Detecting congestion in urban networks based on data fusion

Ayelet Gal-Tzur[1], Yakov Bohadana[2], Yana Barsky[2]

1. Introduction and research goal

Detecting congestion within the urban network is the basis for high quality traffic management. But while researchers and practitioners have been performing this fundamental task for many years, they still regard it as a challenge. Inductive loops, probably the most traditional and widespread information sources for identifying congestion, are considered difficult to install and maintain. The share of modern information sources, especially those based on floating car data, is ever on the rise. However, their reliability for detecting congestion in urban networks is still being examined.

The potential of fusing inputs from several information sources as a means to improve the reliability of congestion detection was identified more than 20 years ago (Sumner, 1991). Ivan and Sethi (1998) analyzed data generated from a simulation of a suburban arterial, and demonstrated that fusing inductive loops and probe vehicle data can improve incident detection. To a great extent, research conducted in recent years has address focused on data fusion for identifying highway traffic states (El Faouzi et al., 2009; Bachmann et al., 2012) and information sources providing similar data, mainly travel times, to address urban networks, (Cheu et al., 2001; Zou et al., 2011; Tettamanti et al. 2014).

Our research aims to determine the potential of fusing several traffic information sources and various data types for prompt and reliable detection of congestion within the urban network.

2. Data collection and preparation

Three types of traffic information sources served as a basis for the research:

- Inductive loops providing volume and occupancy.
- Bluetooth (BT) sensors providing travel times along the link and waiting times near the stop line.
- Travel times obtained from a traffic information provider (TIP) whose data is based on Cellular Floating Car Data derived from mobile phones and location data from GPS equipped devices.

Two types of urban links were investigated:

- A relatively long link (500 m) along Ibn Gabirol arterial in Tel Aviv
- A relatively short link (180 m) which is a minor approach to the Jabotinsky-Ibn Gabirol intersection in Tel Aviv

Over the course of a month, two-minute interval data from each of the information sources was collected. The data was processed in order to identify stable congestion (6 minutes or more). Human assessment of the traffic states for each time interval, conducted by experienced traffic engineers using video imagery, provided the ground truth. The entire data set was divided into two sub-sets, one serving for developing the congestion detection rules and the other for testing their quality.

[1] Ruppin Academic Center, Israel and Transportation Research Institute, Technion, Israel
[2] Transportation Research Institute, Technion, Israel

3. Methodology

Decision trees and logistic regression were used to develop rules for identifying congestion. These methods were applied to the data derived from each information source separately as well as for any combination of the three information sources. Various types of data can serve as input for the detection model, particularly data from previous time intervals which might indicate a congestion build-up process. Hence, an iterative process for finding the best congestion-detection model was conducted for each combination of information sources, as shown in Figure 1.

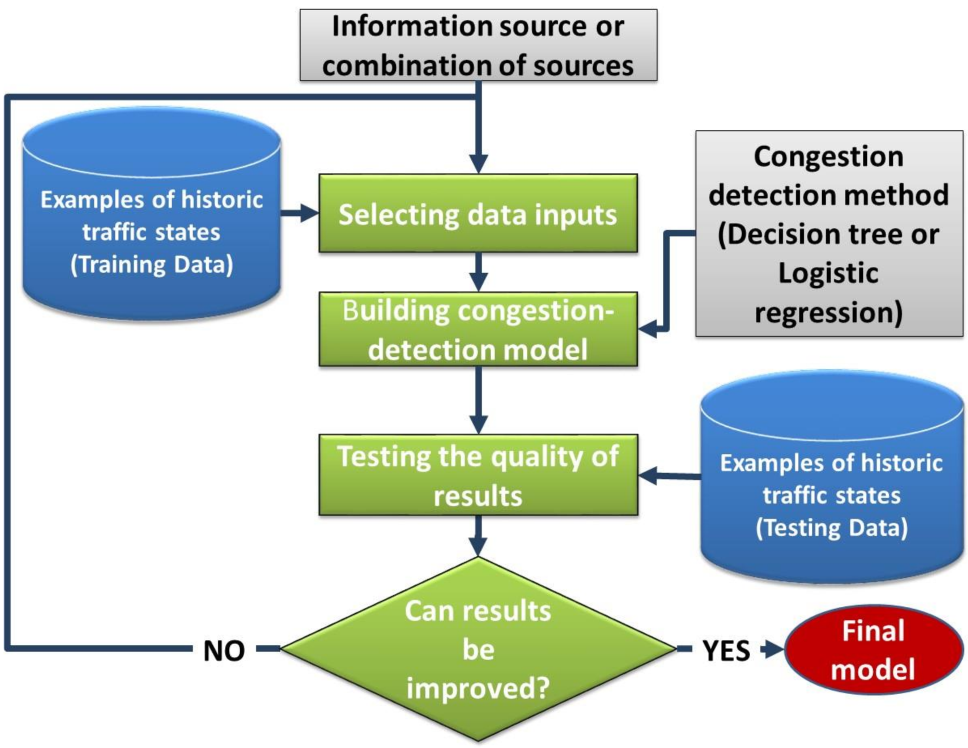

Figure 1- An iterative process for finding the best congestion-detection model

The best detection rules were applied to the testing database. The results were given in a confusion matrix, according to the format given in Table 1, in which the correctly identified Free Flow (FF) and Stable Congestion (SC) states are presented.

		Predicted traffic state	
		Free Flow (FF)	***Stable Congestion (SC)***
Real traffic state	***Free Flow (FF)***	True Negative (TN)	False Positive (FP)
	Stable Congestion (SC)	False Negative (FN)	True Positive (TP)

Table 1 – Confusion matrix derived from applying the detection model to the testing database

Two performance indicators (PIs) were used for evaluating the results of each model, one reflecting the ratio of correctly identified FF conditions (Equation 1) and the other reflecting the ratio of correctly detected SC (Equation 2):

(1) $$\boldsymbol{True\ FF\ Rate\ (TNR) = \frac{TN}{FP+TN}}$$

(2) $$\boldsymbol{True\ SC\ Rate\ (TPR) = \frac{TP}{FN+TP}}$$

Clearly , there is a conflict between the two PIs, as improving one causes deterioration of the other. The right balance between the PIs depends on the priority regime among the different approaches to the intersection.

4. Results and conclusions

Both methods, i.e. decision trees and logistic regression, produced similar results. However, the logistic regression method better enabled balance between the two PIs. Hence, the PI values presented in Tables 2-5 are the results obtained by applying the best logistic regression model to the testing database.

PI	Loop	BT	TIP
TNR	89%	89%	89%
TPR	95%	88%	74%

Table 2 – PI values for the long link along the arterial for each information source separately

PI	Loop + BT	LOOP+ TIP	BT + TIP
TNR	89%	89%	89%
TPR	93%	94%	88%

Table 3 – PI values for the long link along the arterial for fused information sources

PI	Loop	BT	TIP
TNR	91%	92%	93%
TPR	67%	31%	46%

Table 4 – PI values for the minor approach to the intersection for each information source separately

PI	Loop + BT	LOOP+ TIP	BT + TIP	Loop + BT + TIP
TNR	92%	91%	89%	89%
TPR	76%	85%	55%	92%

Table 5 – PI values for the minor approach to the intersection for fused information sources

The above results indicate that induction loops perform best (Table 2 and Table 4). This finding is in line with the findings of previous researches. For the long link along the arterial all three information sources provided relatively good results (Table 2). Given the complex dynamics of traffic flow, it is not surprising that the improvement achieved by fusing several information sources was not substantial (Table 3).

For the short link, i.e. the minor approach to the intersection, the quality of congestion detection by each separate information source was considerably worse compared to that of the long link (Table 4). This phenomenon can be explained by the relatively short distance between the inductive loop and the stop line and the relatively low traffic volumes. However, for this link, the results were significantly improved by fusing several information sources (Table 5).

Acknowledgement

The authors would like to thank the Chief Scientist Unit of the Israel Ministry of Transport for their support in this research

References

Bachmann Chris, Abdulhai Baher, Roorda Matthew and Moshiri Behzad. "Multisensor data integration and fusion in traffic operations and management." Transportation Research Record: Journal of the Transportation Research Board 2308 (2012): 27-36.

Cheu Ruey Long, Der-Horng Lee, and Chi Xie. "An arterial speed estimation model fusing data from stationary and mobile sensors." Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE. IEEE, 2001.

El Faouzi Nour-Eddin, Lawrence Klein, and Olivier De Mouzon. "Improving travel time estimates from inductive loop and toll collection data with Dempster-Shafer data fusion." Transportation Research Record: Journal of the Transportation Research Board 2129 (2009): 73-80.

Ivan John N., and Vaneet Sethi. "Data fusion of fixed detector and probe vehicle data for incident detection." Computer-Aided Civil and Infrastructure Engineering 13.5 (1998): 329-337.

Sumner Roy. "Data fusion in PathFinder and TravTek." Vehicle Navigation and Information Systems Conference, 1991. Vol. 2. IEEE, 1991.

Tettamanti Tamás, MártonTamás Horváth, and István Varga. "Road traffic measurement and related data fusion methodology for traffic estimation." Transport and Telecommunication Journal 15.4 (2014): 269-279.

Zou Jiao, Xiaoxu Bei, and Jian Wu. "Urban Arterial Road Traffic Status Identification Based on Data Fusion." Energy Procedia 11 (2011): 857-862.