Visualizing Temporal-Semantic Relations in Dynamic Information Landscapes

Vedran Sabol, Know-Center Graz, Austria
Arno Scharl, MODUL University, Vienna, Austria

http://www.know-center.at

© Know-Center - gefördert durch das Kompetenzzentrenprogramm
Overview

- Problem & Motivation
- Our approach: dynamic information landscape
- Algorithms
- Future work and Conclusion
- Questions
**Motivation**

- Visualizing different dimensions (aspects) of data sets:
  - topical relations, geographic information, vocabulary, trends...

- Patterns not only within each separate dimension, but also between dimensions
  - Specialised representation for revealing patterns in each dimension
  - Coordinate all views (CMV)

- Developed within the RAVEN project (Relation Analysis and Visualization for Evolving Networks)
  - Builds upon results from previous projects: Geo-Vis, Ontology Browser, Tag Clouds, Information Landscape ...
  - Add analysis of the temporal dimension
RAVEN User Interface Mock-Up

- Interface for non-temporal analysis available (IDIOM project: http://www.ecoresearch.net/climate/)

- Temporal analysis: **Dynamic information landscape**
  - Additional components (optional): ThemeRiver/Timeline Visualisation ...
Dynamic Information Landscape

Information landscape:

- Distance: a measure of (topical) relatedness
- Height: encodes cluster size
- Density: conveys cluster cohesion
- Colour coding: distribution of selected features/metadata

Visualise changes as tectonic processes where hills and islands:

- grow and shrink
- merge and split

Experience with small data sets (MISTRAL and WebRat projects)
Dynamic Information Landscape

Knowledge Planet View

- Alternative to a standard 2D approach
- Project the Landscape onto a sphere instead of a plane
- Continuous space with no edges and preferred directions (axes)
- Point of interest in highest detail, non-relevant information hidden
- Animated, dynamic topography by morphing the 3D geometry between time points
Principle

Place landscape so that it is orthogonal to the time axis of the temporal views

Landscape becomes a slice of a temporal view for the selected time interval

Landscape contains only data from the selected time interval

Navigate along the time axis, landscape adapts smoothly
Algorithms

- Large (up to millions), dynamic data sets
  - → Scaleable, incremental algorithms required
- Clustering algorithm produces a balanced hierarchy
  - K-means with split and merge strategies (ISODATA)
  - Hierarchical Agglomerative or Affinity Propagation for seed initialisation
- Projection algorithm (InfoSky) uses the balanced hierarchy
  - Children of each cluster are positioned with a force-directed placement method
  - Subdivide the area using Voronoi diagrams and proceed recursively
- Fast: 10000 documents in about 1 minute on a modern machine
- Scales well: approx. O(n*log(n))
  - Good results for 100000+ documents

Authorized employees only
Incrementally compute the landscape for predefined, hierarchically organised time intervals

Morph the geometry of the landscape as the selected time interval changes in real time
Examples – Incremental Landscape

- Around 10000 documents
- Objects marked red are removed
- Incremental computation is hard for large data sets
Future Work

- Improve the incremental behaviour of the clusterer
  - Reduce the tendency of making large changes in the partition for small changes in the data set
  - Running time should be closer $O(\#\text{changes})$ than $O(\#\text{documents})$
- Extend the projection algorithm to fully support spherical geometry
- Implement geometry morphing for the knowledge planet
- Integrate everything into a consistent GUI using a coordination framework
- Perform user tests
Conclusion

Dynamic Information Landscape:

- For large data sets visualises two orthogonal aspects of the data within a single view:
  - Complex topical relationships
  - Change information

- Relations and patterns within each aspect and also between both aspects can be understood
  - Potential advantages compared to using multiple views (?)
Questions

With respect to usability and cognitive load on the user what are the advantages/disadvantages of:

- employing one dynamic representation compared to several, static views to convey the same change information?
- projecting the information landscape onto a sphere compared to projecting it onto a plane?

Thank you!

Suggestions and further questions are welcome