

Using Social Sorting to Enhance Email Management

Danyel Fisher, Bernie Hogan, AJ Brush, Marc Smith, Andy Jacobs

Microsoft Research

Redmond, Washington

{danyelf, ajbrush, masmith, andyj}@microsoft.com; bernie.hogan@utoronto.ca

ABSTRACT

We evaluate the use and potential benefits of social sorting as a technique for managing email. We present SNARF, the Social Network and Relationship Finder, a tool which filters and sorts email based on the type of message and the user's past history with an email correspondent, a process we refer to as social sorting. This tool also combines several contemporary techniques for email selection and review, including thread management and multiple simultaneous views. We report on the results of a 4 week field study (N=574), and describe the user behavior that emerged around the tool. Additionally, we present a scale of email overflow to assess users' experience with email generally. While our application did not significantly reduce experience of email overflow, we conclude leveraging past email behavior can be of value to email users for both time-critical triage and peripheral awareness.

Author Keywords

Email overflow, email triage, social metadata, social sorting

ACM Classification Keywords

H.1.2.a Human factors, H.4.3.c Electronic mail

INTRODUCTION

Email is a critical communication medium for many organizations. As employees seek to separate relevant messages from a swath of less relevant messages, they often report feeling overwhelmed [8]. This is particularly the case for individuals with many unread messages and little time. Assisting users in working with email has been addressed in many different ways in the HCI field from innovative interfaces [10] to prioritizing systems [6]. In this paper, we discuss the design and deployment of SNARF (the Social Network and Relationship Finder), an application which employs social metadata and social sorting to reorganize the

HCIC Workshop, 2006

display of incoming email. By ranking authors based on their history of interaction with the user, SNARF supports more effective decisions about which emails are worth reading at a particular time.

Triage is referred to in [8] as the task of sifting through unhandled email and selecting particular messages to handle before others. Many individuals triage first thing in the morning, or after being away from email for a period of time. During these triage periods, users often suffer from having more email than time available; we designed SNARF to assist with these time-critical periods.

Neustaedter et al. [8] report that in those time-stressed periods, many users choose to employ multi-pass strategies to read their email, skimming through to select ones that are particularly easy to handle or particularly important. Additional passes—to choose the next most important message—follow as time permits. Despite the clear inefficiencies of this method, requiring many decisions about which messages are most relevant, it is a method required for triage by most contemporary email tools, which tend to present email in order of receipt, rather than relevance.

In contrast, SNARF presents users with alternative views of their email organized around people using *social sorting*, an ordering based on the history of communication between the recipient and sender. By bringing messages from frequent correspondents to the fore, we believe SNARF allows users to more easily locate relevant messages especially when pressed for time. As such, it reflects the advantages of a multi-pass strategy without requiring users to repeatedly scan their entire inbox.

SNARF can also be used after triage periods to provide a peripheral awareness of new email, highlighting messages from frequent correspondents and those sent directly to the user. Instead of notifications for all new email, SNARF allows users to be passively aware of a select subset of messages as they trickle in during the workday.

To assess the degree to which social sorting is a valuable approach to help people manage their email, we deployed SNARF in a field study in our organization. We evaluated SNARF and the use of social sorting on the following criteria:

- **Usage:** Which features of SNARF were rated as most useful? Which were actually used the most? Did these features assist email triage?
- **Awareness:** Does SNARF help users to maintain awareness of incoming mail?
- **Overflow:** Is continued use of SNARF associated with feeling less overwhelmed by incoming mail?
- **Retention:** Ultimately, which features of SNARF were associated with continued use, and how did these features employ social sorting to that end?

In our field study of 574 users over 4 weeks, we found that participants who kept using SNARF employed triage-oriented views extensively, enjoyed viewing messages by threads, and ‘warmed up’ to the idea of SNARF as a tool for awareness. In the rest of the paper we first discuss related work then describe the SNARF field study and our results. We close with a discussion of the implications of our work.

RELATED WORK

Many projects have attempted to reduce email overflow. Whittaker and Sidner [14] began a thread of research in noting that email is used for a variety of personal information management tasks in addition to communication. They refer to the many different aspects of email—maintaining a calendar, keeping contacts, and driving a working memory—as “email overload.”

The concept of handling email in different phases has been addressed by several different researchers. Takkinen and Shahmehri [11] identify three phases of handling email: “busy mode”, in which users look for only important messages, “cool mode,” in which users use email as it comes in, and “curious mode,” for organizing and catching up on messages. Venolia et al. [13] formalize this structure in terms of distinct tasks, with temporal boundaries: “triage” roughly corresponds to the busy mode, “flow” to cool, and “archiving” and “retrieving” to curious.

Bellotti et al. [2] identified task management as a major role in email, and suggested that feelings of overload may be linked to incomplete tasks within an email store. They find that reported overload corresponds poorly to the number of incoming email messages; instead, they roughly quantify overload as relating to the number of currently-active task “threads”, and the length of time since the last message in those threads.

Nardi et al. [7] suggest that contact management takes a substantial amount of effort, and propose a socially-based interface that shows information oriented around people. In SNARF, we adapt their notion of orienting an interface around participants, although in a very different way.

Other projects have attempted to more generally relieve email stress. Priorities [6], for example, uses machine learning techniques to recommend which message(s) to read next. IBM’s Remail [10] prototype did not focus on

any particular form of task or contact management; rather, it attempts to provide an overview of email from different approaches: it displays conversation threads, displays groups of people by their organizational structure, and connects calendar information to email messages.

Social Sorting of Email

The notion of applying social sorting to online data collections may be read broadly as a form of social navigation. It has been used in particular by Fiore et al. [4] to examine ordering of Usenet news messages. That study found that ordering messages by the people involved in them provided an effective mechanism for highlighting valuable content.

Fisher and Dourish [5] explored another way of placing email within a social context by describing social roles that are visible in social networks of email messages.

Both of these cases present interaction histories from a novel perspective. Both of them are oriented toward retroactive analysis of archives; by contrast, our approach uses archival information to help users make decisions in real time.

DESIGN OF THE PROTOTYPE

The design of SNARF has been discussed in previous work [9]; here, we limit ourselves to a brief overview of its most salient features and concepts in order to explain the work in the field study.

SNARF uses information about past email behavior to display a user’s email from a socially sorted perspective in three main ways:

- Each pane of the main SNARF window presents one *view* showing a subset of available email correspondents¹.
- The senders of messages are socially sorted within each view by the degree to which the user is related to the sender by previous volumes of email exchanged.
- Messages are clustered together and linked to the sender of the message, in order to present a social perspective on the email. Each correspondent is accompanied by the number of unread messages associated with them.

The SNARF Main Window

Figure 1 shows the default configuration of SNARF. The top view, “Unread To/CC me,” shows correspondents whose messages that have explicitly included the user in the To or CC line, and remain unread. The middle view, “Unread

¹ An email address can refer to a person, a distribution list, or an automated generator. For simplicity, we refer to all senders and receivers as correspondents.

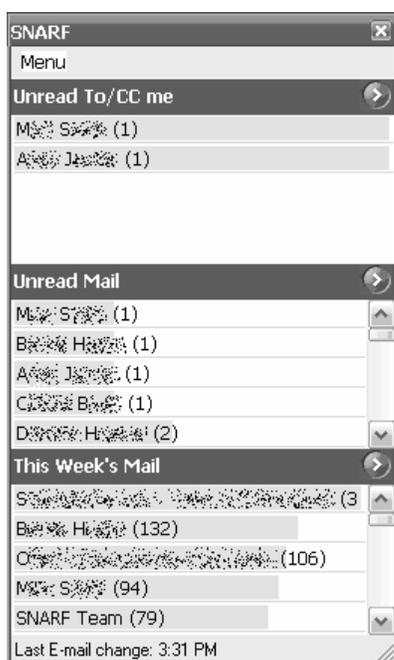


Figure 1. The SNARF Main Window, showing three views.

Mail,” shows *all* correspondents with unread messages that the user received (and thus may have gotten as a BCC or through a mailing list). The bottom view, “This Week’s Mail,” shows all correspondents from whom the user has received messages in the last week, read or unread. Note that the panes are not mutually exclusive; a particular message or person may appear in multiple panes if they match the criteria. Yellow highlights are provided as cues that convey the relative amount of messages from the correspondent within a view.

Conceptually, we separate views into two types. *Personal views* filter for messages that explicitly mention the user, such as “Unread To/CC me.” *Aggregate views* show collections of messages that may not mention the user, such as all unread messages sent to lists (“Unread lists”). Personal views generally contains shorter lists of people with fewer messages apiece; the latter type generally contains more people and more messages.

There are two processes in action here: selection and sorting. First, correspondents are selected for inclusion into a view by the how they addressed the email they sent (either directly or via a list). Second, correspondents are sorted by the number of times that they have interacted with the user in the (configurable) past period. By default, the list of correspondents is ordered by the number of messages that the user has sent to this correspondent. Thus, frequent and recent correspondents bubble to the top, while infrequent correspondents sink to the bottom.

For example, in Figure 1, the user has sent more email messages to *M--- S---* than to *A--- J---*; *M--- S---* is therefore sorted higher on the display than *A--- J---*. Users are able to configure the selection and sorting of these views or add additional views. The configuration options available on this display are discussed in more detail in [9].

It is worth noting that the original design for SNARF [9] showed only one view at a time; users would re-configure the display to match their current task. However, as the design evolved it became apparent that having *more* types of sorts with *fewer* people on the top was important. This reflects the difference between the relevance of a particular sender, and a sender’s message. *M--- S---* may be very important to me, but messages from *M--- S---* sent only to me are often qualitatively different from messages sent by *M--- S---* to a mailing list we both subscribe to. The current display, which can accommodate three views of ten or so names without scrolling on a conventional screen, reflects this design.

These views, like all aspects of SNARF, respond in real time to changes in email state. For example, if a message from a sender to the user is marked as read, the “Unread Messages” view is updated—it will no longer show that sender if no more unread messages remain.

When the user double-clicks on an entry showing just one message, SNARF opens the message in Outlook. If the name is associated with more than one message—for example, there is more than one unread message from that person—then SNARF presents a list of these messages in a second window. Each of these messages is shown with its author, subject, and date; they can each be individually opened in Outlook.

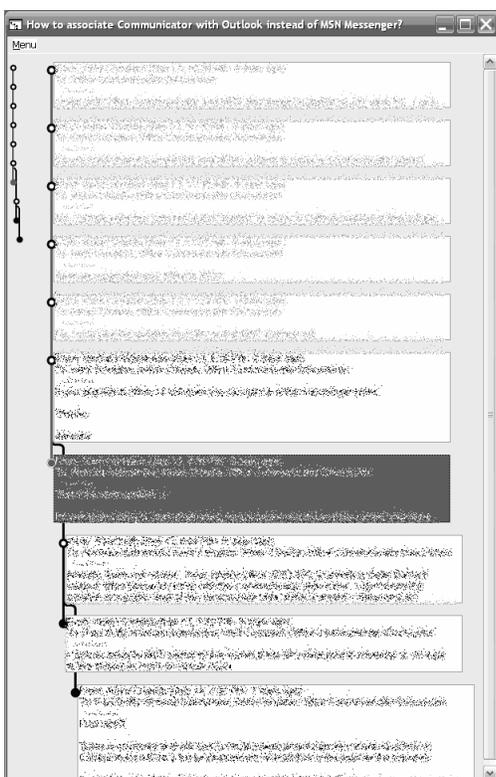


Figure 2. Thread Viewer. Unread messages are dark, while read messages are gray.

Thread Window

Instead of opening messages in Outlook, the Thread Window acts as an alternative message display for reading messages in context. The Thread Window allows a user to see the contents of all the messages in a selected thread in one window, using the display techniques found in Grand Central [12]. Users can quickly read an entire thread, marking messages as “read” as they progress downward. Figure 2 shows the thread view for a ten-message-long thread. This places each message into its appropriate conversational context. Within this display, users can conveniently read the entire conversation at once. (This display also allows users to conveniently delete the entire conversation, too.)

Other Design Aspects

Because the main display of SNARF is fairly narrow and SNARF reflects changing email as it happens, SNARF can function as an awareness tool. With this use, SNARF sits on the side of the screen and allows the user to periodically glance at it to see if anyone interesting has emerged at the top of a view. This differs from conventional notification systems: new email from socially relevant correspondents is indicated by the presence and ranking of the sender in the SNARF display, but doesn’t cause a sound or a distracting visible change. [3] criticizes notifications as being distracting, even when they are not attended to. But unlike the pop-up notifications that many clients provide, SNARF

does not require a quick reaction to click on the pop-up for a new item of interest.

SNARF can also be used as a mailing list reader. Any correspondent—that is, a message sender or receiver—can be manually labeled as a “list.” SNARF provide a “Lists” view that displays only lists and sorts them by the number of messages the user has sent to the list. This allows users to keep up with relevant lists: to view them in one place, and to read their messages in threads.

FIELD STUDY METHOD

In July 2005, we sent out an invitation to use SNARF to 1713 people who had registered their interest at an internal demonstration of the system. We also sent out an invitation to two high-volume mailing lists of employees at our company that specifically target people with interest in advanced and speculative projects. We also know that many of these employees sent copies of the message to other people who may not have received the original invitation.

In total, 574 people ran SNARF at least once within the first two weeks after it was made available. While the project remains ongoing—and more people continue to use it—we restrict the following evaluation to this initial set of users; this allowed us to ensure that everyone in our sample set had access to SNARF for a full four weeks.

Pre and Post-Survey. When SNARF opened, users were presented with a dialog asking them to complete an optional pre-survey related to their experiences with email. We followed up with a reminder message several days later. The survey followed the outlines of [8]; however, it added more detailed information about email habits within the user’s current client. The survey included questions asking users how long they thought they spent triaging email, how much of their email they read, and their experience of stress associated with handling email. A total of 292 people completed the pre-survey, a response rate of 51%.

Four weeks after the last user in our sample started SNARF, we sent out a post-survey. Users were asked to respond regardless of whether they had filled out the pre-survey or had continued to use SNARF after its initial installation. Participants that filled out the post-survey were entered into a lottery for gift certificates. 161 persons filled out the post-survey (response rate of 28%). Of these 161, 122 had also filled out the pre-survey. Combining the pre and post surveys, there were 304 unique SNARF survey respondents.

They identified their job role as largely program managers (23%), developers (17%) and consultants (13%); sales (10%) and software testers (9%) made up the bulk of the remainder. The population was overwhelmingly male (92%).

Post-survey respondents approximated the overall SNARF sample in terms of usage and retention. 4% reported that they could not get SNARF to work, 44% ‘tried it once or

Eight-Point Email Overflow Scale (sorted by factor loadings) with means

1. I feel I spend too much time keeping up with my mail (0.75)
2. Email cuts into the time I wanted to spend on other tasks (0.92) *
3. I have trouble keeping up with email on days I am away from my desk (0.68)
4. I get too much email (1.05) *
5. I spend too much time getting rid of unimportant messages (0.83) *
6. I am satisfied with the strategy I use to keep up with my mail (reverse coded, 0.08)
7. When I return from vacation / time off, I feel overwhelmed when triaging my mail (1.22)
8. Sometimes my emails may get lost or missed (0.47)

* denotes questions on this survey not in [8].

Table 1. Email Overflow scale.

twice’, 29% tried it ‘for a few days’, 11% tried it ‘for a few weeks’ and 13% were still using SNARF.

Email Overflow scale: The previous triage survey [8] contained 8 Likert scale items which worked reliably together as a single email dissatisfaction scale (Cronbach’s alpha=0.88). The values for each of these questions were -2 (strongly disagree) to +2 (strongly agree). For the current study, we modified the questions to focus more specifically on the experience of being overwhelmed by incoming email. Two questions were removed and four added to create a bank of ten questions. Eight of these (five from the previous study and three new ones) worked reliably as a single email overflow factor using maximum likelihood factor analysis, with varimax rotation (Cronbach’s alpha =0.86, p<0.001). The overall mean for this question set is

Usage of SNARF (Daily):

- When SNARF was started and shut down
- What views were shown in SNARF
- How many times users clicked in the SNARF interface: opening messages, message lists, or thread views
- When incoming messages came to SNARF, and when messages changed status (e.g. from read to unread, or were deleted)
- Current folder structure, including number of read and unread messages, earliest message, and latest message per folder

Mail Snapshot (Weekly):

- Weekly anonymized snapshot of the user’s email database, listing the date, sender, and receivers of each message.

Table 2. Behavioral Data from the SNARF Logs

slightly lower than ‘agree’ (0.75, S.D.=.67). The means for the individual items carried over from the previous survey were generally comparable, (i.e. none differed by more than 0.5).

Table 1 summarizes the finalized scale. The highest mean was for question 7, “When returning from vacation...” (1.22, S.D.=0.87), which agrees with the previous study.

Logging. SNARF is instrumented to log a substantial degree of detail about how participants used the program. The log involves two files (table 2). The first, updated daily, records both any changes to the mail store and click data within the program itself. A second file, recorded weekly, contains an anonymized version of the headers of all messages in a user’s mail store.

In this second file, SNARF processed all personal data to produce an anonymous snapshot: senders are numbered from 1 to N, message subjects are encoded with a seeded one-way hash and message bodies are discarded.

While stripping the data inhibits certain types of analysis, these anonymization techniques respect the privacy of our users while collecting data relevant to this particular study.

In order to compare the snapshots of our sample as fairly as possible, we used the first snapshot uploaded. Using this store, we calculated metrics for the six weeks prior to deployment. We found six weeks to be long enough to get good aggregates, but short enough to reduce the likelihood of dramatic changes in the users’ social networks. In the case that SNARF changes an individual’s email behavior, these prior six weeks would remain unchanged. While our recording does track messages that were archived or deleted, it does not accurately track messages expunged from archives; our estimates of message volume may under-count some users’ mail stores.

Unfortunately, due to bugs in our code that surfaced after release, logs from a subset of laptop users were interrupted. We were forced to disregard data from the log files of laptop users for several aspects of the project.

RESULTS ANALYSIS

We begin by presenting a picture of SNARF users and their email stores. Of the 574 users, we received 532 complete snapshots.

Broadly, users had about as many messages coming in directly to them as they sent, not counting mailing list messages: they sent a mean 96 messages (S.D.=75) per week and received a mean of 106 incoming messages (S.D.=95). For most users, lists made up the bulk of their mail: they received an additional 324 messages per week (S.D.=452). All of these distributions were highly skewed: while most users sent and received slightly fewer than the mean, some sent far more. Indeed, the most prolific user in our sample sent approximately 600 messages per week.

Usage

“*[What I liked best was:] seeing my frequent contacts pop up to the top of the list.*”

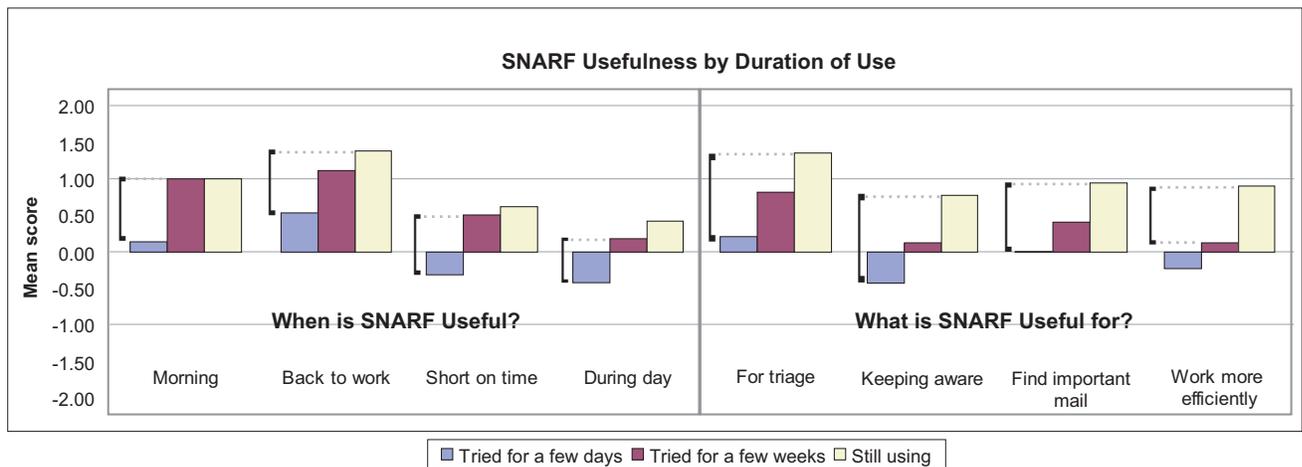


Figure 3. SNARF Usefulness. Brackets indicate the smallest significant ($p < 0.05$) difference between pairs of groups (e.g. for triage, the difference between the lowest group and the highest group is significant).

SNARF is an amalgam of a number of social sorting-oriented features. Which of these were deemed most useful? Which were most frequently used?

SNARF’s usefulness: We asked post-survey respondents to rate the usefulness of a number of SNARF’s features. Only those respondents who said they had used SNARF for at least a few days were asked about features. These 77 respondents were asked about four key times during which SNARF would be useful (“morning”, “coming back to work”, “short on time”, and “during the day”) and four key uses of SNARF (“for triage,” “keeping aware”, “finding important mail,” and “working more efficiently”).

Usefulness was rated on a Likert scale from -2 (strongly disagree) to +2 (strongly agree). SNARF was considered most useful when returning to work after an absence (mean=0.88, S.D.=1.1) and for triage (mean=0.65, S.D.=1.0). We derived the results above using a one-way ANOVA test, looking for differences between the three

groups of users (“few days”, “few weeks”, and “still using”). Differences were significant for all eight questions ($p < 0.01$, two-tailed). We followed up with post-hoc tests to look for specific differences between pairs of groups.

The results are summarized in Figure 3. In general, respondents rated SNARF as most useful when coming back to work after being away. Because triage is most overwhelming when returning to work, this agrees with our concept of SNARF’s utility. The longer people used SNARF, the more useful they found the program overall; the biggest gain was in using SNARF to keep aware of new messages.

Note in this table that “when” questions show little difference between “few weeks” and “still using”. In contrast, “what for” questions feature substantial differences. The people who used SNARF for a few weeks agreed upon *when* the tool was useful; those who stayed for the whole study period had a stronger sense of *why*. In fact, only the users who continued with SNARF rated the tool as useful for helping them work more efficiently.

SNARF’s views: We asked users to evaluate which of five views they found useful. These included the three default views, the optional “Only to Me” view (which shows mail that is addressed only to the user, without carbon-copies) and the optional “List” view (which show messages sent to mailing lists).

As shown in Table 3, respondents clearly preferred views featuring unread messages; particularly, they preferred personal views to others. Personal views contain only messages that are directly addressed to the user. The “Unread only to me” view was not a default setting; users needed to explicitly choose it from a settings dialog. Yet for the users who found it, it was the most likely to be rated as “indispensable.”

Value	Personal Views		Aggregate Views		
	Unread only to me	Unread to/cc me	Unread mail	This week’s mail	List view
Used view (N)	59	76	74	70	52
Distracting	3%	1%	3%	6%	4%
Not useful	5%	4%	11%	36%	14%
Somewhat useful	31%	33%	42%	46%	63%
Very useful	37%	45%	38%	10%	17%
Indispensable	24%	17%	5%	3%	2%
Did not use or find view	19	2	4	7	25

Table 3. Use of Views from the SNARF Logs

The log data indicates that the personal views had fewer messages in them, and that users kept them empty or near-empty. For the personal view “Unread To/CC Me,” two-thirds of its 1293 clicks occurred when the list had 10 or fewer messages. In contrast, two-thirds of the 1142 clicks in “Unread Mail” occurred when the view had 31 or more names in it. Even more dramatically, two-third of the 342 clicks in “This Weeks’ Mail” occurred when the list had over 300 names.

SNARF Threads: The Thread Display places messages within their conversational context by showing messages that are either replies to this message, or that this message replied to. This form of social sorting adds a layer of structure to messages which can help reading and interpretation. In the survey, we asked users to choose three features that they wanted to see improved. This was the most-requested feature for future enhancement. Users who continued to use SNARF were the ones who most wanted improvements to the Thread Viewer (Spearman’s rho=0.33, p<0.05). One user commented that his favorite thing about SNARF was “opening a thread and deleting the whole thread after quickly catching up and responding to the latest in the thread.”

Awareness

OLS Regression on email overload (N=267)

	B	(beta)
(Constant)	-0.51*	
Survey Data		
1. Notifications distract me	0.81***	(0.32)
2. When messages arrive I deal with them right away	-0.50**	(0.14)
3. When triaging I deal with important people first	0.48***	(0.19)
4. Other than SNARF I tried alternatives for email	0.11***	(0.21)
5. Time spent triaging (in minutes)	0.00*	(0.11)
6. Notifies me for all/some/no new messages	-0.11*	(-0.11)
Log Results		
7. Number of distinct people who user has sent messages to	0.00*	(-0.15)
8. Number of distinct people co-addressed on messages with the user	0.00***	(0.45)
9. Number of messages addressed to the user	0.00**	(-0.23)

Adj. R² = 0.29

*p<0.05, **p<0.01, ***p<0.001

Table 4. Model for email overload. Beta represents one standard deviation’s effect on the overall score.

“[What I liked best was:] Quick visibility of new & vocal mail senders.”

SNARF was initially designed as a triage tool; its use as an awareness tool emerged over time. One implication of using this tool for awareness is that users substitute it for other notification methods: if they are glancing at SNARF from time to time, they may remove other notifications.

During the initial survey, we asked all participants whether they had pop-up displays of new mail, or “toast”. We asked again during the post-survey. Some 92 users who responded to both surveys initially said that they display toasts; 22 of those users reported that they turned off the toast during the SNARF study. Users who turned off the toast used SNARF for an average of 14.7 days; users who kept their toast ended sooner, an average of 10 days.

Further tests showed that those who turned off notifications thought that SNARF was more useful during the workday (mean diff.=0.73, p<0.02).

Overflow

SNARF was designed to examine and address the degree to which users were overwhelmed by email. Because the tool reorganizes mail to focus on frequent correspondents, we hoped it would ease the feeling of having too much mail and not enough time. We examined the factors that go into a sense of email overflow, based on the survey, and then examined whether SNARF helped improve overflow.

Modeling Overflow: Using ordinary least squares regression, we built a model of email overflow. The preferred model (shown in Table 4) indicates that overflow is connected to both to behavioral factors (how the user responds to those messages, variables 1-6) and to structural factors (what sort of email comes to the user, 7-9). Users are more likely to suffer from overflow if they are distracted by notifications (1) or if they try to pick-and-choose important messages (3); users are less likely to suffer from overflow if they feel that they can keep on top of their email (2, 5).

Like Belotti et al. [2], we find that overflow is not related to the sheer number of incoming messages: instead, we find that it is *negatively* related to incoming messages that are addressed to the user (9).

It is, however, not related to how many messages are sent by users, how many lists they are subscribed to, nor how frequently the user checks their email.

Overflow and SNARF: Many of the factors in this model are ones that SNARF addresses through social sorting: SNARF allows people to reduce distracting notifications but still deal with relevant messages right away; it allows users to deal with important people first without repeatedly picking through their inbox.

However, we found no significant difference between the overflow levels of users who adopted SNARF and those who

did not. Nor did a comparison between pre-survey and post-survey levels of overflow show a difference.

The discussion section addresses potential reasons why this might be in more detail.

Retention

Despite the fact that SNARF was not successful in reducing overflow, a population of people continued to use SNARF. In this section, we examine what factors were related to people’s continued use of SNARF.

Figure 4 shows the number of days between their first and last execution of SNARF, capped at the 28th day of their use². Roughly one in three users who used SNARF on the day after they installed it continued to use it for the duration of the study. A later analysis revealed that most of the users who were using the program at 28 days continued onward, with minimal further attrition. Responses to the post-survey questions on retention were consistent with the log-based results.

OLS Regression on SNARF Retention (N=70)

	B	(beta)
(Constant)	0.73*	
Average number of recipients per message	-0.14*	(-0.21)
SNARF changed the order I read mail	0.34*	(0.27)
SNARF was useful for keeping aware	0.27**	(0.35)

Adj. R² = 0.34

*p<0.05, **p<0.01

Table 5. Model for SNARF retention

Using data from both the post-survey and the logs, we constructed a model to explain variation in levels of SNARF retention (table 5). The preferred model reduced to three variables. One was based on snapshots: a *negative* correlation with the average number recipients per message. That is to say, people who receive more messages only to them, or to them and few others, were more likely to continue using SNARF. The other two variables were based on survey data: whether the users reported that SNARF was useful for keeping aware, and whether SNARF changed the order in which they read their mail. The preferred model did not include the email overflow scale, any of the usefulness variables, or any other mail volume variables.

² We reanalyzed the log data four weeks after the end of the study date to discover which users continued to use SNARF. These users show up as 28 on attrition curve.

The first variable fits the social sorting concept well. In particular, the default “Unread To/CC” view is most applicable to users when their incoming messages are to them—and few other people. If messages addressed to them have many co-recipients, the messages will still show up, but are less likely to be relevant, and therefore make the view less useful.

The second variable suggests that people who continued to use SNARF were those who found a way to introduce it into their day-to-day work: it allowed them to read the mail in a new, more relevant order.

The last variable reinforces the idea presented in the section on usage: users initially used SNARF for triage, but developed to using it as an awareness tool.

DISCUSSION

By bringing together a number of ways in which social sorting can be used—sorting messages based on their orientation to the receiver; sorting correspondents based on their past history; and sorting conversations based on their structure—we were able to see where social sorting was useful and appreciated. We explore some of these points further, connecting them to email behavior and the design of email clients.

Strategies for handling mail

Neustaedter et al. [8] discuss the main strategies that people use to triage their email as single-pass, where all unread emails are handled one time, or multi-pass where people make several passes through their email picking out relevant messages. Both the data from [8], as well as the pre-survey distributed during this deployment, suggest that being overwhelmed by email is connected to using a multi-pass strategy for email triage. In other words, users who are able to maintain a single-pass strategy for reading incoming email (all other things being equal) feel better about it. This discipline is reflected in strategies advocated by some information overflow consultants [1].

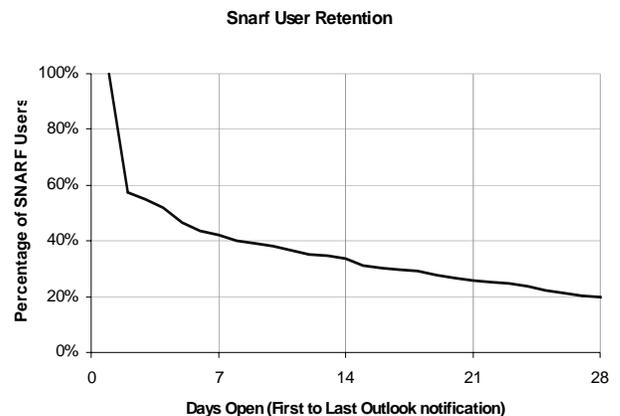


Figure 4. SNARF retention over time

Yet when users are pressed for time, they feel a need to “cherry-pick” their mail, often seeking out messages from people they know are relevant. SNARF cherry-picks for them, first separating messages that are directly to them, then ordering sets of messages by how closely tied the sender is.

The fact that users appreciate this is reinforced by the result that continuing SNARF users reported that they changed the order in which they read email.

We note that many “one-touch” strategies for dealing with email combine *reading* mail with immediate *filing* or *processing* of messages. SNARF does not support the filing and processing phase well. In order to support triage, mail applications must also support this concluding task.

SNARF and Email Overflow

Despite the fact that users found SNARF useful for triage, and that overflow is linked to problems that social sorting might address, SNARF users did not report a measurable improvement in email overflow.

To some extent, this may be inevitable. Mitigating the sense of email overflow may require discipline as well as tools, and a month of a new system may not be sufficient to solve the problem. Indeed, Bellotti et al. [2] noted that email overload is linked to tasks that the user *already knows about*: messages that linger in the inbox because they are not yet handled. While SNARF may reassure the user that they have not received new tasks, their old tasks still remain: and SNARF does not provide a smooth user interface for reviewing email that has already been read.

Conversely, we show that email overflow is negatively linked to incoming messages *directly* to the user. Referring to Bellotti et al. [2] again, we may suggest that those incoming messages are signals of task feedback: fewer messages means less incoming information about tasks, and thus more overload.

We understand the incomplete ties between our triage tool and the reduction of the overwhelm phenomena to be a cue to develop further rather than an indication of failure.

Note on Performance

When SNARF was distributed, some users experienced substantial performance issues. In particular, some users with very large mail stores found it very slow to open messages. This may have discouraged power users, and, indeed, some of those who needed the tool most. Performance was rated as the primary reason for discontinuing SNARF.

This did not affect awareness; however, it may have affected success with triage. Knowing that triage was difficult, some users may have decided to use SNARF as a notification tool instead of as a mail reader. Unfortunately, ‘awareness’ is difficult to mine from SNARF log, as it does not include direct activity from the users.

This issue highlights the benefits of triangulation: what we could not learn from the logs emerged clearly from the survey.

Since the field study, performance has been substantially improved.

SNARF and Conventional Email Clients

It is possible to approximate some aspects of SNARF’s displays on a conventional email client. For example, many clients support “search folders,” a dynamic folder that lists all email that match a given criterion. A search folder of “unread mail, to me” might approximate one SNARF view, at least.

However, the emergent behavior that our users showed, using SNARF for awareness and finding the Thread View useful, suggest that the tool presents different information than standard displays. SNARF presents an awareness display of new email that does not act as a disruptive interruption; it allows an overview of a mailbox without taking up a full screen.

DESIGN IMPLICATIONS

Developing SNARF, and seeing users’ responses to its social sorting mechanisms, has suggested some design directions that may be extended to other systems.

Users have a social history. The ways that they interact with other people can be stored, and then exposed for later use. For example, one current mail client provides an auto-complete based only on the most *recent* sent messages—surely a broader history could provide a more intuitive ordering of potential recipients.

Parallel, multiple perspectives can show more. In many systems, a user has a choice between seeing their information in only one view at a time. SNARF presents several perspectives at once, allowing the user to glance over the most valuable aspects from several angles.

I am my most interesting correspondent. Users found that views that showed messages *to them* were the most valuable. When a system can know that information is particularly concerning to the user, it should be made accessible.

People are a logical grouping. Our users enjoyed seeing *correspondents* as an organizing principle: “I like the concept for being able to group by correspondents,” wrote one user. While some users complained that the names came with too little context, many others found the people an important cue.

CONCLUSION

The SNARF prototype shows that aspects of social sorting are valuable for email triage. SNARF illustrates the utility of an emerging paradigm of information sorting based on social cues. While this information has been available, but

underused, for many years, we are now beginning to understand how it may be productively applied.

Information anxiety is the mother of invention: while it may once have been that a single inbox sorted chronologically was sufficient, a growing number of individuals need more sophisticated means for dealing with incoming mail, as there is simply too much coming in resulting in a genuine need to prioritize.

As a practical day-to-day solution for email handling, SNARF won an small but appreciative audience. This is, we suggest, a tribute to the notion of social sorting.

SNARF utilizes social sorting in several different ways. By partitioning email into different panes, it offers the user an opportunity to ignore irrelevant email without pouring it into multiple complex folders. By ordering people, it allows the most strongly tied contacts to float to the top of lists, allowing people with less strong histories of communication to drift toward the bottom.

Social sorting, then, seems to be a valuable exercise. The meta-information around email could certainly be utilized productively in a number of different ways within contemporary software.

ACKNOWLEDGMENTS

We would like to thank Gina Venolia, for valuable survey and user interface design thoughts; in addition, we would very much like to thank our intrepid users.

REFERENCES

1. Allen, D. (2001) "Getting Things Done: The Art of Stress-Free Productivity." Penguin Putnam: New York
2. Bellotti, V., Duchenaus, N., Howard, M., and Smith, I. (2003). Taking email to task: the design and evaluation of a task management centered email tool. *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems*. New York: ACM Press
3. Cutrell, E., Czerwinski, M., and Horvitz, E. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. *Interact 2001 Conference Proceedings*. Tokyo, Japan. 2001.
4. Fiore, A., LeeTiernan, S., Smith, M. (2002). Observed Behavior and Perceived Value of Authors in Usenet Newsgroups: Bridging the Gap. *Proceedings of the CHI 2002 Conference on Human Factors in Computing Systems*. New York: ACM Press.
5. Fisher, D. and Dourish, P. Social and temporal structures in everyday collaboration. *CHI 2004*.
6. Horvitz E., Jacobs A., Hovel D.. Attention-Sensitive Alerting, Proceedings of UAI '99, Conference on Uncertainty and Artificial Intelligence, July 1999, pp. 305-313.
7. Nardi, B., Whittaker, S., Schwartz, H. (2002) NetWORKers and their Activity in Intensional Networks. *The Journal of Computer-Supported Cooperative Work*. 11, 205-242.
8. Neustaedter, C., Brush, A., and Smith, M. Beyond "From" and "Received": Exploring the Dynamics of Email Triage. *Proceedings of the CHI 2005 Conference on Human Factors in Computing Systems*. New York: ACM Press.
9. Neustaedter, C., Brush, A., Smith, M., and Fisher, D. (2005) The Social Network and Relationship Finder: Social Sorting for Email Triage. *Proceedings of the 2005 Conference on Email and Anti-Spam (CEAS)*.
10. Rohall, S. and Gruen, D. (2002). Remail: A Reinvented Email Prototype. *Proceedings of the CSCW 2002 Conference on Computer-Supported Cooperative Work*.
11. Takkinen, J. and Shahmehri, N. "Café: A Conceptual Model for Managing Information in Electronic Mail." *Proceedings of the Thirty-First Hawaii International Conference on Software Systems*, 1999.
12. Venolia, G.D., and Neustaedter, C. (2003). Understanding Sequence and Reply Relationships within Email Conversations: A Mixed-Model Visualization. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2003)*, ACM Press.
13. Venolia, G.D., Dabbish, L., Cadiz, J.J., and Gupta, A. Supporting Email Workflow. *Microsoft Technical Report TR-2001-88*. 2001.
14. Whittaker, S. and Sidner, C. (1996) Email Overload: Exploring Personal Information Management of Email. *Proceedings of CHI 96 Conference on Human Factors in Computing Systems*. 276-283. New York: ACM Press