Towards Balanced Information Propagation in Social Media

Mahmoudreza Babaei
Max Planck Institute for Human Development

Baharan Mirzasoleiman
Computer Science Department, UCLA

Jungseock Joo
Communication and Statistics Department, UCLA

Adrian Weller
University of Cambridge, The Alan Turing Institute

ABSTRACT

As people increasingly rely on social media platforms such as Twitter to consume information, there are significant concerns about the diversity of news consumption. Users may narrow their attention to posts which reinforce their pre-existing views, which could lead to a more fragmented society. Aiming to combat this, earlier work divided news on a given story into *high consensus* and *low consensus* posts, based on how similar reactions can be expected from users with different political views: high consensus news elicits similar reactions, whereas low consensus news elicits different reactions from readers depending on their political leanings. In this work, we propose and quantify the benefits of a strategy to *spread* high consensus news across readers with diverse political leanings. We first compile a dataset and make the following three key observations: (1) low consensus news is more likely to remain within subgroups of users with similar political leanings, whereas high consensus news spreads more across subgroups; (2) high consensus news posted by neutral publishers spreads more equally across subgroups; and (3) users that get the information from other users instead of the publishers, get an even more biased exposure to news. Then, we propose a strategy that spreads high consensus news through neutral publishers, and quantify the significant decrease in the disparity of users' news exposure. Our extensive experiments on Twitter shows that seeding high consensus information with neutral publishers is an effective way to achieve high spread with little disparity regarding political leaning.

1 INTRODUCTION

People are increasingly relying on social media platforms, such as Facebook and Twitter, to receive news and information [29, 45]. Many news stories are divisive, often posted by polarized publishers, eliciting different reactions from users with different political leanings or pre-existing views, e.g., conservatives or liberals. While various news sources publish high and low consensus news that cover a given story, users often limit themselves to the low consensus divisive stories which can reinforce their prior views [4, 15, 25]. This selective exposure and consumption of divisive information may lead to a more politically fragmented, less cohesive society [12] and the formation of filter bubbles or echo chambers [7, 10, 18, 34].

Due to concerns about societal polarization [40, 43], prior works have proposed exposing users to diverse news stories by nudging users to read other views [32, 35]. While this could be useful for encouraging debate in society, such approaches have been shown to *increase* the chance that users reject stories from other perspectives – perhaps because they believe other publishers and their stories are biased – thereby defeating the purpose [5, 6, 31, 33, 44, 50].

On the other hand, highlighting high consensus news that elicits similar responses from both sides could act as a soothing balm to help bring people together, despite initial ideological differences. Babaei et al. [4] proposed such a complementary approach to increase diversity in users' information consumption by identifying high consensus, yet interesting information. Their system recommends high consensus "purple" posts to both red (conservatives) and blue (liberals) users, hoping to increase users' exposure to cross-cutting news posts, leading to lower societal polarization and lower *segregation* in information consumption [13]. Nevertheless, it still remains unclear how such information is spread across users in a network and how individuals choose to react to it.

In this paper, we investigate users' willingness to share and spread such posts, or the reach of high and low consensus news stories across a diverse audience. We also examine the newsworthiness [19, 48] of both high and low consensus news. Overall, we ask two fundamental questions on a Twitter network: (1) Can high consensus posts help to break filter bubbles (and thus potentially decrease polarization in society)? (2) Can we propose methods to propagate high consensus stories broadly across society? We highlight the following contributions:

I. We compile a novel dataset, which reveals how Twitter users with similar or different political leanings are connected to each other. To do so, we consider a dataset of 400 news tweets posted by 10 publishers containing 80 high and 80 low consensus posts [4]. For every high or low consensus news post, we collected a subset of its 100 random retweeters and for each retweeter we collected a random set of their 100 followers. We compute the political leaning of the 1,616,000 users who either retweeted a high or low consensus news story or were exposed to it. Moreover, to simulate the spread of news in Twitter, we crawl a network of more than 100 million Twitter users. This allows us to compute the political leanings of 69,687 users connected by 2,907,026 links.

II. Using our dataset, we study how individuals with different political leanings get exposed to and retweet high and low consensus news posted by users from various political perspectives. We observe that low consensus news tends to proliferate primarily only amongst users with a particular political learning. In contrast, high consensus news has a higher chance of spreading through the entire network. Importantly, high consensus news posted by a set of neutral publishers spreads more equally across liberal and conservative users than if posted by the same number of a mix of non-neutral publishers.

III. Based on the above observations, we propose a strategy that seeds neutral publishers to expose roughly equal fractions of people with different political leaning to high consensus news with the minimum cost (hoping this may help to break filter bubbles which can trap users). We show that our proposed strategy is more effective than seeding the most influential nodes without taking the political leanings into account.

Our work provides new insights and a complementary tool which may help to reduce filter bubbles, encourage healthier interaction between population subgroups, and lead to a more cohesive society.

2 RELATED WORK

We review related work on diverse and polarized news dissemination, and information propagation in social networks.

2.1 News Consumption Polarization on Social Media

Several recent studies have investigated the dissemination of news in social networks [8], focusing on biases [14], political news [2], and the characteristics of spreaders [26].

Traditionally, professional news organizations played a major role in spreading news by selectively presenting news stories to citizens [41]. Accordingly, news media had a high impact on political issues and public opinions [15, 22]. Several works have focused on understanding how and to what extent news media outlets can impact people and society, such as the White Helmets in Syria [42] and the 2016 US presidential campaign [37, 39].

By examining cross-ideological exposure through content and network analysis, Himelboim et al. [25] showed that political talk on Twitter is highly partisan and users are unlikely to be exposed to cross-ideological content through their friendship network. Other studies also report similar findings such as users' higher willingness to communicate with other like-minded social media users [30].

To understand the political bias in social media better, many researchers have studied political polarization on Twitter by analyzing different groups' behavior. Conover et al. [16] showed that Twitter users usually retweet the users who have the same political ideology as themselves, making the retweeting network structure highly partitioned into left- and right-leaning groups with limited connections between them.

Previous work have mostly investigated news media political bias, and the bias introduced in the content of the news, by different methods such as crowdsourcing and machine learning. [4, 11, 20]. Recently, a complementary approach was proposed by Babaei et al. [4], in which the goal is to inject diversity in users' information consumption by identifying high consensus yet informative news, based on using features such as the publishers' political leaning.

Babaei et al. [4] showed that high and low consensus posts are equally popular and cover broadly similar topics. However, their study did not investigate how low and high consensus news spread through social media, and their potential impacts on readers biased exposure, which are the main concerns of our paper.

In this work, we first investigate how low and high consensus news spread through Twitter. Then we study the effect of spreading high consensus news through users with different political leanings on decreasing the disparity in users' exposure. We show that seeding high consensus information with neutral individuals is the best way to achieve high spread with little disparity regarding political leaning.

2.2 Information Propagation in Social Networks

The process of increasing information propagation and network diffusion by identifying and choosing the optimal set of individuals that utilize social influences to maximize adoption or reception of information in society has been studied widely [21, 23, 27, 38]. The effectiveness of these strategies is studied by Kempe et al. [27] under different social contagion models such as Linear Threshold (LT) and Independent Cascade (IC) models.

The goal of our paper is to propagate news with less disparity amongst users with different political leanings. Two studies on fair influence maximization [1, 46] are the most related to our work. However, their approaches may not be directly applicable to online networks such as Twitter. We show that in social media, the political leaning of the seeds can make a considerable bias in users' exposure to news. In particular, we observe that high consensus news posted by neutral publishers has the lowest disparity for spreading among all users (liberal, conservative and neutral). We use the fair influence maximization method proposed by [1] as a baseline.

In the following sections, we explain our dataset and research design, and discuss our findings.

3 DATASET

In this work, we consider the dataset of 400 news tweets posted by 10 publishers collected in [4] between 9th to 15th May, 2017. The dataset contains 80 low and 80 high consensus news posts.

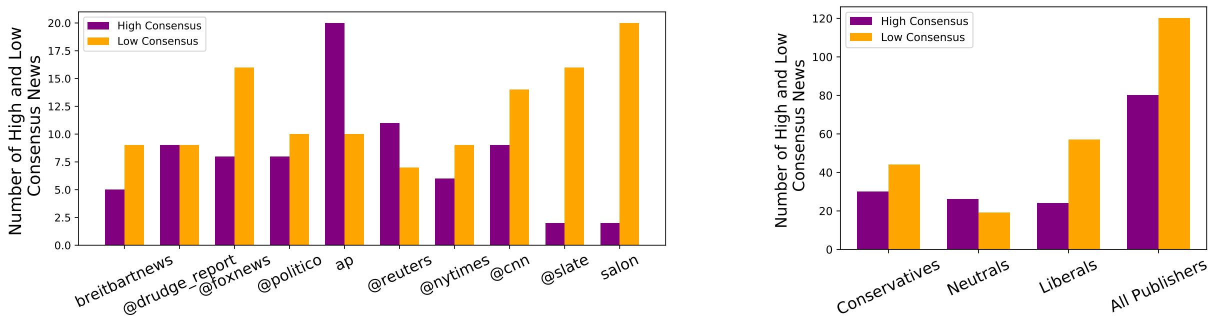
To obtain the political leanings of the users who either retweeted a high or low consensus news or were exposed to it, for every high or low consensus news post in the dataset, we collected a random set of its 100 retweeters. Then for each retweeter, we collected a random set of his 100 followers. Finally, for each of these 1,616,000 users we collected their followees to compute their political leaning.

3.1 High and Low Consensus News Posts

Our news dataset consists of 10 news publishers with different political leanings varying from liberals to neutrals to conservatives: Slate, Salon, New York Times, CNN, AP, Reuters, Politico, Fox News, Drudge Report, and Breitbart News. From each publisher, 40 tweeted news posts are collected during the one week period of 9th to 15th May, 2017. For each news post, the authors in [4] set up a survey in Amazon Mechanical Turk. They asked US AMT workers about their reaction to the post by selecting one out of three options: agreement, neutral, or disagreement. At the end of the experiment, they asked about AMT workers' political leanings: liberal, conservative, or neutral. Based on the above intuition, Babaei et al. [4] proposed to capture the degree of consensus that a social media post is likely to have based on the distribution of the political leanings of the retweeters and repliers of a post, as follows.

$$\text{consensus} = 1 - \left| \frac{\#D_{disagree}}{\#D} - \frac{\#R_{disagree}}{\#R} \right|, \tag{1}$$

where $\#D_{disagree}$ and $\#R_{disagree}$ respectively denote the number of democrats and republicans who disagree with the post, while $\#D$ and $\#R$ are the total number of democrats and republicans.



(A) (B)

Figure 1: Number of high and low consensus news posted on Twitter during 9th-15th May, 2017. (A) shows the number of high and low consensus news for 10 selected publishers, and (B) shows the aggregated result for conservative, liberals, and neutral publishers.

Low Consensus Conservative	Fox News: Schieffer Slams Trump: Comey Firing Reminds Me ofJFK-Oswald ConspiraciesSource, 55 Retweets, 510 Replies, 149 Likes.
	Fox News: @POTUS: "All of the Democrats, I mean, they hated Jim Comey. They didn't like him, they wanted him fired".https://t.co/1ebOtqfIOc 491 Retweets, 293 Replies, 2k Likes.
	Politico: Analysis: Is this a constitutional crisis? Legal experts size up the Comey firing. http://politi.co/2qPEN1c, 210 Retweets, 63 Replies, 254 Likes.
Low Consensus Liberal	The New York Times: What all the Russia investigations have done and what couldhappen nextSource, 131 Retweets, 52 Replies, 226 Likes.
	Salon: Report: Trump "revealed more information to the Russian ambassadorthan we have shared with our own allies" 51 Retweets, 14 Replies, 34 Likes
	CNN: Is Donald Trump the "little boy President"? A @CNNOpinioncontributor takes a closer look at his latest moveshttp://cnn.it/2pE1Uaq, 179 Retwees, 264 Replies, 468 Likes.
High Consensus	Fox News: @johnrobertsFox on firing of James Comey: "This came as a shockto literally everyone, including the @FBI Director." TheFiveSource, 156 Retweets, 259 Replies, 574 Likes.
	AP: BREAKING: Senate intelligence committee invites fired FBI DirectorComey to appear in closed session next Tuesday, 2.7K Retweets, 176 Replies, 5k Likes
	Politico: James Comey told lawmakers he wanted more resourcesfor Russia probe http://politi.co/2r2HxpfSource, 132 Retweets, 40 Replies, 204 Likes.

Table 1: Samples of high and low consensus news posts. First and second rows include low consensus news with conservative and liberal leaning respectively.

Note that high consensus news is different from low attentive news. All stories used in our experiment discuss salient social and political news as opposed to lightweight gossips which may not generate much attention from or disagreement between users due to the chosen topics (see Section 7 for more discussions about the newsworthiness of our news stories). Table 1 shows random sample news that are labeled as low consensus with conservative leaning, low consensus with liberal leaning, and high consensus news respectively. For more examples we refer reader to Table 5 in Appendix.

Figure 1(A) shows the number of high consensus (purple bars) and low consensus (orange bars) news posted by 10 publishers with various political leanings. Figure 1(B) shows the total number of high consensus and low consensus news posted by the same 10 publishers, grouped into liberal, conservative, and neutral categories. It also shows the total number of high consensus and low consensus news posted by all the 10 publishers. We can see that the total number of low consensus posts are considerably higher than the total number of high consensus posts.

As we discuss later, low consensus posts have a much smaller chance of being received by users with different political leanings, which leads to a more politically fragmented society. On the other hand, high consensus posts have a better chance of spreading through communities with various political leanings, and can be utilized to break the filter bubbles.

3.2 Collecting Users' Political Leanings

For every news in our set of 80 high and 80 low consensus news posts, we collected a random set of its 100 retweeters. Then for each retweeter, we collected a random set of its 100 followers. Thus we have 1,616,000 twitter users.

We then inferred every user's political leaning, as a score between -1, +1, using the method of [28], in which, we needed to collect their followees. Inferring the political leaning of a given Twitter user u is based on the following steps – (*i*) generating two representative sets of users who are known to have a democratic or republican bias, (*ii*) inferring the topical interests of u by looking at her followees, and (*iii*) examining how closely u's interests match with the interests of the representative sets of democratic and republican users. Formally,

$$leaning(u) = cos\_sim(I_u, I_D) - cos\_sim(I_u, I_R), \tag{2}$$

where I_u is the interest vectors of user u, and I_D, I_R are normalized aggregate interest vector for the democrat seed set (I_D) and the republican seed set (I_R). Similarity between interest vectors are measured by cosine similarity.

For retweeters with certain political leaning, we calculated the expected fraction of their liberal, conservative, and neutral followers, as is shown in Table 2.

	Liberal	**Conservative**	**Neutral**
Liberal	0.76	0.04	0.2
Conservative	0.045	0.85	0.1
Neutral	0.3	0.27	0.43

Table 2: Expected fraction of liberal, conservative, and neutral followers of retweeters with various political leanings. Rows and columns correspond to retweeters and followers.

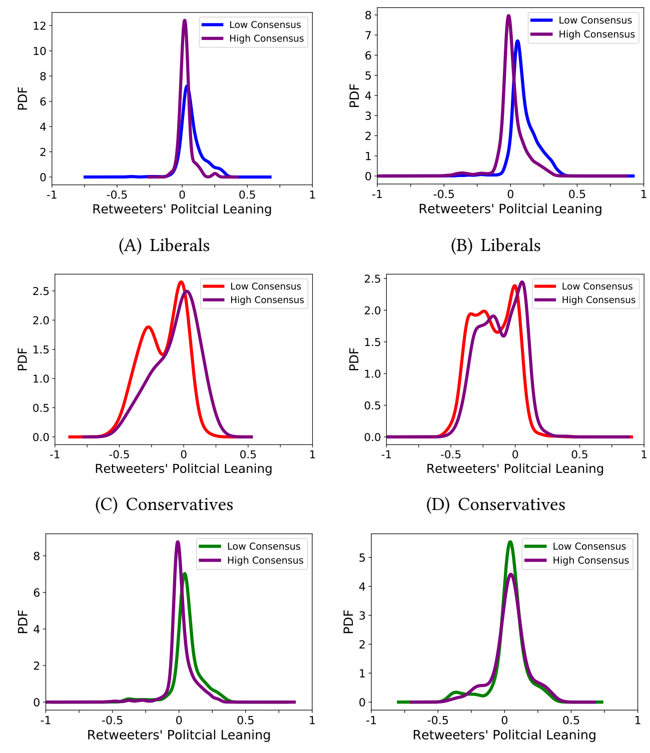
We also estimated the conditional probability that users with different political leanings retweet high consensus and low consensus news post from liberal, conservative, and neutral publishers (given that they retweet) in Table 3. It can be seen that users with a certain political leaning retweet low consensus posts from the publishers with the same political leaning with a very high probability. Interestingly, users retweet high consensus news posts from the publishers with the same political leaning with a smaller probability. On the other hand, there is a very small chance that users with a certain political leaning retweet low consensus posts from the publishers with different political leanings. For high consensus news this probability is larger.

		Retweeters		
		Liberal	**Conservative**	**Neutral**
Publishers	**Liberal**	H: 0.65	H: 0.08	H: 0.27
		L: 0.85	L: 0.04	L: 0.11
	Conservative	H: 0.12	H: 0.68	H: 0.2
		L: 0.08	L: 0.85	L: 0.07
	Neutral	H: 0.34	H: 0.33	H: 0.33
		L: 0.38	L: 0.37	L: 0.25

Table 3: Conditional probability of retweeting a high and low consensus news post (indicated H and L in the table) by users from various political leanings (given that they retweet). Rows and columns correspond to publishers and retweeters. For instance, in the first cell (first row and column), the probability that liberal users retweet high/low consensus news posts published by liberals publishers is 0.65/0.85.

4 THE GAP BETWEEN PROLIFERATION OF HIGH AND LOW CONSENSUS NEWS

In this section, we investigate how high and low consensus news posts spread among users with various political leanings in Twitter. In particular, our goal is to answer the following key questions:



(A) Liberals (B) Liberals

(C) Conservatives (D) Conservatives

(E) Neutrals (F) Neutrals

Figure 2: Distribution of retweeters' political leanings for low and high consensus news posted by publishers with different political perspectives. Distribution of political leanings for a random high and a random low consensus news posted by (A) liberal, (C) conservative, and (E) neutral publishers. Distribution of average political leanings for 100 high and low consensus news posted by (B) liberal, (D) conservative, and (F) neutral publishers. Distribution of political leanings for high consensus news (purple) is more symmetric and centered around 0.

How do individuals with certain political leaning (liberal, conservative, and neutral) get exposed to high and low consensus news posts?

Studying the above key question allows to understand the gap between proliferation of high and low consensus news, and develop strategies to decrease the polarization in the society by breaking the filter bubbles that trap users. We start by investigating users' behavior in retweeting high and low consensus news posts. Then, we discuss how the confirmation bias in retweeting behavior makes the filter bubbles grow larger and promote social polarization.

4.1 Confirmation Bias in Retweeting Behavior

First, we study how users with different political leaning *share* high and low consensus news post. Specifically, we compare how users with different political leanings retweet low and high consensus news posts from publishers with different political perspectives.

Figures 2(A), 2(C), 2(E) show the distribution of the political leanings of all retweeters for one random low consensus and one random high consensus news posted by CNN (liberal publisher), FoxNews (conservative publisher), and Reuters (neutral publisher)[1]. Notice that the distribution of retweeters' political leanings in Figure 2(A), 2(E) has more density in the right (liberal leaning) for the low consensus news. On the other hand, the distribution of retweeters' political leanings in Figure 2(C) has considerably more density in the left (conservative leaning) for the low consensus news. Importantly, the distribution of retweeters' political leanings for high consensus news (purple curve) is more symmetric in all the Figures. Moreover, the mean of the distribution for high consensus news is close to 0.

Next, we consider 80 low consensus and 80 high consensus news posted by the 10 publishers, ranging from liberals to neutrals to conservatives: Slate, Salon, New York Times, CNN, AP, Reuters, Politico, Fox News, Drudge Report, and Breitbart News. For each news post, we consider a set of its 100 retweeters chosen at random. Figures 2(B), 2(D), 2(F) show the distribution of the *expected* political leanings of retweeters of all the low consensus and high consensus news posted by liberal, conservative, and neutral publishers, respectively. Again, the distribution of retweeters' political leanings for high consensus news (purple curve) is more symmetric, and is centered around 0 in all the Figures. In particular, the distribution of retweeters' political leanings for high consensus news posted by neutral publishers has the most symmetric shape around 0.

We summarize our key observations as follows:

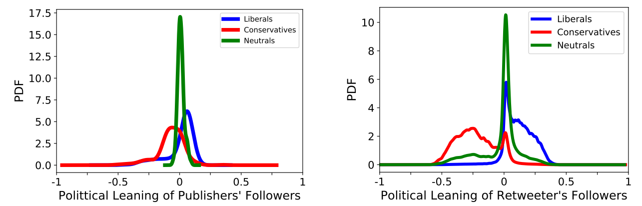
- I. Low consensus news posted by publishers with a specific political leaning (liberal/conservative) are mostly retweeted by users with similar political leanings (Figures 2(B), 2(D)).
- II. High consensus news posted by publishers with a specific political leaning (liberal/conservative) are retweeted by users with various political leanings (liberal/conservative/neutral) (Figures 2(B), 2(D)).
- III. While low and high consensus news posted by neutral publishers spread with lower disparity among users with different political leanings, high consensus news posted by neutral publishers have the highest probability to be spread with minimum disparity among users (Figure 2(F)).

4.2 The Growth of Filter Bubbles in Twitter

Next, we investigate how individuals with different political leanings get *exposed* to high and low consensus news posted by liberal, conservative, and neutral publishers.

Figure 3(A) depicts the distribution of political leanings for followers (level 1) of liberal, conservative, and neutral publishers. We observe that users with conservative or liberal leanings are mostly exposed to news posted by publishers with the same political leaning (the bubble effect). Therefore, the distribution of political leaning for followers of liberal and conservative publishers are skewed to the left and right, respectively. Nevertheless, followers of conservative publishers has a more skewed distribution. This is resulted from the fact that the conservative community is denser, and has fewer connections to liberals and neutrals in Twitter (*c.f.* Table 2). On the other hand, the distribution of political leanings for neutral users is very symmetric and is centered at 0. Hence, neutral users get similar exposure to liberal and conservative view points.

[1]The PDFs have been empirically estimated using kernel density estimation [9]



(A) Publisher's followers (B) Retweeter's followers

Figure 3: Distribution of political leanings for (A) followers of liberal, conservative, and neutral publishers, and (B) followers of retweeters of liberal, conservative, and neutral publishers. As we get farther away from the publishers, the distribution of liberal and conservative followers becomes significantly more skewed (filter bubbles grow larger).

Figure 3(B) shows the distribution of political leanings for followers of retweeters (level 2) of liberal, conservative, and neutral publishers. We observe that while the distribution of political leanings for followers of retweeters of neutral publishers is symmetric and centered around 0, the distribution of political leanings for followers of retweeters of liberal or conservative publishers are extremely skewed. As expected, followers of retweeters of conservative publishers have a more skewed distribution. Interestingly, the skewness of the distributions for followers of retweeters (level 2) is much larger compared to the skewness of distributions for followers of publishers (level 1). This means that filter bubbles in level 2 are larger than those in level 1. Our experiments show that as we get farther away from the publishers, filter bubbles grow even larger (Figure 8).

We summarize our key observations as follows:

- I. Conservatives and liberals get a biased exposure to the news posted in Twitter, while neutrals get similar exposure to liberal and conservative view points (Figure 3(A)).
- II. Users who get the news from retweeters get a significantly more biased exposure, compare to users who get the news from publishers. In other words, as we get farther away from the news publishers, the filter bubbles grow larger (Figure 3(B)).

4.3 Breaking Filter Bubbles

To break filter bubbles, we aim for all individuals to get similar exposure to news stories. Our proposed strategy to break the bubble effect is based on the three key observations discussed earlier: (1) while low consensus news are more likely to proliferate amongst the users with a particular political leaning, high consensus news has a much higher chance of spreading among users with different political leanings; (2) high consensus news posted by neutral publishers has the lowest disparity for spreading among liberal and conservative users; and (3) as users get farther away from publishers, they get a more biased exposure to news. Based on the above observations, we conjecture:

High consensus news posted by neutral users help break the filter bubbles.

High consensus news posted by neutral users achieve high spread with little disparity regarding political leaning. We confirm our conjecture and show the effectiveness of our proposed strategy through an extensive set of experiments later in the paper.

In the following section, we first formulate the problem of information diffusion in social networks. Then, we discuss the problem of finding a near-optimal set of neutral users to seed spreading high consensus news and break the filter bubbles.

5 PROBLEM FORMULATION: INFORMATION DIFFUSION

We start by formulating the information diffusion problem to model the spread of news among individuals with various political leanings in Twitter. We simulate the proliferation of news by assuming that each user can be a publisher. Then we select a set of users and involve them to post news. We represent Twitter by a directed graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of directed edges between the nodes. The nodes in the network are partitioned into three disjoint groups $\mathcal{V} = \{\mathcal{V}_d, \mathcal{V}_r, \mathcal{V}_n\}$, where $\mathcal{V}_d, \mathcal{V}_r, \mathcal{V}_n$ represent users with liberal, conservative, and neutral leanings, respectively. A directed edge (v, u) exists if user v follows user u. When users post tweets, their followers can retweet and spread the tweets in the network. To model spread of information, e.g. news or tweets in Twitter, two well-known classical diffusion models are introduced in the literature [36]: (1) Independent Cascade model (IC) and (2) Linear Threshold (LT) model. In this work, we consider the IC model.

5.1 Independent Cascade model (IC)

In the IC model, information propagates through every edge (v, w) with probability p_{vw}. We have a set of discrete time steps which we denote with $t = \{0, 1, 2, \cdots\}$. At $t = 0$, the initial seed set $S \subseteq \mathcal{V}$ is activated. At every time step $t > 0$, a node $v \in V$ which was activated at time $t - 1$ can activate its inactivated neighbors w with probability p_{vw}. The model assumes that once a node is activated, it stays active throughout the whole process and each node has only one chance to activate its neighbors. The described process stops at time $t > 0$ if no new node gets activated at this time. We note that the IC model is a stochastic process, in which a node u can influence its neighbors w based on the Bernoulli distribution with success probability p_{uw}. A possible outcome of the process can be denoted via a set of timestamps $\{t_v \geq 0 : v \in \mathcal{V}\}$, where t_v represents the time at which a node $v \in V$ is activated.

5.2 Information Diffusion with Low Disparity

Our goal is to find the smallest seed set of users that when post a tweet, it spreads through at least a fraction $Q_p \in [0, 1]$ of liberals ($\mathcal{V}_l$), conservatives ($\mathcal{V}_c$), and neutrals ($\mathcal{V}_n$) in Twitter, where $p \in \{l, c, n\}$. We formulate the problem as follows:

$$\min_{S \subseteq \mathcal{V}} |S| \quad \text{subject to} \tag{3}$$
$$\sum_{p \in \{l,c,n\}} \min\left(f_p(S),\ Q_p \cdot |\mathcal{V}_p|\right) \geq \sum_{p \in \{l,c,n\}} Q_p,$$

where $f_l(.), f_c(.), f_n(.)$ determine the total number users among liberals ($\mathcal{V}_l$), conservatives ($\mathcal{V}_c$), and neutrals ($\mathcal{V}_n$) that are activated as a result of selecting the seed set S. We call $\mathcal{V}_p$ "saturated" by S when $\min\left(f_p(S),\ Q_p \cdot |\mathcal{V}_p|\right) = Q_p \cdot |\mathcal{V}_p|$. When a certain fraction Q_p of individuals with a particular political leaning p are exposed to a news (activated), any new activated individual with political leaning p cannot further improve the utility. This will give individuals with different political leanings a higher chance of being exposed to the news.

We note that the utility function, i.e., $f_p : 2^{\mathcal{V}_p} \to \mathbb{Z}^+$, is a non-negative, monotone, submodular set function [27]. The submodularity is an intuitive notion of diminishing returns, stating that for any sets $A \subseteq A' \subseteq V$ and any node $a \in V \setminus A'$, it holds that:

$$f(A \cup \{a\}) - f(A) \geq f(A' \cup \{a\}) - f(A').$$

Although problem (3) is NP-hard in general [49], for maximizing a submodular function the following greedy algorithm provides a logarithmic approximation guarantee. The greedy algorithm starts from an empty set, add a new node to the set which provides the maximal marginal gain in terms of utility, and stops whenever the desired Q_p fraction of individuals with political leaning p are activated.

5.3 Spreading through Neutrals to Break the Bubbles

In Problem (3), the fraction Q_p can be arbitrary for individuals with different political leanings. However, to break the filter bubbles we wish individuals with different political leanings to get a similar exposure to various news. In other works, we assume similar values for Q_l, Q_c, Q_n. Moreover, the news posted by individuals with neutral leanings have a higher chance of spreading among individuals with liberal and conservative political leanings. Therefore, to break the filter bubbles we aim at finding the smallest subset $S \subseteq \mathcal{V}_n$ that when post a news, at least a fraction Q_p of individuals with political leaning p get exposed to the news. Formally, we have

$$\min_{S \subseteq \mathcal{V}_n} |S| \quad \text{subject to} \tag{4}$$
$$\sum_{p \in \{l,c,n\}} \min\left(f_p(S),\ Q_p \cdot |\mathcal{V}_p|\right) \geq \sum_{p \in \{l,c,n\}} Q_p.$$

6 EXPERIMENTAL RESULTS

In this section we investigate the effect of spreading high consensus news posted by neutral users among individuals with conservative, liberal, and neutral leanings in Twitter. In particular, we show that our proposed strategy is very effective in spreading information among individuals with various political leanings and lowering societal polarization for news consumption. We first describe our instance of Twitter network. We then explain our experimental setup, and present our findings.

Twitter Network. Our network is collected from Twitter in September 2009 [3, 12], and includes: 52 million user profiles, 1.9 billion directed follow links among the users, and 1.7 billion public tweets posted by the users. In order to obtain a static network, we consider the tweets published on July 1, 2009, and filter out users that did not tweet before July 1. After this filtering, we have 70,000 active

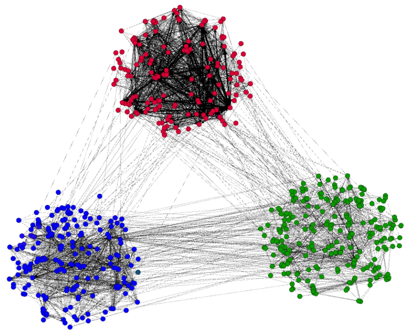


Figure 4: The sample graph from the real Twitter data set collected in 2009. Blue, red, and green nodes indicate users with liberal, conservative, and neutral political leanings.

users. We then extract the strongest connected community, including 69,687 users and 2,907,026 link between them, yielding 19162, 3449, 47076 nodes with liberal, conservative, and neutral leanings, respectively. The average degree of network is 41.5. Figure 4 shows an induced random sample from our final Twitter network.

Sampled Twitter Network. We also created a smaller network by sampling 10% of nodes uniformly at random from our original Twitter network, and connecting the users if they have a connection in the original network. The strongest connected community includes 3,753 users and 6,993 connections with average degree of 1.83. Our sampled Twitter network includes 812 liberals, 186 conservatives, and 2,755 neutrals. Note that the structure of the original Twitter network is very different than the sampled Twitter network. In particular, the sampled Twitter network is significantly sparser than the original Twitter network.

Experimental Setup. For a pair of users $u \in \mathcal{V}_i$ and $w \in \mathcal{V}_j$, we calculate the success probability of activation p_{uw} as the expected fraction of users with political leaning j who retweeted news posted by users with political leaning i. The retweeting probabilities are listed in Table 3.

We apply the greedy algorithm to find a near optimal subset of users that can spread a news over a certain fraction $Q_l = Q_c = Q_n = 0.1$ of liberals ($\mathcal{V}_l$), conservatives ($\mathcal{V}_c$), and neutrals ($\mathcal{V}_n$) in the Twitter network. To evaluate the utility function $f_p(.)$ in Problem 3 and Problem 4, we estimate it by using Monte Carlo sampling [24]. We used 200 samples for this estimation, which yielded a stable estimation of the utility function.

Note that using equal values for Q_l, Q_c, Q_n in Problem (3), we retrieve the fair influence maximization formulation proposed by [1]. In our experiments, we compare our proposed strategy to fair influence maximization.

6.1 Neutrals Can Break Filter Bubbles

In our first set of experiments, we apply the greedy algorithm to Problem (3) and Problem (4) to find the initial set of users to spread news in Twitter. Figure 5 compares the fraction of individuals with liberal, conservative, and neutral political leanings who got exposed to a high consensus news spread through an initial seed set obtained by solving Problem (3) vs. Problem (4). The goal is to expose $Q_p = 10\%$ of individuals with liberal, conservative, and neutral leanings to the news. The top row shows the result on our original Twitter network, and the bottom row shows the result on the smaller sampled Twitter network. Note that the sampled network is much sparser than the original Twitter network.

Figures 5(A), 5(E) show the fraction of exposed individuals when the seeds are selected from the entire network by solving Problem (3). Figures 5(B), 5(F) show the fraction of exposed individuals when the seeds are selected from the users with neutral leanings by solving Problem (4). We note that as more individuals are added to the initial seed set by the greedy algorithm, the disparity in the number of exposed users with different political leanings is much smaller in Figures 5(B), 5(F) compared to Figures 5(A), 5(E). This clearly confirms the effectiveness of our proposed strategy in breaking the filter bubbles.

We note that if we do not take into account the different pattern of diffusion among users of various political leanings, the neutral users may not be the ones that can maximize the spread of information. Figure 6(A), 6(B) compare the fraction of users who were exposed to a high-consensus news, when the seeds are selected from the entire network vs. only the neural users. Here, we assume that all the users spread the news with the same probability of 0.1 irrespective of their political leanings. We see that the news posted by neutral publishers has more disparity and reaches a smaller number of users.

6.2 Neutrals Can Widely Spread the News

Figures 5(C), 5(G) compare the fraction of exposed individuals when the initial seed is selected from the entire network vs. neutrals. There are two interesting observations: the initial seeds selected from neutral users can spread the news even more than users selected from the entire network. Moreover, as we continue the selection process, selected neutral seeds can spread the news as well as the seed set selected from the entire network. This interesting observation confirms the power of neutral users in spreading news in Twitter. As Table 4 depicts, the average number of retweeting of a tweet posted by liberal, conservative, and neutrals are almost equal. There is a interesting observation. High consensus tweets posted by neutrals are retweeted with many democrats and republicans in addition to neutrals. Interestingly, news posted by neutrals is retweeted by an even larger number of users compared to news posted by liberal or conservative publishers.

Figure 7(A) shows the number of users selected with liberal, conservative, and neutral leanings for varying number of seeds selected greedily to solve Problem (3). Figure 7(A) shows the result on the Twitter network, and Figure 7(B) shows the result on the sampled Twitter network. We see that in the set of seeds greedily selected from the entire network, the majority of the users have neutral leanings. This further shows that neutral users are highly effective in spreading information in Twitter. This is consistent with our initial observation, that the news posted by neutrals has a higher probability of spreading among users with different political leanings.

6.3 Neutrals Spread News with Low Disparity

Figures 5(D), 5(H) compare the total disparity of diffusion when the initial seed set is selected from the entire network vs. neutrals. We

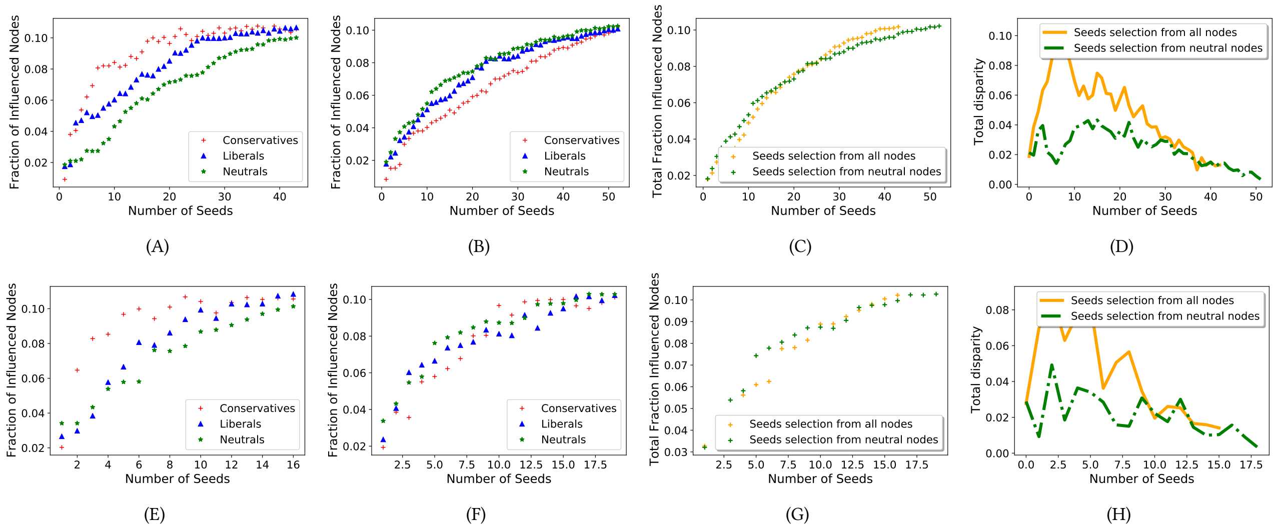
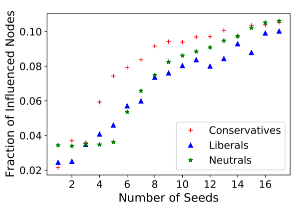
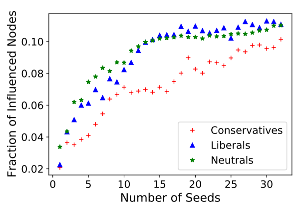


Figure 5: Fraction of individuals with liberal, conservative, and neutral political leanings who are exposed to a high consensus news. Top row shows the result on our Twitter network, and the bottom row shows the result on the smaller sampled Twitter network. (A), (E) show the fraction of exposed individuals when the seeds are selected from the entire network by solving Problem (3). (B), (F) show the fraction of exposed individuals when the seeds are selected from neutral users by solving Problem (4). (C), (G) compare the fraction of exposed individuals when the initial seed is selected from the entire network vs. neutrals. (D), (H) compare the disparity of diffusion when the initial seed set is selected from the entire network vs. neutrals.



(A) Seeds with various leanings



(B) Seeds with neutral leanings

Figure 6: Fraction of users who were exposed to a high-consensus news story, when all users propagate the news with the same probability of 0.1, irrespective of their political leanings in Twitter. (A) shows the result when seeds are selected from the entire network, (B) shows the result when seed are selected only from the neutral users.

define the total disparity as the sum of all disparity (differences) between exposure for each pair of political leanings. Formally, we have: Total disparity=

$$\left|\frac{f_l(S)}{\mathcal{V}_l} - \frac{f_c(S)}{\mathcal{V}_c}\right| + \left|\frac{f_l(S)}{\mathcal{V}_l} - \frac{f_n(S)}{\mathcal{V}_n}\right| + \left|\frac{f_c(S)}{\mathcal{V}_c} - \frac{f_n(S)}{\mathcal{V}_n}\right|.$$

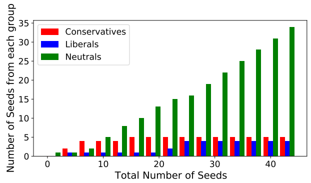
We observe that the total disparity is much smaller when the initial seed set is selected from users with neutral political leanings (Problem (3)) compared to the case when the initial seed set is selected from the entire network (Problem (4)). The difference is larger when the size of the initial seed set is smaller.

		Retweeters			
		Liberal	Conservative	Neutral	Sum
Publishers	Liberal	H: 76 L: 104	H: 9 L: 5	H: 32 L: 15	H:117 L:124
	Conservative	H: 9 L: 10	H: 58 L: 94	H: 18 L: 9	H:87 L:113
	Neutral	H: 45 L: 49	H: 43 L: 47	H: 43 L: 32	H:131 L:128

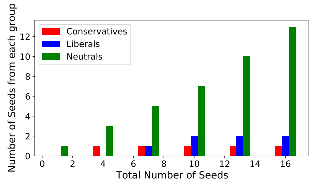
Table 4: Average number of retweeting a high and low consensus news post (indicated H and L in the table) by users from various political leanings. Rows and columns correspond to publishers and retweeters. For instance, in the first cell (first row and column), liberal users on average retweets high/low consensus news posts published by liberals publishers 76 times. We note that news posted by neutrals is retweeted by an even larger number of users, compared to news posted by liberal or conservative publishers.

6.4 Filter Bubbles Grow Larger over Time

Figure 8 shows the fraction of individuals with various political leanings who got exposed to a high consensus news during the diffusion process (IC), for varying number of seeds. More precisely, for a given seed set information diffusion proceeds in discrete time steps $t = \{0, 1, 2, \ldots, \}$. Figure 8 compares the fraction of users with various political leanings who received the news in the first time-step, $t = 1$, and second time-step $t = 2$ in our original Twitter network. Figure 8(A) shows the result when the seeds are selected from the entire network, by solving Problem (3). Figure 8(B) shows the result when the seeds are selected from the users with neutral

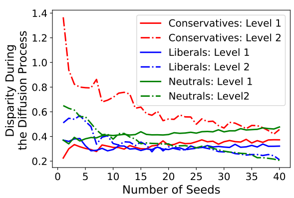


(A) Twitter network

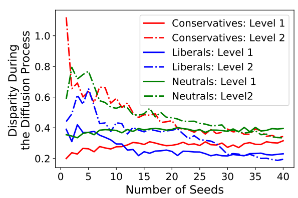


(B) Sampled Twitter network

Figure 7: Number of users selected with liberal, conservative, and neutral leanings for varying number of seeds selected greedily to solve Problem (3). Figure shows the results on (A) the Twitter network, and (B) the smaller sampled Twitter network.



(A)



(B)

Figure 8: Fraction of users who are exposed to the news from different groups in first and second time step of propagation process in the Twitter network. (A) Shows the result when seed are selected from the entire network (Problem 3), (B) shows the result when seed are selected from neutral users (Problem 4).

political leanings, by solving Problem (4). It can be seen that when seeds are selected from the entire network, the disparity becomes larger as the diffusion process continues. On the other hand, the disparity is much smaller when seeds are selected from neutral users.

The above result confirms our observation that the filter bubbles grow larger as the diffusion continues over time. In other words, when the seeds are selected from the entire network, as the we get farther away from the initial set of seeds, the disparity in the number of users with different political leanings who are exposed to the news becomes larger. On the other hand, when diffusion is originated from neutral seeds, users with different political leanings get exposed to the information at the same time. This is crucial while spreading time-critical information, such as health-related information or emergency warnings, in the network.

7 DISCUSSION

Since we propose increasing exposure to high consensus news, we would like to check that such stories carry important information for public discourse. Babaei et al. [4] compared low and high consensus posts on social media by empirically analyzing their properties. They showed that both types of posts are equally popular and cover similar topics. We checked this by analyzing 400 randomly selected posts including examples of high and low consensus news, along with their sources, number of retweets, replies, and likes. See Table 5 in the Appendix for details. We highlight the following observations:

I. For both types of news posts, a variety of news sources exists across the ideological spectrum. Figure 1 also shows several publishers with different political leaning that posts both types of high and low consensus news.
II. On average, high and low consensus tweets are retweeted 158 and 177 times respectively. On average high consensus tweets are liked 532 times, whereas, low consensus news are liked 488 times. Thus, high and low consensus news stories have similar popularity.
III. Galtung and Ruge [19] introduce newsworthiness theory in which they propose several news factors such as frequency, meaningfulness, continuity, etc. Eilders [17] showed that these factors impact news' worthiness. Weber [48] proposes the following hypothesis: "The news factors of a news item influence the level of participation in commenting in an article's comments section". Weber also noted several other factors, such as having a high social impact or being controversial, that may attract more comments as participation [47]. Weber emphasized that if a news story attracts more comments, then it has higher worthiness. Here we can consider the number of replies as participating comments. On average, high consensus and low consensus news stories received 100 and 114 replies (comments), respectively, suggesting that both types of news have similar worthiness.

In summary, we observe that high and low consensus news are similar along multiple dimensions, including variety of news source, popularity, topic covering, and worthiness.

8 CONCLUSION

In this work, we studied the diffusion of news in Twitter. We investigated how users with various political leanings (liberals, conservatives and neutrals) get exposed to low and high consensus news posted by different publishers (e.g. CNN, FoxNews, etc.). We found that (1) while low consensus news stories are more likely to proliferate amongst the users with a particular political leaning, high consensus news has a much higher chance of spreading among users with different political leanings; (2) high consensus news posted by neutral publishers has the lowest disparity for spreading among liberal and conservative users; and (3) as users get farther away from the publishers, they get a more biased exposure to the news. Based on the above observations, we studied the effect of spreading high consensus news through neutral users on decreasing the disparity in users' exposure. Our extensive simulation experiments on Twitter showed that our proposed strategy can be highly effective in decreasing the disparity of information across users with differing views. Our findings may be helpful for breaking filter bubbles and reducing fragmentation in online social media.

REFERENCES

[1] Junaid Ali, Mahmoudreza Babaei, Abhijnan Chakraborty, Baharan Mirzasoleiman, Krishna P. Gummadi, and Adish Singla. 2019. On the Fairness of Time-Critical Influence Maximization in Social Networks. *ArXiv* abs/1905.06618 (2019).
[2] Jisun An, Daniele Quercia, Meeyoung Cha, Krishna Gummadi, and Jon Crowcroft. 2014. Sharing political news: the balancing act of intimacy and socialization in selective exposure. *EPJ Data Science* 3, 1 (2014), 12.

[3] Mahmoudreza Babaei, Przemyslaw Grabowicz, Isabel Valera, Krishna P Gummadi, and Manuel Gomez-Rodriguez. 2016. On the efficiency of the information networks in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. 83–92.

[4] Mahmoudreza Babaei, Juhi Kulshrestha, Abhijnan Chakraborty, Fabrıcio Benevenuto, Krishna P Gummadi, and Adrian Weller. 2018. Purple feed: Identifying high consensus news posts on social media. In *Proceedings of the AAAI/ACM Conference on Artifical Intelligence, Ethics & Society (AIES)*.

[5] Christopher A Bail. 2014. *Terrified: How anti-Muslim fringe organizations became mainstream*. Princeton University Press.

[6] Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115, 37 (2018), 9216–9221.

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

[8] Devipsita Bhattacharya and Sudha Ram. 2012. Sharing news articles using 140 characters: A diffusion analysis on Twitter. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 966–971.

[9] Adrian W Bowman and Adelchi Azzalini. 2004. *Applied smoothing techniques for data analysis*. Clarendon Press.

[10] Engin Bozdag. 2013. Bias in algorithmic filtering and personalization. *Ethics and Information Technology* 15, 3 (01 Sep 2013), 209–227. https://doi.org/10.1007/s10676-013-9321-6

[11] Ceren Budak, Sharad Goel, and Justin M Rao. 2014. Fair and Balanced? Quantifying Media Bias Through Crowdsourced Content Analysis. *http://dx.doi.org/10.2139/ssrn.2526461*.

[12] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and Krishna P. Gummadi. 2010. Measuring user influence in Twitter: the million follower fallacy. In *Proceedings of AAAI Conference on Weblogs and Social Media (ICWSM'2010)* (Washington DC, USA).

[13] Abhijnan Chakraborty, Muhammad Ali, Saptarshi Ghosh, Niloy Ganguly, and Krishna P Gummadi. 2017. On quantifying knowledge segregation in society. *arXiv preprint arXiv:1708.00670* (2017).

[14] Abhijnan Chakraborty, Saptarshi Ghosh, Niloy Ganguly, and Krishna P Gummadi. 2016. Dissemination biases of social media channels: On the topical coverage of socially shared news. In *Tenth International AAAI Conference on Web and Social Media*.

[15] Chun-Fang Chiang and Brian Knight. 2011. Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies* 78, 3 (2011), 795–820.

[16] M. Conover, J. Ratkiewicz, Matthew Francisco, B Gonçalves, F. Menczer, and A. Flammini. 2011. Political Polarization on Twitter. In *Proc. AAAI ICWSM*.

[17] Christiane Eilders. 2006. News factors and news decisions. Theoretical and methodological advances in Germany. *Communications* 31, 1 (2006), 5–24.

[18] Seth Flaxman, Sharad Goel, and Justin M. Rao. 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly* 80, S1 (2016), 298–320. https://doi.org/10.1093/poq/nfw006

[19] Johan Galtung and Mari Holmboe Ruge. 1965. The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of peace research* 2, 1 (1965), 64–90.

[20] Matthew Gentzkow and Jesse M. Shapiro. 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica* 78 (2010), 35–71. Issue 1.

[21] Amit Goyal, Francesco Bonchi, Laks VS Lakshmanan, and Suresh Venkatasubramanian. 2013. On minimizing budget and time in influence propagation over social networks. *Social network analysis and mining* 3, 2 (2013), 179–192.

[22] Tim Groseclose and Jeffrey Milyo. 2005. A measure of media bias. *The Quarterly Journal of Economics* 120 (2005), 1191–1237. Issue 4.

[23] Jason Hartline, Vahab Mirrokni, and Mukund Sundararajan. 2008. Optimal marketing strategies over social networks. In *Proceedings of the 17th international conference on World Wide Web*. 189–198.

[24] W. K. Hastings. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 1 (04 1970), 97–109. https://doi.org/10.1093/biomet/57.1.97 arXiv:https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf

[25] Itai Himelboim, Stephen McCreery, and Marc Smith. 2013. Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication* 18, 2 (2013), 40–60.

[26] Mengdie Hu, Shixia Liu, Furu Wei, Yingcai Wu, John Stasko, and Kwan-Liu Ma. 2012. Breaking news on twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2751–2754.

[27] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *KDD*.

[28] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P. Gummadi, and Karrie Karahalios. 2017. Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. ACM, New York, NY, USA, 417–432. https://doi.org/10.1145/2998181.2998321

[29] Joseph Lichterman. 2010. New Pew data: More Americans are getting news on Facebook and Twitter. *http://tinyurl.com/News-on-Social-Media*.

[30] Zhe Liu and Ingmar Weber. 2014. Is Twitter a public sphere for online conflicts? A cross-ideological and cross-hierarchical look. In *International Conference on Social Informatics*. Springer, 336–347.

[31] Charles G Lord, Lee Ross, and Mark R Lepper. 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology* 37, 11 (1979), 2098.

[32] Sean A. Munson, Stephanie Y. Lee, and Paul Resnick. 2013. Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*. AAAI press, Boston, USA, 419–428.

[33] Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Political Behavior* 32, 2 (2010), 303–330.

[34] Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Group.

[35] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 443–452.

[36] Romualdo Pastor-Satorras and Alessandro Vespignani. 2001. Epidemic spreading in scale-free networks. *Physical review letters* 86, 14 (2001), 3200.

[37] Filipe N Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabricio Benevenuto, Oana Goga, Krishna P Gummadi, and Elissa M Redmiles. 2019. On microtargeting socially divisive ads: A case study of russia-linked ad campaigns on facebook. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 140–149.

[38] Matthew Richardson and Pedro Domingos. 2002. Mining knowledge-sharing sites for viral marketing. In *KDD*.

[39] Marian-Andrei Rizoiu, Timothy Graham, Rui Zhang, Yifei Zhang, Robert Ackland, and Lexing Xie. 2018. # DebateNight: The Role and Influence of Socialbots on Twitter During the 1st 2016 US Presidential Debate. In *Twelfth International AAAI Conference on Web and Social Media*.

[40] David Schkade, Cass R Sunstein, and Reid Hastie. 2007. What happened on deliberation day? *California Law Review* 95, 3 (2007), 915–940.

[41] Pamela J Shoemaker and Timothy Vos. 2009. *Gatekeeping theory*. Routledge.

[42] Kate Starbird, Ahmer Arif, Tom Wilson, Katherine Van Koevering, Katya Yefimova, and Daniel Scarnecchia. 2018. Ecosystem or echo-system? Exploring content sharing across alternative media domains. In *Twelfth International AAAI Conference on Web and Social Media*.

[43] Cass R. Sunstein. 2002. The Law of Group Polarization. *Journal of Political Philosophy* 10, 2 (2002), 175–195. https://doi.org/10.1111/1467-9760.00148

[44] Charles S Taber and Milton Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *American journal of political science* 50, 3 (2006), 755–769.

[45] Jaime Teevan, Daniel Ramage, and Merredith Ringel Morris. 2011. # TwitterSearch: a comparison of microblog search and web search. In *WSDM '11 Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, Hong Kong, China, 35–44.

[46] Alan Tsang, Bryan Wilder, Eric Rice, Milind Tambe, and Yair Zick. 2019. Group-Fairness in Influence Maximization. *arXiv preprint arXiv:1903.00967* (2019).

[47] Patrick Weber. 2012. The virtual get-together. Determinants interpersonal - "o public communication on news websites. *Social Media and Web Science. Frankfurt am Main: DGI* (2012), 457–459.

[48] Patrick Weber. 2014. Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments. *New media & society* 16, 6 (2014), 941–957.

[49] Laurence A Wolsey. 1982. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica* 2, 4 (1982), 385–393.

[50] Thomas Wood and Ethan Porter. 2019. The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* 41, 1 (2019), 135–163.

A NEWSWORTHINESS OF HIGH CONSENSUS NEWS

Here in Table 5 we show some examples of high and low consensus that randomly selected from 400 posts along with their sources, number of retweets, replies, and likes.

High Consensus News	Low Consensus News
BREAKING: Senate intelligence committee invites fired FBI Director Comey to appear in closed session next Tuesday . Source: AP, 2.7K Retweets, 176 Replies, 5k Likes	Report: Trump “revealed more information to the Russian ambassador than we have shared with our own allies” Source: Salon, 51 Retweets, 14 Replies, 34 Likes.
@johnrobertsFox on firing of James Comey: "This came as a shock to literally everyone, including the @FBI Director." #TheFive Source: Fox News, 156 Retweets, 259 Replies, 574 Likes.	Orrin Hatch makes clear the conservative case against Obamacare: Once the public “is on the dole, they’ll take eve… Source: Salon, 113 Retweets, 12 Replies, 10 Likes.
White House calls emergency meetings as global cyberattack spreads http://politi.co/2qgNnW1 Source: Politico, 80 Retweets, 31 Replies, 72 Likes.	Why are Republicans attacking the Census Bureau? Because they don’t want an accurate count of Americans Source: Salon, 691 Retweets, 22 Replies, 680 Likes.
The Latest: US says Russia should be worried about N. Korea missile launch; Japan, US, South Korea discuss threat. http://apne.ws/2r5iNQ1 Source: AP, 167 Retweets, 12 Replies, 93 Likes.	Is Donald Trump the "little boy President"? A @CNNOpinion contributor takes a closer look at his latest moves http://cnn.it/2pE1Uaq Source: CNN, 179 Retwees, 264 Replies, 468 Likes.
White House wants the FBI to complete its investigation into Russia interference in the 2016 election. http://apne.ws/2qso0kS Source: AP, 179 Retweets, 68 Replies, 97 Likes	@SarahHuckabee: "@POTUS over the last several months lost confidence in Director Comey. The DOJ lost confidence in Director Comey." Source: Fox News, 122 Retweets, 94 Replies, 593 Likes.
Donald Trump’s lawyers say he doesn’t have any Russian money “with a few exceptions” http://dlvr.it/P7QvTv Source: Salon, 125 Retweets, 10 Replies, 18 Likes.	Schieffer Slams Trump: Comey Firing Reminds Me of JFK-Oswald Conspiracies Source: Fox News, 55 Retweets, 510 Replies, 149 Likes.
North Korea’s Sunday missile test is what one researcher is calling an "extended middle finger to Trump" Source: CNN, 212 Retweets, 63 Replies, 326 Likes.	Sarah Huckabee Sanders went on Fox last night and, wait for it, said it’s “time to move on” from Russia probe: http://slate.me/2qsCqBO Source: Slate, 28 Retweets, 35 Replies, 48 Likes.
“He knows the last three days have not been good for him”: Sean Spicer’s make-or-break briefing http://politi.co/2r1t8MQ Source: Politico, 59 Retweets, 37 Replies, 112 Likes.	“A fresh start will serve the FBI”: Republicans provide cover for Donald Trump http://ift.tt/2q6C1BM Source: Salon, 107 Retweets, 61 Replies, 88 Likes.
San Diego police: Teen shot and killed left suicide note http://fxn.ws/2ps2UPO #FOXNewsUS Source: Fox News, 52 Retweets, 21 Replies, 84 Likes.	Acting FBI Director contradicts White House claim that fired director James Comey had lost support. http://apne.ws/2r5GJjr Source: AP, 357 Retweets, 56 Replies, 497 Likes.
@HillaryClinton launches Onward Together PAC. Read more: http://fxn.ws/2qlcwz0 Source: Fox News, 144 Retweets, 574 Replies, 94 Likes.	Democrats are now openly talking about impeaching Donald Trump Source: Salon, 105 Retweets, 22 Replies, 153 Likes.
Sources: James Comey told lawmakers he wanted more resources for Russia probe http://politi.co/2r2Hxpf Source: Politico, 132 Retweets, 40 Replies, 204 Likes.	It appears that Trump may have just falsely accused himself of wiretapping himself: http://slate.me/2r9agb8 Source: Slate, 198 Retweets, 25 Replies, 303 Likes.
Trump says ’his decision’ to fire FBI chief, calls him ’showboat’: NBC interview http://reut.rs/2r6gUjg Source: Reuters, 69 Retweets, 62 Replies, 83 Likes.	What all the Russia investigations have done and what could happen next Source: The New York Times, 131 Retweets, 52 Replies, 226 Likes.
Condoleezza Rice: ’When you’re not credible about Syria, you’re not credible about North Korea’ http://fxn.ws/2pudYMd Source: Fox News, 288 Retweets, 80 Replies, 806 Likes.	Republicans are allowing states to drug test people applying for unemployment benefits Source: Salon, 19 Retweets, 10 Replies, 16 Likes.
Senior US official says Trump administration has approved weapons for Kurds. http://apne.ws/2qYRLac Source: AP, 180 Retweets, 45 Replies, 97 Likes.	VP Mike Pence defends firing of FBI Director James Comey, says Trump ’made the right decision at the right time.’ http://apne.ws/2q3fl7Q Source: AP, 65 Retweets, 83 Replies, 76 Likes.
When men and women finish school and start working, they’re paid pretty much equally. But then it all changes. Source: The New York Times, 854 Retweets, 120 Replies, 1.1K likes.	Pres. Trump’s firing of FBI Director James Comey is a "grotesque abuse of power," legal analyst Jeffrey Toobin says http://cnn.it/2q1FQd4 Source: CNN, 803 Retweets, 169 Replies, 1.3 Likes.

Table 5: Examples of high and low consensus news including the source, number of retweets, replies and likes.