Computational Transportation Science: Challenges and Opportunities in Traffic Modeling and Simulation

Cheng Liu, Ph.D
Wei Lu, Ph.D

Geographic Information Science and Technology Group
Computational Science and Engineering Division
Oak Ridge National Laboratory

September 11, 2014
Outline

- Introduction
- Traffic Modeling
- Traffic Simulation (TUMS)
  - Data Processing
  - Transportation Visualization
  - Live Demos
- Conclusions
Computational Transportation Science

- Computational Transportation Science (CTS) is an emerging discipline that combines computer science, data science and engineering with the modeling, planning, and economic aspects of transportation.

- The discipline studies how to improve the safety, mobility, and sustainability of the transport system by taking advantage of information technologies and ubiquitous computing.
Evacuate everyone south of that line.

*The Day after Tomorrow (2004)*
Travel demand modeling

- Four-step Trip modeling
- Tour-based modeling
- Activity-based modeling
- Agent-based modeling

- High resolution demographic data can increase the model accuracy.
Synthetic Population

- An Extraordinary Synthetic Map of Every Household (112,596,600) in America by age, size, income, and race.

- **Data Source:**
  - American Community Survey (ACS)
    - Summary tables by block group
  - Public Use Microdata Sample (PUMA)
    - 5% sample with detailed household information
  - LandScan Dataset
    - 30m by 30m cell population distribution

ACS: Block Group (Education vs Income)

<table>
<thead>
<tr>
<th></th>
<th>&lt;20K</th>
<th>20K – 60K</th>
<th>60K – 100K</th>
<th>&gt;100K</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>40</td>
</tr>
<tr>
<td>College</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>30</td>
</tr>
<tr>
<td>Master</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>20</td>
</tr>
<tr>
<td>Ph.D</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>??</td>
<td>10</td>
</tr>
<tr>
<td>Summary</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

http://www.citylab.com/housing/2013/10/extraordinary-synthetic-map-every-household-america/7375/
## Iterative Proportional Fitting (IPF)

### Geographic Information Science and Technology

---

### PUMA: Education vs Income

<table>
<thead>
<tr>
<th></th>
<th>&lt; 20K</th>
<th>20K – 60K</th>
<th>60K – 100K</th>
<th>&gt; 100K</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>200</td>
<td>300</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>College</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>Master</td>
<td>10</td>
<td>20</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Ph.D</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

---

### Block Group #1: Education vs Income

<table>
<thead>
<tr>
<th></th>
<th>&lt; 20K</th>
<th>20K – 60K</th>
<th>60K – 100K</th>
<th>&gt; 100K</th>
<th>Summary</th>
<th>adjust</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>200</td>
<td>300</td>
<td>100</td>
<td>10</td>
<td>40</td>
<td>40/700</td>
</tr>
<tr>
<td>College</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>80</td>
<td>30</td>
<td>30/380</td>
</tr>
<tr>
<td>Master</td>
<td>10</td>
<td>20</td>
<td>70</td>
<td>50</td>
<td>20</td>
<td>20/150</td>
</tr>
<tr>
<td>Ph.D</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>10</td>
<td>10/65</td>
</tr>
<tr>
<td>Summary</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

---

Managed by UT-Battelle for the Department of Energy
Maximum Entropy

\[
\max \sum_{i,j} w_{ij} \log(w_{ij}) \\
\text{subject to:} \\
\sum_j w_{ij} = O_i \\
\sum_i w_{ij} = D_j
\]

The copula is defined as the joint cumulative distribution function

\[ C(u_1, u_2, ..., u_d) = P[U_1 \leq u_1, U_2 \leq u_2, ..., U_d \leq u_d] \]

Attributes: Household income, size, workers, vehicles, education and travel time

Examples of how to use these data:


Challenges

- **High-Resolution O-D Matrix Assignment:**

  - Given the residential/work areas, block group level travel flow, household attributes (number of people, workers, income, and travel time), and the LandScan population, can we generate the OD matrix for each household?

  - For example, in one LandScan cell, there are 10 persons, 3 family, 4 workers. Each worker in this cell will be assigned to a different cell based on travel time.
Challenges

Deadlock vs Gridlock in microscopic simulation

Prepared by: Dr. Jan-Mou Li
Oak Ridge National Laboratory
Email: lii3@ornl.gov
Challenges - Possible solutions

• **Prevention**
  – Design a tree structure network

• **Avoidance**
  – Steer around deadlock with smart scheduling: Staging

• **Detection & Recovery**
  – Check periodically, such as “Holding” problem
  – Kill a deadlocked process, Alien Adopted My Car (AAMC) in TRANSIMS

• **Do nothing**
  – If deadlock is rare, does it worth the overhead?
  – Manual intervention (kill processes, reboot)
Traffic Simulation (Microscopic)

Washington, DC (2011)

Hurricane Rita (2005)
**TUMS: Toolbox for Urban Mobility Simulation**

**Motivation:**
- A system that can be used by urban planners, emergency managers, transportation professionals to estimate urban mobility performance and decide efficient strategies for special events.

**Goals:**
- Opens Source
  - Both Data and model
  - OpenstreetMap, TANSIMS
- Platform independent
  - Window/Linux/MacOS

**TUMS (Toolset for Urban Mobility Simulations)**
- TUMS integrates *open data, mobility models, open-source programs* to generate different urban mobility applications.

---

Input Data:

- LandScan™ and OpenStreetMap
  - 30m * 30m for US & 1km * 1km for global
  - Daytime & Nighttime
  - High resolution road networks

- 3634 links, 2608 nodes, 7718 activity locations, 1062 stop/yield signs, 261 traffic signals, and 21 shelters.


Data Processing Example: Alexandria, VA
Spatial Resolution Difference

Traffic Analysis Zones (62) vs. LandScan USA Population Cells (5657)
Daytime and Nighttime

Temporal Resolution Difference

LPC Daytime (5657)  
LPC Nighttime (4522)
Left-Hand Traffic

163 countries/areas on the right
79 countries/areas on the left

How to use right-hand models to simulate left-hand traffic?

- Data in WGS84 projection (Lon/Lat).
- Simulation uses UTM projection (Cartesian coordinates with x, y)

Mirror technique:
- Flip x coordinates;
- Run Simulation;
- Flip x coordinates back.
- Switch left/right for traffic analysis

\[
\begin{align*}
    x_{\text{min}} &= x_0 - \frac{\text{radius}}{2} = 5,000,000 - \frac{6,378.137}{2} = 1,810,931.5 \\
    x_{\text{max}} &= x_0 + \frac{\text{radius}}{2} = 5,000,000 + \frac{6,378.137}{2} = 8,189,068.5 \\
    x_{\text{left}} &= x_{\text{min}} + x_{\text{max}} - x_{\text{right}} = 10,000,000 - x_{\text{right}}
\end{align*}
\]
Challenges

- **OSM data quality**
  - Geocoding
  - Topology
  - Data competence
  - others

- **High-resolution data for transportation**
  - LandScan vs. TAZ
  - OSM vs. Highway network
Results for TAZ vs. LPC

TAZ underestimated the evacuation clearance time

Alexandria city road network

<table>
<thead>
<tr>
<th>Road network type</th>
<th>Major</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>1061</td>
<td>2608</td>
</tr>
<tr>
<td>Number of links</td>
<td>1278</td>
<td>3634</td>
</tr>
<tr>
<td>Number of activity locations</td>
<td>1644</td>
<td>7718</td>
</tr>
<tr>
<td>Number of un-signalized nodes</td>
<td>110</td>
<td>1062</td>
</tr>
<tr>
<td>Number of signalized nodes</td>
<td>182</td>
<td>261</td>
</tr>
</tbody>
</table>
HWB outperformed the other two in evacuation performance. The higher data resolution, the more accurate results.
• Traffic Simulation Models (Evacuation Scenarios):
  ➢ Travel Demand modeling
    ➢ # of evacuation people & departure time
    ➢ Weibull distribution
  ➢ Trip Distribution Modeling
    ➢ Origin-Destination Matrix
    ➢ Super Node Trip Distribution Algorithm
  ➢ Traffic Assignment Modeling
    ➢ Activity-based traffic assignment with more accurate representation of vehicles accessing problems

• Global applications
  ➢ The same simulation model for any locations.
Super Node Trip Distribution Algorithm

- **Input:** Evacuation Area A; Road Network G(V, E); Destinations D
- **Output:** O-D matrix
- **Goal:** Min \{Travel Cost_i, i is the i^{th} LPC\}
- Minimize computation time


Managed by UT-Battelle for the Department of Energy
High Resolution Visualization
Link-based visualization to assess transportation network performance
Microscopic Visualization (Demo)

Vehicle-based visualization to identify choke points, traffic controls, etc.
Challenges

- High Performance Computing for Traffic Simulation
  - Current running time can be 16 hours.
  - Message Passing Interface (MPI) based Parallel computing
  - Multiple-threads based computing
  - Hybrid (MPI + OpenMP) computing

Challenges

- **Microscopic Transportation Visualization**
  - Big Data storage
  - Network Communication
  - Interactive Operation

• **Destination Compliance Behavior during Evacuation**

• Studies on evacuation modeling with traveler information and compliance behavior through macroscopic simulations reveals that traveler compliance behavior affect evacuation efficiency.

• What is the impact of high resolution zone structure on travelers’ compliance behavior (destination choices) in evacuation?

Case study

One-to-one destination assignment

Multiple destination assignment

\[ k^{th} \text{ shortest destination, } D\{k, k + 1\} \geq 1 \text{ mile} \]

\[ k = 1, 2, 3 \]
# Evacuation Time

<table>
<thead>
<tr>
<th>Population Resolutions</th>
<th>Compliance Level</th>
<th>Percent Evacuated</th>
<th>Evacuation Times in Minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>25%</td>
<td>50%</td>
</tr>
<tr>
<td>TAZ</td>
<td>100%</td>
<td>114</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>70% (2)</td>
<td>112</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>50% (3)</td>
<td>114</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>30% (4)</td>
<td>119</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>0% (5)</td>
<td>125</td>
<td>204</td>
</tr>
<tr>
<td>LPC</td>
<td>100% (6)</td>
<td>120</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>70% (7)</td>
<td>120</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>50% (8)</td>
<td>119</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>30% (9)</td>
<td>121</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>0% (10)</td>
<td>120</td>
<td>223</td>
</tr>
</tbody>
</table>

(1) Only 99.26% of the trips arrived at destinations in 10 hours. 99% is used instead of 100%.
(2) Only 98.85% of the trips arrived at destinations in 10 hours.
(3) Only 96.21% of the trips arrived at destinations in 10 hours.
(4) Only 92.48% of the trips arrived at destinations in 10 hours.
(5) Only 86.50% of the trips arrived at destinations in 10 hours.
(6) Only 89.09% of the trips arrived at destinations in 10 hours.
(7) Only 87.94% of the trips arrived at destinations in 10 hours.
(8) Only 88.07% of the trips arrived at destinations in 10 hours.
(9) Only 86.23% of the trips arrived at destinations in 10 hours.
(10) Only 85.87% of the trips arrived at destinations in 10 hours.
High resolution data is not significantly sensitive to travelers’ compliance behavior.
More challenges in CTS

- Multimodal Simulation (Transits / Pedestrian)
  - Simulate dense urban areas
- Social Media data for travel demand modeling
  - Generate trip chain based O-D matrix
- Connected Vehicles Simulation
  - Impacts on transportation energy system

Conclusions

- CTS integrates the core concepts in each domain, but also introduces many challenges.

- Two workshops
  - ACM SIGSPATIAL International Workshop on Computation Transportation Science
  - DAGSTUHL CTS Seminar

- Six Institutes
  - University of Illinois at Chicago
  - Argonne National Laboratory
  - Aalborg University
  - National University of Singapore
  - Technische University Braunschweig
  - The University of Tennessee

- Not just another name of ITS!
Contacts & Questions

- Dr. Cheng Liu
  - liuc@ornl.gov

- Dr. Wei Lu
  - luw4@ornl.gov