Quantifying City-Scale Resilience to Extreme Congestion Events

Dan Work
Civil and Environmental Engineering and Coordinated Science Laboratory
University of Illinois at Urbana–Champaign

Joint work with Brian Donovan (UIUC) & Dr. Jong Lee (NCSA)

[Research Sponsors: NSF, NCSA, FHWA RSI UTC, IDOT]
[Open source code and data available at GitHub and publish.illinois.edu/dbwork]

Goal: Empirically quantify resilience to extreme congestion events

- Extreme congestion events characterized by:
  - Changes in travel demand
  - Loss of cyber components
  - Damage to physical components

- Assess transportation network performance quantitatively
  - Recovery time
  - Deviation deviations from typical conditions

Go to “war” with the sensors you have, not the ones you want – D. Estrin

- 697,000,000 taxi trips (2010-2013)
- FOIL data: New York City Taxi and Limo Commission
- Data structure:
  - Only two GPS points
    - Start of trip
    - End of trip
  - Meter distance and travel time
  - Driver ID and Car ID
  - Fare data
- Download: publish.illinois.edu/dbwork/open-data

Filtering the data

(7.5% data errors, 11% uninformative)
High resolution traffic estimation with coarse GPS data

- Compute hourly traffic paces via iterative link travel time estimator
  - **Routing**: Given link speeds, route all taxis and compute predicted travel times
  - **Link travel time correction**: Compare predicted travel times with true travel times, then adjust the link travel times to minimize the travel time prediction error

Scaling to multi-year datasets

- NYC road graph
  - 24,000 nodes
  - 63,000 links

- Typical hour: 20,000 trips

- Computing traffic estimates
  - Bottleneck: routing step
  - Accuracy: ~2 min error per trip
  - Saved estimates: publish.illinois.edu/dbwork/open-data

Storing traffic data in a vector

- Build a column vector of traffic paces
- Each element corresponds to the average pace in a given hour during the four year period

\[
M_1 = \begin{bmatrix}
  M_{11} \\
  M_{21} \\
  \vdots \\
  M_{m1}
\end{bmatrix}
\]

- Pace on link 1 (Wed noon-1pm, 1/6/2010)
- Pace on link 2 (Wed noon-1pm, 1/6/2010)
- \( \vdots \)
- Pace on link \( m \) (Wed noon-1pm, 1/6/2010)

[Donovan, Lee, Work, IEEE T-ITS 2016 (in review)]
Building a traffic data matrix

- \( n = 52 \) (weeks/year) \( \times 4 \) (years) copies of traffic on a Wednesday from noon-1pm over a four year period

\[
M = \begin{bmatrix}
M_{11} & M_{12} & \cdots & M_{1n} \\
M_{21} & M_{22} & \cdots & M_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
M_{m1} & M_{m2} & \cdots & M_{mn}
\end{bmatrix}
\]

Facts about traffic data matrices

- If traffic is perfectly periodic and repeatable:
  - all columns of \( M \) are identical
  - \( M \) is a rank 1 matrix

- Real traffic matrices
  - Have missing data
  - Have estimation errors
  - Have some columns that are far from typical (e.g., 12/25/2013 is a federal holiday)
Detecting outliers via Robust PCA

• **Problem:** Given Data Matrix M, find a decomposition composed of:
  - \( M \): original data matrix
  - \( L \): low rank matrix
  - \( C \): column sparse matrix indicating outliers

• **Solved via convex optimization:**
  
  \[
  \begin{align*}
  \text{Minimize:} & \quad \|L\|_* + \lambda \|C\|_{1,2} \\
  \text{Subject to:} & \quad \|M - (L + C)\|_F \leq \varepsilon
  \end{align*}
  \]

• **Key outcome:** Any nonzero column in \( C \) is an outlier

• Extreme points in the low dimensional subspace are also identified as outliers via Mahalanobis distance.

<table>
<thead>
<tr>
<th>Event</th>
<th>Start Time</th>
<th>Duration (Hours)</th>
<th>Max Pace (Min/Mi)</th>
<th>Min Pace Dev (Min / Mi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hurricane Sandy</td>
<td>2012-10-28 21:00</td>
<td>134</td>
<td>2.26</td>
<td>-1.54</td>
</tr>
<tr>
<td>Snowpocalypse</td>
<td>2010-12-26 13:00</td>
<td>109</td>
<td>4.24</td>
<td>0.33</td>
</tr>
<tr>
<td>Blizzard</td>
<td>2011-01-31 10:00</td>
<td>47</td>
<td>2.01</td>
<td>0.34</td>
</tr>
<tr>
<td>Hurricane Irene</td>
<td>2011-08-27 13:00</td>
<td>43</td>
<td>0.65</td>
<td>-1.65</td>
</tr>
<tr>
<td>Blizzard</td>
<td>2010-02-10 06:00</td>
<td>32</td>
<td>0.65</td>
<td>-1.03</td>
</tr>
<tr>
<td>Blizzard</td>
<td>2013-02-08 06:00</td>
<td>27</td>
<td>1.53</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

Summary and future perspectives

- Quantifying extreme events
  - Hurricane Sandy: ~5 days of recovery time
  - Worst congestion after the hurricane hit.
  - Re-entry process is significantly slower than evacuation.

- Long term goal: Improve transportation policy
Quantifying City-Scale Resilience to Extreme Congestion Events

Dan Work
Civil and Environmental Engineering and Coordinated Science Laboratory
University of Illinois at Urbana–Champaign

Joint work with Brian Donovan (UIUC) & Dr. Jong Lee (NCSA)

[Research Sponsors: NSF, NCSA, FHWA RSI UTC, IDOT]
[Open source code and data available at GitHub and publish.illinois.edu/dbwork]