Understanding Patterns of Media Use in an Activity-Centric Collaborative Environment

David R Millen, Michael J. Muller, and Kate Ehrlich

IBM T. J. Watson Research

One Rogers Street

Cambridge, MA 02142

{david_r_millen, michael_muller, katee}@us.ibm.com

1. Background

We began this research program to address a gap in existing collaborative environments. Ad hoc collaboration systems, such as email and chat, are lightweight and flexible. They provide good support for short-term, dynamic communication needs. However, collaborative activities which extend over longer periods of time, or over larger numbers of participants, are notoriously difficult to conduct over email or chat. Discussion databases and structured workspaces are more appropriate to the larger-scale collaborations, but they are difficult to set up. It is particularly difficult to begin with an informal collaboration (e.g., in a chat), and then extend that collaboration to a somewhat larger membership (perhaps in email), and then to extend that collaboration again into a more formal, large-scale environment (a discussion database or a workspace).

A second problem is that most collaborative environments support one or at most two types of documents as "first class objects." For example, most instant messaging environments support text (first class object) and a few graphic indicators, and perhaps URLs (second class object); email supports text (first class) plus attachments or URLs (second class); discussion databases support text (first class) plus attachments or URLs; and so on. Logically, a complex activity may contain diverse resources, such as simple text, formatted text, spreadsheets, graphics, and executable modules, but these diverse documents are either (a) spread across a number of specialized repositories, or (b) attached to a smaller number of first class objects such as texts (email messages, discussion database entries).

This diversity of shared resources means that people must monitor and participate in multiple shared venues, dividing their attention and their effort across multiple storage media and multiple communication/notification paradigms. Even if they are successful in this divided attention and context management task, they face difficulties in choosing which medium to use for any new collaborative activity.

A third problem, the *scope and dynamism* of sharing, is derived from the other two problems. Most shared environments are restricted to either asynchronous sharing (shared file systems, discussion databases, email) or synchronous sharing (online meetings, chats). There is a trend toward providing both shared asynchronous stores and shared synchronous experiences (e.g., [5]).

1.1 ActivityExplorer

ActivityExplorer (AE) is a prototype client that manages multiple types of shared items (we will use the word "item" for any type of document or other shared object) in an environment featuring structured collections of items, synchronous and asynchronous sharing of items, presence (awareness of members' online status), access control at the level of items or item collections, item status (whether an item is currently being accessed by one or more members), and notifications of selected actions by other members.

Members use ActivityExplorer to organize items into one or more item collections, called activity threads (Figure 1). Each person is likely to be a member of more than one activity thread, and therefore AE also provides a list view of all activity threads to which the member has access permission. This list view may optionally show *all* items in all threads to which the member has access privileges, in which case the list shares some attributes of the multiple-object email-centric view of the Thrasks research program [1]; see also [7].
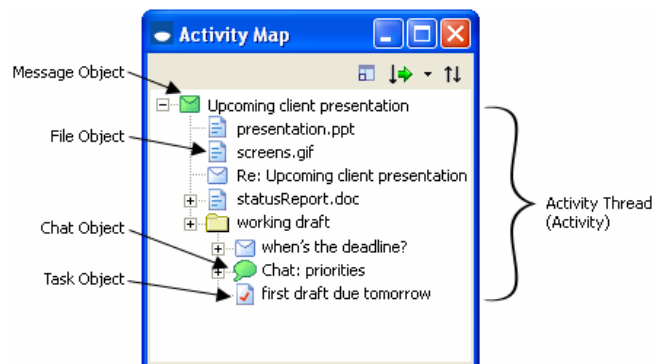

Figure 1. .An Activity Thread including four of the six types of items.

Each activity thread contains one or more items in a tree structure (Figure 1). Any type of item may serve as either parent or child in the tree structure. Items may be created by any of several simple operations: creating a *new activity thread*, including its root item; creating a child item to an existing item in a thread; dragging a document (including an email message) from outside of AE into AE, and dropping the document on the name of a person with whom it is to be shared. Items may be rearranged in the tree, or into other trees (different activity threads), as people's needs change.

Each item has a list of members, who are the people who have access permissions on that document. A user may create an item as the child of the current item, in which case the default is for the child to inherit the membership of the parent item. Membership may also be edited manually for each item, or for an item and all of its child items (i.e., from the branch "out" to all of its leaves), or for the entire activity thread in which the item occurs. In the research prototype, all members have equivalent status and

permissions (we recognize that a production version of this concept may require more defined roles for members). In keeping with conventional security approaches, people are unable to see items of which they are not members.

The current version of ActivityExplorer supports six item types: **messages** (similar to email), **files** (formatted documents prepared with specialized tools, such as word processors or spreadsheet programs), **persistent chats**, **tasks** (action items with externally visible *completed/non-completed* status), **screen-shares** (including graphical annotations), and **folders**. Each type of item appears as a first class object in the activity thread, as shown in Figure 1. Users may easily find an item by its name, instead of searching for a message that contains or references the item. Future versions of ActivityExplorer are likely to support search across diverse item types.

In summer 2003, we conducted a field trial of AE in a research group of 33 people [9], including 14 interns and their mentors, plus five additional researchers. In this paper, we are specifically interested in understanding the differences in media use (item use) by participants in the field trial. Social network analysis has been selected as the analytical framework because the relationships among persons within the trial are directly observable via their communications interaction. An important theoretical concept is the strength of social ties between two people, or among the members of a team or working group. The 33 members of the community using ActivityExplorer constructed different interaction patterns while using over 2000 "items" to construct over 400 activities. We may use these interaction patterns, and their directionality (*from* document author, *to* other members of the document), to analyze social networks of communication among the six media of the six item types.

2. Research Methods

At the beginning of the field study, we requested intern-mentor pairs to learn about and use AE as a tool to prepare for their initial project poster in the beginning of their internship. We also encouraged people to use the AE tool other work activities for the remainder of the summer. The data for this report comes from the period July 3 through September 2, 2003

For this investigation, we analyzed a user level "action log." 14 specific user actions (e.g., create, read, delete, or add a new member) were logged. For each action, the following data was captured in the log: date/time, item number, activity thread, and the ID of the user.

Our understanding of the use of the AE was also informed by interviews with the student interns. We talked with seven of the 14 interns about their experiences with AE. Interviews covered several broad topics (beginning to use AE, intern-mentor work, collaborating with other interns, responding to notifications), but were deliberately open-ended to allow opportunistic story-telling and collection of interesting items or item sequences. These interviews were conducted by one member of the team, and took place in each intern's work area. Interns could illustrate their points by bringing up a particular activity thread or object on the screen, or by showing a resulting paper artifact.

3. Results

The first set of analyses presented here will show that the choice of various media types was not consistent across all individuals in the trial. We used social network analysis to investigate media use differences among the trial population. To do this, we mined the log files to discover the media use for each media type for every pair combination. Every time a particular media type (e.g., chat, file, and message) was used, the strength of that media type relationship was incremented for each of the members. The direction of the relationship link was from author to reader. For example, if person A authored and shared a message with persons B and C, then the media type "message" link from A to B and from A to C would be increased.

We are admittedly using this social network analysis of media use for exploratory purposes as a way to visualize the media patterns for the trial participants. It is our hope that it will guide further investigation and increase our understanding of the how work groups interact using a collaborative tools like the AE (for similar analyses applied to email, see [2, 11]).

We have presented selected network graphs for several of the media types below (see Figures 2 and 3). For consistency, we have included the complete pool of mentors and student interns who were part pants in the trial. At the same time, for clarity we have excluded, in these network graphs, several other members of the trial.

3.1 Individual Differences in Media Use

The sociogram presented in Figure 2 shows the strength of relationship among trial participants based on the number of *message* items shared. While there were generally a significant number of messages shared among many of the participants, there were several members of the group who are clearly isolated.

The sociogram in Figure3 shows the interactivity among the trial participants using the *screensharing* media type. Not all of the trial participants used this media type and we were curious to the kinds of work that was supported with screen sharing. The two heaviest users of screen sharing were engaged in showing colleagues" how to use" various software modules or to show specifically the changes needed to some software under development.

3.2 Similarities of media use

In order to see whether the various media types were used with similar frequency by participants in the trial, we computed the correlations between the matrixes for each media type. As seen in Table 1, there are small to moderate, reliable (p < .01) correlations for each of the various media types. It appears that the behaviors for creating/reading messages, files and chats are slightly more consistent than for screen and task items.

	Files	Chat	Screen	Tasks
Messages	0.49	0.38	0.24	0.17
Files		0.52	0.33	0.24
Chat			0.19	0.25
Screen				0.27
Tasks				

Table 1 Correlation between cells in different media matrices. (All correlation values p < .01)

3.3 Distributions of media use

Another way to understand media use is to look at the relative number of times an individual initiate or authors a communication event in various media versus the number of times he or she is the

recipient (e.g., reader). In Figures 4 and 5, we have shown a scatter plot of trial participants as a function of the number of posts versus reads. In social network analytic terms, this is a representation of the number of inflow/outflow vertices for each node in the graph. In each of the charts, we have plotted the line at which the number of inflows is equal to the number of outflows.

Several aspects of Figures 4 and 5 are noteworthy. First of all, the scales of the axes are quite different, revealing the different degree of use for each of the media types supported by the AE. In Figure 5 a fairly large group of participants read more often than they contributed, while a smaller number of individuals were exceptionally prolific message writers (i.e., EE and L). A cluster of nonparticipants is also visible; they neither wrote nor read the message discourse. Somewhat surprisingly, we did not find the traditional "lurker" profile, someone who actively reads a large number of items, but rarely contributes.

For chat conversations, Figure 5 shows a slightly more even distribution of participation. The distribution shows an outlier (U) who initiated significantly more chat events than she was invited to. In this case, the participant initiated a single, very large group chat which accounted for much of the asymmetry.

3.4 Group difference in media use

The final results to be presented here include an analysis of differences in media use by two of groups in the trial: interns and mentors. In Table 2, we present the network density measures for important partitions of the matrix. The density measure reflects the "average" number of media items. The students were considerably more likely than the mentors to communicate among themselves (intern to intern) using messages. On the other hand, interns and mentors used messages with the same frequency to communicate with members of the other group. The higher density of files, screen sharing, and chats from mentors to interns most likely reflects the sharing of information (files, screens) or initiated work-related chats.

	Interns to Interns	Mentors to Mentors	Mentors to Interns	Interns to Mentors
Messages	6.5	2.1	3.5	3.6
Files	0.87	0.74	1.6	0.77
Chat	0.13	0.17	0.25	0.14
Screen	0	0.17	0.1	0
Tasks	0	0.16	0.05	0

Table 2. Density measures for group partitions of media matrices.

4. Discussion

Sociograms offer one way of visualizing patterns of sharing. While social network analysis and resulting sociograms have been used as a research tool for interactive systems for some time (e.g., [6], the application of network information is only now emerging in software applications. Research systems like Contact Map [10] and web-based network tools like Friendster [3] and Linkedin [8] are good examples of the use of social network information in application design.

The sociograms that we have seen for media use suggest that informal collaborative tools like AE could also use network information in novel and useful ways. We were able to easily identify specific role behaviors and core information providers (and consumers) through various sociograms based on AE usage data. Team leaders and program managers may be able to use the sociograms to understand organizational needs and opportunities. Media-based network information could also be mined and used to indicate the media preference for individuals, or perhaps even suggest the media choice (or template of choices) for a specific set of work activities.

The sociograms of interaction reveal individual variation in the use of various media support by the AE tool. These differences can be explained in several ways. A role-based assignment (i.e., intern coordinator) resulted in a distinctive central node and high degree of file sharing. Interest in a common activity (i.e, development of a particular software application) resulted in a distinctive pattern of sharing of task items. And finally, individual communication preferences may explain the differences observed in message and chat frequencies among trial participants.

5. REFERENCES

[1] Bellotti, V., Ducheneaut, N., Howard, M., Smith, I., "Taking Email to Task: The Design and Evaluation of a Task Management Centered Email Tool," *Proc. ACM CHI 2003*, Ft. Lauderdale, Florida, 2003.

[2] Fisher, D., "Email in the social workscape. Position paper at CSCW 2002 workshop, *Redesigning email for the 21st Century*.

[3] Friendster, http://www.friendster.com.

[4] Geyer, W., Vogel, J., Cheng, L., Muller, M, "Supporting Activity-Centric Collaboration through Peer-to-Peer Shared Objects,": *Proc. ACM Group 2003*, Sanibel Island, FL, USA, November 2003.

[5] Groove Networks, http://www.groove.net

[6] Haythornthwaite, C., "Tie strength and the impact of new media," *Proc HICSS 2001*.

[7] Kaptelinin, V., "UMEA: Translating Interaction Histories into Project Contexts," *Proc. ACM CHI 2003*, Ft. Lauderdale, Florida, 2003, 353-360.

[8] Linkedin, http://www.linkedin.com.

[9] Muller, M.J., Geyer, W., Brownholtz, B., Wilcox, E., and Millen, D.R., "One-hundred days in an activity-centric collaboration environment based on shared objects." *Proc. ACM CHI 2004*.

[10] Nardi, B.A., Whittaker, S., Isaacs, E., Creech, M., Johnson, J., and Hainsworth, J., "Integrating communication and information through ContactMap. *Communications of the ACM*, *45* (4). (2002).

[11] Tyler, J., Wilkinson, D, and Huberman, B.A., "Email as spectroscopy: Automated discovery of community structure within organizations." Technical report available at http://www.hpl.hp.com/research/idl/papers/email/ (2004)
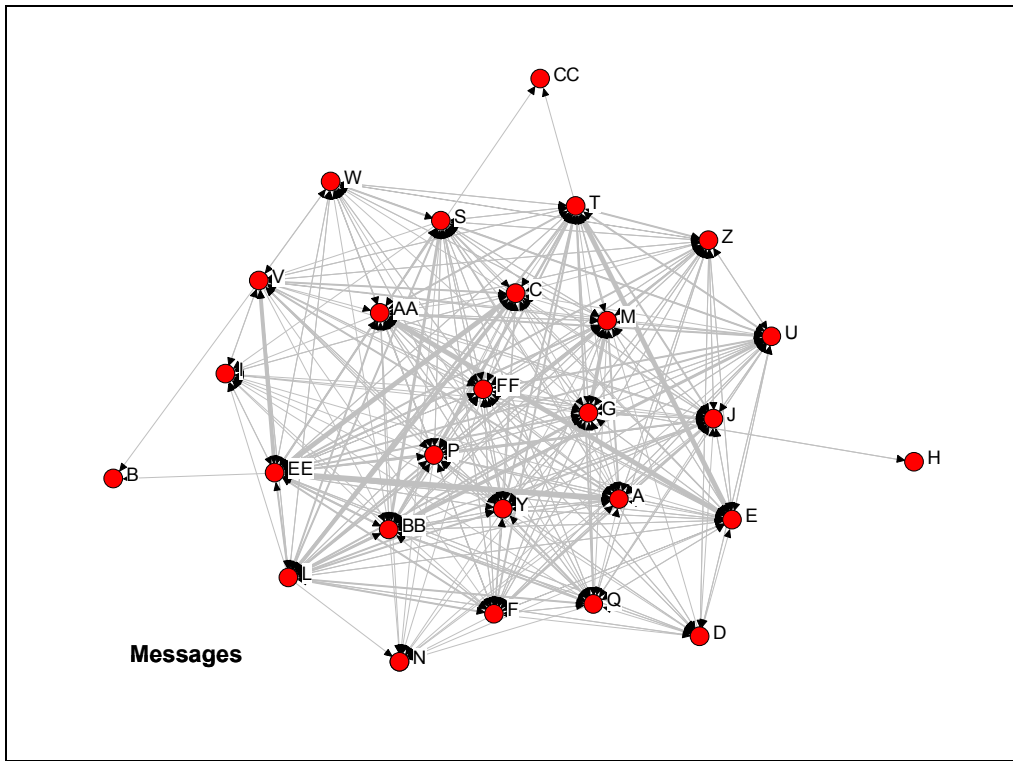
Figure 2. Sociogram of message interactions during the field trial.
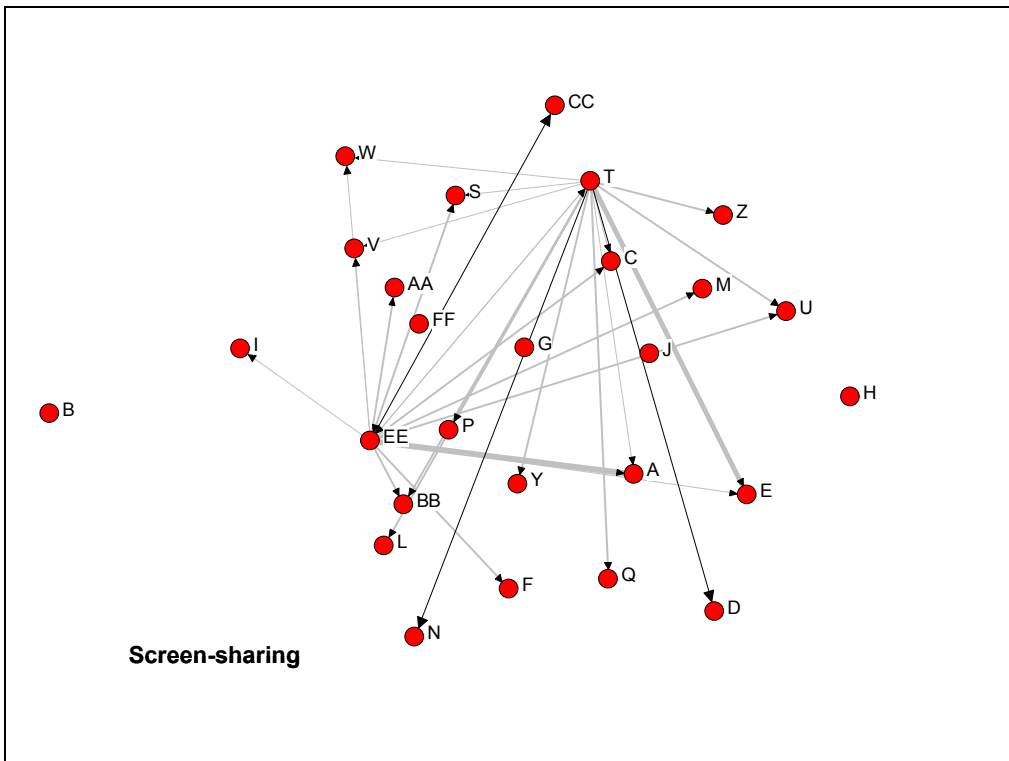

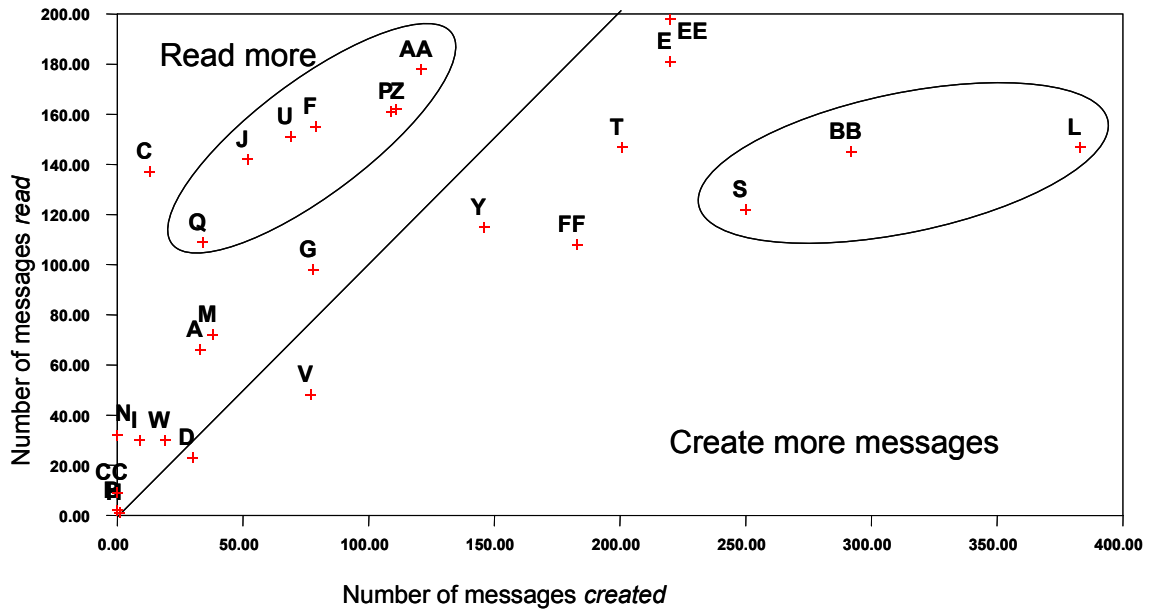**Figure 3. Sociogram of screen-sharing interaction.**

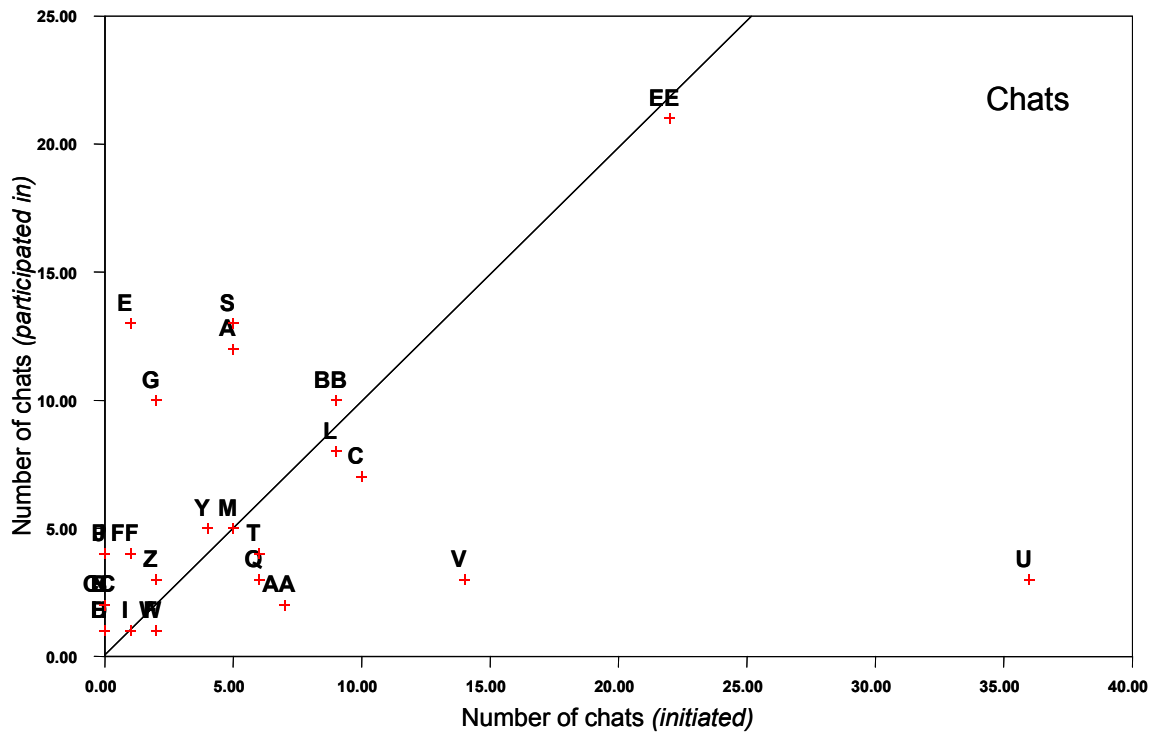**Figure 4. Inflows/outflows (creates/reads) for message items.**

Figure 5. Inflows/outflows (create/participate) for chat items.