LSTB/Spanner Revisited

March 6, 2017

Rationale

My proposed project would revisit previous work on a technique for creating low stretch trees with layout free edge bundling. I have some formal experience researching network data/graph visualization for a financial analytics company in Burnaby, BC, and have been generally interested in the topic for the past few years. I have read and reviewed a good bit of published research on network visualization. This document will briefly outline existing research, the original LSTB paper submissions and their documented shortcomings, and finally, a work package to revise and resubmit LSTB for publication.

Background

Graph visualization is a complex research field, with diverse recommendations and outcomes among different types of data. Lee, et al. (2006) detail the visual attributes of a graph in their task taxonomy paper, listing Nodes, Links, Paths, Graphs, Clusters, and Groups as the key objects and components. The authors also adapt a list of low-level tasks from Amar, et al. (2005), including Retrieve Values, Filter, Compute and Derive Value, Find Extremum, Sort, Determine Range, Characterize Distribution, Find Anomalies, Cluster, Correlate, Find Adjacent Nodes, Scan, and Set Operation. Their paper puts forth the argument that all complex, high-level tasks are comprised simply by a combination of these low-level tasks. Furthermore, viewer engagement with graphs is highly dependent on the characteristics of the dataset, in combination with the chosen task(s). For instance, viewers may simply compute and derive values and/or find extremum in a small, sparse dataset; however, a large and dense dataset can be better suited for clustering, scanning, or filtering. In the proposed project, we will focus on visualizations and tasks for medium to very-large data sets, and in particular, data sets that are prone to visual clutter and occlusion when represented in a graph layout.

LSTB originated as a Computer Science Ph.D. dissertation at the University of British Columbia. In her paper, McKnight (2015) cites two predominant techniques for reducing visual clutter in the representation of graphs: adjusting vertices and edge bundling. One contribution from this research was the operationalization of “layout-free” versus “layout-based” edge bundling practices, where geometric vertex positions either are, or are not used in various bundling algorithms. McKnight then introduced a novel algorithmic system, which outputs explicit bundles independently of any pre-existing hierarchy specifications. This eliminated the need
for a layout-based bundling procedure, without rendering it impossible. The system introduces the use of low-stretch spanning trees for edge routing during bundling. In a two-step process, once these novel bundles are determined, any vertex layout method could be used to draw the graph. Additionally, McKnight’s visual encoding implementation allowed viewers to interact with the bundles by repositioning vertices on the fly, as well as by hovering over and highlighting individual edges or bundles of edges. The ability to render graph data in a low-stretch tree structure was described as the major visual advantage/unique impact of this technique.

McKnight’s dissertation was eventually revised and submitted to Graph Drawing, and then to Pacific Vis. Despite the contributions conveyed by the authors, including a novel algorithmic technique, visualization framework, and interactive features, both submissions were rejected. The overwhelming concern expressed by reviewers was a lack of motivation and clear use cases for the LSTB system methodology. The current project aims to deliver a stronger motivation, by identifying and elucidating which kinds of data are best suited for the LSTB system and visual encoding framework, to subsequently execute novel visualizations, and then test viewer interaction with them.

Proposed Work

Large-scale graphs often suffer from visual clutter problems, particularly as occlusion inhibits conveying the complexity of their internal structure. Research from Adai, et al. (2004) showed that a minimal spanning tree of a network can be useful to guide its layout and produce vertex coordinates during visualization. This work was primarily focused towards massive biological data, such as protein-protein interactions and gene-function. In many cases, when the inherent geometric layout of a graph is bad, the resulting visualization is so densely cluttered that no task from Amar et al. (2005) can be successfully executed. In short, a graph display with excessive layers of occlusion is not useful. The proposed LSTB extension will focus on untangling this type of clutter, in order to facilitate success in a high-level path-tracing task.

Data sets to be evaluated:

- Vast Challenge 2013 Computer Network Data:

- Vast Challenge 2012 Bankworld Data:

Projected Milestones and Schedule

1. March 7th, 2017: Finalize proposal and project plan.
2. March 10th, 2017: Contact previous authors and discuss implementation/code/etc.
3. March 15th, 2017: Select final data sets (at least 3) for evaluation.
5. March 25th, 2017: Begin to finalize viewer tasks and design for behavioral pilot study (path tracing in old systems vs. LSTB system visualization).
6. April 1st, 2017: Begin running experiment, collecting data.
7. April 11th, 2017: Data analysis and begin write up.
References


Other Notable Resources...

