Community Detection for the Twittersphere during the Kavanaugh Confirmation Hearings

Antonio Aguilar, Meena Chetty, Rachel Hirshman

December 9, 2018

Abstract

Existing literature surrounding polarization in social networks suggests that communities within these networks are highly partisan, with little interaction across political communities. We study how cross-ideological interaction occurs in social networks and when these distinct, or even non-distinct, communities arise. We focus on Twitter during the confirmation of Supreme Court Justice Brett Kavanaugh, a recent political event that incited conversation between political parties.

We build retweet networks and mention networks for four key dates during the confirmation over a dataset of tweets containing relevant keywords and hashtags. We perform community detection using Louvain and label propagation algorithms in an attempt to replicate Adamic et al. [1] and specifically Conover et al.'s [3] finding about the polarization of these networks. Our hypothesis is that swing senators, a subset of the political elites we will choose, will serve as bridge nodes between communities of opposing leaning. We find that this is true in the case of mention networks, while retweet networks are far more politically segregated and serve as an indication of political alignment. Temporal analysis also shows that those who remain in conversation with regard to the Kavanaugh confirmation over time engage in more cross-ideological interaction.

1 Introduction

Especially recently, scholars have been fascinated by the prevalence of echo chambers, or distinct ideological communities, on social media. Numerous articles seek to determine the existence of these communities and understand their broader societal implications. For the most part, existing literature surrounding polarization in social networks suggests that communities within these networks are highly partisan, with little interaction across political communities. Our analysis does not dispute this reality, and in fact it supports these findings. However, what we deem to be more interesting to analyze are actually those limited cross-community interactions that do exist. What nodes are reaching beyond their own ideological space? Do these cross-community interactions increase or decrease over time? These are some of the questions we seek to explore in our analysis.

To understand factors that influence cross-ideological interaction, we perform temporal analysis on our Twittersphere of choice. We choose key dates during the period of the Kavanaugh confirmation and build separate networks for each of those dates. The dates of interest to us are September 27 (Kavanaugh and Blasey Ford testify in Congress), September 28 (Judiciary committee votes), October 4 (FBI investigation concludes), October 6 (Senate confirms Kavanaugh). It is on these dates when there is significant political discussion on Twitter and likely activity to sway swing senators. To understand the polarization within these networks and to understand the composition of the neighbors of swing senators, we perform community detection. we leverage two community detection algorithms - Louvain and LPA - and compare the results to each other. We find that there is very little discrepancy between the community determinations of the two algorithms on both the retweet and mention graphs across all four dates.

To understand cross-ideological interaction in our networks, we will compare interaction between predominantly liberally classified and predominantly conservatively classified communities over time (i.e. across our sample of date-based networks). Finally to test our hypothesis regarding the political orientation of the nodes mentioning or retweeting swing senators, we use the communities built through Louvain modularity optimization to assess the ratio of liberal to conservative nodes in the set of neighbors for each swing senator.

2 Related Work

2.1 Adamic et al., 2005

Adamic et al. focus on measuring the interaction between liberal and conservative blogs leading up to the 2004 presidential election. The authors gathered a dataset of blogs and balanced the dataset by taking 700 of the largest liberal blogs and 700 of the largest conservative blogs.

From this dataset, the authors built a network of blog activity based on a citation network structure, where one blog cites another if it links to it. The blogs are then assigned a rank based on in-links and out-links, similar to PageRank. The authors also assign pairs of blogs similarity metrics based on content in blog posts. Once they settle on a dataset of the "most popular" 20 liberal and 20 conservative blogs according to the metrics described above, they generate a directed, multi-edged graph of these blogs. The authors then implement a pruning algorithm until there no longer exists a link between a node corresponding to a liberal blog and a node corresponding to a conservative blog.

We employ a similar method using preidentified liberal and conservative Twitter accounts to analyze our network, as discussed below. Additionally, we replicate the pruning algorithm to understand how it performs on a network of Twitter accounts as opposed to blog posts. This helps us verify whether Twitter networks similarly separate into liberally and conservatively leaning supernodes.

2.2 Barberá et al., 2015

Barberá et al develop a correspondence analysis-based method of ideological point estimation that works comparably to a more complicated Bayesian method which is intractable for medium to large graphs of the size of social media networks. They show that the results obtained with their simpler, vectorized method are very highly correlated with the more computationally expensive estimates.

They use the decision to follow another Twitter user to estimate ideology, since that is an expensive signal which is often given to users who align with one's own political beliefs. We use the retweet as a signal, which has been shown to have the kind of polarization that allows ideological score estimation methods to work. Although it proved unfeasible for us to similarly employ correspondence analysis on our retweet graphs, we do leverage Barberá et al's approach to choosing an initial seed of political elites.

2.3 Conover et al., 2011

Conover et al. focus on analyzing political Twitter leading up to the 2010 midterm congressional elections. They seed a sample of 355 million relevant tweets with the two most popular political hashtags on Twitter at the time: #p2 ("Progressives 2.0") and #tcot ("Top Conservatives on Twitter"). They identified the set of co-occurring hashtags for each seed and ranked those using a Jaccard coefficient. After choosing the 55 most pertinent ones, they kept a corpus of 252,300 relevant tweets.

Two networks were assembled with Twitter users as nodes. In the *retweet* network, an edge was placed from A to B whenever user A retweeted content originally from user B. The *mention* network laid an edge if A mentioned user B. In order to establish the large-scale structure of these networks, the researchers performed community detection using label propagation and a greedy hill-climbing algorithm. They were able to conclude that the retweet network contains two clusters of users who primarily bounce around their own content but that the mention network is not similarly clustered and is far more heterogeneous.

In this paper, we replicate the generation of two separate networks - a *retweet* network and a *mention* network. We also expand upon the work of Conover et al. to identify the importance and existence of inter-group connections. Specifically, we are interested in understanding better how two ideological groups are linked and by whom (i.e. which node), and how these links vary over the duration of a major event such as the Kavanaugh confirmation, as this is analysis that the authors do not acknowledge or perform.

3 Approach

3.1 Data

We have downloaded a massive publicly available dataset of tweets related to the Kavanaugh confirmation from pushshift.io. The dataset gathers all tweets between September 22 and October 9, 2018 that present one of the following keywords or hashtags: 'Kavanaugh', #Kavanaugh, 'Supreme Court', #KavanaughHearings, #KavanaughHearing and #kavanaughNomination. The corpus totals 56 million tweets, 3.2 million unique accounts are included within it, and it takes up 315GB of data uncompressed.

Rather than working with all 56 million tweets, we have decided to perform temporal analysis on the tweets by looking at specific key dates from the trial. Doing so not only makes the analysis more manageable, but it is also an approach that no authors to our knowledge have done in a comprehensive manner. Therefore, we look at snapshots of the network throughout the nearly three week period. The dates of interest to us are September 27 (Kavanaugh and Blasey Ford testify in Congress), September 28 (Judiciary committee votes), October 4 (FBI investigation concludes), October 6 (Senate confirms Kavanaugh). For each of these dates, we build mention and retweet networks and perform the same analysis described below on both networks to allow for direct comparison across dates and between the mention and retweet networks.

In this dataset, we have access to the user who posted the tweet, the content of the tweet (including mentions to other accounts), and whether the tweet was a retweet.

We also pre-identified liberal and conservative users by labeling Twitter handles of all U.S. congressmen as liberal or conservative based on their party affiliation. Additionally, we added to this list by including liberal and conservative Twitter accounts with strong followings as identified by news and media sources such as statsocial.com, which has identified the top 100 most influential left-leaning and right-leaning Twitter handles according to their follower base, similarly to how Barberá et al created their own seed set. [2]

3.2 Methods

3.2.1 Networks

Our dataset allows for us to build two basic networks using this data: one based on mentions and one based on retweets, replicating the work of Conover et al.[3] For each date in our analysis, we build the following graphs.

We first built a mention network. The nodes of the network represent all Twitter users that have either authored a tweet or been mentioned in the content of a tweet in our dataset. The nodes are Twitter handles (e.g. @hillaryclinton). We build a directed graph using these nodes to understand the components in our network where an edge of weight 1 exists from user A to user B if user A has mentioned user B in a tweet. If user A mentions user B multiple times, we increment the weight of the edge accordingly.

We then built a retweet network. The nodes of the network represent one of two accounts: a Twitter username that has retweeted another account or a Twitter username that was retweeted by another account. The nodes are again Twitter handles. Again, we built a directed graph where there is an edge between node A and node B if node A retweets node B. The graph is weighted according to how many times node A has retweeted node B, or in the directed case, how many times node A and node B have retweeted each other.

In order to better understand the directionality of our graphs, we also built undirected versions of the graphs described above. A comparison of the number of edges in both the directed and undirected graphs for the mention and retweet networks suggests that for the most part mentions and retweets are both unidirectional. That is, only in a few cases does node A retweet/mention node B AND node B retweet/mention node A. For the purpose of our community detection methods, we treat our graphs as undirected since we are trying to simply measure cross-ideological interaction which can go both ways.

3.2.2 Connected components

We first begin our analysis by looking at strongly connected components in the directed retweet and mention networks.

For each strongly connected component, we identify the number of pre-identified liberals and conservatives in the largest component of each graph to understand what this component represents. We then perform the rest of our analysis on this SCC for each network.

3.2.3 Pruning

In order to better assess the communities and key nodes within the graph, we perform pruning on the full graphs and the SCC, similar to what was performed by Adamic et al.[1]. Edges are removed between nodes if the edge weight is less than or equal to 2. Subsequently nodes that are now disjointed from the core graph given that their edges to other nodes have been removed are completely removed from the graph. Various network metrics and visual representations are then recomputed. This allows us to focus on the more highly connected regions of the graph and more easily visualize the graph.

3.2.4 Community Detection

We implement two community detection algorithms to more robustly understand the political leaning of unseeded nodes in the networks we have built. Community detection as a method also helps us to identify the nodes in our seed which most commonly engage on Twitter with nodes in the opposite community. We then compare these two algorithms to determine which performs better for this problem space.

Louvain Modularity Optimization: The Louvain method for community detection is a greedy maximization algorithm that maximizes modularity in two steps. First, Louvain starts small by assigning nodes to neighboring communities and measuring changes in modularity. The node is then assigned to the community which maximizes the change in modularity. Second, Louvain aggregates the nodes it has assigned to each community into one node. This process is then repeated until no increase in modularity can be achieved. Modularity is defined as:

$$Q = \frac{1}{2m}\Sigma_{i,j}\big[A_{ij} - \frac{k_i k_j}{2m}\big]\delta(c_i, c_j)$$

Label Propagation Algorithm (LPA): We also perform label propagation on the retweet and mention networks. We adapt preexisting implementations from Github and networkx for our purposes. We begin by assigning every node a label of its own id. We then process all nodes in the graph in a random order. Every node is iteratively assigned the label that appears most frequently amongst its neighbors; if there is a tie, it is broken randomly. We continue this process until every node's label no longer changes.

Upon performing the two community detection algorithms on the SCC from each date of interest for both mention and retweet networks, we then use the presence of nodes from our seed set to determine which communities are "liberal" and which communities are "conservative". We then employ various analysis techniques and measurements on these temporal communities to understand and compare them. The results of this analysis are covered later in this paper.

3.2.5 Bridge nodes

We define a bridge node as any node which connects to a node in the the opposite political community (as determined by community detection techniques) with an edge weight greater than or equal to 2. Given that one retweet or one mention does not carry much significance, we have chosen to add this edge weight restriction, similarly to our reasoning behind pruning. Additionally, we choose to focus only on bridge nodes that are within our seed as this provides more interesting and tangible qualitative analysis.

3.2.6 Swing Senators

In addition to analyzing bridge nodes, we are also interested in understanding better the political ideological composition of the nodes interacting with swing senators during this confirmation hearing, to verify or negate our hypothesis outlined in the abstract. The senators we focus on are: Jeff Flake, Susan Collins, Bob Corker, Joe Manchin and Lisa Murkowski. We also include analysis of President Donald Trump. We perform this analysis temporally. We first begin by determining who the neighbors are of each of these Senators for each date. Then, using the liberal and conservative communities generated by Louvain community detection for each date, we calculate how many individuals from each of those communities is interacting with the Senator. Our reasoning for using the Louvain algorithm to analyze swing senators as opposed to LPA is because, in the mention networks, LPA actually generates a single massive community as opposed to more distinct and modularized ones; we cover this in detail in our analysis below. Finally, we compare the political ideological composition of the nodes interacting with the Senator to the party to which the Senator belongs and the way in which the Senator voted on the final vote, and analyze how these numbers vary across our temporal snapshots.

4 Results and Findings

Date	Sept 27	Sept 28	Oct 4	Oct 6
Retweet Nodes	31446	22631	8469	6925
Retweet Edges	423487	314081	94621	66653
Mention Nodes	68909	71278	3627	2180
Mention Edges	1157474	1250618	20644	12497

Table 1: # Nodes and # Edges in SCCs of Mention and Retweet Networks

4.1 Retweet Network

Overall, we see that as time goes on, retweet activity on Twitter related to the Kavanaugh confirmation decreases. This is evident in Table 1, which details the number of nodes and edges in the SCC of each graph from each date.

4.1.1 Community Detection

Date	Sept 27	Sept 28	Oct 4	Oct 6
Louvain	0.47	0.44	0.47	0.50
LPA	0.43	0.42	0.46	0.36

Table 2: Modularities of Retweet Network with Louvain and Label propagation algorithms

Interestingly, however, despite graphs of decreasing size, the Louvain community detection algorithm produces more distinct communities over time on the retweet network. Over time, the Louvain modularity of the SCC increases (Table 2). Additionally, the modularity of the SCC as determined by Louvain is consistently greater than the modularity determined by LPA. The modularity trend from LPA is also opposite that of Louvain where the modularity is decreasing over time (Table 2). Because Louvain is optimizing for modularity and converges based on this criteria, we do expect this trend.

Upon executing the Louvain and LPA methods on each date graph, we then determined how many nodes from our seed set where in each community the algorithm outputted. There is very little discrepancy between the number of seed nodes in the communities produced by Louvain compared to those produced by LPA, as seen in Table 6.

October 6th is a particularly unique case where the largest community generated by Louvain does not actually contain any nodes from our seed set. Originally, we thought this could be due to the stochastic nature of Louvain. However, upon recomputing the Louvain communities numerous times, it became evident that there is in fact a community of nodes on Oct. 6th that are distinct from any nodes in our seed set. Therefore, as evidenced by the graph below, there are three large communities, two of which we can determine to be ideologically distinct based on our seed set, and a third which we cannot classify based on our analytical approach.

4.1.2 Bridge nodes

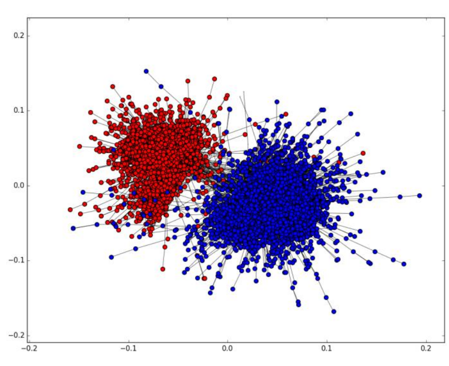
In conducting an analysis of bridges nodes in each of the date retweet graphs, we see that there are more nodes in the conservative community that interact with nodes in the opposite community than nodes in the liberal community (Table 4). For example, on Sept. 27, only one of our seeded nodes in the liberal community, the New York Times, has an edge of weight greater than or equal to 2 to the conservative community. Whereas there are 12 seeded nodes in the conservative community with edge weights greater than or equal to 2 to the liberal community. This pattern suggests that more liberals are interacting with conservative elites than vice versa which makes intuitive sense because most of the key voices during the confirmation hearing were conservatives whether that be congresspeople, news personalities, or other elite politicos.

Of particular interest is who these bridge nodes actually are. One node that appears as a bridge node on three of the four dates is Senator Orrin Hatch. The relative ubiquity of Senator Hatch as such a bridge node suggests that these cross-community interactions are not as much engagement as they are political tools. Senator Hatch is very pro-Kavanaugh, Kavanaugh is even in his profile picture on Twitter, and incredibly active on Twitter. Therefore, those involved in the hearing are not engaging cross-ideologically as a means to cross barriers, rather the bridge nodes just had a really strong presence in their respective communities so the other side sought to leverage and counter what they said to influence their own base.

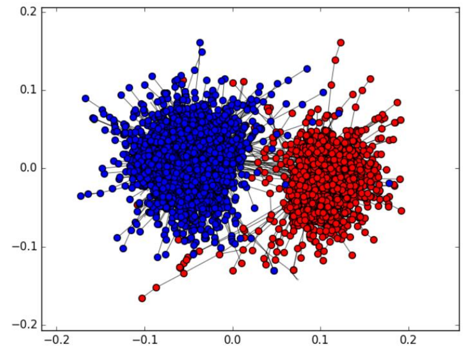
4.1.3 Retweet Network Visualizations

In order to better conceptualize the liberal and conservative communities in the graphs and how the structure of the communities changes of time, we have visualized the SCCs using a spring layout. Before generating the visualizations, we prune the SCC using the pruning algorithm described in the methods section. The visualizations clearly demonstrate two distinct communities - one liberal and one conservative. The nature of the two communities and their separation differs over time, but on the whole the two communities cluster together, unlike what is seen in the mention network discussed in the next section.

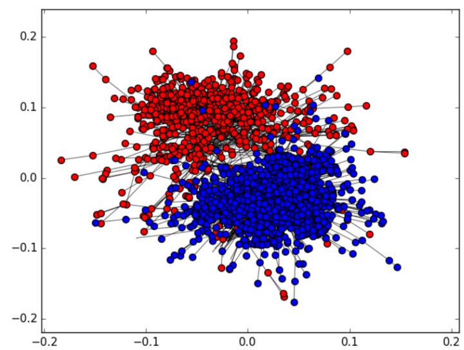
The one evident anomaly is October 6th, whose community characteristics are discussed above. The visualization demonstrates that the the liberal and conservative communities are relatively sparse as compared to the largest community in which there are no nodes from the seed set. A path for future research would be to better understand why this community developed and how it interacts with the liberal and conservative communities; unfortunately that research is beyond the scope of this paper.



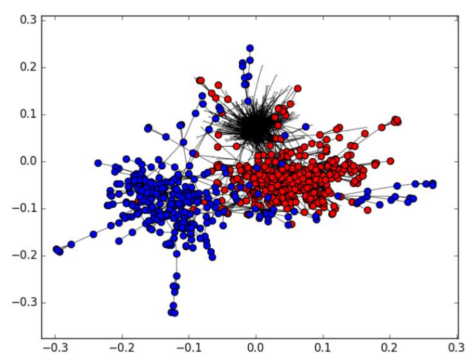
September 27th Pruned SCC - Retweet Network



September 28th Pruned SCC - Retweet Network



October 4th Pruned SCC - Retweet Network



October 6th Pruned SCC - Retweet Network

4.2 Mention Network

We cannot really analyze the Oct. 6 mention graph at least from a modularity perspective because none of our seeds are in the largest communities of the SCC for the mention network. Therefore, for the mention network the focus will be on Sept. 27, Sept. 28 and Oct. 4 for any analysis which involves Louvain. The LPA analysis covers all four dates.

As in the retweet network, mention activity on Twitter related to the Kavanaugh confirmation decreases across our temporal snapshots as seen in Table 1, indicating a peak in activity for both mentions and retweets on September 27 at very the beginning of the controversy.

This is somewhat counter-intuitive but could be a result of a decrease in engagement due to exhaustion or Twitter user being deterred by the strong political nature of the confirmation.

4.2.1 Community Detection

Date	Sept 27	Sept 28	Oct 4	Oct 6
Louvain	0.46	0.43	0.56	0.54
LPA	0.02	0.02	0.43	0.43

Table 3: Modularities of Mention Network with Louvain and Label propagation algorithms

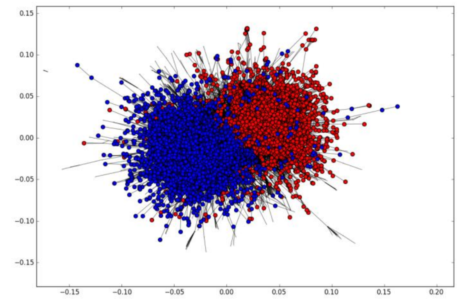
As in the retweet network, the modularity of the Louvain algorithm is consistently far greater than that of LPA for the mention network. Again, this is because Louvain optimizes for modularity. An interesting trait of the LPA's community detection on the mention network is that it consistently detects one massive community that contains a majority of both our liberal and conservative seed sets across all temporal snapshots. This confirms that the communities in mention network are in fact far more heterogeneous and far less segregated based on political leaning. Thus, LPA is not helpful in understanding interactions between preexisting communities of strong certain political leanings, since the mention network inherently contains nodes of both dominant political communities. This lack of clear separation between political leaning is especially evident in the September 27 and September 28 graphs as seen in Table 3, which have modularity 0.02.

4.2.2 Bridge Nodes

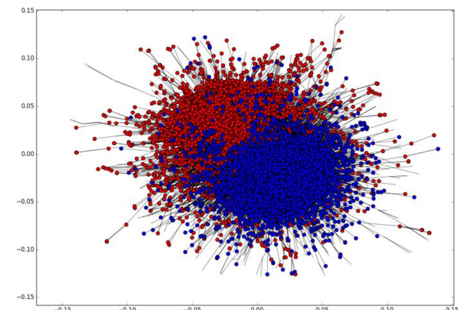
Contrary to the retweet network, there are approximately the same number of bridge nodes in both the liberal and conservative communities when communities are detected via the Louvain algorithm. The greater number of conservative bridge nodes is reflected in the October temporal analysis, again suggesting that overall more liberally leaning people were interacting with conservative political elites than were conservatively leaning people interacting with libearl political elites.

4.2.3 Mention Network Visualizations

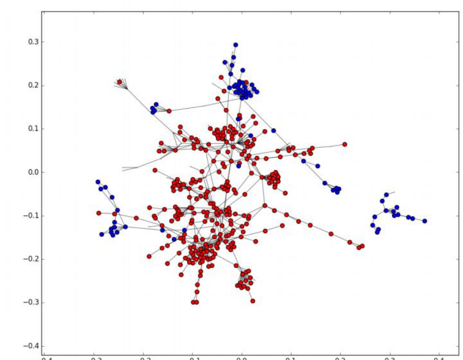
These visualizations confirm the fact that the mention network is far less segregated based on political leaning. The communities in these visualizations were identified by the Louvain algorithm, and labeled based on our seed set as previously described. Similar to the retweet network, the visualization is produced on a pruned SCC.



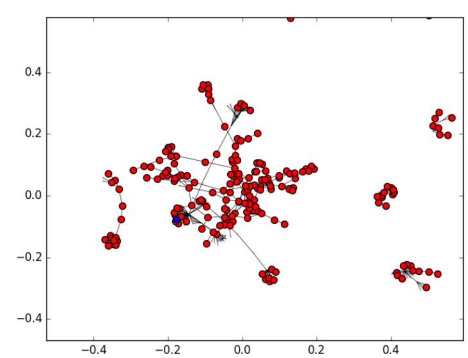
September 27th Pruned SCC - Mention Network



September 28th Pruned SCC - Mention Network



October 4th Pruned SCC - Mention Network



October 6th Pruned SCC - Mention Network

The fact that the mention network is far less politically segregated than the retweet network suggests that it may be more optimal for cross-ideological interaction analysis due to the fact that more liberal and conservative nodes are linked than in a more politically segregated network. We can verify this through our analysis of swing senators, which is discussed in the section.

4.3 Comparing across Networks

4.3.1 Bridge Node Analysis

By looking at the visualizations of the SCCs as well as the modularities for graphs on the same date, it is evident that the retweet network has more distinct communities than the mention network. Further confirming the smaller distinction between communities in the mention network, we also see that there are many more bridge nodes across every date in our analysis in the mention network as compared to the retweet network (Tables 4 and 5). The bridge node distribution between the largest liberal and largest conservative communities identified in the SCC of each network is as follows:

Date	Sept 27	Sept 28	Oct 4	Oct 6
# Liberal Nodes	1	1	2	0
# Conservative Nodes	12	2	5	2

Table 4: Bridge Nodes of Retweet Network

Date	Sept 27	Sept 28	Oct 4	Oct 6
# Liberal Nodes	31	40	1	0
# Conservative Nodes	33	38	7	0

Table 5: Bridge Nodes of Mention Network

4.3.2 Community Detection Comparisons

As seen in Table 6 and as previously discussed, the Louvain and LPA algorithms generate very similar community concentrations in the retweet network. This is not the case for the mention network. While Louvain does detect separate political communities, LPA tends to generate one large community for the mention network, and that community consistently contains more conservative seed nodes than liberal ones; thus, this community is conservatively leaning on the whole as seen in Table 7. This suggests that the mention networks of Kavanaugh related tweets during the confirmation period were very conservatively skewed, which makes sense given that many of the major political players during this event were strongly conservative, such as Kavanaugh and Trump themselves.

Date	Sept 27	Sept 28	Oct 4	Oct 6
Louvain Liberal Size	21895	14903	5047	2578
LPA Liberal Size	21567	15093	5155	2208
Liberal Overlap	21358	14632	4933	2171
Louvain Cons. Size	9273	7144	3211	2406
LPA Cons. Size	9255	7105	2908	4179
Cons. Overlap	9035	6996	2171	2337

Table 6: Intersection between Liberal and Conservative communities generating by Louvain versus Label Propagation on Retweet Network

Date	Sept 27	Sept 28	Oct 4	Oct 6
Louvain Liberal Size	32032	29376	1865	N/A
LPA Liberal Size	0	0	0	N/A
Liberal Overlap	0	0	0	N/A
Louvain Cons. Size	23145	27770	952	N/A
LPA Cons. Size	67128	69330	1318	N/A
Cons. Overlap	23077	27388	388	N/A

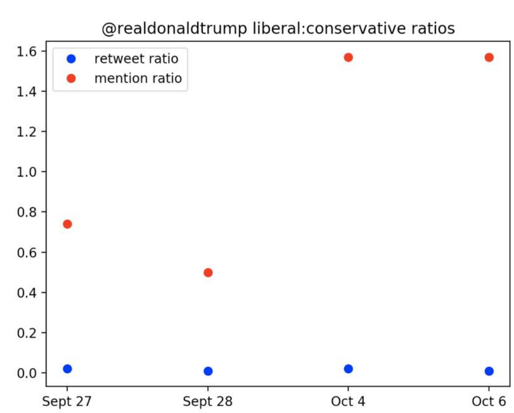
Table 7: Intersection between Liberal and Conservative communities generating by Louvain versus Label Propagation on Mention Network

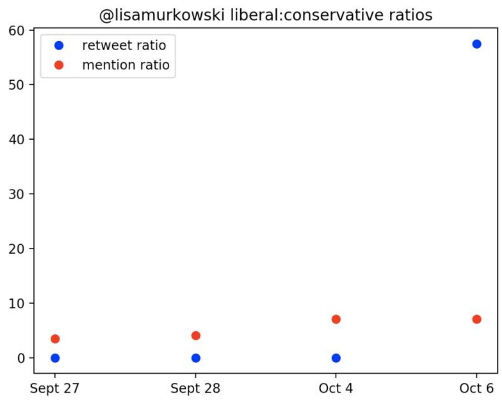
4.3.3 Swing Senator Analysis

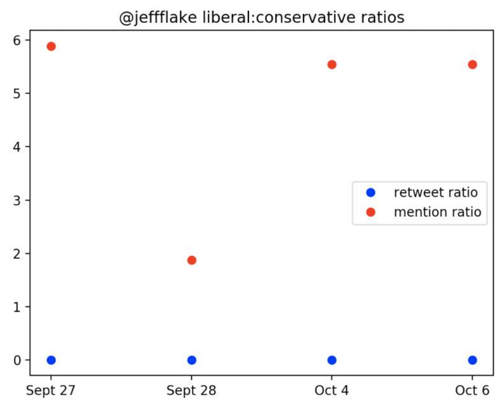
As previously mentioned, the fact that community detection on the mention network generated less segregated political communities leads us to hypothesize that there is more cross-ideological interaction in the mention networks. For each swing senator, we computed the number of neighbors in a liberally classified community and the number of neighbors in a conservatively classified community, and calculated the liberal:conservative ratio. We performed this on both the mention and retweet networks. Some of the most interesting results are shown in the following figures for Trump, Murkowski, and Flake.

As hypothesized, the ratio of liberal:conservative neighbors for swing senators is far greater in the mention network than in the retweet network. In many cases, there are 3-7 times as many more liberal neighbors than conservative neighbors. The fact that so many more liberally leaning Twitter users are interacting with conservative political elites than conservatively leaning Twitter users in the mention network suggests that cross-ideological interaction is pervasive in the mention network. The same pattern is not usually true of the retweet network, as indicated by the figures below. From this, we can conclude that in the Twittersphere, cross-ideological interaction is pervasive in mention networks, while retweets function as more of an endorsement as described in our analysis of Barbera et al's work. [2]

Another interesting trend visible in these figures is that cross-ideological interaction of liberal Twitter users with conservative political elites (i.e. our swing senators) increases over time. As previously mentioned, the size of the mention network decreases over time. In coupling these two facts, we can conclude that while activity decreases across our time period of interest, those who remain involved as the activity dies down tend to engage in more cross-ideological interaction.







We also consistently found the above figures and swing senators to be among the top five nodes of highest degree in each of our networks; they were on the very tail end of the power law distributions of these graphs. (Most notably, Jeff Flake was mentioned over 115,000 times on September 28, the day he was confronted by protesters, including survivors of sexual violence, for several minutes in an elevator on Capitol Hill.) Below are the nodes of highest degree in the October 4 mention graph:

1. Donald Trump: 34526
2. Jeff Flake: 24552
3. Susan Collins: 23581
4. Lisa Murkowski: 21124

5 Conclusion

In this paper, we have presented an analysis of retweets and mentions on Twitter during the time period of the Kavanaugh confirmation. We have taken a novel temporal approach by assessing the presence, or lackthereof, of distinct liberal and conservative communities on four key dates during the confirmation using two community detection algorithms - Louvain modularity optimization and label propagation analysis. We also expand upon previous work on political polarization to better understand cross-ideological interactions over time. We find that those individuals retweeted the most by the opposing ideological group are highly active on Twitter and hold an extreme position in their own ideological group, such as Senator Orrin Hatch. Finally, given the uniquely polarizing nature of the Kavanaugh confirmation, we also direct analysis to key individuals - swing senators and President Trump - to prove our hypothesis that those individuals interacting with these senators (particularly in the mention network) are disproportionately from the opposite ideological community. Additionally, though the size of mention and retweet networks decrease over time with respect to the Kavanaugh confirmation, those who remain in conversation over time engage in more cross-ideological interaction across communities.

6 Contributions

Antonio - Literature review, power law fitting (not included in final paper), graph statistics and poster
Meena - Mention network, community detection comparisons, swing senator analysis, label propagation, writeup
Rachel - Retweet network, bridge nodes, graph visualizations, raw data parsing, Louvain, pruning, writeup
Link to our codebase:
https://github.com/14meenac/kavanaugh_network

References

[1] Adamic, L. A., and Glance, N. The political blogosphere and the 2004 u.s. election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery* (New York, NY, USA, 2005), LinkKDD '05, ACM, pp. 36–43.

[2] Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., and Bonneau, R. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science 26*, 10 (2015), 1531–1542. PMID: 26297377.

[3] Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Menczer, F., and Flammini, A. Political polarization on twitter, 2011.