Integrating State DOT traffic volume data with private vendor archived travel time data:
Moving toward a more precise approach
Regional Transportation System

Road Network:
• 147 million vehicle miles each day
• 26,000 miles of roads, including 177 miles of toll roads
• 4,800 bridges

Transit Network:
• 732,000 trips daily
• 13% of commuters ride transit
• 250 bus routes
• Commuter rail: 390 miles of track, 150 stations
Overview

- Background
- Objective
- Common conflation methods
- Issues with the conflation methods
- Suggested approach
- Conclusions
- Issues/Next steps
Investigating roadway performance using attributes from data sources with different segmentation, georeferencing and spatial accuracy

New Jersey Pilot Study: Testing Potential MAP-21 System Performance Measures
- Annual Hours of Delay (Vehicles and Persons)
- Reliability Index
Two major travel corridors of NJ

- I-78 | Interstate Highway
- NJ 18 | Arterial
Background: Data

Travel Time (Speed) Data: Hourly Average

Hourly Vehicle Volume
Objective

- Difference in Feature Representation
- Positional Inaccuracy
- Segmented Differently
- Different Geometry
Common conflation methods

- Spatial Approaches for Conflating GIS Roadway Datasets
  - Li and Liu, 2014

- Improving Mobility Information with Better Data and Estimation Procedures
  - TTI, 2011

Method 1: Linear Referencing

1. Generate route layer from layer 1
2. Locate features along routes for layer 2
3. Create linking table
4. Join layers through linking table

Method 2: Spatial Join

1. Extract start and end points from layer 1
2. Consolidate adjacent points
3. Join layers using ‘Spatial Join’
4. Split layer 2 on the basis of points
5. Generate buffer around layer 2
Common conflation methods

- Conflation Methodologies to Incorporate Consumer Travel Data into State HPMS Datasets
  - Green et.al, 2013

- Efficient Integration of Road Maps
  - Safra et.al, 2006

Method 3: Midpoint Snapping

1. Reducing existing network
2. Midpoint conversion of layer 1
3. Snapping of layer 1 to layer 2

Method 4: Only Algorithms

```
AproxMatching(M1, M2)
Input: Two road maps M1 and M2
Output: A matching μ of the polylines in M1 and the polylines in M2, with respect to a condition χ on nodes and a semantics □ ∈ {AND, OR}

1: for i = 1, 2 do
2:   let Si be all the pairs (n, t), such that n is an endpoint of t in Mi
3:   sort Si according to the coordinates of the nodes
4: for each node n in Si do
5:   let degree(n) be the number of lines for which n is an endpoint
6:   let Ni be the set of nodes in Si that satisfy χ
7: μn ← Ø
8: for each pair of nodes n1 ∈ Ni, n2 ∈ N2 do
9:   if (n1 is the nearest neighbor of n2 in N1) □ (n2 is the nearest neighbor of n1 in N2) then
10:      add (n1, n2) to μn
11:      add (n2, n1) to μn
12:      if n1 is not empty do
13:         pop an element (n, t) from I
14:         let j be the index opposite to i, i.e., if i is 1 then j is 2 and if i is 2 then j is 1
15:         for each (n, t) ∈ Si and (n', t') ∈ Sj s.t. m(n', t') then
16:            add (n', t') to μj
17:            if (n'.t') is not in V then
18:               add (n'.t', j) to f
19: for each (n', t') ∈ Si and (n', t') ∈ Sj s.t. m(n', t') then
20:            add (n', t') to μj
21:            if (n'.t') is not in V then
22:               add (n'.t', j) to f
23:               add (n, t, i) to V
24: return μ
```

Match-Lines(S1, S2, μn)
Input: Two sets S1 and S2 of pairs of a polyline and an endpoint, a matching μn of nodes
Output: A matching μ of the polylines in S1 and the polylines in S2
Issues with the conflation methods

- Positional inaccuracy led to the problem of finding the correct search radius/search tolerance in method \(\text{method 1}\) & method \(\text{method 2}\)
Issues with the conflations methods

- Imprecise start and end mile posts to segments of layer 2 in method ❷ & method ❸
Method 2: Spatial Join

- Extract start and end points from layer 1
- Consolidate adjacent points
- Split layer 2 on the basis of points
- Generate buffer around layer 2
- Join layers using ‘Spatial Join’

- **Erroneous re-segmentation**
  - Low search radius will split layer 2 on the basis of only some points generated from layer 1
  - High search radius will split layer 2 on the basis of non-corresponding points generated from layer 1
Method 2: Spatial Join

Extract start and end points from layer 1 → Consolidate adjacent points → Split layer 2 on the basis of points → Generate buffer around layer 2 → Join layers using ‘Spatial Join’

- Difficult to assess the correct number of split segments
- Manually validate splitting

I-78 West fragmented into 75 segments
Erroneous fragmentation led to 134 segments
Correct fragmentation led to 159 segments
Buffering and spatial join is an iterative process as identified by the literature.
Issues with the conflation methods

Method ❶: Only Algorithms
No tools to conduct the process efficiently.

Method ❷: Midpoint Snapping

- Reducing Existing Network
- Midpoint Conversion of layer 1
- Snapping of layer 1 to layer 2

As mid point of a line is only considered, parts of layer 1 are not joined to the corresponding parts of layer 2.
Conflation method steps

Method 1
- A. Create route event layer
- B. Locate features along routes
- C. Create linking table
- D. Join attributes on basis of linking table
- E. Extract start and end points | Create point layer
- F. Split segments on basis of corresponding layer
- G. Create buffer
- H. Spatial Join
- I. Conversion to midpoints
- J. Snapping midpoints to line segments

Method 2

Method 3

Method 4

K. Algorithms
• Use a non-segmented layer 3 for conflation
• Layer 3 should have characteristics in common with layers 1 and 2
Select, re-arrange & re-use the steps

E. Extract start and end points | Create point features for layer 1 (TMC layer)

Create point features from start and end vertices of line feature

Create point features from start and end latitude and longitude

<table>
<thead>
<tr>
<th>tmc</th>
<th>road</th>
<th>direction</th>
<th>intersection</th>
<th>state</th>
<th>county</th>
<th>zip</th>
<th>start_latitude</th>
<th>start_longitude</th>
<th>end_latitude</th>
<th>end_longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>103+05931</td>
<td>Bordentov</td>
<td>NORTHBOUNDE</td>
<td>Turnpike Access</td>
<td>NJ</td>
<td>BURLINGT</td>
<td>8016</td>
<td>40.088726</td>
<td>-74.824901</td>
<td>40.103143</td>
<td>-74.79118</td>
</tr>
<tr>
<td>103P05931</td>
<td>Bordentov</td>
<td>NORTHBOUNDE</td>
<td>Turnpike Access</td>
<td>NJ</td>
<td>BURLINGT</td>
<td>8016</td>
<td>40.103143</td>
<td>-74.79118</td>
<td>40.104432</td>
<td>-74.788222</td>
</tr>
<tr>
<td>103+05931</td>
<td>Bordentov</td>
<td>NORTHBOUNDE</td>
<td>Hornberger Ave</td>
<td>NJ</td>
<td>BURLINGT</td>
<td>8505</td>
<td>40.104432</td>
<td>-74.788222</td>
<td>40.112373</td>
<td>-74.769785</td>
</tr>
<tr>
<td>103-05935</td>
<td>Bordentov</td>
<td>SOUTHBOUND</td>
<td>Hornberger Ave</td>
<td>NJ</td>
<td>BURLINGT</td>
<td>8505</td>
<td>40.126438</td>
<td>-74.738207</td>
<td>40.112474</td>
<td>-74.76986</td>
</tr>
<tr>
<td>103-05938</td>
<td>Bordentov</td>
<td>SOUTHBOUND</td>
<td>Turnpike Access</td>
<td>NJ</td>
<td>BURLINGT</td>
<td>8016</td>
<td>40.112474</td>
<td>-74.76986</td>
<td>40.10379</td>
<td>-74.789978</td>
</tr>
</tbody>
</table>
J. Snap point features of layer 1 (TMC layer) to line features of layer 3 (SLD layer)

Warning! Find the correct search radius for snapping
F. Split segments of layer 3 (SLD layer) on basis of point features of layer 1 (TMC layer)
H. Spatially join layer 3 (SLD layer) and layer 1 (TMC layer) to transfer all the attributes to one layer – layer 4.
B. Locate features along routes | Provide correct SRIs and MPs to features of layer 4

<table>
<thead>
<tr>
<th>Shape</th>
<th>SRI</th>
<th>MP_START</th>
<th>MP_END</th>
<th>DIRECTION</th>
<th>SLD_NAME</th>
<th>RTE_LABEL</th>
<th>SRI_OLD</th>
<th>rd_TMC9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120+05903</td>
</tr>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120P05903</td>
</tr>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120P05903</td>
</tr>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120P05903</td>
</tr>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120P05903</td>
</tr>
<tr>
<td>Polyline M</td>
<td>00000078__</td>
<td>0</td>
<td>67.830002</td>
<td>West to East</td>
<td>I-78</td>
<td>78</td>
<td>00000078</td>
<td>120P05903</td>
</tr>
</tbody>
</table>

Recalculated start and end mileposts

SRI
TMC Codes
Precisely apportioning volume data - through SQL queries

Relationship between layer 4 (layer having attributes of both TMC & SLD layer) and layer 2 (CMS layer)

- Extension

- Containment

- Partial Overlap
- Extension

- Containment

- Partial Overlap
Conclusions

• Use of non-segmented layer 3 reduces the difficulty of re-segmenting already segmented line features

• ‘Snapping’ at the beginning reduces the difficulty of finding the ‘search radius/search tolerance’ multiple times in the process
Conclusions

- Calculating precise mileposts for all segments facilitates the precise apportioning of the volume data through SQL queries.

- The process precisely conflates two datasets with different positional accuracy, feature representation, feature segmentation and geometric representation, in terms of both spatial and non-spatial attributes.
Issues / Next steps

- Automate the process
- Conflate the entire road network at the same time including intersections

NJDOT/NJIT effort
Thank You

Sutapa Bhattacharjee
sbhattacharjee@njtpa.org

Defining the Vision. Shaping the Future.

NJTPA
NORTH JERSEY TRANSPORTATION PLANNING AUTHORITY
One Newark Center, 17th Floor
Newark, NJ 07102
(973) 639-8400
www.njtpa.org