CONFLATION: EDGEMATCHING TOOLS AND WORKFLOWS

DAN LEE
Agenda

Conflation Overview and Geoprocessing Tools

Edge Matching Workflow

➢ Demo

Conclusions and Future Work
Conflation Overview and Geoprocessing Tools
When using multi-source spatial data together

Common obstacles in analysis and mapping:

➢ Spatial and attribute inconsistency caused by differences in data collection and modeling
➢ High cost to fix the problems
Conflation reconciles multi-source datasets and optimizes data quality and usability

Between overlapping datasets:
- Detect feature changes (differences) through feature matching
- Make spatial adjustment and attribute transfer

Between adjacent datasets:
- Detect and resolve feature conflicts and disconnections through edge matching and alignment

Ultimately:
- Maintain an unified and seamless dataset – enriched and up-to-date
- No longer live with various imperfect datasets
- Rely on the data to perform analysis and quality mapping with confidence

What does it take to achieve the goals?
Our initial focuses

Develop highly automated tools in Geoprocessing framework

- Starting with linear features (roads, parcel lines, rivers, etc.)
- Aiming at high accuracy (not promising 100%)
- Providing information to facilitate post-processing

Build workflows

In ArcGIS 10.5.1 and Pro 2.0

Have you used these tools in ArcMap?

TW session (01:30pm – 02:45pm, Thursday, Room 31A)
Conflation Tools and Workflows: An Introduction
Edge matching (EM) tools for adjacent datasets

Based on proximity, topology, and continuity analysis, as well as attributes information

Generate Edgematch Links (GEL)
- From source features to adjacent features

Followed by Edgematch Features (EF)
- Connects features guided by the established links
Options for connecting features

<table>
<thead>
<tr>
<th></th>
<th>Input Features Only</th>
<th>With Adjacent Features</th>
<th>With Border Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOVE_ENDPOINT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
</tr>
<tr>
<td>After</td>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADD_SEGMENT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
</tr>
<tr>
<td>After</td>
<td><img src="image10" alt="Graph" /></td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADJUST_VERTICES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td><img src="image13" alt="Graph" /></td>
<td><img src="image14" alt="Graph" /></td>
<td><img src="image15" alt="Graph" /></td>
</tr>
<tr>
<td>After</td>
<td><img src="image16" alt="Graph" /></td>
<td><img src="image17" alt="Graph" /></td>
<td><img src="image18" alt="Graph" /></td>
</tr>
</tbody>
</table>

- Input feature
- Edgematch link
- Adjacent feature
- Border feature
- Adjusted input feature
- Adjusted adjacent feature
- Adjacent feature as reference
- New ending location
Align Features

Based on proximity, topology, and similarity analysis, as well as attributes information.
Edge Matching
Workflows
Conceptual workflow
Example edgematching of adjacent datasets
Goal - make two adjacent datasets properly connect

Source features

Adjacent features

Together
Results

Move endpoint
Conflation workflow in real world scenarios

- **Preprocessing**
  - In same projection
  - Data validation
  - Selection of relevant features

- **Conflation and evaluation**
  - Conflation tools
  - Workflow tools

- **Postprocessing**
  - Queued review
  - Interactive editing
Supplemental tools and guidelines for download

http://www.arcgis.com/home/item.html?id=36961cde1b074f1f944758f6abec87cc
You can also search by “conflation” at arcgis.com to find the download.

We are improving the add-in toolbar and workflow tools.
Demo: Real world scenario
Breakdown of the conceptual workflow into sub-steps

- **Step 6a**: QA
- **Step 6b**: QA
- **Step 7**: QA
Demo data overview

Two road datasets (an area in Alabama):

- **EdgeRoads** – 7576 features
- **GISRoads** – 3634 features

Both datasets:

- Contain roads that are within 1 km to borders
- Have inconsistent road names
GEL result

Generated 454 links; midpoints of links were created for visualization purpose. Borders were not in the process, but displayed for reference.
EM_CONF in output

- 100 (matched with no ambiguity)
- 50 (spatially matched with unmatched attributes)
- < 50 (spatially matched with some ambiguity and weak continuity)
GEL evaluation results

EM_CONF < 33: 134 links
Intersecting links: 33 locations
Potential missing links: 62 source dangle locations

It's time for inspection ...
Inspection and editing of edgematch links

Reviewed:
- 33 locations of intersecting links \(\text{NEAR\_DIST} \geq 0\)
- 98 low EM\_CONF links \((\text{EM\_CONF} < 33) \text{ AND } (\text{REV\_FLAG IS NULL})\)
- 62 source dangle locations (near links)

Summary:
- 388 (~85%) of total 459 links were good (54 were flagged for recheck)
- 71 (~15%) of total links were modified, removed, or added

<table>
<thead>
<tr>
<th>OBJECTID</th>
<th>FREQUENCY</th>
<th>REV_FLAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>267</td>
<td>&lt;Null&gt;</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Added</td>
</tr>
<tr>
<td>3</td>
<td>66</td>
<td>Good</td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>Modify</td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>Recheck</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Remove</td>
</tr>
</tbody>
</table>
What happened to the SRC_FID and TGT_FID of the added or modified links?
Edgematich Features
Edgematic result

Review flagged locations …
Edgematching of adjacent datasets workflow completed!

### Automated processing

<table>
<thead>
<tr>
<th>Step</th>
<th>Processing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 6a</td>
<td>6.52 sec</td>
</tr>
<tr>
<td>Step 6b</td>
<td>4.09 sec</td>
</tr>
<tr>
<td>Step 6c</td>
<td>2.15 sec</td>
</tr>
<tr>
<td>Total</td>
<td>12.76 sec</td>
</tr>
</tbody>
</table>

### Interactive processing (not counting final review)

<table>
<thead>
<tr>
<th>QA Links</th>
<th>Time (2-3 review counts per minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Count (locations or feature groups)</td>
<td>~ 193</td>
</tr>
<tr>
<td>Edit Count (field values)</td>
<td>192</td>
</tr>
<tr>
<td></td>
<td>~ 1 - 1.6 hrs.</td>
</tr>
</tbody>
</table>
Conclusions and Future Work

Thanks to:

- Department of Public Works (DPW), Los Angeles County, USA.
- Resource Management Service, LLC, Birmingham, AL, USA.
- All others who supported us along the way.
Edge matching can be done more efficiently now

It takes a workflow:

- Use the best practice in preprocessing.
- Run automated tools to obtain highly accurate results and evaluation information.
- Interactively review and edit the results. The time is worth-investing.
Consider conflation a higher priority

Study the tools and workflows; understand the results

➢ Start with small test areas

Customize the workflows for your organizations

➢ Improve data quality and usability
➢ Bring new live and value to your data

Work with broader communities

➢ Data sharing and collaboration
➢ Seamless analysis and mapping

Please send us your feedbacks and share your stories ... 😊
Future work

New tools and enhancements
- Split Line By Match tool (for overlapping datasets)
- Better feature matching
- Tools for other feature types

Integrated processing and inspections
- Design of Conflation Manager is underway

Formalization of workflows
- Common scenarios (e.g. multi-scale data updating, linking buildings of different scale)
- Incorporation of other data sources (imagery, lidar, GPS)
- Contextual conflation (spatially related features)
Conflation Manager (ConfMgr)
Conflation in multi-scale data updating and mapping

*DLM* – digital landscape model; *DCM* – digital cartographic model

Source 1

Source 2

Both

Conflated

Generalized

Conflated & updated

Re-generalized

*DCMs*

*DLMs*
Recent papers


➢ Lee D (2015), Using Conflation for Keeping Data Harmonized and Up-to-date, ICA-ISPRS Workshop on Generalisation and Multiple Representation, 2015, Rio de Janeiro, Brazil. 


Please Take Our Survey on the Esri Events App!

Download the Esri Events app and find your event

Select the session you attended

Scroll down to find the survey

Complete Answers and Select “Submit”

Your feedback allows us to help maintain high standards and to help presenters

Thank you for attending!