

Social Networks for End Users

Danyel Fisher

*University of California, Irvine
Survey Paper for Advancement to Candidacy
August, 2003*

1. Introduction

Social network analysis is a tool for understanding the shapes and contexts of group interactions. By mapping the relationships between groups of people—a map in which nodes are individuals and ties are the collected or measured relationships between them—analysts can develop both qualitative and quantitative descriptions of the group’s relationships (Wasserman and Faust 1994; Grannovetter 1973). Network analysis has been applied to a wide variety of organizations, as well as to other forms of social relations, and has found use in sociology, organizational behavior (as in Ahuja and Carley, 1998), and epidemiology (as in Morris and Kretzschmar, 1997), to name only a few. Recently, network analysis has also become better-known in more far-flung fields and the popular press; examples have been seen in physics (where it is a tool within complexity theory) and biology (Jeong, Tombor et al. 1999). The tools of network analysis, including identifying people central to networks, locating network “holes” (Burt 1992), and matching network structures, have been found to be broadly applicable.

A growing literature has begun to apply the analytical methods of the field to the computer- and social-networks in computer-supported collaborative work (CSCW). Wellman (1996), for example, argued that the patterns of intercommunication between members of electronic communities made up parts of a social network. CSCW has had a long-standing interest in the particular concerns of how people get their work done, and in the social structures that help them do that work. However, the majority of network analysis work has come from the perspective of an external observer, rather than from the perspective of individual members of the network. Even ego-centric network analyses (as described in Wellman’s 1993 review of the field) are largely reviewed and studied from without.

Top-down network analysis is a continuing theme within CSCW. Wellman (Wellman, Salaaf et al. 1996) highlights the connections between social networks and computer networks, pointing out that CSCW can be used to support a variety of network configurations, ranging from tight cliques (as in video teleconferences) to loose connections (as in email). It has been used in a variety of projects, ranging from Eveland and Bikson’s (1986) early work with email at the RAND corporation to recent projects examining the interrelationship between communication patterns and organizational structure (MacDonald 2003).

This discussion will examine the use of social network analysis in CSCW for end users. There is a particular challenge in this: CSCW, as a design field, is accustomed to examining projects from the user’s perspective, and working with users to design future iterations of software. Social network analysis, arising from sociology, is more accustomed to dealing with groups of people, from the outside. Finding places where network analysis directly addresses the problems of CSCW, then, is a particular challenge—and focusing on the end-user only limits the scope yet further.

It is reasonable to ask whether there are ways to introduce network analysis more tightly into the “socio-technical design circle” (as O’Day, Bobrow et al. 1996 put it). Can network analysis have a direct value to the members of the network, rather than only the indirect results of an analyst making structural recommendations and changes. Can, in other words, aspects of a network be

presented to a user, either directly or after some steps of algorithmic processing, and will the user be able to make sense of it?

Users might well be interested in information derived from their networks: the raw material for network analyses is often collected directly from users, and the material is often quite salient to them. A card-sort, for example, elicits broad clusters of “alters”—other members of the network—from a participant. Other forms of interview, including matrices be filled in by participants, or interaction observed by researchers (as in Gibson 1999) depend on collecting information by the participants. Yet they are often involved in only the data collection phases of the network analysis process: it is their responses that are used to generate the network. Participants, though, are left out of later stages.

One way to handle this might come from ego-centric network studies. In them, partial networks are collected from respondents without an attempt to connect them back to other respondents’ information. Instead, the respondent’s own understanding of his or her network becomes the object of study (Wellman 1993). The analysis from such an ego-centric study might, individually, be of interest to the respondent; however, such results are usually aggregated over large groups and studied collectively.

The question, though, of how relevant these networks can be for users can become more relevant in a CSCW context. While hand-prepared networks might not present new information, automatically-elicited networks often contain unfamiliar material. It is possible to algorithmically prepare and present networks to users based on automated collection from CSCW applications, such as email or chat applications.

In order to highlight the variety of types of network analysis prevalent in the field, this paper reads the notion of “network” broadly, emphasizing the notion of a group of alters connected to an actor, while disregarding some aspects of network topology. We consider a network generally as a set of users and the relationships between them; the network may be homogeneously composed of people, a bipartite graph of people and the artifacts around which they interact, or even more heterogeneous graphs.

Many technologies share underlying philosophies, if different implementations. In order to cover the scope of social networks within the field, it will be impossible to thoroughly cover all subfields in which networks are broached. Instead, we aim to approach the core insights of the different areas of research, choosing several projects from each that illustrate the general concepts at hand.

We constrain the discussion to places where end-users benefit from the network analysis fairly directly: while there certainly are plausible effects on lower-level employees from an analysis presented to a manager, we are more concerned about results that can be presented to users. At the most direct, a visualization showing an employee’s position in an organizational network would certainly qualify; we draw the opposite line at positions where a human must intercede to interpret the network, such as situations in which design decisions are rolled into a network analysis.

It is less important, however, how explicit the network must be. An ego-centric network, less its ties, becomes a contact list or an address book—a list of alters, possibly annotated with

information about the context of interaction. In order to manage a comprehensive survey of networks in CSCW, several variants of this list will be discussed. We begin with a discussion of *contact management*, understood as a form of egocentric network management. When a user manipulates a contact list—whether through a rolodex, a palm pilot, or a cell phone autodial list—they are working within one of the most prominent forms of network traces available. This is a restricted list—it loses much of the topology of the network, leaving only the immediate alters—but provides interesting information. Where the activity of members of a contact list becomes visible, the list becomes an *awareness system*.

Nor need the network necessarily be known to the user. Analysis on some sorts of networks is done on implicit relations, such as “is interested in similar topics”. A generalization of the social network concept, that of considering a graph with weighted similarity measures between every user and every other user, leads to meaningful visualizations and useful graphs at the heart of *collaborative filtering* research. Collaborative filtering is one aspect of *social navigation*, an attempt to derive information for users of a system based on collective actions.

Social navigation is an approach to network analysis in which activity in the network is analysed, then presented collectively as a supplement to information-seeking. The second part of this paper discusses methods of mining social networks based on *social navigation*. These methods typically include ways to cluster and analyze nodes of users—in the former case, examining their behavior; in the latter, examining their data. Last, this paper will explore various network *visualizations* that show allow users to visualize networks.

The various areas discussed—contact information, awareness, knowledge management—are all very broad; it would be infeasible to cover them all in great detail. Rather, this essay strives to select examples of particular interest, and then discusses the range of variation in other related systems.

1.1. A Framework for Discussing Network Research

In order to categorize and discuss the technologies, we ask several questions about them. These form rough axes to categorize and extract relevant features from the technologies.

- + Is the network *ego-centric*, or *top-down*? Ego-centric networks are of limited scope; they may look no further than the immediate set of alters. Top down networks involve a more thorough data collection, often reconciling several data sets in order to reconstruct a thorough graph.
- + Is the network *entered*, or *discovered*? In the former case, a system might solicit network information from users; in the latter, the network might be automatically extracted from online data sources.
- + Is the network *single-mode*, tying users directly to other users, or is it *multi-modal*, tying users to each other by way of shared artifacts?
- + Is the network *displayed* directly to users, or presented in a numerically *processed* form?
- + Are the members of the network *known to the user*, or are they maintained *anonymously*?

- + Are the members of the network the same as the *end-users* of the network? In general, this paper is concerned with systems where this holds; however, we will introduce others for contrast.

This questions will form a framework for discussing research on networks for end-users.

2. Ego-Centric Analysis within CSCW

The simplest form of a social network is an ego-centric network. Indeed, when people ordinarily attend a “networking” event or speak of their “social network”, they ordinarily mean their immediate contacts and friends-of-friends, and not those at greater distance. This manifests itself within the CSCW world as a contact list, or as an address book. A contact list is simply an ego-centric social network, without interconnections.

In this section, we will review general contact management systems. We will then discuss awareness systems, which we understand as contact management with presence information. Last, we extend awareness into social navigation.

2.1. Contact Management

From this perspective, Rolodexes, address books, instant message buddy lists, cell phone number lists (Berg, Taylor et al. 2003; Grinter and Eldridge 2003) and email contact lists (Ducheneaut and Bellotti 2001) are all personal views of networks. Address books are a traditional source and destination for contact information. For example, both cell phones and email programs store contact lists as a convenience to the user: an index into frequently-dialed numbers and often-sent messages. Personal information management (PIM) devices often synchronize with these address books to build a minimal social network, flattened into a small number of contexts: “work” and “personal,” for example.

To place contact management within the rest of the framework, most of these networks are ones that were *entered* by their maintainers; in fact, managing a contact list is a substantial amount of work. The members are all *known to the user* and non-anonymous, are *presented* directly back to the user. Last, there is, generally, only one user of the address book.

Where traditional address books—such as those stored by email programs or PIM—collected and managed contacts in a set of categories, some forms enhance these networks with activity or presence information. It is common for instant message tools to indicate other active users on the contact list (conventionally known as a “buddy list”), often providing some measure of their activity (Isaacs, Alan et al. 2002; Grinter and Palen 2002; Herbsleb, Atkins et al. 2002). Users leave “away messages” messages for their contacts that indicate their current activities or feelings; the system also indicates whether the buddy is logged into the system, and whether they are idle, showing when a user hasn’t been active on their computer for a time.

These network ties may be outgoing links only (as in AOL Instant Messenger’s—AIM’s-- implementation), or may be constrained to bidirectional links (as in Yahoo’s implementation, which prevents users from having someone on their buddy list who hasn’t approved them.)

A union of egocentric lists can be used to construct a full network; with the right collection tools, then, the underlying network can be constructed. BuddyZoo¹ collects the network of AIM Buddy Lists from users who send in their buddy lists; it then allows users to see visualizations of their buddy lists, see popularity measures, and count network distance between screen names.

Contact lists ordinarily provide no further information about a particular contact than these simple metrics. Some research projects, however, have explored further dimensions and attempt to articulate the network more thoroughly. AwareNex (Tang, de Bonte et al. 2000) provides additional information about where a user is located, and attempts to provide up-to-date information about how a user can be best contacted. However, some recent work has begun to suggest that these systems might be enhanced with additional activity awareness. Thus Berg, Taylor et al (2003) suggest that the list might be presented differently to users; dynamically sorted, animated, and clustered based on how a given participant is interacting with his friends. Berg's work also suggests a richer notion of activity; her work proposed a variety of ways in which the contacts might highlight their movements and actions.

Contact lists are not necessarily individual artifacts, unshared by other users. Grinter (2003) discusses teens sharing their phones to read each other's address books. One of their study participants reports that he would borrow others' phones and re-insert his name, irritated that he

had been removed. A well-filled address book was treated as a sign of social status; when a participant in their study would receive a new phone, they would often initialize it by copying an entire number list out of another's phone.

As the contact list begins to expand, categorization becomes difficult—and a categorized contact list bears strong resemblance to a card-sorted social network. Whittaker and Jones (Whittaker, Jones et al. 2002) discuss the difficulty their interviewees had

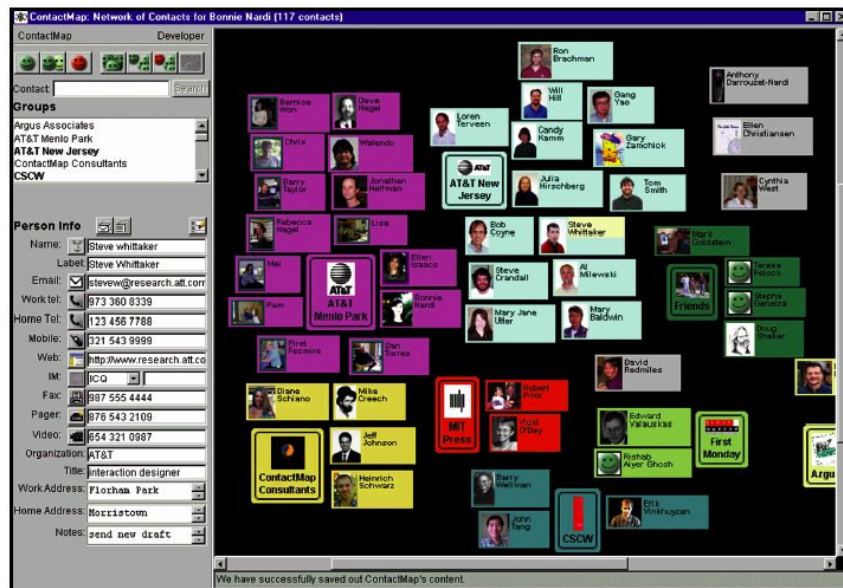


Figure 1. ContactMap. Note that the users, with their pictures, have been clustered into groups.

tracking the associations between their contacts, and remembering who they were in touch with. The ContactMap system (Nardi, Whittaker et al 2002) was built as a partial solution to this problem. It gets close to visualizing a network on the user's desktop as a way of managing contact information. It initially collects a list of names from an email record; it then allows its

¹ <http://www.buddyzoo.com>

user to cluster the names as they will into explicit groups. This then becomes an enhanced information manager: it is possible to view contact information for each person or group, and to send email messages to the entire group. The ContactMap also functions as an index into the past correspondence, thus unifying the notion of the visible map with the actual interaction. It does not, however, match this access to any form of awareness or activity. The ContactMap, like the buddy list and the phone quick-dial list, is a fixed one; names can be manually added to it, but it does not adaptively modify its lists based on external changes.

2.2. Awareness and Presence

A contact list with active information becomes an awareness tool. Awareness, a popular theme within CSCW, was described by Dourish and Bly (1992). They describe awareness as “knowing who is ‘around’, what activities are occurring, who is talking with whom, and provides a view of one another in the daily work environments.” This definition does not require computer support

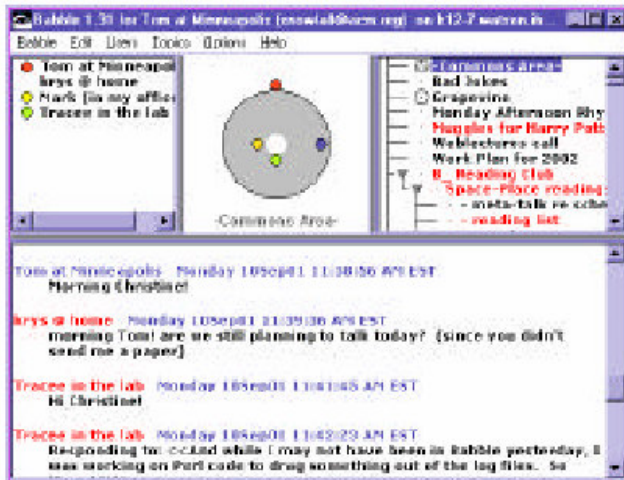


Figure 2: Babble

nor networks; an intern in the corner cubicle has good awareness of who passes into the building. Heath and Luff (1992) describe the subtle techniques used by a pair of coworkers communicating about emergency management to each other and to their various co-workers on the London Underground. CSCW systems support awareness by relaying subtle information about what remote co-workers are doing.

Within the network framework, awareness systems match contact lists: ego-centric connections, entered by end users, allow other known users to be displayed on screen.

In the CSCW world, awareness tools make it

possible for users to learn more information about the activities of a group of correspondents. Schmidt (2002) attempts to arrive at a taxonomy of awareness, emphasizing both active forms (such as Heath and Luff’s emergency team telegraphing important facts to each other) and passive forms (in which users gain so-called peripheral awareness with tools that suggest, rather than show, the activity of their partners). Emphasizing the work practice that leads to awareness, he writes that “‘Awareness’ is not the product of passively acquired ‘information’ but is a characterization of some highly active and highly skilled practices.” Portholes (Dourish and Bly 1992), for example, is a peripheral system that allows users to glance into an office some distance away. Thus, a casual glance would inform a user whether their collaborators are busy in a meeting—or idly tapping their thumbs, and that it might then be a good time to reach them. The “skilled practice,” in this case, is then one of interpreting availability, activity, and interest from the video images.

Gutwin and Greenberg (1998, 1999) point out that peripheral awareness allows users to communicate more efficiently and to get more context about what other users are doing. Their

awareness system gave users the ability to have rough location and activity information on their collaborators in one experimental condition; in another condition, users had only their own information. Users completed their tasks more quickly with awareness, and were more able to communicate with their collaborators.

Further information can come not only from who is available, but who is interacting with the system. The NyNex Portholes system (Lee, Girgensohn et al. 1997), for example, provided lists of both “people in my portholes viewer” (which allowed a user to see all users who were connected to the system) and an overview of “who is looking at me.” This seems to be more fair, by a sense of reciprocal awareness, and their users liked knowing who was watching them. From a network point of view, the NyNex system then allows users to perceive incoming links as well as outgoing ones—they can be aware of how their network is shaped.

Ackerman and Starr’s Espresso (1995) provides a global level of information across the network. In their system, users can see not only the messages that are sent around in a series of chat rooms, but can see some global statistics of the system, such as how many messages are being sent and who is sending them. This allows users to anticipate which sections of the system are active enough to be worth watching. These activity cues allow more active participation.

Similarly, Babble (Bradner, Kellog et al. 1999) combines social activity indication with a simple visualization across a limited team. Based on a series of principles of transparency and mutual visibility (Erickson, Smith et al. 1999; Erickson 2003), the system allows team members to know who is contributing to a persistent conversation. The system splits conversation into a series of themed “topics”, and so users can also tell which conversations are have current updates. Teams tend to set up independent rooms for interaction; similarly, group members affiliated with particular interests often persist within a single topic. Babble, then, defines an affinity network from person to topic, and annotates that network with information about which people are most and least active, and which topics they are active within.

A particularly relevant form of workspace awareness is the “Anchored Conversations” system (Churchill, Trevor et al. 2000). With that system, users could attach fragments of conversation, such as instant message transcripts, to documents. Documents that had been so annotated would then carry with them information about who was interested in them, as editors and commentators. Those annotations describe, in a very practical sense, a notion of who is interested in a particular topic and thus highlight an affinity network.

2.3. Social Navigation and Collaborative Filtering through Awareness

One closely-related cousin of awareness is the field of social navigation (Shardanand and Maes 1995; Hook, Benyon et al. 2003). In a social navigation system, the actions of a group of people (such as book purchasers, or movie reviewers, or web surfers) are collected and used to provide insight to later people doing the same things. Users, then, have a degree of awareness of the actions of other users in the systems. In Wexelblat’s Footprints (2003), for example, links on web sites begin to show signs of change as users read links more often. The system monitors the paths that surfers take through the web and feeds them back to later users, constructing recommended paths and popular ways of finding information. In related systems, documents can begin to show

their own age as users work with them, allowing later readers and editors to notice popular, important, and heavily-modified parts (Hill, Hollan et al. 1992).

Within the framework, social navigation continues to present an ego-centric view. The network is often a discovered and anonymous one, as users wander between sites seeing the accumulated tracks of previous users. The network is bi-modal; users are joined to the documents they share, and don't see each other directly.

A related technology, collaborative filtering (as Amento, Terveen et al. 2003, and collected at Resnick and Varian 1997) is based on the principle that similar users are likely to enjoy similar things. Users' preferences placed into a large, multidimensional space, and are clustered automatically based on proximity within the space. Nearby users are given access to collective information about each others' preferences. Amazon can, at a user level, describe the notion of "people like you" who recommend particular books or ideas. Collaborative filtering, then, defines an implicit network across the set of all participating users. As we will see, this sort of network can be mined: users at far points can be seen as isolates; users further within the network can be seen as central.

Domingo (2001) considers the implicit network from collaboration systems to be a social network, and analyzes it accordingly, searching for communities that may be interested in a particular marketing tactic. The network ties are those of "influence," and a user is said to influence another if the recommendation of the one applies to the other. This is, he suggests, in some ways a purer network than other forms; a collaborative filtering system that compares strangers can be fairly sure that the users are joined over no other media. This network, then, is generated independently from friendship or geographical networks. It is no longer quite a "social" network, although it is understood as a way to interpret and locate influential people.

3. Mining Social Networks

Results from the automatic analysis of a social network can be fed back to users. One genre that has produced a series of results for end users comes from the idea of mining network structure to provide information about the underlying data. Some of these networks, drawn from web pages, are only implicitly social, but the results that can be mined from them highlight ways that results can be presented. For example, pages that have structurally-relevant positions within the networks can be placed higher in a list of returned values.

Page and Brin (et al. 1998) introduced the PageRank system, which has been turned into a popular search engine, Google. With the PageRank algorithm, web pages are assigned values based on their location within the network: well-ranked sites are defined recursively as those that are referred to by other well-ranked sites. A related algorithm, Kleinberg's (1998) hubs-and-authorities model also highlights good sources of information. Kleinberg's model considers a network to be divided into two groups, "hubs" and "authorities." Again, the definition is a structural one: a good hub is one that points to good authorities; while a good authority is one that is pointed to by good hubs. Users of the algorithm have developed prototype tools to find sections of the web on particular themes. These can lead to interfaces that highlight navigation cues through the space of web pages. One application of these technologies is the TopicShop (Amento, Terveen et al. 2003) system. It uses so-called "social data mining" to sort

out important pages from large lists, and then allows users to gather and cluster the data in ways that are particularly relevant to them.

Social networks may also be derived from other socially-driven web pages, such as weblogs. In a broadly non-academic trend, a series of weblogs have ranked each other using PageRank-like algorithms², automatically rated each others' importance in the "blogosphere" by mutual links, and displayed the results publicly as indices into the weblogs. Weblog authors can, and do, check these ratings to see how their weblog ranks; web surfers check the rankings as ways to find their own reputation.

Other systems search weblog links to find a consensus on important pages. If enough weblogs point to a particular site, than it must be of some importance—something that has captured weblog community's collective imagination. The network this searches is a bipartite one (the "hubs" are all weblogs; the "authorities" are the pages they search for, and the hubs aren't ranked) but the interface does allow users to know who is important³.

3.1. *Expertise Location*

One popular use of network information is for expertise location. Based on the observation that the expertise in many organizations is broad enough that the organization does not know what it knows, expertise location attempts to describe and locate expertise within an organization. One mechanism for locating this knowledge is to produce a directory of employees and their areas of expertise. These directories can be assembled in a variety of ways; however, they still retain the social challenges of fostering cooperation between employees. Network analysis attempts to help with both of these problems, by examining both what communities the employees are connected to, and by examining what knowledgeable employees are close enough to be a friend-of-a-friend. In an example of the latter, Tyler (2003) uses email network to identify close-knit interest communities at a research lab, separating groups on the basis of their "betweenness." The communities were found by examining pairs of users who emailed each other frequently.

Expertise location uses global networks, not ego-centric ones, based on information accumulated across an organization. The information can be accumulated through automatic means (as in Tyler), or manually. It handles closed organizations; even if the members may not be familiar with each other, they can find out information about each other through business directories and organizational charts.

Some uses of expertise location have been very explicit about their use of networks. For example, Busch (Busch, Richards et al. 2001) uses social network analysis to try to detect

² Information about the Blogosphere Ecosystem can be found at

<http://www.myelin.co.nz/ecosystem/>
<http://www.truthlaidbear.com/EcoFAQ.php>

³ Various weblog rankings can be found at

<http://www.daypop.com>
<http://www.popdex.com>
<http://blogdex.media.mit.edu>

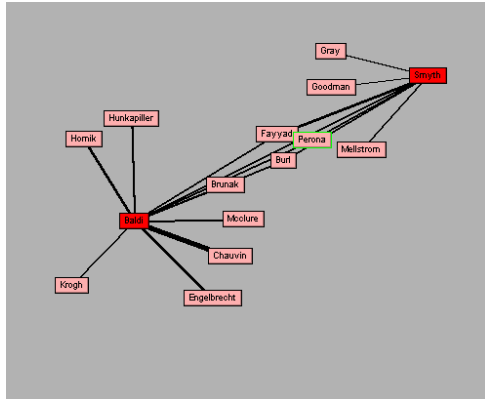


Figure 3: ReferralWeb

patterns of tacit knowledge. Their network is collected from a name-generator instrument; their analysis seeks out top collaborators and the ways that they collaborate with each other. They then generate diagrams of networked collaboration and try to find crucial players within the organization. In the study described in the paper, they find that the group secretary forms a crucial gate-keeping role between two dense networks within the office. Busch presents the network as an analytical tool, but does not present a tool for end-users to work with the network themselves.

Implicit networks—derived from socially published information—can be tied to more explicit user information in order to produce results that can be meaningful to users. ReferralWeb (Kautz, Selman et al. 1997) follows traditional network co-citation models, but allows end-users to browse between different members of the network. Each author is associated with past co-authors for papers (and other names that may co-occurred on web pages), as well as a list of keywords. ReferralWeb then makes an explicit and visible social network from this implicit information. The ReferralWeb interface allows users to view portions of the network, and thus can find important authors in the field, locate shortest paths between important authors, and search for communities of knowledge—groups of authors who work on similar topics and are closely connected to each other. Since the ReferralWeb system uses publicly-available information and posts results to public websites, the users of the ReferralWeb network are not the same as the participants in the network.

The IKNOW system (Contractor, Zink et al.) also captures social information to display networks to users. It collects information about areas of knowledge and network ties from within a restricted community of users (who either enter their data manually, or have it scanned off of home pages). It then presents information in the form of several different networks, based on common interests, common referrals (to external web sites or other users), and common vocabulary. It presents the information back to users as networks, and, like ReferralWeb, displays networks as well as network measures to the users

3.2. Reputation Systems

A number of systems use network analysis as a way to synthesize a “reputation” figure. In some systems, a well-thought-of person has more weight in rating others; in others, recommendations are only valid within a ‘network of trust’, the people who have been rated, and who rate those raters. The former is exemplified by systems such as Slashdot’s “karma”⁴ system. In the framework presented above, reputation systems accumulate information from a network of connections. They use information entered by users, and largely do not explicitly display the networks they use to calculate their connections. Members of the network know each other, and are the end users of the system.

⁴ <http://slashdot.org/moderation.shtml>

That genre of reputation system was tested by a series of experiments at Microsoft Research. Jensen, Davis and Farnham (2002) extracted network information to present reputation information to users. They then were able to test the system on a series of users, asking them to predict how likely they would be to interact with certain alters based on their online reputation. Their study found that reporting similar interests was a strong determinant of a user's desire to enter a chat, while implicit location on a network (as derived from network ties) was less useful.

The "Advogato" system⁵ evaluates the abilities of open source programmers to work with each other by extending what the system refers to as a "web of trust." Programmers who have witnessed other programmers' skills are able to rate them on a scale ranging from unskilled to expert. Users of the system—the programmers—can then view the ratings of other programmers with whom they have not personally collaborated by being shown a compound path rating. (Thus, if I rate a programmer as "apprentice", and he rates a collaborator of his as "expert", this review will have lower credibility in the system than the equivalent rating by a programmer who I have rated well.)

A similar network is embodied in the key exchange system of private-key infrastructure (PKI) systems. Within the secured world of PKI networks, all information must be transferred either by hand, or encrypted with a known key. These keys, though, are information themselves. Once a user has collected a small number of keys by hand that she trusts, she may collect other keys safely only by getting them from trusted users. These new keys can then also be trusted. The recommendations and degrees of trust, then, implicitly imposes a network of connections⁶. In practice, the space is structured as a strongly-connected hub; a small number of trusted repositories and users are well-rated and act as certifying authorities for most other users (as shown by Guardiola, Guimera et al. 2002).

Other sorts of network-based reputation input and output systems are available. The LambdaMOO Cobot (Isbell, Kearns et al. 2000) wandered about its online world, collecting "social statistics": what users interacted with other users, and how much. This information was relayed back to the users through a social player, the Cobot, who could answer simple queries such as "Who is like me"? Cobot would respond with messages like

cobot [to Gabaldon] : Here are your relationships with Sparklebug ... Sparklebug is ranked 7 on your list of playmates. You are ranked 19 on Sparklebug's list. ...

(Isbell, Kearns et al. 2000)

Although the Cobot collected pair-wise ties across a number of different modalities (who hugs whom, who talks to whom, and so on)—perfect for network data—the information was not explicitly grouped into network information.

⁵ <http://www.advogato.org/trust-metric.html>

⁶ <http://bcn.boulder.co.us/~neal/pgpstat/>
<http://dtype.org/keyanalyze/explanation.php>

A related series of online tools, among them “Friendster”⁷ and “SixDegrees,”⁸ present networks directly to users. Users of the software select a list of friends by mutual ties, and then can view their immediate friends, the friends of their friends, and so on. Both services overlay a dating and connection service over the network; users can search their friends-of-friends (to within three or four removes) to find dates or new friends, and can arrange introductions through the sometimes-tenuous chains that connect them. There is, however, no built-in structural analysis: users are not automatically rated based on their friends’ connections, for example.

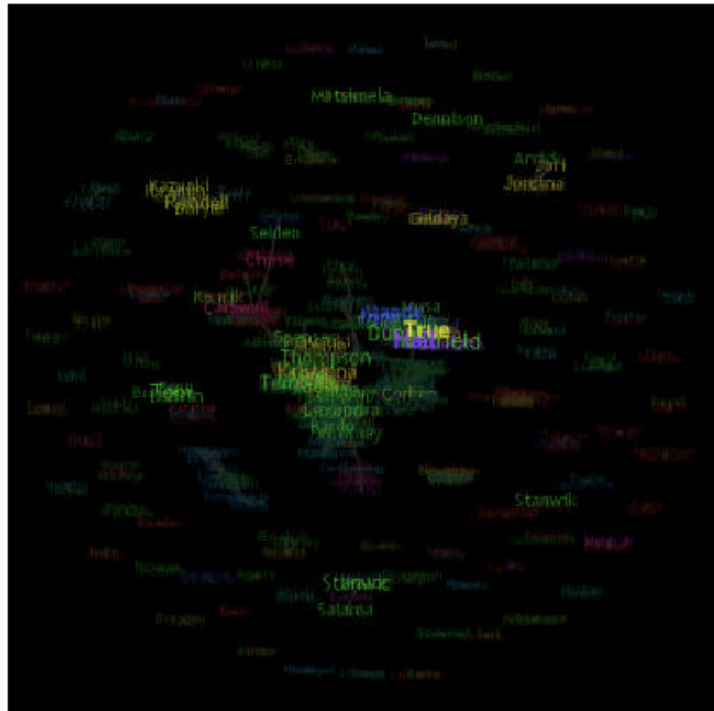


Figure 4: Social Network Fragments

4. Social Visualizations

The last set of uses for network information for end users have come from visualization areas. Social networks, seen by users, become in themselves interesting: users find their own networks interesting; networks serve as a tool for recall, for storytelling, and for introspection. Networks (as shown with ContactMap) can be a tool for sorting and organizing information.

These networks are both ego-centric and top-down, and are both entered explicitly by users and discovered. The information in them, though, is displayed directly to users; the members are known to each other. All networks discussed here are meant to be used by the people who are depicted: the visualizations are meant for end-users.

Most of these systems describe visualizations of collaborative social spaces: spaces where users interact with each others. They then try to visualize some aspects of the conversations within these spaces, understood as parts of a network of conversation. We begin with email-based systems, and then discuss systems that work in online conversation spaces.

4.1. *Email Visualizations*

Farnham (2002) has presented a variety of systems that develop social networks out of email records. One project centered on the implicit networks that can be derived from co-membership on email lists, and visualized networks based on that information. The networks are portrayed as

⁷ <http://www.friendster.com/>

⁸ <http://www.sixdegrees.com/>

traditional social networks; nodes are connected to each other explicitly. In addition, the project extracted a list of “closest co-workers,” and annotated corporate information pages with the list of close connections. A second project highlighted the networks that can be derived from an email database; using data extracted from contact lists in the inbox, the system displayed connected networks of users. This view roughly clustered the list of contacts into groups, and allowed the user to change various thresholds of connection in order to make the groups more or less connected.

A series of projects from MIT Media Lab have explored different aspects of conversational network visualization. Social Network Fragments (boyd 2002 pg. 75-94) illustrates a single users’ email spool, visualized as a social network. The network, assembled through a combination of automatic collection and manual labeling, draws several different sorts of ties, ranging from “is weakly aware of” to “trusts.” In this case, “weakly aware” is a general tie for pairs of users who have both been carbon-copied on a single message. The most trusting tie, in contrast, is reserved for those who have been placed on a “BCC” (blind carbon-copy) list; those are the contacts who can be given non-reciprocal access to the names of others.

The “PostHistory” project⁹ visualizes activity on a contact list—specifically, in an email archive—but declines to explicitly show network interactions.

4.2. *Visualizing Interactive Spaces*

While mapping email networks is a study, often ego-centric, of a localized set of interactions, visualization projects have also striven to show the size of interactive spaces. A number of projects—collected, for example, at the Atlas of Cyberspace¹⁰ (Dodge and Kitchin 2000)—have attempted to illustrate aspects of the topography of cyberspaces. Here, we particularly emphasize online conversation spaces—interactive spaces that have a particular emphasis interaction.

“PeopleGarden” (Xiong and Donath 1999) attempts to create glyphs of users as individual flowers; their shared interaction patterns are seen as a garden. Each flower-shaped glyph represents the time since the user last posted to the group, the number of messages to which the user received responses, and the amount of time a user has been in the group.

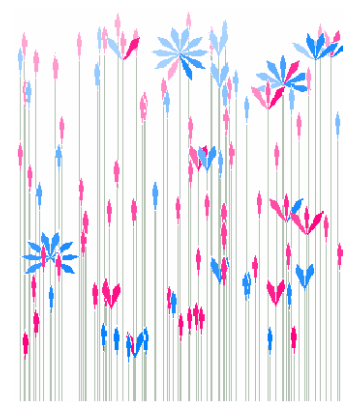


Figure 5: PeopleGarden



Figure 6: VisualWho

⁹ Viegas, Fernanda. PostHistory.

<http://web.media.mit.edu/~fviegas/posthistory/>

¹⁰ The Atlas of Cyberspace. (<http://www.atlasofcyberspace.com>)

“VisualWho” (Donath 1995) shows individual participation on mailing lists—much as does Farnham’s visualization—but doesn’t attempt to find linkages between individuals. Instead, it broadly pulls users into regions of a display based on their affiliation with mailing lists, and their affiliation with individuals on mailing lists. The “VisualWho” display is sensitive to some forms of awareness; it dims the display of members who are not logged in, allowing the user to locate those who are active.

Other views of conversational networks include Sack’s Conversation Map (2003), which visualizes Usenet newsgroups. This tool shows the implicit networks derived from which users have responded to other users. They are meant both as tools for analysts, and eventually to be placed in the hands of end-users—members of the newsgroup can view their own participation, and can begin to understand the interactions within the newsgroup. Similarly, Ducheneaut’s Open Source Visualizer (Ducheneaut, 2003) shows Open Source mailing list communication. Group members can review their own contributions to the mailing list, can examine important conversation themes, and can understand how the project breaks up into implicit teams.

Smith’s “Netscan”¹¹ (2003) is a real-time web service that also dispatches “newsgroup report cards.” While it does not show inter-user network ties, it does highlight relevant network statistics about people interacting. It provides information, per user and per group, on how many people responded to messages, how many groups they posted to, and other similar statistics. “Netscan” also contains a visualization based on cross-posts; pairs of groups that receive many messages posted to both groups are thought to be more closely related, and get stronger ties.



Figure 7: Conversation Map (l) and Netscan (r)

¹¹ <http://netscan.research.microsoft.com>

AT&T's Clan Graph (Terveen, Hill et al. 1999) attempts to present a hybrid between displaying a network directly, and collecting information based on the network. The system organizes web sites based on their conceptual (and geographical) adjacency. Interlinked web pages are sorted into sites (based on common URL) and combined into clusters (based on network interconnections). The system then presents back to users a network of web pages; as the user moves from one page to another, the system presents concentric circles of pages relevant to their initial query.

5. Conclusion

Social networks are frequently used as an analysis technique, but are less frequently directed toward end users. The field of CSCW seems like a fruitful place to apply network techniques to end-user concerns: designers can develop applications that collect networks from user interaction, process them into results, and feed them directly back to users. While this is promising, the few forays into social network analysis have not yet fully caught on as an end-user design technique. This is partially result of cultural differences: social networks is a technique from analytical fields, and particularly sociology; the tracks of CSCW that build systems are more accustomed to seeing themselves as design-based fields.

There are, however, many techniques within CSCW that use information analogous to social network analysis: they instead measure the contact lists of users (and collect an ego-centric perspective); they examine a user's set of choices in a recommender system. Many of these techniques could be strengthened by a more thorough examination of network analysis. This paper illustrates a number of opportunities for future research: many currently-existing systems could be enhanced by either understanding and using network information, or by presenting it to end users.

This paper has highlighted a series of features of networks that are worth recapitulating. This paper has examined the value of automatically-generated networks, and has found a series of systems that use these successfully. These networks are comprehensible by users, and their results are useful. We have also highlighted awareness, and with it the dynamics of networks that change with current activity. Last, we have taken a particular interest in tapping ego-centric sources. While multi-modal networks seem to be rarer, the heterogeneous network of links between people and the documents they work with are implicit themes in systems like IKNOW and other network view programs.

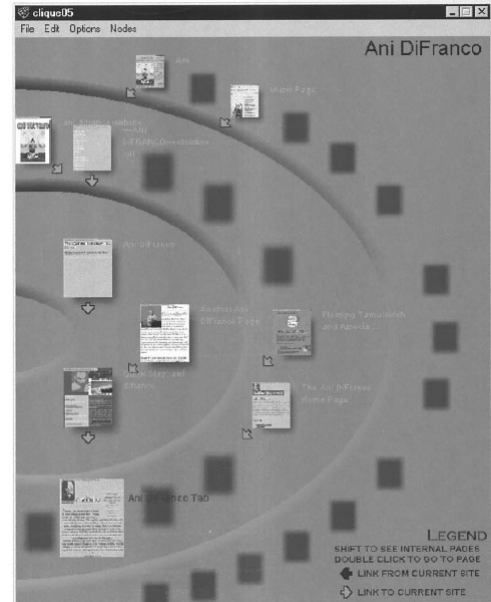


Figure 8: ClanGraph

References

- Ackerman, M. S. and B. Starr (1995). Social Activity Indicators: Interface Components for CSCW Systems. Proceedings of the 1995 ACM Symposium on User Interface Software and Technology (UIST '95).
- Ahuja, M. and Carley, K. M. (1999). "Network Structure in Virtual Organizations", joint issue of Organization Science, Nov./Dec and Journal of Computer-Mediated Communication, 3/4 1998.
- Amento, B., L. Terveen, et al. (2003). From PHOAKS to TopicShop: Experiments in Social Data Mining. From Usenet to CoWebs: Interacting with Social Information Spaces. C. Lueg and D. Fisher. London, Springer.
- Amento, B., L. Terveen, et al. (2003). "Experiments in Social Data Mining: The TopicShop System." ACM Transactions on Computer-Human Interaction **10**(1): 54-85.
- Berg, S., A. S. Taylor, et al. (2003). Mobile phones for the next generation: device designs for teenagers. Proceedings of the 2003 ACM Conference on Human Computer Interaction (CHI '03), Ft. Lauderdale, FL.
- boyd, d. (2002). Faceted Id/entity: Managing representation in a digital world. Media Lab, School of Architecture and Planning, Massachusetts Institute of Technology: 118.
- Bradner, E., W. A. Kellog, et al. (1999). The Adoption and Use of 'BABBLE': A Field Study of Chat in the Workplace. Proceedings of the European Conference on Computer-Supported Collaborative Work (ECSCW '99), Copenhagen.
- Burt, R. S. (1992). Structural Holes: The social structure of competition. Cambridge, MA, Harvard University Press.
- Busch, P. A., D. Richards, et al. (2001). Visual Mapping of Articulable Tacit Knowledge. Australian Symposium on Information Visualisation, Sydney, Australian Computer Society.
- Churchill, E., J. Trevor, et al. (2000). Anchored Conversations: Chatting in the Context of a Document. Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems (CHI '00).
- Contractor, N., D. Zink, et al. (1998). "IKNOW: A tool to assist and study the creation, maintenance, and dissolution of knowledge networks." Community Computing and Support Systems: 201-217.
- Dodge, M. and R. Kitchin. (2000). Mapping Cyberspace. New York: Routledge.
- Domingos, P. and M. Richardson (2001). Mining the Network Value of Customers. Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, ACM Press.
- Donath, J. (1995). Visual Who: animating the affinities and activities of an electronic community. Proceedings of the third ACM international conference on Multimedia, San

- Francisco, California, ACM Press.
- Dourish, P. and S. Bly (1992). Portholes: supporting awareness in a distributed work group. Proceedings of the 1992 ACM Conference on Human Factors in Computing Systems (CHI '92), Monterey, California, New York: ACM Press.
- Ducheneaut, N. and V. Bellotti (2001). E-Mail as Habitat. Interactions. **8**: 30-38.
- Ducheneaut, N. (2003). "The Sociotechnical Construction of Open Source Software." Unpublished PhD Thesis, UC Berkeley School of Information Management and Systems.
- Isaacs, E., W. Alan, et al. (2002). The character, functions, and styles of instant messaging in the workplace. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, ACM Press: 11--20.
- Erickson, T. (2003). Designing Visualizations of Social Activity: Six Claims. Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI '03), Ft. Lauderdale, Florida, ACM Press.
- Erickson, T., D. Smith, et al. (1999). Socially translucent systems: Social proxies, persistent conversation, and the design of "Babble." Proceedings of the 1999 ACM Conference on Human Factors in Computing (CHI '99).
- Eveland, J. D. and T. K. Bikson (1986). "Evolving electronic communication networks: an empirical assessment." Proceedings of the 1986 ACM Conference on Computer Supported Cooperative Work 3: 91-101.
- Farnham, S., A. Turski, et al. (2002). MSR Connections: Exploring Who Knows Whom through Social Networks. Submitted to CSCW 2002.
- Gibson, D. R. (1999). Taking Turns and Talking Ties: Conversational Sequences in Business Meetings. Sociology Department. New York, Columbia University: 222.
- Grannovetter, M. (1973). "The strength of weak ties." American Journal of Sociology 78: 1360-1380.
- Grinter, R. and M. Eldridge (2003). Design for the socially mobile: Wan2tlk?: everyday text messaging. Proceedings of the 2003 ACM Conference on Human factors in Computing Systems.
- Grinter, R. and L. Palen (2002). Instant messaging in teen life. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, ACM Press: 21--30.
- Guardiola, X., R. Guimera, et al. (2002). "Macro- and Micro-structure of Trust Networks."
- Gutwin, C. and S. Greenberg (1998). Effects of Awareness Support on Groupware Usability. Proceedings of the 1998 ACM Conference on Human Factors in Computing Systems, New York: ACM Press.
- Gutwin, C. and S. Greenberg (1999). "The effects of workspace awareness support on the usability of real-time distributed groupware." ACM Transactions on Computer-Human Interaction **6**(2): 243-281.

- Heath, C. and P. Luff (1992). "Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms." Computer Supported Cooperative Work: The Journal of Collaborative Computing. 1(1-2): 69-94.
- Herbsleb, J., D. L. Atkins, et al. (2002). Introducing instant messaging and chat in the workplace. Proceedings of the 2002 ACM Conference on Human Factors in Computing Systems.
- Hill, W., J. D. Hollan, et al. (1992). Edit Wear and Read Wear: Text and Hypertext. Proceedings of the 1992 ACM Conference on Human Factors in Computing Systems (CHI '92), New York: ACM Press.
- Hook, K., D. Benyon, et al., Eds. (2003). Designing Social Information Spaces: The Social Navigation Approach. Readings in CSCW. London, Springer.
- Isbell, C. L., M. Kearns, et al. (2000). "Cobot in LambdaMOO: A Social Statistics Member." AAAI '00.
- Jeong, H., B. Tombor, et al. (2000). "The large-scale organization of metabolic networks." Nature. 407. 5 October, 2000. 651-654.
- Jensen, C., J. Davis, et al. (2002). Finding Others Online: Reputation Systems for Social Cyberspaces. Proceedings of the 2002 ACM Conference on Human factors in Computing Systems, Minneapolis, Minnesota, ACM.
- Kautz, H., B. Selman, et al. (1997). "Referral Web: combining social networks and collaborative filtering." Communications of the ACM 40(3).
- Kleinberg, J. M. (1998). Authoritative Sources in a Hyperlinked Environment. Proceedings of the 1998 ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, ACM Press.
- Lee, A., A. Girgensohn, et al. (1997). NYNEX Portholes: Initial User Reactions and Redesign Implications. Proceedings of the 1997 ACM Conference on Supporting Group Work, Phoenix, AZ.
- MacDonald, D. W. (2003). Recommending collaboration with social networks: a comparative evaluation. Proceedings of the 2003 ACM Conference on Human factors in Computing Systems.
- Morris, M. and Kretzschmar, M. (1997) "Concurrent partnerships and the spread of HIV." AIDS. 11(6):641-648.
- Nardi, B., S. Whittaker, et al. (2002). "ContactMap: Integrating Communication and Information Through Visualizing Personal Social Networks." Communications of the ACM. 45(4): 89-95.
- O'Day, V. L., D. G. Bobrow, et al. (1996). The Socio-Technical Design Circle. Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work (CSCW '96), Cambridge, Mass., ACM Press.
- Page, L., S. Brin, et al. (1998). The PageRank Citation Ranking: Bringing Order to the Web, Stanford University.

- Resnick, P. and H. R. Varian (1997). "Recommender Systems." Communications of the ACM **40**(3): 56-58.
- Sack, W. (2003). Conversation Map: A Content-based Usenet Newsgroup Browser. From Usenet to CoWebs: Interacing with Social Information Spaces. C. Lueg and D. Fisher. London, Springer.
- Schmidt, K. (2002). "The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW'." Computer Supported Cooperative Work: The Journal of Collaborative Computing. **11**: 285-298.
- Shardanand, U. and P. Maes (1995). "Social Information Filtering: Algorithms for Automating "Word of Mouth"." Proceedings of the 1995 ACM Conference on Human Factors in Computer Systems (CHI '95).
- Smith, M. (2003). Measures and Maps of Usenet. From Usenet to CoWebs: Interacing with Social Information Spaces. C. Lueg and D. Fisher. London, Springer.
- Tang, J. C., A. de Bonte, et al. (2000). Instant messaging: products meet workplace users. Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, ACM Press: 363.
- Terveen, L., W. Hill, et al. (1999). "Constructing, Organizing, and Visualizing Collections of Topically Related Web Resources." ACM Transactions on Computer-Human Interaction **6**(1): 67-94.
- Tyler, J., D. M. Wilkinson, et al. (2003). "Email as Spectroscopy: Automated Discovery of Community Structure within Organization." (Draft) Communities and Technologies.
- Wasserman, S. and K. Faust (1994). Social Network Analysis. Cambridge, Cambridge University Press.
- Wellman, B. (1993). "An Egocentric Network Tale." Social Networks **17**(2): 423-436.
- Wellman, B., J. Salaaf, et al. (1996). "Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community." American Review of Sociology **22**: 213-238.
- Wexelblat, A. (2003). Results from the Footprints Project. Designing information spaces: the social navigation approach. K. Hook, D. Benyon and A. J. Munro. London, Springer.
- Whittaker, S., Q. Jones, et al. (2002). Contact management: identifying contacts to support long term communication. Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work.
- Xiong, R. and J. Donath (1999). PeopleGarden: Creating Data Portraits for Users. Proceedings of the 12th annual ACM symposium on User interface software and technology, Asheville, NC., ACM.