Developing a National Highway Network from ARNOLD and HPMS

Gary Baker and Peter Wilke
April 24, 2019
Outline

• Background and Motivation
• Building the Network
• Results and Testing
Background and Motivation

- The U.S. DOT has maintained and distributed the National Highway Planning Network (NHPN) since the early 1990s.
- Over the years additional networks have been created and maintained for different purposes, most notably, the Freight Analysis Framework, the National Highway System.
- And then there’s ARNOLD.
- Most agree that the USDOT needs a single national highway network.
The U.S. DOT needs a single national road network that:

- Can be used for national and regional mapping, modelling, and analysis
- Meets the needs of a wide variety of users
- Is used by all U.S DOT modal administrations (e.g. FHWA, FRA, FTA)
- Is consistently updated
- Can be freely distributed
- Can be better linked with alternative data sources (e.g. Waze, NPMRDS)
- Can realize the synergies and momentum from everyone working with and improving upon a single network!
Background and Motivation

Why use ARNOLD and HPMS?

• Provided directly by the states
• Updated annually
• Has key attributes needed for modeling
  • AADT, functional class, number of lanes
• Geometry-rich
• Relatively uniform format
• No license restrictions
Outline

• Background and Motivation
• Building the Network
• Results and Testing
Building the Network: Very High-level  

1. Loading HPMS records and shapefiles  
2. Removing routes with only functional class 7 or null records  
3. Fixing minor spatial gaps  
4. Splitting at intersections and operational changes  
5. Assigning mileposts to features using M values  
6. Pushing down attributes from LRS to road segments  
7. Subsetting to functional class 4 and facility types 1 and 2  
8. Appending manual features  
9. Clipping at state borders  
10. Generating automated dangle connectors  
11. Generating state border connections  
12. Merging all states and connectors
Challenges

- Detailed source datasets need to be dropped as first step to manage network size
- Some states need to be broken into two parts for initial processing steps (CA, TX)
- 49 different sources, each with numerous subtleties
- Missing geometries
- Geometry disconnects
- Coincident geometries
- Potentially incorrect (or realigned) geometry
- Unconnected features
- Missing attributes (e.g. null functional class)
- Multiple HPMS events for segment
- Averaging AADT values

- Collapsing consecutive identical HPMS records on single segment.
- Breaks caused by functional class subset
- Inconsistent M-values (e.g. out of order, 0)
- Esri feature to line bug which overwriting coincident M-values
- Unsplitting lines
- Splitting lines at intersections and operational changes
- Dual carriageway inconsistencies
- Grade separation
- Complex interchanges
- State borders
Geometry disconnects in resulting network can be due to a number of different reasons.
Geometry Updates: Disconnects

Thousands of geometry disconnects exist nationally.

Example of an extremely small disconnect in the network.

A 0.6 meters (2 foot) disconnected road segment.
Geometry Updates: Manually Appending Features

Feature in red was missing from input geometry received from the state.

A new feature is manually appended to improve flowability.
Geometry Updates: Dangles

- Dangles are dead-end nodes within a network.
- Legitimate dangles
  - International borders
  - Actual dead-end roadway
  - A dangle may exist where a functional class 1-4 roadway becomes a functional class 5+ roadway
- Other dangles can highlight gaps in the network.
- Identifying and connecting dangles is a major step in enhancing the connectivity of the network.
Geometry Updates: Dangle to Dangle Connectors

• This category of connectivity gaps attempts to connect ‘dangle pairs’—dangles which are within the search tolerance of exactly one other dangle.

• Code to automatically create a connector between them.
Geometry Updates: Dangle to Road Connectors

- Connects any remaining dangles to the nearest non-neighboring route within the search tolerance.

- Various rules have been developed in an effort to avoid false connectors.
Attribute Assignment: Multiple HPMS Events

- Features may have multiple HPMS records—how best to assign attributes?
- Initially took the highest value seen on segment.
- Then switched to all attributes from longest record (indicated with arrow for the feature to the left).
- Now splitting features between intersections where non-AADT attributes change between adjacent records, if no change then a weighted average is used for AADT
For some states multiple HPMS records exist for a single feature because of very fine level of change in AADT as shown at left. The feature would be assigned a weighted average value.

Figure below shows AADT difference by mileage for functional class 4 road segments with multiple HPMS events.
Attribute Assignment: M-value Challenges

Esri Feature to Line bug

- Coincident features with different route IDs and M-values of 27.806 and 144.509 at same coordinate
- After running Esri Feature to Line tool both have M-value of 144.509 - causes incorrect milepost assignment
- Esri bug documented

Non-consecutive M-values on same route

- Multipart feature has non-consecutive M-values of 79.428 and 201.382 at same coordinate
- One feature should have an end milepost of 79.428 and the next a start of 201.382
- Instead, both are assigned 79.428 - causes incorrect milepost assignment
Multipart route ID with inaccurate M-values that cause incorrect mileposts and HPMS attributes to be assigned. HPMS events for route shown below.
Attribute Assignment: Inconsistent Func. Class

Variation is seen in functional class both within single routes (shown left) as well as across state borders (right).
Latest prototype included only facility type 1 and 2 features. Removing facility types 4 (ramps) and 6 (non-inventory directional) reduced the network size without sacrificing flowability. Shown on right is an area south east of Worcester, MA.
Attribute Subsets: Functional Class

Investigating potential inclusion of functional class 5 is important to model one level of detail past intended use. Functional class 5 provides much greater coverage in rural areas.
Other Geometry Considerations: Grade Separation

Example of unrealistic split of non-at-grade geometries, not currently being handled.

Google Maps showing north-south road is an underpass beneath I-90.
Route in red crosses black state border multiple times, and HPMS data only available from one of two states.

Only one dangle is created on northern side of the border, results in incorrect connector being drawn.
Other Geometry Considerations: State Borders

Connecting neighboring states at borders taking into account proximity and attributes.
Outline

• Background and Motivation
• Building the Network
• Results and Testing
## Field Name Description

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBJECTID</td>
<td>ArcGIS object identifier</td>
</tr>
<tr>
<td>Shape</td>
<td>Geometry type</td>
</tr>
<tr>
<td>YEAR_RECOR</td>
<td>Year of data</td>
</tr>
<tr>
<td>STATE_CODE</td>
<td>Two digit state FIPS</td>
</tr>
<tr>
<td>ROUTE_ID</td>
<td>Route identifier</td>
</tr>
<tr>
<td>COUNTY</td>
<td>County code</td>
</tr>
<tr>
<td>FUNC_CLASS</td>
<td>Functional class</td>
</tr>
<tr>
<td>ASSUMED_FUNC_CLASS</td>
<td>Binary to indicate functional class was assumed</td>
</tr>
<tr>
<td>FAC_TYPE</td>
<td>Facility type</td>
</tr>
<tr>
<td>NHS</td>
<td>NHS designation</td>
</tr>
<tr>
<td>AADT</td>
<td>Average annual daily traffic</td>
</tr>
<tr>
<td>AADT_COMBO</td>
<td>Combo unit truck AADT</td>
</tr>
<tr>
<td>AADT_SINGLE</td>
<td>Single unit truck AADT</td>
</tr>
<tr>
<td>THRU_LANES</td>
<td>Number of through lanes</td>
</tr>
<tr>
<td>SPEED_LIMIT</td>
<td>Road speed limit</td>
</tr>
<tr>
<td>UR_CODE</td>
<td>Urban code</td>
</tr>
<tr>
<td>AADT_AVERAGED</td>
<td>Binary to indicate AADT values were averaged because of multiple HPMS events along feature</td>
</tr>
<tr>
<td>M_VALUES_FLAG</td>
<td>Binary to indicate inaccurate/inconsistent M-values received in input data and attributes may be inaccurate</td>
</tr>
<tr>
<td>REINSERTED</td>
<td>Facility type</td>
</tr>
<tr>
<td>NHS</td>
<td>1 or 2 indicates feature was reinserted for connectivity after facility type subset</td>
</tr>
<tr>
<td>COMMENTS</td>
<td>Optional state comments</td>
</tr>
<tr>
<td>LEN_MILES</td>
<td>Feature length in miles</td>
</tr>
<tr>
<td>Shape_Length</td>
<td>ArcGIS length field (meters)</td>
</tr>
</tbody>
</table>

### Network Comparisons

<table>
<thead>
<tr>
<th>Functional Class</th>
<th>New Network</th>
<th>FAF4</th>
<th>HM-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>47,069 miles</td>
<td>47,436 miles</td>
<td>48,191 miles</td>
</tr>
<tr>
<td>Class 1 &amp; 2</td>
<td>65,664 miles</td>
<td>62,963 miles</td>
<td>66,824 miles</td>
</tr>
<tr>
<td>Classes 1-4</td>
<td>464,213 miles</td>
<td>365,398 miles</td>
<td>468,882 miles</td>
</tr>
<tr>
<td>Classes 1-5</td>
<td>1,000,692 miles</td>
<td>433,606 miles</td>
<td>1,005,705 miles</td>
</tr>
</tbody>
</table>
Testing Network Flowability

• To be useful for many analysis the resulting network must be flowable and result in reasonable routes.
• How do we assess the quality of flows on this new network?
• We can flow the same OD pairs on the new network and another network we know flows reasonably well (i.e. FAF) and compare the resulting routes.
• Which OD pairs do we use for flow testing?
While population based OD pairs and methodology of flowing between all points informed the location of network breaks, route overlap was large, run times were long, and many areas weren’t being well tested.

- Populated areas > 50,000
- 54,615 routes– 24 hours solve time
Defining Origin - Destination Pairs

- Using county pairs from FHWA’s 2008 Auto Business Traveler Analysis Framework, the top 20% of county-to-county OD pairs originating from each county are used.
- Points are placed at the county’s highest populated place.
- Results in greater feature coverage and shorter run times.

76,020 routes – 3 hour solve time
## Network Flow Testing

<table>
<thead>
<tr>
<th>STATE</th>
<th>New Net Miles Flowed*</th>
<th>FAF4 Miles Flowed*</th>
<th>HM-20 Miles*</th>
<th>New Net Flowed % of HM-20</th>
<th>FAF4 Flowed % of HM-20</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>7,616</td>
<td>6,794</td>
<td>22,968</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>AR</td>
<td>6,809</td>
<td>6,155</td>
<td>20,162</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>AZ</td>
<td>2,891</td>
<td>2,860</td>
<td>11,580</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>CA</td>
<td>7,732</td>
<td>7,231</td>
<td>43,515</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>CO</td>
<td>5,622</td>
<td>5,652</td>
<td>15,657</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>CT</td>
<td>789</td>
<td>77</td>
<td>4,206</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>DC</td>
<td>54</td>
<td>53</td>
<td>298</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>DE</td>
<td>298</td>
<td>316</td>
<td>1,152</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>FL</td>
<td>6,425</td>
<td>6,116</td>
<td>19,268</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>GA</td>
<td>10,586</td>
<td>9,974</td>
<td>27,285</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>IA</td>
<td>7,404</td>
<td>7,110</td>
<td>24,926</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>ID</td>
<td>3,306</td>
<td>3,165</td>
<td>10,921</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>IL</td>
<td>8,331</td>
<td>7,925</td>
<td>28,197</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>IN</td>
<td>6,509</td>
<td>5,778</td>
<td>19,759</td>
<td>33</td>
<td>29</td>
</tr>
<tr>
<td>KS</td>
<td>8,359</td>
<td>7,732</td>
<td>32,494</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>KY</td>
<td>6,550</td>
<td>5,456</td>
<td>12,799</td>
<td>51</td>
<td>43</td>
</tr>
<tr>
<td>LA</td>
<td>4,217</td>
<td>4,075</td>
<td>11,707</td>
<td>36</td>
<td>35</td>
</tr>
<tr>
<td>MA</td>
<td>1,219</td>
<td>1,222</td>
<td>8,148</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>MD</td>
<td>1,593</td>
<td>1,560</td>
<td>5,769</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>ME</td>
<td>1,529</td>
<td>1,466</td>
<td>5,718</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>MI</td>
<td>6,609</td>
<td>6,095</td>
<td>32,647</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>MN</td>
<td>8,563</td>
<td>8,089</td>
<td>30,118</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>MO</td>
<td>9,120</td>
<td>8,008</td>
<td>28,286</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>MS</td>
<td>6,663</td>
<td>6,064</td>
<td>20,378</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>MT</td>
<td>5,547</td>
<td>5,482</td>
<td>14,228</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>NC</td>
<td>7,738</td>
<td>6,697</td>
<td>19,752</td>
<td>39</td>
<td>34</td>
</tr>
</tbody>
</table>

*Mileage includes FC 1-4 and FC-5 rural features
Network Flow Testing

Highest percent of HM20
(Kentucky – 51%)

Lowest percent of HM20
(Massachusetts 15%).

Note: Kentucky has 120 counties and Massachusetts has 14.
Assessing Network Flowability

Route mileages between the networks are compared. Outliers are identified and mapped.
Assessing Network Flowability

- Image shows a break in the network identified by comparing flow results to the FAF4.
- It can be hard to identify the location of the break causing the problem. We built some tools to help us quickly locate them.
- Shown at arrow is facility type 4 feature that was removed from the network, causing the break.
Challenges to Assessing Network Flowability

**FAF4**
- Typically less network detail.
- Functional class can differ from prototype network

**Prototype**
- Typically more network detail (perhaps too much).
- Functional class assignment may differ from FAF.
- Connectivity breaks.
• There have been many challenges but significant progress has been made in developing a non-proprietary flowable national network from state provided ARNOLD and HPMS data.
• This is still a work in progress.
• We would appreciate any input that you may have.
Thank you!

Questions?