Human-Computer Interaction -- INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 813-816

DriftCatcher: The Implicit Social Context of Email

Andrea Lockerd, Ted Selker

MIT Media Lab, 20 Ames St., Cambridge, MA, USA

<alockerd, selker>@media.mit.edu

Abstract: This work uses Artificial Intelligence (AI) in an electronic communication interface to improve people's ability to focus on important communications. Many social cues that allow people to function naturally with their social networks offline are not obvious in Computer Mediated Communication (CMC). This work uses automatic social network analysis to bring some cues to CMC that foster the user's coherent understanding of the resources of their personal communication network. The DriftCatcher email client analyzes and presents some social qualities of messages, presenting users with the main idea, letting them *catch the drift* of the social undertones of their email. Our user study shows that social context in an email browser significantly increases a person's ability to make decisions about the value and importance of messages, demonstrating the value of AI techniques in improving human-human communication.

Keywords: Artificial Intelligence, Social Network Analysis, and Computer Mediated Communication

1 Introduction

When entering a social setting, you instantly scan the room, see who is there, and make mental notes: "I haven't seen her lately", "there's Bob with someone I've not met". This example characterizes the way people naturally use social network analysis in face-to-face interactions. Today's Computer Mediated Communication (CMC) tools lose the social cues of this scenario, making it harder to maintain ones' social network online. The email client in particular is a tool that has not changed significantly in years. Most of the functionality available in modern email clients (forwarding, folders, sorting by header data) was already present in MSG, written in 1975 and widely considered the first email client.

In people's daily use of email, an opportunity arises for modeling aspects of their social network. This work submits that automatic social network analysis can bring social cues to an email interface that foster the user's coherent understanding of the people and resources of their communication network.

One might assume that human communication, with its subtlety and nuances, is an area Artificial Intelligence (AI) is far from understanding. Attempts have been made to apply AI and Social Networks theories to the problem (Kautz, 1997), (Foner, 1997), (Vivacqua, 2000). One of the best examples is Coordinator, an application that identifies patterns of organizational speech and the action that speech tends to induce, letting it better support collaboration (Flores, 1988). A few aspects of our work set it apart from prior work. We are interested in social networks from a single user's point of view rather than an organization. Our system has automatic social network modeling (i.e. no user profile is needed), and it attempts to recognize not only the task related resources but also the *social* resources exchanged between people in the messages they send.

While a machine may never come to be perfect in its analysis of the social nuances of human communication, this work demonstrates that even a partial understanding of these nuances can improve an email interface. The following is a system that reasons about social networks, recognizes social characteristics of human relations, and presents an augmented communication interface (Lockerd, 2002).

2 Implementation

This section details the system architecture components: the DriftCatcher email client, incoming mail processing, and the social network modeling of the SocNetServer.

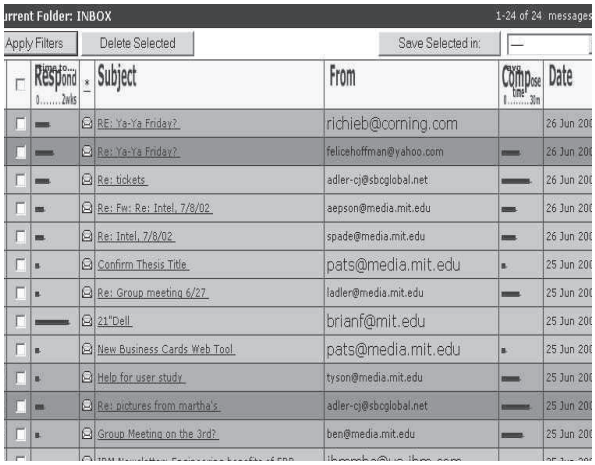

Figure 1: DriftCatcher email client

2.1 DriftCatcher: A Social Mail Client

The American idiom 'catch the drift' means to realize the main idea. With the DriftCatcher email client, users *catch the drift* of their personal communication network. DriftCatcher is an extension of the Emumail webmail client (www.emumail.com) and is social in two dimensions: data collection and display. It utilizes the SocNetServer to add social context cues, which let users deal with communication in a social context in addition to the temporal context of current mail browsers. It also sends information about a user's behavior with their social network back to the SocNetServer.

2.1.1 Collecting Implicit Social Behaviors

An email client, used regularly, is in a position to collect information dynamically about how the user behaves with the people in their personal network. It sees a number of implicit behaviors that characterize online relationships, such as how long users spend reading and writing messages, or how long they take to reply to a message once it's been read. Currently, the DriftCatcher client sends information to the SocNetServer about read and compose time for each message (with a timeout if there is no typing for a few minutes), and the SocNetServer incorporates this into its model of the user's online social network.

2.1.2 Socially Augmented Display

DriftCatcher displays the inbox along social dimensions as well as a typical temporal dimension. We considered more radical interface designs and visualizations, but decided that it was more important to preserve people's familiarity with the current email browser paradigm. The interface is a typical email browser in which each message line is augmented with social context meta-data (Fig 1), which will be discussed in detail shortly.

2.2 Incoming Mail Handling

The email of DriftCatcher users is processed on its way to their mail server with Procmail, a Unix mail processing utility. The script queries the SocNetServer for *social context* statistics about the message and its sender, and then adds this information to the message header. Once a user puts the script on their mail server, the system starts tracking their personal network and marking email with social context. The information is accessible when they use DriftCatcher to view their email, but the network modeling continues regardless of email client choice.

2.3 SocNetServer

The SocNetServer embodies the personal network analysis behind the DriftCatcher client. A person's social network consists of the set of people with whom they have ties, connections between them, and social resources they exchange. While a true model of a person's social network would include more than email interactions, SocNetServer recognizes elements likely to be relevant to an online communication interface. It models the dynamic personal network of each user and has generic models of social resources exchanged in email. Clients communicate with the SocNetServer using XML-RPC (www.xml-rpc.com).

2.3.2 Modeling Social Resources of Email

SocNetServer has static statistical models of social resources people exchange in email (i.e. *informing, inquiring, sharing, planning, intimate, etc.*). The attempt to understand the social resources as well as the dynamic network features is a key element that sets this work apart from related works.

We built Support Vector Machine (SVM) classification models, which are particularly good at learning to classify in a large feature space with sparse data (Witten, 2000). The process starts with labeled examples of the intended classification. With email, the corpuses available for this are from public mailing lists and newsgroups. While these do have some variance of relationship and social context, it was necessary to collect a new corpus of data for the domain of personal email. We collected data from 8 subjects emailing each other over a month, yielding 550 email messages, which were then labeled with social resource meta-data. Thirty labels were established by considering of how

Bales' Interaction Process Analysis best translates from physical to online interaction (Bales, 1950); eight of these were represented in the corpus: *informing, inquiring, interest, keepInTouch, planning, sharing, intimate*, and *supportive.* The final SVM models use the feature set: terminating punctuation, frequency of punctuation, time and date related words, URLs, including the old message in the new, and the frequency of emoticons used. In cross-validation testing with 10% holdout, a standard technique used to judge the success of the models, these models had 50-70% accuracy.

2.3.1 Personal Network Modeling

Some features of a social network are dynamic and unique to a user: network structure, frequency of contact, symmetry of contact, response times, time spent reading and composing messages. In addition to having the social resource models, SocNetServer collects these dynamic elements over time to model each user's personal online social network.

2.4 Social Context in the Interface

As mentioned previously, the interface (Fig. 1) is a typical email browser with social context meta-data for each message line. The social context information in the message header allows DriftCatcher to augment the presentation of each message in the following ways.

The font size of sender's names varies based on tie strength. Weak ties are bigger than strong, in accordance with the finding that weak ties are better for finding new information and gaining access to other networks (with new resources and social capital) (Granovetter, 1982). Tie strength is based on frequency and symmetry of contact.

With each message, DriftCatcher displays the average time the user has taken to compose messages to this sender (0-30+ minutes), based on messages composed with DriftCatcher. This feature is expected to be useful when, for instance, a user has only a few minutes and wants to find a few messages that can be dealt with quickly.

As shown in Figure 1, the leftmost column is an indication of the time left to respond to this message (0-2 weeks). This indicates how long the user should take to respond to this message if they want to reciprocate the response pattern of the sender. The default for a new contact is two weeks, and this changes as a response behavior is established.

The background color of a message reflects the social resource classification. Green, yellow, pink, orange, and blue represent *Inquire*, *KeepInTouch*, *Interest*, *Planning*, and *Inform/Share* respectively.

3 Evaluation

We conducted two user studies to explore the following:

- Did the system learn to classify social resources into similar categories as a group of humans?
- Do people have a better sense of the value of communication with the DriftCatcher client?

3.1 Social Resource Classification

Six participants labeled a set of 72 messages with the eight categories of social resources; in 82% of the cases the participant's labeling matched each other. The SVMs then classified the same set of messages.

In 68% of the cases at least one person's labeling matched the machine classification. When looking just at the 82% in which there was participant consensus, the machine matched in 49%. In general, the machine was more generous in giving a message a particular label, yielding many false positives. To some extent, this was expected given our small dataset. The question then becomes whether even a partial understanding of the social nuances of messages can improve an email interface.

3.2 Effects of the DriftCatcher Client

To test whether people have a better sense of the value of communication when using the DriftCatcher client, volunteers were given this scenario and task:

> *Today is your 1st day as an administrative assistant. One of the people you support is returning in 5 minutes. Go through the 24 new messages in her inbox, and choose 3 messages she should deal with first. She considers email priorities (in no particular order): People trying to make plans or things that affect her schedule; People asking for a favor or advice; Timely responses, especially to people with whom she has a close relationship*

The task involves stepping into the social context of another person's inbox, and making a judgment given too short a time to read all 24 messages. Since participants spent a short time with the interface, this shows immediate effects; we believe success here is promising for longitudinal effects.

The study was conducted with 30 participants and was counter balanced presenting inboxes with and without social context cues. The results from one-way ANOVA tests with each inbox (Table 1) support that the DriftCatcher client helped people with the task, not by increasing the number of messages they could attend to, but by increasing the *value* of those attended to. The inboxes were not

standardized or tested for equivalence, which we believe explains the non-uniform significance findings; nonetheless, significance was found with two of the three inboxes in each of the following cases. The percentage of messages read from a "close relation" as well as requiring a "quick response" increased significantly with the social client. As did the percentage of "inquiring", "informing", or "planning" messages read. As this is a preliminary result, further work needs to be conducted to confirm or deny these findings.

In addition to these quantitative results we conducted a survey of the participants for qualitative feedback. Almost every participant liked the idea of social context annotation of their email. Mentioned most frequently as particularly useful were the reciprocating response time feature and the closeness of relationship indication.

	Box 1	**Box 2**	**Box 3**
Tot. Read F(2,27)	1.71, p<.2	.32, p<.73	.18, p<.84
%Close F(2,27)	2.71, p<.09	6.1, p<.01	.53, p<.6
%Quick F(2,27)	5.89, p<.01	3.57, p<.04	.52, p<.6
%Inquire F(2,27)	81.38, p<.00	.45, p<.64	7.1, p<.00
%Inform F(2,27)	3.22, p<.06	23.88, p<.00	29.91, p<.00
%Plan F(2,27)	2.24, p<.13	8.71, p<.00	18.51, p<.00

Table 1: One-way ANOVA tests for each inbox comparing groups presented with social annotations against those without.

4 Conclusion

As people come to depend ever more on electronic communication for maintaining and managing their relationships, computers will need to support and augment this interaction. We have shown that an email client that comments graphically on social relevance can significantly improve a user's ability to attend to messages useful to their tasks.

The SocNetServer compiles personal social network information based on email interactions and has statistical models of email social resources. The DriftCatcher email client utilizes the SocNetServer's intelligence to help users understand their social network. Our user study shows that when users are given the social client and asked to pay attention to scheduling, inquiries, close relations, and timely responses, they appropriately succeed in reading a greater percentage of messages of the following categories: of close relations, needing quick responses, and involving planning or inquiry.

Acquiring more training examples and building better models of email social resources is an obvious area for future work. Soon users will be able to train the system dynamically, "showing" it examples of various social resources, thereby allowing the models to continually improve.

This work promises to extend the implication of computers and AI, improving human-human communication and relationship management. The possibilities for such technology are extensive and controversial. Do we want to trust computers to do a better job than us at deciding how to prioritize activities, communications, and relationships? If we depend on such tools will we loose our social skills and intuition? These questions of future research warrant deep consideration from computer scientists, designers, and sociologists.

References

Bales, R.F. "Interaction process analysis", Reading MA: Addison-Wesley, 1950.

Emumail, www.emumail.com.

Flores, F., Graves, M., Hartfield, B. & Winograd, T. "Computer Systems and Design of Organizational Interaction", *ACM Trans. Info. Sys.*, Vol. 6.2, 1988.

Foner, L. "Yenta: Multi-Agent Referral System", *1st Intl. Conf. on Autonomous Agents*, 1997.

Granovetter, M. "The Strength of Weak Ties", *Social Structure and Network Analysis*, Ed by Marsden, P. and Lin, N. Beverly Hills: Sage, 1982, (105-130).

Kautz, H., Selman, B., Shah, M. "Referral Web: Combining Social Networks and Collaborative Filtering", *C ACM*, v40, 3, Mar 97, 63-65.

Lockerd, A.L. "Understanding Implicit Social Context in Electronic Communication", MIT Masters Thesis, Sept. 2002.

Vivacqua, A. & Lieberman, H. "Agents to Assist in Finding Help", *Proc of the ACM Conference on Computer Human Interaction*, April 2000.

Witten, I. & Frank, E. "Data Mining", Morgan-Kaufmann, 2000.

XML-RPC, www.xml-rpc.com.