ECI 256 Final Project

Clark Fitzgerald

December 2016

1 Abstract

Goal- Find all the active fixed bottlenecks in the 2016 PeMS data. Quantify when and where they exist. Rank them in order of severity. Distinguish temporary from permanent bottlenecks. Compare with highway incident data to determine which are caused by accidents.

To do all this we'll need heuristics and well defined measures for comparison. It would be best to take this from the established literature.

Note- These are initial thoughts before doing the literature review.

One general heuristic for this is to find all places that the velocity v or density ρ is much lower upstream than downstream for at least some period of time T_{min}. Then it's reasonable to say there is a bottleneck between those two mile markers.

We can do this with a two stage approach. First look at every timestamp and check if the average velocity for station i is below some threshold v_i while station $i+1$, the next one downstream on the freeway has traffic moving at velocity greater than v_{i+1}. If this condition persists for longer than t_{min} then call it a bottleneck between station i and $i+1$. This can then be used to check dimensions of the bottleneck in time and space. There we go- that's a great title- Spatial temporal dimensions of traffic bottlenecks.

Choosing an appropriate v_i here may be difficult- so it would be best to use something different for each station.

We may want to forget about detecting bottlenecks if accurate data are too far apart in space or time- ie. 5 miles is probably too far.

2 Review

[Chung et al., 2007] describe traffic characteristics at three fixed bottlenecks.

[Chen et al., 2004] describe systematic identification of bottlenecks from 5 minute loop data. We want to do a similar thing with 30 second data on a larger scale. They use velocity measurements, which are not always reliable / available. So it might be better in our case to use occupancy and flow.

The say: "Five-minute data provide sufficient resolution for this analysis because the traffic features sought are on the order of 30 min or more." So maybe we look for finer features.

[Hall and Agyemang-Duah, 1991] use flow and occupancy since velocity is not available.

> Averages of flow and occupancy across the three lanes were used. The two tended to vary together in the period before congestion aud diverge during the congested period. Determining the exact beginning and end of congestion was, however, difficult from these numbers, so the ratio of occupancy to flow was used. Three values of the ratio were tested for the threshold level: 1.0, 1.1, and 1.2. A ratio of 1.0 gives a longer duration of bottleneck flows, some of which were very low, suggesting that demand was below capacity. A ratio of 1.2 excludes sustained periods (10 min) of high flows (5,800 vehicles/hr or more). A ratio of 1.1 or above persisting for 3 min was selected as the criterion for the identification of the start of a queue.

[Wieczorek et al., 2010] applies this to Oregon data.

[Zhang and Levinson, 2004] show that Queue Discharge Flows QDF's, normal around 2K passenger cars per lane per hour.

Tue Nov 15 08:52:50 PST 2016

A different approach would be to treat the occupancy data as *images*. We can use opencv to do this. For each freeway we can take the following steps:

1. Use R to read the raw file and convert to an image with many missing values. This means expansion to a uniformly spaced grid.
2. Interpolate missing values.
3. Threshold the image, so for a pixel with density $\rho > \rho_0$ we define it as high density, otherwise low density. We'll have to experiment to see what the appropriate value of ρ_0 is but I suspect around 0.3.
4. Once this has been done for every day we can compute an average thresholded image showing the areas of high density. This will show recurring bottlenecks. May need to threshold this also.
5. Compare each day with the average. The difference will show the unusual patterns that occurred on just one day. Might have to do some denoising here- eliminate points that are not part of a cluster.

Once we have the denoised difference we can do the following:

1. (Optional) Detect shapes that are flat on top, since this is the distinguishing feature of a bottleneck.

2. Compute centroid, bounding boxes, and area which will quantify the impact in terms of space and time.

3. Join these features to incident data. This likely will require some text analysis.

Then we can answer questions such as:

1. How many traffic events occur which have no associated incident data? And what sort of events were they probably?

2. What is the impact of an event of type X on a given section of highway? Something along the lines of: When traffic flow is 1500 veh per lane per hour in a two lane freeway a collision involving exactly two vehicles typically creates congestion lasting 10-15 minutes which propagates back 2-3 miles.

3. How can we model the distribution of traffic incidents, ie. Poisson with some parameters.

But how can these results be useful more broadly? More accurate simulations. Input to real time routing. Impacts of planned construction events.

Empirically quantify the impact of various flow levels on congestion. I suspect that for a given type of accident total delay experienced by travelers increases superlinearly as a function of (flow / capacity). Assuming we're looking at uncongested regions. If this holds then it can be used as an argument that we need to try to keep (flow / capacity) under some level. This could mean taking measures such as toll roads or tax credits for employers who find ways to keep employees off roads during peak commute hours. My favorite- an argument to make cars more expensive (like tolls and mileage tax) and use money to improve public transportation.

References

[Chen et al., 2004] Chen, C., Skabardonis, A., and Varaiya, P. (2004). Systematic identification of freeway bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board*, (1867):46–52.

[Chung et al., 2007] Chung, K., Rudjanakanoknad, J., and Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, 41(1):82–95.

[Hall and Agyemang-Duah, 1991] Hall, F. L. and Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation research record*, (1320).

[Wieczorek et al., 2010] Wieczorek, J., Fernández-Moctezuma, R., and Bertini, R. (2010). Techniques for validating an automatic bottleneck detection tool using archived freeway sensor data. *Transportation Research Record: Journal of the Transportation Research Board*, (2160):87–95.

[Zhang and Levinson, 2004] Zhang, L. and Levinson, D. (2004). Some properties of flows at freeway bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board*, (1883):122–131.