CONFLATION PROCEDURE TO COMBINE SPEED AND VOLUME DATASETS

INTRODUCTION

This documentation introduces the conflation process to both technical and non-technical audiences so that managers can get a general understanding of the conflation process while technical staff can take the steps described and perform the conflation of networks. For each step in the process, there is a general description, the tools necessary to complete the step, and any important points to keep in mind.

This paper describes the entire conflation process for spatially joining the attributes of the INRIX roadway network ("Join Map") which is based on the Traffic Message Channel (TMC) mapping network to the HPMS roadway network ("Base Map") based on the individual roadway shapefiles provided by each of the state departments of transportation that supply data to FHWA to compile the HPMS database. The two roadway networks contain various levels of roads including freeways, tollways, arterial streets and many lesser streets. For purposes of producing the NPMRDS road network, only the NHS roads were considered.

DATA PREPARATION

The reason that the conflation process is both necessary and tedious is that very rarely do the roadway networks that are being conflated actually overlay perfectly. Often times the centerlines for the roadbeds do not align and additional resources must be used to match the proper roadway segments together. They rarely will have any common attributes that make a "table join" possible either. It takes a great deal of resources to ensure that the proper alignment of the two networks occurs. It is possible to eliminate some of the roadway network that might not be necessary for the analysis. Some pre-processing of the data can save a great deal of time during the conflation process. For example, in conflating for the NPMRDS road network, only the functional classes corresponding to the NHS are included. The elimination of all other functional classes will shrink the size of the GIS shapefile and speed up the processing time. Also, some records might be incomplete so these roadway links in the shapefile can be eliminated as well.

It is important to note that one of the first decisions that needs to be made during the conflation process is the determination of which roadway network is going to be the master network. The master is the roadway network which gets the attributes from the other network loaded on it. Generally, the

master network should be the network that more closely aligns with the purpose for the analysis. For example, if a state department of transportation is performing a congestion analysis on a corridor, then the master network should be the DOT's roadway inventory network rather than the speed network. This will ensure that the results of the analysis can be shown in the terminology and mapping units that correspond to the DOT rather than to the speed data road network. There are some cases where the speed data network might be preferable as the master. One example might be where GIS shapefile that is being matched to the speed shapefile has some quality concerns. It might be preferable to conflate to the "better-mapped" network for purposes of displaying the results. This decision is an important one and needs to be contemplated prior to beginning the conflation process.

So some key questions to answer during the preparation portion of the analysis include:

- Can any records be eliminated such as excluding certain functional classes to speed the processing time?
- Are there any records with known data concerns that can be eliminated prior to processing?
- Which mapping network is going to be the base network?

Here are the typical steps to prepare the data for conflation:

1. Clip out the region of the "Base Map" layer using a state polygon layer (use the "select by location" method and select the "have their centroid in" option).

2. Export the selected "Base Map" to a new layer. (Tool: right click the "Base Map" layer on the layer list and select Data>Export Data...>Export: Selected Features). Do step 1 and 2 to the "Join Map" layer if needed.

3. Project the new "Base Map" layer to a coordinate system consistent with the "Join Map" layer. If necessary, project both "Base Map" and "Join Map" layers so that the default map unit is foot or meter instead of decimal degree. The following steps assume that all layers are projected. (Tool: ArcToolbox>Data Management Tools>Projections and Transportations>Feature>Project)

4. Prepare the attribute table of the "Join Map" for processing:

   a. Add an ORN_LENG field (use Double for data type) to the table. This field will store the original segment lengths. Right click the header of new column and select "Calculate

Geometry..." to calculate the segment length. (Tool: open attribute table and click Options>Add Field...)

b. Add an ORN_FID field (use Long Integer for data type) to this table and copy the FID values to the newly added field. This field stores the original FID permanently and will serve as a unique ID necessary during the final quality control step (see Step 12 for more information). The original FID field, which is automatically populated by ArcGIS when the layer is loaded, is dynamic, meaning a FID value is not permanently associated with the same segment.

THE CONFLATION PROCESS

The conflation process can appear very complex to persons not familiar with GIS systems. Because of this complexity, the process is described at a very high level and also at the technical level to actually perform the operations necessary to complete the process. This example used in this paper to describe the process is the one that would be used to produce the NPMRDS road network. The INRIX Traffic Message Channel (TMC) speed map ("Join Map") is conflated onto the HPMS roadway inventory data map ("Base Map").

High Level Description

Below are listed the steps necessary to carry out the conflation process. These steps, at a very high level, are as follows:

- Convert the "Join Map" into a point layer showing the starting and ending point of each link (Exhibit 2)
- Create a buffer around each point (Exhibit 3)
- Consolidate and eliminate any points that overlap or are adjacent (Exhibit 4)
- Break the "Base Map" lines into shorter lines using the "Join Map" points as breakpoints
- Create a spatial buffer around each "Join Map" link in order to join the appropriate "Join Map" link with each split "Base Map" link (Exhibit 5)
- Join the "Join Map" links to the "Base Map" links
- At this time, the conflation process is complete; however, there may be links that did not get matched so this process may need to be repeated on unmatched links to determine if there is a match that did not fit the parameters used for matching in the first iteration.

- Combine all layers if more than one iteration was needed
- Perform quality control checks on the conflation results

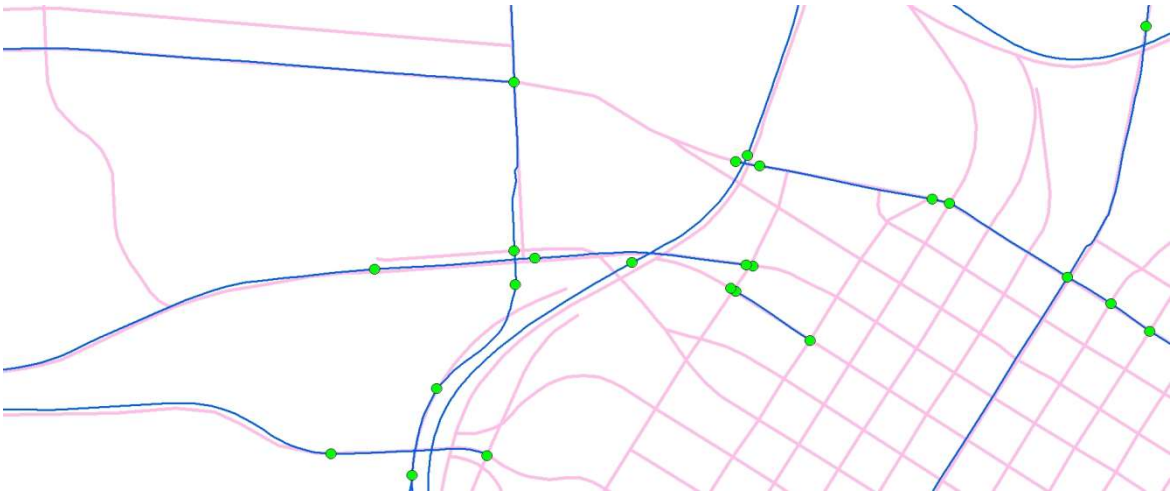

Exhibit 2. Example of a Point Layer with Starting and Ending Points

Exhibit 3. Create Buffers Around the Points

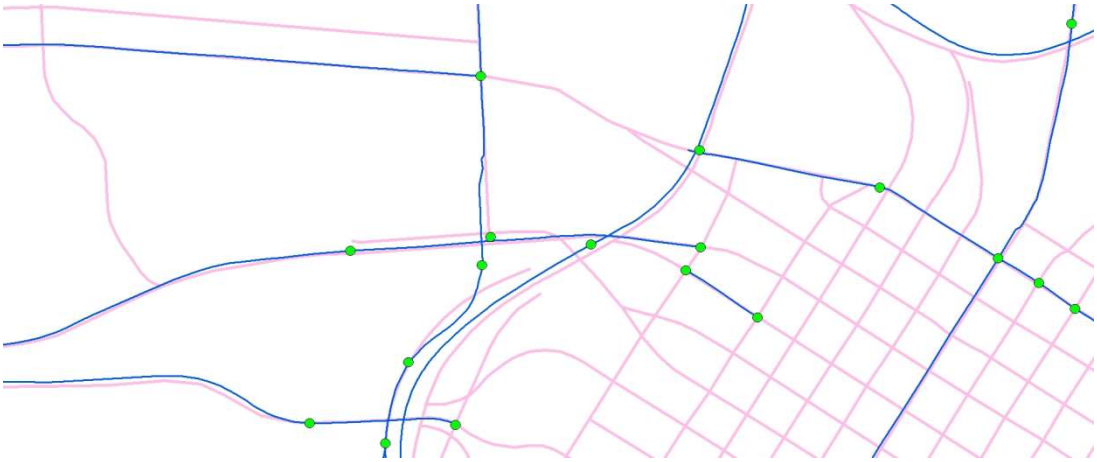

Exhibit 4. Break the "Base Map" Layer

Exhibit 5. The Spatial Buffer on "Join Map" Segments

The quality control step is crucial to ensure that the analysis will produce the desired results. Obviously data errors exist in any database and could have been present in either the "Base Map" or "Join Map" datasets prior to conflation. However, some errors can result directly from the conflation process. Because of the potential for data errors created due to conflation, it is necessary to spend some manual quality control time to ensure

the overall accuracy. Some examples of issues that can arise during quality control include:

- "Base Map" segments that should have been conflated with corresponding "Join Map" segments but were left unprocessed. This could result from very poor or inconsistent mapping along a stretch of roadway
- "Base Map" segments that should not have been conflated with any "Join Map" segments but were matched due to their proximity to other "Join Map" segments.
- "Base Map" segments that were conflated with inappropriate "Join Map" segments. This happens in areas with HOV lanes, freeway frontage roads, and at some complex interchanges.