajectories could be combined more effectively with fixed sensor data to reproduce trajectories of all vehicles.

We propose a novel fusion technique to estimate vehicle trajectories using multi-sensor traffic data based on concepts of traffic engineering which can fully utilize probe trajectory information. Once vehicle trajectories are estimated, they could be used for several applications such as travel time estimation and signal coordination. In this paper application of the methodology for signal coordination is demonstrated in the end.

2. Background and literature review

Existing data fusion techniques for travel time estimation can be categorized as follows [1]: *i)* Statistical based, *ii)* Probabilistic based and, *iii)* Artificial cognition based. Among statistical techniques, weighted combination of travel times from different sources is the most common approach. Weights are generally derived from variance-covariance estimation errors by applying methods such as "voting technique" [2].

Probabilistic approaches such as Bayesian approach [3], Dempster-Shafer inference [4] were used to tackle the problem of data fusion for the purpose of travel time estimation. El Faouzi [3] describes application of Bayesian approach to combine travel times estimated from conventional loop detectors and probe vehicles on urban routes. Mathematical basis for applying Dempster-Shafer inference to improve travel time estimation by fusing estimated travel times from toll collection stations and conventional loop detectors is demonstrated by El Faouzi *et al.*[4]. Ivan *et al.*[5] applied artificial neural networks (ANN) to develop arterial incident detection models that fuse probe vehicle and fixed detector data. Some researchers proposed data fusion frameworks which include different fusion stages. Combinational fusion techniques may use statistical, probabilistic and even artificial cognition concepts in several stages before producing the final output from different data source. Choi and Chung[2] proposed an algorithm for fusing multiple data sources to generate a representative value in terms of link travel time focusing on detector and GPS probe vehicle data. Their methodology involves voting technique, fuzzy regression, and Bayesian pooling method.

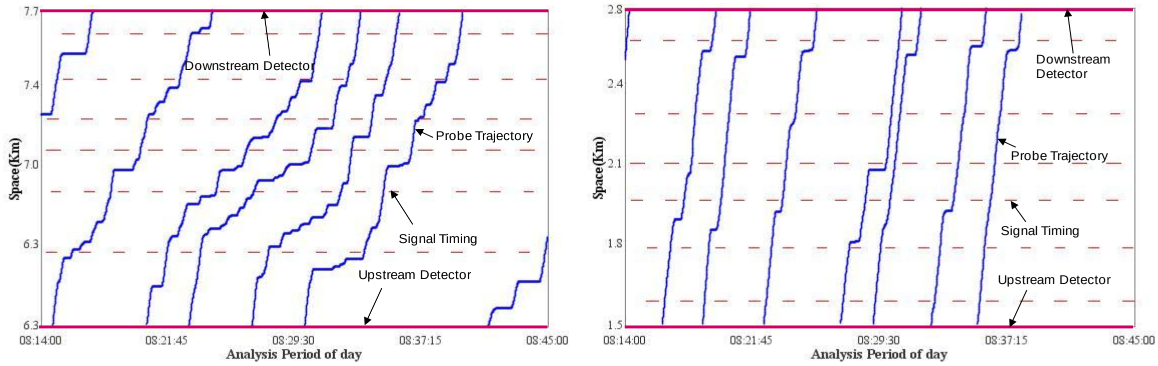
Existing data fusion methodologies basically relay on statistical concepts and do not benefit from principles of traffic engineering to fill the gaps in data. Berkow *et al.*[6] proposed a fusion method to combine the data derived from loop detectors and probe buses to improve travel time estimations on urban routes. They developed an algorithm to identify congested time periods using detector data which were used later to reproduce bus trajectories. However, they did not consider the delay at signalized intersections within their study area.

Proposed methodology in the current research considers delay at signalized intersections and implements principals of traffic engineering to combine probe and detector data to reproduce vehicle trajectories on urban arterials. The next section introduces study area, available data and analysis period.

3. Study area and available data

As shown in **Figure 1**, a 1450m stretch of an urban route in Tokyo was considered for the purpose of this study. The study area is a portion of the test bed dedicated for advanced research in real time traffic management by VICS (Vehicle Information and Communication System) center in Japan. This route is a two-way, single lane and includes six signalized intersections. Signal timing parameters, loop detector records and probe data were collected on October 28, 2009 from 07:00 am to 19:00pm. However, for the purpose of this study a shorter analysis period is chosen for the morning peak period from 08:14 to 08:45. Traffic volumes (pulse data) were collected using loop detector installed in the upstream and downstream of intersections. Signal timing parameters including the duration of each phase and cycle length for all signalized intersections were also available. Probe data include

*Keywords: Data fusion, Probe data, Signal coordination
**Master course student, Institute of Industrial Science, the Univ. of Tokyo (4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Tel: 03-5452-6419, FAX: 03-5452-6420, Email: neatly60@iis.u-tokyo.ac.jp)
***Member, Dr. Eng., Institute of Industrial Science, the Univ. of Tokyo (4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Tel: 03-5452-6419, FAX: 03-5452-6420, Email: babak@iis.u-tokyo.ac.jp)
****Member, Professor, Dr. Eng., Institute of Industrial Science, the Univ. of Tokyo



a) Westbound **b)Eastbound**

Figure 1: Study area and location of signalized intersections and loop detectors

time and travelled distance (position) at 1 second intervals were collected using GPS equipped vehicles. The travelled distance of probe vehicle was measured from a start kilo post (origin) of the study area. Location of intersections, signal timing patterns and probe trajectories of east bound and west bound route are shown in **Figure 1(a) and 1(b)**.

4. Methodology

Proposed methodology to estimate vehicle trajectories is based on the kinematic wave theory originally developed by Lighthill, Whitham and Richards (LWR) in 1950s [7)]. They described a theory of one-dimensional wave motion which could be applied to certain types of fluid motion or to highway traffic flow. The key feature of the LWR theory was that there is some functional relation between the flow q and the density k which might vary with location x but not with time t. Assuming no entering or exiting traffic, the conservation equation implies:

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \tag{1}$$

Equation (1) can be written as:

$$w(q(x,t),x)\ \frac{\partial q(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \tag{2}$$

where, $w(q,x) = \partial k(q,x)/\partial q$ and the 1/w is called the "wave speed".

Newell [8, 9 and 10)] combined the concept of cumulative curves with LWR theory and extended it to the 3D kinematic wave theory in 1993. Given N(x, t) as cumulative number of vehicles at location x and time t, Newell suggested evaluating N(x, t) rather than q(x, t):

$$\begin{aligned} dN &= (\partial N/\partial x)dx + (\partial N/\partial t)dt \\ &= -kdx + qdt = (-k+qw)dx \end{aligned} \tag{3}$$

According to equation (3), if one knows $N(x_0, t_0)$ and $q(x_0, t_0)$ at some boundary point, one can determine $N(x, t)$ at all points along the same cumulative curve. In the special case of a homogeneous road section, k, q and w are all constant and, as a result of equation (3), so is dN/dx. If we interpret $N(x, t)$ as a surface in a three-dimensional (N, x, t) space, the surface $N(x, t)$ is a surface generated by a family of straight lines. The value of $N(x_0, t_0)$ and $q(x_0, t_0)$ at any boundary point would determine a line in the surface $N(x, t)$. From appropriate boundary data one could easily construct such a surface and, thereby, determine $N(x, t)$ everywhere.

The 3D kinematic wave theory has been used for several applications such as dynamic network assignment [11)], traffic simulation [12)], etc. Recently Daganzo [13 and 14)] proposed an efficient calculation method based on variational formulation of the 3D kinematic wave theory to evaluate forward and backward waves within the traffic flow. We employed Daganzo's method to estimate vehicle trajectories and validated the proposed methodology on the test bed of this study.

Given $N(x_0, t_0)$ and $q(x_0, t_0)$ at some boundary point, variational theory provides a solution to estimate $N(x, t)$ by treating the problem as capacity constrained optimization problem. Flow at any point in time-space is bounded from above by q_{max}, the capacity. A similar capacity constraint should also hold if the road is viewed from a rigid frame of reference that moves with speed x'. In this case the capacity relative to the frame (the "relative capacity") is the maximum rate at which traffic can pass an observer attached to the frame. An observer that moves with speed x' next to a traffic stream with density k and flow q is passed by traffic at rate $q\text{-}kx'$. In simplified variational theory, fundamental diagram is triangular and the relative capacity function (cost function) is linear. In this case the maximum cost is:

$$r = (1 + w/u)q_{\max} \tag{4}$$

where u and w are forward wave speed and backward wave speed, respectively. In variational theory the problem is solved approximately by overlaying a dense but discrete, network with short straight valid paths as links, and the following two properties: *i*) the slopes of links branching from each node which represent range of wave speeds with sufficient resolution and ii) link costs.

As shown in **Figure 2**, the solution domain can be modeled with a network at which the mesh resembles triangular fundamental diagrams with identical steps in time-space plane. Each node in the network represents

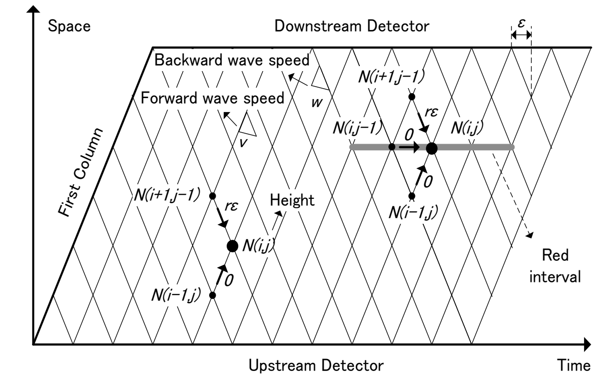


Figure 2 Solution network

the height of the three dimensional cumulative surface at a specific point. Cost rate along forward wave is 0, which means there is no change in cumulative number of vehicles along forward wave. Cost rate along backward wave is the maximum allowable change in cumulative number of vehicles ($r.\varepsilon$). If there is any red interval in the solution domain, the cost rate of that path becomes 0. In this case, red interval creates a shortcut in the network along which cumulative number of vehicles is constant.

Considering the three dimensional shape of the cumulative surface, trajectory of vehicle *i* would be the contour level *i* on the three dimensional surface where *i* is an integer number. For the test bed of this study, forward wave speed was estimated from probe data during free-flow conditions while saturation flow rate was estimated using detector data.

It is assumed that there are no vehicles entering or leaving the study area from midway intersections. Proposed methodology is applicable when there is no passing allowed (first-in, first-out conditions).The test bed of this study is a single lane facility. Saturation flow rate, jam density, free-flow-speed and passing times of the vehicles were estimated this single lane.

5. Calculation steps

(1) Constructing the network

The process starts with defining the network. Considering a triangular fundamental diagram, given the forward wave speed u, jam density k_j and saturation flow rate q_{max}, backward wave speed w is estimated using equation (5). The horizontal distance between nodes, *tstep*, is an input variable. Considering computational straightforwardness, time step of 1 second is recommended. The vertical distance between network nodes, *sstep*, can be estimated using equation (6).

$$w = \frac{q_{\max}}{k_j - \frac{q_{\max}}{u}} \qquad (5)$$

$$sstep = \frac{u \times w \times tstep}{u + w} \qquad (6)$$

(2) Setting boundary conditions

Each node in the network presents the height of the cumulative surface. To set initial conditions, the height along the first column is assumed to be 1. The next step is assigning appropriate heights for the network boundaries. Considering the passing times of the vehicles recorded by detectors, cumulative traffic counts at the upstream and downstream are assigned as the height to the nodes along the lower and upper boundaries of the network.

(3) Reference probe trajectories

When probe trajectories are used as reference to estimate other trajectories, additional treatments are necessary. A probe trajectory can be interpreted as a contour on a three dimensional surface of cumulative curves. Therefore a constant height should be assigned to the nodes in the vicinity of the probe trajectory. Since cumulative vehicle counts at upstream are known a constant height can be estimated once intersection point of the probe trajectory with the lower boundary of the network is known.

(4) Signal timing constrains

Red intervals create shortcuts in the network as shown in **Figure 2**. In other words, during red intervals, it would be possible to move from node (*i*, *j*-1) to node (*i*, *j*) without incurring any cost. Therefore it is necessary to distinguish all the nodes in the network that represent red intervals. Since positions of signalized intersections are known, considering signal timing data, such a process can be accomplished.

(5) Optimization

Optimization is performed to find the height of each node in the network. Considering equations (4) and (7) the link cost can be estimated using equation (8). Optimization process begins from upstream node and continues along each column until the last node in downstream.

$$\varepsilon = \frac{u}{u + w} \times tstep \qquad (7)$$

$$link\ cost = r\varepsilon = q_{\max} / tstep \qquad (8)$$

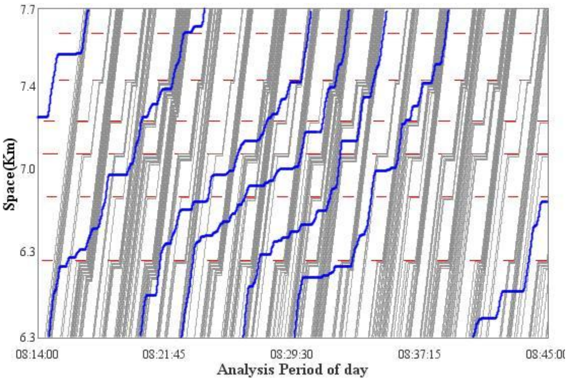
In general, height of node (*i*, *j*) is estimated from equation (9). If node (*i*, *j*) represents red intervals in the time-space plane, then considering the shortcut effect of red intervals, height is estimated from equation (10).

$$N(i,j) = Min[N(i-1,j), N(i+1,j-1) + link\ cost] \qquad (9)$$

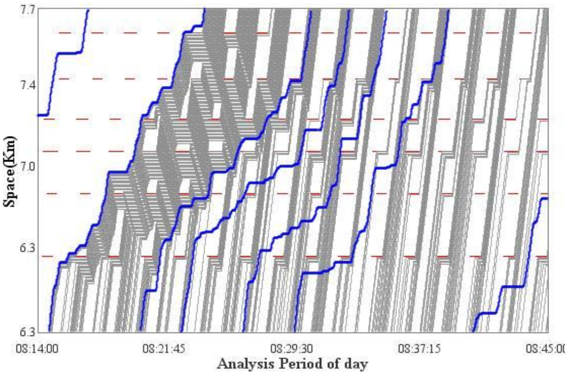
$$\begin{aligned} N(i,j) = Min[&N(i-1,j), N(i,j-1), \\ &N(i+1,j-1) + link\ cost] \end{aligned} \qquad (10)$$

(6) Trajectory Estimation

Considering the three dimensional shape of the cumulative surface, trajectory of vehicle *i* would be the



a) No probe trajectory as reference



b) First full probe trajectory as reference

Figure 3 Estimated trajectories (Westbound)

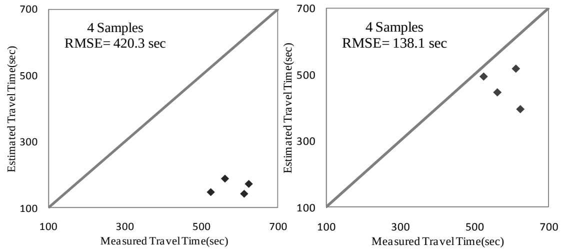
contour level *i* on the three dimensional surface where *i* is an integer number. For the test bed of this study, forward wave speed was estimated from probe data during free-flow conditions while saturation flow rate was estimated using detector data.

6. Analysis Results

Figure 3a), shows estimated trajectories when none of probe trajectories were used as reference. Although forward and backward waves are modeled correctly, estimated trajectories still do not match with corresponding probe trajectories. Since height of the first column in the network was set to 1, impact of traffic conditions from earlier time intervals is not reflected in estimated trajectories.

Since a probe trajectory reflects real traffic conditions, using a probe trajectory as a reference to set up initial boundary conditions would improve accuracy of estimations. **Figure 3b)** shows estimated trajectories when only the first full probe trajectory is used as reference. Impact of backward waves from earlier time intervals is clear in **Figure 3b)**.

Considering significant improvements in the accuracy of estimated trajectories after using the first full probe trajectory as a reference, each probe trajectory was used as a reference to estimate trajectories of the proceeding vehicles. Afterwards the estimated trajectory corresponding to the next probe trajectory was considered and its travel time was compared with the travel time of the respective probe vehicle. **Figure 4**



(a) Without any probe as reference (b) With all probes as reference

Figure 4 Travel time estimation improvements after using probe trajectories as reference

shows the comparison results. While proposed method underestimates travel times without using probe trajectories as reference, when probe trajectories are used as reference, estimated travel times significantly improve and estimation error becomes less than before.

After model validation, application of proposed methodology for signal coordination is briefly demonstrated in the next section.

7. Application for Signal Coordination

Proposed methodology was validated for travel time estimation considering probe trajectory as reference. Hereafter application of proposed methodology for signal coordination is demonstrated. The main objective is to minimize delay on both directions. As shown in **Figure 1** it is clear that westbound route suffers more from extended delays compared with the eastbound. This may be due to bad signal coordination or heavier traffic demand in this direction. Loop detector data shows traffic volume in the westbound is just slightly higher than that in the eastbound. So it can be expected that a better coordination result in less delays in both directions. Original cycle length for all signals for this study period is 120 sec but range of green split is varied in between 55% to 69%. In order to have better coordination, offset value of each downstream signal with respect to upstream signal for both direction is estimated separately keeping the original cycle length and green splits for each signal.

First trial for offset change is started with ideal offset that means dividing distance in between to intersection by free flow speed. And as estimated vehicle trajectories are already been obtained so it is easy to see visually where and how many vehicles which are coming form upstream intersections are stopped by the red interval at downstream intersection. So by a little effort we can obtain optimum offset for each intersection regarding to the upstream intersection.

As it is mentioned before offset setting is done separately for both direction. **Table 1** shows the strategies implemented regarding different offsets for the eastbound.

For each strategy average travel time is estimated using estimated trajectories within the zones shown in **Figure 5(a)** and **Figure 5(b).** Then average travel time is compared with free flow travel time of the corresponding sections and delay is calculated.

Offset values shown in **Table 1** are implemented one by one for both east and west directions. Average delay

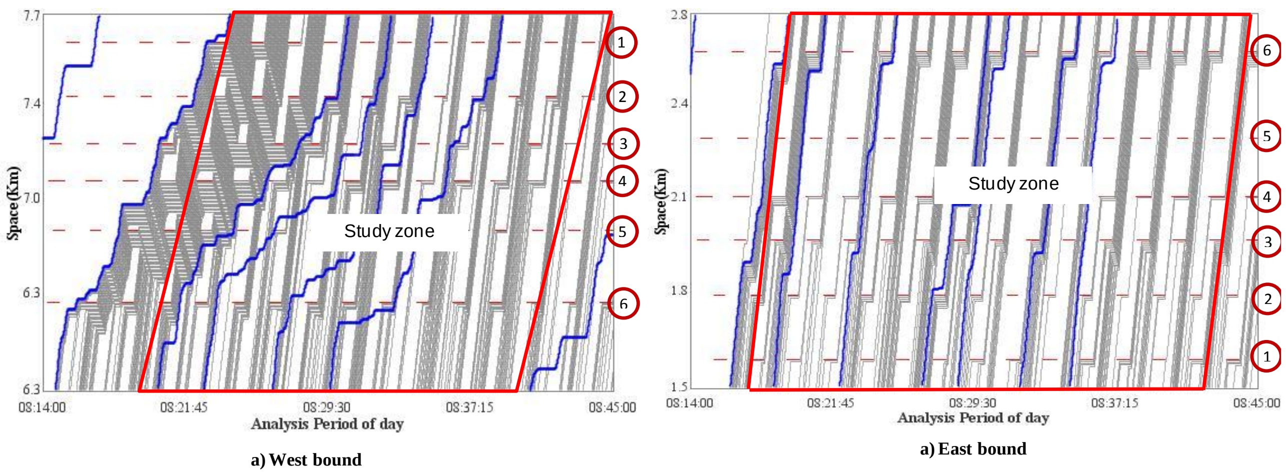


Figure 5: Average travel time estimation zone using first full probe as reference for both routes

for each direction and total delay for both directions are calculated for each Strategy. **Figure 6** shows that strategy E is the most efficient offset setting for both routes as total delay is the minimum.

8. Conclusions and future works

Most of existing data fusion techniques use probe trajectories to simply estimate travel time and neglect the rich trajectory data which represent traffic conditions in time and space. More over majority of data fusion techniques are merely based on statistical methods without considerations for traffic engineering theories. In contrast, proposed methodology is based on 3D kinematic wave theory of traffic flow and can fully employ probe trajectory data to estimate vehicle trajectories. Such features culminate in reliable travel time estimation with minimal input data requirements. In addition, estimated trajectories can be used for various applications such as travel time estimation and dynamic signal coordination.

The methodology can still be improved and generalized in several ways. In this study vehicles coming in or going out in the middle of the section were not considered. Neglecting incoming or outgoing vehicle may result in significant over-underestimation. It was also assumed that first in-first out (FIFO) condition holds throughout the study area. Such an assumption might not be true for multilane facilities where passing is allowed. Tackling these issues require adjustment of the fundamental theory as well as introducing extensions to the proposed methodology which are considered in the future directions of this research. In addition, signal coordination process demonstrated in this paper was based on an iterative procedure which does not guarantee optimum coordination in both directions. Development of an efficient signal coordination algorithm is another issue which should be considered in the future extension of the current research.

Table 1 Offset values between consecutive intersections on eastbound (seconds)

Strategy	Int. # 2	Int. # 3	Int. # 4	Int. # 5	Int. # 6
A	18				
B	18	20			
C	18	20	16		
D	18	20	16	17	
E	18	20	16	17	9

Acknowledgments

Authors are very grateful to the VICS center, and i-Transport Lab for their kind assistance in providing them with valuable data necessary for this study.

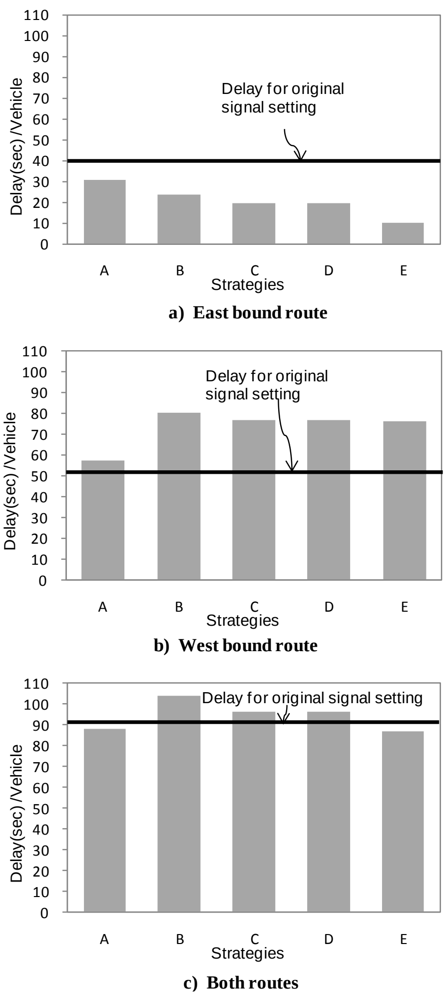


Figure 6: Directional and total delay for each strategy

References

1) EL Faouzi, N. E.; "Data Fusion in Road Traffic Engineering: An Overview, Proceedings of Multisensor", Multisource Information Fusion: Architectures, Algorithms, and Applications, Vol. 5434, pp.360~371, Orlando, 2004
2) Choi, K. and Chung, Y.; "A data Fusion Algorithm for Estimating Link Travel Time", ITS Journal: Intelligent Transportation Systems Journal, Vol. 7, pp.235~260, 2002
3) EL Faouzi, N. E.; "Bayesian and Evidential Approaches for Traffic Data Fusion: Methodological Issues and Case Study", TRB 85th Annual Meeting, Washington D.C., 2006
4) El Faouzi, N. E., Klein, L. A, De Mouzon, O.; "Improving Travel Time Estimates from Loop and Toll Collection Data with Dempster-Shafer Data Fusion", TRB 88th Annual Meeting, Washington D.C., 2009
5) Ivan, J. N., Schofer, J. L., Koppelman, F. S. and Massone, L. L. E.; "Real-Time Data Fusion for Arterial Street Incident Detection Using Neural Networks", Transportation Research Record, Vol. 1497, pp. 27~35, 1995
6) Berkow, M., Monsere, C. M., Koonce, P., Bertini, R. L. and Wolfe, M.; "Prototype for Data Fusion Using Stationary and Mobile Data Sources for Improved Arterial Performance Measurement". TRB 88th Annual Meeting, Washington D.C., 2009
7) Lighthill, M. J. and Whitham, G. B.; "On kinematic waves. I: Flood movement in long rivers. II: A theory of traffic flow on long crowded roads", Proceedings Royal Society, London, Vol. A229, pp.281~345, 1955
8) Newell, G.F.; "A simplified theory on kinematic wave in highway traffic, part I: general theory", Transportation Research Part B, Vol. 27(4), pp.281~287, 1993
9) Newell, G.F.; "A simplified theory on kinematic wave in highway traffic, part II: queuing at freeway bottlenecks", Transportation Research Part B, Vol. 27 (4), pp.289~303, 1993
10) Newell, G.F.; "A simplified theory on kinematic wave in highway traffic, part III: multi-destination flows", Transportation Research Part B, Vol. 27 (4), pp.305~313, 1993
11) Kuwahara, M. and Akamatsu, T., "Dynamic equilibrium assignment with queues for a one-to-many OD pattern", Proceedings 12th International Symposium on Transportation and Traffic Theory, pp. 185–204 Berkeley, 1993
12) Daganzo, C.F.; "The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory", Transport. Res. B 28, pp. 269–287, 1994
13) Daganzo, C.F.; "A variational formulation of kinematic waves: Solution methods", Transportation Research Part B, Vol. 39, pp. 934~950, 2005
14) Daganzo, C.F.; "A variational formulation of kinematic waves: basic theory and complex boundary conditions", Transportation Research Part B, Vol. 39 (2005), pp.187~196, 2005