Representing Spatio-Temporal Events Across Scales

P. Raposo, A. C. Robinson

GeoVISTA Center, Department of Geography, The Pennsylvania State University, 302 Walker Building, University Park, PA 16802
Email: [paulo.raposo; arobinson]@psu.edu

1. Introduction

Recent advances in text mining and geocoding make it increasingly easy to collect, extract, and categorize events from unstructured text sources such as news articles. The resulting datasets can include events with references to multiple locations; for example, a political protest in a specific part of a city, which is related to a country as well as transnational political situations. There can also be multiple temporal references in these event data. Both aspects present challenges in designing interfaces to geovisual analytics systems to support pattern recognition and reasoning.

Here we present progress on developing dynamic cartographic representation techniques to support visual analytics of space-time event data, with a specific focus on how to handle spatial references at multiple scales. First, we demonstrate how a global hierarchical tessellation can be used as a event aggregation technique reconfigurable to multiple spatial resolutions. Second we explore the use of data-driven tint bands to represent country and regional-level place mentions in event datasets using political boundaries. Through case study examples using events gathered from news reports on conflict and political events in Syria, we show how both techniques can be blended together in order to explore event data across multiple spatial scales in the STempo visual analytics environment. Finally, we present ideas for future research and highlight some of the challenges we have encountered so far in using these cartographic approaches with spatio-temporal event data.

2. Background

A significant body of previous research has explored the general need for new techniques to support decision making using spatio-temporal data in visual analytics systems (Andrienko et al. 2007). Geovisual representation techniques that have been proposed for event data include the creation of event trajectories and clusters (Andrienko et al. 2011), which can then be visualized using space-time cubes (Kraak 2008). Many related examples focus on event data that correspond to human, vehicle, or animal movement. In our case, we focus on events that correspond to human activity that are not characterized by raw movements alone. Attributes of event data gathered from news articles can include political motivations, references to other events taking place in the past, present, or future, and other forms of context (e.g., geographical, temporal, and social). Representations for exploring these data must go beyond showing physical progression in space and time from one place to another. Instead, we must be able to explore the complex contexts associated with multiple events, and to be able to compare events of different types (protest events vs. diplomatic events, for example).

In earlier work we have developed an approach that uses T-pattern analysis (Magnusson 2000) for identifying significant co-occurrences in space-time event data from news articles (Peuquet et al. 2015), and shown how these T-pattern results can be explored to support reasoning in a coordinated multiple-view geovisualization environment we call STempo. In the sections that follow we highlight related work and our next steps to enhance the utility of the mapping features in STempo—a key area of need identified by a recent user evaluation conducted of STempo (Robinson et al. 2016).

3. Visual Methods for Representing Large Spatio-Temporal Event Collections

In order to visualize data at two different geographic levels of precision, we use a hierarchical global tessellation combined with cartographic tint bands. Both are used in the STempo environment as choropleths; the tint bands allow us to represent data at a national level of precision, while the tessellation allows for multiple bin sizes for aggregations of data geocoded to higher spatial precision.

3.1 Using a Quaternary Triangular Mesh for Event Data Representation

In the 1980s, Geoffrey Dutton proposed the use of a Quaternary Triangular Mesh (QTM) for supporting hierarchical representation in geographic data structures (Dutton 1989). The QTM uses an octahedron projected onto the Earth, which each triangular facet subdivided into four more triangles, recursively. One of the advantages of the resulting triangular mesh is the efficiency with which addressing systems can be applied to it. As each facet of a QTM is a triangle, facets can be combined to form larger shapes such as quadrilaterals and hexagons, two commonly used binning geometries in thematic mapping. Also, since a QTM subdivides recursively, it inherently stores or represents spatial data at multiple spatial resolutions. The combination of high performance for data retrieval, recomposable tessellation, and the ability to map data to multiple resolutions motivated us to explore using a QTM to support visual analytics of event data from news articles, where large collections of observations require analysis at multiple scale perspectives in real time.

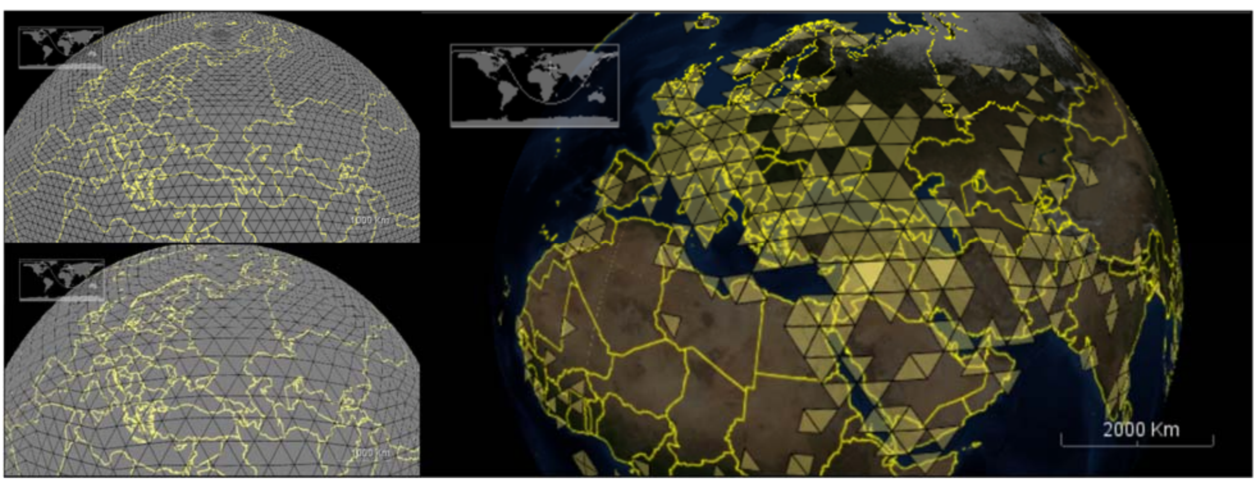

Figure 1. Using a QTM to Represent Event Data in STempo

Figure 1 provides an example of our visualization using a QTM. At left, QTM subdivision levels 5 and 6 are shown to illustrate how the method can use variably-sized triangular bins, thereby enabling multi-resolution analysis and visualization. On the right, sample data is visualized using a QTM by changing the opacity of each triangular spatial bin to reflect the number of news stories geocoded to a point within it, thereby producing a density or heat map. The data can be re-binned and redrawn to any level of the QTM (i.e., any subdivided triangle size).

QTM bins are triangles, and therefore introduce an anisotropic sampling bias due to their non-uniform orientations. Also, while individual bins can be aggregated into forms such as hexagons, these aggregated forms cannot be equivalent in area over large portions of the Earth. Thus a QTM, while appearing spatially uniform, is not, and therefore does not provide a perfectly regular sampling or representing geometry. The same is true, however, of square bins or pixels over large portions of the Earth. We believe the sampling bias introduced by a QTM is out-weighed by its benefits for global-scale tessellation.

3.2 Representing Nation-Level Place Mentions Using Tint Bands

Tint bands are a standard cartographic representation where the margins around countries or other polygonal bodies are colored or shaded. They are commonly used in reference maps to differentiate neighbouring countries using different colors. We employ them here as a graduated symbol, setting the tint band opacity for a given country to reflect the number of news stories geocoded to that country (and not to a finer spatial resolution). Since tint bands use only the marginal area of a country polygon, they are an efficient means of representing an areal value in terms of how much map space they occupy. Hence, they introduce relatively little map clutter, and cooperate well with other cartographic symbol types. Figure 2 provides an example of our use of tint bands to represent nation-level data.

Figure 2. Using Tint Bands to Show Place References in Event Data in STempo.

As used here, tint bands have the same problems choropleth maps do, in that non-normalized numbers shown across variably-sized polygons provide a visualization that is potentially misleading, at least to a user who isn't trained to ignore differences in area.

We use QTM bins and tint bands together as a means of representing data at different spatial resolutions on the same display. When data can be geocoded precisely (i.e., to a city or neighborhood), we bin and represent it with a QTM, and when data is geocoded to the spatial precision of a country, we bin it by country and represent it with tint bands. Figure 3 provides an example of such a combined data view. In this case, tint bands are blue to represent politically-categorized data events at a national level, and QTM bins are red to represent military-categorized data events at more spatially precise levels.

Figure 3. Using a QTM and Tint Bands to Show Place References in Event Data in STempo.

4. Looking Ahead

We have shown here how a QTM and tint bands can be used to visualize different spatial dimensions for complex spatio-temporal event datasets. A QTM has the potential theoretical advantage of being recomposable into hexagonal and quadrilateral shapes. In practice, we have noted that sampling bias in the geometry of a QTM is one potential limitation associated with using this method.

Tint bands can also be used to show event density among country or regional political units, which can help provide a more effective overview that leverages place hierarchy knowledge gained by geocoding places in event references using a system like Geonames which supports hierarchical references to smaller scales. As described above, tint bands become problematic in the same way that traditional choropleth maps do when looking at locaions with significant areal size differences, an effect exacerbated by some map projection distortions. This is effect is partly mitigated by using a interactive virtual globe, as we do in STempo, because the user can easily pan and zoom, bringing areas of interest into the area under the user's gaze, where distortion is minimal. On the other hand, tint banding may be helpful to users who want an overview at the country or regional level, and at the same time would benefit from retaining uncluttered geographic context in the form of labels, shaded relief, and other geographic data within the polygonal areas of countries or regions.

In future work we plan to implement and evaluate the readability of multiple tint bands arranged concentrically for representing multivariate data. We also plan to implement spheroidal Voronoi tessellations as an alternative global data binning method, making it available to a user as a complement to a QTM.

References

Andrienko, G., N. Andrienko, C. Hurter, S. Rinzivillo & S. Wrobel. 2011. From movement tracks through events to places: extracting and characterizing significant places from mobility data. In *IEEE Symposium on Visual Analytics Science and Technology (VAST 2011)*, 161-170. Providence, RI.

Andrienko, G., N. Andrienko, P. Jankowski, D. Keim, M.-J. Kraak, A. M. MacEachren & S. Wrobel (2007) Geovisual analytics for spatial decision support: setting the research agenda. *International Journal of Geographical Information Science,* 21**,** 839-857.

Dutton, G. 1989. Modelling locational uncertainty via hierarchical tessellation. In *Accuracy of Spatial Databases,* eds. M. Goodchild & S. Gopal, 125-140. Londong: Taylor & Francis.

Kraak, M.-J. 2008. Geovisualization and Time – New Opportunities for the Space–Time Cube. In *Geographic Visualization*, 293-306. John Wiley & Sons, Ltd.

Magnusson, M. S. (2000) Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments, & Computers,* 32**,** 93-110.

Peuquet , D. J., A. C. Robinson, S. Stehle, F. A. Hardisty & W. Luo (2015) A method for discovery and analysis of temporal patterns in complex event data. *International Jounal of Geographical Information Science,* 29**,** 1588-1611.

Robinson, A. C., D. J. Peuquet, S. Pezanowski, F. A. Hardisty & B. Swedberg (2016) Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data. *Cartography and Geographic Information Science***,** 1-13.