

An Analysis of Political Views on Blogs

December 1, 2008 Update

Todd Sullivan

todd.sullivan@cs.stanford.edu

1 Introduction

The blog crawler that provides our data is still unavailable. Thus, for our last update, we do not have any data around the elections and only have data up to October 28. In this update we briefly examine the friend networks of bloggers with political opinions. We also compare our results with poll data and the actual election results, showing that the blog data provides a reasonable estimate of the actual results at times.

2 Friend Networks

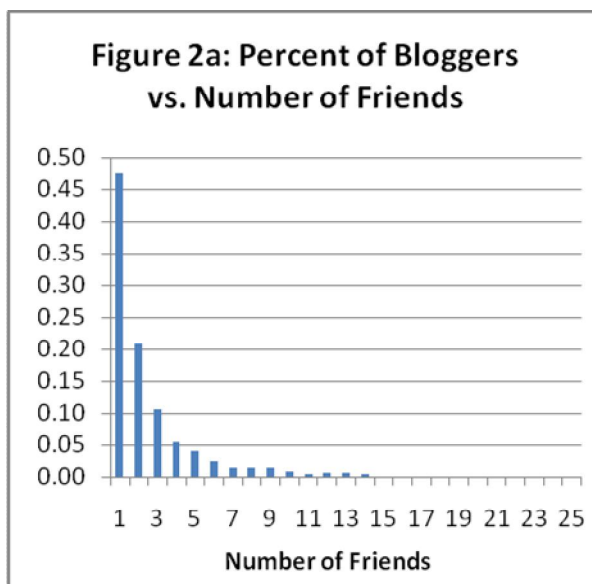
Our data contains limited friend information for the LiveJournal network. We extracted all of the friend links between bloggers that have expressed an opinion about at least one of the classes. For each blogger, we summed the blogger's FeelScores for all classes that the blogger has an opinion on, with the FeelScores for the McCain, Palin, Republican, and Conservative classes multiplied by -1. We then marked bloggers with a positive sum as Democrats and bloggers with a negative sum as Republicans.

Using only bloggers that have expressed an opinion through their blog (not through listing interests or joining communities), there are 523 Democrats and 620 Republicans that are friends with at least one other blogger with an opinion. These 1,143 authors have 1,575 mutual links between each other.

One might expect the authors to largely stay within their own groups, with Democrats befriending Democrats and Republicans befriending Republicans. We found the opposite

to be the case. 64% of all mutual links are between a Democrat and a Republican, 23.6% are between two Democrats, and 12.4% are between two Republicans. Interestingly, even though the data contains 1.19 Republicans for every Democrat, there are almost twice as many Democrat-Democrat links as there are Republican-Republican links.

The majority of these bloggers are friends with only one other political blogger in the network. The percentage of political bloggers that have a given number of political blogger friends follows a power law as shown in Figure 2a. Comparing the frequencies of frequencies with respect to Republicans and Democrats, we see that Democrats are slightly more likely to have more political blogging friends, as shown in Figure 2b.



While 64% of all links are cross-ideology links, using a macro-average, for any given blogger in the group 50.9% of the blogger's friends are of the same ideology. Here, for

each blogger we are calculating the percentage of political blogger friends that have the same ideology as the blogger. We then average these percents to reach 50.9%. Using this metric, Republicans are slightly more likely to be friends with Republicans (51.7%) than Democrats are likely to be friends with Democrats (49.9%). Of course, this difference is small enough that it may simply be due to the fact that there are 100 more Republicans in the group than Democrats.

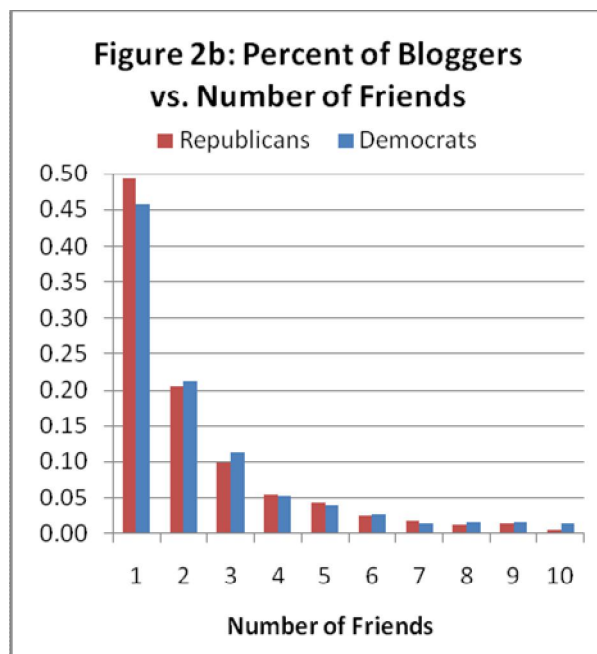
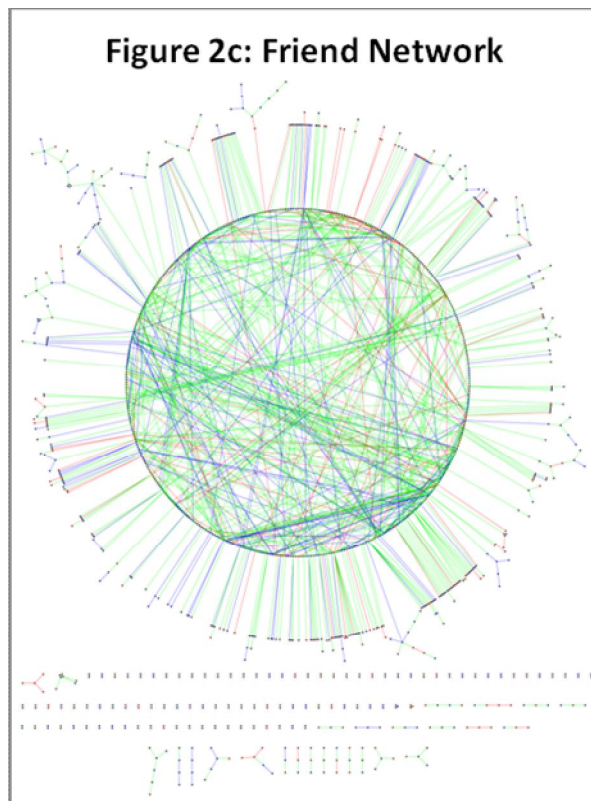


Figure 2c shows a visualization of the friend network consisting of all political bloggers on the LiveJournal network that are friends with at least one other political blogger. Green lines indicate a Democrat-Republican friendship, red lines indicate a Republican-Republican friendship, and blue lines indicate a Democrat-Democrat friendship. We see that the majority of authors are connected to each other in some way. The isolated groups are shown at the bottom of the figure, with most of these groups having a size of two.

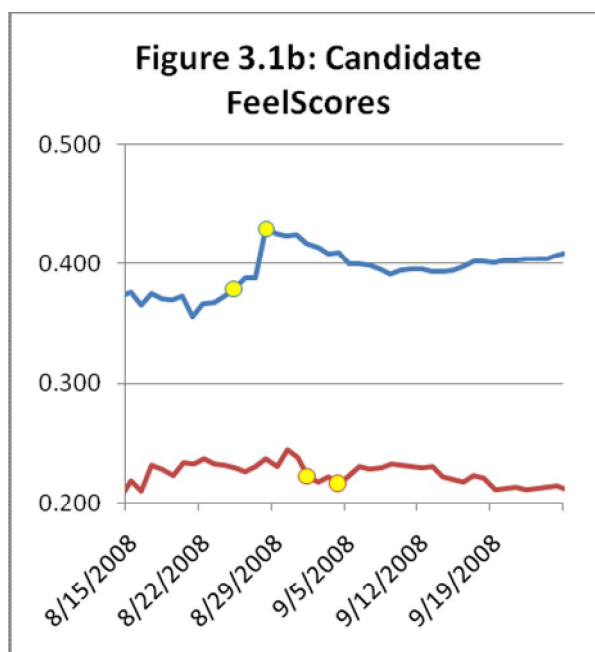
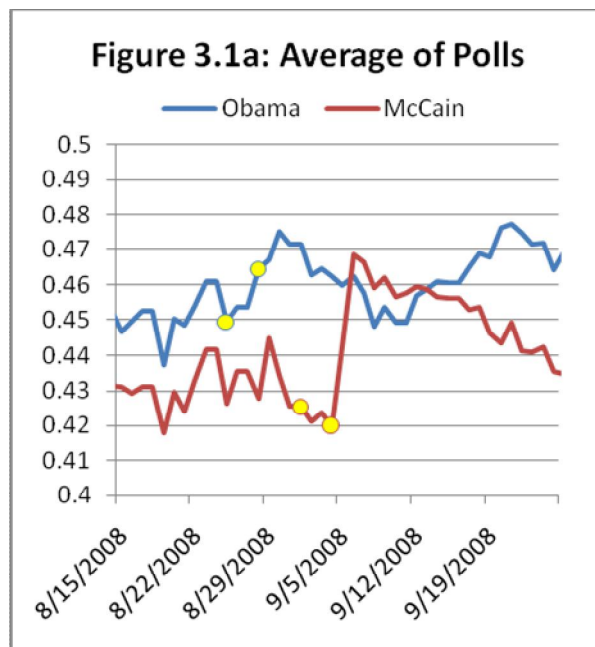


3 Poll Data

We had hoped to match poll data as well as the actual election results with our data around election time. Since we only have data up to October 28, we are limited in our analysis options.

3.1 National Conventions

Figure 3.1a shows averages of nation-wide polls compiled by electoral-vote.com during the Democratic and Republican National Conventions. The start and end of each candidate's national convention is marked as yellow dots. Both candidates show a spike in popularity due to their respective convention, and both spikes are followed by a gradual decline. While our FeelScores do not exhibit these exact trends, Obama's curve follows the rise and gradual fall trend, as seen in Figure 3.1b.



3.2 Election Popular Vote

Obama received 53% of the popular vote in the election on November 4 while McCain received 46%. Since we only have blog data up to October 28, we used all bloggers' opinions on October 28 to calculate the expected popular vote percentage for each candidate.

To determine which candidate an author will vote for, we sum the FeelScores for all classes that the author has expressed an opinion about, with the McCain, Palin, Republican, and Conservative FeelScores inverted (multiplied by -1). We then divide the sum by the number of classes that the author has expressed an opinion about. Thus our final score is a number between -1 and 1 where a number greater than 0.1 indicates a vote for Obama, a number less than -0.1 indicates a vote for McCain, and a number between -0.1 and 0.1 indicates an undecided voter.

We calculate this score for every blogger and assign half of the undecided authors to Obama and half to McCain. At the end, we have mapped all political bloggers to a vote for Obama or McCain, and can calculate a popular vote estimate for each candidate. Table 3.2 shows the result of this process using the blog-based FeelScores for each author.

The result for each row is generated from a sum of the FeelScore classes listed in the row's first column. In each instance, all classes are given equal weight. Our closest result to the 53%/46% in the actual election is when summing the FeelScores for the presidential and vice presidential candidates, which is within one point of the true outcome.

As we add more classes, the popular vote shifts closer towards McCain but never gives McCain a majority. This shifting to McCain is due to the generally more negative blog-based FeelScores for the Liberal and Democrat classes. When including the vice presidential candidates, the drop is due to the fact that almost no one has an opinion about Biden while many authors have an opinion about Palin, and most of the opinions are positive.

| Table 3.2: Popular Vote | | | |
|---|--------------|--------------|--------------|
| Classes Used | Obama | McCain | # of Authors |
| Obama McCain | 56.5% | 43.4% | 3,790 |
| Obama-Biden vs. McCain-Palin | 53.6% | 46.4% | 4,470 |
| Obama-Biden- Democrat vs. McCain-Palin- Republican | 51.8% | 48.2% | 5,217 |
| Obama-Biden- Democrat-Liberal vs. McCain-Palin- Republican- Conservative | 51.2% | 48.8% | 5,832 |

3.3 Election Exit Polls

Table 3.3a shows CNN exit poll results by gender, as well as our own predictions for authors that have listed their gender. Our predictions use the Obama, McCain, Biden, and Palin classes with equal weights in the sum for each author. Our female prediction is within one point of CNN's exit poll while our male prediction is overoptimistic for Obama.

Table 3.3b shows CNN exit poll results by age with our predictions using the same classes and weights as before. In this case, our predictions are not nearly as close to the exit poll data. Our data still has the same trend as the exit polls with younger people preferring Obama while older people are more neutral or slightly towards McCain.

| Table 3.3a: Exit Polls by Gender | | | |
|----------------------------------|-------|--------|--------------|
| CNN Exit Poll Results | | | |
| Gender | Obama | McCain | # Polled |
| Female | 56% | 44% | 9,453 |
| Male | 51% | 49% | 8,740 |
| Predicted Results | | | |
| Gender | Obama | McCain | # of Authors |
| Female | 55.4% | 44.6% | 261 |
| Male | 54.1% | 45.9% | 636 |

| Table 3.3b: Exit Polls by Age | | | |
|-------------------------------|-------|--------|--------------|
| CNN Exit Poll Results | | | |
| Age | Obama | McCain | # Polled |
| 18 to 29 | 67% | 33% | 3,210 |
| 30 to 44 | 53% | 47% | 5,083 |
| 45 to 64 | 50% | 50% | 6,689 |
| 65+ | 46% | 54% | 2,854 |
| Predicted Poll Results | | | |
| Age | Obama | McCain | # of Authors |
| 18 to 29 | 60.4% | 39.6% | 476 |
| 30 to 44 | 56.8% | 43.2% | 294 |
| 45 to 64 | 57.0% | 43.0% | 135 |
| 65+ | 33.3% | 67.7% | 36 |

4 Conclusion

Our analysis in this update was limited due to not having blog data after October 28, yet even using all opinions from October 28 we were able to make predictions on the popular vote that are close to the actual results of the election and exit polls. We also briefly examined the friend networks of political bloggers and showed poll results that follow the trend that we have mentioned many times with the public's opinion of a candidate rising sharply due to the candidate's national convention and then slowly dropping off. In our next report, we will compile all of our results from these updates into a cohesive summary of our most interesting findings while studying political views in blogs.