Review of “The dynamics of mass interaction”

Paper reviews Usenet, as the largest conversational application, about the nature of its interactions. Descriptive data about newsgroup demographics, communication strategies and interactivity is provided and predictions from the common ground model of communication are derived to test the predictions. The paper selects 500 active newsgroups for the case study. It defines active groups as those which had at least 180 messages over 6 months, equivalent to one message per day. However, considering the fact that the paper discusses 2.15 million messages produced by 659,450 people over 6 months, we should expect 0.6 messages per person per month. Knowing that each group has several active posters, we could reach the conclusion that many people are not participating in posting and are just listeners. This would result in the fact that only certain number of people are seriously involved in the discussion and would reject the two initial expectations of H7 & H8 without the need of further analysis.

Also the paper gathers data over a 6 month period from July to December. Considering the fact that many schools and universities are less active during the summer period, using only data from 6 months could affect the result, since many students, whom I think are the majority of people using Usenet during 90s, would have more free time to participate. Expanding the time frame from 6 months to a whole year would lead to more accurate results.

The paper also relies on the fact that high proportion of initiating messages would indicate repeated failures to start conversations, or prevalence of conversational dead-ends. However, my own experience has shown me that when thread depths increased, after certain level, some people will start thinking that the discussions has become too long to follow and the current discussion has completely been different than the initial one. These people would start a new thread at this point. However, others usually would ask them to continue with the previous thread rather than using the new one. The success of each thread, the new one vs. the previous one, depends on the number of people agreeing with either cases of continuing with the previous thread or starting a new one. So I think it is not through to say that all the new threads that are not followed in depth are failed. Contents need to be further checked to see if the new threads have been somehow followed in any of the existing ones or not. In next part I will try to further discuss this.

Also the paper only relies on the header information and ignores the thread content. I think using content would lead to more realistic results. Techniques, such as more frequent words used in each thread, can be used to gain more accurate idea about the main content of each thread.

The paper also highlighted that their findings show massive inequalities between different people in a specific newsgroup and that the Internet is not an egalitarian forum. I disagree with these findings; since people being not involved in the discussion in one newsgroup might be verbose participants in other newsgroups. Internet gives people equal chance to participate in the discussion; however it is up to them to decide where and when to participate. People would prefer to participate in discussions when they know enough about it. So it is not true to say that Usenet conversations do not strictly involve repeated levels of mass participants. Listeners in a group could be active participant in other groups.

Finally the discussion about great number of cross-posting could be strengthened by the fact that some people would re-write some of the discussion in one group into another. I think this should also be considered as cross-posting, since it is the same content being repeated in different groups. As mentioned before, using content and key words in this case, rather than just relying on thread topics, could be used to cover the problem.
Review of “Studying Cooperation and Conflict between Authors with history flow Visualizations”

The paper discusses the wiki pages with special emphasize and case study on Wikipedia. Wiki pages are those where every visitor can edit each page freely. The paper introduces a new visualizing tool, called history flow, and use it to reveal patterns within Wikipedia pages.

One drawback of the proposed visualization tool, history flow, could be the fact that usually the first draft of a page is revised not only to add/remove some parts but also to re-order current available information. In the used algorithm where texts from different revisions are connected to each other with a line, in case of moving of a sentence to another location within the same paragraph problems occur. It is not mentioned in the paper whether such shifting is considered as removing a part and adding some new parts or it would be considered as old texts being shifted. Since the paper states reviewing different revisions of a document in a paragraph level as a drawback to the current used source controls systems, it should be the second item which is valid. Then, moving and re-ordering current sentences in a long text could result in too much intersections of relating lines. This issue could be even more significant considering the paper’s finding of the fact that the person creating a page generally sets the tone of the article on that page and, therefore, his text usually has the highest survival rate. Keeping the initial sentences and adding new sentences to that requires re-ordering and moving of the previous sentences in most of the times and this is why this fact can be combined with the above mentioned problem. However the other paper’s finding that people tend to delete and insert text more frequently than moving text in an article can somehow reduce the effect of the problem.

I really liked the paper’s discussion on detecting mass deletion and the finding that in reality the instances of mass deletion are fixed so quickly that they cannot be seen when revisions are spaced by date. In other words, all vandalism of this type would almost immediately be corrected.

I also have the idea that maybe the content review of the pages can be used to find idiosyncratic copies. In other words, if the system detects an ongoing argument between several authors by removing or revising a part of a page and then the previous author rewriting his previous content again, the system could review the content to recognize the keywords of the paragraph. These keywords can then be studied to see if they are biased and one-sided. In this case the system can automatically remove these sentences. The authors of the paper almost use the same method to check the obscene words. The same thing can be done for revisions containing repetitive one-sided or biased phrases. This strategy can efficiently be combined with current Wikipedia knowledge about the fact that “neutral point of view” (NPOV) authors usually sign in with their names when adding a non-biased content.

It is also interesting to see the authors’ suggestion from their observations that in Wikipedia pages, when people cannot convince others of why their edits are valid via the comments they leave, the discussion escalates into the talk pages. This is why 17% of all Wikipedia’s’ pages are of talk pages.

It is also interesting to know that Wikipedia provides rapid turnovers where news can be implemented with a speed that is impossible in a printed version. Also the fact that pages are not going to stabilize over time in term of size reveals that it is expanding and would continue to grow further in future.
Review of “Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia”

The paper introduces the Wikipedia as an open-content encyclopedia. It argues that an important factor of its success is the guiding editorial policy of neutral point of view. However, I think this might not be true since the biased writer won’t consider themselves biased. They, most of the time, believe themselves as people writing the truth; but inside their mind they are so confident in their beliefs that they write one-sided. Besides this, on topics of conflict between nations, members of each nation would revise the page in a way that adopt to their national belief. I believe this would ultimately shift the content of the page toward the nation with more population. As an example, in case of the conflict between Chinese people over Republic of China with Taiwanese, more population of China could distort the truth. In cases of this kind, people won’t consider themselves as biased; they just re-write the page in a way they have been taught by traditions.

The authors chose Wikipedia participants to interview based on looking at pages that list site-wide editing activity and placed recruiting message on their personal talk pages. This way of choosing samples lack from the task that enormous number of messages on their personal talk could be a sample of being involved in an argument with others. This could either be due because of their biased writings or their effort for vandalism. It would be better if the content of the talk pages was content checked to see if they use strong words for denying others and fighting over their one-sided opinion or it is just a normal discussion.

The fact that Wikipedia pages could become featured article was new to me. However since the content of the page is going to be changed over time, after a while it could completely be different than the original one. It is nice to have it on the main page for a day where people can see that a specific page has been reviewed and approved. But keeping their status as features status for always is not going to be a good idea. I think one solution is to change their status from featured article whenever a certain percentage of their content has been changed.

Also the fact that arbitrators can ban individual’s IP or ID from accessing certain sections of the site, could result in more aggressive actions from the banned person. Since one can, almost easily, use other IPs (even if his own is static) than his previous used one, banning based on the IP is not a good idea. I think some studies need to be done in this issue. In case the results show that vandalism is a great problem in the system, which means it consumes too much effort from others, and also if the automatic page recovery algorithms, such as using visualization tools, are not going to being effective, then new solutions should be introduced in the system. An example of such solution could be assigning trustworthy to people. Newbie would start from lowest level of trustworthy. They could be graded by other senior participants based on their acts and could improve their level of trustworthy by being contributors. Certain restrictions, such as intruding a mentor to review the changes made by the newbie, could also be applied to the system.

Papers discussed it in brief that the Wikipedia results would come up in many Google searches. I think this is one of the main reasons for success of Wikipedia. I personally get familiar with Wikipedia first through Google search. After several times being referred to the same website for definition of a topic by Google, I started to check the website by myself even before googling the topic. I think the role of search engines in the success of Wikipedia should be further studied and discussed.