ISR Institute for Software Research
University of California, Irvine

Studies of Automated Collection of Email Records

Danyel Fisher
University of California, Irvine
danyelf@ics.uci.edu

Paul Moody
Lotus Research
paul_moody@lotus.com

August 2002

ISR Technical Report # UCI-ISR-02-4

Institute for Software Research
ICS2 210
University of California, Irvine
Irvine, CA 92697-3425
www.isr.uci.edu

Studies of Automated Collection of Email Records

Danyel Fisher
Institute for Software Research
University of California, Irvine
danyelf@ics.uci.edu

Paul Moody
Lotus Research
paul_moody@lotus.com

ISR Technical Report # UCI-ISR-02-4

August 2002

Abstract: There are few quantitative techniques for directly measuring email use patterns. This paper describes an automated tool that, with a user's permission, reads their mail database to create a one-time snapshot and gathers relevant structural and behavioral information. We successfully collected important statistics about message threading, folders, and mail volume. Our techniques are relevant to the further development of mail systems, and to future studies of email behavior.

Studies of Automated Collection of Email Records
Technical Report UCI-ISR-02-4

Danyel Fisher
Information & Computer Sciences
University of California, Irvine
Irvine, CA 92697
danyelf@acm.org

Paul Moody
Lotus Research
55 Cambridge Parkway
Cambridge, MA 02142
paul_moody@lotus.com

ABSTRACT
There are few quantitative techniques for directly measuring email use patterns. This paper describes an automated tool that, with a user's permission, reads their mail database to create a one-time snapshot and gathers relevant structural and behavioral information. We successfully collected important statistics about message threading, folders, and mail volume. Our techniques are relevant to the further development of mail systems, and to future studies of email behavior.

KEYWORDS
Email, asynchronous communication, filing, conversation threads, study techniques, methodology

INTRODUCTION
Email is the leading form of online conversation. It functions as a "habitat"—as a personal information space handling todo lists, calendars, and important communication [3]. Users find value in its currency as well as its importance as an archive [16]. Email users seem to receive increasing amounts of email annually [13] as the technology becomes an important one for both office and interpersonal communication. With many relationships and business connections maintained dominantly over email, it is important to understand how people are actually using their email to deal with message flow and overload.

Figure 1: a message thread, time running left to right. Messages are represented by nodes; two messages are joined if the latter is a response to the former. Note that there are three branches.

A series of previous studies have ethnographically observed small groups of users organizing their office email. Each of [2, 5, 16] have produced fieldwork-based projects discussing email use. They have examined the amount of email that users receive, and have discussed the roles that email takes in their lives. A large-scale, quantitative approach can allow an examination of the quantities of email that users send, and can suggest ways to deal with overload. Some of the lingering questions in email can be found in the sidebar.

We describe our study of the email records of a number of information workers at a high technology company. We have taken advantage of increasingly-powerful scripting tools on mail software to allow client-side scripts to generate automated response.

We found that an automated approach allowed us to easily learn a great deal of information from email boxes: we were able to analyze threading patterns, and to see some of the indicators of mail overload.

CURRENT QUESTIONS IN EMAIL RESEARCH
How much email do users receive? How much do they keep?

Does keeping more email lead to a feeling of being overloaded?

How do users respond when they suffer from overload?

What can relieve overload?

What role can threading serve in organizing email discussion?

Although email is implicitly private, and thus hard to study, we were able to provide user-controlled tools that allowed anonymous participation, and were thus able to find a number of users who were able to help us out.

This technique is valuable because it allowed us to directly access the current state of mail inboxes There were also unexpected side effects: some users continued to use the tool even after the research ended, as they found it was a valuable method of reflection.

THE TEMPTATION OF ELECTRONIC TRAILS
This paper outlines some of the important design decisions behind the data collection tool in order to collect useful data and gain user acceptance, and presents the results we were able to obtain with it.

Electronic interactions have the nice property that they can be programmed to leave behind a trail, and the trail can be studied later. The advantages of automation are dramatic. Use can be measured without video cameras or lengthy interviews. Different sites can be compared to each other using a common set of measures, and samples can be taken

at a large scale. Last, automation allows the user to immediately receive feedback from the experiment.

We believe that this ability to study email artifacts will allow us to directly review email behavior, and therefore to approach common problems, as well as to address, on a larger scale, some of the issues raised in the other studies.

HANDLING OVERLOAD

We wish to study how users handle their flows of email, their filing and sorting methods, and their successes and failures. Our primary interest is in studying email overload, the perpetual difficulty that email boxes seem to grow more quickly than users can keep up with. Many studies, both recent [1, 3, 16], and older [12] have suggested overload is the top problem facing email users today: [3], for example, refers to email as a "serial killer" app. A recent survey, done by the Lotus Corporation (described in [15]), found three major failings of email. First, email users complained that they are *overwhelmed* by the volume of email they receive. Second, users complained that useful information *gets lost*, or can't be found when they need it. Last, email creates an imbalanced expectation of responsiveness—the receiver of a message cannot respond as quickly as the sender can send a message out.

Research papers have been hesitant to suggest remedies for overload. While several filtering [6] possibilities are available, informal preparatory interviews suggested that few if any users, even the technically proficient ones, use automatic filtering extensively. Balter [1] suggests easy-to-use full-text-search on a growing corpus of messages. Full text is available on our research platform, Lotus Notes; however, it does not resolve overload problems.

One possible way to reduce the impact of overwhelming message flow might be through threaded conversation. With online conversation, the direction of a threaded conversation can be explicit (see Figure 1).

Threaded Conversation

Threaded conversation has been a successful technique for managing traffic on Usenet groups. For those systems, the conversational *thread* is the basic unit of online conversation. Early newsgroup software supported the notion of explicitly tracking threads, a set of related comments on a single idea. By creating a new message as a reply to a previous one, some messages could be tagged as a response to some previous comment. (Obviously, some messages are not in reply, as they start new threads). Today, threading is common in a large number of discussion fora, and has even been extended to synchronous chat [11].

Many Usenet readers allow readers to view their interaction as a set of threads. However, Usenet groups are a shared online space. In contrast, email is a private activity, located in the ephemeral space shared only by those members who participate in a particular conversation.

Threading generally refers to conversational elements tied together chronologically: a message gets a response, and that response is responded to. For the purposes of this project, we considered only a subset of all possible threads. Rather than try to link conversations semantically, we only examined header information. Messages that are tagged as replies to each other, and that arrive in order, or messages that share a title and a group of participants, are labeled as threaded. We were willing to call a pair of messages—a note and a single response—a thread, as there are certainly advantages to clustering a question with its answer.

Explicitly drawing out threading is a way to understand how extensive a conversation has become, to observe turn taking, and to make sorting easier: the messages on a single thread are likely to be on a similar topic. A few email programs support reading mail in a conversation view. These views typically order series of messages with the same title based on when the first message in sequence arrived, making this a difficult way to monitor their inbox as new messages appear in older sections of the inbox – often out of view. This is inconsistent with the notion of an inbox as a list of current events[1], and informal interviews have suggested that many users turn away from this presentation.

However, there are difficulties that accompany thread propagation. Conversational analysis work has highlighted the importance of taking turns in communication [4]. Online, these turns can be confusing when new members are brought suddenly into a conversation [7] highlights users' confusion at being forwarded one message in a part way into a long series, and thus thrust mid-way into a long thread.

Not all messages can, or should be, threaded. Periodic announcements, notifications, and schedules are often "one-off" messages, not meant as a part of a conversation.

We believe, then, that threads are most useful when their participants *do not* feel thrust into the middle of a thread, but rather generally participate from the start. Similarly, if a user must be added, it would be desirable to be able to forward an entire thread at once, as context, to new users.

METHODOLOGY

It can be very hard to study email. Email is fundamentally private. We felt it was important to, above all else in the study, respect that privacy for our participants. Automated techniques are dangerously prone to being invasive, so we engineered our system to keep users informed. Participants can often be resistant to having their work habits studied in too much detail, especially with something both so personal and time consuming as email.

Further, studying a technology like email can be a slow process. Despite the technology's ubiquity, many of the people in the workgroup environments we're interested in approach their email as a series of interruptions [3], rather than waiting for a specific time to check mail.

Ethical Difficulties and User Concerns

We needed to respond to a number of important user concerns in our project. We wanted to respect the privacy

of our subjects, and therefore to not inspect much of their email. Similarly, much internal corporate mail is confidential to the company. Rather than risk gathering sensitive information, we decided to stick only to header information–and even this presented challenges in how to protect their privacy.

We encrypted all personally-identifiable fields with the standard MD-5 hash algorithm. Because hashing produces the same value for each name across all mailboxes, we could track instances of messages and authors across all mail without compromising our participants' anonymity.

This allowed us to both match threads between users, by matching time stamp, author, and subject; and analyze social networks, by matching "to" and "from" lists of different authors.

We sent out the applet to our study participants, who then had the choice of whether to run it. Although the information could have been extracted from the central server, we wanted to ensure consent.

Although continually monitoring mailboxes would give us a powerful data set, we were unable to build something that could be run for a long period of time without interfering with the user's usual work. Therefore, we collected only a single, individual snapshot of a mailbox. The snapshot has certain myopias. We were unable to gain a true picture of how much email a user received, instead we got a fairly accurate image of the amount of email that they kept per day. This still seems to be an important number: it measures the amount of email that gets stored by the user.

This did present us with a wide range of results—some users seemed to delete most of their daily mail, leaving us only crumbs; others carefully archive everything.

Procedure

We constructed an algorithm to reconstruct threads within a mailbox, based on internal evidence from messages including date of delivery and title similarity. For each message that the user had sent or received in the last 60 days and had kept in their active mail file, we sorted it into its appropriate thread. The specific attributes we collected can be found in Table 1. In addition to data on saved mail, the application included a survey with a dozen questions about online expertise, comfort with technology, and email overload.

Table 1: Data collected from participants

For the mailbox:
The range of dates of messages
Count of outgoing messages over the last day, week, month, and last two through six months
For each folder, the count of messages in the folder and the range of dates in the folder
For the last two months:
For each message in a thread, the sender*, the list of participants*, the title*, the date, the message size, and the number of attachments.
The count of messages exchanged with each correspondent
* Indicates information that was scrambled

In general, the application required about five minutes of supervised execution time (during which most subjects filled out the online survey) and an additional twenty minutes of unsupervised time (during which many subjects went to lunch). We filed the results that were sent to us into a database, separating it by individual threads and users.

Our project came from a particularly privileged perspective: we were able to distribute it among a sample of participants who all used the same version of the same mail program. This should be possible in other environments, as technology support groups often attempt to provide a uniform computing environment. Certainly, other mail clients can be scripted to deliver similar information.

Detailed study of one inbox

In order to confirm these results, one of us also engaged in a series of detailed studies of a single inbox. These reviews were used to confirm that our algorithm had an extremely low rate of mis-threads and under-threaded conversation. In general, it successfully identified messages that had legitimate responses. We also found that very few threads included messages that were not among the most important messages for that day.

FINDINGS

We distributed the application embedded within an email message to a selection of users at a large technology firm, and asked them to forward the application to friends of theirs. All users had previously indicated an interest in new email technology; most were researchers. All participants were located at the main office of the company, which was located in three nearby buildings.

74 users completed the survey and sent in their data files. After reviewing the results, we discarded several obvious technical errors where the tool had failed to read the mail file. For some questions, we also removed responses from recent hires, as they had only been in the company a few months, and had not yet have formed stable mail strategies. In the end, we have N=57 usable responses.

Hypothesis 1: *Users will vary in their organizational schemes; however, there will be few important changes since Whittaker and Sidner [16].*

We first wished to examine whether our results were consistent with [16] or Ducheneaut and Bellotti [3]. These studies discussed the challenges of personal archiving and filing. Folders are a principal way of accomplishing this. With some clients, it is possible to merely use full-text search on email inboxes, possibly obviating the need for folders. [16] observed that creating a folder is cognitively

difficult, but that users merely allowing their inboxes to accumulate may find the sheer size of their inboxes frustrating. [3] suggests that effective users will have shallow folder hierarchies.

Accepted Hypothesis 1. Whittaker discusses three major organizational schemes: "non- filers", who tend to use text search to manipulate their massive inboxes; "frequent filers", who constantly work to keep their inbox short, and "spring cleaners", who keep periodically sweep their inboxes to clean out detritus.

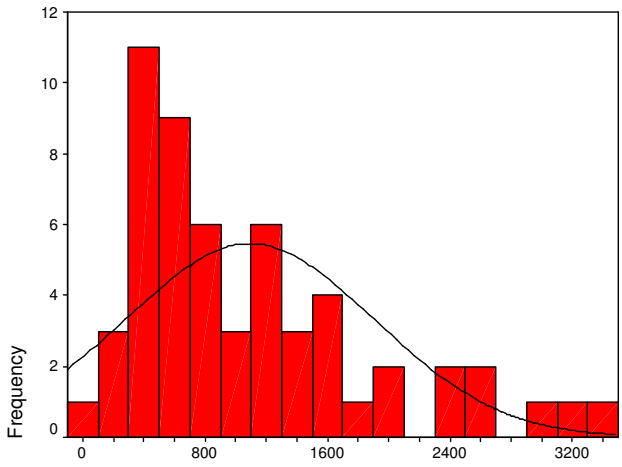
Because our dataset is a snapshot, it was difficult to distinguish between non-filers and frequent filers, and we found no particularly clear distinction between these categories. The results contained information on how much mail, in total, had been kept over the last six months, and across all users, a fairly consistent amount of mail was kept over time. There was one notable exception. One participant was a particularly thorough cleaner, the vast majority of whose entire mailbox seemed to date back less than one week: this person's mailbox showed a ruthless recent pruning, assuming the flow of incoming mail was fairly constant.

Our study suggest that the once well-defined strategies that Whittaker observed are becoming less distinct: our data[1] shows a continuity of approaches.

There were a few exceptional users who maintained particularly deep hierarchies of folders. One, for example, maintained 937 separate folders (half of which contained under five messages), storing 11690 messages. Most users kept somewhere between 100 and 6000 messages in these folders, falling smoothly on a (logarithmic-scaled) curve. Despite Balter's [1] notion that most efficient users should keep shallow hierarchies with no more than 25 folders, we found that only 21% of users kept a flat strategy. 40% had a two-layer structure; the remainder had three or more layers. Users kept a median value of 73 folders; 15% of our sample had over 100 folders. (Values ranged from 3 to 937; mean=73.11, σ=133.45). We also asked users of their opinion of their mail flow. Did they feel behind on their email, or interrupted by email? We hoped that this would give us an indicator of a user's feeling of overload. In fact, we found a positive correlation between users who reported that they "feel behind" in their email, and who kept many messages in folders ($r=.388$, $p<.003$). Feeling behind also correlated ($r=.457$, $p < .001$) with keeping more mail during the previous two months. This may be consistent with Whittaker & Sidner's conclusion that filing often fails as mail flow increases.

[1] Due to technical difficulties, we were able to count neither the number of messages in just the inbox, nor the total messages in the mailbox over all time. However, we were able to figure out the number of messages in the last six months—a number which is not directly comparable with the number of messages in folders.

Figure 2: Histogram, number of users against the total number of messages touched in the last two months across all folders.



We tentatively accept Hypothesis 1.

Hypothesis 2: *Locality is not overcome: most connections remain dominantly local, within the same team or office.*

We also wanted to learn how much communication was concentrated within the work group. Wellman [14], as well as other sources [8], suggest that most online interaction occurs locally, supplementing interoffice memos, visits, and phone calls. Despite a world-spanning technology, most users send email over short distances, within their offices, work groups, or companies. Are these conversations, in fact, dominantly local, or remote?

Accepted Hypothesis 2: *Locality is not overcome: most connections remain dominantly local, within the same team or office.*

Because the survey was anonymous, we don't have detailed information about the locations of participants' desks. However, we did categorize messages by coming from "inside the company" (that is, by sharing the same domain name as the user), from "inside the work group," "from the manager," or "outside the company". Since the work-group was self-identified (and then cross-referenced through a corporate white pages), there might be some inaccuracy of in vs. out of group categorization.

Most users communicated mainly with other people inside the company, and largely within their groups. We found that half of our sample population received just 25% of their email from out of the company, while nine of ten users got under half of their email from outside the company.

Half of our users got a quarter or more of their mail from their work group. There was a slight negative correlation (r=-.314, p<.018) between the amount of mail that comes from within the group, and how much mail is kept: this may suggest that mail in groups is more ephemeral.

Hypothesis 3: *Threading is prevalent, and connects a substantial portion of inboxes; therefore, threading could be a good way to summarize online conversation.*

Accepted Hypothesis 3. We had good data to examine threading in email. Our users kept between 85 and 3318 messages in the two-month sample, showing a tremendous range of activity levels. Most fell to the lower half of that range: the median was 821 messages. Table 1 shows that one-third of messages were threaded, falling into 94 conversations.

Threads weren't, by and large, long. 56% of all the threads we reviewed were only two nodes—a call and a response; 87% of threads are 4 messages or less. However, all users still had a fairly large number of threads that were noticeably longer.

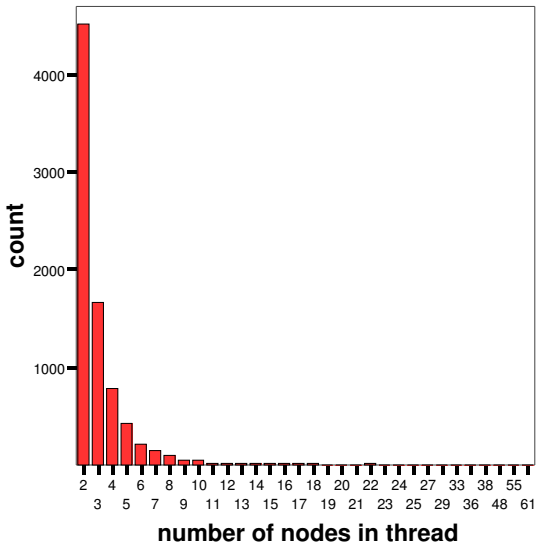
We calculated how much smaller inbox volume could be if only there was only one entry per thread. Collapsing the threads to single entries would lead to mailbox being reduced by about 23%, this reduction in perceived volume could help users deal with their feeling overwhelmed: not only would the inbox appear smaller, but messages would be placed in context.

Hypothesis 4*: In general, online conversations will tend to be fairly easy to use. For most threads,* ***(a)*** *the number of participants will remain somewhat constant, and* ***(b)*** *responses to messages will be fairly timely*

For a threading to be a useful tool for reducing mail overload, two things must be true of message threading. The first is that threads represent a sufficiently large portion of an inbox, so that it will be worthwhile to connect messages by threading. The second is that the conversation structure is sufficiently consistent in participation, and time range that it would make sense for two users to discuss, and perhaps exchange or forward conversation threads

(accepted) a: We attempted to test this by examining all threads in the database. Turnaround was very good: most messages were answered in a day. 70% of all two-part exchanges—a message and a response—were within one day. If a response didn't come the next day, it would arrive the day following (78%). 95% were covered within a week. We also examined messages that had gotten more than one reply. For those threads, we examined the "maximum time to wait"—that is, the longest chronological gap between two messages. Again, responses were quite prompt. 51% of these conversations waited no more than one day between messages, 60% waited no more than two days. By the time two weeks had passed, 90% of messages had replies.

Figure 3: Histogram of nodes in threads, across all users.



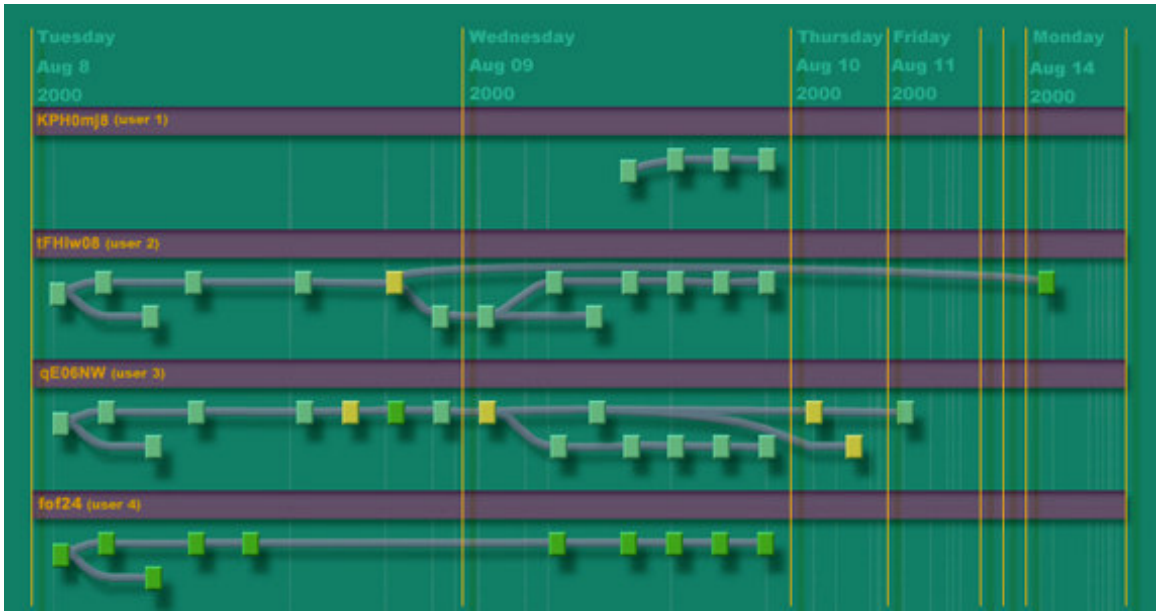
(rejected) b**:** To look at the consistency of participation, we looked at how many people were added, or removed, from a thread over time. Although this figure is uncertain—for example, people can respond from a different address—we applied a number of rather conservative estimates to our data. We found that over a third (36%)of the threads experienced at least one shift during the duration of the conversation. This was rather high: apparently, populations on threads are not particularly constant. This reflects poorly, too, on the next hypothesis.

Hypothesis 5*: Users will want to maintain conversational*

Table 2: Statistics for message flow for the last two months. (N=57)

	Incoming	Outgoing	Total messages	Total Threaded Messages	Total Threads	Percentage of messages threaded	New total, if threaded	Savings, if threaded
Mean	181.07	909.71	1090.79	401.98	129.88	34.6%	818.68	23%
Median	138.00	670.50	821.00	283.00	94.50	35.8%	629.50	23%
Std. Deviation	177.44	703.23	816.45	356.25	111.53	9.7%	599.14	7.2%
Minimum	11	68	85	26	10	7%	69	4%
Maximum	1010	2817	3318	1501	463	49%	2402	35%

Figure 4: Four views of the same thread.



example. This thread clearly illustrates that there *are* radically different views of the same information. While one participant saw a conversation of four messages in an evening, another saw ten messages two days. Two others saw a far longer, and branchier, structure of eighteen messages lasting over a day.[2]

flow, and will actually be more *able than Usenet groups to keep threads fairly linear.*

Some research [9] has suggested that conversations on Usenet newsgroups become rapidly multi-threaded and change topics frequently. In email, this may lead to threads' becoming hard to understand, as users need to juggle multiple conversations at once. This has been a problem, and often leads to fragmentation of topics. However, the email infrastructure is less liable to posting delays, especially across local groups, and the participants are likely to know each other. This might lead to more unified conversations. (For example, the conversation in Figure 1 is only slightly branchy, as it shows a single branch off of an otherwise linear conversation).

Rejected Hypothesis 4. We tracked the branchiness of threads and found that two-thirds of online conversations with more than two messages were 'branchy'—that is, didn't display a linear structure. Branchiness was positively correlated (**r** = .576, p<.03) with the number of nodes in the conversation. Most threads, however, are short.

One of the unique capabilities of our approach was the ability to track conversations between several different participants. By looking for messages with the same date, subject, and recipient list, we were able to watch participation shifts on several different threads. We were fortunate that our data contained some threads that were shared between participants. Figure 3, for example, shows one thread simultaneously observed by four of our participants (This visualization is related to those discussed in [10]).

The visualization shows four different users' views of a single conversation thread. These four views are the sets of messages that each user received during the course of the conversation. User 1 was added substantially later, for

This visualization gives a sense for the varied ways that threads could be seen by different users. In this image, we see a tree of messages, stretching over time from left to right. Color codes are based upon the relationship of the participant to the messages – for example in the second user's data a green message would have been from someone in their group. name. Note that the first user ("KPH0mj8'") sees only a part of the conversation, but misses out on both the introductory context and the following messages. In contrast, "tFHIw08" (user 2) and "qE06NW" (user 3) both share much—but not all—of the conversation.. What is crucial to understand is that multiple users are seeing the "same" conversation, but with radical differences in perspective. There is no single canonical thread; rather, there is a continued conversation that adds and removes individual users.

Note that one message is missing on August 8 from user 2, a message which is highlighted in yellow for user 3. Interestingly, that indicates a point where user 3 intentionally excluded user 2 from a followup to the thread, and yet the next day includes user 2 on another message they sent out in this thread. These sorts of shared views were not uncommon: among the twenty-five thousand messages that were parts of threads, we found that almost two thousand of them were seen by several different participants. These multiple views help us construct a single view that highlights changes that occur to the thread.

[2] We have reviewed the header information for all the messages. The differences in views are *not* an artifact of the 'snapshot' technique, but rather reflect differences in carbon-copy lists. For the purposes of this discussion, the colors of nodes are arbitrary.

DESIGN IMPLICATIONS

These are real issues for email. Five years after Whittaker & Sidner [16]—and a decade after *Seven Steps to Better Email* [12]—users are increasingly overwhelmed with email. Our sample population kept 13 (but some as many as 60) messages a day for as long as six months. Just as previous studies have shown slowly growing numbers, there's no reason to believe that this number won't continue to increase over time.

The issues we raised at the start of this paper are still relevant. 33% of our respondents reported that they "often" felt they had trouble finding messages they had once had; the data under hypothesis one suggests that users are overwhelmed by the incoming messages, too.

Threading seems to be a useful way to understand and review large conversations quickly. Based on this work, a design group has developed a prototype, figure 5. Our data under hypothesis three suggests that email inboxes—if viewed compressed, with one entry per thread—could drop in size as much as a third for this group. It would also help separate out informative mail—scheduling, mass messages, and other assorted mail—from the important conversational communication: the conversation forms an easy to identify thread, while the announcements often stand alone.

It is clear that email based conversation is, in some ways, very different from the posted conversations on message boards. The population of conversation shifts radically, and moves between reading groups. In contrast, a discussion on a public board is always open to the same group of readers. There may be side-conversations, but those are based on an individual response, and not a large-scale redirection. Indeed, it seems that in all but the most trivial of conversations it would be difficult to track down the "canonical" set of messages that form the core of the conversation—different readers have different perspectives.

Potentially Powerful Study Tool

Visualizing threads makes for a useful tool. Some of our participants who received a custom tool that allowed them to see their threads, enjoyed having this new view of their email. Not long after we distributed the tool, some users contacted us, asking to turn off the encryption functions so they could see which messages were connected to other via our threading. They had downloaded the application and checked it periodically. We are also exploring social network implications of the tool, as interpersonal contact becomes explicit through this technology and the visualizations.

In the off-line world, turn-taking has been examined by [4] in a labor-intensive study of hundreds of business meetings. He found strict rules for interaction that governed who could speak in what order, and thus found a set of rules for participation shifts in conversation. This sort of tool could be a valuable online counterpart.

Figure 5: A prototype threading mailbox view.

Unusual Circumstances

Lotus Notes offers certain features, like full text search, that make some strategies work well. It is possible that in a mailer that didn't have some of these features, the system—and the users—would respond differently. However, since this study showed little evidence that those features really make a crucial difference, we aren't concerned.

One limit in our study is the lack of any data on the contents of the email messages. We intentionally moved away from content information, choosing instead to emphasize the external, formal attributes of messages. Content analysis is a complex process, raises privacy concerns, and requires more study than we were able to arrange. Future studies may, however, want to start to examine the specific roles of individuals, and to isolate users from out of the group.

Because of the limitations of the snapshot methodology, we would be particularly interested to see future studies that took repeated snapshots. This would allow the system to examine messages that were removed from archives over time.

CONCLUSIONS

Our results suggest that users are interested in tools that help them collect, introspect, and work with their mail easily. We have found that automation is a powerful way of collecting data, and that it can be done, even in this private context, in a way that does not compromise confidentiality or user comfort.

We have found continued evidence that users struggle to keep up with their mail. They keep hundreds of folders, and dozens of messages per day. Threading could be an effective personal archiving tool, and a useful personal view of an inbox. However, as there is no canonical thread, it could be difficult to isolate the important set of messages to be forwarded.

We are less inclined to believe that the answer is more advanced filters, instead we feel that methods of organizing and summarizing data are the needed changes. It is important to remember that collaboration is distributed, so client-side features cannot be too radically different from today's email or they risk becoming incompatible. Precedent has shown that the best-accepted tools are backward compatible. However because teams seem to work heavily internally (Hypothesis 2), it is possible to have advanced email tools that provide features to support their internal communication.

A variety of data sets can be gathered with automated collection. Designers interested in understanding how their systems are used in the medium term should attempt automated collection to watch what features are adopted over time. Researchers trying to understand turn taking, group decision processes, or business interaction may also gain insight from these archives of communication. Content analysis may lead us to information about how decisions are made or who know what and whom – all of which are important knowledge management issues.

Our approach has continued to be successful for us. We've been able to use this as an infrastructure for a more detailed, non-anonymous follow-up study.

ACKNOWLEDGMENTS

We thank our study participants, who were kind enough to provide us unfettered access to a truly valuable resource, and Michael Muller, for extensive comments on a previous version of this text

REFERENCES

1. Balter, O., "A keystroke-level analysis of e-mail message organization." in *Proceedings of CHI 2000*, (2000), 105-112.

2. Bellotti, V. and Smith, I., "Informing the Design of an Information Management System with Iterative Fieldwork." in *Proceedings of DIS 2000*, (Brooklyn, NY, 2000), 227-237.

3. Ducheneaut, N. and Bellotti, V. "E-Mail as Habitat" *Interactions*, 2001, 30-38.

4. Gibson, D.R. "Taking Turns and Talking Ties: Conversational Sequences in Business Meetings" *Sociology Department*, Columbia University, New York, 1999, 222.

5. Mackay, W.E. "Diversity in the Use of Electronic Mail: A Preliminary Inquiry." *ACM Transactions on Office Information Systems*, *6* (4). 380-397.

6. Maes, P. "Agents that reduce work and information overload". *Communications of the ACM*, *37*. 31-40.

7. Markus, M.L. "Finding a Happy Medium: Explaining the Negative Effects of Electronic Communication on Social Life at Work." *Transactions on Office Information Systems*, *12* (2). 119-149.

8. Olson, G.M. and Olson, J.S. Distance Matters. *Human-Computer Interaction*, *15* (2/3). 139-178.

9. Osborne, L.N. Topic Development in USENET Newsgroups. "*Journal of the American Society for Information Science"*, *49* (11). 1010-1016.

10. Rohall, S., Gruen, D., Moody, P. and Kellerman, S. "Email Visualizations to Aid Communication." *IEEE Symposium on Information Visualization 2001*.

11. Smith, M., Cadiz, J.J. and Burkhalter, B. "Conversation Trees and Threaded Chats." *Proceedings of CSCW 2000*.

12. Terry, D.B. "7 Steps to a Better Mail System," Xerox Corporation, Palo Alto Research Center, Palo Alto, 1990.

13. The User Group. "The User Group: Three Hours a Day Wasted on Irrelevant Emails; Research from the User Group Shows Information Overload is a Threat to the Economy", M2 Communications, London, 2000.

14. Wellman, B. "Ties and Bonds". *Connections*, *22* (1). 12-18.

15. Werbach, K. "The Future of Email". *Release 1.0*, *19* (6). 24.

16. Whittaker, S. and Sidner, C., "Email overload: exploring personal information management of email." in *Human factors in computing systems*, (Vancouver, Canada, 1996), ACM, 276-283.