Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis

Adam Perer, Ben Shneiderman
Human-Computer Interaction Lab &
Department of Computer Science
University of Maryland
College Park, MD 20742
[adamp,ben]@cs.umd.edu

ABSTRACT
Although both statistical methods and visualizations have been used by network analysts, exploratory data analysis remains a challenge. We propose that a tight integration of these technologies in an interactive exploratory tool could dramatically speed insight development. To test the power of this integrated approach, we created a novel social network analysis tool, *SocialAction*, and conducted four long-term case studies with domain experts, each working on unique data sets with unique problems. The structured replicated case studies show that the integrated approach in *SocialAction* led to significant discoveries by a political analyst, a bibliometrician, a healthcare consultant, and a counter-terrorism researcher. Our contributions demonstrate that the tight integration of statistics and visualizations improves exploratory data analysis, and that our evaluation methodology for long-term case studies captures the research strategies of data analysts.

Author Keywords
Information visualization, statistics, social networks, evaluation, case studies, exploratory data analysis

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
The growing popularity of social network analysis (SNA) can be seen in best-selling books such as Malcolm Gladwell's, "The Tipping Point", Albert-László Barabási's, "Linked", and Duncan Watt's "Six Degrees." Our research focuses on how diverse users conduct social network analysis and how designers can better support their needs.

Social network analysis is inherently complex since analysts must understand every node's attributes as well as relationships between nodes. There are many statistical algorithms which reveal nodes that occupy key social positions and form cohesive social groups [32]. However, it is difficult to find outliers and patterns in strictly quantitative output. In these situations, information visualizations can enable users to make sense of their data, but the visualizations are often hard to interpret because of overlapping nodes and tangled edges.

Our strategy in designing *SocialAction* is to integrate both statistics and visualizations to enable users to quickly derive the benefits of both. Statistics are used to detect important individuals, relationships, and clusters. Instead of tabular display of numbers, the results are integrated with a network visualization in which users can easily and dynamically filter nodes and edges. The visualizations simplify the statistical results, facilitating sensemaking and discovery of features such as distributions, patterns, trends, gaps and outliers. The statistics simplify the comprehension of a sometimes chaotic visualization, allowing users to focus on statistically significant nodes and edges.

Evaluating systems for information visualization tools is problematic because controlled studies may not effectively represent research strategies. Information visualization can differ from other fields of HCI since systems are designed to be exploratory: the set of tasks users may want to perform is not known. For these reasons, we designed a novel methodology to evaluate *SocialAction* with four case studies involving researchers who worked on their own data with their own problems.

In summary, our contributions are:

1. To demonstrate that tightly integrating statistics and visualization improves exploratory data analysis, enabling users to generate significant discoveries.
2. To show that long-term case studies provide an in-depth understanding of how researchers conduct exploratory data analysis. Our four structured and replicated case studies (political analyst,

bibliometrician, healthcare consultant, and counter-terrorism researcher) confirmed the value of integrating statistics with visualization, while providing guidance for redesign of our tool.

RELATED WORK

Social network visualizations have been used to aid SNA since its inception [9], because they provide a natural way to communicate connectivity and promote fast pattern recognition by humans. However, there are challenges when visualizing networks [8, 13]. There are many layout algorithms that place nodes and links to minimize link crossings and adhere to aesthetic principles. These algorithms fall short, however, when the number of nodes is larger than several hundred and the large number of overlapping links makes it hard to judge connectivity [27].

There are a number of software tools designed to help analysts understand social networks. Tools such as KrackPlot [16], Pajek [7], UCINET [3], and visone [4] focus their efforts on statistical analysis and feature limited interaction in their visualizations. Other systems, such as NetDraw [2] and Tom Sawyer [1] focus their efforts on visualization, but lack many statistical algorithms of importance to analysts. These tools feature an impressive number of analysis techniques which users can perform on networks. However, they often are a medley of statistical methods and overwhelming visual output that leaves many analysts uncertain about how to explore their networks in an orderly manner. SNA is a deductive task, and a user's exploratory process can be distracted by having to navigate between separate statistical and visualization packages.

We provide a thorough review of projects focusing on improving interactive exploration with social networks in [17]. More recent work includes Greenland, which augments a node-link diagram with a MDS scatterplot of statistical graph signatures [33]. NodeTrix uses a hybrid approach of node-link diagrams, which show the structure of a network, and adjacency matrices, which highlight communities [12].

There have been many studies evaluating information visualization systems using controlled experiments [6]. However, Plaisant has recently initiated a challenge to information visualization researchers to rethink their evaluation strategies and choose approaches that consider the nature of exploratory tasks [20]. In this spirit, Shneiderman and Plaisant propose Multi-dimensional In-depth Long-term Case studies (MILCs) to study the tasks of information visualization system users [26]. Their methodology suggests working closely with expert users and performing in-depth observations to capture users' creative activities during exploration.

The novelty of MILCs is apparent from a review of the 132 papers in the 2005-2007 IEEE Information Visualization and the 2006-2007 Visual Analytics Science & Technology Conferences. Only 39 papers had any user evaluation and each tested users for less than 2 hours of tool usage. Furthermore, all but 9 of these tests used domain novices who were given standard tasks.

Saraiya et al. identified characteristics of insight, arguably the primary purpose of visualization tools. By pairing tools with experts and measuring the number of insights reached, they empirically evaluated five visualization tools [23]. However, these evaluations did not capture long-term insights as the evaluation sessions lasted only a few hours. Saraiya et al. followed up this work by performing long-term case studies with experts to address two key characteristics missing from their previous approach: motivation and significance [24]. Their work provides insight into the practices of actual data analysts which have implications for both design and evaluation of information visualization systems. The InfoVis contest also allows long-term analysis but the evaluation is informal [21]. The VAST contest improves upon this by making ground truth available [10].

SOCIAL NETWORK ANALYSIS & *SOCIALACTION*

Our interviews with social network analysts, both in academia and industry, suggest that statistical analysis is the most commonly used technique when attempting to interpret social networks. Although visualizations are common in their research publications and reports, they are typically created after the analysis is complete for communicative purposes. However, the most effective visualizations are those that are meticulously hand-crafted.

These exploratory practices might seem surprising, as there is evidence that humans are better at analyzing complex data with images rather than with numbers [5]. Social network data is extremely complex, as the dimensionality of the data increases with each relationship. However, those familiar with network visualizations might sympathize with these statistically attuned practitioners. Network visualizations are typically a tangled set of nodes and edges, and rarely achieve "NetViz Nirvana" (the ability to see each node and follow its edges to all other nodes). Network visualizations may offer evidence of clusters and outliers, but in general it is hard to gather deeper insights from static visualizations.

Our first argument is that it is hard to find patterns and trends using purely statistical methods. Our second argument is that network visualizations usually offer little utility beyond a small set of insights. So what should a social network researcher do? Use both – in a tightly integrated way.

Integrating Statistics with Visualization

The Visual Information Seeking Mantra ("Overview first, zoom and filter, then details on demand") [25] serves as guidance for organizing the complex tasks of a social network analyst. At the first step, analysts begin with an overview of the network both statistically and visually (Figure 1a). Measurements of the entire network, such as

density, diameter and number of components, are presented alongside a force-directed layout of the network. The visualization gives users a sense of the structure, clusters and depth of a network, while the statistics provide a way to both confirm and quantify the visual findings. If the network is small, or the analysts are interested purely in the topology of the network, this step may be enough.

A more capable analyst will wish to gain a deeper understanding of the individual elements of the network. Users can use statistical importance metrics common in social network analysis to measure the nodes and edges. For instance, an analyst can rank the nodes by degree (the most connected nodes), betweenness (the gatekeepers), closeness (well-positioned nodes to receive information) or other metrics. After users select a metric, a table lists the nodes in rank order. *SocialAction* assigns each node a color, ranging from green (low ranking) to black (average ranking) to red (high ranking). This helps illustrate each node's position among all ranked entities. The network visualization is updated simultaneously, as well, and paints each node with the corresponding color. Users now can scan the entire network to see where the important nodes reside (Figure 1a).

To gain further insights, *SocialAction* allows users to continue on to step 2 of the Visual Information Seeking Mantra("filter and zoom"), where most other social network analysis packages strand users. Panning and zooming naively is not enough to empower users. Zooming into sections of the network force users to lose the global structure, and dense networks may never untangle. *SocialAction* allows user-controlled statistics to drive the navigation. Users can dismiss portions of the network that do not meet their criteria by using range sliders. Filtering by attributes or importance metrics allows users to focus on the types of nodes they care about – while simultaneously simplifying the visualization (Figure 1b).

After analysts make sense of global trends through statistical measurements and visual presentations, their analyses often are incomplete without understanding what the individual nodes represent. Contrary to most other network visualizations, labels in *SocialAction* are always present. The controls for font size and length allow the analyst to decide their emphasis. In line with step 3 of the Visual Information Seeking Mantra's "Details on Demand", users can select a node to see all of its attributes. Hovering over a node also highlights each node's edges and neighbors, achieving "NetViz Nirvana" for the node of interest (Figure 1c).

These examples are just a few of *SocialAction*'s many statistical and visual features. *SocialAction* also offers other sophisticated ways of analyzing networks, such as:

- ranking edges (to identify powerful versus weak relationships)

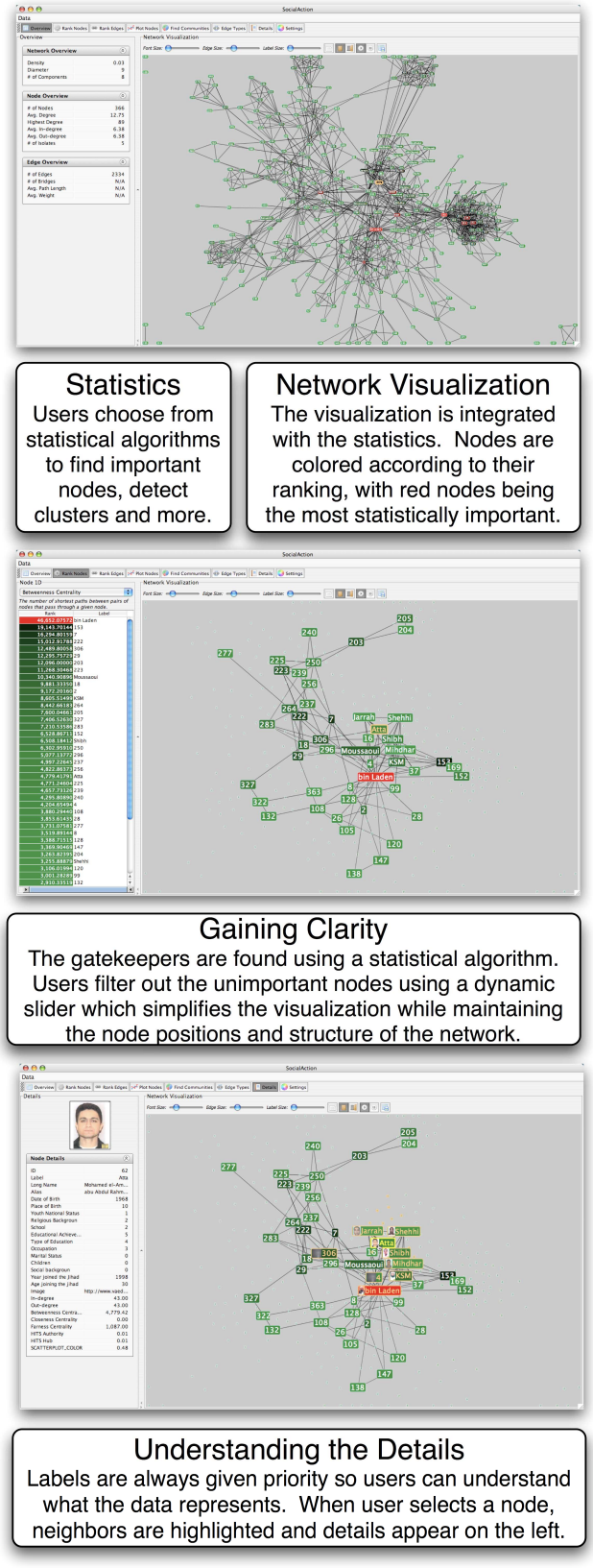

Figure 1. Exploring a social network in *SocialAction*. This figure features the "Global Jihad" terrorist network from our 4th case study. In order to protect sensitive information, node labels have been anonymized except for those individuals publicly identified in the Zacarias Moussaoui trial.

- plotting nodes into more comprehensible scatterplots (to find patterns and outliers)
- enabling clustering algorithms (to find well-connected communities)
- indicating multiplexity of social ties (to analyze different edge types, such as friends versus enemies).

Due to limited space, we save the description of these features until the Case Studies section in which we demonstrate their utility in practice.

SocialAction 2.0 is a significant rewrite of *SocialAction* 1.0, described in [17]. The statistical and visualization algorithms have been optimized to support real-time interaction with large networks of interest to our case study partners (10,000-100,000 nodes). The system is implemented in Java and uses the *Prefuse* [11] toolkit for the visualizations.

In summary, bringing together statistics and visualization is an elegant solution for exploratory data analysis. The visualizations simplify the statistical results, improving the comprehension of patterns and global trends. The statistics, in turn, simplify the comprehension of a sometimes chaotic visualization, allowing users to focus on statistically significant nodes and edges.

EVALUATION METHODOLOGY

Traditional laboratory-based controlled experiments have proven to be effective in many user interface research projects. When new widgets, displays, interaction methods, or input devices are being developed, controlled experiments can compare two or more treatments by measuring learning times, task performance times, or error rates. Typical experiments would have 20-60 participants, who are given 10-30 minutes of training, followed by all participants doing the same 2-20 tasks during a 1-3 hour session. Statistical methods such as t-tests and ANOVA are applied to show significant differences in mean values. These summary statistics are effective, especially if there is small variance across users.

However, because domain experts work for days and weeks to carry out exploratory data analysis on substantial problems, their work processes are nearly impossible to reconstruct in a laboratory-based controlled experiment, even if large numbers of professionals could be obtained for the requisite time periods. A second difficulty is that exploratory tasks are poorly defined, so telling the users which tasks to carry out is incompatible with discovery. Third, each user has unique skills and experience, leading to wide variations in performance which undermine the utility of summary statistics. In controlled studies, exceptional performance is seen as an unfortunate outlier, but in case studies, these special events are fruitful critical incidents that provide insight into how discovery happens. Fourth, we wanted more than quantitative analyses of the tool. We also wished to hear about the problems and frustrations users encountered as well as their thrilling tales of success [15]. For such reasons, we turned to structured and replicated case study research methods to collect supporting evidence for our conjecture that integrating statistics with visualization would facilitate discovery for social network analysts.

Inspired by the goals of MILCs [26], we developed a methodology for studying the effectiveness of *SocialAction*:

1. Interview (1 hour): This initial phase involves an interview to understand the intentions of the participant. The achievement of the intentions acts as one benchmark of success at the end of the study. Furthermore, this phase acts as an opportunity for observers to decide if participants are appropriate candidates for the study. This evaluation was limited to knowledgeable domain experts conducting serious research with well-defined goals.
2. Training (2 hours): Users participate in a training session with the software developers. The participants are expected to use *SocialAction* to find insights during this practice analysis session. After the training session, users have access to a brief instruction manual.
3. Early use (2-4 weeks): Participants install *SocialAction* in their workplace where they load their own data relevant to their research goals. Each week, observers visit the participants' workplaces to interview them regarding their progress. For case studies involving remote locations, interviews occur over the phone. In the tradition of action research [31], the developers try to accommodate participant needs by modifying and adding features to the software to meet critical needs.
4. Mature use (2-4 weeks): This phase features more hands-off, "ethnographic"-style observation. No further improvements are made to the software despite requests from participants. Similar to phase 3, researchers visit each participant's workplace or conduct phone interviews. The software developers continue to provide technical support as needed.
5. Outcome (1 hour): This exit interview provides participants a formal chance to explain how the software impacted their research. The participants revisit their original intentions from Phase 1 and rate each intention based on the level of achievement.

CASE STUDIES

In order to validate our claims, we conducted four case studies of users with diverse skill sets, domains of knowledge, and social network expertise. The participants were not recruited, but instead sought out *SocialAction* on their own after facing challenges in making sense of social networks. The descriptions of the case studies below only discuss a fraction of the participants' insights but are representative of their overall experience.

Case Study 1: Senatorial Voting Patterns

Congressional analysts are interested in partisan unity in the United States Senate. For instance, *Congressional Quarterly* calculates such unity by identifying every vote in which a majority of Democrats voted opposite a majority of Republicans, and then counts, for each senator, the percentage of those votes where they voted with the party. This metric can be useful for tracking an individual senator's party loyalty from year to year, but it does not tell much about the overall patterns in the body. Chris Wilson, then an associate editor for the *US News & World Report,* was interested in voting patterns among United States senators.

Wilson was seeking to uncover senatorial patterns, such as strategic, bipartisan, or geographic alliances in the data set. Wilson spent significant effort mining voting data from public databases, but was unable to find such distinct patterns through his normal methods of analysis.

Wilson believed social network analysis could yield the answers he sought. His data included voting results for each senator during the first six months of 2007, beginning when the Democratic Party assumed control of the chamber with a one-seat majority. A social network can be inferred from co-occurrences of votes. Before contacting us, Wilson tried to visualize this data in KrackPlot [16], ManyEyes [30] and NetDraw [2] but did not manage to find any interesting patterns.

Early Use

From the data, Wilson constructed the network such that when a senator votes with another senator on a resolution, an edge connects them. The strength of each edge is based on how often they vote with each other (e.g., Barack Obama and Hillary Clinton voted together 203 times, whereas Obama and Sam Brownback voted together only 59 times). This leads to a very dense network because there are certain uncontroversial resolutions that all senators vote for (e.g. Resolution RC-20, a bill commending the actions of "the Subway Hero" Wesley Autrey). All senators are connected, which leads to a visualization of a huge, tangled web. *SocialAction*'s interactive statistics empower users to dig deeper, without forcing users to choose an arbitrary cut-off before analysis begins.

SocialAction allows users to rank edges according to importance metrics. Wilson used this feature to compare network visualizations by dynamically filtering out relationships with low importance rankings. For instance, the 180-vote threshold (about 60 percent voting coincidence) is shown in Figure 2a. Partisanship is strong even at this fairly low threshold, and the Republican senators who are most likely to vote with Democrats (Collins, Snowe, Specter, and Smith) are evident. This suggests that, in this particular Senate, although both parties are partisan, Republicans are less so than Democrats.

As the threshold increases, the bipartisan edges diminish (Figure 2b). Another unexpected consequence was that the Democrats stay more tightly unified than the Republicans as the threshold increases. Wilson believed this interaction beautifully illustrated the Democratic caucus's success in keeping members in line, an important fact when reviewing legislative tactics. The integration of statistics and visualization made this discovery possible.

Mature Use

In order to determine patterns of individual politicians, Wilson used the statistical importance metrics of *SocialAction*. The capability to rank all nodes, visualize the outcome of the ranking, and filter out the unimportant nodes led to many discoveries. Wilson stated, for instance, that the *betweenness centrality* statistic turned out to be "a wonderful way to quantitatively measure the centers of gravity in the Senate". *SocialAction* made it evident that only a few senators centrally link their colleagues to one another. Wilson was also able to use the interactive clustering algorithms of *SocialAction* to "uncover geographic alliances among Democrats". These findings are just a sample of the sorts of insights that eluded Wilson prior to his analysis with *SocialAction*.

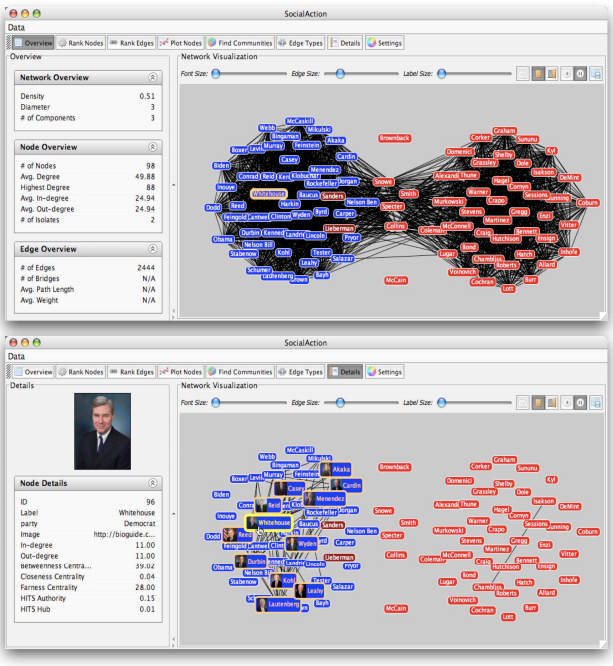

Figure 2. The social network of the U.S. Senators voting patterns. Republicans are colored red, Democrats blue and Independents maroon. In the top image (a), the partisanship of the parties appeared automatically (180 vote threshold). In the bottom image (b), the threshold is raised to 290 votes. The Democrats' relationships are much more intact than the Republicans. Details-on-demand are provided for Senator Whitehouse, the senator with the highest degree at this threshold.

Outcome

Wilson was thrilled with the discoveries that *SocialAction* helped reveal. The tight integration of statistics and visualization allowed him to uncover findings and communicate them to his peers both at his publication and on Capitol Hill. *SocialAction* received so much attention internally that the magazine hopes to replicate some of its functionality for its online readers. This will provide readers with further data analysis opportunities, in the spirit of [29]. Since the case study, Wilson has moved to Slate Magazine but still uses *SocialAction* for investigative reporting. So far, analysis from *SocialAction* has led to an interactive feature analyzing the social networks of steroids users in Major League Baseball [19], with more stories planned for the future.

Case study 2: Knowledge discovery for medical research

The National Library of Medicine (NLM) maintains PubMed, a search engine with access to more than 17 million citations in the health sciences. A recently revised feature of PubMed is the related article search. This feature aims to improve knowledge discovery by linking together critical information that may be missed by keyword searching. When users reach a citation of interest, five related articles are suggested on the screen. Sophisticated information retrieval algorithms generate these recommendations automatically. Jimmy Lin, a Ph.D. expert in information retrieval, led the project at NLM.

Lin and his colleagues sought to understand the usefulness of the recommendation algorithm. A successful algorithm would allow users to browse the document collection using the related articles links and reach other relevant documents. A network of documents can be created by linking together each document with its recommendations from the algorithm. The network's structure is important, since isolated documents without links from other relevant documents cannot be reached by browsing. Lin hoped to gain deeper insights about the usefulness of the algorithms by using SNA to explore the recommendation network.

Early Use

For the experimentation with *SocialAction*, Lin used data from a TREC genomics test set [14]. This set was chosen because there was ground truth on the relevance of documents (such as results for the query "what is the role of the gene GSTM1 in the disease Breast Cancer"). Lin then generated document networks, where for each known relevant document, the top five related documents were linked (e.g., the suggestions from the related article search in PubMed). Upon loading the network for the first time in *SocialAction*, a eureka moment occurred. Lin proclaimed, "This figure is exactly what I wanted to see!"

Two phenomena were immediately noticeable from the visualization. First, relevant documents tend to cluster around each other (notice the dense red cluster in the middle of the network in Figure 3). This supports the cluster hypothesis in information retrieval, which proposes that relevant documents tend to be more similar to each other than to non-relevant documents [28]. However, there were also a number of isolated islands of documents (notice the disconnected, star-shaped clusters in Figure 3). These represent documents that would be unreachable by users when using the related article feature, undermining the goals of that feature.

Lin used a variety of the exploratory features of *SocialAction*. For instance, he used the importance rankings for nodes to find the most suggested articles, or the gatekeeper articles that bridge two clusters together. However, Lin's initial goal was to characterize the effects of the related article search, as opposed to refining the algorithm. Thus, Lin focused mostly on overall network statistics (such as number of disconnected components, density, and diameter) to quantify the output of the retrieval algorithm. Figuring out which statistics are useful is often an under-surveyed problem of analysis tools. *SocialAction*'s design, which supports users quickly iterating through measurements while maintaining a constant visualization, served a useful role in this exploration.

Lin also requested additional features for *SocialAction*, such as the capability to calculate statistics for nodes with certain attributes (e.g., the number of relevant documents linked from each relevant document). Since Lin also was interested in using the statistical information to inform his retrieval algorithm, an exporter for the statistics was built.

Mature Use

With the requested features implemented, Lin used *SocialAction* to study 49 different query networks. Each of

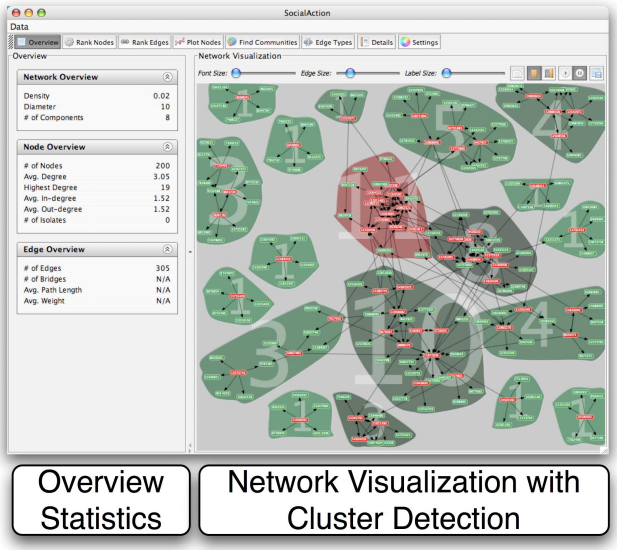

Figure 3. The recommendation network of a query on PubMed documents. Relevant documents are red, non-relevant are green. The community algorithm highlights closely-connected clusters in the network. Communities are color-coded by the percentage of relevant documents and labeled by the number of relevant documents.

the networks had varying properties (number of suggested articles, number of relevant documents, density). The integration of statistics and visualization allowed Lin to quickly explore the networks, spending less than a few minutes on each network after becoming comfortable with *SocialAction*. This exploratory investigation led to the visual insight that networks with more relevant documents (red nodes) clustered together tend to have fewer the disconnected components.

Lin also used the clustering features of *SocialAction* to find tight-knit groups of articles that are highly similar to each other. Figure 3 shows the network components broken down into smaller communities using the hierarchical clustering algorithms available in *SocialAction* [17]. Each community is surrounded by a bubble colored based upon statistical information chosen by users (in this case, the average number of relevant documents). This visual evidence supports the cluster hypothesis Lin sought to confirm. *SocialAction* allows users to control the size of the clusters, digging deeper and deeper into the closest-knit groups. However, while this feature allowed Lin to advance his exploration, he chose to leave these results out of his analysis due to the subjective nature of cluster size.

Outcome

Using *SocialAction*, Lin and his colleagues were able to better understand the performance of their retrieval algorithm. The analysis showed that users can access most of the relevant documents by clicking on the related article links (e.g., without having to go back to the search results and reformulate a query). However, they also identified isolated clusters, which represented relevant documents that were not reachable by browsing. The results of this analysis led to a submission of a high-quality research article. The exploratory nature of *SocialAction* allowed the researchers to measure their algorithms even though they had no prior knowledge of which SNA statistics would be useful. They also believe *SocialAction* will be a useful tool for verifying the effectiveness of new recommendation algorithms for PubMed.

Case Study 3: Engaging Hospital Trustee Networks

A Northeastern healthcare insurer is interested in engaging hospital boards in their region to speak loudly about healthcare quality. They are using social network analysis to help inform and prioritize this initiative.

They hired Bruce Hoppe, a professor at Boston University, who also serves as a consultant aiding businesses in optimizing their operational networks. He uses social network analysis to accelerate business results and has experience with many Fortune 500 companies. Despite having a repertoire with over 8 social network analysis software tools (including [2, 3, 4, 16]), he has yet to find a suite that achieves his needs in exploring data effectively. For this reason, he was interested in integrating *SocialAction* into the workflow of his latest project.

Early Use

Hoppe began using *SocialAction* to analyze the board interlock network of over 500 organizations (such as hospitals, businesses, and non-profits) provided by the healthcare insurer. These 500 organizations had a total of almost 8,000 board members. Hoppe was pleased that *SocialAction* could load all of the data at once and provide an overview of the whole network. In general, he was used to cropping data before analysis.

After seeing the large network, the healthcare insurer asked Hoppe to focus on a subset of the network: the hospitals and their boards of trustees. Unlike other SNA tools, *SocialAction* allows users to compare different but related varieties of statistical measures on a scatterplot. When Hoppe noticed this feature, he became interested in the relationship of *degree centrality* and *betweenness centrality*: to what extent were trustees sitting on many boards also the gatekeepers who connected many diverse hospitals. The scatterplot enabled Hoppe to quickly spot patterns in the healthcare network and the important outliers (Figure 4). In particular, a relatively unknown "Trustee 527" (anonymized for confidentiality) emerged as a focus of attention due to her unique position of few hospital connections but nonetheless a key gatekeeper in the network. The integration of statistics and visualization provided Hoppe with inspiration for a report delivered to the healthcare insurer.

Mature Use

The healthcare insurer was informed by the report which Hoppe provided after his analysis with *SocialAction*. The network was filtered according to meaningful SNA metrics and had become more comprehensible. Now present were

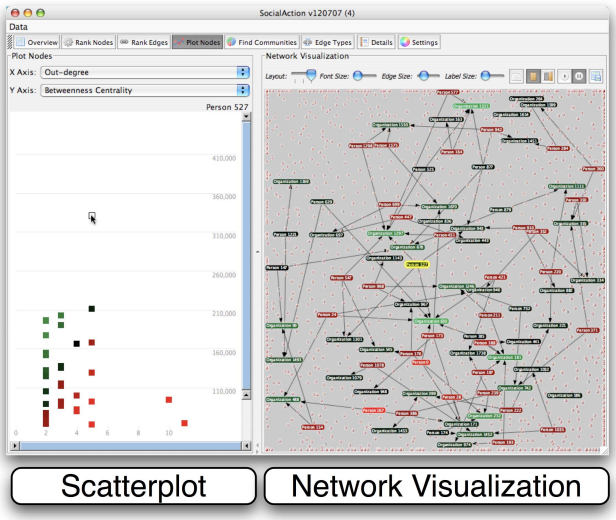

Figure 4. The scatterplot allows users to compare multiple SNA metrics (In this example, out-degree and betweenness). An obvious outlier is selected in the scatterplot, who was a trustee that had the highest betweenness (key gatekeeper) despite having only 4 connections. The names of trustees and organizations are anonymized to protect this confidentiality.

labels which allowed viewers to focus on the connections between hospitals and trustees. This transparency of the underlying data led the healthcare insurer to question the data. In fact, they found gaps in their data. These data discrepancies are being corrected, so Hoppe was forced to temporarily halt his exploration. However, the visualizations and filtering power of *SocialAction* allowed him to interpret these critically important data issues during analysis. Hoppe suspects a purely statistical approach to analysis might have missed these details.

Outcome

Hoppe used *SocialAction* as his main exploratory tool during his consulting work for the Northeastern healthcare insurer. In his monthly reports, he often included insights and visualizations resulting from his use of *SocialAction*. These findings have had significant impact with the healthcare organization. They now have a better understanding of the region's hospital trustee network and are working to make sure it informs their quality initiative. However, Hoppe admitted "I like having a medley of complex and ad-hoc tools. I am much more likely to recommend SocialAction to my clients – who need one simple approach to network exploration – than I am to adopt it as my own primary SNA tool." SocialAction lacked certain features critical to his needs, such as additional statistical measures, comprehensive map-editing for nodes (e.g., size, label, and color), the ability to save these edits for future updating, and the ability to export the final results as vector graphics for high resolution presentations.

Case Study 4: Group Dynamics in Terrorist Networks

The National Consortium for the Study of Terrorism and Responses to Terror (START) is a U.S. Department of Homeland Security Center of Excellence. START has a research team around the world which "aims to provide timely guidance on how to disrupt terrorist networks, reduce the incidence of terrorism, and enhance the resilience of U.S. society in the face of the terrorist threat". One member of this team is James Hendrickson, a criminologist Ph.D. candidate, who is interested in analyzing the social networks of "Global Jihad".

Previous research has pointed to the importance of radicalization informing and sustaining terrorist organizations. While the radicalization process has been well described from a psychological standpoint, he believes theories examining the group dynamics of terrorism have largely failed to properly measure the size, scope and other dynamics of group relations. Hendrickson proposes to systematically compare the density and type of relationships held by members of the "Global Jihad" to evaluate their predictive ability in determining involvement in terrorist attacks. Marc Sageman, a visiting fellow at START, assembled a database of over 350 terrorists involved in jihad when researching his best-selling book, "Understanding Terror Networks" [22]. Hendrickson plans to update and formally apply social network analysis to this data as a part of his Ph.D. dissertation.

Early Use

The Sageman database has over 30 variables for each terrorist. Among these variables are different types of relationships, including friends, family members, and educational ties for religion. Hendrickson proposes that the types of relationships connecting two individuals will hugely affect their participation in terrorist attacks. Hendrickson began this analysis using UCINET [3] and was able to analyze some of his hypotheses. However, he believed UCINET did not facilitate exploration and generating new hypotheses easily. Hendrickson initially was skeptical of using visualizations for his analysis. He prefers being able to prove statistical significance quantitatively rather than relying on a human's judgment of an image. The quick access to the statistical counterparts of *SocialAction*'s visualizations eased his concerns.

In particular, *SocialAction*'s multiplexity feature aided Hendrickson's exploration. *SocialAction* allows users to analyze different relationship types without forcing users to load new data sets. The visualization shows the selected relationship edges but keeps node positions stable in order to aid comprehension. The statistical results are also automatically recomputed based on the newly selected

Figure 5. The multiplexity of the "Global Jihad" social network is demonstrated. The upper visualization (a) shows the Friendship network, with bin Laden the most popular individual. The bottom network (b), showing religious ties, offers a much different view of the terrorist organization.

structure. For instance, only the 'Friend' relationships among Jihadists are selected in Figure 5a. (Compare this to the denser Figure 1a, which shows all relationship types.) The nodes here are ranked by degree, so red nodes have the most friends. Jihadists Osama Bin Laden and Mohamed Atta (known for his role in the 9/11 attack) are ranked the highest. However, when the religious ties are invoked, a different set of key jihadists emerge (Figure 5b).

Mature Use

After analyzing the statistical attributes of nodes, Hendrickson became interested in understanding the individuals' attributes. As an example, he was interested in answering questions like, "Does an individual's socioeconomic status or education level impact their position in the terrorist network?" Like statistical rankings in *SocialAction*, users can rank and filter based upon attributes. Hendrickson filtered out individuals without college degrees, without religious backgrounds, or without engineering expertise and analyzed the results. The combination of nodal attributes with statistical filtering and plotting streamlined his accustomed workflow. He suspects he may not have been as free thinking if it wasn't for the ease of exploration in *SocialAction*. This analysis inspired Hendrickson to think of new, not-yet-coded attributes, to test additional hypotheses. He is currently augmenting Sageman's database with new attributes so he can look for patterns in *SocialAction*, visually and statistically.

Outcome

Hendrickson's experience with *SocialAction* has led to new inspiration for his dissertation thesis. Although he had access to the dataset long before the case study began and conducted analysis with other SNA software, the integration of statistics and visualization allowed exploration in new, interesting ways. As a result, the START center is interested in making *SocialAction* the default network analysis tool for internal and external users who wish to access their databases. They are currently integrating a specialized version of *SocialAction* into their online global terrorism database.

DISCUSSION AND FUTURE WORK

The four case studies have provided evidence that exploratory data analysis improves with integrated statistics and visualization. Tools to support the generation of hypotheses are sometimes overlooked. *SocialAction* provides the participants with the freedom to load all of their data to identify global trends. Instead of removing data blindly, users can filter the data according to statistical principles of social network analysis. This provided our participants with a level of comfort lacking with other tools.

In order to integrate statistics and visualization, the emphasis on interactivity is high. The layout algorithm updates in real-time, allowing users to see animated, emergent trends. Furthermore, the statistical algorithms also are optimized for real-time interaction to reduce hesitance from computing additional measures. *SocialAction* only contains a subset of social network analysis algorithms. These algorithms were carefully chosen by analyzing the most common metrics used in social network journals. However, users have the ability to import statistical measurements from other programs if any required measurements are missing.

In addition to providing evidence to support our hypothesis, the case studies also acted as a stimulus for pushing the technology forward. The implementation was not driven for a controlled study, but rather to handle a wide range of use by inquisitive researchers. It forced the implementation to operate on real, large datasets. This results in *SocialAction* being a tool that can be used by professional researchers solving their research problems. The authors are exploring the use of systematic yet flexible guides to help domain expert users through complex exploration of data over days, weeks and months [18]. The developers are using the findings from the case studies to improve *SocialAction* and make it available to the public for use by social network researchers. The authors also plan to conduct additional case studies to document use in other domains.

CONCLUSION

Each of the case study participants relied on *SocialAction*'s integration of statistics and visualization. Chris Wilson, the associate editor of US News & World Report, reached insights by ranking edges, allowing him to distinguish powerful and weak relationships in the partisan U.S. Senate. Jimmy Lin, at the National Library of Medicine, used the overview statistics and visualization to characterize his recommendation algorithms. Bruce Hoppe, a consultant for a major healthcare provider, used the statistical scatterplots to find patterns and outliers that will have impact on the operational network of the company. Finally, James Hendrickson used multiplexity identification to statistically analyze different types of relationships in the "Global Jihad" networks.

Each of these case study participants started without prior knowledge about which features of *SocialAction* would help them achieve their goals. However, the tight integration of statistics and visualization allowed them to rapidly try each feature until patterns were evident. Although the case studies do not produce quantitative measures of the interface, they reveal the novel strategies used by the analysts when statistics and visualization were integrated.

Using *SocialAction* as evidence, we believe a tight integration of statistics and visualization can aid exploratory data analysis beyond networks. Complex data sets, including temporal, hierarchical, geographic, and tabular data, often produce overwhelming visualizations and statistics. Analysis tools that tightly integrate these facets, like *SocialAction*, can aid researchers exploring the frontier of their domains.

ACKNOWLEDGMENTS

We thank Georg Apitz, Bita Azhdam, Ben Bederson, Allison Druin, Jonathan Grudin, Catherine Plaisant, Vibha Sazawal, Dave Wang, Mary Whittaker, our case study participants, and the anonymous reviewers for helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. 0633843.

REFERENCES

1. *Tom Sawyer Visualization*. Tom Sawyer, (2007).
2. Borgatti, S. *Netdraw 2*. Analytic Technologies, (2007).
3. Borgatti, S., Everett, M. G. and Freeman, L. C. *UCINET 6*. Analytic Technologies, (2007).
4. Brandes, U. and Wagner, D. *visone - Analysis and Visualization of Social Networks* In *Graph Drawing Software*, M. Junger and P. Mutzel. Springer-Verlag (2003).
5. Card, S. K., Mackinlay, J. D. and Shneiderman, B. *Readings in Information Visualization: Using Vision To Think*. Morgan-Kaufman (1999).
6. Chen, C. and Czerwinski, M. Empirical Evaluation of Information Visualizations: An Introduction. *Intl. Journal of Human-Computer Studies*, 53(2000), 631-635.
7. de Nooy, W., Mrvar, A. and Batageli, V. *Exploratory Social Network Analysis with Pajek*. Cambridge University Press, Cambridge (2005).
8. Di Battista, G., et al. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, New Jersey (1999).
9. Freeman, L. C. Visualizing Social Networks. *Journal of Social Structure*, 1, 1 (2000).
10. Grinstein, G., et al. The VAST 2006 Contest: A tale of Alderwood. In *Proc. IEEE Symp. on Visual Analytics Science and Technology*(2006), 215-216.
11. Heer, J., Card, S. K. and Landay, J. A. prefuse: A Toolkit for Interactive Information Visualization. In *Proc. ACM Conf. on Human Factors in Computing Systems*(2005).
12. Henry, N., Fekete, J.-D. and McGuffin, M. J. NodeTrix: a Hybrid Visualization of Social Networks. *IEEE Trans. on Visualization and Computer Graphics*, 13, 6 (2007), 1302-1309.
13. Herman, I., Melancon, G. and Marshall, M. S. Graph visualization and navigation in information visualization: A survey. *IEEE Trans. on Visualization and Computer Graphics*, 6, 1 (2000), 23-43.
14. Hersh, W., et al. TREC 2005 genomics track overview. In *Proc. 14th Text Retrieval Conf.*(2005).
15. Hughes, J., et al. The role of ethnography in interactive systems design. *interactions*, 2, 2 (1995), 56-65.
16. Krackhardt, D., Blythe, J. and McGrath, C. KrackPlot 3.0: An Improved Network Drawing Program. *Connections*, 17, 2 (1994), 53-55.
17. Perer, A. and Shneiderman, B. Balancing Systematic and Flexible Exploration of Social Networks. *IEEE Trans. on Visualization and Computer Graphics*, 12, 5 (2006), 693-700.
18. Perer, A. and Shneiderman, B. Systematic Yet Flexible Discovery: Guiding Domain Experts through Exploratory Data Analysis. In *Proc. Intelligent User Interfaces (IUI)*(2008), 109-118.
19. Perer, A. and Wilson, C. The Steroids Social Network: An Interactive Feature on the Mitchell Report. *Slate Magazine*, (2007).
20. Plaisant, C. The challenge of information visualization evaluation. In *Proc. Advanced visual interfaces*. ACM Press (2004), 109-116
21. Plaisant, C., Fekete, J. D. and Grinstein, G. Promoting Insight Based Evaluation of Visualizations: From Contest to Benchmark Repository. *IEEE Trans. on Visualization and Computer Graphics*, 14, 1 (2008), 120-134.
22. Sageman, M. *Understanding Terror Networks*. University of Pennsylvania Press, Philadelphia (2004).
23. Saraiya, P., North, C. and Duca, K. An Evaluation of Microarray Visualization Tools for Biological Insight. In *Proc. IEEE Symp. on Information Visualization*. IEEE Press (2004), 1-8.
24. Saraiya, P., et al. An Insight-based Longitudinal Study of Visual Analytics. *IEEE Trans. on Visualization and Computer Graphics*, 12, 6 (2006), 1511-1522.
25. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. In *Proc. Visual Languages*(1996), 336-343.
26. Shneiderman, B. and Plaisant, C. Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies. In *Proc. Beyond time and errors Workshop (BELIV)*. ACM Press (2006), 1-7.
27. van Ham, F. *Interactive Visualization of Large Graphs*. Technische Universiteit Eindhoven, 2005.
28. van Rijsbergen, C. J. *Information Retrieval*. Butterworths, London (1979).
29. Viegas, F. B. and Wattenberg, M. Communication-Minded Visualization: A Call to Action. *IBM Systems Journal*, 45, 4 (2006).
30. Viégas, F. B., et al. Many Eyes: A Site for Visualization at Internet Scale. In *Proc. IEEE Symp. on Information Visualization*(2007).
31. Warmington, A. Action Research: Its Methods and its Implications. *Journal of Applied Systems Analysis*, 7(1980), 23-29.
32. Wasserman, S. and Faust, K. *Social Network Analysis: Methods and Applications*. Cambridge University Press (1994).
33. Wong, P. C., et al. Graph Signatures for Visual Analytics. *IEEE Trans. on Visualization and Computer Graphics*, 12, 6 (2006), 1399-1413.