1 Visualizing Email Archives

This summary paper makes arguments about how people use email and email clients. It argues that much of the everyday use of email is for social and personal purposes, rather than practical or task-focused purposes. This realization informs possible design directions for email clients, notably that they should become better at displaying social information and that the client itself should form a good habitat for the user.

I wanted to start by saying that some of the flattening effect of email is actually beneficial sometimes. I don’t always want to be reminded of the extended social status and history of whoever I’m communicating with; sometimes what is most important is just their relationship to me, or even just the content of the email. The sparsity of email has likely been a driving factor for the sense of global community on the internet. I’m not saying that social network visualizations destroy this effect, just that the lack of explicit social structure in email has not been entirely negative.

For most email users today, the suggestions the paper makes about tagging over folder-based categorization have largely been implemented. The most common system is simple dynamic user-defined tags or automatic structured tags, such as GMail’s labels. Cheap searching enhances the ability to retain knowledge of a large email archive. I think putting too much structure into tags is probably overkill, and may have the same negative effects as folders.

Affective analysis is something that email definitely lacks. Even though we have some fairly advanced algorithms for detecting spam, most clients don’t do any similar processing for emotional content. A similar benefit would be the ability to summarize email, or even to pick the most salient concepts and use them as a first level of search. Something like PostHistory that used concepts instead of contacts would be very interesting.

I don’t know if this just a difference with my personal email, but cc and bcc don’t tend to offer much information in the way of social networks (beyond the first level of association). It might be useful to correlate email with another data source such as a web service explicitly designed to create networks of people. The addition of affect information or correspondence frequency could richen these visualizations, but the email by itself seems to be too weak to handle the job.

Converting email clients from tools to habitats is an interesting proposition, especially since so many traditionally thick-client applications are migrating to the web. The web is inherently interconnected by hyperlinks, and any piece of semantic information in email or other personal archives is potentially a link, anchor, or search term. The amount of context which can be automatically culled from information is greater. Part of the goal of an “online habitat” for email should be to incorporate this extended context in such a way that the user’s ability to “cultivate” the archive is maximized. Balancing this goal with that of increasing social visibility is an important tradeoff.
2 Social Network Fragments

The chapter presents some theoretical background on social networks, especially relating to identity and information management, then presents the Social Network Fragments visualization. The visualization scans a user’s email archives (after the user has categorized their contacts into a few groups) and presents a graphical view of their social network by finding connections between all contacts.

The theoretical parts were very interesting. I’d never considered the argument that being a bridge was important to information management or security. The fewer of your contacts who know each other (i.e., the more pairs of contacts for whom you are a bridge), the more they have to communicate through you, or at least the more information flows through you in the long run. This actually explains a variety of weird social phenomena, like why we have middlemen or why some people seem to be natural “hubs.” This behavior could indicate an instictive ability to maintain structural holes.

This is something that should certainly be visible in the visualization of Social Network Fragments. I’m surprised at the level of interconnectedness that is shown in the example images. It’s hard for me to see how the graph could end up looking like anything but star-shaped or partially wheel-shaped. I guess if each contact was included in some but not all of a set of CC’d emails, that might do it.

The determination of link strength largely makes sense. The shakiest part of that seems to be the use of BCC. I’ve never personally BCC’d an email if there was also a non-BCC recipient. Sometimes email information gets shared in other ways, such as through references in instant messaging or physical communication, but never on the level of the entire email. That said, it seems to work as far as producing a network with interesting structural features. Since the parameters are adjustable, there’s not any harm in including many different types of relationships and measures of link strength.

I really like the ability to zoom in and out. It would lend itself nicely to multitouch displays, increasing the element of interactivity and explorability. The history view is sufficient for playing the network through time, but I found the placement of the colored squares counterintuitive. The outermost square is largest and most visible, so I though it should be the one to indicate the highest level of trust (though I understand the inner/outer metaphor).

The other things which seemed odd were mentioned the reading, such as doing layout on every time slice and indicating the difference between link types in some way that doesn’t distort the 2D distance metric.
3 Public Displays of Connection

This paper discusses the construction of different social networking sites from 2004 (which is surely more than a decade in internet years). It talks about why users of social networks display their connections and what social or practical function this has, and how the design of social networking sites affects this behavior.

I thought the four properties of connections in social networking sites were interesting. Especially interesting is that two of the most-used networks, Facebook and Twitter, have opposite approaches to connections: Facebook connections are (potentially) private, nuanced, and contextualized; only the mutuality property holds. In contrast, all Twitter links are mutual, public, unnuanced, and decontextualized. It seems likely to me that this indicates a qualitative difference in the reasons and uses of the sites’ social networks. Facebook has a highly social function, but often this social function is highly contextualized to group contacts into logical groups. Twitter is a “flat” network, with no inherent grouping functions; the links serve more as an indicator of interest and information flow. While interest can follow the closeness of a relationship, it also follows the celebrity of the poster. Links in In that sense, Twitter networks are less display-oriented than Facebook networks.

The shortcomings of social networks mentioned early in the paper, namely the possibility of creating a persona which is false yet appears to be well-connected, are problems that have often been addressed in reputation networks. Well-known methods exist to find clusters of nodes which are artificially inflating the reputation of a single target node (for instance, in order to bolster its false identity claims).

A lot of the points made by the paper indicate that reputation is often treated like a quantifiable resource by those who have to consider the consequences of their actions. So-called “social capital” has been around a long time, but what online networks add to the equation is the ability for outsiders to quickly obtain a numerical estimate of any user’s reputation. When links are public, it is extremely difficult to create structural holes in one’s social network without using completely autonomous pseudonyms. The implicit tradeoff is between information power (created by using pseudonyms to maintain structural holes) and social capital (created by having a single avatar which can accumulate the reputation of all the user’s interactions).

The paper notes that improved technology is unlikely to increase the number of strong personal ties, but does increase the number of weak ties, making the user’s extended network more heterogenous and increasing their information base. Of course, this increases the number of duplicitous elements likely to appear in a user’s network, as the paper mentions. The more contacts a user has, the more susceptible they are to spam and misinformation. However, I feel that the source of these problems is also part of the potential solutions. As each user’s network grows larger, the network as a whole grows more connected. This enables some of the tools of reputation networks and security to be applied in order to search for various kinds of attacks, even before they’re carried out.
4 The Image of the Environment

The book discusses the problems of urban design in the context of the populace’s image of the city. The central focus of the discussion is on the clarity of this image, meaning the ability of its residents to pick out salient features and differentiate between areas to help them get from place to place and to increase their familiarity and enjoyment of the city. It was not immediately obvious to me what the connection was to the visualizations we normally discuss, until I realized that the problem of altering the urban environment to produce a particular image is the same as the problem of present abstract information with visual cues, but from the opposite end. The same background theories apply: differentiation, landmarks, identity, paths, containment, rhythm, and scale, among others. The city planner’s problem is to predict the kind of abstractions and images people are likely to make from their interactions with the city, and design the city so that these mental images are as useful and pleasing as possible. The visualization’s purpose is to explicitly design the abstraction to draw attention to facets of the raw data (the city itself, in this analogy). In both cases, the goal is to increase the utility of the information or structures that are inherent in the data set.

Some of the metaphors we use when discussion human-computer interfaces are even taken from physical actions which fit directly into the city metaphor: tracing a path (walking/driving), browsing (exploring), zooming out (taking a broad view or panorama), eye candy (scenery), boxes and windows (boundaries/edges), labels (signs and symbols), alignment (regularity/linearity), animation (rhythm/melody/flow). Another connection is that a visualization, like a city, is a habitat. It has easily and commonly found elements. Making the habitat a good one involves a balance between familiarity and exploration.