Integrating Statistics and Visualization for Exploratory Power:

From Long-Term Case Studies to Design Guidelines

Adam Perer and Ben Shneiderman ▪ *University of Maryland*

Evaluating visual-analytics systems is challenging because laboratory-based controlled experiments might not effectively represent analytical tasks. One such system, SocialAction, integrates statistics and visualization in an interactive exploratory tool for social network analysis. This article describes results from long-term case studies with domain experts and extends established design goals for information visualization.

Visual analytics (VA) is different from other fields of human-computer interaction because VA systems are designed to be exploratory: the set of tasks users want to perform might not be known. Evaluating such systems is problematic because controlled studies might not effectively represent analytical tasks. One such VA system, SocialAction, aims to improve exploratory data analysis of social networks.[1–3] SocialAction integrates both statistics and visualizations to enable analysts to benefit from both types of analysis. Statistics help identify important individuals, relationships, and clusters throughout the network. Network visualizations integrate the statistical results, allowing analysts to dynamically rank and filter the network. Visualizations simplify the statistical results by highlighting patterns, gaps, and outliers. Statistics simplify the typically dense network visualizations, allowing analysts to focus on statistically significant nodes and edges.

In order to understand whether SocialAction's integrated approach helps analysts, we designed an evaluation methodology to evaluate SocialAction with case studies involving researchers who worked on their own data with their own problems. In this article, two case studies are presented that reveal that an integration of visualizations and statistics can lead to valuable insights during analysis. However, there are few information visualization tools that provide an integrated approach to data analysis (for more on this, see the "Why Combine Visualization and Statistics?" sidebar, next page). In order to promote such an integration, we reflect on data from our case studies and present specific guidelines for VA system designers.

The Evaluation Methodology

Our methodology is inspired by the goals of Multidimensional In-Depth Long-Term Case Studies (MILCs).[4] Although many other methodologies for qualitative case studies exist, we believe this is the first long-term, structured, and replicated approach that takes into account the unique demands of information visualization users. (For more on evaluation techniques for information visualizations, see the related sidebar, page 42.) Our methodology consists of the following five phases.[2]

Interview (1 hour)

This initial phase involves an interview to understand the intentions of the domain experts. 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 whether the domain experts are appropriate candidates for the study.

Why Combine Visualization and Statistics?

The visual-information-seeking mantra of "overview first, then details on demand"[1] offers some guidance to designers who strive to support interactive exploration of information visualizations. Similarly, John Tukey's mantra of "exploratory data analysis" has been the basis for successful statistical analysis techniques.[2] Tukey's techniques and those he inspired allow analysts to answer questions such as, "What is a typical value?" "What is the distribution?" "What is the percentile?" and "What are the outliers?"

Given these two widely cited techniques, you might question the need for additional design goals. However, the evidence is mounting that a combination of interactive and statistical techniques is necessary for successful information visualization systems. There are few information visualization tools that integrate this dual-front approach to solving data analysis problems. One such dual-front approach, described by Deborah Swayne and her colleagues, highlights the promise of an integrated system.[3] A more recent example, NodeXL, integrates network visualizations inside Microsoft Excel.[4]

In particular, our four long-term case studies reveal that such a combined approach can produce more valuable insights. However, the lack of specific design guidelines perhaps limited our original paper's usefulness to future researchers, designers, and practitioners. So, we developed the guidelines presented in the main article.

References

1. B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization," *Proc. IEEE Symp. Visual Languages*, IEEE Press, 1996, pp. 336–343.
2. J. Tukey, *Exploratory Data Analysis*, Addison-Wesley, 1977.
3. D.F. Swayne, A. Buja, and D.T. Lang, "Exploratory Visual Analysis of Graphs in GGobi," *Proc. 3rd Int'l Workshop Distributed Statistical Computing* (DSC 03), 2003; www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/SwayneEtAl.pdf.
4. M.A. Smith et al., "Analyzing (Social Media) Networks with NodeXL," to be published in *Proc. 4th Int'l Conf. Communities and Technologies* (C&T 09), 2009.

Training (2 hours)

Users participate in a training session with the software developers. The domain experts 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.

Early Use (2–4 weeks)

Domain experts install SocialAction in their workplace, where they load their own data relevant to their research goals. Each week, observers visit the domain experts' workplaces to interview them regarding their progress. For case studies involving remote locations, interviews occur over the phone. In the tradition of action research, the developers try to accommodate domain experts' needs by modifying and adding features to the software to meet critical needs.

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 domain experts. Similar to the early-use phase, researchers visit each domain expert's workplace or conduct phone interviews. The software developers continue to provide technical support as needed.

Outcome (1 hour)

This exit interview provides domain experts a formal chance to explain how the software impacted their research. The domain experts revisit their original intentions and rate each intention on the basis of the level of achievement.

Applying the Methodology

In our case studies, we limited our evaluations to knowledgeable domain experts conducting serious research with well-defined goals. We interviewed all the domain experts at their workplace except for one who was at a remote location. Notes from each interview were recorded as field notes, and they were later transcribed digitally.

At each interview in the early-use and mature-use phases, the domain experts recalled their insights from that week's efforts using SocialAction. Typically, they shared their notes and screenshots for each insight with the interviewer. They often used SocialAction to demonstrate how they made their discovery. One downside of our evaluation was that we required the domain experts to manually step through their paths to discovery, because SocialAction did not automatically log the actions of each domain expert. We recommend that future researchers using our methodology should also provide logging and replaying features.

In the early-use phase, the domain experts shared why they felt certain insights were being impeded by missing features. All feature requests were compiled into a master list, and we estimated how long each feature would take to implement. The domain experts could then prioritize the features that they would prefer to have before the next interview session.

In the closing interview, we shared a list of the insights from our field notes and ensured they were documented correctly. The domain experts then discussed with us how SocialAction aided or impeded their analytical process throughout the case study. They also discussed their future plans and expected outcomes.

Two Case Studies

In order to validate our claim that SocialAction improves visual analysis of social networks, we conducted four case studies of domain experts with diverse skill sets, domains of knowledge, and social network expertise.[2] The domain experts were not recruited but sought out SocialAction on their own after facing challenges in making sense of social networks.

Owing to limited space, we describe only the case studies involving senatorial voting patterns and knowledge discovery for medical research. (The other two involved a consultant studying healthcare operational networks and a counter-terrorist researcher studying group dynamics in terrorist networks. Full descriptions of these case studies appear elsewhere.[2]) The descriptions of these case studies discuss only a fraction of the domain experts' insights but are representative of their overall experience.

Senatorial Voting Patterns

Congressional analysts are interested in partisan unity in the US Senate. For instance, the *Congressional Quarterly* calculates such unity by identifying every vote in which a majority of Democrats voted opposite a majority of Republicans, and then counting for each senator the percentage of those votes in which he or she 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.

Chris Wilson, an associate editor of *US News & World Report* at the time of our earlier study, was interested in voting patterns among US senators. He was seeking to uncover senatorial patterns, such as strategic, bipartisan, or geographic alliances, in the data set. He 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 (SNA) 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, ManyEyes, and NetDraw but found no 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 (for example, 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 (for example, 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 them to choose an arbitrary cut-off before analysis begins.

SocialAction's interactive statistics empower users to dig deeper, without forcing them 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 1a (page 44). 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 (see Figure 1b). 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.

Evaluating Information Visualizations

In the field of human-computer interaction, traditional laboratory-based controlled experiments have proven effective for many research projects. Controlled experiments are effective when comparing treatments of new widgets, displays, interaction methods, or input devices by measuring learning times, task performance times, or error rates. These experiments typically contain 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. The results are then generally applied to summary statistical methods, such as t-tests and ANOVA (analysis of variance), to highlight significant differences.

However, laboratory-based controlled experiments are less compelling for information visualization and visual-analytics (VA) research. VA systems are often designed for domain experts who work for days and weeks to carry out exploratory data analysis on substantial problems. These types of tasks are nearly impossible to reconstruct in a controlled experiment for a variety of reasons. First, it is difficult to recruit large numbers of analysts for a potentially lengthy time period. Second, exploratory tasks are poorly defined, so commanding users to perform specific tasks is not compatible with discovery. Third, users have diverse skills and unique experience, leading to wide variations in performance that might 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, researchers might desire more than what quantitative analyses provide. Hearing about the problems and frustrations users encounter as well as their thrilling tales of success is useful for research. For these reasons, we suggest a methodology of structured and replicated case study research methods to collect supporting evidence.[1]

The novelty of structured and replicated case studies is apparent from a review of the 132 papers in the 2005–2007 IEEE Information Visualization Conferences and the 2006 and 2007 IEEE Symposium on Visual Analytics Science and Technology (see Figure A). Only 39 papers had any user evaluation, and each tested users for less than two hours of tool usage. Furthermore, all but nine of these tests used domain novices who were given standard tasks.

However, Catherine 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.[2] In this spirit, Ben Shneiderman and Plaisant propose Multidimensional In-Depth Long-Term Case Studies to study the tasks of information visualization system users.[3] Their methodology suggests working closely with expert users and performing in-depth observations to capture users' creative activities during exploration. Of course, there is a long history of qualitative analysis and case studies, but here we highlight work that focuses on the unique demands of information visualization users.

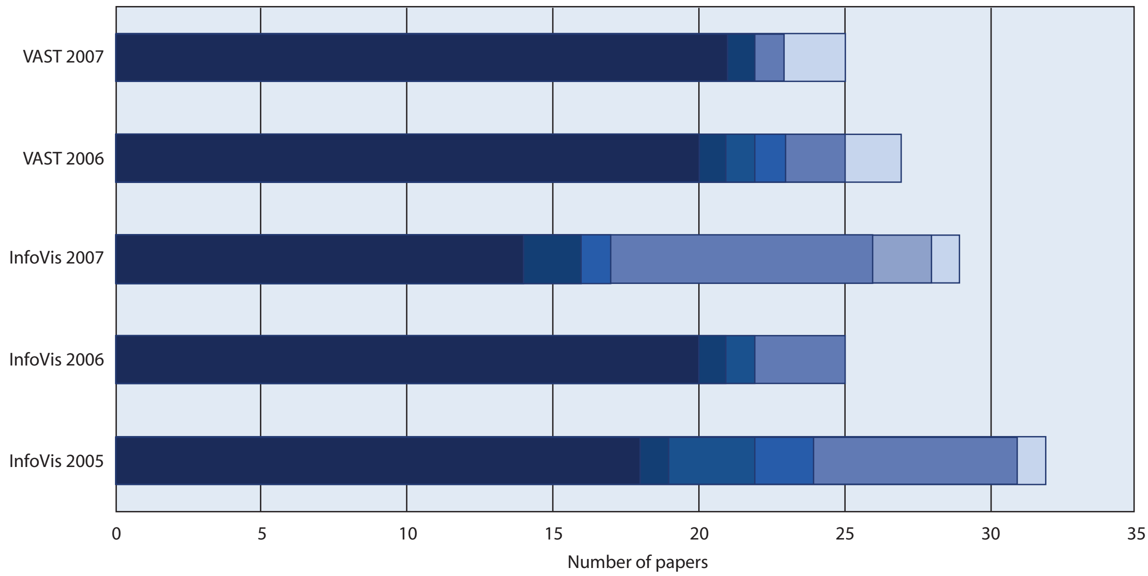
Purvi Saraiya and her colleagues identified characteristics of insight, arguably the primary purpose of visualization tools. By pairing tools with experts and measuring the number of insights reached, they empirically evaluated five visualization tools.[4] However, these evaluations did not capture long-term insights because the evaluation sessions lasted only a few hours. Saraiya and her colleagues 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.[5] Their work provides insight into the practices of actual data analysts, which have implications for both the design and evaluation of information visualization systems.

Jinwook Seo and Shneiderman also conducted three long-term case studies with domain experts, as well as a survey, that helped show the Hierarchical Clustering Explorer's efficacy and suggested improvements to it.[6] Victor González and Alfred Kobsa used a six-week case study with weekly interviews that illustrated that information visualization tools are most powerful when they are complementary to the workflow of analysts.[7] Finally, we conducted four replicated, structured, and long-term case studies to collect evidence that integrating statistics with visualization would facilitate discovery for social network analysts.[1] These experiments and the resulting reflections are described in the main article.

References

1. A. Perer and B. Shneiderman, "Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis," *Proc. Conf. Human Factors in Computing Systems* (CHI 08), ACM Press, 2008, pp. 265–274.

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 *betweenness centrality* (an SNA statistic that attempts to quantify the gatekeepers—that is, the bridges between communities) 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



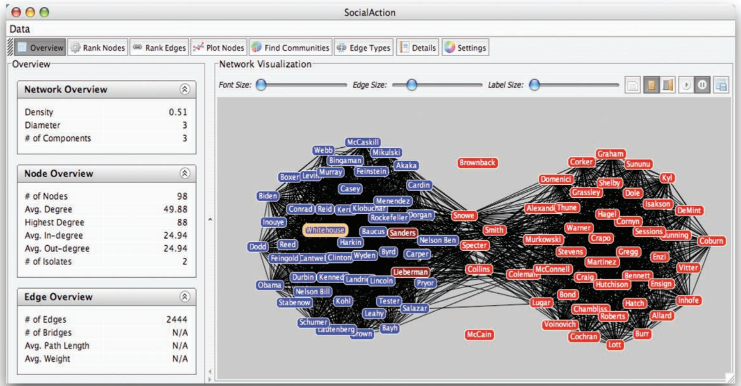
	InfoVis 2005	InfoVis 2006	InfoVis 2007	VAST 2006	VAST 2007
■ No evaluation	18	20	14	20	21
■ No evaluation possible	1	1	2	1	1
■ Case study	3	1	0	1	0
■ Usability	2	0	1	1	0
■ Controlled study	7	3	9	2	1
■ Logging	0	0	2	0	0
□ Survey	1	0	1	2	2

Figure A. Evaluation methods in papers from the 2005–2007 IEEE Information Visualization Conferences (InfoVis) and the 2006 and 2007 IEEE Symposium on Visual Analytics Science and Technology (VAST). Most papers featured no analysis, most evaluations were controlled studies, and no papers had structured, replicated case studies.

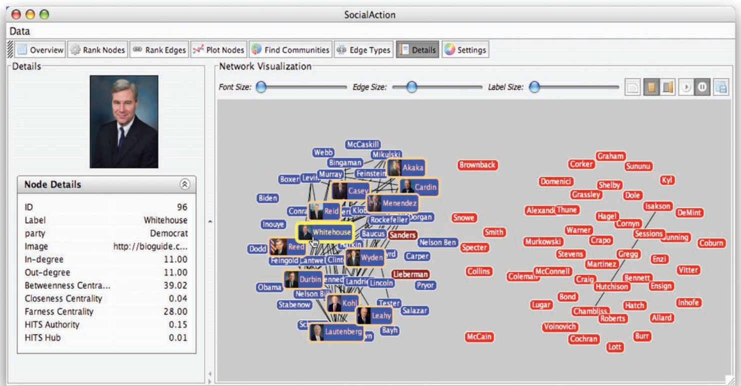
2. C. Plaisant, "The Challenge of Information Visualization Evaluation," *Proc. Working Conf. Advanced Visual Interfaces* (AVI), ACM Press, 2004, pp. 109–116.
3. B. Shneiderman and C. Plaisant, "Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-Depth Long-Term Case Studies," *Proc. 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, ACM Press, 2006, pp. 1–7.
4. P. Saraiya, C. North, and K. Duca, "An Evaluation of Microarray Visualization Tools for Biological Insight," *Proc. IEEE Symp. Information Visualization* (INFOVIS 04), IEEE CS Press, 2004, pp. 1–8.
5. P. Saraiya et al., "An Insight-Based Longitudinal Study of Visual Analytics," *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 6, 2006, pp. 1511–1522.
6. J. Seo and B. Shneiderman, "Knowledge Discovery in High-Dimensional Data: Case Studies and a User Survey for the Rank-by-Feature Framework," *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 3, 2006, pp. 311–322.
7. V. González and A. Kobsa, "A Workplace Study of the Adoption of Information Visualization Systems," *Proc. 3rd Int'l Conf. Knowledge Management* (I-KNOW 03), 2003, pp. 92–102; http://i-know.tugraz.at/previous/i-know03/papers/ivkm/Gonzalez_IVKM.pdf.

the sorts of insights that eluded Wilson prior to his analysis with SocialAction.

Outcome. Wilson was impressed 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 *US News & World Report* hopes to replicate some of its functionality for its online readers. This will provide readers with further data analysis opportunities. Since the case study, Wilson has moved to *Slate* but still uses Social-Action for investigative reporting. So far, analysis from SocialAction has led to an interactive feature



(a)



(b)

Figure 1. The social network of the US Senators' voting patterns (98 nodes, 4,753 edges).[2] Republicans are red, Democrats blue, and Independents maroon. (a) When the threshold is 180 votes, the partisanship of the parties appears automatically. (b) When the threshold is raised to 290 votes, the Democrats' relationships are much more intact than the Republicans. Figure 1b shows details for Senator Whitehouse, the senator with the highest degree (that is, number of relationships) at this threshold. (Source: ACM; used with permission.)

analyzing the social networks of steroid users in Major League Baseball, with more stories planned for the future.

Knowledge Discovery for Medical Research

The US National Library of Medicine (NLM) maintains PubMed (www.ncbi.nlm.nih.gov/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 might 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, an information retrieval expert, 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-article links and reach other relevant documents. A network of documents can be created by linking each document to the related articles that the algorithm recommended. The network's structure is important, because isolated documents without links from other relevant documents cannot be reached by browsing. Lin hoped to gain deeper insights about the usefulness of the algorithm by using SNA to explore the recommendation network. The recommendation network is not a social network. However, it demonstrates that although SocialAction is designed to support SNA, it also allows users to explore and interpret non-social networks, such as communication, financial, biological, and citation networks.

Early use. For the experimentation with SocialAction, Lin used data from the Text Retrieval Conference genomics test set. This data 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 (for example, 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 2). This supports the cluster hypothesis in information retrieval, which proposes that relevant documents tend to be more similar to each other than to nonrelevant documents. However, there were also a number of isolated islands of documents (notice the disconnected, star-shaped clusters in Figure 2). These represent documents that would be unreachable by users through 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. 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 the 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 (for example, the number of relevant documents linked from each relevant document). Because 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 the networks had varying properties (number of suggested articles, number of relevant documents, and 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 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 2 shows the network components broken down into smaller communities using the hierarchical clustering algorithms available in SocialAction.[1] Each community is surrounded by a bubble colored on the basis of 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, although this feature allowed Lin to advance his exploration, he chose to leave these results out of his analysis owing 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 (for example, without having to go back to the search results and reformulate a query). However, as we mentioned before, Lin and his colleagues also identified isolated clusters, which represented relevant documents that were not reachable by browsing.

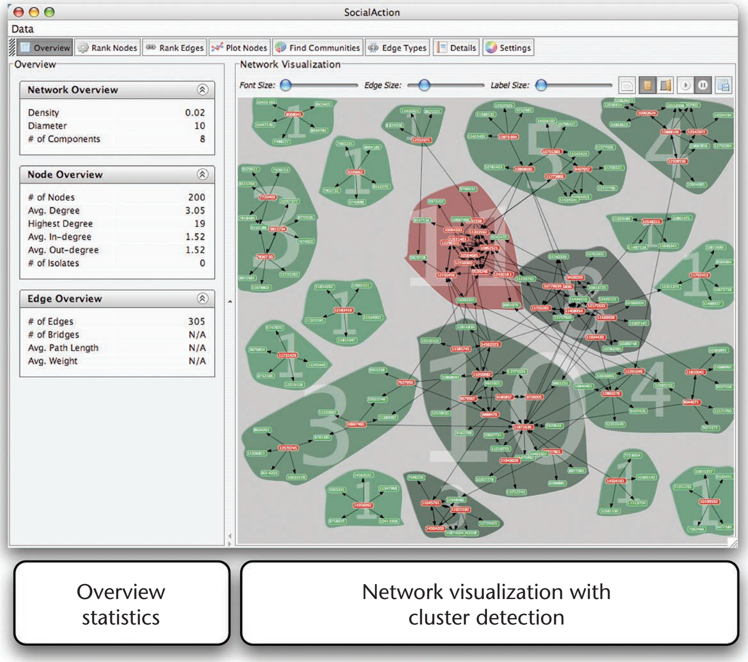


Figure 2. The recommendation network of a query on PubMed documents (200 nodes and 305 edges).[2] Relevant documents are red; nonrelevant are green. The community algorithm highlights closely connected clusters in the network. Communities are color coded by the percentage of relevant documents and labeled by the number of relevant documents. (Source: ACM; used with permission.)

The results of this analysis led to publication of a high-quality research article in a prestigious information retrieval journal. 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.

Summary of the Case Studies

These case studies show that the integration of statistics and visualization can improve exploratory data analysis by supporting the explorative and creative tasks of analysts. Interactive techniques are a key part of the design to make both the statistical and visual components comprehensible. Without interactions such as ranking and filtering, statistical output and network visualizations might be too complex to interpret.

However, in most information visualization systems, these interactive techniques often focus on inherent attributes rather than computed attributes from statistical techniques or data mining. In order to make the lessons learned from SocialAction more applicable to a broad range of researchers and designers, we developed our set of guidelines for enhancing interactive visualization with computed attributes.

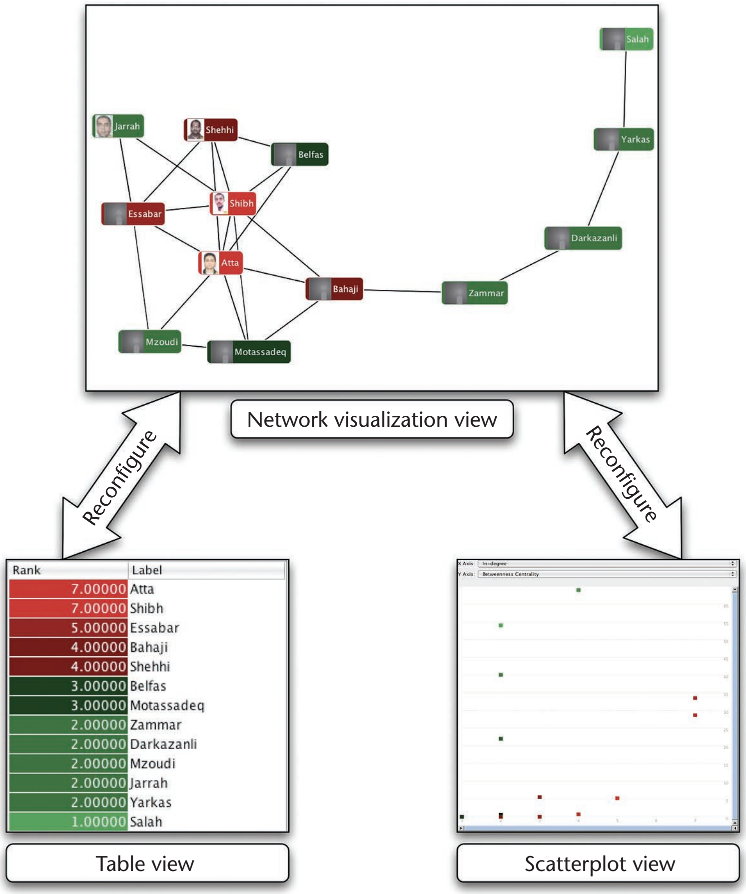


Figure 3. The *reconfigure* interaction with computed attributes, as implemented in SocialAction. The network visualization is reconfigured to tables and scatterplots. A table view, ranked by a computed attribute, allows users to focus attention on important nodes. A scatterplot view, presenting two computed attributes, allows users to compare multiple importance rankings in a visualization that has been shown effective at highlighting patterns, gaps, and outliers.

Design Guidelines for Information Visualizations with Computed Attributes

Humans can be quite good at scanning data, recognizing patterns, and remembering images. However, as data grows larger and more complex, it is clear that interaction is necessary to present data interpretable by humans. Ji Soon Yi and his colleagues proposed seven categories of interaction techniques for information visualization systems: *reconfigure, connect, encode, select, filter, abstract/elaborate, and explore*.[5] These categories suggest ways that users can navigate through complex information landscapes. However, each of these techniques usually relies on inherent attributes of the data.

For instance, in a social network of male and female students, users could select students of interest, explore by following the paths of relationships to other students, or filter on the basis of gender. However, the inherent attribute-based interactions might not support the needs of certain users. If the tasks are to find the gatekeepers, communities, and most popular students, an algorithmic approach might be faster and more precise. Thus, it seems to make sense that the visualizations should be augmented with these computed attributes if they are relevant to users' tasks.

We used the seven interaction categories as the basis for our guidelines on augmenting visualization with computed attributes. In each of these interactions, users should be in complete control and not have to rely on interpreting a black box of automatic algorithms. Of course, to integrate computed attributes, information visualization tools will need to be more sophisticated. Users must navigate both the visualization and the statistical algorithms. In order to aid the design of such tools, we've developed guidelines for navigating the statistical algorithms as well.

Reconfigure: Providing Statistical Perspectives

The reconfigure interaction is common in information visualizations to provide users with a different perspective on the data. For instance, Spotfire (http://spotfire.tibco.com) provides eight different perspectives to visualize tabular data, such as scatterplots, bar charts, and heat maps, allowing users to easily switch between representations that best suit their task at hand. In a similar spirit, the Table Lens allows users to sort and rearrange columns of tabular data to highlight different patterns in the data.[6] However, most of these techniques focus on reconfiguration based on inherent attributes, not computed attributes.

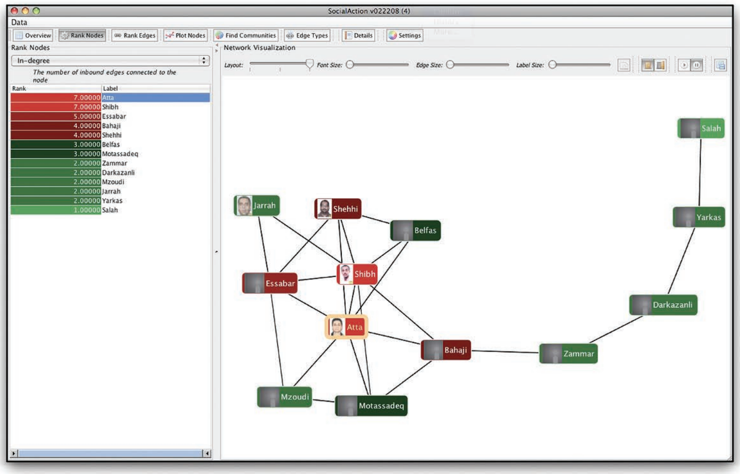
The complexity of network visualizations is an example of when a reconfiguration can be useful. Whereas others have tried reconfiguring network visualizations into trees or matrices, a reconfiguration based on computed attributes can provide an even simpler view of the data while highlighting statistically interesting properties. SocialAction reduces the network visualization into tables and scatterplots. A table view, ranked by a computed attribute, allows users to focus attention on important nodes. A scatterplot view, presenting two computed attributes, allows users to compare multiple importance rankings in a visualization that has been shown effective at highlighting patterns, gaps, and outliers. The reconfigurations in Figure 3 illustrate how the reconfigure interaction can be made even more powerful by computed attributes.

Connect: Coordinating Statistics and Visualization

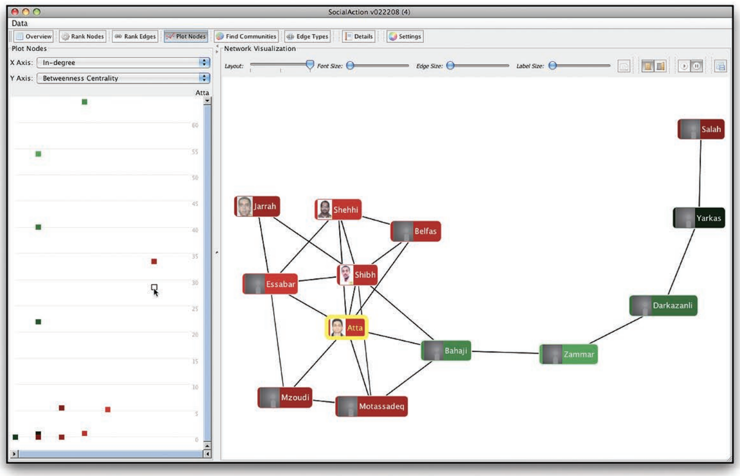
The connect interaction is in synergy with the reconfigure interaction. It aims to highlight associations and relationships between data items. As the reconfigure interaction suggests, it can be advantageous to users to see different displays of the same data. The connect interaction lets users view these displays at the same time. A common technique in this category is brushing, which allows users to select a data item in one view and see the item in multiple views. Brushing is most often used in coordination with different projections of inherent attributes of the data. For instance, if a Spotfire user selects a data point in a scatterplot, the corresponding data point will also be highlighted in an associated bar chart.

Providing reconfigured views of computed attributes is a good first step, but connecting these views is essential. Users should be able to browse the data visualization and the computed-attribute views in a coordinated manner. At times, the inherent attributes might reveal discoveries, and at other times, the computed attributes will help reveal clues during analysis.

SocialAction connects these two views together using side-by-side displays that are coordinated (see Figure 4). If a data point is in one view, the same point is represented in another view. Users can brush from one view to the other. If users wish to find nodes with certain structural properties, they can choose an algorithm that detects those nodes in a sorted table instead of being forced to visually scan a complex visualization. If users care about multiple structural properties, a scatterplot can saliently show the intersection between them. However, computed attributes might not always measure what users are seeking. By connecting both views, users can judge the utility of the algorithms and reflect on the algorithms' impact on their tasks. Thus, the connect interaction with computed attributes allows users to both learn about the data and the quality of the algorithms. SocialAction also supports connect by providing coordinated visual encoding (described in the next section).



(a)



(b)

Figure 4. The *connect* interaction with computed attributes, as implemented in SocialAction: The (a) table view and (b) scatterplot view are connected with the network visualization, providing side-by-side views, brushing operations, and coordinated visual encoding.

Encode: Representing Computed Attributes

A popular use of the encode interaction is to allow users to use color, size, fonts, shapes, and orientation to visually encode data points with attributes of interest. This allows visualizations to keep their spatial arrangement constant while visually presenting additional data about each of the nodes. This is a widely used technique in many information visualization systems; however, most encoding focuses on inherent attributes. Visually encoding with computed attributes is a natural extension and a convenient way to augment visualizations.

SocialAction follows this design goal by using color to encode results from algorithms. By encoding nodes and edges with computed attributes, SocialAction lets users easily find entities with certain statistical features (for example, the most popular individuals or the gatekeepers). By default, color is defined along a red-black-green spectrum. Values with the highest ranking are red, those in the middle are black, and those with the lowest ranking are green. SocialAction assigns these values along this gradient on the basis of the value of a computed attribute. When users select a different computed attribute, the colors update appropriately. Noncolorblind human eyes can easily distinguish between red and green, suggesting that this color spectrum is effective. These computed-attribute encodings can provide clues about the topology even when the topology is too cluttered and dense to make sense of, as demonstrated in Figure 5 (next page).

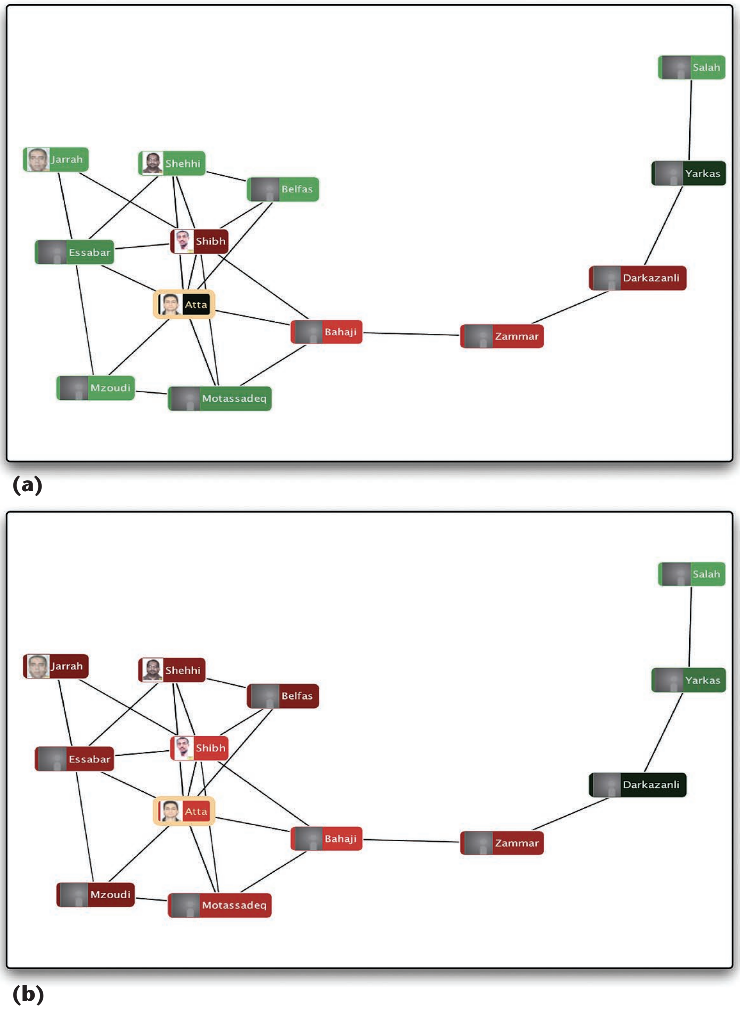


Figure 5. The *encode* interaction for computed attributes, as implemented in SocialAction: nodes visually encoded with (a) betweenness centrality, a computed attribute that measures for gatekeepers, and with (b) closeness centrality, a computed attribute that measures network access.

Select: Marking Interesting Computed Attributes

The select interaction allows users the ability to mark a data item as interesting. After users find and mark a data point as interesting, the selection should be persistent. This is a particularly important design goal when many different computed attributes are available to users. During exploratory data analysis, users might not know which computed attributes will lead to insights. The ability to efficiently switch between various computed attributes should be supported. But ultimately, users might care about the effect of each available algorithm on particular data points. When both computed-attribute values and layouts can change, the select interaction can be important for analysts.

This design goal is demonstrated in SocialAction, where users have the ability to select a node at any time (see Figure 6). They can choose a node in either the network visualization or the statistical views. No matter whether they change the layout of the network or compute a new statistical measure, both views will keep track of the previously selected node. Keeping selection information persistent is important to allow users to be more adventurous when trying additional statistical algorithms. Users shouldn't have their exploration feel constrained by the system; instead, the system should give them the freedom to creatively select and analyze particular data points of interest.

Filter: Focusing on Important Data

Filtering can decrease the complexity of the visualization by removing data points that aren't immediately relevant to the task of users. Many information visualization systems allow users to filter out certain data points on the basis of inherent attributes, and are most effective with dynamic queries and range sliders. However, few systems allow users to remove data points that are deemed statistically less important. Allowing users to filter by computed attributes is one way to achieve this goal. Although computed attributes can be displayed with visual encoding and coordinated with statistical views, the resulting visualizations might still be too complex to comprehend. However, the ability to filter out data that is relevant according to task-related computed attributes is an effective way of reducing complexity.

The design goal of filtering by computed attributes should be integrated into all systems wishing to give users more control over the visualization display. As with inherent attributes, users should be able to use dynamic queries and range sliders to see how the filtering process affects their data. SocialAction follows this design goal, which is extremely important in social network visualizations, which are typically incomprehensible when they have more than 50 nodes and edges.

Figure 7 (page 50) shows how a social network visualization is made more comprehensible by filtering according to statistical rankings—in this case, betweenness centrality. Figure 7a shows the network unfiltered; Figure 7b shows the filtering enabled. In Figure 7b, the important nodes can be read, and the edges between them are apparent.

Abstract/Elaborate: Reducing or Increasing Detail

When information visualizations are too dense, abstracting the data into higher-level components can be useful. Inversely, when the information visualizations are too sparse, elaborating details can be effective. By abstracting or elaborating in statistically significant ways, users might understand the data more effectively.

Clustering is one statistical technique that adds or reduces detail. For example, in a social network visualization, it can be difficult to discern where various communities of tightly connected individuals exist. Abstracting and elaborating can both be used to display the statistical findings from the clustering.

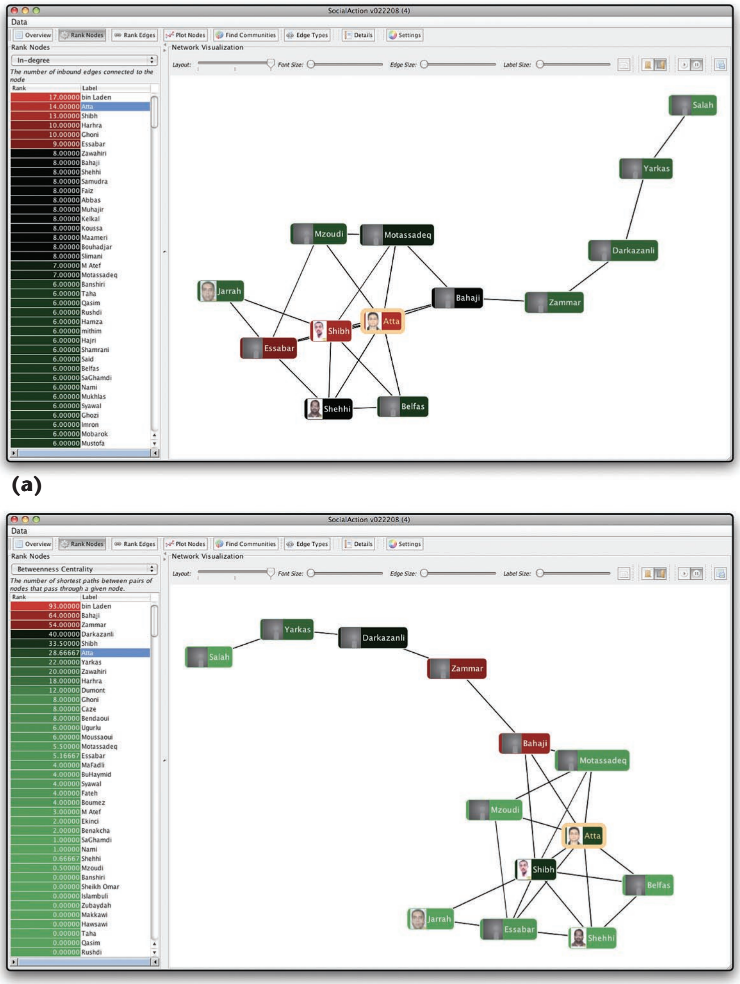
Figure 8a (page 51) shows a social network visualization using only a force-directed layout. Figure 8b shows how SocialAction incorporates community information into the visualization by surrounding each community with a polygon. This new information allows users to see which nodes belong to which community and which relationships span multiple communities. Compared to Figure 8a, rich information previously hidden from the human's eyes is now presented thanks to statistical algorithms. This community information could also be used to simplify the visualization by turning each community into a metanode.[1]

Explore: Reaching Insights

Algorithms from statistics and data mining are often created to find interesting properties of data. These results can act as suggestions for exploration by users. However, there are often a variety of algorithms for measuring a data set. For example, SPSS Statistics (www.spss.com/spss/alpha.htm), a leading statistical-analysis tool, includes over 80 sophisticated statistical techniques. So, being able to quickly assess the usefulness of an algorithm is important.

One way to help users do this is to organize techniques according to tasks the users are trying to accomplish, rather than present a lengthy list. In SocialAction, we organized the numerous SNA techniques into six tasks related to finding important individuals, relationships, and communities.[3] These tasks were based on knowledge gained through interviews with analysts who described their methodologies. Obviously, different types of users might require different tasks. User interfaces should be robust enough to support multiple user types. This robustness might add an additional layer of complexity to user interface design. However, focusing on tasks rather than features will allow users to focus on discoveries rather than navigating menus.

In addition, in SNA, nearly every researcher seems to invent a centrality measurement to suit his or her needs. Analysts generally have a finite amount of time to analyze their data, so tools should bring their attention to the most useful SNA algorithms.



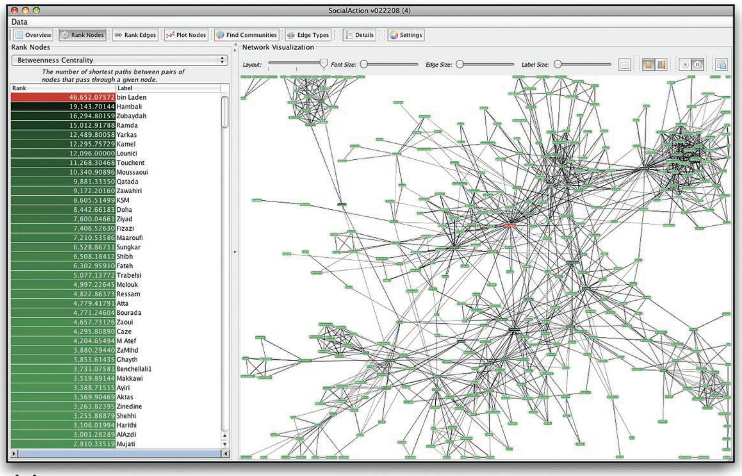
(a)

(b)

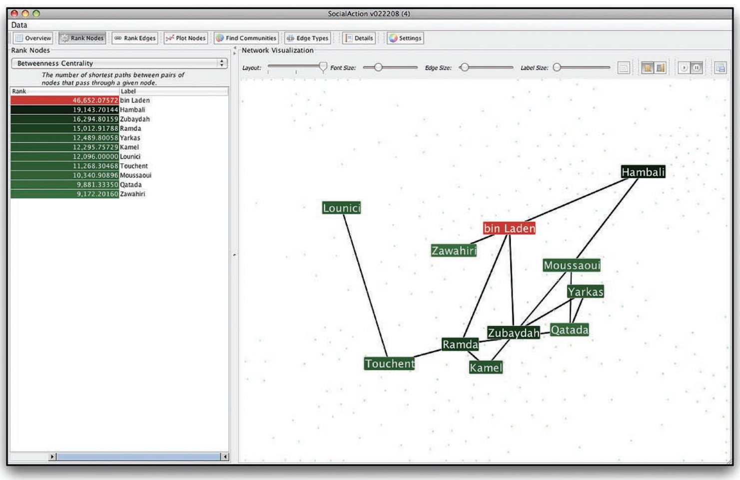
Figure 6. The *select* interaction for computed attributes, as implemented in SocialAction: In (a), a central actor is marked as interesting. In (b), the network layout has changed and a different computed attribute was chosen. However, the same actor is still marked owing to the persistence of the select interaction.

So, when developing SocialAction, we reviewed and tabulated the use of ranking algorithms in popular social network journals, such as *Connections*. The popularity of the algorithms' use provided guidelines for which ones we implemented. Their popularity also formed the basis for their order in the interface. The effect of this is twofold: explorers will more often select algorithms that are respected by their peers, and users can quickly access the algorithms they most likely care about.

Algorithms should be optimized to run in real time. If the algorithms are too slow, algorithmic results should be precomputed whenever possible. Another strategy is to log which algorithms are most often run by the users and to run them on a background thread when the user is performing less-CPU-intensive actions.



(a)



(b)

Figure 7. The *filter* interaction for computed attributes, as implemented in SocialAction: (a) the complete social network before filtering and (b) filtered nodes with high betweenness centrality. This highlights how filtering allows users to focus on nodes with interesting characteristics instead of them becoming lost in the chaotic overview.

Statistical techniques can yield valuable discoveries, but typical data analysis tools support only opportunistic exploration that might be inefficient and incomplete. When the number of tasks is large and the algorithms are complex, guides can help domain expert users through complex exploration of data over days, weeks, and months. In fact, we believe this idea is so important we describe a refined architecture that uses *systematic yet flexible* (SYF) design goals to guide domain expert users through complex exploration of data over days, weeks, and months in a conference paper.[3]

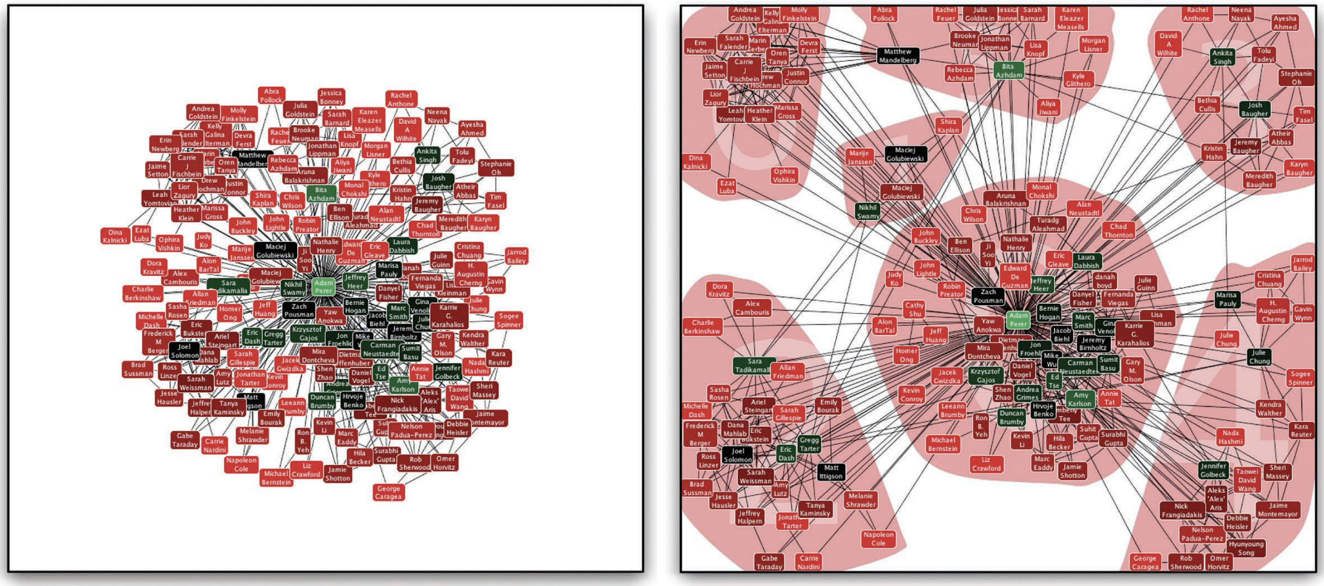
Our case studies clearly show that SocialAction led to new insights and discoveries. These creative discoveries might have been lost or undermined in traditional experiments with summary statistics. However, the case studies do not show the extent to which SocialAction was responsible for the discoveries. As we mentioned before, during the interviews, the domain experts often highlighted their great moments of discovery. Unfortunately, moments of limited success were recorded or remembered less frequently, which provided less useful feedback about negative aspects of the design.

Logging user actions is an obvious extension to the methodology. Although logs alone will not capture the full story of exploration, they can be used in conjunction with interviews to refresh users' memories as well as figure out a quantitatively accurate version of where users spent most of their time. Logs might highlight users getting lost during exploration or never using certain features that might have led to insights. Logging will hopefully serve as a tool to improve the accuracy of reporting on insights and as a reference for reporting on failures. Logging that allows users to annotate important states during analysis and later replay them, as described in the SYF architecture,[3] should yield advantages for data collection during evaluation as well.

Here, we've focused mostly on social networks. However, integrating statistics and visualization goes beyond social networks—statistical algorithms and data-mining results can aid in analyzing temporal, hierarchical, and multidimensional data. The design goals and examples can extend to these and other complex data types—to emphasize the power of an integrated approach.

References

1. A. Perer and B. Shneiderman, "Balancing Systematic and Flexible Exploration of Social Networks," *IEEE Trans. Visualization and Computer Graphics*, vol. 12, no. 5, 2006, pp. 693–700.
2. A. Perer and B. Shneiderman, "Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis," *Proc. Conf. Human Factors in Computing Systems* (CHI 08), ACM Press, 2008, pp. 265–274.
3. A. Perer and B. Shneiderman, "Systematic yet Flexible Discovery: Guiding Domain Experts through Exploratory Data Analysis," *Proc. 13th Int'l Conf. Intelligent User Interfaces* (IUI 08), ACM Press, 2008, pp. 109–118.
4. B. Shneiderman and C. Plaisant, "Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-Depth Long-Term Case Studies," *Proc. 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, ACM Press, 2006, pp. 1–7.



(a) (b)

Figure 8. The *elaboration* interaction for computed attributes, as implemented in SocialAction: a social network visualization (a) using only a force-directed layout and (b) after elaboration with computed attributes from a community detection algorithm. This screenshot highlights how elaborated details on a social network allow analysts to see structure that was impossible to see without them.

5. J.S. Yi et al., "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," *IEEE Trans. Visualization and Computer Graphics*, vol. 13, no. 6, 2007, pp. 1224–1231.
6. R. Rao and S.K. Card, "The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus + Context Visualization for Tabular Information," *Proc. Conf. Human Factors in Computing Systems* (CHI 98), ACM Press, 1998, pp. 329–399.

***Adam Perer** is a research scientist in the Social Technologies Group at IBM's Haifa Research Lab. His research interests include human-computer interaction, information visualization, social technologies, and social networks. Perer has a PhD in computer science from the University of Maryland. Contact him at adamp@il.ibm.com; http://perer.org.*

***Ben Shneiderman** is a professor in the University of Maryland's Department of Computer Science and the founding director of the university's Human-Computer Interaction Laboratory. His research interests are network visualization and electronic health records. He's the lead author of* Designing the User Interface: Strategies for Effective Human-Computer Interaction *(5th ed., Addison-Wesley, 2009). Contact him at ben@cs.umd.edu; www.cs.umd.edu/~ben.*

For further information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.