Bahador Saket and Jackson Toeniskoetter
CS 544
Prof. Carlos Scheidegger
Final Project Progress Report
April 21, 2015

Datasets: We have available to us every data set found at http://gmap.cs.arizona.edu/datasets. This includes the co-authorship network among International Symposiums on Graph Drawing authors from 1994-2015, a network of 3204 books constructed using Amazon's "customer's who bought this also bought" information, and many more.

Each dataset is processed at http://gmap.cs.arizona.edu into an image file, which contains labeled elements corresponding to the nodes and links in the data. For our project, we choose to process each dataset as an SVG file, so that it can be conveniently manipulated using JavaScript and d3. Since the data is presented in a clean format, we currently perform no actual cleaning of the data.

Progress: We can parse the svg file containing each dataset into a graph data structure. We support selecting individual and groups of nodes and edges. Zooming and panning have been implemented using the JavaScript library SVG-Pan-Zoom. Brushing has been implemented using d3.

Current Work: We are implementing the selection of subgraphs, which will allow the user to make diverse comparisons between two subgraphs using a variety of graph metrics. The final goal is to produce a visualization tool which will support interactive exploratory analysis of the maps produced by GMap, and how well the map corresponds to the underlying network.

The data is initially presented using the map generated at http://gmap.cs.arizona.edu. Filtering will be implemented by allowing the user to select which nodes and links belong to each subgraph using a combination of brushing and mouse clicking. Users will be able to select which features of each subgraph they wish to compare, and will be able to extract information about each node in a subgraph and display it by hovering over the node. Features will include graph metrics such as size, density, as well as highlighting of paths between nodes or subgraphs.