Establishing Classification and Hierarchy in Populated Place Labeling for Multiscale Mapping for The National Map

Stephen Butzler, Cynthia Brewer and Jim Thatcher
RA: Doug Minnigh

Gould Center, Geography
Pennsylvania State University
The Problem

Effective dynamic labeling of place through scale with database driven mapping

Current federal standard - GNIS point data
- Limited attribute data for label hierarchy
- Placement of label limited to anchor point

Project Goals
- Additional attribute data for improved hierarchy
- Place polygons for improved placement
Portion of USGS ScaleMaster work (24K, 50K, 100K...)

[Map showing various locations and landmarks in St. Louis, Missouri, with highlighted areas and labels such as 'The Ville', 'Brooklyn', 'St. Louis', 'Eads Square Park', 'Carr Square', 'Kosciusko']
289K: With our basic hierarchy... and with no place hierarchy
578K: With our basic hierarchy... ...and with no place hierarchy
Basic Approach

Census Polygons
- Initial categorization by Census Class Codes (FIPS-55)
- Added attribute data and added additional hierarchy

GNIS Points
- At large scale uses points for places not represented by polygons
- Added attribute data and added additional hierarchy
Class codes further define census polygons:

Class **C** — incorporated places
Class **M** — military places
Class **P** — pop. place = incorp. place
Class **T** — MCD (e.g. township)
Class **U** — CDP

Note: We ranked MCDs coextensive with CDPs (T/U) higher than sole MCDs (T).

*Pittsburgh*
Census Data Types

2009 TIGER/Line Shapefiles (geography)
• **Incorporated** Places, Census Designated Places (CDPs)
• Minor Civil Divisions (MCDs), no CCDs used
• **Economic Places**
  (NOTE: 2007 Economic Census definitions of Economic Place include all places with EITHER 5000 residents OR 5000 employees)

2002 Survey of Business Owners (attributes)
• part of Economic Census (every 5 years)
• provides data on employment numbers (used to create our classifications)
• place-level data for the 2007 **SBO** was not yet available at time of analysis
Top 4 of 5 classes in R1

2.3M
### PA Place Hierarchy

<table>
<thead>
<tr>
<th>Place</th>
<th>C</th>
<th>T</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pittsburgh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West Mifflin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bethel Park</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mount Oliver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penn Hills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cranberry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reserve</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mount Pleasant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastonville</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shadyside</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### CO Place Hierarchy

- Denver
- Aurora
- Englewood
- BowMar
- Columbine
- Roxborough Park
- Sullivan

### Table

<table>
<thead>
<tr>
<th>C</th>
<th>T</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorpor.</td>
<td>Economic Place</td>
<td>MCD</td>
</tr>
</tbody>
</table>

*U4 and U6 GNIS points (no polygons)*

*Unique dashed outline*
Evaluating Our Results

1. Are labels placed on maps?
2. Does the additional label hierarchy provide value?

We compared our results to *The National Map Viewer* and *Google maps* at three scales ...
Pittsburgh at 1:144,448 using our place classing with The National Map Viewer and Google maps comparisons
Yes... different sets of other features labeled and different label sizes among compared maps interfere with precise comparisons—this is pilot work.
Pittsburgh at 1:577,791 using our place classing with The National Map Viewer and Google maps comparisons
Evaluation: Label Inventory Process

289K placenames

6 pages of placenames

Google map

TNM Viewer

PSU
Labels compared to PSU map

PSU map **did worse** than Google and/or TNM Viewer

PSU map **did better** than Google and/or TNM Viewer

Many more township names on PSU map

Incorp. places not on PSU map

Incorporated and economic place
MCD, CDP and economic place
Incorporated place
Other MCD/CDP

number of labels — 30 25 20 15 10 5 0 5 10 15 20 25 30 35 40 45

- 144K
- 289K
- 578K
Other Possible Attributes

Additional hierarchy with census data

• In pilot, used SBO “**Number of Employees**” but there are other indicators ...

• For places:
  – Total population

• For economic places:
  – Total number of firms
  – Total revenue
Labeling results of original methodologies
- Place, MCD and Economic Place
- 2002 SBO Data
- Number of Employees

Labeling results of altered methodologies
- Place and MCD
- 2000 Decennial Census Data
- Total Population

Map with labeled data for Pittsburgh area.
Labeling results of original methodologies

- Place, MCD and Economic Place
- 2002 SBO Data
- Number of Employees

Labeling results of altered methodologies

- Place and MCD
- 2000 Decennial Census Data
- Total Population
Further Work

Next phase of project
• Simplified version of methodology (Population data)
• Updated versions of TIGER/Line Shapefiles
• Addressed issues in existing methodology
  1. Additional Hierarchy for GNIS points
     Adding attribute data to establish hierarchy
  2. Urban/Rural Divide
     Introduced partition according to Census defined Urbanized Area and apply two separate labeling schemas
At large scale GNIS points have limited hierarchy

At small scale GNIS point labels are competing with polygon labels
Data Types
2010 TIGER/Line Shapefiles
• Block level geography
2010 Decennial Census
• Population density

Additional Method 1 - Process
Processing
Create continuous population density surface
• Join population density data to block polygons
• Collapse block polygons to points
• Using Inverse Distance Weighting create interpolated surface of population density
Spatial join GNIS points to new surface

Additional Method 1 - Process
Hierarchy of GNIS points
- Three levels under polygon labels
  - Oakland
  - Hazelwood, Shadyside
  - Glenwood, Carrick, North Side

Additional Method 1 - Samples
Increased control within GNIS point layer
- Oakland, East Liberty, Banksville

Additional Method 1 - Samples
Additions to Methodology

Next phase of project

• Simplified version of methodology (Population data)
• Updated versions of TIGER/Line Shapefiles
• Addressed issues in existing methodology

1. Additional Hierarchy for GNIS points
   Adding attribute data to establish hierarchy

2. Urban/Rural Divide
   Introduced partition according to Census defined Urbanized Area and apply two separate labeling schemas
Disparity of attribute data
- Does not allow hierarchy to emerge in rural areas
- At decision points causes rural labels to drop faster

Additional Method 2 - Justification
Disparity of attribute data
• Does not allow hierarchy to emerge in rural areas
• At decision points causes rural labels to drop faster

Additional Method 2 - Justification
Additional Data
2000 TIGER/Line Shapefiles
• Urbanized Areas and Urbanized Clusters
Generalize only Urbanized Areas for regional decisions
- MCD, Place, GNIS points relationship to generalized Urbanized Area
- Develop alternate set of labeling rules for Urban/Rural
Underlying Generalized Urbanized Area
• Urban Schema – Pittsburgh, Plum, Monroeville
• Rural Schema – New Castle, Butler, Kittanning, Indiana

Additional Method 2 - Samples
Underlying Generalized Urbanized Area
- Urban Schema – Pittsburgh, Plum, Monroeville
- Rural Schema – New Castle, Butler, Kittanning, Indiana

Additional Method 2 - Samples
Next steps

• Test the methodology to see if easily adaptable to other states
  – With eastern states that have similar administrative parameters
  – Update western states model and rerun
• Incorporate method into an overall cartographic design with competing content through scale
• Evaluate
Acknowledgements

USGS Center of Excellence for Geospatial Information Science (CEGIS), 2007 to present

Resources
Project resources:  
ScaleMaster.org
Gould Center:  
NationalMapping.us
CEGIS-USGS:  
cegis.usgs.gov

Thanks:
Wes Stroh  
Penn State
Babs Buttenfield  
CU-Boulder
Lynn Usery  
Larry Stanislawski  
USGS-CEGIS
Charlie Frye  
ESRI