A REAL-TIME DATA FUSION FRAMEWORK FOR CORRIDOR TRAVEL TIME ESTIMATION WITH MULTIPLE DATA SOURCES

Xuechi Zhang[1]; Masoud Hamedi, PhD[2*]; Ali Haghani, PhD, Prof[1], Yeming Hao[1]

[1]Department of Civil and Environment Engineering, University of Maryland, College Park;

[2]Center for Advanced Transportation Technology (CATT), Maryland, masoud@umd.edu.

UNIVERSITY OF MARYLAND

17-04974

ABSTRACT

Due to technological improvements in recent years, the available independent data sources for reporting travel time on freeway and arterial corridors has increased. Combining and fusing data from different sources and turning data into useful information is an important part of any integrated transportation system. A successful travel time fusion scheme improves real-time data coverage and accuracy, and is capable of reporting reliability by quantifying underlying uncertainties. This paper presents a multi-source real-time travel time fusion framework, based on Dampster-Shafer evidential theory (D-S theory). Data source credibility and real-time data variance extracted for each heterogeneous data sources are considered as time-dependent variables in the proposed model. The model is independent of underlying technology behind each sources and the number of input sources is not limited. An application of the fusion framework by using real-world data collected from three sources on a Maryland freeway corridor is presented and discussed. The results from case study show that the fusion scheme is successful in merging data along with reliability indicator, and its resolution can be adjusted with respect to the intended applications.

CONTACT

Xuechi Zhang
University of Maryland
Email: xuechi@umd.edu
Phone: 240-893-5119

Travel Time Measurements

- **Identification and Re-identification Method**
  - Use identification and re-identification time stamps of the vehicles to estimate travel time between consecutive detection stations installed along the road;

$$\bar{T}_k = \frac{\sum_{i=1}^{N}(t_i^B - t_i^A)}{N}$$

  - Applications: Bluetooth and WiFi detections.

- **Stationary Detector Based Estimation Method**
  - Use weighted speed data collected at each single point to estimate travel time of the adjacent segment.

$$T_l(k) = \frac{d_1}{v_1(k)} + \sum_{x=2}^{m} \frac{2(d_x - d_{x-1})}{v_x(k) + v_{x-1}(k)} + \frac{L - d_m}{v_m(k)}$$

  - Applications: Loop Detector, and Microwave Radar.

- **Continuous Vehicle Tracking Method**
  - Use average speed data reported by location aware and internet enabled mobile devices to estimate travel time of a particular path.

$$\bar{T}_l(k) = \sum_{i=1}^{n} \frac{1}{|S_i|} \sum_{t \in S_i} \hat{T}_i(t)$$

  - Applications: Probe travel time archived by INRIX.

Dempster-Shafer Theory Based Fusion Model

Real-time Information Credibility
Recognize real-time data reliability before further fusion (i.e. historical v.s. real-time).

- Re-identification and Tracking Methods:

$$\alpha_t = 1 - (1 - \beta)^{N_t}$$

- Stationary Detector Method:

$$\varphi_t = \begin{cases} 1 \cdot e^{-b \cdot \sigma_k}, & N \geq N^* \\ e^{-a\left(\frac{L}{N} - \frac{L}{N^*}\right)} \cdot e^{-b \cdot \sigma_k}, & N < N^* \end{cases}$$

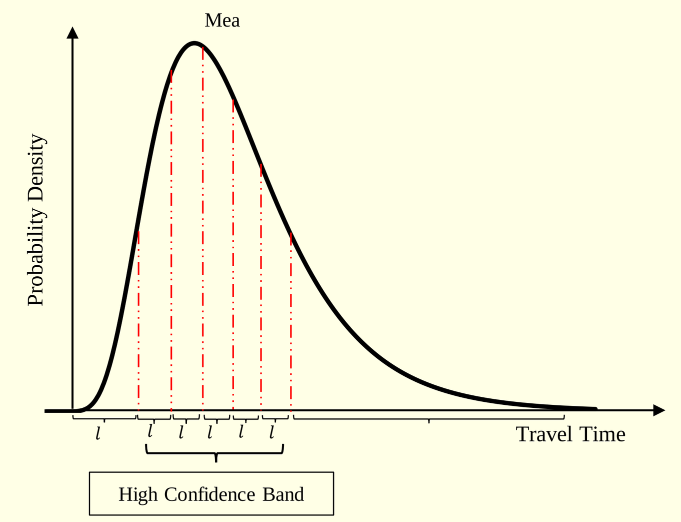


Figure 1. Distribution Based Proposition Generation.

- Dempster-Shafer evidential theory, is a probability based data fusion scheme, also called D-S theory.
- .
- D-S Theory based Fusion Model:

$$m(l_q) = K \sum_{\forall a_i^{x1} \in S_{x1}} \sum_{\forall a_i^{x2} \in S_{x2}} \cdots \sum_{\substack{\forall a_i^{xm} \in S_{xm} \\ a_i^{x1} \cap a_i^{x2} \cap \cdots \cap a_i^{xm} = l_q}} m_{x1}(a_i^{x1}) \cdot \cdots \cdot m_{xm}(a_i^{xm})$$

$$K^{-1} = 1 - \sum_{\forall a_i^{x1} \in S_{x1}} \sum_{\forall a_i^{x2} \in S_{x2}} \cdots \sum_{\substack{\forall a_i^{xm} \in S_{xm} \\ a_i^{x1} \cap a_i^{x2} \cap \cdots \cap a_i^{xm} = \emptyset}} m_{x1}(a_i^{x1}) \cdot \cdots \cdot m_{xm}(a_i^{xm})$$

$$where, m(a_i) = \omega \cdot [F^r(a_i^u) - F^r(a_i^l)] + (1 - \omega) \cdot [F^h(a_i^u) - F^h(a_i^l)], m_x(\phi) = 0, and \sum_{i=1}^{k} m_x(a_i) = 1$$

Multi-source Travel Time Fusion Framework

For each time window (e.g. 5-min or 15-min):

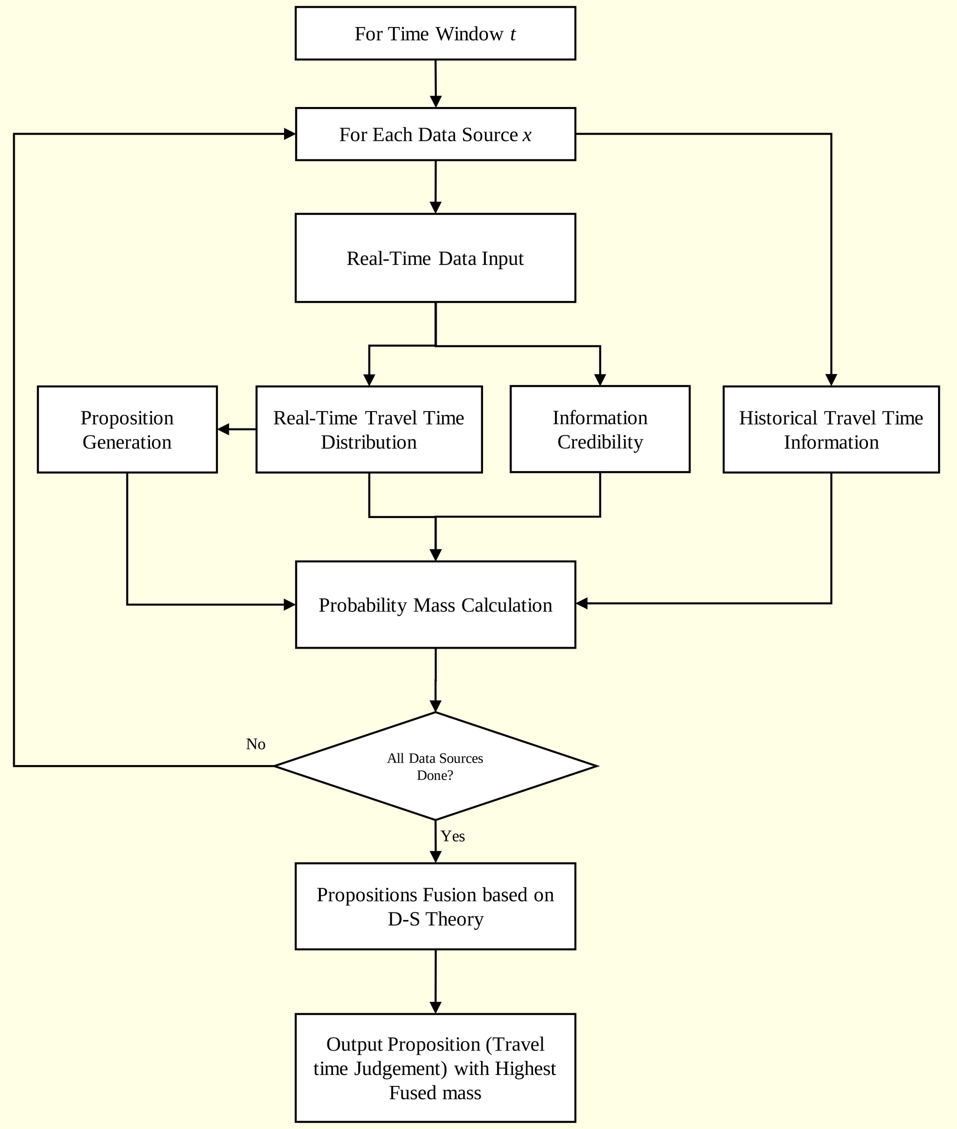
Step 1: At each time window, estimate the travel time distribution and the information credibility function for each data source based on their detection samples.

Step 2: For each data source, generate a set of travel time judgement propositions according to the estimated distributions in step 1. The propositions consists of k high likelihood propositions $\{l_{m1}, l_{m2}, \ldots, l_{mk}\}$ around the mean value and two propositions $\{l_{o1}, l_{o2}\}$ belonging to the low confidence bands representing other possibilities.

Step 3: Calculate the probability masses for the propositions of each data source, by considering both the real-time information and the historical reference and with respect to the information credibility function.

Step 4: Apply the combination rules of D-S evidential theory to calculate the fused probability mass of each optional proposition contributed by all data sources.

Step 5: Output the travel time proposition with the largest fusion mass as the final result, representing the travel time value of the current window.



Estimation Uncertainty

While most applications use a single value as average travel time, the variation associated with the travel time samples within the corresponding interval is very useful to understand and analyze travel time uncertainty.

- Uncertainty for Re-identification and Tracking Methods:

$$\sigma_k^2 = \begin{cases} 0, & N \leq 1 \\ \frac{1}{N-1} \sum_{i=1}^{N} (T_k^i - \bar{T}_k)^2, & N \geq 2 \end{cases}$$

- Uncertainty for Stationary Detector Method:

$$\sigma_k = \begin{cases} NA, & m = 1 \\ \frac{L}{\min\{v_x(k) | x = 1, \ldots, m\}} - \frac{L}{\max\{v_x(k) | x = 1, \ldots, m\}}, & m \geq 2 \end{cases}$$

Unlike the traditional way to represent travel time variation by sampling variance, the above equation conservatively depicts the travel time variation by computing its lower bound and upper bound at a given time period.

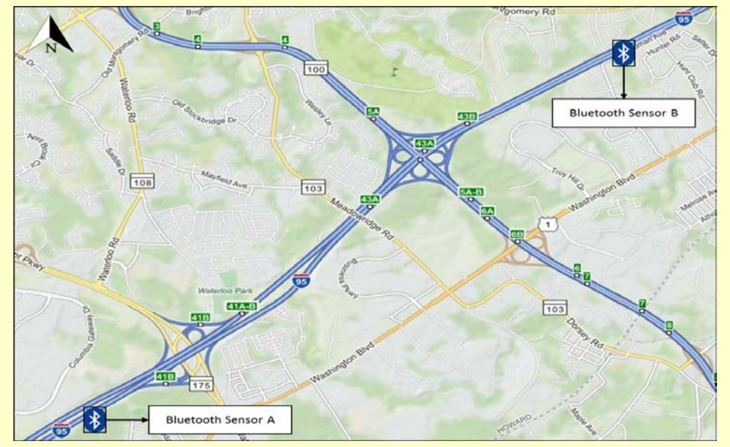


Figure 2. Map of the Study Segment.

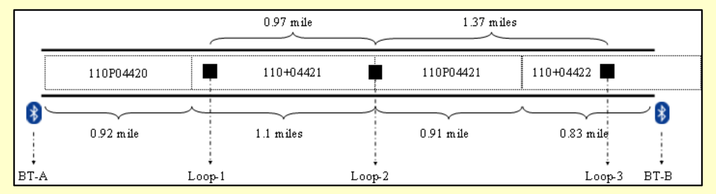


Figure 3. Horizontal Alignment of the Study Segment

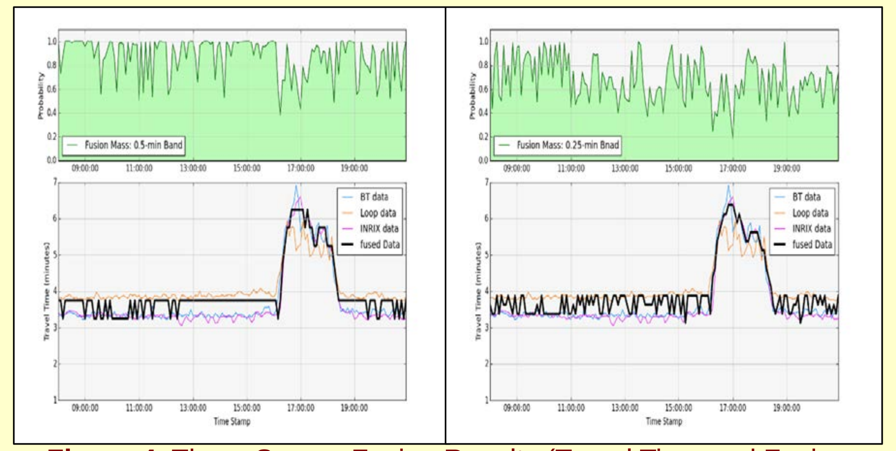


Figure 4. Three-Source Fusion Results (Travel Time and Fusion Score) on Feb 24, 2014: (a) 0.5-min Interval Based Fusion, (b) 0.25-min Interval Based Fusion.

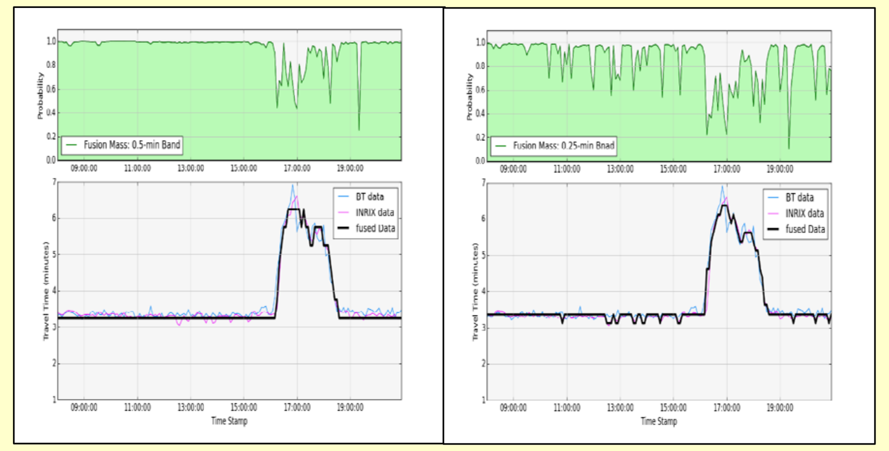


Figure 5. Two-Source Fusion Results (Travel Time and Fusion Score) on Feb 24, 2014: (a) 0.5-min Interval Based Fusion, (b) 0.25-min Interval Based Fusion.

Table 1. Classification Rate of 0.5-Minute Based Fusion Results

Time Period	0.5-min Interval Based Two-Source Fusion Classification Rate (0.5-min Band)		0.5-minute Interval Based Three-Source Fusion Classification Rate (0.5-min Band)		
	Bluetooth	Probe	Bluetooth	Stationary	Probe
0 - 8 AM	66.5%	94.5%	56.7%	27.9%	72.3%
8 - 10 AM	84.6%	93.2%	64.5%	30.3%	69.5%
10 - 12 AM	85.3%	96.9%	64.1%	27.8%	73.6%
12 - 2 PM	85.8%	97.0%	52.1%	34.0%	67.8%
2 - 4 PM	71.0%	92.6%	47.2%	33.2%	67.6%
4 - 6 PM	63.6%	93.0%	54.3%	31.6%	82.6%
6 - 8 PM	76.1%	95.1%	60.0%	39.4%	71.1%
8 - 0 PM	78.6%	97.0%	51.8%	46.3%	52.8%

Table 2. Classification Rate of 0.25-Minute Based Fusion Results

Time Period	0.25-min Interval Based Two-Source Fusion Classification Rate (0.5-min Band)		0.25-minute Interval Based Three-Source Fusion Classification Rate (0.5-min Band)		
	Bluetooth	Probe	Bluetooth	Stationary	Probe
0 - 8 AM	76.3%	95.1%	67.7%	31.2%	76.0%
8 - 10 AM	85.7%	96.2%	64.4%	31.9%	73.1%
10 - 12 AM	87.0%	98.1%	66.5%	30.6%	75.0%
12 - 2 PM	87.0%	98.6%	58.1%	32.6%	70.9%
2 - 4 PM	73.6%	96.9%	50.5%	32.7%	74.7%
4 - 6 PM	65.1%	97.3%	55.5%	33.9%	86.3%
6 - 8 PM	81.4%	97.9%	61.1%	41.4%	75.8%
8 - 0 PM	81.8%	97.4%	59.9%	38.0%	65.1%

CONCLUSIONS

- Proposed a multi-source travel time fusion framework based on Dempster-Shafer (D-S) evidence theory;
- Confidence score evaluating the fused travel time reliability can be statistically generated by the fusion model;
- Real-time information credibility and sampling variation for each type of travel time collection method are formulated and considered;
- More datasets with ground truth value will be used to comprehensively test the proposed data fusion framework.

SELECTED REFERENCES

1. Dion, F., & Rakha, H. (2006). Estimating dynamic roadway travel times using automatic vehicle identification data for low sampling rates. Transportation Research Part B: Methodological, 40(9), 745-766.
2. Shafer, G. (1976). A mathematical theory of evidence (Vol. 1). Princeton: Princeton university press..
3. Zhang, X., Hamedi, M. and Haghani, A., 2015. Arterial Travel Time Validation and Augmentation with Two Independent Data Sources. Transportation Research Record: Journal of the Transportation Research Board, (2526), pp.79-89.