SHORT-TERM TRUCK TRAVEL TIME ESTIMATION USING A ROBUST HYBRID MODEL

Abolfazl Karimpour, Graduate Research Assistant (karimpour@email.Arizona.edu)
Amin Ariannezhad, Graduate Research Assistant (ariannezhad@email.Arizona.edu)
Yao-Jan Wu, Ph.D., P.E. (Advisor) (yaojan@email.Arizona.edu)

Smart Transportation Lab, The University of Arizona

1. Motivation

- Travel time is the key factor for many traffic management and intelligent transportation systems, as well as being the key to understanding the mobility performance measures.
- Travel time is the best indicator of the road transportation system condition.
- Efficient and reliable freight movement is a critical aspect to US economy.
- Accurate travel time estimation for commercial trucks is important for freight scheduling and operations.
- Accurate truck travel time can bring supportive information for both private and public agencies.
- Freight travel time is becoming increasingly important for state's Department of Transportation (DOT). DOTs can use freight travel time to estimate truck mobility performance measures such as truck travel time index, delay, and buffer time.

2. Objectives

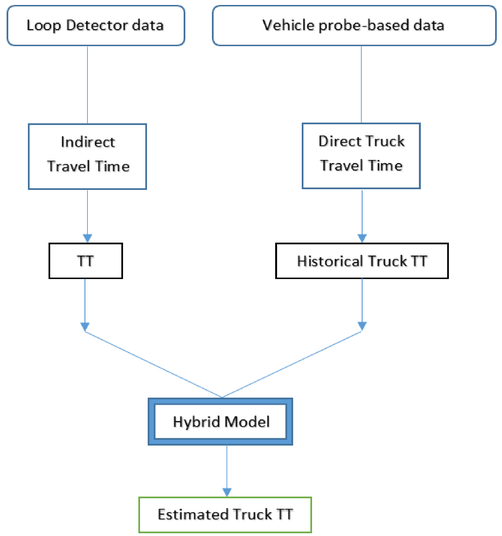
- Developing a new hybrid model to estimate truck travel time for regions where truck data is limited or is not available.
- Estimating truck travel time based on indirect travel time from loop detectors and historical truck travel time.
- The structure of the hybrid model will allow us to incorporate other data sources than loop detectors, which makes the model transferable to other application
- Increasing the robustness of the estimation by incorporating prior knowledge.

3. Backgrounds

- Recent emerging technologies, such as mobile phones and vehicles with Global Positioning Systems (GPS), offer encouraging enhancement in providing traffic travel time information. The data collected from these new technologies are commonly knowns as the probe vehicle data.

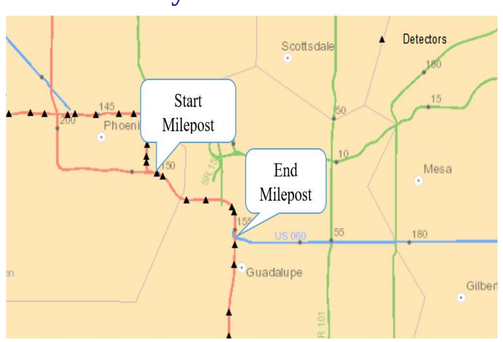
- Many truck companies have their own GPS-based travel time data to track their commercial vehicles, the truck companies are less willing to share their information due to privacy issues.
- Fused data have lots of advantages compared to data from only one source. Using fused data will increase the estimation robustness. Also, the spatial coverage of data from multiple sources is more than using single data sources

4. Methodology



The proposed hybrid model consists of two different modules. The first module is the estimation module, and the second is the prior knowledge module. The estimation task is handled in the first module with the help of recursive least square (RLS) filter. The prior knowledge module is incorporated to increase the estimation accuracy while facing anomalies in the data. Maximum a posterior estimator is used in the second module.

5. Study Site and Data



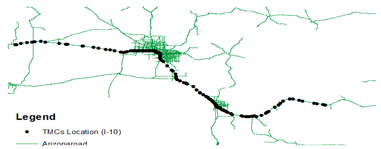
Interstate 10, MP 150 to MP155-EastBound

To verify the effectiveness of the proposed hybrid model, the traffic data collected on Interstate 10 (I-10), Arizona was used in this study. The Eastbound direction on the corridor was selected for three reasons:

1) It has recurring congestion during morning and afternoon peak hours.
2) Non-recurring congestion happens frequently due to high volume traffic leaving the Phoenix central business district (CBD).
3) Two types of data, the loop detector data and probe vehicle data, can be collected concurrently.

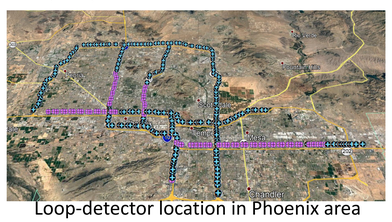
Two types of data are used in the proposed hybrid model:

Probe-Vehicle-Based Travel Time: The National Performance Management Research Dataset (NPMRDS) is provided by FHWA. This dataset provides vehicle probe-based travel time data for vehicles and trucks separately. The travel time information is broadcast by Traffic Message channels (TMCs). The TMCs are located approximately on average one TMC per mile in urban areas



TMCs location on Interstate 10

Loop-based Travel Time: Dual loop detectors are placed approximately one mile apart on each lane. Traffic flow, average speed and occupancy are collected at a 20-sec intervals.

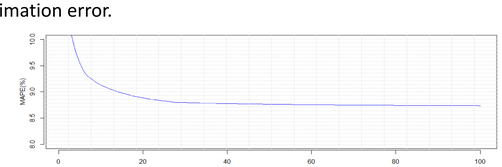


Loop-detector location in Phoenix area

6. Experimental Results

Calibrating Initial Model Coefficients

The performance of the RLS algorithm is closely associated with the initial values chosen for the model. To illustrate the importance of initial values, the model is run for 100 times to obtain the estimation error.



Model Performance

The trained model is used to estimate average peak hour truck travel time for April 12 through April 30, 2016 , except the weekends. The estimated values are then compared to the real truck travel time calculated from probe-vehicle data.

Date		12-Apr	13-Apr	14-Apr	15-Apr	18-Apr	19-Apr	20-Apr	21-Apr	22-Apr	25-Apr	26-Apr	27-Apr	28-Apr	29-Apr
Average Peak Hour Truck TT	MAPE	11.2	9.9	10.1	11	11	11.7	13.8	11.8	13.5	12.5	12.1	11.2	20.3	10.3
	MAE	0.24	0.22	0.22	0.19	0.23	0.23	0.31	0.25	0.26	0.278	0.3	0.29	0.31	0.16
	RMSE	0.19	0.17	0.16	0.17	0.18	0.19	0.25	0.2	0.22	0.21	0.2	0.17	0.21	0.14

The proposed trained hybrid model for one day can estimate peak hour truck travel time for a month with approximately on average MPAE of 11%. This proves the effectiveness of the proposed method while dealing with missing data or data unavailability.

Model Robustness

To show how these two modules help each other in case of anomalies in the data, following experiment is done: The data from 7 AM to 7:15 AM in April 11 is exposed to outliers. That is, a constant value is added to all the data gathered from loop detectors for this period, then two different models are used to estimate truck travel time. First, the hybrid model is used to estimate average peak hour truck travel time. Second, the first module of the hybrid model (estimation module) is solely used to estimate the average peak hour truck travel time.

	MAPE (%)		MAE (Sec.)		RMSE (Sec.)	
Average Peak Hour Truck TT	Hybrid Model	RLS	Hybrid Model	RLS	Hybrid Model	RLS
	10.81	13.49	49.99	63.1	66.03	79.19

The hybrid model is more robust and reliable while dealing with anomalies in databases. It is noteworthy that a perfect database without any outlier is far out of reach, therefore the robustness and reliability of the model is very crucial.

7. Conclusion

- The truck travel time is estimated by fusing indirect travel time from loop detectors and historical truck travel time from NPMRDS.
- The proposed method can be beneficial in cases when truck data is limited or unavailable. Also, since prior knowledge is incorporated to the hybrid model, the model robustness and reliability is improved compared to other estimator filters.
- The model shows approximately 8.7 percent error while estimating truck travel time.
- The robustness of the proposed model is compared with a simple estimation filter. The proposed model showed 3% more accuracy while estimating truck travel time in case there exist anomalies in the database.