CS224W Project Report*

Kaiping Chen, Lucas Alexander DiSanto Sanchez

December 12, 2016

1 Introduction

With the rapid development of the Internet in the past decades, the Internet has become a significant tool for civic discourse (Holt 2004; Puopolo 2001). An increasing number of social network sites, including Facebook, Twitter and Reddit, have become high-traffic venues for discussion on any and all topics. As a result, more web users are exposed to and choose to engage in online discussion about social and political issues more frequently. Despite this rapid growth of the social network platforms, scholars still have little understanding about the factors that influence the quality of online political discourse. In this paper, we are particularly interested in the effect of network factors: how the nature of an online network influences the civility and the diversity of online deliberation (Kushin and Kitchener 2009). As Himelboim 2011 reveals, patterns of online participation follow the power-law distribution and are particularly manifested when the group size is large. This hierarchical social network structure constrains people's equal rights to be informed and to participate in the online discourse. Besides the power-law structure which leads to inequality, online social networks also tend to fuel the formation of the echo chamber, where people expose themselves to opinions that they agree with. Consequently, the like-minded people cluster together in their homophily (McPherson, Smith-Lovin, and Cook 2001; Centola 2010). Thus, there is a paradox there: on one hand, the Internet provides many more channels for people to express opinions and to expose themselves to diverse information; on the other hand, it foments the dominance of **the powerful few** and **the reinforcement of the like-minded**. To offer evidence towards this understudied topic of how network structure influences the quality of online discourse, our paper examines all of the posts in the Reddit/Politics during the year 2014. For each post, we generated two types of graph: the author (node) graph and the comment tree graph. We computed the power-law distribution, reciprocity, the breath and depth of the conversations of each Reddit/Politics post. We utilized Natural Language Processing to compute the (emotional) polarization of all the comments of each post and the similarity of all the comments of each post to measure the quality of the discourse.

Our main findings are: 1). We did not find a severe power law distribution among the majority of the posts. Instead, the network structure on Reddit/Politics is consisted of different clusters and extensive interaction within the cluster. 2). When the interaction is more extensive, the discourse tends to be more emotionally polarized but more diverse as well. 3). The power law distribution contributes to a more emotionally polarized discourse.

2 Related Work

Scholars have contrasting views on the structure of social network platforms. Some argue that the power-law network structure hinders the equality of online discourse. Due to the fact that online discussion is unrestricted and the social network is very large in size, the discourse is usually dominated by the few and as a result, the content of the discourse echoes with the perspectives of the powerful few (Hill and Hughes 1999; Wilhelm 2000; Zaphiris and Sarwar 2006). Differently, other scholars (Gómez, Kaltenbrunner, and López 2008) found that some online social network platforms, such as the Slashdot community, do not suffer the power-law problem. These platforms have a moderation system that award good quality comments

*We would love to sincerely thank Will Hamilton for offering us valuable suggestions during the whole project.

and punish uncivil comments. Due to this moderation system, posts tend to receive many replies with an exponential decay in the number of replies to each comment at each level in the comment tree.

We chose to study the network structure of Reddit/Politics because this subreddit is a platform in between restricted and unrestricted platforms as debated by the above scholars. It is restricted because of its voting system to award popular posts and to punish uncivil posts. It is unrestricted because unlike users on Slashdot, who care about the scores they get on their reviews and comments, users on Reddit/Politics might not care as much about whether their comments are down voted or up voted. More importantly, Reddit's scoring system does not inform us about the quality of the discussion happening within each post. A popular post can be an echo chamber where many users with extreme opinions express their like-minded opinions and upvote each other.

Most of the literature on studying social network platforms focuses on investigating their network structure, while less of it investigates the discourse quality and even less explores the relationship between the two. Therefore, our paper contributes to the literature on online social networks by exploring how the network structure influences the quality of discussion: *What kind of network (interaction) structure can facilitate a less polarized and more diverse conversation among users of online social media platforms?*

3 Method

3.1 Generating Network Graphs

Our dataset consists of all the comments in the Politics subreddit during the year 2014 (N= 2,333,826 comments, comprising 70,856 posts), attained from from snap.stanford.edu. The dataset contains information about each comment: most notably, the comment text, user id, comment id, parent comment id, score, and time stamp.

The data must be cleaned before it can be used to generate graphs that represent posts. Because the dataset consists of all comments in the Politics subreddit during a certain time span, it contains comments that were part of posts made before 2014. Since these posts were made before 2014, we do not know if they also contained comments from before our dataset began, and thus we discarded all comments from such posts, thus ensuring we only process posts that were created in the year 2014 and are thus complete.

We generated two kinds of graph for each post. One is the author-author-relationship structure of each post, or the "**author graph**". In this graph, nodes represent Reddit users who authored comments, and a directed edge from node A to node B indicates that user A commented on a comment made by user B. This graph is a multigraph with directed edges, so user A can make multiple comments to user B, and those comments are distinguishable from comments from user B to user A. The other kind of graph we made for each post is the "**comment graph**" of each post. This graph is also a directed graph, but in this graph nodes represent comments, and an edge from node A to node B means that comment A is a reply to comment B. This comment graph has one or more connected components, each of which has a tree structure we will call a "**comment tree**". The root node of each such comment tree is what we will call a "**top-level comment**", meaning a comment that is a direct reply to the post itself, and therefore has no parent comment.

By creating a comment graph in addition to an author graph, we are able to capture a wider variety of metrics about the conversation structure (i.e.,interaction structure) in each post. The author graph allows us to capture metrics about the interactions among users, including the clustering coefficient and the degree to which the distribution of in-degrees follows a power law. The comment graph allows us to examine the structure of the conversation itself, most notably the breadth and depth (or height) of the comment tree, as well as the number of comment trees and the nestedness[1] of the post.

3.2 Measuring Network Structure

We computed the basic network properties for the author graph for each post: the total number of nodes, the total number of edges, the total number of identifiable clusters[2], the percentage of nodes that have 0 or 1

[1]We define the nestedness of a post to be the sum of the heights of all the comment trees in a post.

[2]We use the python package "community" to identify the number of clusters in a post and to compute the modularity of each post. We found that 50% of the posts do not have a community structure, with mod < 0.3. The python package for community detection is located at: http://perso.crans.org/aynaud/communities/

in-degree, the average InDegreeOutDegreeRatio of all the nodes, and the clustering coefficient of the graph. We also computed for the author graph of each post two additional interesting properties: the degree to which the node-degree distribution follows the power law, and the reciprocity among nodes. It is important to note that we used the Gini coefficient to estimate the extent to which the power-law applies to the degree distribution. The rationale for using the Gini coefficient as a proxy is that the Gini distribution is very similar to the power law distribution. A power-law distribution describes"a function that decreases as k to some fixed power, such as k^{-2} in the present case (Easley and Kleinberg 2010)." The Gini coefficient measures the inequality among values of a given frequency distribution and describes how heavy-tailed a distribution is. Distributions that are more unequal are also more power-law-like (heavy-tailed). The perfect equality is when the Gini coefficient is equal to 0 and the perfect inequality is when the Gini coefficient is equal to 1. Therefore, the larger the Gini coefficient, the more power-law-like the distribution is. We computed the Gini coefficient[3] for each post and we also assigned the author network of each post to one of three categories: Inequality (Gini coefficient 0.6 or above), Medium (Gini coefficient between 0.3 and 0.6), and Equality (Gini coefficient 0.3 or below). The other additional interesting property we calculated for the author graph was the reciprocity between the authors in a post. This is a measure of how frequently an author A who replies to an author B's post also has a response from author B on one of his comments somewhere in the same post. The reciprocal responses could potentially be in the same chain of comments, or they could be in separate parts of the conversation - what we are trying to capture with this metric is mutual interaction between two authors. Reciprocity ranges from 0 to 1 and can be interpreted as the percentage of comments in a post that are between a pair of authors who have each responded to a comment made by the other at least once.

We also computed the network properties for the comment graph for each post. CommentCount is the total number of nodes in the comment graph. AvgTreeBreadth measures branching and breadth of conversation, and is the average number of children over all comments that are not leaf-nodes; i.e. the average in-degree over all non-leaf nodes. AvgTreeDepth is the average tree height over all separate components (comment trees) in the comment graph, and represents conversation length or depth. BreadthDepthRatio is defined as AvgTreeBreadth divided by AvgTreeDepth, and gives a sense of the trade-off a post makes in breadth and depth of conversation. A ratio value that is larger than 1 means that conversations in a post have more breadth than depth, indicating that the conversation goes in many different directions and explores many perspectives, but on average goes into less on each one. A ratio value that is smaller than 1 means that conversations in a post have more depth than breadth, indicating that the conversation structure happens among fewer topics but with more depth on each of them. Nestedness is the sum of the tree depths over all comment trees in the post, and is similar to AvgTreeDepth, but does not normalize by the total number of trees. Lastly, CommentTreeCount is the number of separate unconnected components in the comment graph, or in other words the number of comment trees in a post. Note that this is equivalent to the number of top-level comments in the post.

3.3 Measuring Discourse Quality

Discourse, for our purposes, refers to the communication and conversation going on within a post. Therefore, to be considered as a discourse, there should be at least be two edges[4] in each post. We filtered out all posts from our data that have fewer than two edges. This exclusion strategy also helps us exclude posts that have only no nodes, 1 node and 2 nodes. It decreases the outlier effect when we calculate network properties.

We define a high quality discourse as a conversation that expresses diverse perspectives and is not overly emotional. If the comments are very emotional, this means that commenters discuss the topic with each other not with reason, but with extreme sentiments. Similarly, if the comments are very similar among each other, this means that the conversation either happened among like-minded individuals. It could also be the case that the conversation participants are not necessarily like-minded to begin with, but instead they

[3]we use the python code in this source for computing the gini value for each post: http://www.npcompleteheart.com/posts/calculating-a-gini-coefficient.html

[4]Note that the author graph and the comment graph for a given post have the same number of edges. The number of edges in each represents the total number of comments minus the number of top-level comments. This means that we do not have edges in the author graph between each top-level comment author and the author of the post. This is partly because of the limitation in our dataset that we do not have access to the author of each post; only the author of each comment.

tended to agree with each other and mimic others' viewpoints. A high quality online discourse should be as rational and as diverse as possible.

To measure the similarity of the comments in a given post, we used the python scikit-learn package and applied the Tf-idf algorithm to compute the cosine similarity value among any of the two comments in a post. The similarity score of a post is obtained through averaging all the of similarity scores among all pairs of comments. The similarity score is a number between 0 and 1. The larger the number, the more similar the comments are of a given post.

To measure the emotional polarization of the discourse, we used the SentiStrength[5] software to compute the sentiment of each comment. Emotionally polarized discourse means that there is a substantial number of comments at or near each end of the distribution of emotion. The epitome of a polarized distribution is the bimodal distribution (Figure 1d). To mathematically measure polarization, Schmitt 2016 points out that polarization can be measured by "the squared distances of parties' ideological positions to the ideological center." In fact, this squared distance is very similar to the standard deviation. As in Figure 1a (normal distribution, low polarization), the standard deviation is low. In Figure 1b and Figure 1c where there is more skew, standard deviation may be slightly larger than in Figure 1a. Figure 1d is the case with the largest standard deviation because most comments are far away from the mean. Therefore, the emotional polarization of the discourse varies like standard deviation: the larger the standard deviation, the more polarized. To approximate a post's emotional polarization, we calculated the sentiment score of each comment and then calculated the standard deviation of the sentiment scores for the post.

3.4 Analyzing Network Structure and Discourse

After filtering out posts with fewer than two edges, there were 16,166 posts remaining. The unit of analysis is a post. The independent variables are all the network properties. We have two dependent variables: the emotional polarization of the discourse and the similarity of the discourse.

We used forward selection algorithms to select the best suited model and ran an Ordinary Least Square regression for each dependent variable.

4 Results

4.1 The Network Structure of Reddit/Politics

As we noted in the Methods section, we exclude posts that cannot be defined as a discourse (posts with fewer than two edges) because when a post has very few edges, it distorts our network property computations and it cannot be considered as contributing true discourse.

Table 1 reports the basic statistics on the distribution of network property values over all the posts in our processed dataset. The mean value of the Gini coefficient is 0.41 and the median is 0.44, indicating that there is no severe power law distribution in the Reddit/Politics network. Table 2 further shows that among all the posts, only 16% have a severe power law distribution (Gini coefficient above 0.6). The mean reciprocity value is 0.55, indicating that 55% of comments in the post were between pairs of authors who have each, in the same post, commented on a comment made by the other. The mean modularity is only 0.23, which indicates a very weak community structure in Reddit/Politics. Examining the conversation structure using the comment graph, we found that the mean number of comment trees per post is 4.45, which means that on average, there are 4 conversation clusters in each post. The mean BreadthDepthRatio is 0.33, indicating that the conversation structure is generally more horizontal than vertical in Reddit/Politics. This means that each comment gets more replies on average, and the conversation is more likely to cover a broader array of perspectives, but overall, ideas are pursued to a lesser depth and explored in less detail.

Table 3 reports the basic statistics on the discourse indicators among all the posts. The mean value of similarity is 0.04, indicating that there are very diverse perspectives in Reddit/Politics. Rather than an echo-chamber, Reddit/Politics is more a platform for the expression of multiple perspectives. The mean value of the emotional polarization is 0.99. This is not a high level of polarization, because the sentiment score distribution ranged from -4 to 4. A normal distribution of polarization with a standard deviation of

[5]http://sentistrength.wlv.ac.uk

0.99 could fit 95% of its mass on one side of the sentiment score scale (between -4 and 0, or 0 and 4). The way we define polarization, if all nodes share the same kind of sentiment (for example, positive sentiment), it is not a very polarized distribution. If we center the normal distribution at 0, 95% of a normal distribution's mass could fit approximately between -2 and 2, which are moderate scores. Since 95% of the mass would be between those two moderate scores, most of the scores would be smaller than moderate. Even if the distribution were a bimodal distribution with mean 0, with a standard deviation of 0.99, its peaks would each be less than 2 away from 0: moderate-to-low scores. Thus, a mean polarization of 0.99 conveys an overall small amount of polarization.

We also explored the network structure of the post that receives the largest number of comments in our dataset: 203 comments. Figure 3 is a visualization of the author graph of this post. Figure 3 shows that there is a very focused and reciprocal discussion among some authors, but it is hard to conclude that this network exhibits a power law degree distribution because the center and majority of conversation does not happen among only a handful of authors, but rather in a substantial cluster that includes a significant proportion of all authors in the graph. We also observe that there is a moderate number of individuals at the peripheral of the network, having fewer interactions with the core and having some isolated interactions among each other.

4.2 The Relationship between Discourse Polarization and Network Structure

Table 4 reports the regression analysis results on the effect of network properties on the emotional polarization of the Reddit/Politics discourse. We had several interesting findings.

First, we found that the number of total edges in a post is negatively associated with the extent of emotional polarization of the discourse and the coefficient is significant at the 0.01 level. This indicates that larger posts and more interaction among users contribute to a less emotionally polarized discourse. Nevertheless, the factor of "Total Edges" neither informs whether the interaction happens among a few individuals and is dominated by the few; nor informs whether people reply to each other extensively or reply to each other just once and do not engage in further discussion. In order to capture in more detail the interaction structure on Reddit, we examine the impact of the indicators of PowerLawGini and BreadthDepthRatio.

In terms of the impact of power law, we found that there is a positive association between PowerLawGini and the emotional polarization of the discourse, and that the coefficient is significant at the 0.01 level. This indicates that with the increase in the discourse dominated by the few, emotional polarization also increases. This is reasonable because the existence of power law restrains the expression of diverse perspectives and reinforces the opinion of the few, thus contributing to a more sentimental discourse rather than a rational discourse. We also categorize the power law distribution of the posts into "Unequa", "Medium" and "Equal" based on the value of the Gini coefficient. Figure 2 compares the extents of emotional polarization under different levels of power law network structure. Using the t-test, we found that compared to the equal structure, the medium structure increases the emotional polarization by 15% (p-value = 2.2e-16); compared to the medium structure, the unequal structure increases the emotional polarization by 4.36% (p-value = 2.824e-13).

In terms of the impact of the conversation structure, the indicator BreadthDepthRatio tells us whether the communication happens more horizontally or more vertically. "Horizontal" means that a comment receives many replies but there are few further replies to the replies, thus the tree is short and wide. This indicates an interaction that is more like a broadcasting network structure, with many individuals replying to a comment but very little interaction among the individuals. "Vertical" means that a comment, though it receives fewer replies, has extensive follow-up discussion following those replies. The coefficient of the BreadthDepthRatio is negative and is significant at the 0.01 level, indicating that when the interaction becomes more horizontal than vertical, emotional polarization decreases. This implies an interesting phenomenon: when the interaction is less deep, emotional polarization is less likely to happen. Yet, when the interaction is more extensive, the discourse is more likely to polarize in sentiment.

Although we found that a more extensive interaction contributes to a more polarized discourse, it is important to note that the depth of the conversation only captures the amount of replies toward each comment but **not** whether the reply is reciprocal or not. We can imagine two interaction situations: one is a comment receives many replies and further replies to the replies but these replies are from different individuals; the other is that a comment receives many replies and the further replies are the responses from

the previous level of nodes to their descendants. When we interact reciprocity with BreadthDepthRatio when forming our regression analysis, we found that holding the BreadthDepthRatio constant, the more reciprocal the interaction is, the less polarized the discourse is.

4.3 The Relationship between Discourse Similarity and Network Structure

Table 5 reports the regression analysis results on the effect of network properties on the similarity of the Reddit discourse.

First, the adjusted R square of the model is much higher than the regression model on discourse polarization, indicating that network properties have more explanatory power on the extent of similar (diverse) the discourse than the extent of emotional polarization.

Second, we again examine the impact of power law and conversation structure on discourse similarity. The coefficient of the PowerLawGini is negative and significant at the 0.01 level, indicating that when the conversation is more dominated by the few, the similarity among discourse unexpectedly decreases. Conventional expectation is when the discourse is dominated by the few, it is more likely for people to have similar ideas because they are exposed to the comments by the few rather than being exposed to diverse perspectives. However, the negative coefficient informs us that on Reddit, even if the conversation is dominated by the few, people still express their own perspectives rather than tending to agree with and thus mimic the discourse of the few.

In terms of how the conversation structure affects discourse similarity, we found that the coefficient of the BreadthDepthRatio is positive and is significant at the 0.01 level. This indicates that a more horizontally-oriented but less vertically-oriented (extensive) communication increases the similarity of the discourse. This means that without extensive communication, people tend to agree with each others' ideas, or repeat the same things. This is reasonable because a more extensive communication helps people to discover the differences in their perspectives through giving people more opportunities to talk about their perspectives and exchange opinions more extensively with others. Different from the negative impact of extensive interaction on emotional polarization, we found that extensive interaction can have a positive impact on facilitating diverse perspectives. In particular, the interaction term of BreadthDepthRatio further shows that if the extensive conversation is reciprocal, it is more likely to have a diverse discourse.

Therefore, the above regression results reveal the **double-edged** effect of the vertical conversation structure: on one hand, a more extensive conversation can reinforce people to become more emotional; on the other hand, a more extensive conversation can facilitate the expression of multiple perspectives. This double-edged effect echoes with the theory of deliberation and conflict resolution. When people are given more opportunities to engage extensively with others, conflict can deepen because people can become more emotional, yet conflict can also soften because people can understand more about each others' perspectives. The above regression results also reveal that when the network structure becomes more dominated by the few, the discourse is more likely to be more emotional.

5 Conclusion

Through examining the relationship between network structure and discourse quality on Reddit/Politics 2014, our paper offers the first systemic analysis that investigates how network structure affects the emotional polarization and similarity of online discourse. Our paper reveals the double-edged effect of the vertical conversation structure on discourse quality. Extensive interaction can make individuals become more emotional but can also make individuals understand each other more and thus allow more opportunities for diverse perspectives to be expressed. In particular, if the interaction is reciprocal, it mutates the emotional polarization and strengthens the opportunity for diverse views to be expressed.

One limitation of the paper is that we did not explore the omitted variables to control them in our regression analysis. Except for the network structure, other things such as the composition of individuals in different posts can also affect discourse quality.

In order to complement this limitation, our future agenda is to explore the causal relationship between network properties and discourse quality. It can be approached in two ways. First, we can use experimental method to randomly assign individuals into different network structures and compare the quality of their

comments. The second approach is to observe the discourse quality over a certain period of time. If we can find evidence that the individuals who discuss on Reddit/Politics do not vary over time, i.e., the community demographics are the same over time, then since we have the time stamp for each post, we can utilize time series analysis methods to explore how changes in the network structure over time affect discourse quality.

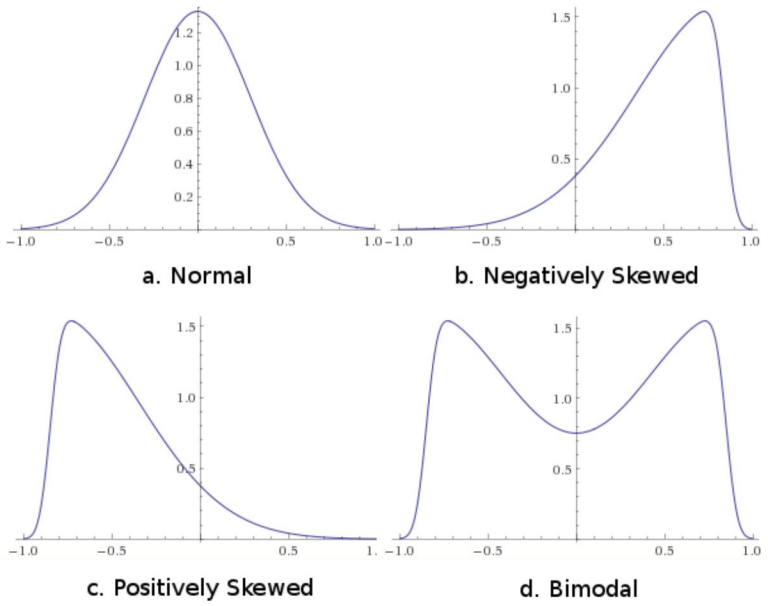

Figure 1: Distributions with Varying Levels of Skew and Bimodality

Table 1: Network Property Distributions

	Min	Median	Mean	Std	Max
AuthorCount	1.00	6.00	7.86	6.65	105.00
TotalEdges	2.00	4.00	7.99	10.34	187.00
LowDegreePercent	0.00	0.57	0.53	0.26	1.00
InDegreeOutDegreeRatio	0.00	0.76	0.70	0.37	2.20
ClusteringCoefficient	0.00	0.00	0.01	0.06	1.00
PowerLaw	0.00	0.44	0.41	0.19	0.88
Reciprocity	0.00	0.55	0.50	0.35	1.00
Modularity	0.00	0.17	0.23	0.24	0.87
clusterSize	1.00	3.00	4.13	3.15	53.00
CommentCount	3.00	9.00	12.44	12.41	203.00
AvgTreeBreadth	1.00	1.11	1.25	0.36	6.00
AvgTreeDepth	1.00	4.00	5.57	5.67	115.00
BreadthDepthRatio	0.01	0.33	0.43	0.47	6.00
CommentTreeCount	1.00	4.00	4.45	3.52	64.00
Nestedness	0.08	1.00	1.62	1.78	53.00

Table 2: Percentage of Different Levels of Power Law

Power Law category	Inequality	Medium	Equality
Percentage	0.16	0.62	0.23

Table 3: Discourse Property Distributions

	Min	Median	Mean	Std	Max
Similarity	0.00	0.03	0.04	0.03	0.71
Emotional Polarization	0.00	1.00	0.99	0.32	2.45

Table 4: Discourse Polarization and Network Properties

	Dependent variable:
	Polarization
AuthorCount	0.009^{***} (0.001)
TotalEdges	0.0002 (0.002)
LowDegreePercent	-0.017 (0.023)
InDegree	0.033^{***} (0.007)
InDegreeOutDegreeRatio	-0.015 (0.014)
ClusteringCoefficient	0.077^{*} (0.044)
PowerLawGini	0.282^{***} (0.023)
Reciprocity	0.029 (0.018)
BreadthDepthRatio	-0.039^{***} (0.007)
CommentCount	-0.003 (0.002)
Nestedness	-0.003 (0.003)
Reciprocity:BreadthDepthRatio	-0.174^{***} (0.032)
Constant	0.863^{***} (0.020)
Observations	16,166
R^2	0.066
Adjusted R^2	0.065
Residual Std. Error	0.310 (df = 16153)
F Statistic	95.131^{***} (df = 12; 16153)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5: Discourse Similarity and Network Properties

	Dependent variable:
	Similarity
AuthorCount	-0.001^{***} (0.0002)
TotalEdges	0.0002 (0.0002)
LowDegreePercent	-0.012^{***} (0.002)
InDegree	-0.001 (0.001)
InDegreeOutDegreeRatio	0.003^{**} (0.001)
ClusteringCoefficient	-0.012^{***} (0.005)
PowerLawGini	-0.029^{***} (0.002)
CommentCount	-0.00001 (0.0002)
BreadthDepthRatio	0.007^{***} (0.001)
Reciprocity	0.006^{***} (0.002)
Nestedness	0.001^{***} (0.0003)
BreadthDepthRatio:Reciprocity	-0.003 (0.003)
Constant	0.053^{***} (0.002)
Observations	16,166
R^2	0.116
Adjusted R^2	0.116
Residual Std. Error	0.032 (df = 16153)
F Statistic	176.917^{***} (df = 12; 16153)
Note:	*p<0.1; **p<0.05; ***p<0.01

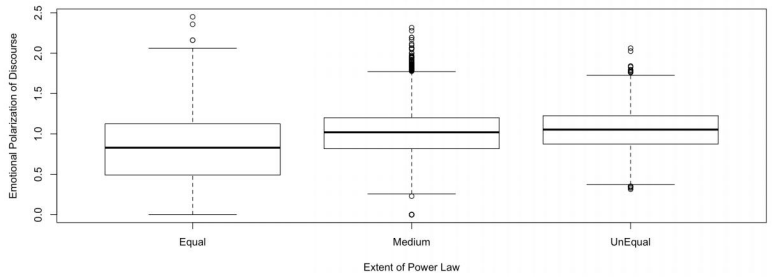

Figure 2: The Extent of Discourse Polarization under Different Levels of Power Law

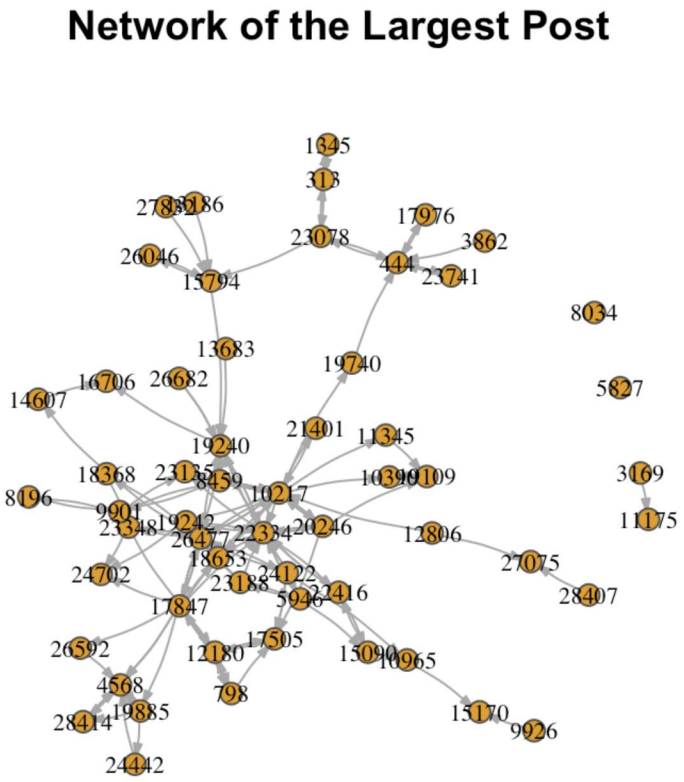

Figure 3: The Network Structure of the Largest Post: Reddit/Politics 2014

References

Centola, Damon (2010). "The spread of behavior in an online social network experiment". In: *science* 329.5996, pp. 1194–1197.

Easley, David and J Kleinberg (2010). "Power laws and rich-get-richer phenomena". In: *Networks, Crowds, and Markets: Reasoning about a Highly Connected World. Cambridge University Press.*

Gómez, Vicenç, Andreas Kaltenbrunner, and Vicente López (2008). "Statistical analysis of the social network and discussion threads in slashdot". In: *Proceedings of the 17th international conference on World Wide Web.* ACM, pp. 645–654.

Hill, Kevin A and John E Hughes (1999). *Cyberpolitics: Citizen activism in the age of the Internet.* Rowman & Littlefield Publishers, Inc.

Himelboim, Itai (2011). "Civil Society and Online Political Discourse The Network Structure of Unrestricted Discussions". In: *Communication Research* 38.5, pp. 634–659.

Holt, Richard (2004). *Dialogue on the Internet: Language, civic identity, and computer-mediated communication*. Greenwood Publishing Group.

Kushin, Matthew J and Kelin Kitchener (2009). "Getting political on social network sites: Exploring online political discourse on Facebook". In: *First Monday* 14.11.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001). "Birds of a feather: Homophily in social networks". In: *Annual review of sociology*, pp. 415–444.

Puopolo, Sonia (2001). "The Web and US senatorial campaigns 2000". In: *The American Behavioral Scientist* 44.12, p. 2030.

Schmitt, Johannes (2016). "How to Measure Ideological Polarization in Party Systems".

Wilhelm, Anthony G (2000). *Democracy in the digital age: Challenges to political life in cyberspace*. Psychology Press.

Zaphiris, Panayiotis and Rifaht Sarwar (2006). "Trends, similarities, and differences in the usage of teen and senior public online newsgroups". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 13.3, pp. 403–422.

Team Member Contribution

Kaiping: Problem formulation, review of related work, writing code to measure network property and discourse similarity, exploring methods to measure power law and community structure, data analyses, writing up the report.

Lucas: Problem formulation, review of related work, writing code to generate author and comment graphs for the posts, writing code to measure emotional polarization of discourse and developing methods to evaluate polarization, filling in sections and revising the report.