

Finding Beautiful Insights in the Chaos of Social Network Visualizations

Adam Perer

My purpose throughout is to interpret the material by juxtaposing and assembling the notations into a unified, coherent whole.

—Mark Lombardi, 2000

MARK LOMBARDI WAS PERHAPS THE PERFECT NETWORK LAYOUT ALGORITHM. As an artist intent on communicating complex networks of financial and political scandals, he diligently drew networks where nodes never overlap, edges rarely cross, and the connections are smooth and curvy (Figure 10-1). This amount of grace and sensitivity is rarely present in the visualizations of social networks created by computational means. While advanced computational layout algorithms may be grounded in physical models of springs and forces, they rarely highlight patterns and trends like Lombardi's drawings do. This chapter details my attempts to empower users to dig deeper into these chaotic social network visualizations with interactive techniques that integrate visualization and statistics.

Visualizing Social Networks

The increasing amount of digital information in modern society has ushered in a golden age for data analysis. Ample data encourages users to conduct more frequent exploratory data analyses to explain scientific, social, cultural, and economic phenomena. However, while access to data is important, it is ultimately insufficient unless we also have the ability to understand patterns, identify outliers, and discover gaps. Modern databases are simply too large to examine without computational tools that allow users to process and interact with the data.

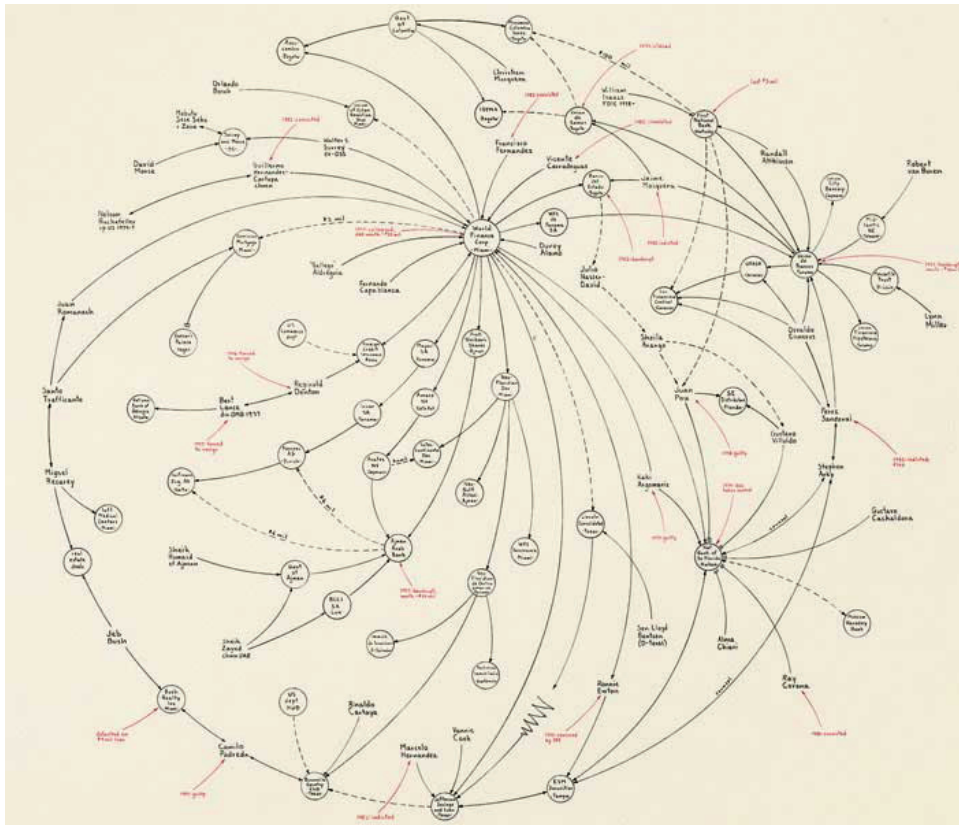


Figure 10-1. An example of the artist Mark Lombardi's hand-drawn social networks: "World Finance Corporation, Miami, Florida, c. 1970-79 (6th Version)" (1999); image courtesy of PIEROGI Gallery, Brooklyn, NY.

Our most powerful sensory receptors—our eyes—have far more bandwidth and processing power than our receptors for smell, sound, taste, or touch. Presenting data through information visualizations is therefore an effective way to take full advantage of the strong capabilities of our most powerful human perceptual system. However, choosing an effective presentation is challenging, as not all information visualizations are created equally. Not all information visualizations highlight the patterns, gaps, and outliers important to analysts' tasks, and furthermore, not all information visualizations "force us to notice what we never expected to see" (Tukey 1977).

A growing trend in data analysis is to make sense of linked data as networks. Rather than looking solely at attributes of data, network analysts also focus on the connections between data and the resulting structures. My research focuses on understanding these networks because they are topical, emergent, and inherently challenging for analysts. Networks are difficult to visualize and navigate, and, most problematically, it is difficult to find task-relevant patterns. Despite all of these challenges, the network perspective remains appealing to sociologists, intelligence analysts, biologists, communication theorists, bibliometricians, food-web ecologists, and many other professionals. The growing popularity of social network analysis (SNA) can be seen in, and inspired by, popular bestselling books such as Malcolm Gladwell's *The Tipping Point* (Back Bay Books), Albert-László Barabási's *Linked* (Plume), and Duncan Watts's *Six Degrees* (Norton). Countless analysts wish to analyze their network data, but there are few mature or widely used tools and techniques for doing so.

Network analysts focus on relationships instead of just the individual elements that can explain social, cultural, or economic phenomena; how the elements are connected is just as important as the elements themselves. Prior to the social network analysis perspective, many analysts focused largely on inherent individual attributes and neglected the social facet of behavior—i.e., how individuals interact and the influence they have on one another (Freeman 2004). Using newer techniques from the social network community, analysts can find patterns in the structures, witness the flow of resources or messages through a network, and learn how individuals are influenced by their surroundings.

In practice, social network visualizations can be chaotic, particularly when the network is large. Visualizations are useful in leveraging the powerful perceptual abilities of humans, but cluttered presentations, overlapping edges, and illegible node labels often undermine the benefits of visual exploration. In these situations, interactive techniques are necessary to make sense of such complex static visualizations. *Inherent* attributes are the attributes that exist in the dataset, such as gender, race, salary, or education level. Interactions such as zooming, panning, or filtering by the inherent attributes of nodes and edges can simplify complex visualizations. Unfortunately, such techniques may only get users so far with complex networks and may not tell the whole story, particularly in small-world networks where dense connections will rarely untangle (van Ham 2004). Inherent attributes lack the structural, topological information critical to social network analysts. Our major contribution is to augment information visualizations with *computed* attributes that reflect the tasks of users. Computed attributes can be calculated from relevant statistical importance metrics (e.g., degree or betweenness centrality), clustering algorithms, or data mining strategies.

This approach of leveraging computed attributes is particularly valuable for social network analysts, as they have also come to believe that inherent attributes do not tell the whole story. In fact, an approach taken by many social network analysts is to ignore inherent attributes during exploration to avoid bias, and to only focus on the data's structural properties. For social network analysts, computed attributes can be calculated with a rich set of statistical techniques, from sociology to graph theory, that allow analysts to numerically uncover interesting features within their networks. Analysts might seek a tight-knit community of individuals, or the gatekeepers between them, or the most centrally powerful entities; there are a variety of sophisticated algorithms for finding these traits.

Most visualization tools aim to project complex data into comprehensible views. However, few tools assist users by providing computed attributes that highlight important properties of their data. Users can switch back and forth between statistical and visualization packages, but this can result in an inefficient flow in the analysis process, which inhibits discovery.

SocialAction is the software tool Ben Shneiderman and I created to explore these issues (<http://www.cs.umd.edu/hcil/socialaction>). It provides meaningful, computed attributes on the fly by integrating both statistics and visualizations to enable users to quickly derive the benefits of both. *SocialAction* embeds statistical algorithms to detect important individuals, relationships, and clusters. Instead of presenting statistical results in typical tabular fashion, the results are integrated in a network visualization that provides meaningful computed attributes of the nodes and edges. With computed attributes, users can easily and dynamically filter nodes and edges to find interesting data points. 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 sometimes-chaotic visualizations, allowing users to focus on statistically significant nodes and edges. The presence of these rich interactions within one consistent interface provides a fluid, efficient, visual analytic system that allows users to focus on insights and generating hypotheses rather than managing a medley of software packages. I'll walk you through this rich interaction of statistics and visualization later, but let's begin with the motivation for why this is necessary.

Who Wants to Visualize Social Networks?

My fieldwork with social network analysts, both in academia and industry, suggests that pure statistical analysis is the most commonly used technique when attempting to interpret social networks. Although network visualizations are common in research publications and reports, they are typically created for communicative purposes after the analysis is complete and not necessarily visualizations used during the exploratory analysis.

A history of the use of visual images in social networks is described in “Visualizing Social Networks” (Freeman 2000), including one of the earliest known examples of a social network visualization by Jacob Moreno in 1934. In Figure 10-2, the triangle nodes are boys and the circle nodes are girls. Without knowing any details about who the individuals in this classroom are, one quickly learns from the visualization that 1) boys are friends with boys, 2) girls are friends with girls, 3) one brave boy chose a girl as his friend (although this was not reciprocated), and 4) there is an isolated group of two girls. This visualization typifies how a legible and well-positioned network can explain the social structure of individuals.

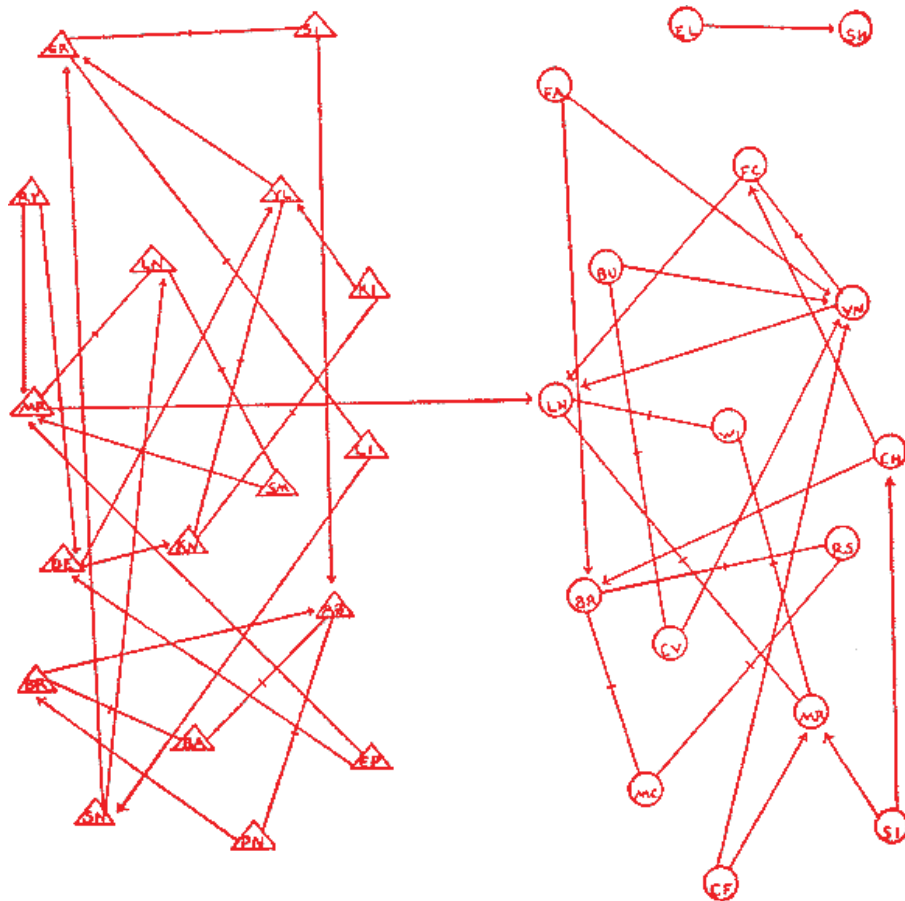


Figure 10-2. One of the earliest social network visualizations: Jacob Moreno's Friendship Choices Among Fourth Graders (Moreno 1934).

Social network data is extremely complex, as the dimensionality of the data increases with each relationship. Those familiar with network visualizations might sympathize with these statistically attuned practitioners, as it is very difficult to design a useful network visualization when the number of nodes or edges is large. Large network visualizations are typically a tangled set of nodes and edges, and rarely achieve “NetViz Nirvana” (a phrase coined by Ben Shneiderman to describe 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 these complex visualizations.

My first argument is that it is hard to find patterns and trends using purely statistical methods. My 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—to arrive at beautiful visualizations. The design of SocialAction centers on this goal.

The Design of SocialAction

Structural analysts have proposed numerous measures for statistically assessing social networks. However, there is no systematic way to interpret such networks, as those measures can have different meanings in different networks. This is problematic, as analysts want to be certain they are not overlooking critical facets of the network. In order to make exploration easier, I interviewed social network analysts and reviewed social network journals to tabulate the most commonly used measurements. I then organized these measures into six user-centered tasks: Overview, Rank Nodes, Rank Edges, Plot Nodes, Find Communities, and Edge Types. In the following sections, I’ll describe each of these tasks and their associated features in detail. However, let’s first begin with an illustration of the main goals of the process.

Shneiderman’s Visual Information-Seeking Mantra—“Overview first, zoom and filter, then details on demand” (Shneiderman 1996)—serves as guidance for organizing the complex tasks of a social network analyst. Analysts begin with an overview of the network, both statistically and visually; see Figure 10-3(a). Measurements of the entire network, such as the density, diameter, and number of components, are computed and 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 apply statistical importance metrics common in social network analysis to measure the nodes (also known as vertices) and edges (also known as links). For instance, analysts can rank the nodes by degree (the most connected nodes), betweenness (the gatekeepers), closeness (nodes that are well positioned 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, painting each node with the corresponding color. Users now can scan the entire network to see where the important nodes reside; see Figure 10-3(a).

To gain further insights, SocialAction allows users to continue on to step 2 of the Visual Information Seeking Mantra (“filter and zoom”). This is where most other social network analysis packages strand users. Panning and zooming naïvely is not enough to empower users: zooming into sections of the network forces 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; see Figure 10-3(b).

After analysts can make sense of global trends through statistical measurements and visual presentations, but their analyses often are incomplete without an understanding of what the individual nodes represent. Contrary to most other network visualizations, labels are always present in SocialAction. The controls for font size and length allow the analysts to decide their emphasis. In line with step 3 of the Visual Information Seeking Mantra (“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; see Figure 10-3(c).

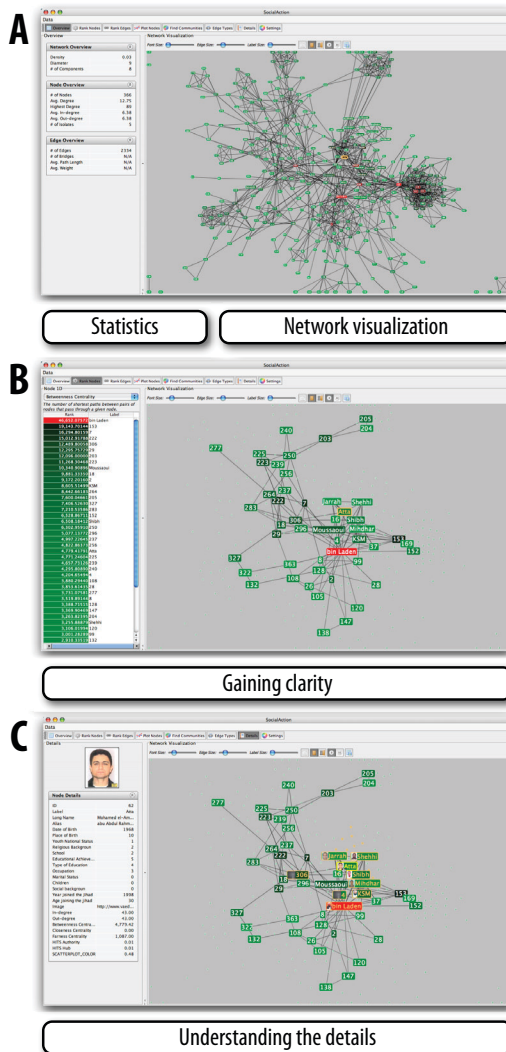


Figure 10-3. (a) The Statistics side of the interface allows users to choose statistical algorithms to find important nodes, detect clusters, and more. The Visualization side is integrated with the statistics. Nodes are colored according to their ranking, with red nodes being the most statistically important. (b) The gatekeepers are found using a statistical algorithm. Users filter out the unimportant nodes using a dynamic slider that simplifies the visualization while maintaining the node positions and structure of the network. (c) Labels are always given priority so users can understand what the data represents. When a user selects a node, neighbors are highlighted and details appear on the left.

For another, albeit more lighthearted, example, let's take a look at my personal social network on Facebook. If I visualize the connections using a standard network layout algorithm, I get a Jackson Pollack–like mess; there is something intriguing about the mess, but it certainly lacks the grace of a Lombardi piece. However, if I make use of some statistics (in this case, a clustering algorithm designed to detect communities), I get a much more sensible output. What used to be a bunch of tangled nodes and edges is now my social network grouped into meaningful categories. I can see clusters of my high school friends, my college friends, my graduate school friends, my Microsoft colleagues, and so on (Figure 10-4). An image devoid of meaning becomes beautiful, thanks to our dear algorithmic friends.

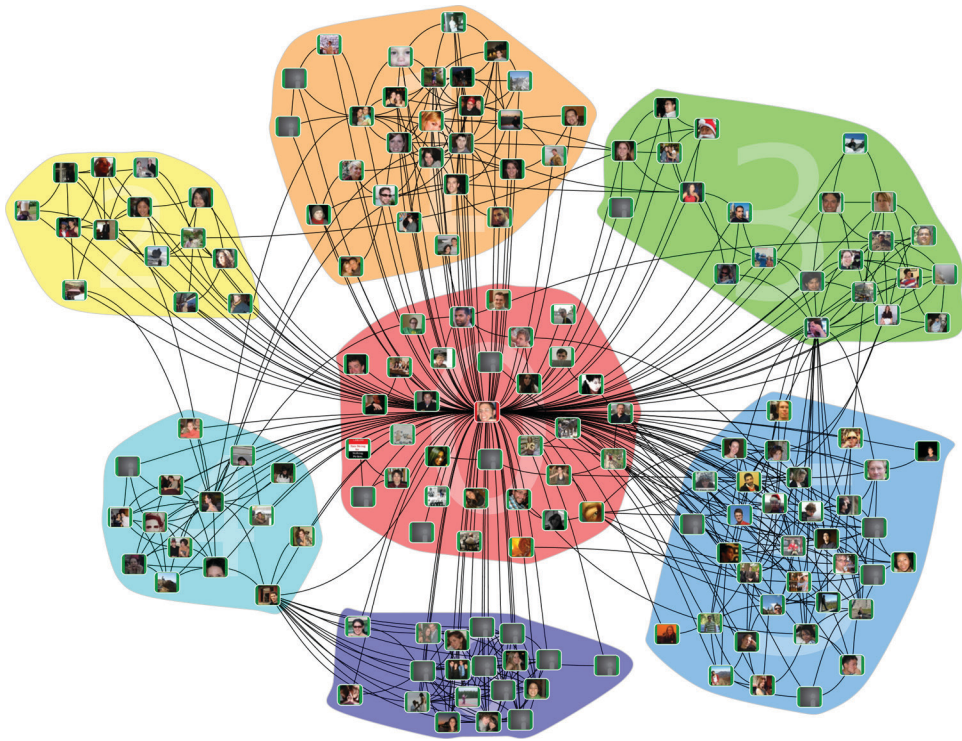


Figure 10-4. A visualization of my Facebook social network. By running a clustering algorithm on top of the network, seven meaningful communities of friends were found representing different facets of my life. Without clustering, the network was too tangled to provide any meaning.

In summary, bringing together statistics and visualizations yields 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 the sometimes-chaotic visualizations, allowing users to focus on statistically significant nodes and edges.

Case Studies: From Chaos to Beauty

Ultimately, what makes network visualization beautiful? An 18th-century Scottish philosopher, David Hume (1742), wrote:

Beauty is no quality in things themselves. It exists merely in the mind which contemplates them; and each mind perceives a different beauty.

However, Hume's view of beauty was contested. A Scottish associate, Henry Home (Lord Kames), believed that beauty could be broken down to a rational system of rules.

When it comes to visualizations based on underlying data, I side with Lord Kames. Insights offered are the measure of success for a beautiful visualization. Analysts may be seeking to confirm their intuitions, detect anomalies or outliers, or uncover underlying patterns. Chris North, a professor at Virginia Tech, characterizes insights as complex, deep, qualitative, unexpected, and relevant findings. While a helpful characterization, the impression is that measuring insights is perhaps as complicated as measuring beauty. Traditional laboratory-based controlled experiments have proven to be effective for many scientific tests, but do they work for insights? For instance, if I invented new display or input widgets, controlled experiments could compare two or more treatments by measuring learning times, task performance times, or error rates. Typical experiments would involve 20–60 participants each 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 would then be applied to check for significant differences in mean values. These summary statistics are effective, especially if there is small variance across users.

However, how does someone break insights into a set of measurable tasks? The first challenge is that analysts often work for days or weeks to carry out exploratory data analyses on substantial problems, and their work processes would be 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 by their nature poorly defined, so telling the users which tasks to carry out would be incompatible with discovery. Third, each user has unique skills and experience, leading to wide variations in performance that would undermine the utility of the 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, I wanted more than quantitative analyses of the tool; I also wished to hear about the problems and frustrations users encountered, as well as their thrilling tales of success. For such reasons, I turned to structured and replicated case study research methods to decide if SocialAction could generate beautiful visualizations.

The following sections summarize a few of my case studies of real analysts using SocialAction to visualize their own data. In homage to Mark Lombardi, I have chosen here to report on the covert networks of politicians and terrorists.

The Social Network of Senatorial Voting

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 in which that senator voted with his or her party. This metric can be useful for tracking an individual senator's party loyalty from year to year, but it does not reveal much about the overall patterns in the body.

Chris Wilson, then an associate editor for the *US News & World Report*, became interested in voting patterns among United States senators in 2007. Wilson set out to uncover senatorial patterns such as strategic, bipartisan, or geographic alliances in the dataset. He spent significant effort mining voting data from public databases, but was unable to find any 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.

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 led to a very dense network, because there were certain uncontroversial resolutions that all senators voted for (e.g., Resolution RC-20, a bill commending the actions of "the Subway Hero" Wesley Autrey). All the senators were connected, resulting in a visualization resembling a huge, tangled web.

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 10-5. Partisanship is strong even at this fairly low threshold, and the Republican senators who were most likely to vote with Democrats (Collins, Snowe, Specter, and Smith) are evident. This visualization suggests that in this particular Senate, although both parties were partisan, Republicans were less so than Democrats.

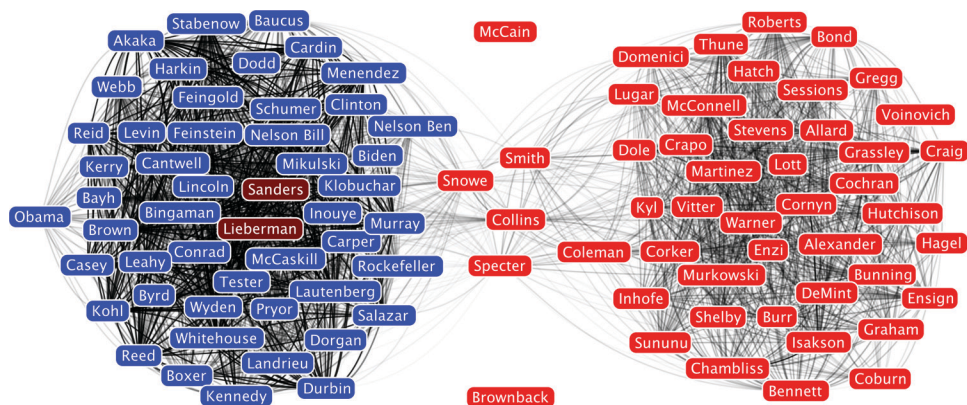


Figure 10-5. This visualization shows the voting patterns of U. S. senators during 2007. The red Republicans are on the right and the blue Democrats are on the left, with two Independents. Links indicate the similarity of voting records, revealing that Democrats had stronger party loyalty during 2007. Four Republican senators from Northeastern states often voted with Democrats. McCain and Brownback were campaigning for the presidency and did not vote often enough to be connected.

Another unexpected revelation was that the Democrats appeared to stay more tightly unified than the Republicans as the threshold increased, as evidenced by the much denser and darker connections on the Democrat side. Each edge is slightly transparent, but the constant overlapping of Democrats yields a dark mass, whereas the Republican side is much sparser. 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.

To determine the voting patterns of individual politicians, Wilson used SocialAction's statistical importance metrics. 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 had eluded Wilson prior to his analysis with SocialAction.

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 the *US News & World Report* and on Capitol Hill. SocialAction received so much attention internally that the magazine hopes to replicate some of its functionality for its online readers. Since completing the case study Wilson has moved to *Slate* magazine, but he still uses SocialAction for investigative reporting. Analysis using

SocialAction has already led to an interactive feature analyzing the social networks of steroids users in Major League Baseball (<http://www.slate.com/id/2180392>), and more stories are planned for the future.

The Social Network of Terrorists

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 world-wide research team that “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 PhD 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, Hendrickson believes theories regarding the group dynamics of terrorism have largely failed to properly measure the size, scope, and other dynamics of group relations. He proposes to systematically compare the density and type of relationships held by members of 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 bestselling book, *Understanding Terror Networks* (University of Pennsylvania Press). Hendrickson plans to update and formally apply social network analysis to this data as a part of his PhD dissertation.

The Sageman database has over 30 variables for each suspected terrorist. Among these variables are different types of relationships, including friends, family members, and educational ties. Hendrickson hypothesized that the types of relationships connecting two individuals will hugely affect their participation in terrorist attacks. He began his analysis using UCINET and was able to analyze some of his hypotheses. However, he believed UCINET did not facilitate exploring and generating new hypotheses. Hendrickson initially was skeptical of using visualizations for his analysis. He preferred being able to prove statistical significance quantitatively rather than relying on a human’s judgment of an image. However, he says 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 datasets. 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 structure. For instance, only the “Friend” relationships among jihadists are selected in Figure 10-6(a). (Compare this to the denser Figure 10-3(a), 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 attacks) are ranked the highest. However, when the religious ties are invoked, a different set of key jihadists emerges; see Figure 10-6(b).

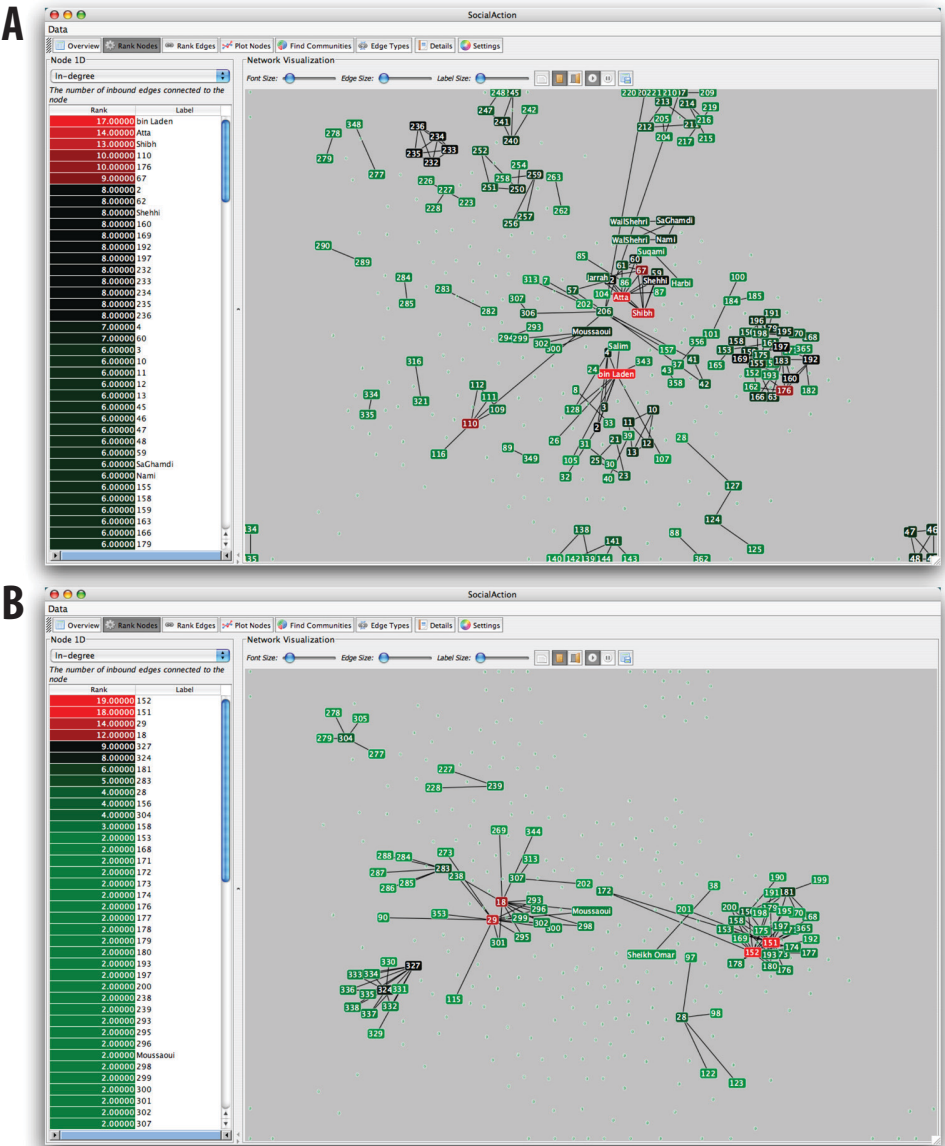


Figure 10-6. The multiplexity of the “Global Jihad” social network is demonstrated. The upper visualization (a) shows the Friendship network, with bin Laden the most popular individual. The bottom network (b), showing religious ties, offers a much different view of the terrorist organization.

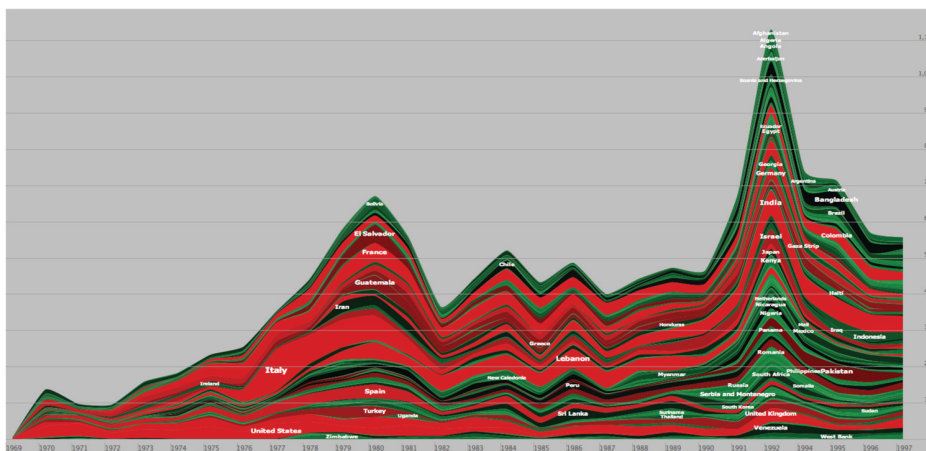
After analyzing the statistical attributes of nodes, Hendrickson became interested in understanding the individuals' attributes. For example, he was interested in answering questions like, "Does an individual's socioeconomic status or education level impact his position in the terrorist network?" Of course, social network data does not allow users to infer causation, but it may show correlation. Like statistical rankings in SocialAction, users can rank and filter based upon attributes. Hendrickson filtered out individuals without college degrees, religious backgrounds, or engineering expertise, and analyzed the results. The combination of nodal attributes with statistical filtering and plotting streamlined his accustomed workflow, and he commented that he might not have been as free in his thinking if it weren'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.

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 had conducted analyses with other SNA software, the integration of statistics and visualization in SocialAction 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 its databases.

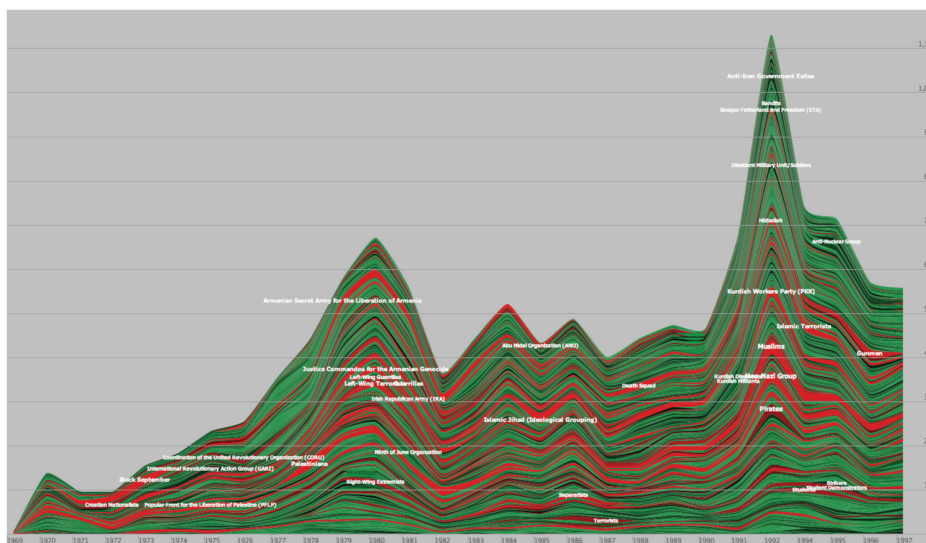
One other use of SocialAction by the START center was to look at networks that evolve over time. In their global terrorism network, nodes can be connected based on whether two people committed a terrorist attack in the same area, or used the same weapons, or came from the same region. Edges can also have temporal characteristics; for example, an edge could represent an attack in a certain year. The types of edges used depend on what types of questions the analyst is trying to answer. In tandem with a network diagram, users can see a stacked histogram, as in Figure 10-7. Each node is represented as a line and each column represents an edge type. The node's thickness in each column represents the node's ranking in the network of that edge type. The color is based on the node's overall ranking across all edge types.

In Figure 10-7, two stacked histograms are shown that demonstrate the evolution of a terrorist network over time. This particular network had two types of nodes: terrorist groups and the countries in which they had committed attacks. The country nodes are alphabetized and stacked in Figure 10-7(a), whereas all the terrorist groups appear in Figure 10-7(b). The thickness of the node at each year is based on the node's degree in the network. Nodes are colored based on their degree (red implies high degree, green implies low degree) and are labeled in their peak years (there is a clear peak of attacks in 1992). Various trends can be interpreted from this image, such as that Italy had many different groups attacking in the earlier years, whereas India had peak activity in the later years.

A



B



172 BEAUTIFUL VISUALIZATION

In 2007, the temporal visualizations shown in Figure 10-7 were on display at the New York Hall of Science as a part of the Competition on Visualizing Network Dynamics (<http://vw.indiana.edu/07netsci/>). I'll end this chapter with a quote by one of the judges that emphasizes some of the goals of SocialAction and perhaps the essence of creating beautiful visualizations:

Networks are best read if they are not only "technically accurate" and visually attractive but when they employ a type of rendering that creates a landscape. That creates a bridge for the uninitiated audience to cross into the field of expertise. Dataland travels have now become so enjoyable, they may soon appear as special fare destinations at a travel agency near you. Perer's visuals make that trip into the land of terror networks absurdly attractive. Having intellectual entertainment and visual pleasure with terrorism analysis is perhaps one way to diffuse the very essence of terror—by analyzing it without being terrified. And in the end it leads to a hopefully more rational dealing with it, which is the opposite of what terrorism is trying to instill.

—Ingo Günther,
Tokyo National University for Fine Arts & Music, Japan

References

- Freeman, Linton. 2000. "Visualizing Social Networks." *Journal of Social Structure*. <http://www.cmu.edu/joss/content/articles/volume1/Freeman.html>.
- Freeman, Linton. 2004. *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver, BC, Canada: Empirical Press.
- Hume, David. 1742. *Essays: Morale, Political, and Literary*.
- Moreno, J.L. 1934. *Who Shall Survive?* Washington, DC: Nervous and Mental Disease Publishing Co.
- Shneiderman, Ben. 1996. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In *Proceedings of the IEEE Symposium on Visual Languages*. Washington, DC: IEEE Computer Society.
- Tukey, John. 1977. *Exploratory Data Analysis*. Boston: Addison-Wesley.
- van Ham, Frank. 2004. "Interactive Visualization of Small World Graphs." In *Proceedings of the IEEE Symposium on Information Visualization*. Washington, DC: IEEE Computer Society.

