Mapping the Invocation Structure of Online Political Interaction

Manish Raghavan
Cornell University

Ashton Anderson
University of Toronto

Jon Kleinberg
Cornell University

ABSTRACT

The surge in political information, discourse, and interaction has been one of the most important developments in social media over the past several years. There is rich structure in the interaction among different viewpoints on the ideological spectrum. However, we still have only a limited analytical vocabulary for expressing the ways in which these viewpoints interact.

In this paper, we develop network-based methods that operate on the ways in which users share content; we construct *invocation graphs* on Web domains showing the extent to which pages from one domain are invoked by users to reply to posts containing pages from other domains. When we locate the domains on a political spectrum induced from the data, we obtain an embedded graph showing how these interaction links span different distances on the spectrum. The structure of this embedded network, and its evolution over time, helps us derive macro-level insights about how political interaction unfolded through 2016, leading up to the US Presidential election. In particular, we find that the domains invoked in replies spanned increasing distances on the spectrum over the months approaching the election, and that there was clear asymmetry between the left-to-right and right-to-left patterns of linkage.

1 INTRODUCTION

Political interaction has long constituted a key use of social media, and there is a correspondingly rich history of research into its structure — one that extends the much longer history of scholarship on the role of media in the political process [6, 14, 23, 25].

A crucial issue in this line of work is the extent to which political interaction on social media takes place primarily among users who are ideologically similar, or whether it reaches across the political spectrum. Early analysis of political blogging indicated a clustered structure, with a high density of linkage among ideologically similar blogs and a lower density of linkage between blogs with strongly differing views [1]. Subsequent work, looking at platforms that arose further into the evolution of social media, suggested that a more complex structure was developing, in which homophily in views remained a powerful force, but where the platforms were providing users with some level of cross-cutting exposure [3, 4].

These questions are important, as they ask whether online political interaction consists of opposing sides who engage with each other, or well-separated clusters who are isolated in "echo chambers" or "filter bubbles" [11, 27, 30].

The answers to such questions depend intrinsically on which types of interactions are being considered. Existing work in the online domain has implicitly focused on two standard forms of interaction: page-to-page interaction, expressed by hyperlinks among documents [1]; and user-to-user interaction, expressed by communication among people on social platforms [3, 4, 9]. Each of these induces a network on a set of entities — sources and users respectively — which can then be analyzed relative to an underlying political spectrum.

Networks of sources invoked by users. Here we consider a different type of political interaction network, defined as follows. When a user u shares a page A, and a user v replies by sharing a page B, there is not simply an interaction between users u and v; an interaction is also induced between pages A and B. As reshares develop into a widespread style of social media content production [8, 10, 17, 24], the ability of users to deploy page references as proxies in their discussion becomes an activity requiring very low effort, and we find through a large-scale analysis of Twitter data leading up to the 2016 U.S. Presidential election that such A-B interactions are widespread: users regularly invoke links in this back-and-forth fashion when they interact with each other.

These *invoked interactions* between pages A and B are fundamentally different from both user-to-user and page-to-page interaction networks. Unlike free-form user-to-user interactions, they create logical relationships among the information sources, not just the consumers and sharers of these sources. But they are not like traditional page-to-page interactions either, because they are not based on a hyperlink from A to B, and they are not in general determined by the authors of either A or B; it is the readers *deciding how A and B should be used in discussion* who are determining the logical link between them. In this sense, invoked interactions between A and B are not directly under the control of the authors of A and B; they form a kind of "revealed interpretation" of A and B once they are released into social media.

The idea that replying to page A with page B can create a semantically meaningful connection between A and B formed the basis of an elegant technique due to Frigerri, Adamic, Eckles, and Cheng [12] for identifying pages to debunk widely circulated rumors. Drawing on the fact that snopes.com is a heavily-used site for evaluating Internet rumors, they demonstrated that many instances of a pair (B, A), where A is a page appearing in Facebook posts and B is a page on snopes.com, serves as strong evidence that B is providing a judgment on the credibility of A. In this way, scanning the replies to posts where A occurs for pages residing on snopes.com provides an automated method for identifying a page B that can help users evaluate the veracity of A.
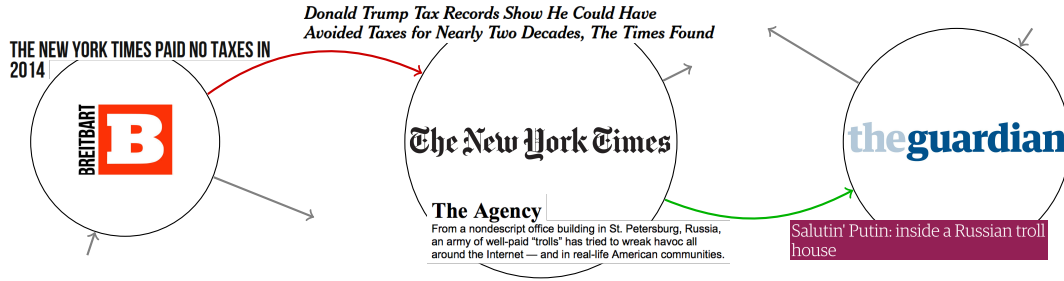
Figure 1: A section from an example invocation graph containing instances from our data. In each case, an article from domain *B* in response to an article from domain *A* contributes to the link from *B* to *A*. The link from Breitbart to The New York Times comes from articles like the pair shown here, a Breitbart article on the Times' taxes in response to a Times article on Donald Trump's taxes. Other links demonstrate an interaction where one article supports another: for example, a New York Times article shared in response to a Guardian article on Russian information operations contributes to the link from The Times to The Guardian.

Given that invoked interactions between pages *A* and *B* more generally are voluminous, transcending any one particular domain or use case, what do we learn if we consider the set of all such interactions as a network in its own right, latent in a social media platform?

The present work: The structure of invocation graphs. We are interested in understanding the global structure of the network of invoked interactions, and developing methods that can probe this structure, particularly for questions related to political interaction. Because our primary focus in this work will be at the granularity of news and blogging *sources* rather than the individual pages they produce, we will consider this network at the level of domains: using a large Twitter dataset covering all of 2016 up to the U.S. Presidential election, we say that an *invoked interaction* from domain *B* to domain *A* occurs whenever a user replies to a tweet containing a page from domain *A* with a page from domain *B*.

We define the *invocation graph* on a set of domains of interest as follows: for all pairs of domains *A* and *B* where there is at least one invoked interaction from *B* to *A*, we include a weighted directed edge (*B*, *A*) whose weight is equal to the number of such invoked interactions. Because we want to study how portions of the invocation graph reflect aspects of political interaction, we choose the node set (i.e. the domains of interest) using a preprocessing step that only includes domains that were extensively retweeted with the official Twitter accounts of Hillary Clinton and Donald Trump, and then apply some further filtering heuristics that we describe in the next section.

After this filtering, we have an invocation graph that reflects the ways in which Twitter users interacted with one another by invoking content from different politically relevant domains over the course of 2016. Figure 1 shows a few domains of such an invocation graph with some sample interactions from our data. We can see these interactions can be supportive, like an in-depth New York Times report on Russian information operations shared in reply to an article from The Guardian on the same subject, or adversarial, like a Breitbart article on the New York Times' taxes shared in response to a Times article on Donald Trump's taxes. These replies

exhibit rich structure and give a sense for how complex political interactions unfold on social media

We can thus return to some of the initial motivating questions and ask how this interaction was structured relative to a political spectrum containing support for Clinton on the left and support for Trump on the right. We do this using a political spectrum induced from the data, rather than relying on external domain knowledge. There are a number of ways to do this [5, 11, 14] that yield broadly consistent results, and we employ a method of Benkler et al (described further in the next section) that bases the spectrum on the relative frequency of co-tweets with the Clinton and Trump Twitter accounts [5].

Embedding the Invocation Graph in a Political Spectrum. Having thus located the domains on a political spectrum, we now have an *embedded* version of the invocation graph: the nodes, representing domains, are embedded in a one-dimensional spectrum, and the weighted directed edges span pairs of points in this spectrum. We can now ask how the edges are distributed across distances on the spectrum, ranging from short-range interactions that connect domains of similar political orientation to long-range interactions that reach across sides. In examining these questions, it is crucial to recall that these links among domains are not defined by hyperlinks in the content on these domains, but instead via the replies made by users when they invoke this content in discussions: it is not that the domains are replying to each other, but that they are being invoked in replies by users. The data thus reflects choices made by the consumers of the content, rather than by authors of the content.

In Section 3, we propose a set of methods to analyze how the edges of the invocation graph span the underlying spectrum. A core component of these methods is, for a domain *x*, to consider its *out-link distribution* — the distribution of "landing points" on the political spectrum for all links out of *x*. As we move from left to right along the spectrum, tracking the out-link distributions of domains we encounter along the way, do the means of these distributions tend to move from left to right as well, or do they tend to move inversely from right to left? The former case would indicate a positive correlation in the locations of the source and target of an invoked interaction (*B*, *A*), suggesting links are used

to connect to similar sides of the political spectrum; the latter case would correspond to a negative correlation and hence connections across the spectrum, with domains on the left being invoked to reply to domains on the right, and vice versa.

It is not a priori obvious which type of correlation we should expect to see; and as a reinforcement of this fact, we find that the nature of the correlation actually inverts over the course of 2016 leading up to the U.S. Presidential election. In the early parts of 2016 we have a positive correlation, with politically similar domains being invoked to reply to each other; but by the time we reach the months directly preceding the election, this same correlation measure has become negative, indicating that most of the linking is now crossing the spectrum. We verify this effect using multiple measures, including one in which we compare the trends across the spectrum to what we'd observe in a *randomly rewired* version of the embedded graph.

We also propose a set of methods to identify inherent *asymmetries* in the patterns of linkage: do replies from left to right have the same structure as replies from right to left? Using our measures, we find strong asymmetries in the 2016 Twitter data, with domains on the right side of the spectrum having a disproportionately high rate of out-links in the invocation graph and domains on the left side of the spectrum having a disproportionately high rate of in-links. This right-to-left flow in the replies persists across the entire time range, and is a key characteristic of the structure.

Since a recurring theme in our analyses is the way in which replies increasingly engaged opposite sides of the political spectrum as 2016 went on, it is interesting to ask whether we see a similar effect in a more traditional user-to-user interaction graph, with nodes corresponding to users and directed edges to replies from one user to another. To explore this, we adapt the techniques developed for the invocation graph to a user-to-user graph built from Reddit. Specifically, we analyze a snapshot of Reddit's politics subreddit, r/politics, for the same period of 2016 up to the election; we classify users by whether they had posted to the Clinton or Trump subreddits, and then look at the rate of replies among different types of users. We find that the trend on Reddit closely tracks the trend in the invocation graph built from Twitter, with increasing linkage between the two sides as the election approached.

Overall, our methodology suggests that the invocation graph on domains, and its embedding into a one-dimensional spectrum, captures important aspects of political interaction on social media — the tendency of users to interact by invoking links to authored content, and the use of these interaction patterns to thus reveal relationships among the content based on usage in everyday discussion.

2 BASIC DEFINITIONS

We begin with a Twitter dataset containing aggregate-level information about tweet-reply pairs. For each month from January to November 2016 (the US Presidential election was held on November 8, 2016), the dataset consists of pairs of domains x_1 and x_2 along with an accompanying count, the number of times a tweet containing a page from domain x_1 was posted in reply to a tweet containing a page from domain x_2. In addition, for each month, we have an auxiliary dataset of co-occurrences: for each domain x, the number of times a user posted a tweet containing a page from x on the same day that the user retweeted Hillary Clinton's or Donald Trump's personal Twitter account. Finally, we have the number of retweets of Clinton or Trump for each month.

The invocation graph we construct from this data is a directed graph with domains as vertices, where each domain corresponds to a news source. We draw an edge $x_1 \rightarrow x_2$ if a tweet containing a URL from domain x_1 is posted in reply to a tweet containing a URL from domain x_2. The weight of this edge is the number of such tweet-reply pairs. On first inspection, the most prominent feature of this graph is that self-loops (consisting of links from a domain x to itself) have much higher weight than other edges. Since our goal is to examine political interactions between domains, we remove all self-loops from the graph.

Isolating Political Domains. The first issue we encounter is that Twitter contains a wide range of URLs, not just pages from politically relevant domains. We could select only known political domains by whitelisting them, i.e. only considering the subgraph over a predefined set of domains; however, this approach will inevitably miss out on influential but less well-known news sources.

On the other hand, there are challenges to a completely unsupervised approach. URLs on Twitter are dominated by social media sites (e.g. twitter.com, facebook.com) as well as content-hosting sites (e.g. imgur.com, bitly.com) which produce virtually no content of their own, but instead host user-uploaded content such as images, links, and text. While the usage of these content-hosting sites would be interesting to study, this is outside the scope of our work.

We begin by blacklisting several known social media and content-hosting domains and remove them from the graph. However, there are plenty of domains that appear on Twitter that are not politically relevant, and we cannot individually remove each such domain. To filter out such domains, we need some measure of *political engagement* for each domain. We can construct such a measure by using the observation that politically relevant domains should frequently co-occur with known political entities – in our case, the official Twitter accounts of Hillary Clinton and Donald Trump.

Our measure of political engagement for a domain, then, is simply the number of times a user posted a tweet with that domain on the same day that he or she retweeted either Clinton's or Trump's official Twitter account. Intuitively, the more politically engaging a domain is, the more it will co-occur with these political entities. With this proxy, we can select domains with high political engagement, excluding social-media and content-hosting domains.

As a final filter, we require that each domain have an edge of at weight least W to some other domain in the political subgraph. This restricts our attention to the most actively used political domains. Based on this, we can formally define the invocation graph. Every domain in the invocation graph

- is not blacklisted (social media and content-hosting domains)
- has political engagement above some threshold p
- has at least one edge to another domain in the invocation graph with weight at least W

We require the invocation graph be connected and contain some seed domain. Any choice of popular American news outlet would

yield exactly the same result, as they all belong to the same connected component. (We use nytimes.com, but the choice is immaterial.)

Based on this definition, the algorithm to construct the invocation graph is as follows: begin with the full graph, remove all nodes that are either blacklisted or have political engagement below p, and run a breadth-first search beginning at nytimes.com following only edges of weight at least W. In practice, we find that the values we use for p and W don't affect our results much, and all of the results presented here use $p = 10000$ and $W = 100$. Because we study the change in this invocation graph over time, we build a new graph G^m for each month m from January to November 2016.

A Political Spectrum. In order to characterize the political nature of this graph, we need some way of organizing the domains along a political spectrum. Drawing on techniques from [5] and our definition of political engagement, we define the quantities $P_T(x|C)$ and $P_T(x|T)$ for each domain x as the empirical probabilities that a user posted a tweet containing a URL from domain x given that, on the same day, he/she retweeted Clinton's or Trump's official account respectively. Intuitively, if a user has retweeted Clinton on a given day, the domains that he/she invokes are more likely to be on Clinton's end of the political spectrum, and the same is true for Trump. Figure 2 shows the resulting values, where the blue line is $P_T(x|C) = P_T(x|T)$. Interestingly, most domains lie above this line, suggesting a difference in the populations of users retweeting Clinton and Trump respectively.
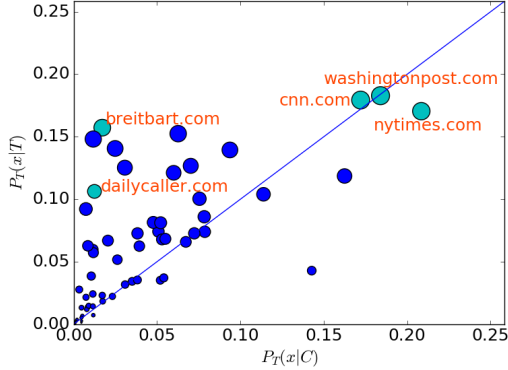

Figure 2: $P_T(x|C)$ **vs.** $P_T(x|T)$ **for September 2016**

Furthermore, we can condense this information into a single *political score* for each domain:

$$s_T(x) = \frac{P_T(x|T)}{P_T(x|C) + P_T(x|T)} \quad (1)$$

Note that $s_T(x) \in [0, 1]$, and the larger $s_T(x)$ is, the closer x is to the Trump end of the spectrum. Throughout this section, we use the spectrum built on January-September 2016.

3 METHODOLOGY

With these definitions, we can analyze various properties of the invocation graph, and in particular, its interaction with the political

spectrum. Do usage patterns of news articles and other political content differ across the political spectrum? How do these patterns change leading up to the election? Using the invocation graph in tandem with the political spectrum, we can shed light on these questions.

Out-link Distributions. A key characteristic we study is the *out-link distribution* for a domain x: the distribution of positions on the political spectrum where edges originating from x land. A domain's out-links describe the way in which it is used to reply to other domains. For example, if a domain has many out-links to other domains near it on the political spectrum, one might expect that it is being used to reinforce a particular point of view; however, if it has many out-links to the opposite end of the spectrum, it may instead be used to disagree or argue with an opposing viewpoint.

To draw upon this intuition, we ask how x's out-link distribution varies based on its political score $s_T(x)$. As a baseline for comparison, we compute the *global* out-link distribution — the distribution over the political spectrum of where edges in G land. Comparing to this baseline will give us some insight as to how x's linking pattern differs from the "average" linking pattern.

However, since G contains no self-loops, x doesn't link to itself, while the global out-link distribution contains links to x. To prevent this from biasing the comparison, we instead compare x's out-link distribution to the out-link distribution of $G \backslash x$, that is, G with x and all edges into or out of x removed.

Thus, x's out-link distribution distribution $d(x, G)$ is a distribution over $[0, 1]$ assigning probability mass to $y \in [0, 1]$ proportional to the weight of edges from x to x' such that $s_T(x') = y$. The global out-link distribution $D(G \backslash x)$ can then be expressed as

$$D(G \backslash x) = \frac{\sum_{x' \in G \backslash x} d(x', G \backslash x) \cdot \text{vol}(x', G \backslash x)}{\sum_{x' \in G \backslash x} \text{vol}(x', G \backslash x)},$$

where $\text{vol}(x', G \backslash x)$ is the total weight of edges leaving x' in $G \backslash x$. In other words, $D(G \backslash x)$ is the weighted average of $d(x', G \backslash x)$ for $x' \in G \backslash x$.

We make the comparison between these distributions formal as follows: for each x, let $\mu_{\text{out}}(x)$ be the weighted average political score of the domains that x links to, so that $\mu_{\text{out}}(x) = \mathbb{E}[d(x, G)]$. This gives an estimate of what types of domains x is used to reply to — if $\mu_{\text{out}}(x)$ is close to 0, then x is used primarily to reply to Clinton-related domains, while if it is close to 1, then x is used primarily to reply to Trump-related domains. In a slight abuse of notation, let $\mu_{\text{out}}(G \backslash x)$ be the weighted average political score of all endpoints of edges in $G \backslash x$, so $\mu_{\text{out}}(G \backslash x) = \mathbb{E}[D(G \backslash x)]$. Then,

$$\delta_{\text{out}}(x) = \mu_{\text{out}}(x) - \mu_{\text{out}}(G \backslash x) \quad (2)$$

is a measure of how much x deviates from other domains towards the Trump end of the political spectrum.

From past work documenting *homophily* in online political activity [1], one might expect that a domain will be used primarily to engage with politically similar domains, implying that $s_T(x)$ is positively correlated with $\delta_{\text{out}}(x)$. Furthermore, it is not a priori obvious whether this trend should become more or less pronounced leading up to the election — do the opposing parties increasingly converse within themselves, or do they engage with one another?
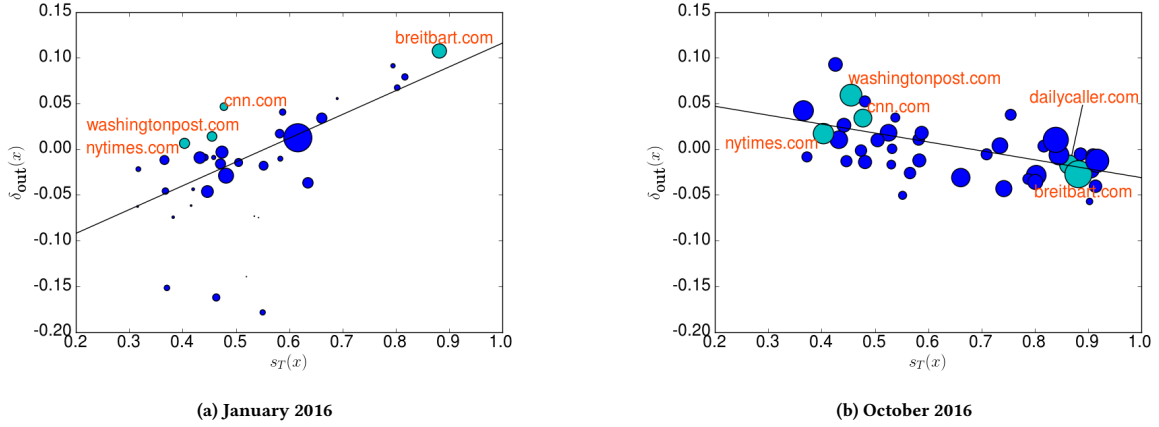
(a) January 2016

(b) October 2016

Figure 3: Correlation between $s_T(x)$ and $\delta_{\text{out}}(x)$ for January and October

Figure 3 shows that in January, domains are more likely to have links in the graph to politically similar domains, so $s_T(x)$ and $\delta_{\text{out}}(x)$ are positively correlated. However, by October, this correlation has reversed: on average, domains are being used to reply *across* the political spectrum instead of to politically similar domains. Replies seem to be adversarial, as with the breitbart.com → nytimes.com edge in Figure 1.

To understand this change in correlation over time, we define $a(G^m)$ to be the slope of the correlation between $s_T(x)$ and $\delta_{\text{out}}(x)$ for month m. Figure 4 shows that $a(G^m)$ decreases significantly from January to October 2016. This suggests that as the election drew nearer, the fraction of interaction between opposite ends of the spectrum increased, and domains from the two opposing sides were actually being invoked more often to reply to each other.
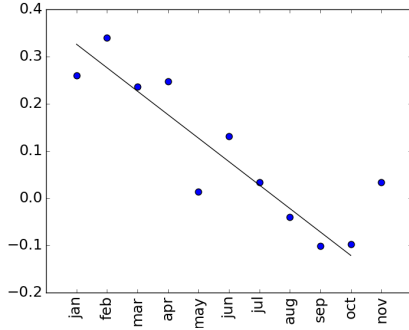

Figure 4: $a(G)$ for January-November

Edge Lengths and Crossing Points. Another way of characterizing the political aspects of the invocation graph is to consider the lengths and locations of the edges on the political spectrum. As the interaction between opposing viewpoints increases, we expect to see longer edges in the invocation graph that cross the political spectrum instead of staying on one side. Figure 5 shows that this trend holds true between January and October — while most links

had length close to 0 in with respect to the political spectrum in January, by October, many longer edges were present.
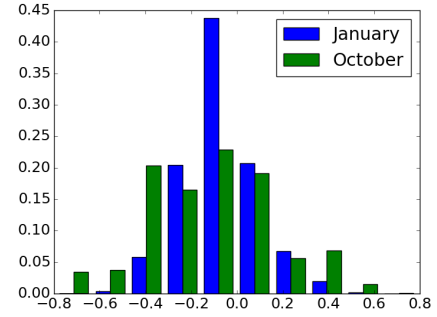

Figure 5: Edge length distributions for January vs. October

To enable us to visualize where on the spectrum edges lie, we make the following definitions. For a point $y \in [0, 1]$, let $f_{\rightarrow}(y, G)$ be the number of edges (x_1, x_2) such that $s_T(x_1) < y < s_T(x_2)$. Similarly, let $f_{\leftarrow}(y, G)$ be the number of edges such that $s_T(x_1) > y > s_T(x_2)$. In other words, $f_{\rightarrow}(y, G)$ is the number of edges crossing y from left to right on the political spectrum, and $f_{\leftarrow}(y, G)$ is the number of edges crossing y from right to left.

In order to interpret these functions, we need a baseline to compare against. A natural baseline in such scenarios is the *randomly rewired* graph $\hat{G}$, the idea of which goes back to [26]. In $\hat{G}$, every vertex has the same indegree and outdegree as in G, but each edge has a randomly chosen endpoint. Note that this can create self-loops, which G does not have by construction; however, the number of self-loops in $\hat{G}$ is small (often 0 for a given randomization) and therefore their effect on this analysis is negligible. Figure 6 shows $f_{\rightarrow}$ and $f_{\leftarrow}$ for G and $\hat{G}$, where the values shown for $\hat{G}$ are in expectation. In January, both $f_{\rightarrow}$ and $f_{\leftarrow}$ are dominated by their rewired counterparts, while in October, the opposite is true. This suggests that at the beginning of the year, domains were used to

(a) January 2016

(b) October 2016
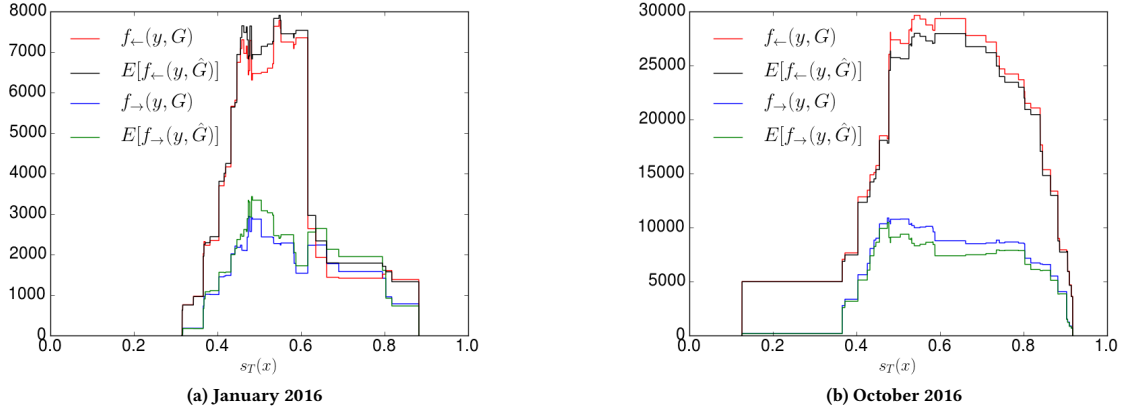
Figure 6: $f_\rightarrow(y)$ and $f_\leftarrow(y)$ for G and $\hat{G}$

reply to politically similar domains, resulting in shorter edges than a random baseline, while closer to the election, domains were used to reply across the political spectrum. This comparison allows us to get a sense for how actual behavior deviates from random behavior, and how this deviation changes over time.

Asymmetry in Out-links. Another striking feature of Figure 6 is that $f_\leftarrow$ dominates $f_\rightarrow$, showing that many more links crossed right-to-left than left-to-right. Intuitively, it seems that a disproportionate number of edges originate on the right and end on the left, corresponding to right-leaning domains being used to reply to left-leaning domains. We can make this precise by defining $r(x)$ as the ratio of indegree$(x)/($indegree$(x) + $outdegree$(x))$ and analyzing how $r(x)$ changes with $s_T(x)$. Figure 7 shows that $r(x)$ is negatively correlated with $s_T(x)$, meaning that domains on the right end of the political spectrum produce a disproportionate number of out-links compared to domains on the left end. In other words, domains on the right are more often used to reply to other domains, while domains on the left are more often the recipients of replies.
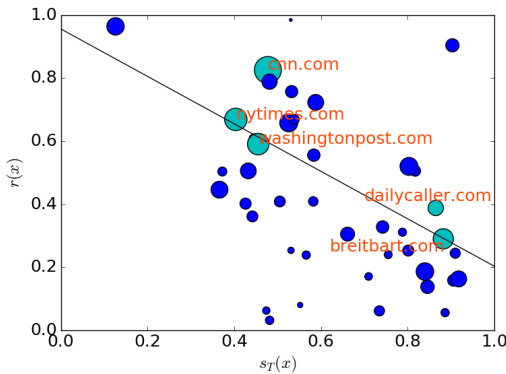
Figure 7: $s_T(x)$ vs. $r(x)$ for October

4 COMPARING TO THE USER LEVEL

Having established a set of results for the structure of invocation graphs on Twitter, we would like to verify that our findings are qualitatively consistent with what we see in more traditional user-to-user communication graphs on social media. Since our Twitter dataset doesn't contain information about individual users, we instead turn to a publicly available Reddit dataset[1]. Reddit is a community discussion website organized into posts, or "submissions," and comments on those submissions. Comments are threaded, so that a comment is either in reply to a top-level post or to another comment. The data consists of every post and comment from Reddit in 2016, along with its author's username. Reddit is subdivided into forums for particular topics called *subreddits*. We focus on three subreddits in particular: r/politics, r/hillaryclinton, and r/The_Donald, which are devoted to politics, Hillary Clinton, and Donald Trump respectively. All three subreddits were among the most active subreddits during 2016. Note that in addition to studying user-to-user dynamics on Reddit, in principle we could also use Reddit data to replicate our Twitter invocation-graph analysis at the domain level; however, it turns out that Reddit contains too few comments with URLs for robust domain-level trends to emerge.

Interactions at the User Level. In order to test for analogous trends at the user level to those we found at the domain level, we need to modify our methods. In particular, we now need some kind of political information about users. Whereas in Section 2 we anchored the political spectrum to the official Clinton and Trump Twitter accounts, here we anchor our notion of political affiliations to r/hillaryclinton and r/The_Donald. Since most users are active in at most one of the two subreddits, we have a simpler notion of a political score: we define a set U_C of users who posted in r/hillaryclinton but not r/The_Donald, and we define a set U_T of users who posted in r/The_Donald but not r/hillaryclinton. There are 22,164 users in U_C and 281,334 users in U_T (more than

---

[1]https://files.pushshift.io/reddit/

	Twitter	Reddit
1	donaldjtrump.com	thegatewaypundit.com
2	thegatewaypundit.com	zerohedge.com
3	breitbart.com	breitbart.com
4	dailycaller.com	donaldjtrump.com
5	zerohedge.com	dailycaller.com
6	foxnews.com	dailymail.co.uk
7	nypost.com	foxnews.com
8	dailymail.co.uk	nypost.com
9	thehill.com	bbc.co.uk
10	politico.com	theguardian.com
11	cbsnews.com	cbsnews.com
12	nbcnews.com	cnn.com
13	cnn.com	thehill.com
14	washingtonpost.com	nbcnews.com
15	bbc.co.uk	huffingtonpost.com
16	theguardian.com	washingtonpost.com
17	nytimes.com	nytimes.com
18	huffingtonpost.com	politico.com
19	politifact.com	newsweek.com
20	newsweek.com	politifact.com
21	hillaryclinton.com	hillaryclinton.com

Table 1: Comparison of Twitter and Reddit political spectra

ten times the size of U_C). We assume that most users in U_C are pro-Clinton, while most users in U_T are pro-Trump (this is consistent with the explicit ground rules for participating in these subreddits).

Validating Political Information from Subreddits. We verify that r/hillaryclinton and r/The_Donald contain strong signal about political orientation by adapting our methodology from Twitter to build a spectrum over the domains on Reddit, and then comparing this spectrum to the one built from Twitter. To do this, we define $P_R(x|C)$ to be the empirical probability that a post or comment in r/hillaryclinton contains a URL from domain x (and analogously for $P_R(x|T)$ and r/The_Donald). As in (1), we can define a political score from Reddit as

$$s_R(x) = \frac{P_R(x|T)}{P_R(x|C) + P_R(x|T)}. \quad (3)$$

Table 1 shows the orderings of the Twitter and Reddit spectra for 21 domains. The Spearman rank correlation [28] (a measure of squared distance) between the two orderings is 0.871 (compared to a maximum of 0.757 over 10,000 randomly shuffled orderings). Thus the two settings align well, demonstrating that our notion of political affiliation is adaptable to Reddit, and that it contains strong and consistent signal.

Building a User-to-User Graph on Reddit. With this notion of political affiliation, we can now investigate some basic properties of political discourse on Reddit. Just as we first restricted our attention to political domains on Twitter, here we restrict our attention to the main political subreddit r/politics. Since only a subset of the users in r/politics are in either U_C or U_T, we focus on comment-reply pairs in which both users involved are in one of U_C or U_T. There are 4 possible types of interaction: $U_C \rightarrow U_C$, $U_C \rightarrow U_T$, $U_T \rightarrow U_C$, and $U_T \rightarrow U_T$ where $p \rightarrow q$ means that a user from p posts a reply to a comment from a user in q. Let $n_{p \rightarrow q}$ be the number of $p \rightarrow q$ interactions in a given time period. We

organize the data into sliding windows of 30 days. In the following plots, the value at a particular date represents the 30-day window ending at that date. Figure 8a shows comment counts throughout 2016 averaged over the next 30 days, and Figure 8b shows what fraction of the comments were each type of interaction.

Figure 8b shows that a steadily rising number of comments are cross-cutting (e.g., are of types $U_C \rightarrow U_T$ and $U_T \rightarrow U_C$) from the beginning of 2016 up until the election in early November, followed by a return to a baseline rate. Figure 8c further reinforces this point, showing a strong negative slope in the ratio of edges between users of the same political leaning and users of different political leanings from January until the November 8 election. These results are consistent with our findings our Twitter, where an increasing fraction of political interactions reach across the political spectrum on Twitter up until the election.

To interpret the significance of these results, we adapt our random rewiring technique to the user level by randomly reassigning users to comments, preserving the invariant that each user still posts the same number of comments as in the original data. Whether we randomize globally (a comment is randomly attributed to any user from all of 2016) or each month (a comment is attributed to a randomly chosen user from the same month that comment was written), the results are the same – the observed slope is significantly more negative than the minimum over 100 random trials.

Using our Reddit dataset, we've shown that our invocation graph methodology can be adapted to analyze traditional user-user communication graphs. In particular, we've used co-occurrences to determine political information about both users and domains. The month-to-month trend found at the user level mirrors our findings on Twitter — in the months leading up to the election, online political interaction increasingly reached across the political spectrum.

5 TWO-DIMENSIONAL ALIGNMENT

Our formulation of the political spectrum in this work indicates that it has a natural *two-dimensional* structure, with one dimension corresponding to co-occurrence probabilities with content related to one candidate, and the other dimension corresponding to co-occurrence probabilities with content related to the other candidate.

In Section 4, when we established that our invocation graph methodology extends naturally to the traditional user-level setting, part of our analysis involved measuring how well the one-dimensional spectra $s_T(\cdot)$ and $s_R(\cdot)$ shown in Table 1 align. Here we consider how to measure the alignment of the corresponding two-dimensional spectra.

To make this comparison, we need to account for the fact that axes may have different scales (e.g. posts on Reddit contain URLs at a different rate than on Twitter). This means we need to scale $P_R(x|C)$ and $P_R(x|T)$ in order to find the "best match" between the two spectra. We formalize this as the following optimization problem, minimizing the squared ℓ_2 distance of each pair of points:

$$\min_{a,b:a,b \geq 0} \sum_{x \in D} (P_T(x|C) - aP_R(x|C))^2 + (P_T(x|T) - bP_R(x|T))^2 \quad (4)$$

where D is the set of domains. Since all points lie in the first quadrant, we can drop the constraint $a, b \geq 0$. Note that (4) can be
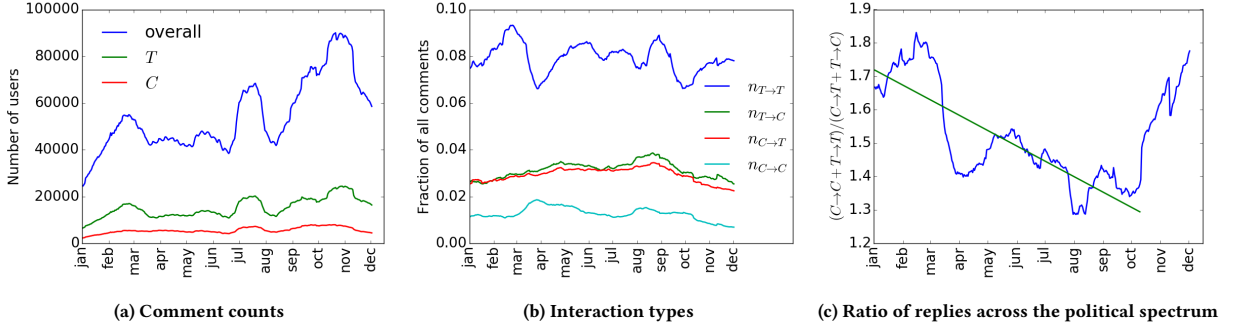
(a) Comment counts   (b) Interaction types   (c) Ratio of replies across the political spectrum

Figure 8: User-level trends from the Reddit dataset. In Figure 8c, a decrease in $(n_{C\to C} + n_{T\to T})/(n_{C\to T} + n_{T\to C})$ shows more interaction reaching across the spectrum.

separated into 2 identical optimization problems of the form

$$\min_c \sum_{x\in D} (u_x - cv_x)^2. \qquad (5)$$

If u and v are the vectors $[u_x : x \in D]^\top$ and $[v_x : x \in D]^\top$ respectively, then this can be written as $\min_c \sum_{x\in D} \|u - cv\|_2^2$. This is convex and has derivative

$$\frac{d}{dc}\|u - cv\|_2^2 = \frac{d}{dc}(u - cv)^\top(u - cv) = 2v^\top u - 2cv^\top v.$$

Setting this equal to 0, we find that (5) has the solution $c = \frac{v^\top u}{v^\top v}$.

Using this, we can scale the Reddit spectrum to compare it to the Twitter spectrum, producing the plot shown in Figure 9. The spectra roughly align, with the domains in approximately the same positions for both Twitter and Reddit.
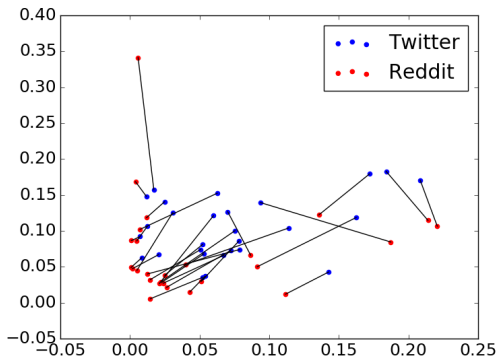

Figure 9: Scaled spectrum comparison

If instead we wish to minimize ℓ_1 distance between pairs of points, we get the optimization problem

$$\min_{a,b:a,b\geq 0} \sum_{x\in D} |P_T(x|C) - aP_R(x|C)| + |P_T(x|T) - bP_R(x|T)| \quad (6)$$

Again, this results in two separate optimizations of the form

$$\min_c \sum_{x\in D} |u_x - cv_x|. \qquad (7)$$

The subgradient of each term is

$$\begin{cases} \{v_x \cdot \text{sign}(cv_x - u_x)\} & u_x \neq cv_x \\ [-v_x, v_x] & \text{otherwise} \end{cases}$$

We sort the x's by increasing u_x/v_x, so $i < j \implies u_{x_i}/v_{x_i} \leq u_{x_j}/v_{x_j}$, and then choose i such that

$$\sum_{j\leq i} v_{x_j} > \sum_{j>i} v_{x_j} \text{ and } \sum_{j<i} v_{x_j} < \sum_{j\geq i} v_{x_j}.$$

Essentially, i balances the v_{x_j}'s, splitting them into two sets such that adding v_{x_i} to either set gives that set the larger sum of the two. Thus, the subgradient of (7) contains 0 for $c = \frac{u_{x_i}}{v_{x_i}}$, making this the solution to (6).

We can also use this same alignment technique to compare spectra between months. Figures 10a and 10b show that in absolute terms, the spectra produced by January and February on Twitter look quite different. In fact, after scaling the axes, the spectra align very well (Figure 10c, using ℓ_1-minimization). One way to see this is to compare the quality of alignment in our data the quality of alignment among shuffled versions of the spectra for January and February, where the domains labeling the points are randomly permuted; we find that the alignment for the real data has much lower cost than the typical alignment for shuffled data. The disparity between the unscaled spectra seems to indicate an influx of users in February who retweeted Clinton but didn't necessarily engage in other sorts of political activity; this left the relative position of other domains in the two-dimensional spectrum unaffected.

We also note this type of scaling doesn't affect the one-dimensional orderings in Table 1. We can rewrite the political score as $s_T(x) = \frac{1}{1+\tan\theta_T}$, where θ_T is the angle of x from the $P_T(x|C)$ axis. Since scaling the axes preserves the ordering of these angles, it also preserves the one-dimensional rankings.

6 FURTHER RELATED WORK

Our work builds on rich literatures in online social media, online news, information diffusion, and the landscape of online political interaction, and it also draws on a long history of work studying the role of media in politics.

(a) January 2016
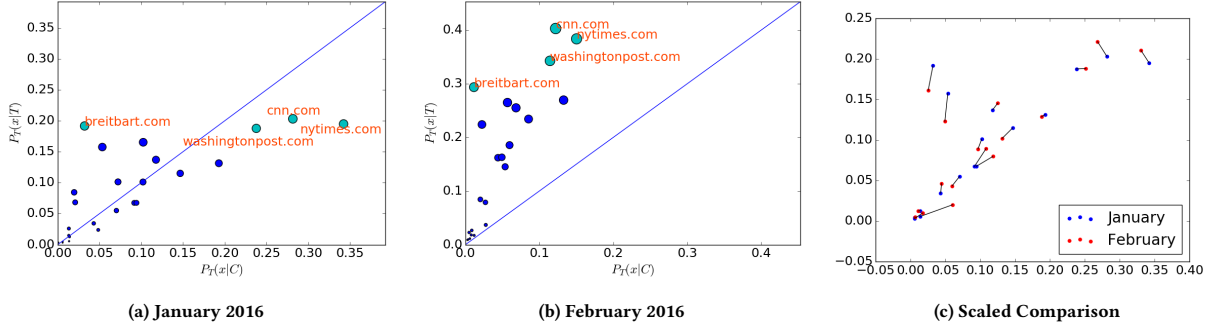
(b) February 2016

(c) Scaled Comparison

Figure 10: Spectrum Comparison for January and February 2016 on Twitter

With access to information now abundant, a growing line of work has investigated how this impacts consumption of political information. As mentioned in the introduction, a key issue has been measuring the extent to which online political interaction crosses ideological divides, or whether it stays relatively confined in "filter bubbles" or "echo chambers" [3, 11, 13, 15, 27, 31]. Concerns about the Web potentially catalyzing balkanization, or fragmentation into isolated ideological divisions, date from the 1990s [32]. Early empirical work on the structure of online political interaction through blogging identified strong ideological partitions evident in large-scale analysis of linking patterns [1], whereas analyses of more recent social media platforms finds evidence for a more complex structure, in which both ideological entrenchment and exposure to more diverse content is promoted on the Web [3, 11].

More broadly, our paper relates to the extensive work on the structure of information sharing on the Web [4], as well as theoretical work on how network structure affects information flow [18, 20]. This structure has been quantified in many ways, particularly by organizing shares into information cascades [8, 17], where nodes are typically people and edges between people indicate if one person directly shared a piece of content with another person. In contrast, our work here introduces and studies invocation graphs, where nodes are information sources and edges indicate that a user shared content from one source in response to content from another source.

The role of the media in politics is the subject of an active field of study [19, 29]. Particularly relevant to our work here is the study of how news and public opinion spread through social networks, including the early theory of two-step flow [21]. More recently, there has been a spirited debate about the impact that "influencers" have on these processes [2, 7, 16, 22, 33].

7 CONCLUSION

In this work we have introduced *invocation graphs*, together with a set of techniques for analyzing them, as a means of probing the structure of online political interaction. In combination with previous methods for inducing a political spectrum from data, we develop methods for measuring several important phenomena. In particular, we analyzed a natural embedding of the invocation graph in the political spectrum, and asked how its edges are distributed across this spectrum—whether they are sequestered in ideological pockets or whether they span larger ideological distances. Applying

these techniques to political interaction on Twitter in the months leading up to the 2016 US Presidential election, we observed that political interaction via the invocation graph became increasingly cross-cutting as the election neared. We also developed methods to analyze whether there are inherent asymmetries between how the right and the left engage each other via replies. Applying our techniques to Twitter, we found that edges in in the invocation graph more consistently went from sources on the right to sources on the left than in the other direction.

It is worth emphasizing a critical feature of invocation graphs, which is that they are composed of *invoked interactions* as opposed to direct interactions. Although on a surface level invocation graphs may resemble the hyperlink graphs that form a basic staple of Web analysis, they are actually quite different, as the links are generally not under the nodes' control. News sources publish content, and then it is up to the readership to determine how these sources connect to each other in the invocation graph. In this sense, relative positions and functions of news domains in invocation graphs are indicative of how the public actually uses them in online political discussion. Beyond the ideological territory news sources may try to explicitly claim for themselves, invocation graphs position these sources according to their roles in political interaction.

There are a number of important directions that remain to be pursued. First, it is intriguing to see how interaction both in the invocation graph on Twitter and the user-to-user interaction graph on Reddit became more and more ideologically cross-cutting as the election approached, given that homophily would suggest that most links should link to nearby points on the spectrum. Is there a systematic way to relate this trend to an underlying level of polarization, so that changes in the structure of the embedded invocation graph might provide insight into polarization and how it evolves? Second, beyond the two large social media datasets we considered, it would be illuminating to apply our methods in other settings as well. In particular, does online debate increasingly cross ideological divides in the run-up to milestone events in general, or was this specific to the 2016 US Presidential election? In general, we believe that applying and extending our methods for online political interaction using invocation graphs contains many promising directions for future work.

REFERENCES

[1] Lada A. Adamic and Natalie Glance. 2005. The Political Blogosphere and the 2004 U.S. Election: Divided They Blog. In *Proceedings of the 3rd International Workshop on Link Discovery (LinkKDD '05)*. ACM, New York, NY, USA, 36–43. https://doi.org/10.1145/1134271.1134277

[2] Eytan Bakshy, Jake M Hofman, Winter A Mason, and Duncan J Watts. 2011. Everyone's an influencer: quantifying influence on twitter. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 65–74.

[3] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.

[4] Eytan Bakshy, Itamar Rosenn, Cameron A. Marlow, and Lada A. Adamic. 2012. The Role of Social Networks in Information Diffusion. In *Proc. World Wide Web Conference*.

[5] Yochai Benkler, Robert Faris, Hal Roberts, and Ethan Zuckerman. 2017. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. *Columbia Journalism Review* 1, 4.1 (2017), 7.

[6] W. Lance Bennett. 1996. *News: The politics of illusion*. Longman.

[7] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and P Krishna Gummadi. 2010. Measuring user influence in twitter: The million follower fallacy. *ICWSM* 10, 10-17 (2010), 30.

[8] Justin Cheng, Lada A. Adamic, P. Alex Dow, Jon M. Kleinberg, and Jure Leskovec. 2014. Can cascades be predicted?. In *23rd International World Wide Web Conference, WWW '14, Seoul, Republic of Korea, April 7-11, 2014*. 925–936. https://doi.org/10.1145/2566486.2567997

[9] Michael Conover, Jacob Ratkiewicz, Matthew R. Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. Political Polarization on Twitter. In *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2847

[10] P. Alex Dow, Lada A. Adamic, and Adrien Friggeri. 2013. The Anatomy of Large Facebook Cascades. In *Proceedings of the Seventh International Conference on Weblogs and Social Media, ICWSM 2013, Cambridge, Massachusetts, USA, July 8-11, 2013*. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6123

[11] Seth Flaxman, Sharad Goel, and Justin Rao. 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly* 80 (2016).

[12] Adrien Friggeri, Lada A Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor Cascades. In *ICWSM*.

[13] R Kelly Garrett. 2009. Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication* 14, 2 (2009), 265–285.

[14] Matthew Gentzkow and Jesse M Shapiro. 2010. What drives media slant? Evidence from US daily newspapers. *Econometrica* 78, 1 (2010), 35–71.

[15] Eric Gilbert, Tony Bergstrom, and Karrie Karahalios. 2009. Blogs are echo chambers: Blogs are echo chambers. In *HICSS'09. 42nd Hawaii International Conference on System Sciences*. IEEE, 1–10.

[16] Malcolm Gladwell. 2006. *The tipping point: How little things can make a big difference*. Little, Brown.

[17] Sharad Goel, Ashton Anderson, Jake M. Hofman, and Duncan J. Watts. 2016. The Structural Virality of Online Diffusion. *Management Science* 62, 1 (2016), 180–196. https://doi.org/10.1287/mnsc.2015.2158

[18] Benjamin Golub and Matthew O Jackson. 2009. How homophily affects learning and diffusion in networks. (2009).

[19] Doris A Graber and Johanna Dunaway. 2017. *Mass media and American politics*. Cq Press.

[20] Matthew O Jackson and Leeat Yariv. 2006. Diffusion on social networks. *Economie publique/Public economics* 16 (2006).

[21] Elihu Katz. 1957. The two-step flow of communication: An up-to-date report on an hypothesis. *Public opinion quarterly* 21, 1 (1957), 61–78.

[22] Elihu Katz and Paul Felix Lazarsfeld. 1966. *Personal Influence, The part played by people in the flow of mass communications*. Transaction Publishers.

[23] Bill Kovach and Tom Rosenstiel. 1999. *Warp Speed: America in the Age of Mixed Media*. Century Foundation Press.

[24] Ravi Kumar, Mohammad Mahdian, and Mary McGlohon. 2010. Dynamics of Conversations. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 553–562.

[25] Paul F. Lazarsfeld, Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. Duell, Sloan, and Pearce.

[26] Michael Molloy and Bruce Reed. 1995. A critical point for random graphs with a given degree sequence. *Random structures & algorithms* 6, 2-3 (1995), 161–180.

[27] Eli Pariser. 2011. *The Filter Bubble: What the Internet is Hiding from You*. Viking.

[28] Charles Spearman. 1904. The proof and measurement of association between two things. *The American journal of psychology* 15, 1 (1904), 72–101.

[29] John Street. 2010. *Mass media, politics and democracy*. Palgrave Macmillan.

[30] Cass Sunstein. 2007. *Republic.com*. Princeton University Press.

[31] Cass R Sunstein. 2001. *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton University Press Princeton, NJ.

[32] Marshall Van Alstyne and Erik Brynjolfsson. 1996. Could the Internet balkanize science? *Science* 274, 5292 (1996), 1479.

[33] Duncan J Watts and Peter Sheridan Dodds. 2007. Influentials, networks, and public opinion formation. *Journal of consumer research* 34, 4 (2007), 441–458.