Integration of folksonomies into the process of map generalization

Meysam Aliakbarian and Robert Weibel
GIScience Center, Department of Geography
University of Zurich

AGILE 2016
Workshop on Generalisation and Multiple Representation
Agenda

- Motivation
- Background
- Methodology
- Results
- Outlook
Motivation

- User Generated Content (UGC) is growing in:
  - quantity
  - diversity
  - quality
- Most of UGC content have a geographic component
- Cartographic products based on UGC content need to focus more on the hidden knowledge.
Background – Pillar I

Folksonomies

- Taxonomies formed by tagging behavior of user
- Studies are on relation \( (y) \) between users, resources and tags:
  - \( F = f(u,r,t,y) \)
- First examples are based on tagging behavior of user in social bookmarking systems.

Background – Pillar II

OSM

- Features are stored in form of nodes, ways, relations
- Users enrich features with tags (key, value)
- Tagging policies are agreed upon users in OSM wiki
- Shared agreed tags can be taken as users common agreement on features
- Folksonomy:
  - features -> resource
  - tags -> tags
Background – Pillar III

Semantic similarity

• Notion of similarity (relatedness) between concepts
• Level of commonality is being measured
• Different measures are available in the literature:
  • Jaccard
    \[
    \text{sim}_{\text{Jaccard}}(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}
    \]
  • Dice
    \[
    \text{sim}_{\text{dice}}(X,Y) = \frac{2|X \cap Y|}{|X| + |Y|}
    \]
  • Cosine
    \[
    \text{sim}_{\text{cosine}}(X,Y) = \frac{X \cdot Y}{\|X\|\|Y\|}
    \]
Background – Pillar IV

Generalization

• Taking constraint-based modeling as state of the art
• Three-fold relation between *constraints, measures* and *relations* (Steiniger and Weibel 2007)
Methodology

Semantic similarity in Generalization

• Calculating feature-feature similarity
  • Taking a feature as a central feature
  • Calculating similarity of other features (feature-feature)
  • Including the measures into generalization operator
• Notion of similar/dissimilar
  \[ s = \text{Sim}(X, Y) : \begin{cases} 
  s \leq \alpha \rightarrow \text{similar} \\
  \alpha \leq s \leq \beta \rightarrow \text{similar} \\
  s > \beta \rightarrow \text{test if } X = Y
\end{cases} \]
• Including values in similarity measure
  \[ \text{Sim}_{\text{KeyValue}}(X, Y) = \frac{2|K_X \cap K_Y|}{|X| + |Y|} \frac{|V_X \cap V_Y|}{|K_X \cap K_Y|} \]
Methodology

Modification of Generalization operators

• Selection
  • Selecting semantically similar features
  • Selecting semantically dissimilar features
Methodology

Modification of Generalization operators

- Aggregation
  - Aggregating semantically similar features to a new feature
- Spatial constraint is needed
Results

Search feature:
- amenity: restaurant
- cuisine: turkish
- name: Hazev

- amenity: restaurant
- name: Carluccio's
- opening_hours: Mo-Th 07:30-23:00;
- phone: +44 20 7...
- Website: www.carluccios.com

- amenity: fast_food
- cuisine: sandwich
- name: Subway
- amenity: fast_food
- cuisine: sandwich
- name: Baguette...

- amenity: fast_food
- cuisine: asain
- name: Lemongr...
Outlook

Outlook of the study

• Working on other generalization operators
• Tag-Tag analysis
• Spatial-Semantic combination
• Property inheritance for features
Thank you!

Meysam Aliakbarian
meysam.aliakbarian@geo.uzh.ch

Robert Weibel
robert.weibel@geo.uzh.ch
References
