Structured Time-sensitive Mail Search

Nathan Matthews
nmatthe2@uiuc.edu
February 18, 2009

1 Purpose

The purpose of this email visualization is to help visualize how your contacts fit into the landscape of your projects, interests, and social groups. Automatic labels and searchable keywords form the basis of the distinguishing characteristics. For instance, GMail’s label system, which incorporates subject, author, and keyword search, could be imported. Alternatively, a simple text search could be used.

2 Cluster/Label View

Given a set of labels, the program scores each email in the user’s archive according to how well it matches these criteria. It then sums over each unique contact to find the contact’s score for each label, and a cutoff is applied to arrive at a finite set of contacts. In addition, emails to and from mailing lists are included, so a contact can be included even if there is never a direct mail between the user and the contact. The user themselves is scored based on email sent.

The mapping of contacts to labels is done in a time-sensitive manner so that the evolution of the social networks, as well as the user’s own migration between topics or projects can be seen dynamically. In this view (Figure 1), contacts are represented as circles whose area corresponds to the amount of mail sent. Mailing lists are represented by rectangles. In the likely case that the circles overlap (for instance, because they are associated very strongly with just one label), they are displaced from the center. The positioning between the labels should therefore be seen as only approximate and relative.

The gray lines between circles indicate the amount of email between those two contacts (similarly for lines between contacts and lists). If possible, in the implementation it would be nice if the gray lines were separated into a double, bidirectional line, to indicate the amount of traffic in each direction.
Figure 1: Cluster/network view of email over time
3 Single-Label Activity View

In this view, a single label, topic, or keyword is chosen. Contacts are scored as before, by how active they are within this label. As before, the scoring is time-sensitive. The contacts are plotted in a stacked area graph, which allows the user to see the relative activity of different contacts. Although it is not shown in Figure 2, the area under a contact’s activity line contains a marginal plot of each email sent, represented by a vertical line of higher-intensity color on the exact date/time the email was sent. Moving the mouse over a contact’s area highlights the area, and moving the mouse over one of these lines displays a summary of the email.