Social Coherence in Computer-Mediated Communication

John C. Paolillo
School of Library and Information Science
011 Main Library, Indiana University
Bloomington, IN 47405 USA
+1 812 855 3247
paolillo@indiana.edu

Keywords
Social network analysis, computer mediated communication

INTRODUCTION

Among the many ways of identifying coherence in discourse, a particularly useful one considers different social dimensions of coherence, as revealed through the unfolding of the exchange of messages, turns or other units of information. When we communicate, whether face-to-face, through message systems, databases or any other technical affordance, we engage in an inherently social act. Social acts of communication are the vehicle for the organization, overt or tacit, of our collective action, and many other aspects of the social world. Thus, through observing and analyzing patterns of communication, it is possible to reveal patterns of social coherence at the level of both the group and the individual. These patterns are especially helpful in visualizing and understanding social processes at work in large scale interactions, where other methods find themselves adrift in a sea of data, without sufficient means to generalize about the processes of interaction. Suitably constructed visualizations of social coherence in the exchange of messages among the participants in a discussion can focus the attention of researchers and designers to socially significant sites of interaction, which can be interpreted in terms of the social architecture of the discourse studied, or which may be examined in greater detail for further insight.

In this position paper, I offer several examples of patterns of social coherence in CMC drawn from my own research, where patterns of interaction inherent in the communication exchanges are extracted and represented as social network diagrams. I attempt to illustrate the principles guiding my approach and the kinds of information that it reveals, as well as a few of its potential applications.

SOCIAL NETWORKS AND COHERENCE

The approach to social coherence adopted here comes from the theory of social networks, an area of research in sociology in which the relations among members of a group are of primary interest. Social networks can be constructed out of nearly any kind of relational data, such as common membership in clubs or organizations, affinities for certain objects or activities, or exchanges of gifts, material goods, or symbols, among many other things. All of these relations express different aspects of social organization or structure, which can be regarded as a system of roles, relations, evaluations, and actions lending unity and coherence to the action of a collection of individuals, much the way that direction, form and purpose are imparted to a flock of birds.

The form of visualization most characteristic of social network analysis is the social network diagram, in which points or nodes represent individuals at some level of analysis, and lines are drawn between the points indicating links between various individuals. Lines may have arrowheads, indicating a direction of relationship (e.g., A gives something to B), resulting in a directed social network diagram, and they may have different weights, to indicate different strengths of relationship. Social network diagrams offer an easy to read map of the social relations among members of a group.

In terms of structural analysis, one generally finds two types of approaches in the literature: clique analysis and positional analysis. Clique analysis tried to identify sets of mutually inter-related members in a larger group. Partitioning a group of people into cliques is a useful way of identifying which people are most likely to share certain kinds of attributes, knowledge, or behaviors. Clique analysis alone isn't always able to identify the sources of such shared characteristics, however, and it is often supplemented with positional analysis. Positional analysis seeks to classify the members of a group into equivalence classes in terms of their relational characteristics. Members of a clique, since they are mutually related to one another, tend to hold similar social positions. But positional analyses can also place group members at common degree of remove from a clique into an equivalence class, even though they have no mutual relationship. Thus, positional analysis reveals patterns of structural coherence that do not depend on direct relationships among participants.

The two types of analysis, clique and positional, are provide complementary information about a network, and for a complete understanding it is desirable to conduct both analyses on the same data and interpret them together. My

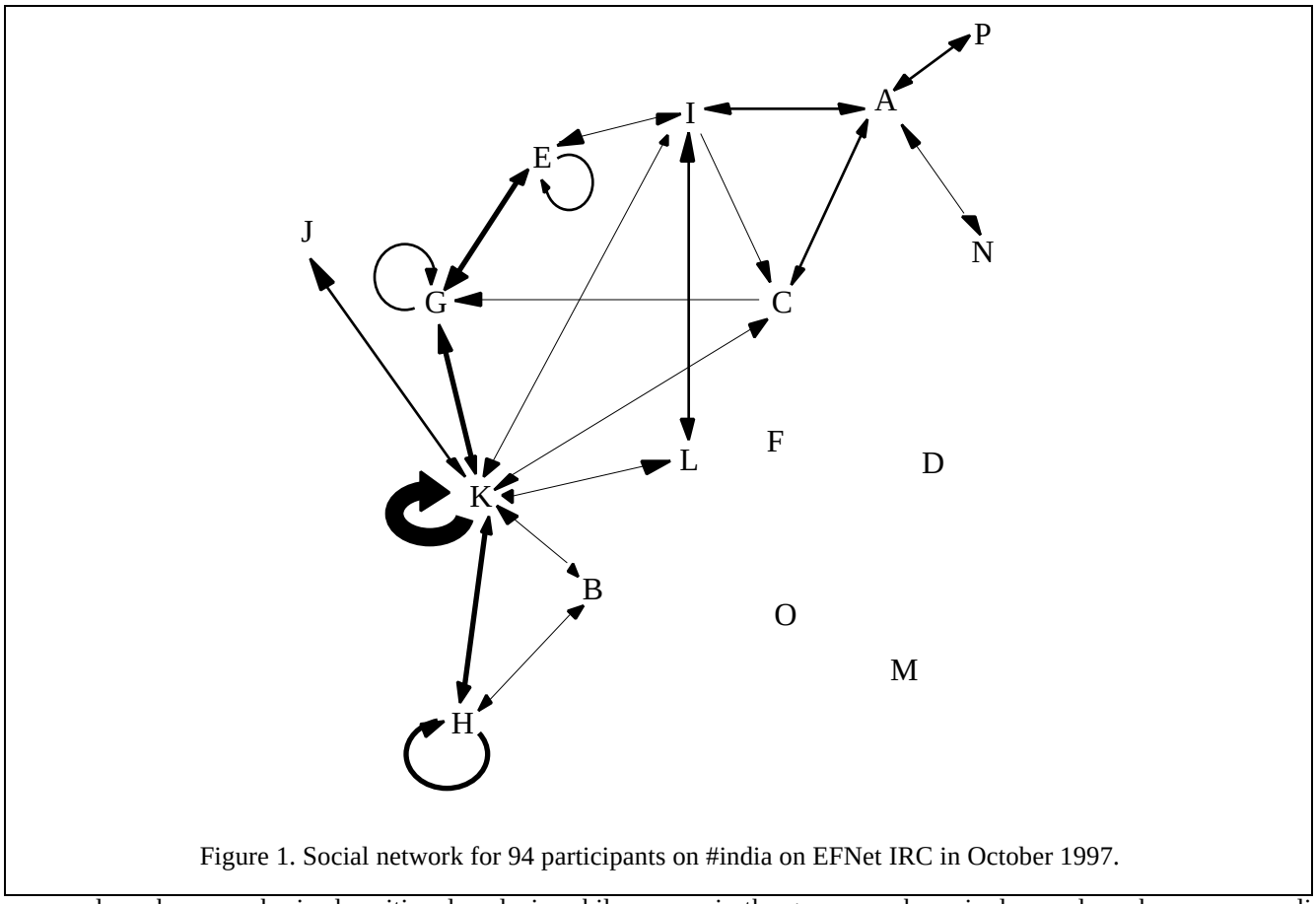

Figure 1. Social network for 94 participants on #india on EFNet IRC in October 1997.

own analyses have emphasized positional analysis, while those of others favor clique analysis. Hence, in this paper I focus on positional analysis, in order to illustrate its methods and the sorts of results it can yield.

FROM CONVERSATION TO NETWORKS

Social network methodology gives many different ways of establishing relations among members of a group and most commonly this is done via some extrinsic means, such as through interviews or surveys conducted with the group members. Computer-mediated communication offers a unique opportunity for social network analysis since we often have a comprehensive record of the messages exchanged among the members of a group. The exchange of a message is a nearly atomic social act, since relationships between individuals are typically manifested out of dozens, hundreds or thousands of such acts. Thus, messages give us a more fine-grained view of social relationship than is possible through traditional methods. In addition, messages provide a measure of social relationship in large-scale communication networks such as Usenet, where survey and interview methods would be laborious and intractable.

Message exchange data can be extracted readily from corpora of CMC (or monitored in real-time CMC interfaces). Likewise, positional analyses can be accomplished readily using correlational techniques such as factor analysis. In a factor analysis, group members can be grouped into equivalence classes according to their relations with others in the group: each equivalence class shows up as a distinct factor, so each member loads most strongly on one factor. These factors or classes are then equated with social positions. Since factor analysis extracts significant co-variation, there is always one residual class whose members' interactions are not strongly correlated with those of any other member; this class usually represents a peripheral social position. Sometimes it is useful to conduct distinct factor analyses on different arrangements of the same data, and to collate the resulting factors for a richer set of social positions than would come from a single analysis. For example, participants in a discussion group might be classified according to who addresses them on the one hand, and according to who they address on the other. Cross-classification according to these two characteristics yields social positions representing their total communicative behavior.

To construct a social network diagram indicating the relation among these positions, the messages exchanged among the group members need to be aggregated by social position. When this is done, the strength of the link between any two social positions can be weighted by the number of messages for each link. Sometimes, further analysis using log-linear modeling can be helpful in setting the weights to take account of the number of participants in each position. Drawing the social network diagram can proceed either by algorithm or by hand; since the requirements of my social network analyses tend to be more complex than the more common display algorithms

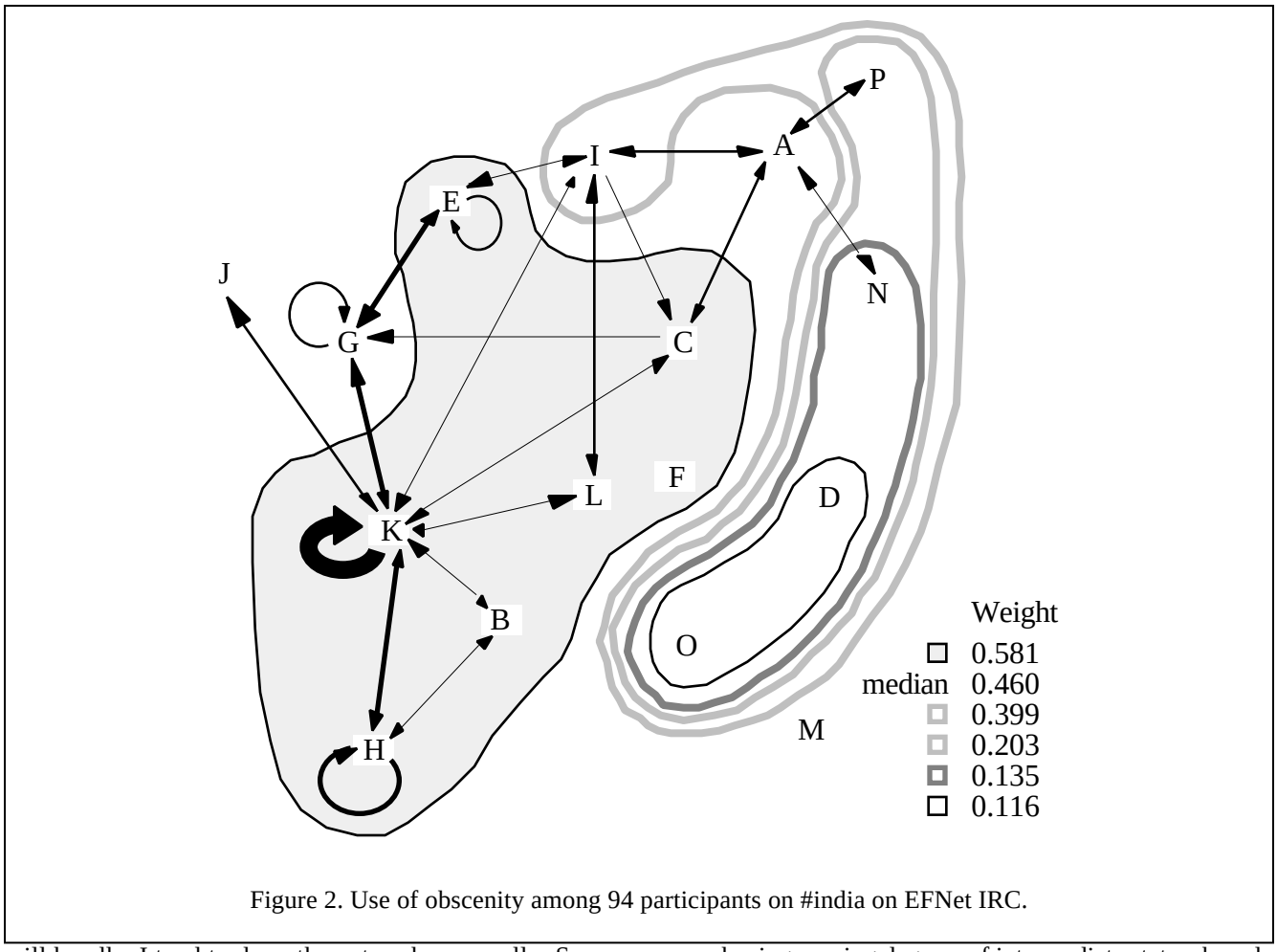

Figure 2. Use of obscenity among 94 participants on #india on EFNet IRC.

will handle, I tend to draw the networks manually. Specialized algorithms could readily be fine-tuned and programmed into a user interface if required.

NETWORK VISUALIZATIONS

The factor analysis technique is very powerful and can work with a number of different arrangements to produce different views of the same data. Depending on the data and the arrangement, different interpretations are available to the researcher. In my own use of the technique, I have applied factor analysis to IRC turn-exchanges to classify chatters in terms of who they address and who addresses them, and to Usenet cross-posting data to ascertain clusters of newsgroups with substantial traffic in common. Representative visualizations are given in figures 1 and 2.

Figures 1 and 2 represent the interaction among different members of an IRC channel #india on EFNet IRC for a 24-hour period in Fall 1997. The 94 most frequent participants are represented in this diagram, where sixteen social positions represent the total communicative behavior of these participants. Since group size varies from one member (N) to 38 members (A), a log-linear model was used to assist in setting the line weights.

Figure 1 clearly indicates the central status of position K, and peripheral status of positions such as D, F, M and O. Other positions, such as G, H, I and C, can be interpreted as having varying degrees of intermediate status, based on their linkage to K, either directly or via intermediaries. These positions, and their corresponding memberships, are not immediately evident from a qualitative reading of the IRC logfile, though it seems likely that long-time participants on EFNet #india have knowledge about the participants that corresponds to what is represented in Figure 1. Thus, positional analysis of the message data enriches the information available for interpreting a record of CMC.

The positions observed are useful in other ways as well: Figure 2 presents the same network as Figure 1, though this time with isobars representing the probability of the use of obscene language (a common occurrence on this channel) superimposed, as predicted by a log-linear model. This visualization shows a clear correspondence between the use of obscenity and social position; while many interpretations can be offered for this fact (see Paolillo 2001), such correspondence are potentially useful to analysts studying CMC, as well as to users who may wish to exercise certain kinds of choice over their computer-mediated interactions.

CONCLUSION

Positional social network analyses based on message exchange data in CMC offer readily-interpreted an potentially useful visualizations of interactional coherence in CMC. Positional analysis can be conducted based on a CMC corpus, or on real-time data, using algorithmic or manual techniques. The visualizations make available information that is not otherwise self-evident from qualitative or other forms of quantitative analysis of the messages. At the same time, these techniques augment other social networks , which may be based on extrinsic data sources, or which may use clique analysis in place of positional analysis.

REFERENCES

Degenne, A. and Forse, M. 1999. *Introducing Social Networks*. Thousand Oaks: Sage Publications.

Donath, J.; K. Karahalios, and F. Viegas. 1999. Visualizing Conversations. *Proceedings of the 32nd Hawaii International Conference on System Sciences.* Los Alamitos, CA: IEEE Computer Society.

Paolillo, J. Forthcoming (2002). "'Conversational' codeswitching on Usenet and Internet Relay Chat". To appear in S. Herring, ed., *Computer-Mediated Conversation*. Cresskill, NJ: Hampton Press.

Paolillo, J. 2002."Democratic Participation in the Discursive Management of Usenet". *Proceedings of the 35th Hawaii International Conference on Systems Sciences*. Los Alamitos, CA: Institute of Electrical and Electronics Engineers (IEEE) Computer Society.

Paolillo, J. 2001. "Language variation in the virtual speech community: A social network approach to Internet Relay Chat". *Journal of Sociolinguistics*. Oxford: Basil Blackwell. Preprint at http://ella.slis.indiana.edu/~paolillo/research/variation.pdf

Paolillo, J. 2000."Visualizing Usenet: A factor analytic approach". *Proceedings of the 33rd Hawaii International Conference on Systems Sciences.* Los Alamitos, CA: Institute of Electrical and Electronics Engineers (IEEE) Computer Society.

Paolillo, J. 1999."The virtual speech community: Social network and language variation on IRC". *Journal of Computer-Mediated Communication* 4.4 http://www.ascusc.org/jcmc/

Paolillo, J. 1996."Language choice on soc.culture.punjab". *Electronic Journal of Communication/Revue Electronique de Communication*, 6(3). http://www.cios.org/ and http://ella.slis.indiana.edu/~paolillo/research/paolillo.publish.txt

Sack, W. 2000. Discourse diagrams: Interface design for very large-scale conversations. *Proceedings of the 33rd Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society.

Tacq, J. 1997. Multivariate Analysis Techniques in Social Science Research: From Problem to Analysis. Thousand Oaks, CA: Sage.

Wasserman, S. and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Appendix

(a) John C. Paolillo is an associate professor of Information Science and Informatics at Indiana University. He earned his Ph.D. in Linguistics from Stanford University, and has conducted social and linguistic research on computer-mediated communication since 1995. He specializes in the application of quantitative methods and statistical models to data from language use and social interaction.

(b) Linguistics, Information Science, Quantitative methodologies and statistical models.

(c) See references section under Paolillo

(d) The Sociable Media Group (MIT Media Lab: Judith Donath) http://smg.media.mit.edu/