A DATA FUSION TECHNIQUE TO ESTIMATE TRAVEL TIME FROM SPARSE AVI AND PROBE DATA ON URBAN STREETS [*]

by Julien MONTEIL[**], Babak MEHRAN[***] and Masao KUWAHARA[****]

1. Introduction

Availability of travel time information is of great importance for congestion management on urban streets. Continuous travel time information is useful for trip planning and real-time routing. In addition, it enables traffic engineers to dynamically evaluate efficiency of different signal timing plans in order to minimize overall delay and emission on urban arterials.

Travel time on urban streets can be estimated directly using probe sensors (e.g. probe vehicles) or indirectly from fixed sensor data (e.g. loop detectors). Traffic data from various sources have different accuracies and limitations. For instance, probe vehicles can provide spatial traffic information and direct measurements of travel time. However, their frequency is restricted and there is always GPS communication errors included in the data. On the other hand fixed sensors can record traffic data contentiously but only at fixed locations.

Alternatively, data fusion techniques can be applied to increase accuracy, reliability and robustness of travel time estimation. Fusion of traffic data from multiple sensors results in both qualitative and quantitative benefits such as extended spatial and temporal coverage, increased confidence and reduced uncertainty.

Most existing fusion techniques to estimate travel time on urban streets are merely based on statistical methods without considerations for traffic engineering concepts. Moreover, despite the significance of signal timing parameters, only few fusion methodologies consider signal timing parameters as inputs for travel time estimation. Probe data contain much richer information of vehicle trajectories than solely 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 information could be combined more effectively with fixed sensor data to reproduce trajectories of all vehicles. This paper aims to establish a novel fusion technique to reproduce vehicle trajectories according to signal timing parameters and concepts of traffic engineering which allows reliable estimation of travel time on urban streets.

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[2)] describes the 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.

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

3. Study area and available data

As shown in **Figure 1**, a 521m stretch of an urban route between Fukazawafudou intersection and Komazawakoen intersection in Tokyo was considered for the purpose of this study. This route is a single lane facility and includes four signalized intersections and several un-signalized intersections. The road segments between consecutive signalized intersections are referred to as link A, link B and link C.

---

[*]Keywords: *Travel time estimation, Data fusion, Probe data, AVI data*
[**]Master course student, Université de Lyon ENTPE, (3, rue Maurice Audin 69518 Vaulx-en-Velin Cedex, Tel: (+33) 4-7204-7070, Email: julien.monteil@bvra.entpe.fr)
[***]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

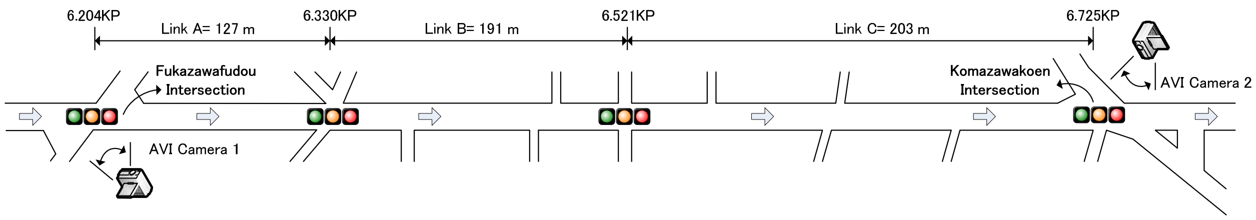

Figure 2: Study area and location of signalized intersections and AVI cameras

A pair of AVI cameras was installed at the entrance and exit points of the study area which are named as AVI 1 (upstream) and AVI 2 (downstream). The number plate of vehicles entering or leaving the study area was recorded by AVI cameras and matched to estimate travel time during analysis period. AVI data before and after matching were available.

There were 20 probe taxis used to collect probe data on the study area. Probe data include time and travelled distance (position) at 1 second intervals. The travelled distance was measured from a parking lot (origin) used by probe taxis in the vicinity of the study area. The positions of signalized intersections are represented relative to the origin in **Figure 1**.

Signal timing parameters including the duration of each phase and cycle length for all signalized intersections were provided by Tokyo Metropolitan Police Department.

Signal timing parameters, AVI records and probe data were collected on September 1st, 2006 from 6:16 am to 11:41 am. However, for the purpose of this study a shorter analysis period from 7:13 am to 7:26 am was considered.

4. Preliminary analysis

Figure 2 represents vehicle trajectories derived from taxi probe and AVI data. Since cycle length is varying at each intersection most of the times, duration of green and red intervals were extracted from signal timing data for each cycle and considering the position of intersections, signal timings are presented graphically in **Figure 2**. There is a 3 second yellow interval before red which is not presented in **Figure 2**. For simplification, the first 2 seconds of the yellow interval is treated as green and the last 1 second is considered as red.

AVI trajectories does not match with signal timings as they were just drawn using the entrance and exit times recorded by AVI cameras.

Since probe data were available every 1 second, it was possible to draw contentious probe trajectories. As it is indicated by small circles in **Figure 2**, there are some inconsistencies between probe trajectories and signal timings as some probe vehicles pass signalized intersections during the red intervals. Such inconsistencies are due to GPS positioning error which has a random characteristic. Modification of these errors will be discussed later in this paper.

Considering **Figure 2**, it is obvious that existing data are insufficient to estimate travel times continuously. Sometimes there are no probe vehicles passing during a whole cycle. On the other hand, the matching rate of AVI cameras are low and AVI data can be used only to estimate travel times between AVI cameras. It is not possible to estimate travel times on smaller portions of the study area using AVI data.

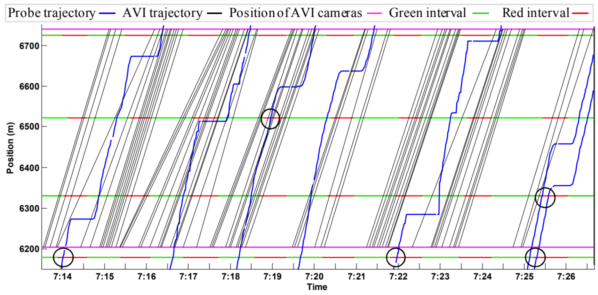

Figure 1：Probe and AVI trajectories

The main concept of the proposed methodology is to combine probe and AVI data according to principles of traffic engineering and signal timing parameters in order to define the stopping and starting shockwaves near signalized intersections. Afterwards, queuing vehicles and delay zones upstream of signalized intersections are identified and the trajectories of all vehicles are reproduced. Availability of vehicle trajectories enables estimation of travel times between different locations on the study area and provides an opportunity to adjust positioning errors included in probe data.

5. Methodology

Figure 3 shows the general framework of the proposed methodology. The basic assumption is that traffic condition is under-saturated and vehicles travel at free flow speed (FFS) on each link. Since number plate records from AVI 1 and AVI 2 cameras are available, it is possible to know the time at which vehicles enter or leave the study area. As a result, considering the FFS for each link and entrance time of vehicles, parallel trajectories are drawn from the entrance point (location of AVI 1 camera). FFS at each link is estimated using probe data. The average FFS for link A, link B and link C is estimated to be 29.2 km/h, 29.6 km/h and 34.6 km/h respectively. The FFS is relatively low and slightly different at each link considering the length of the link, geometry and number of un-signalized intersections.

It should be noted that all the vehicles recorded by AVI 1 camera does not necessarily reach the end of the study area at AVI 2 location and they may leave the route midway. On the other hand, some new vehicles may enter the study route from the intersections. However, since such data are not available, it is assumed that all entering vehicles continue their trip until the end of the route and do not leave the study area. Moreover, it is assumed that there are no queues caused by turning vehicles at any of un-signalized intersections on the study area.

Parallel trajectories are continued until the first signalized intersection. Considering signal timings shown in **Figure 2**, there are two possibilities for individual trajectories: *i*) vehicle trajectory reaches the

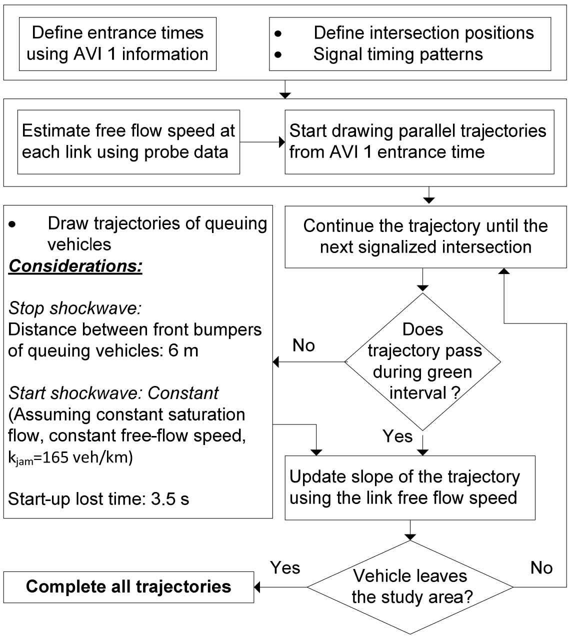

Figure 3：Framework of the proposed methodology

intersection during green interval while there is no queue, and *ii*) vehicle trajectory reaches the intersection during red interval or when there is a queue. The only treatment for trajectories crossing the intersection during green interval (no queue) is to update their slope according to downstream link FFS. In contrast, trajectories arriving at the intersection during red interval or when there is a queue; cannot pass and should wait until the beginning of green or dissipation of existing queue.

To draw queuing vehicles' trajectories during red interval stopping shockwave, start-up lost time and starting shockwave is considered. For stopping shockwave, the jam density (k_j) is assumed to be 165 veh/km. In other words, the distance between the front bumpers of queuing vehicles is assumed to be 6 m. Free flow condition trajectories are terminated at the intersection stop line (for the first arriving vehicle) or 6 m before the preceding queuing vehicle, and the slop of the trajectory is changed to zero (horizontal lines) to reflect stationary queuing conditions.

The start-up lost time which is the additional time, consumed by vehicles in a queue because of the need to react to the initiation of the green phase and to accelerate is assumed to be 3.5 s.

The starting shockwave (u_s) is estimated using the following equation:

$$u_s = \frac{SF}{k_d - k_j} \qquad (1)$$

where, SF is saturation flow rate, k_d is traffic density downstream the intersection and k_j is jam density (165 veh/km). The average saturation flow rate was estimated to be 1770 veh/h using AVI 1 and AVI 2 data during green intervals. k_d was estimated through dividing the saturation flow rate by the free flow speed on the link downstream on each intersection. Since all independent variables in equation (1) are constant, the staring shockwave is constant for each signalized intersection.

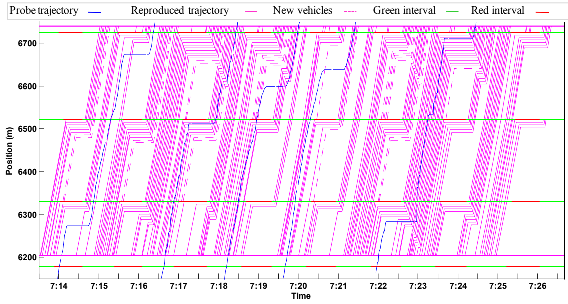

Figure 4：Reproduced vehicle trajectories compared with probe trajectories

After defining the starting shockwave, trajectories can be continued from the beginning of the next green interval (considering start up lost time). Since vehicles are entering a new link, slope of trajectories should be updated considering the FFS on the downstream link.
The procedure is repeated for all entering vehicles until they reach the exit point of the study area.

6. Outputs and validation

Figure 4, represents vehicle trajectories reproduced according to the proposed methodology. Free flow areas, queuing areas, starting and stopping shockwaves are conspicuous in **Figure 4**.

Since the lengths of the links between signalized intersections are short (**Figure 1**), sometimes the arriving traffic volume to downstream intersection is as same as the saturation flow of the upstream intersection. As a result, starting shockwaves and stopping shockwaves are parallel in most of the cases.

As mentioned earlier, there was no information available about the vehicles which leave the study route midway. However, comparing AVI 1 and AVI 2 records, it was possible to identify some new vehicles entering the route from midway intersections. It was assumed that new vehicles are entering the study area only from signalized intersections. Trajectories of these new vehicles are represented by dotted lines. Considering the exit time of new vehicles recorded by AVI 2, trajectories of new vehicles were inspected reversely to find the most likely intersection from which they entered the route. Afterwards, new trajectories were added according to the proposed methodology starting from the entering intersection.

Performance of the methodology regarding travel time estimation is evaluated in **Figure 5**. There were 67 vehicle records matched between AVI 1 and AVI 2 cameras during analysis period. After matching, the travel times corresponding to matched vehicles were estimated. Since trajectories of matched vehicles were also reproduced in **Figure 4**, considering the exit time of reproduced trajectories, travel times of matched vehicles were estimated from **Figure 4** and compared with measured values from AVI cameras. The Root Mean Square Error (RMSE) is estimated to be 5.7 s which is not far from expectations considering several assumptions made in the proposed methodology. According to **Figure 5**, proposed methodology slightly underestimated travel times. The possible reasons for such inaccuracies are discussed here.

There are several un-signalized intersections within the studied rout. Considering that the studied route is a

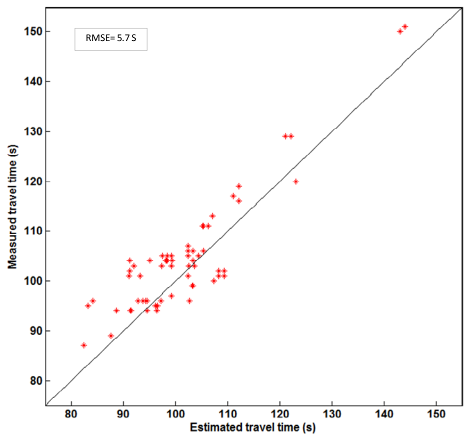

Figure 5：Performance of proposed methodology

single lane facility, turning vehicles drastically affect the movement of through traffic. However, it was assumed that vehicles travel at free flow speed along each links.

7. Adjustment of probe trajectories

There are some inconsistencies between probe trajectories and reproduced trajectories in **Figure 4**. As it was explained in section 4, most of the inconsistencies are due to GPS positioning errors included in probe data. In addition to GPS positioning errors, other measurement errors might be included in the probe data used for the purpose of this study. While collecting probe data a predefined algorithm was used to generate position data when there were communication errors. In the algorithm, the diameter of the probe car's tire was used to measure travelled distance when position data cannot be retrieved due to GPS communication errors. Estimated travelled distance at each 1 second was added to the last position value to produce position data dynamically. It is obvious that the tire diameter can be change according to tire air pressure which depends on many factors such as temperature. This kind of positioning error is cumulative and increase by time.

To make a better use of probe data for further research, probe trajectories were adjusted according to reproduced trajectories. As presented in **Figure 6**, inconsistent probe trajectories were first shifted vertically to the nearest appropriate location where horizontal portions of the probe trajectory matches with reproduced queuing vehicles' trajectories. Afterwards, if the subject probe trajectory crosses any of the intersections during the red interval, it is horizontally shifted to the left until the end of the green interval which is the nearest likely point at which the probe vehicle could cross the intersection.

8. Conclusions and future directions

Most existing data fusion methods for travel time estimation are based on statistical theories and do not consider traffic engineering principles to make a better use of available data. On the other hand, intersection delay and signal timing parameters are neglected in most of data fusion algorithms.

In this paper, a novel methodology is proposed to fuse AVI and probe data in order to estimate reliable travel times on urban routes with several signalized intersections. The methodology considers signal timing parameters and employs concepts of traffic engineering such as shockwave analysis to reproduce vehicle

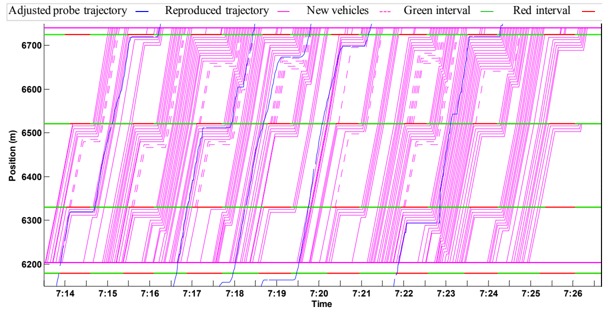

Figure 6：Adjusting probe vehicle trajectories using reproduced trajectories

trajectories. Proposed methodology was validated and showed promising results.

There are still some issues that should be addressed to improve the accuracy and applicability of the proposed methodology. The basic assumption in the proposed methodology is that vehicles travel at free flow speed on links between signalized intersections. Yet, such an assumption is not realistic and probe data could be used more efficiently to adjust travel speed on each link. In this research probe data were available at each 1 second. However, in practice probe data are collection in much longer intervals (e.g. 1 minute). The analysis period only includes under-saturation conditions which is not the case in most of busy urban routes during peak periods. Queue spill back from downstream intersections could be considered to expand the applicability of the methodology.

Once vehicle trajectories are reproduced on an urban corridor, in addition to travel time estimation, it is possible to evaluate efficiency of different signal timing patterns to minimize delay. Moreover, emission levels can be estimated given vehicle trajectory data. Then signal timing parameters can be optimized to minimize overall emission on a subject urban corridor.

Acknowledgments

The authors are very grateful to DENSO Co. Ltd. and Tokyo Metropolitan Police Department for their kind assistance in providing them with valuable data necessary for this study.

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, 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*, 7, 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* 1497, 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.