Capturing Team Dynamics Through Temporal Social Surfaces

Peter A. Gloor
Center for Digital Strategies, Tuck School of Business, Dartmouth, Hanover NH
Center for Coordination Science, MIT, Cambridge MA
pgloor@mit.edu

Abstract

This paper introduces Temporal Social Surfaces. Temporal Social Surfaces show the dynamic evolution of social relationships in groups. Changes in measures of social network analysis such as betweenness centrality are plotted as a curve for each actor. The curves for all actors are combined in a single picture. Smoothing the resulting surface leads to a single-picture overview of social dynamics of a group. Temporal Social Surfaces are computed using two variants of a "sliding time window" algorithm with TeCFlow (Temporal Communication Flow Visualizer). TeCFlow is a system for creating movies of temporal information such as e-mail archives, Web logs, and Web search results. TeCFlow takes as input time series of records, and generates an interactive movie of the changes in the resulting directed graph.

1. Introduction

In a weather-forecasting system, weather patterns consisting of time series of recently collected data are fed into a computer model to predict the weather patterns of the next five to fifteen days. In our project we are aiming to do the same for the evolution of interaction patterns in networks. As with weather patterns used to predict sunshine and thunderstorms, communication flows allow for predicting positive and negative developments in groups of people. Temporal visualizations of social interactions can be extremely valuable as an early warning system, showing high-pressure systems, impending storms, and other relationships in groups that are difficult to anticipate through other means. By analyzing and aligning business processes and knowledge flow, organizations get a unique opportunity to increase the productivity of knowledge workers through greater creativity, efficiency, and quality.

We have been developing the TeCFLow environment for the visual identification and analysis of the dynamics of communication in virtual spaces for the last three years [11,12]. Similarly to a weather map, our tool plays back an interactive movie depicting the interaction between actors in a network based on "communication acts" stored in their communication archive. We have applied our tool to the dynamic analysis of social networks based on e-mail archives and the analysis of changes over time in the link structure of Weblogs and other Web sites [12]. By comparing dynamic interaction flows with the output produced by virtual teams, we aim to identify characteristic dynamic communication patterns of high-performing groups of knowledge workers based on the work done, among others, by [6,7].

2. Related Work

There has been substantial research dedicated to computing and visualizing social networks as directed graphs or adjacency matrices. AGNI [23], ucinet [3], pajek [2], KrackPlot [13], VisOne [5], Agora [15] and NetMiner [19] are some of many tools for social network analysis (SNA). Zoomgraph [1] is a recent system for social network analysis that combines graph visualization with a built-in programming language for graph manipulation and analysis. Sonia [17] is one of the few tools that temporally visualizes social networks by creating movies of animated graph structures. McGrath and Blythe [16] have shown that motion has a positive effect on the accuracy of viewers' perceptions of change in status in a social network. In earlier studies Gloor [10] and Lawrence, Badre, and Stasko [14] investigated the use of animation for the understanding of algorithms. PieSpy visualizes the evolution of social networks of chat users over time [18].

3. Design and Architecture of TeCFlow

TeCFlow takes as input communication archives and automatically generates static and dynamic visualizations of the communication networks. The dynamic visualization consists of an interactive movie showing the evolution over time of the communication network within the group. Active relationships are displayed in a sliding time window, with inactive relationships decaying over time. TeCFlow also calculates and plots the evolution of group betweenness centrality and density [24] over time to discover interesting events in the lifetime of a virtual team.

Actor betweenness centrality is the number of geodesic paths (shortest paths) that pass through this node. We count the number of "times" that any node needs to access this given node to reach any other node in the network by the shortest path. More precisely, if g_{jk} is the number of geodesic paths between j and k and g_{jik} is the number of paths between j and k that pass through i, then g_{jik} / g_{jk} is the proportion of geodesic paths between j and k that pass through i. The sum BC_i = g_{jik} / g_{jk} for all j,k pairs is the betweenness centrality for node i. Normalized betweenness centrality is obtained by dividing simple betweenness by its maximum value. We use Freeman's index [24] for quantifying the overall level of betweenness in the set of actors BC_{group}, which summarizes the actor normalized betweenness indices:

$$BC_{group} = \frac{\sum_{i=1}^{g}[BC_*^{'} - BC_i^{'}]}{(N-1)}$$

where BC_*' is the largest realized normalized actor betweenness index of the set of actors and N is the number of nodes in the network. DC_{group} is the normalized group degree centrality, where DC_* is the largest simple actor degree and N is the number of nodes in the network:

$$DC_{group} = \frac{\sum_{i=1}^{N}[DC_* - DC_i]}{[(N-1)(N-2)]}$$

Group density of a network is defined as the proportion of ties present in relation to all ties possible:

$$D = \frac{l}{N(N-1)/2}$$

l is the number of present edges; N is the number of nodes in the network.

The interactive TeCFlow movie can be stopped anytime to drill-down into the messages that are currently exchanged between actors. Multiple e-mail addresses can be combined into an online personality, reflecting the fact that people frequently use different e-mail addresses.

We have implemented an open architecture. Communication messages are processed locally in three steps. In the first step, the messages are parsed and stored in decomposed format in a MySQL database. In the second step the database can be queried to select messages sent or received by a group in a given time period. In the third step the selected communication flows are represented in a visual browser using static and dynamic views [11,12] and netgraphs [23]. TeCFlow has been developed in Java and is available for free download from http://www.ickn.org/ickndemo.

4. The "Sliding Time Frame" Algorithm for Temporal Visualization

To display the evolution of communication patterns over time, we developed a dynamic visualization algorithm where the layout of the graph is automatically recalculated every day, resulting in an interactive movie. The simplistic approach would have been, for any given day, to base the graph structure on the communications that occurred during this day. However, this approach does not take into account interactions that happened before this particular day, and would result in a jerky animation of low quality. For our dynamic visualization, we therefore propose a new algorithm: the sliding time frame algorithm, where we are always looking at a time interval consisting of a flexibly chosen number of days.

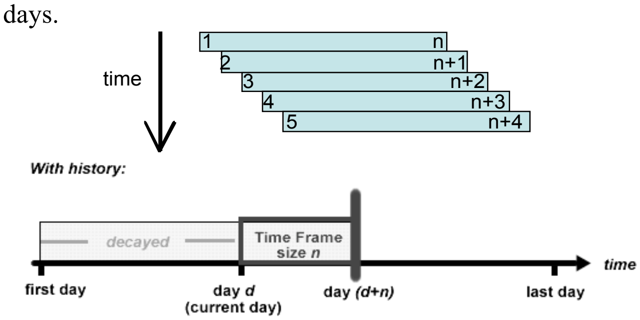

Figure 1. Sliding time frame algorithm in "with history" mode

The basic idea of the sliding time frame algorithm is to display active ties between actors in a sliding time frame covering a flexibly selected interval of n days starting from the current day d the visualization is showing (figure 1). The window frame moves forward day by day, and new ties (i.e. e-mail messages exchanged) are subsequently added to the graph each day until the desired width of n days of the sliding time frame is reached. The time window displays all messages exchanged within the current time frame. In "with history" mode, messages sent before the current time frame are included in the layout of the graph. Once an e-mail message has been sent, it will influence positioning of the actors in the graph for the rest of the animation. After the message moves out of the n-days wide time frame, it is displayed in the visualization as a dimmed out arc. This reflects persistent ties that stay active once they have been established for the remainder of the lifetime of the team.

In "no history" mode, only the edges in the current time frame are included (figure 2), i.e. only the e-mail messages exchanged within the time frame are used to calculate the graph. Day d is the current day that the visualization is showing and the current time frame is [d, d+n]. Only communications inside the current time frame are calculated and displayed. This means that the lifetime of ties (i.e., e-mail messages) is limited to the width of n days of the time frame, afterwards links decay and disappear. In this version of our system the tie decay function is binary, ties have a fixed lifespan of n days and then cease to exist.

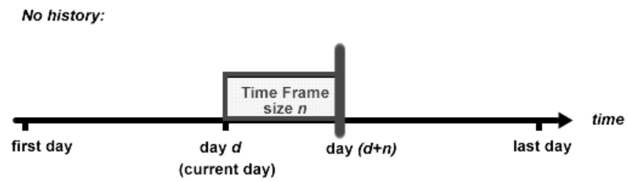

Figure 2. Sliding time frame algorithm in "no history" mode

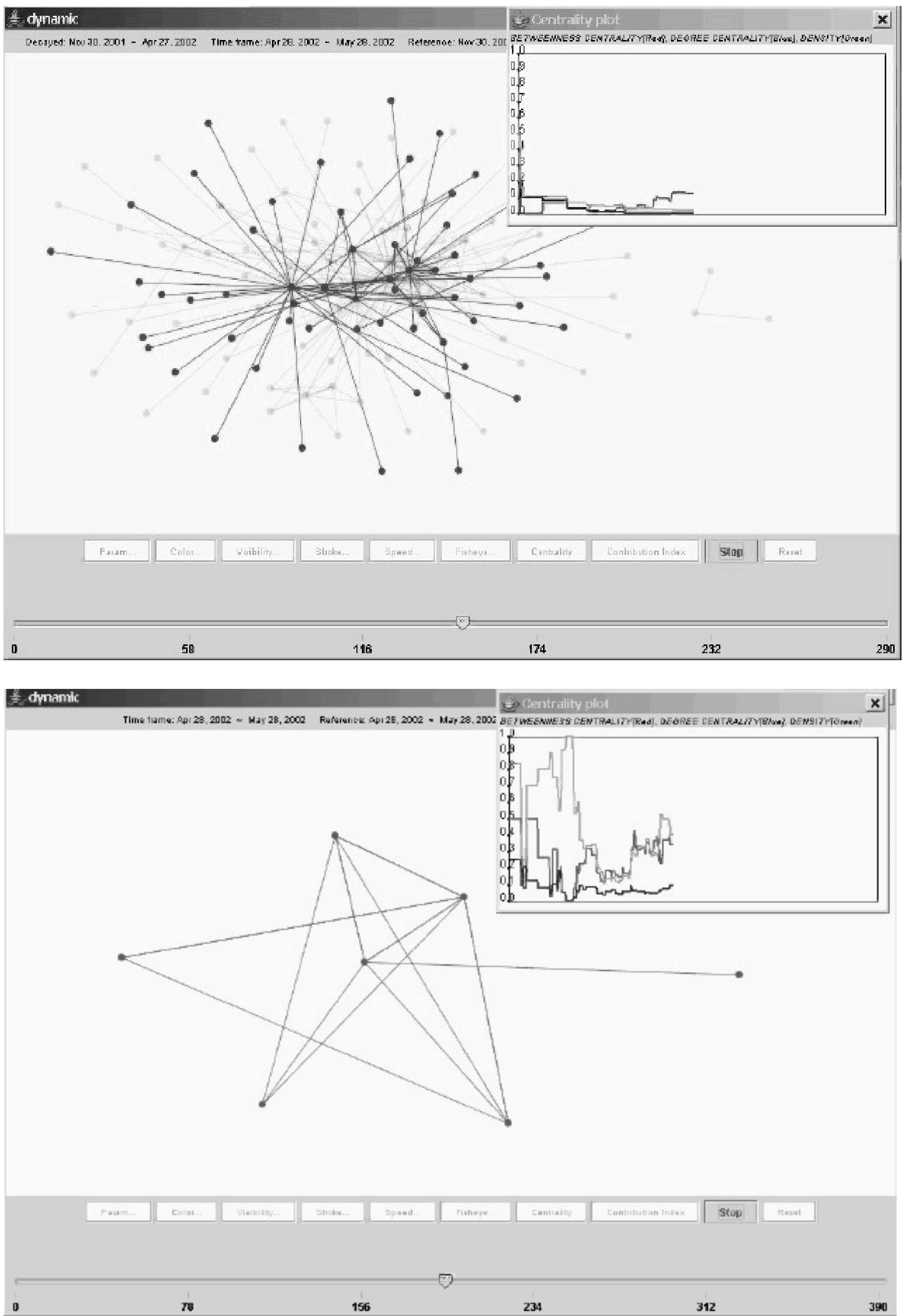

Figure 3. Temporal view with history (top) and without history (bottom)

Figure 3 gives an example, comparing "with history" and "no history" views of the same data set. It displays a snapshot of the communication flow of a subset of the online community described in section 5. In this example, messages exchanged between Nov 30, 2001 and May 28, 2002 are shown. Figure 3 shows the "with history" and "no history" screen shots of May 28, 2002. The top view of figure 3 displays accumulated links since Nov 30, 2001 "with history". The time frame is set to 30 days, which means that all the links between April 28, 2002, and May 28, 2002, are shown in full color, while the links older than 30 days are dimmed out. The bottom view of figure 3 shows the same data sets, but "without history." This means that only the links within the 30-day time window April 28, 2002, and May 28, 2002 are calculated and shown. Note the differences in group betweenness centrality, group degree centrality, and density in the small window. In "with history" mode, centralities get smoothened out, as the social network structure changes less and less over time. In "no history" mode, centralities oscillate wildly, displaying the centrality of the graph made up of the messages exchanged in the last 30 days.

5. Temporal Social Surface - Innovation

Looking at the evolution of group betweenness centrality over time allows the observer to get a quick overview of the temporal dynamics of a group. Observers need to look for peaks and troughs in the temporal evolution of group betweenness centrality and density to find changes in the collaboration patterns of groups.

The sample data set in this paper consists of an e-mail archive of a virtual consulting practice with 200 members of a global consulting firm, covering the time period from mid-2000 to early 2002. It is composed of the ego-networks of the practice leader and the practice coordinator (i.e. their e-mailboxes). Those e-mailboxes are taken as an approximation of the organizational memory of the consulting practice, as the practice leader and the coordinator were informed of all major events in the practice. The major advantage of this data set is that the author was intimately involved in the analyzed work processes. The disadvantage is that the mailboxes of the practice coordinator and the practice leader do not include direct one-to-one communication among the practice members bypassing the practice coordinator or the practice leader.

Figure 4 illustrates evolution of group betweenness centrality (GBC) of the consulting practice. The first peak in GBC in figure 4 appears when a new innovative concept is developed by the core group. The second peak appears when the brochure with the new service offering is published, the final peak denotes the official marketing launch of the new service offering.

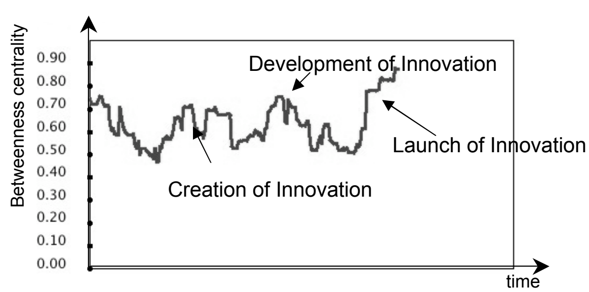

Figure 4. GBC plot

Compared to the condensed plot of the evolution of betweenness centrality, adding a curve for each individual actor and extending the curve into a surface conveys much richer information. In order to get a straightened out surface, the betweenness values for each time vector are sorted by increasing value. The price to pay for this better readable surface is the loss of traceability for individual actors.

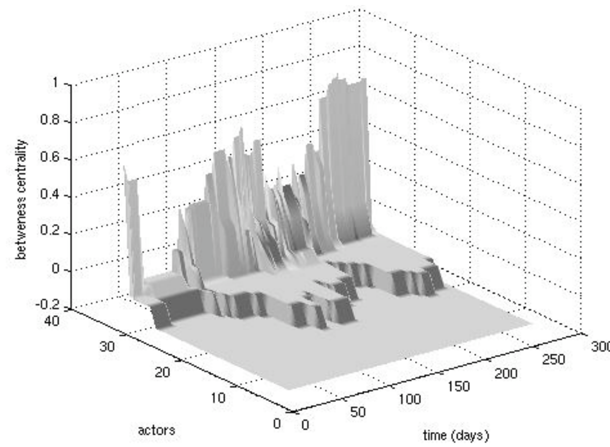

Figure 5. Temporal Social Surface: Innovation

Figure 5 shows the changes in betweenness centrality of 40 actors over a period of 300 days. Figure 5 illustrates team dynamics of the core team developing the innovative consulting service offering. This picture has been computed in "no history" mode, with a time window of 50 days. In figure 5, the majority of group members is inactive at any given time, but observers are easily able to identify peaks in activity of individuals and of large parts of the group as "elevated planes" and "peaks" in the surface. There are five phases in the lifecycle of this group. After initial kick-off by a few central actors, a larger group of about ten people communicates with low centrality. Afterwards there is a first period where collaboration of about 5 central members with about 20 peripheral people peaks. In the next phase, total participation goes down, as does centrality. In the final phase, there are two or three highly central members, communicating with about 20 peripheral participants.

6. Temporal Social Surface - Learning

The dataset for this example consists of a subset of the e-mail messages of the previous section. Communications on the subject of organizing a global Web-based seminar ("Webinar") have been collected; allocation of messages to the dataset was done manually. The archive includes 607 messages exchanged among 80 actors, covering a time period of about 190 days. The Webinar was prepared by a small team of self-selected members of the consulting practice over a multi-month period. One main speaker then delivered the Webinar during one hour, assisted by his team members. The audience of the Webinar was spread out globally, and had the opportunity to ask questions to the speakers via e-mail during the talk. Questions that could not be covered during the talk were answered in the next few days. Because of overwhelming demand, the team decided to revise and rerun the Webinar a few weeks later. The team worked together on some minor changes until the seminar was delivered again.

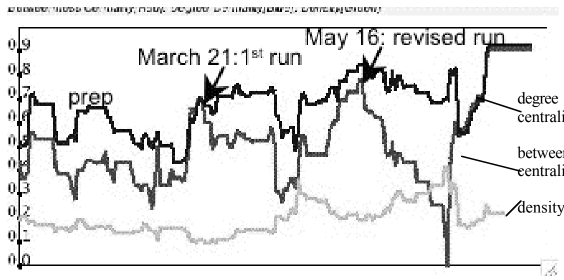

Figure 6. Group Betweenness Centrality of Webinar

Figure 6 illustrates changes in group betweenness centrality. The three phases in the organization of the Webinar can be recognized, with a decline in group betweenness centrality (dark line) for the organization of the first and second run of the Webinar, and a spike in group betweenness centrality (when the speaker delivers the event). In the preparation phase group betweenness centrality is relatively high because of the core/periphery structure [4] of the core group involved in a dialogue with potential speakers for the event.

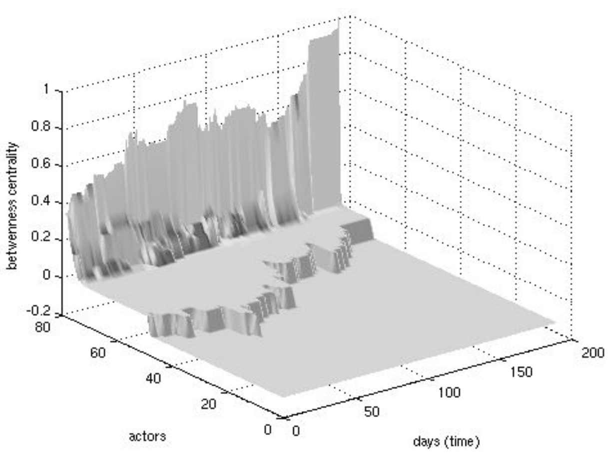

Figure 7. Temporal Social Surface: Learning

Compared with the temporal evolution of group betweenness centrality displayed in figure 6, the temporal social surface shown in figure 7 conveys much richer information. The four phases of initial preparation, first run, second preparation, and wrap-up can again be distinguished, but now we can also see that the second peak in attendance (the actors on the first plateau) is smaller than the number of attendants of the first run of the Webinar. We also see that during the first run there was a considerable number of active actors involved (the dark slope at the foot of the high peaks) besides the main speaker. The final wrap-up in phase four is done by two to three actors with high centrality, communicating with around ten people.

7. Temporal Social Surface – Interest

The last example illustrates the temporal social pattern of the diffusion of the innovation. A small core group of initially 2 people is joined in a steady growth pattern by another 80 people over a period of 300 days.

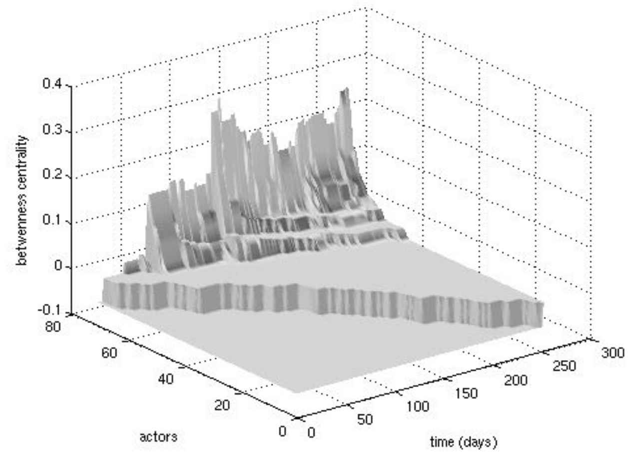

Figure 8. Temporal Social Surface: Interest

We can recognize 4 levels of membership, illustrated by the three “plateaus” in figure 8. The large majority are passive lurkers with a betweenness centrality of 0. This means that those people switch from being not involved at all to being passive recipients of message. They are on the first plateau. There is a slowly growing group of increasingly central people on three levels of active membership. On the second plateau there is a group of 3 to 5 people of modest betweenness centrality. Then there is another small group of constant size on centrality plateau three. Finally there is a growing group of 3 to 10 significantly more central actors with relatively high betweenness centrality.

8. Future Work and Conclusions

The temporal visualization of communication networks through Temporal Social Surfaces offers a novel visual way to discover different phases in the life cycle of virtual communities. It conveys insights that might be difficult to obtain by other means. The visual approach permits to find periods of low and high group betweenness centrality, and to identify potential periods of high productivity and information dissemination. It needs to be complemented by other contextual cues to obtain a full understanding of the activities, such as interviews with community members and a content analysis of the messages exchanged.

Users have found our pictures intuitively useful to gain a quick overview of group dynamics. We are currently working on a more systematic study comparing analysis of longitudinal social networks by conventional means with our dynamic approach to get a more in-depth understanding of strength and weaknesses of our method.

Our continuing goals are to gain deeper insights into the evolution of online group dynamics and developing a theory of member roles in virtual communities using more detailed communication pattern analysis.

9. Acknowledgements

I am grateful to Yan Zhao, the main developer of TeCFlow, and Thomas J. Allen, Hans Brechbuhl, Scott Dynes, M. Eric Johnson, Rob Laubacher, Fillia Makedon, and Thomas W. Malone for their help and encouragement. This project has been supported by the MIT Center for Coordination Science, the Center for Digital Strategies at Tuck at Dartmouth, and the Devlab at Dartmouth.

10. References

[1] Adar, E. Tyler, J. Zoomgraph, HP Laboratories, Palo Alto, 2003. www.hpl.hp.com/research/idl/projects/graphs/zg2003.pdf

[2] Batagelj , V. & Mrvar, A. Pajek - Program for Large Network Analysis. Connections Vol. 21, No. 2, 47-57. 1998.

[3] Borgatti, S., Everett, M. & Freeman, L.C. UCINET IV, Version 1.0, Columbia: Analytic Technologies, 1992.

[4] Borgatti, S.P. & Everett, M.G. Models of Core/Periphery Structures. Social Networks 21: 375-395. 1999.

[5] Brandes, U. & Wagner, D. visone - Analysis and Visualization of Social Networks. In Michael Jünger and Petra Mutzel (Eds.): Graph Drawing Software, pp. 321-340. Springer-Verlag. 2003.

[6] Bulkley, N. & Van Alstyne, M. Why Information Should Influence Productivity. (forthcoming) In The Network Society: A Cross-Cultural Perspective. Manuel Castells, ed. Edward Elgar Publishing. 2004.

[7] Cross, R. & Cummings, J. Relational and Structural Network Correlates of Performance in Knowledge Intensive Work. Academy of Management, Seattle, WA. Paper published in Proceedings. 2003.

[8] Danah, B. Potter, J. Viegas, F. Fragmentation of identity through structural holes in email contacts. Proc. Sunbelt XII. 2003.

[9] Fruchterman, T.M.J. Reingold, E.M. Graph drawing by force directed placement. Software: Practice and Experience, 21(11), 1991.

[10] Gloor, P. AACE–Algorithm Animation for Computer Science Education. Proceedings of the 1992 IEEE Workshop on Visual Languages, 1992.

[11] Gloor, P. Laubacher, R. Dynes, S. & Zhao, Y. Visualization of Communication Patterns in Collaborative Innovation Networks: Analysis of some W3C working groups. Proc. ACM CIKM 2003, New Orleans, Nov. 5-6. 2003.

[12] Gloor, P. Zhao, Y. TeCFlow - A Temporal Communication Flow Visualizer for Social Networks Analysis, ACM CSCW Workshop on Social Networks. ACM CSCW Conference, Chicago, Nov. 6. 2004.

[13] Krackhardt, D. Blythe, J. & McGrath, C. KrackPlot 3.0 User's Manual. Pittsburgh: Carnegie-Mellon University. 1995.

[14] Lawrence, A. Badre, A. Stasko, J. Empirically Evaluating the Use of Animations to Teach Algorithms, Proceedings of the 1994 IEEE Symposium on Visual Languages, St. Louis, MO, October 1994, pp. 48-5

[15] Mazzocchi, S. Virtual Community Dynamics, 2003. http://www.betaversion.org/~stefano/papers/AC2003-agora.pdf

[16] McGrath, C. Blythe, J. Do You See What I Want You to See? The Effects of Motion and Spatial Layout on Viewers' Perceptions of Graph Structure Authors. Journal of Social Structure JoSS, Vol. 5, No. 2. 2004.

[17] Moody, J. McFarland, D. Bender-deMoll, S. Dynamic Network Visualization: Methods for Meaning with Longitudinal Network Movies. 2004. http://www.sociology.ohio-state.edu/jwm/NetMovies/Sub_CD/dynamic_nets_public.html

[18] Mutton, P. PieSpy Social Network Bot. http://www.jibble.org/piespy/ (Accessed 14 October 2003)

[19] NetMiner, Cyram Co. Ltd., 2003, http://www.netminer.com

[20] Smith, M. Fiore, A. Visualization components for persistent conversations, Proc ACM CHI. 2001.

[21] Stark, D. Vedres, B. The Social Times of Network Spaces: Sequence Analysis of Network Formation and Foreign Investment in Hungary, 1987-2001, Columbia University. 2004.

[22] Van Alstyne, M. Zhang, J. EmailNet: A System for Mining Social Influence & Network Topology in Communication. University of Michigan working paper. 2003.

[23] Varghese, G. Allen, T. Relational Data in Organizational Settings: An Introductory Note for Using AGNI and Netgraphs to Analyze Nodes, Relationships, Partitions and Boundaries. Connections, Volume XVI, Number 1 & 2, Spring. 1993.

[24] Wasserman, S. Faust, K. Social Network Analysis. Cambridge University Press, 1994.