Visualizing Email Archives – draft
Judith Donath

Reading this paper sparked a lot of innovation in my head. The first section, “Why visualize email archives?”, basically lists the reasons current archives fail. It then goes on to say directions that archives should go. As a result, there is just a black box in the middle for how to get from the current failures to the goal state of archives. The second section, “What information is in email?” gives the tools available to developers when creating visualizations. Together, the two sections contain so much information that it seems strange that email archive visualizations have not become mainstream. I hope that reading the section on current work has not boxed in my ideas.

This paper denounces the folder metaphor commonly used in favor of use-defined attributes (tags). Folders are, “cognitively-difficult (to file)” (pg. 6), it says, while tagging will, “provide more detailed and extensive annotation data ... it would be cognitively easier than the current method” (pg. 6). One practical use of tagging can be seen on delicious, where users keep track of website bookmarking by giving each bookmark a list of tags they want to use to describe the page with. That way, a page can be found later by searching for any of its associated tags. However, I have not seen e-mail clients moving away from the folders metaphor. Gmail has made an attempt to add tags using its “labels” feature. Sadly, it is not very “detailed” or “extensive”.

The problem with tagging is how much learning would be required for it to work well. For delicious, it works because users have already decided that the bookmark they want to save is important enough to put in the time to tag it for later. Individual emails are not always that important. In fact, it is so easy to send an e-mail, I would say that even excluding spam messages, most e-mails a person receives on a day to day basis are not important. This is probably why tagging is not added to more e-mail clients. Tagging all of one’s e-mail would take up a considerable amount of time.

An answer to this problem would be some sort of smart e-mail client that learns a user’s preferences and tags accordingly. Gmail has a very lite version that allows users to automatically filter and tag messages, but again there is a problem of time consumption. Filters can be very specific and provide the “detailed and extensive annotation data” that this paper is advocating, but it comes at the expense of taking a considerable amount of time to set up the filter. Easier, much broader filters can be set, but they provide minimal amount of data, like messages from family members, or work.

Until appropriate learning algorithms are researched, I would like to advocate a more robust filing system. Why not add photos, music, or other outside information to an archive? Allowing users to recall other more memorable incidences happening around the same time could give people easier recall of the information they were looking for in an archive. Additional information would hopefully add a lot more of an emotional element by helping remind user’s of the context of messages. This would create more of a habitat for archiving information.
The beginnings of this paper, where the theories of social networks are described, make me a little sad. It does not seem right to reduce people to such base terms. Like learning the tricks that con artists use, I am left feeling a little unclean. Studying the shape of social networks, and learning about the differences between strong and weak ties is fine. I also have no problem with information flow, and determining the paths that information takes as it spreads. Even, finding out which individuals tend to be the most important bridges of communication is fine, if it is trying to give a community a picture of itself. I also do not have a problem with such research being used maliciously. However, these theories seem to be put forth from the standpoint that people are very self serving, and are constantly looking for personal gain. Using terms like “advantageous” and “controlling information” sounds rather subversive. One should not have to try and out manipulate their peers in a sort of information war.

There seems to be an assumption in this information that separate networks care about the information held in each. I cannot think of much information gain coming from the fact that someone is the sole tie between two groups. Presumably there is a reason that the two networks are separated in the first place. Another assumption seems implicit in these theories is that someone who is the sole tie between two communities is known in both communities and has power within them. For instance, being a member of a community in which you are not considered important would give someone very little opportunity to use their ties to their advantage.

I think that this paper has a very objective perspective on spring based systems. We have touched on the problems with spring based systems in class, but they are well documented here. For instance, it is hard to show high dimensionality in a spring based system, and as a result clustering does not always occur because of a similarity between object. It may instead form from object being pushed away from objects that they are not related to, and into other clusters. On the other hand spring based systems can show data well when objects can be easily reduced to one or two dimensions. In perhaps the best case, developers can use a plot along an x and y axis along with color to show information.

It is not very clear how easy it is to see a network or individual over time. For each slice are people maintained in the same radian along the circle and simply moved further or closer depending on weight? If they are not kept in the same radian each slice needs to be scanned for each individual person which seems like a bit of a hassle. Also, I wonder if there is any significance to each slice being two weeks long. Perhaps two week intervals are necessary to gain enough information for a reasonable visualization with a good number of connections.
This paper talks at length about the social networks and their associated information. Personally, I do not like the notion that the only reason for social networks is to gain “social capital”. It is upsetting to think that people are making connections so that they can leverage those connections later. People should not be objects. This is not some game of pokemon where having the right people that you can call at any time makes you a better person than someone else. And it is certainly not the case that people should be scored and ranked.

It is, however, good to know that there are people who think this way. It explains why I have had my Facebook page scrutinized for job interviews. What makes this sort of cross over from the virtual world to the real world even worse is the fact that despite being a site that is not especially “businesslike” like LinkedIn. If your real name is attached to an account your information is still looked at as if you were on a “businesslike” site. Can an individual not have a social life on the internet that is both personal and public? Everyone should be allowed to have a work life, and a social life. Putting something slightly off color, like “clogging toilets”, under my interests on a social site, should not have any impact on my work life.

People that subscribe to the kind of thinking that this paper describes make me want to get rid of any account I have on a social networking site. Why is it so strange to just want to keep in touch with friends? I am not however, looking to exploit them for favors. Nor am I looking to expand my network. I just do not want to lose the network that I have. I have too many “connections” to keep track of easily, so having a Facebook account makes keeping track of those relationships relatively easy.

Context is an incredibly important aspect when it comes to social interaction, but it seems that any interaction is lumped under the same context. With search engines being able to bring back deeper data on searches, and the increasing trend to maintain a constant web presence across multiple portals, the concept separate online presences is being lost. This trend is terrible, and needs to be addressed. Hopefully, as more people grow up that have been using the internet from childhood, they will not expect themselves to be held to the same standards and their younger selves, and by association will allow for the internet to have separate coexisting circles.