Systematic Yet Flexible Discovery of Social Networks

Adam Perer, Ben Shneiderman

Department of Computer Science

Contact: {adamp, ben}@cs.umd.edu

hcil.umd.edu/socialaction

The increasing availability of digitized information encourages users to conduct more frequent and complex exploratory data analyses. The basic string search or SQL query are no longer adequate for advanced users who seek to understand patterns, discern relationships, identify outliers, and discover gaps.



Figure 1 – The SYF infrastructure integrated into *SocialAction*, a social network analysis tool.

Data mining strategies, cluster analysis, and search engine results are helpful tools for such exploration, which typically takes days, weeks, or months. Advanced users may be trying to sift through gigabytes of genomic data to understand the causes of inherited disease, to filter legal cases in search of all relevant precedents, or to discover behavioral patterns in social networks with billions of people. For these challenging tasks, users must conduct repeated searches, combine results, and consult with colleagues. As they grow familiar with the data, they move from divergent conjectures to more careful hypothesis testing so as to collect evidence supporting their emerging insights.

Current tools can produce useful nuggets of information, but advanced users are increasingly aware of the need to shift from opportunistic discoveries to more systematic approaches. A *systematic* approach guarantees that all measures, dimensions and features of a data set are studied. Such an approach guides novices, ensures analysts of completeness, and facilitates cooperation during analyses that may take weeks or months. However, a wholly strict guide would undermine the *flexible* needs of an analyst, as they will inevitably wish to directly pursue insights based on past successes, new information, fresh hypotheses, or unproductive directions.

## Social Network Analysis

We demonstrate the benefits of SYF by integrating it into a tool for social network analysis, *SocialAction*. Sociologists, intelligence analysts, communication theorists, bibliometricians, food-web ecologists, criminologists and numerous other professionals are interested in understanding social networks. Network analysts focus on relationships instead of just their individual elements; how the elements are put together is just as important as the elements themselves. In many previous studies, sociologists focused largely on behavioral attributes and neglected the social facets of behavior (how individuals interact and the influence they have on each other). Using newer techniques employed by the social network community, analysts can now find patterns in the structure, witness the flow of resources through a network, and learn how individuals are influenced by their surroundings.

# Accompanying Video

[www.cs.umd.edu/hcil/pubs/video2007.shtml](http://www.cs.umd.edu/hcil/pubs/video2007.shtml)

# PAPERS

1. Adam Perer, Ben Shneiderman. Systematic Yet Flexible Discovery: Guiding Exploratory Data Analysis.  *HCIL Tech Report HCIL-2007-9*. Under submission. (2007)*.*
2. Adam Perer, Ben Shneiderman. Balancing Systematic and Flexible Exploration of Social Networks. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 12(5), 693-700. (2006).