ForumReader is a forum visualization specially tailored for flash forums. An example of a flash forum is the comments section of Slashdot.org articles. Flash forums differ from traditional forums in that they have shallow thread depths and only exist for a finite amount of time. ForumReader has a scatter plot view and a thumbnail thread view. The scatter plot uses a clustering algorithm to cluster similar messages together on a graph. The thumbnail view creates tiny thumbnails of each forum post and displays them all while highlighting important thumbnails.

ForumReader is a novel approach to visualizing flash forums. I feel traditional web views are inadequate for conveying meaning. Either there are too many posts shown to find interesting information, or too many posts are hidden away from view. ForumReader allows for a more customizable view of flash forums. I believe the concept of showing all the forum data with some shading mechanism can help users parse through the forums to find pertinent information. This type of visualization also makes flash forums more accessible to average users.

ForumReader is a good visualization for reading flash forums. I question its usefulness on newer topics though. A newly created flash forum will not have much information to visualize. The scatter plot will be sparse and bare. There will be few thumbnails and even fewer shaded thumbnails from searches. The entire ForumReader program appears to be based around visualizing existing flash forums. Given the transient nature of flash forums, I wonder how effective ForumReader is for people wishing to contribute to the forums. Even still, ForumReader is definitely useful for reading and searching archived or existing flash forums.

This may not be much of an issue, but it appears to be difficult to tell which threads are replies to which posts in the thumbnail view. Perhaps when the program is full screen this is not a problem. The only screenshots shown in the paper are rather small and appear tiny on my laptop. It is difficult to tell exactly how far indented each thumbnail is. Another reason why this may not be important is due to the characteristics of a flash forum. Flash forums by nature have very shallow thread depths. In essence, it is not all that important to know which posts are replies and which are new topics.

I do not like the idea of displaying every post in the Text view screen. Flash forums often hide posts which have a low rating. It is just too much information to display and can take a long time to scroll through. I think it would be better to only display a few posts’ text in this view at a time and add buttons to go to the next set of posts. Perhaps each thread could be displayed in the Text view with next/previous thread buttons. This would eliminate the extremely long length of text which makes finding anything next to impossible.

I find it strange that users found the scatter plot and the thumbnail view to be useful by themselves, but were not as good together. Since a user could simply ignore either of the tools, one would have expected the combination of the two tools to be at least as good as each tool individually. The sum of the parts in this case is actually worse than the parts alone. I believe the users may have been too fascinated with how the two visualizations played off each other and less concerned with the actual threads. The more indirection created, the less likely people will pay attention to the actual data.
This paper describes a visualization program called CodeSaw. This program visualizes open source code repositories. It graphs individual user’s contributions to the project over time. In the graph, shaded areas above the line indicate code contributions, and shaded areas below the line indicate email contributions. The paper describes the iterative process of developing CodeSaw. The first iteration of CodeSaw showed too much information. Finally, interviews with project developers provide anecdotal evidence of the usefulness and limitations of CodeSaw.

CodeSaw has a very simple design. It is easy to tell what the visualization shows and how to interpret the data. I like the spatial messaging system as a way to explain the graph. If there is an extended period of inactivity for a user, they can write a message explaining that they were away or why they were not able to work. One question I have is about the coloring scheme used. Does each project contributor get assigned a unique color? Can the user set specific colors for different contributors?

I like the idea of CodeSaw. I believe its features are most useful for unorganized large scale software development. In highly structured environments, there is a command structure and accountability. In this structure, it rapidly becomes known who is and who is not working. In small scale projects, it is also easy to know who is contributing to a project based on subversion logs. The CodeSaw visualization helps when there are many contributors and little formal community structure. I believe it shows off who is not pulling their weight on a project.

Having a year long timeframe is a decent length of time. One feature I would like to see, though, is dynamically setting the length of time. Yahoo! Finance has a good visualization for displaying stock market data on a user set time scale. In Yahoo’s system, the data scales to the length of time shown. If over a couple days of data are to be shown then the data is aggregated on a daily scale. If only one day is to be visualized, then the data is aggregated on an hourly scale.

A couple of the interviewed people mentioned that CodeSaw does not adequately show how active a community is involved in a project. They believed it unduly showed a project as lonely. One person suggested using the forum post and bug reporting data. I agree that forum posts could be treated as emails and that data could be aggregated into the timeline under the mail graph. Patches could also be aggregated under the code graph. This would incorporate more of the community data into the visualization.

Although all the data aggregated is public, I feel this visualization may act as a deterrent to entering a project. Some people may feel subconscious about not contributing enough to a project and may opt not to contribute at all. It could also be undue pressure on some individuals to contribute more so they can be more a part of the group. Finally, it could lead to discrimination if certain contributors feel more entitled because they have contributed more. These people could look down on the opinions and thoughts of lesser contributors. In the end, this could create an inflamed work environment filled with envy and paranoia. I have read about Wikipedia power users being paranoid about the motives of new editors to that system. A similar outcome could occur here.
This paper is about the emerging genre of weblogs. In this paper, blogs are categorized and characterized. The bloggers themselves are also characterized based on information provided on the blog pages. The overall conclusion of this paper is that blogs are primarily used as public online diaries. They are largely a rehashing of personal homepages. The authors conclude that blogs are progressing toward a form of asynchronous text-based communication.

I found the work by Krishnamurthy on splitting blogs into four quadrants to be interesting. I could easily see where websites that I visit on a regular basis fit into the diagram. Nearly all the websites I visit would be classified in the community half plane. I wonder if the most popular blogs are found to be more community oriented. The study shows that most blogs are personal and do not link to other outside sources. The study does not mention whether these personal blogs are heavily trafficked. This information is probably impossible to discover based solely on reading the blogs. Unless the blog itself mentions its own popularity, there would be no way of telling which blogs are more heavily visited than others.

The authors excluded the most popular and perhaps most common form of blogging from their study. They purposefully ignored LiveJournal and Diaryland blog pages. They believed that these blog sites would overshadow their results. If these two sites are so common as to overshadow the data of other blog sites, then these two sites are more indicative of the state of blogging as it currently stands. I find it strange that the authors would simply ignore data because of its popularity.

The paper makes repeated reference to the fact that blogs are viewed as interactive websites “oriented towards external events”. I do not have this false misconception, and I do not believe it is a common misconception. I think most people would agree that blogs are something akin to online diaries. Many bloggers are self-aggrandizing individuals with nothing more to say than a simple recap of their own boring lives. Successful bloggers on the other hand may have more interactive sites which are catered towards external events. This article makes no distinction between the common blogger and successful bloggers.

The outcomes from the study appear to match what is expected. The stereotypical teen blogger is usually a high-school girl and the stereotypical adult blogger is a politically active male. Since the study showed 58% of teen bloggers are girls and 63% of adult bloggers are males, this profiling appears to fit. One concern I have with this data is that it relies on the bloggers own page for information. If a blogger lies about their own identity, then the numbers for the study would be distorted. Since it is common for people to masquerade on the internet as someone other than who they say they are, the data found in the study is suspect.