Congestion Maps
Boston University
Center for Information and Systems Engineering
Division of Systems Engineering

Project

Congestion Maps is a web-based inference engine for visualizing and analyzing congestion and travel times in traffic networks. Its main objective is to inform urban planners, policy makers, fellow researchers, and the public at large about the increase in traffic conditions (congestion and travel times) in the Eastern Massachusetts network.

These maps are part of the project "A Dynamic Optimization Framework for Connected Automated Vehicles in Urban Environments" funded by the National Science Foundation and carried out at the CODES and NOC Labs and the Center for Information and Systems Engineering at Boston University.

Data description

The dataset contains millions of observed vehicle speeds on more than 11,000 road segments in eastern Massachusetts including collector roadways, arterial roadways, and expressways. The data was gathered by the private firm INRIX and made available to us by the Central Transportation Planning Staff (CTPS) of the Boston Metropolitan Planning Organization (MPO). We have obtained both 2012 (50 GB) and 2015 (130 GB) datasets. This collection of data contains average vehicle speeds on a minute-by-minute basis for the entire year. The data was collected in real-time from mobile phones, connected cars, trucks, and other fleet vehicles equipped with GPS sensors (vehicle-probe technology).

We also use the MassDot 2016 Road Inventory database which contains engineering characteristics (number of lanes) and typical traffic patterns of each road segment. The database covers more than 500,000 road segments in eastern MA.

Processing

We analyze data for the three-month period between April and June, for both 2012 and 2015. We are interested in calculating three key metrics: average speed time, travel time and congestion. All of these metrics are obtained for each road segment and for each time period (later defined as time slots). We begin by defining the relevant notation to explain how each metric was evaluated.

Notation

- We define the following time slots: *AM* (7:00 am - 9:00 am), *MD* (11:00 am - 01:00 pm), *PM* (5:00 pm - 7:00 pm), *NT* (9:00 pm - 11:00 pm), *All* (12:00 am - 11:59 pm) . and let $\mathcal{I}$ denote the index set containing those slots. In particular $i \in \mathcal{I}$ corresponds to one of these slots, e.g., the AM period.

- For each $i \in \mathcal{I}$, we let T_i denote the set of all minutes contained in the time slot. Let $t \in T_i$ correspond to a specific minute in time slot i.

- Let $a \in \mathcal{A}$ denote a road segment.

Data

Our data consists of three main variables:
v_{ait} which is the average speed (miles/hr) on a road segment a in time instance i at time t.
l_a which is the length in miles of road segment a.
m_a which is the capacity (veh/hr) of the road segment a.

Calculations

Step 1: Calculate average speed by instance

$$v_{ai} = \frac{1}{|T_i|} \sum_{t=1}^{T_i} v_{ait}, \qquad \forall a \in \mathcal{A}, \forall i \in \mathcal{I}. \tag{1}$$

where T_i is the set of indices pointing to each speed observation of road segment a and time period i, and where $|T_i|$ denotes the cardinality of the set T_i.

Step 2: Calculate travel time

$$\tau_{ai} = \frac{v_{ai}}{l_a}, \qquad \forall a \in \mathcal{A}, \forall i \in \mathcal{I}. \tag{2}$$

where τ_{ai} is the travel time or cost of utilizing road segment a.

Step 3: Calculate density

Using the Greenshield model [2], it follows that density is a linear function of speed. Then:

$$d_{ait} = d^{max}(1 - \frac{v_{ait} - v_a^{min}}{v_a^{max} - v_a^{min}}). \tag{3}$$

Let v_a^{min} denote the minimum speed at segment a, which we set to 1 mph for all a. We also let v_a^{max} be the corresponding free flow speed estimated as the 85% percentile of the road segment's speeds. In order to find this percentile we build a histogram by randomly sampling more than 5,000 observations for each road segment. Finally d^{max} is the *jam density* value which is estimated as twice the value of the road segment's capacity

Step 4: Calculate congestion

$$C_{ai} = \frac{1}{|T_i|} \sum_{t=1}^{T_i} \mathbb{1}\{d_{ait} \geq m_{ai}\}, \qquad \forall a \in \mathcal{A}, \forall i \in \mathcal{I}, \tag{4}$$

where $\mathbb{1}$ is the *indicator function*; equal to one when the condition is satisfied and zero otherwise.

These expression is estimating the fraction of time in a specific time period (e.g. AM) when the road segment a is above its capacity. This corresponds to the common notion of "traffic jam".

Step 5: Group data by Zip Code

To group the statistics by Zip Code, we calculate a weighted-sum for each statistic.

Avg. Speed

$$v_Z = \frac{\sum\limits_{a \in Z} l_a v_{ai}}{\sum\limits_{a \in Z} l_a}, \qquad \forall i \in \mathcal{I}. \tag{5}$$

Where Z is the set of all arcs in zip code Z.

Zip Congestion

$$C_Z = \frac{\sum\limits_{a \in Z} l_a C_{ai}}{\sum\limits_{a \in Z} l_a}, \qquad \forall i \in \mathcal{I}. \tag{6}$$

References

[1] Zhang, J., Pourazarm, S., Cassandras, C. G., Paschalidis, I. C. (2018) *The Price of Anarchy in Transportation Networks: Data-Driven Evaluation and Reduction Strategies.* Proceedings of the IEEE, 106(4). https://doi.org/10.1109/JPROC.2018.2790405

[2] Adolf D. May (1990) *Fundamentals of Traffic Flow.* Prentice - Hall, Inc. Englewood Cliff New Jersey 07632, second edition