An et al. *EPJ Data Science* 2014, **2014**:12
http://www.epjdatascience.com/content/2014/1/12

EPJ.org

EPJ Data Science
a SpringerOpen Journal

RESEARCH Open Access

Sharing political news: the balancing act of intimacy and socialization in selective exposure

Jisun An[1*], Daniele Quercia[2], Meeyoung Cha[3*], Krishna Gummadi[4] and Jon Crowcroft[5]

[*]Correspondence: jan@qf.org.qa; meeyoungcha@kaist.edu
[1]Qatar Computing Research Institute, Majlis Al Taawon Street, Doha, Qatar
[3]Graduate School of Culture Technology, KAIST, 291 Daehak-ro, Daejeon, Republic of Korea
Full list of author information is available at the end of the article

Abstract

One might think that, compared to traditional media, social media sites allow people to choose more freely what to read and what to share, especially for politically oriented news. However, reading and sharing habits originate from deeply ingrained behaviors that might be hard to change. To test the extent to which this is true, we propose a *Political News Sharing (PoNS)* model that holistically captures four key aspects of social psychology: *gratification*, *selective exposure*, *socialization*, and *trust & intimacy*. Using real instances of political news sharing in Twitter, we study the predictive power of these features. As one might expect, news sharing heavily depends on what one likes and agrees with (*selective exposure*). Interestingly, it also depends on the credibility of a news source, i.e., whether the source is a social media friend or a news outlet (*trust & intimacy*) as well as on the informativeness or the enjoyment of the news article (*gratification*). Finally, a Twitter user tends to share articles matching his own political leaning but, at times, the user also shares politically opposing articles, if those match the leaning of his followers (*socialization*). Based on our PoNS model, we build a prototype of a news sharing application that promotes serendipitous political readings along our four dimensions.

Keywords: news sharing; political news; political diversity; social media; Twitter

1 Introduction

Media bias has been widely studied in cultivation theory. This holds that popular media such as newspapers, television, and now the Internet have the power to influence our view of the world and set our day-to-day norms. Media bias - appearing as either selecting what to report or choosing a slant on a particular report [1, 2] - matters because it affects the political beliefs of the audience, alters voting behavior [3, 4], and has negative societal effects like increasing intolerance of dissent and creating segregated and polarized communities [5].

Since social media sites have been recently used to share news stories at a global scale [6–8], they promise to connect millions of individuals who hold very diverse political views [9] and diversify their media consumption [10]. Unfortunately, in social media, people's news consumption patterns have not changed much compared to those in traditional media - people tend to avoid information that conflicts with their views, resulting in the old-fashioned problem of media bias, even reinforcing what is known as the filter bubble [11].

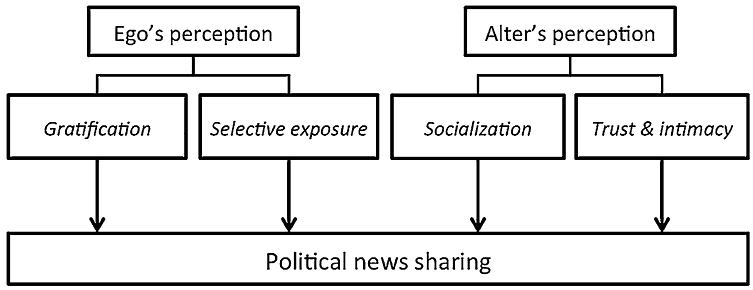

Figure 1 Overview of political news sharing (PoNS) model.

The choice of what to read and share is a process determined by a number of psychological factors such as cognition and motivation. Investigating them thoroughly will lead us to understand the media bias problem better and develop a tool mitigating this effect. One set of theories is related to the 'ego's perception' and includes two main factors: *gratification*, suggesting that people read and share news to satisfy their desires such as informativeness and entertainment [12] and *selective exposure*, suggesting that people like to read information in agreement with their views and avoid conflicting information [13]. Another set of theories is associated with 'alter's perception' and focuses on social aspects of information sharing: whether the person who passes the information is credible (trust & intimacy) and whether the person who will receive the news would like it (socialization). These two sets of theories have not been considered together, and we will do so here.

To this end, we propose a Political News Sharing (PoNS) model based on the two popular perspectives as theoretical foundation. The PoNS model is graphically summarized in Figure 1. There are four major factors that might impact news sharing: gratification, selective exposure, socialization, and trust & intimacy. We have evaluated the PoNS model with more than 150,000 cases of sharing political news in Twitter. Using data of twenty-four popular news media outlets and twenty-one million Twitter users, we study the predictive power of the four factors separately and collectively. More specifically, we make the following contributions:

1. We investigate the extent to which Twitter users are exposed to political diversity. We find that 90% of the users receive information from news media of only one political leaning - that is, most people do not subscribe to politically diverse media outlets. On the other hand, their friends' retweets lead them to diversify their news consumption, in that, 41% of the users are exposed to politically diverse news.
2. We test which factors motivate people to share news. The most important factor is the source's credibility: a user is 49% more likely to retweet news coming from media sources (original tweet) than news from other users (retweeted one). However, when sharing *political* news, people prefer those from friends (Trust & Intimacy). The second strongest factor is exposure: with an extra exposure to a news article, a user is 23% more likely to retweet the news (Gratification).
3. Political news is not generally considered to be a retweeting subject, but when people share, they mostly retweet articles they agree with (agreeable news), confirming the key role of selective exposure theory. We also find a weak evidence that when the articles is interesting to their followers, people share political news reflecting views different from their own (socialization).

4. These findings provide a holistic view of how people share political news. The first finding suggests that the formation of echo-chambers resulting from subscriptions to traditional media outlets is countered by the more serendipitous news sharing happening among users. Also, the fact that followers hold a certain influence over a user is not surprising, if one considers that people are influenced by peers who are up to three (social network) hops away from them [14]. Based on these four generic factors that motivate political news sharing, we demonstrate a new way of visualizing news articles that gives users a fine control over the PoNS' four dimensions.

2 Background

Researchers in media communication have long been studying the media effect, and we review some of those studies below.

2.1 Media bias and its consequences

Media bias has been shown to have negative societal consequences (e.g., intolerance of dissent, political segregation, group polarization) [5]. Republicans and Democrats read different newspapers and books [15] and geographically sort themselves by choosing in which neighborhoods to live [16]. Media slant changes people's beliefs, for example, in whom to vote [3, 17]. Group polarization is prevalent not only in the offline world (e.g., in the form of geographic sorting) but also in the online world. Blogs reflecting different political views rarely link to each other [18], and online news consumption is also biased, much like offline news [19].

A few recent studies examined how people exchange political content in online social networks. [20] has looked at Twitter use of U.S. political parties. [21] has shown a retweeting network of political hashtag that shows a clear segregation of two political parties; however they have found active interactions across those two parties in a mention network. Related to this work, [22] has reported that political discussions taking place in Twitter can go to extreme easily. We build upon this work and expand it by determining to which extent Twitter users segregate themselves into echo chambers, and what could be done about it.

2.2 News sharing in social media

Due to its popularity and the data's easy accessibility - Twitter data is publicly available - research on Twitter has been flourishing for the last few years. Kwak *et al.* [23] studied the topology of the Twitter graph, finding a non-power-law follower distribution, a short effective diameter, and low reciprocity. Other studies have provided insights into the patterns of user participation in Twitter by looking into the use of Twitter as a medium of information spreading, including sharing URLs and reporting news [24], posting local news [25], and promoting political views [26]. Despite a large body of research on information sharing being conducted, news articles published by media sources have less been examined.

A news article published in social media can reach many more audience members than media outlets' direct followers, passed through the social network. Consider the example of a news article exchanged by Twitter users in Figure 2. It shows how the same news could propagate with different sentiments (i.e., positive and negative comments). The article is originally tweeted by the Washington Times and is then received by followers of the media source, who might have different political views. Among the followers, two of them

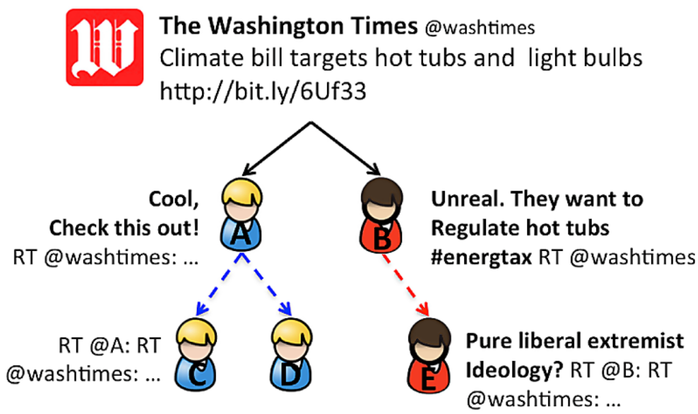

Figure 2 Pictorial flow of news in Twitter.

might decide to retweet the article to their followers. In the toy example, user A is in tune with the political view of the Washington Times and he adds a comment ('cool. check this out'). Then, user A's followers, user C and user D, get to receive the Washington Times' original tweet along with user A's positive recommendation. However, not all news articles are positively recommended. In the case of user B, he may decide to retweet the news, but with a negative comment ('Unreal. they want to regulate hot tubs #energtax' and his follower user E might further decide to add another negative comment 'pure liberal extremist ideology?') before he retweets.

2.3 Motivations of news sharing

A number of theories in media and communication research have been suggested to understand why people consume news. Gratification theory states that satisfying audiences' social and psychological needs is the key to attracting and keeping those audiences [12]. Specifically, desires such as entertainment, interpersonal communication, information learning, escapism, and surveillance are the general factors that are associated with news consumption on the Internet [27–30]. The few studies that have focused on content sharing activities in online communities found that gratification, social interaction, reciprocity, and self-identifications are strongly related to why people share knowledge online [31]. On the other hand, as an attempt to understand how people manage opinion conflicts, selective exposure theory hypothesizes that individuals tend to favor information that reinforces pre-existing views while avoiding contradictory information [13].

In the context of social media study, number of exposures has been widely considered as a proxy of social impact. Social impact theory states one's belief, motive, behavior changes as a result of presence or actions of other individuals' [32]. The first principle is that it is a multiplicative function of the strength, immediacy, and number of sources present in the environment. A number of studies have examined the impact of exposure in relation to whether it motivates people to share a piece of information. Previous work has found a strong evidence that the number of exposures is strongly related to a hashtag adoption [33] and a rapid growth of hashtags [34].

However, in this work, we use the number of exposures as a proxy to measure informativeness of a tweet. Social impact theory also states that the relative impact of each additional person decreases and when an individual is a part of a group, the impact of sources is divided among individuals exposed and is therefore reduced - as more people exposed, less likely to change their behavior. The theory also relates to a theory of 'diffusion of responsibility' in a socio-psychology phenomenon whereby a person is less likely to take responsibility for action or inaction when others are present [35, 36], hinting that one may feel less obligated to share a piece of information as more exposures happen. Thus, we do not use the number of exposures as directly connected to the motivation of sharing. We rather consider it as an additional context of the tweet.

3 Political news sharing model

When ideas spread, there are always three parties: the person(s) creating the news, the person passing on the news, and the person(s) receiving it. Considering how these three parties influence motivations of a person in sharing news, we identify four major factors: gratification, selective exposure, socialization, and trust & intimacy. Followed by relevant prior literature, we discuss corresponding Twitter specific measures of each factor.

3.1 Gratification

We assume that how a user perceives a given article contributes to whether or not he will share it. More precisely, informativeness describes the extent to which news shared can provide users with relevant and valuable information. A user would perceive a piece of information as valuable when he is exposed to it multiple times, resulting in motivating him to share it further. Related to this, entertainment is another key factor for understanding news consumption [29]. People are more likely to share news that they deeply care about and are interested in. To study why users retweet news articles, we derive two predictors from gratification theory:

1. *F1 numexposures* denotes how many times the retweeter repeatedly gets exposed to the article.
2. *F2 topic-interesting-me* reflects the extent to which the retweeter is interested in the article's topic. To compute it, we create, for each user, his *interest-vector* by considering each article the user posts, classifying the article's categories, and aggregating the classifications of all the user's articles into a unique *interest-vector*. The classification consists of 12 categories and is performed by the Alchemy Application Programming Interface (http://www.alchemyapi.com/), which is a popular text-mining web service that classifies news articles in a number of topic.

3.2 Selective exposure

Selective exposure theory states that individuals tend to favor information that reinforces pre-existing views and avoid contradictory information. In this theory, individuals are likely to choose political articles with opinions that fit with and support what they already know. Various studies have examined whether selective exposure exists in news consumption and consistent evidence has been found across a variety of media [37]. Below are the two predictors associated with the selective exposure theory:

3. *F3 political* reflects whether the article is about politics or not (binary).
4. *F4 leaning-matched-me* indicates that the retweeter's political views match those of the media outlet that published the article (binary). This factor is considered only if the news article is about politics.

3.3 Socialization

Socialization plays a critical role in determining whether a user will share a news article. This is particularly true in social networking sites. The experience that a user has in sharing news articles depends on the context created by the user's peers (e.g., only few reactions on what has been shared, being ignored). To encourage discussions or idea exchange and to ultimately enrich the social media experience, users might consider what their online friends like to read or agree with.

We thus expect that one retweets news articles that are relevant to one's followers. Hence the two predictors related to socialization are:

5. *F5 topic-interesting-followers* indicates the extent to which the retweeter's followers are interested in the article's topic. This is computing based on the average similarity between the categories of the article (as per Alchemy categories) and the interest-vectors of one's followers.
6. *F6 leaning-matched-followers* represents a fraction of retweeter's followers whose political views match that of the article. This factor is considered only if the news article is about politics.

3.4 Trust and intimacy

In social media, news articles are shared not just by news outlets but also by users. People easily turn their ears to a piece of information or an opinion coming from their 'friends.' The stronger the relationship, the more easily people accept what friends share. Through peer influence, users might receive news articles reflecting views different from their own. We would like to examine whether trust and intimacy affect a user's decision to share a given article. Trust in the sender might also impact one's willingness to retweet the article. To capture the impact of friendship on retweeting news articles, we consider a number of measures between a user who tweets an article and his follower who receive it. The predictors related to trust and intimacy are:

7. *F7 fromfriend* indicates whether a news article comes from one of the retweeter's friends or from a media source (binary).
8. *F8 mutualfriend* is a measure of whether the user and the propagator(s) are friends with each other (i.e., have a mutual relationship).
9. *F9 difference-in-followers* is the difference between the retweeter's number of followers and the propagator's. A friend having a greater number of followers may be a public figure or influential.
10. *F10 sharedfollowers* is the number of common followers between retweeter and propagator. Having more common followers may mean that the two have common interests.
11. *F11 sharedfollowees* is the number of common followees between retweeter and propagator.
12. *F12 sharedleaning* reflects whether the retweeter's political views match those held by the propagators.

4 Data

4.1 Collecting Twitter data

Twitter was created in 2006 and it has been rapidly growing, attracting 255M monthly active users [38]. In Twitter, the users share content composed from 140-character text messages called *tweets*. Users can choose whom to *follow* - a social relationship in Twitter is not necessarily mutual. Hence, topologically, a Twitter network is a directed graph: an individual has a number of 'followees' whom he follows and 'followers' who follow him. A user will receive all tweets posted by his followees. Unless a user sets his privacy setting as 'private' explicitly, all tweets he posts are visible to the public by default.

For our analysis, we gathered publicly available information from Twitter. We firstly identified a set of news media sources by consulting both the website http://newspapers.com (which listed the top 100 newspapers in the USA) and Twitter's 'Browse Interest' directory (its news directory) [39]. From these two lists, we generated a list of news providers, including mainstream news outlets such as the New York Times and CNN. We also included individual journalists and anchors as they are known to have a large audience and play a prominent role as news providers. We only considered US-based news media outlets, a total of 22.

Using the Twitter API, we obtained all follow links to media sources and their corresponding tweets for an 8-month period (from January to August 2009). To efficiently identify the consumption behavior of news on Twitter, we focus on the set of news media tweets that contain a URL. Through the Twitter API, we collected all tweets that contain any of the URLs posted by the 22 media sources. Not all of these users were directly following media sources. For each user who posted, retweeted, or replied to those URLs, we also gathered his follow links.

The resulting dataset includes 22 media sources with 55,777 tweets with a total of 8,793,507 subscribers. Among all tweets, there were 42,483 tweets containing URLs, 397,640 retweets, 21.4 million Twitter profiles, and 720 million directed follow links. For convenience, Table 1 shows a summary of the data collected from all media sources.

Table 1 Summary of the three media sources under study

Account	Subscribers	Tweets	URLs	RTs of URLs	Mondotimes
abcnews	16,397	3,800	3,729	10,412	left
ariannahuff	23,912	185	94	894	left
davidgregory	1,115,405	575	159	2,361	left
huffingtonpost	54,418	4,186	4,174	29,385	left
jdickerson	953,993	1,469	413	14,501	left
maddow	1,091,269	127	116	12,316	left
nbcnightlynews	12,602	2,118	2,105	3,234	left
nprnews	116,834	1,956	1,848	30,825	left
nprpolitics	1,272,479	2,803	2,342	20,238	left
nprscottsimon	887,009	893	68	763	left
nytimes	1,755,740	5,676	5,527	91,379	left
theearlyshow	6,873	1,524	1,413	4,191	left
todayshow	108,481	1,672	1,050	26,291	left
washingtonpost	27,196	1,903	1,617	9,619	left
andersoncooper	319,257	3,528	3,436	23,495	center
cnnbrk	2,596,796	524	240	32,131	center
jackgraycnn	587,758	3,109	368	26,598	center
richardpbacon	819,312	1,793	224	16,116	center
foxnews	100,272	6,401	6,361	15,699	right
chicagotribune	17,588	1,010	1,010	15,610	right
usnews	4,747	4,239	4,233	7,545	right
washtimes	6,954	2,025	1,956	4,037	right
all	8,793,507	55,777	42,483	397,640	

4.2 Extracting political discourse

To categorize the URLs in our tweets, we use, again, the Alchemy API. We use this API because it has been shown that it entails superior classification performance compared to other popular classifiers [40]. Given a URL, Alchemy extracts the associated text and returns featured words, the main topic, and a confidence value for the categorization which scales from 0 to 1 representing the API's degree of belief that the text pertains to that category. The main topic is chosen from the following 12 topics: Arts Entertainment, Business, Computer Internet, Culture Politics, Gaming, Health, Law Crime, Recreation, Religion, Science Technology, Sports, and Weather. We excluded URLs that are categorized as 'None' (e.g., video live streaming or personal photos) and URLs that have low confidence values (<0.5 on Alchemy's scale of [0,1]).

Out of 42,483 URLs from the 22 media sources, 23,017 URLs were successfully classified. For these categorized news articles, 41% of them have been retweeted at least once, where culture_politics is the mostly popular category, where 73% of articles in culture_politics has been retweeted at least once, followed by entertainment (68%), and science_technology (57%).

Next, to classify news outlets into liberal, conservative, or center, we consulted the website http://www.mondotimes.com and used the Americans for Democratic Action (ADA) scores of media sources [1] that is widely used for comparing media bias across different outlets [2]. The ADA score measures a media outlet's political bias based on the number of times the outlet cites various think-tanks and other politically-oriented groups. The score is on the scale from 0 to 100, where 0 indicates a strong conservative tendency. Four media outlets (Fox News, Chicago Tribune, U.S. News & World Report, and Washington Times) were classified as right-wing, five (including CNN) as center, and fourteen (including Huffington Post, NPR News, and New York Times) as left-wing. As we are interested in how different political opinions reach users having different political views, we chose to focus on left and right media outlets (18 in total), since they have a clear political stance.

4.3 Inferring political leaning of users

We inferred the political leaning of each user based on the set of media outlets that the user subscribed to. To reduce noise in the data, we only considered users who tweeted more than 5 times in the last three months of our data collection period (this leaves us with 2.9M users). Then we filtered out users who follow only one media source under the assumption that they are less interested in news reading through social media. After this pruning, 419,446 users were still left.

To infer the political leaning of individual users, we have used their subscriptions to media outlets, under the conservative assumption that one's political leaning can be determined only if all media outlets the user follows exhibit the same political leaning [41]. Recent study has shown this mapping method is valid [42]. To analyze the restrictions introduced by this assumption, we randomly picked 30 left-leaning users and 30 right-leaning users in our dataset, and asked them their political leanings. We received 22 and 16 responses from left and right-leaning, respectively, retaining a response rate of 63%. Among those 38 people who answered, we found 30 users (78.9%) were classified correctly (16 left-leaning users with 72.7% matching rate and 14 right-leaning users with higher matching rate of 87.5%). With such a high level of accuracy, we choose to use the conservative rule of thumb to assign political leanings of users. These users accounted for 380,568 or 90.7% of our users. Most were left-wing (88%) and only 44,943 users (12%) were right-wing.

Table 2 Summary of the two compared datasets

	Active followers	Classified URLs	Retweets (Retweeters)
Original dataset (14 left & 4 right)	419,446	22,179	154,078 (68,225)
Balanced dataset (4 left & 4 right)	146,480	11,289	60,524 (27,597)

The balanced dataset is a subset of the original dataset, where we balance the number of left and right media sources.

Because our original dataset includes many more left-wing media sources, it may over-represent left-leaning users. Even though they are over-represented compared to the general population, they are still representative of the Twitter population. We therefore generated a 'balanced dataset' that includes the same number of media sources from each political leaning (four of each are selected randomly). A summary of the two datasets is shown in Table 2. The numbers reported in the row 'balanced dataset' correspond to the mean values across 20 different reshuffled versions of the balanced dataset. All our analyses have been carried out on both datasets, and the corresponding results remain consistent between the two, suggesting that the ways in which we select outlets do not impact the results.

5 Status quo of media bias

Retweets from friends can expose individuals to diverse political views. To test the extent to which this is the case, we map retweets back to the original tweets by tracking URLs, which do not change from tweet to retweet. By consolidating all tweets containing the same URL, we build a propagation tree for each news article.

5.1 Top news covered by left and right media

From our eighteen media sources, 14,568 URLs were categorized as political news articles. These URLs spawned 31,473 retweets. 17.5% of users engaged in political news propagation, and users who follow both left and right media sources (following four media outlets at least) were three times more likely to propagate political news than others.

To give a sense of which political news stories are shared in Twitter, we listed the top stories in Table 3. One can see that the left and right media have a different tone of voice even on the same topic. For example, on an issue regarding North Korea, left-wing media reported 'North Korean Leader Pardons, Releases U.S. Journalists' while right-wing media said 'North Korea Threatens to Wipe Out U.S.'

5.2 Exposure to diverse opinions

To investigate whether Twitter users live in echo chambers or not, we examined what users receive and what they decide to promote by retweeting. More specifically, we initially consider two main sources of news articles (i.e., media sources a user follows and his Twitter friends) and compute the political diversity of news articles coming from the two sources based on the Shannon Index, which is defined as the (political) entropy of the news articles associated with the user. It is $-\sum_{i=1}^{S}(p_i \log p_i)$, where S is the total number of possible political preferences, and p_i is a proportion of news articles that reflects the ith political preference. If a user's articles reflect all political views to the same extent, then the Shannon Index (the user's political diversity) is maximum. While the Shannon Index

Table 3 The top 10 mentioned political news articles from the left-leaning and the right-leaning media sources in Twitter

Rank	Left-leaning media	Source	# of RT
1	The President's Opening Remarks on Iran	nytimes	203 (*l*100, *r*50, *b*53)
2	Cheney Is Linked to Concealment of C.I.A. Project	nytimes	113 (*l*97, *r*2, *b*14)
3	Sarah Palin Resigning as Alaska's Governor	nytimes	50 (*l*38, *b*12)
4	N. Korean Leader Pardons, Releases U.S. Journalists	nytimes	48 (*l*38, *b*10)
5	'Military Coup' Underway In Iran	nytimes	47 (*l*39, *r*3, *b*5)
6	Health Care Hecklers & the Rise of Right-Wing Rage	nytimes	43 (*l*41, *b*2)
7	Conservatives Don't Know He's Joking	nytimes	42 (*l*38, *b*4)
8	N.Y. Assembly Passes Gay Marriage Bill	nytimes	39 (*l*39)
9	10 Most Offensive Tea Party Signs From Tax Day Protests	nytimes	39 (*l*34, *b*5)
10	Rick Perry Calls For Fed Help With Swine Flu	nytimes	37 (*l*28, *b*9)

Rank	Right-leaning media	Source	# of RT
1	North Korea Threatens to 'Wipe Out' U.S.	foxnews	40 (*l*2, *r*20, *b*18)
2	Obama Claim of AARP Endorsement 'Inaccurate'	foxnews	30 (*r*16, *b*14)
3	House leaders drop their plans to buy fancy jets	foxnews	30 (*r*20, *b*10)
4	WH Says Girl Chosen at 'Random' to Speak at Town Hall	foxnews	28 (*l*2, *r*18, *bb*8)
5	Pelosi Calls Health Care Critics 'Un-American'	foxnews	28 (*r*14, *b*14)
6	Latino Leaders Call for Illegal Immigrants to Boycott Census	foxnews	24 (*l*2, *r*12, *b*10)
7	Outbursts, Hot Tempers Fill Town Hall Meetings	foxnews	24 (*r*16, *b*8)
8	Obama: Recovery Will Take Years Not Months	foxnews	24 (*r*20, *b*4)
9	AARP Faces Backlash From Seniors Over Health Care Reform	foxnews	22 (*r*16, *b*6)
10	Palin to stump for conservative Democrats	foxnews	20 (*r*6, *b*14)

The table also shows the news source of the article and the number of retweets of the article from different political groups, where *l* stands for liberals, *r* for conservatives, and *b* for others.

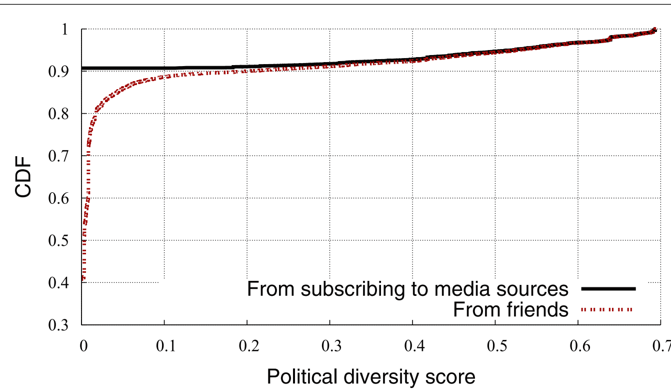

Figure 3 CDF of political diversity score. The plot shows the distribution of political diversity score from direct subscription (solid line) and that from social media friends (dotted line).

is popularly used as a measure of diversity, the resulted values may be biased as it does not take the sample size into consideration. To solve the bias problem, we apply Miller-Madow correction technique [43].

5.2.1 Diversity from subscribing to media sources

For each user, we consider the media sources the user follows, determine their political leanings, and compute their overall political diversity score using the Shannon index. Across all users, we find that the distribution of political diversity is skewed. Figure 3 shows that 90% of users have a political diversity score of 0, meaning that they only subscribe to media sources of their political leaning, and only 10% have diversity score greater than 0, meaning they subscribe to at least one media source whose political view is different from their own. Based on the classification of different 'user types' in previous studies [44], we could classify the users in the 90% group as either *challenge-averse* (i.e., users who seek out affirming opinions but reject the idea that they avoid challenging items) or *support-seeking* (i.e., users who are primarily interested in opinions that are similar to their own), while the remaining 10% would be classified as *diversity-seeking* (i.e., users who are interested in considering opinions that challenge their own).

To then distinguish who is support-seeking among the 90%, we consider which of these users have retweeted news articles containing political views different from their own. People in the support-seeking category do not like political diversity, yet they do not mind receiving a few articles they disagree with. We find that, among the low-diversity users in the 90% group, 86% are challenge-averse and 14% are support-seeking. Compared to the previous work suggesting that there was no evidence of the existence of support-seeking individuals [44], we observed three very distinct groups: users who do not subscribe media outlets nor share articles contrasting their political views, users who occasionally share articles even if they are in conflict with their views, and users who enjoy diverse opinions.

5.2.2 Diversity from friends

Having looked at the political diversity introduced by the media outlets users subscribe to, we now examine the diversity introduced by their Twitter 'friends'. Thus, for each user, we consider the news articles the user receives not only from media sources but also from friends. We then compute the diversity of political views contained in those articles using, again, the Shannon index.

We find that the distribution of political diversity score among users is still skewed as seen in Figure 3. However, there is a crucial difference: now the proportion of users with political diversity score of 0 drops from 90.7% to 40.7%, suggesting that social media friends are a primary source of political diversity in Twitter. At the population level, the geometric average of political diversity shows a 7.13-fold increase (the same goes for the politically balanced dataset in which the increase is even higher - it is 12.24-fold). We also find that the higher the diversity from direct media subscription, lower the changes in the diversity from friends with Pearson's correlation coefficients of $r = -0.29$ ($p < 0.0005$) and with a Spearman's correlation coefficients of $r = -0.14$ ($p < 0.0005$).

To see which set of media outlets gets more exposure through social media friends, we select users who follow: (1) only left-leaning media sources; and (2) only right-leaning media sources. We find that both left-leaning (55.6%) and right-leaning media outlets (56.7%) profit from the social network (Figure 4), reaching more than half of non-subscribers in our dataset. With a balanced set, 45.3% of left-leaning and 47.8% of right-leaning users are exposed to media outlets having views different than their friends'. The row percentages slightly change between the two cases yet they are comparatively the same and lead to the same conclusion: left-leaning media outlets reach non-subscribers slightly less than right-leaning media outlets do.

6 Evaluation of PoNS model

Having observed that users are exposed to politically diverse news articles far more by their friends than directly by the media sources, we now test the PoNS model and examine which of the four factors (gratification, selective exposure, socialization, and trust & intimacy) are better associated with the chance of sharing political news. We use a logistic

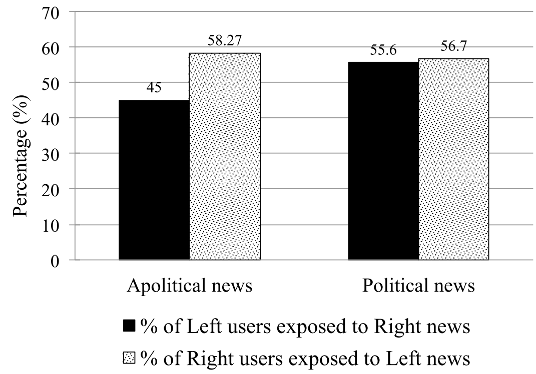

Figure 4 The probability that political news stories reach the other political group through social media friends. The bar plot shows an increase in political diversity after incorporating indirect media exposures through social media friends.

binomial regression, which models the probability that a user retweets a given news article based on twelve predictors extracted across the four factors. All predictors undergo a logarithmic transformation, when necessary (i.e., when they are skewed). The dependent variable is thus:

$$y_i = \begin{cases} 1 & \text{if user } i \text{ retweeted,} \\ 0 & \text{if user } i \text{ did not retweet.} \end{cases}$$

Since our data only includes positive cases - that is, the cases when people share the news articles - we need to augment our dataset with negative cases (by under-sampling them): we do so by adding an equal number of negative cases - that is, with a set of random news article-and-user pairs. By construction, the resulting sample is balanced (the response variable is split 50-50), and the accuracy of a random prediction model would thus be 50%. We model a retweeting probability as a linear combination of the predictive variables, plus terms for interactions. We use the first 7 and a half months of our data to calculate the independent variables and use the last two weeks of data for the test, which had 14,309 retweeting cases. Adding the same number of random negative cases, we use 28,618 cases to build the model.

The results of the logistic regressions are reported in Table 4. The coefficients reported tell us the extent to which the corresponding predictors explain the retweeting behavior. The p-values indicate the extent to which coeffiecients are statistically significant. To show how well the model fits the data, we use Hosmer-Lemeshow test of 'goodness-of-fit' and report χ^2 and its p-value. Please note that with Hosmer-Lemeshow test, the higher the p-value of the model, the better the model fits the data. The Hosmer and Lemeshow's (H-L) goodness of fit test divides subjects into deciles based on predicted probabilities, then computes a chi-square from observed and expected frequencies. Then a probability (p) value is computed from the chi-square distribution to test the fit of the logistic model. If the H-L goodness-of-fit test statistic is greater than 0.05, as we want for well-fitting models, we fail to reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model's estimates fit the data at an acceptable level. That is, well-fitting models show non-significance on the goodness-of-fit test, indicating model prediction that is not significantly different from observed values.

Table 4 Logistic regression results for retweeting news

Predictive variables	Original	Revised
Gratification		
F1 numexposures	0.93***	0.93***
F2 topic-interesting-me	–0.09	-
Selective exposure		
F3 political	–0.80***	–0.89***
F4 leaning-matched-me	0.72***	0.91***
Socialization		
F5 topic-interesting-followers	0.57***	0.49***
F6 leaning-matched-followers	–0.14	-
Trust and intimacy		
F7 fromfriend	–2.09***	–1.47***
F8 mutualfriend	0.13*	-
F9 difference-in-followers	0.05***	-
F10 sharedfollowers	–0.00	-
F11 sharedfollowees	–0.00***	-
F12 sharedleaning	–0.37***	–0.38***
F13 politicsfromfriend	0.31***	0.35***

Both models pass the goodness-of-fit tests: Original ($\chi^2 = 11.1629$, $p = 0.45$) and Revised ($\chi^2 = 13.2977$, $p = 0.39$). Signif. codes: 0 (***) 0.001 (**) 0.01 (*) 0.05 (.).

Logistic regression coefficients cannot directly be interpreted on the scale of the data as models are nonlinear on the probability scale. To ease the interpretation of the logistic regression coefficients β, one could apply the 'divide by 4' rule which can be applied if the probabilities (i.e., values of the outcome variable) are close to 0.5, that is the case for our data [45]. To see how, take a predictor x (e.g., whether or not the article is about politics), its regression coefficient β_x, and the outcome variable y_i. From the idea that the slope of the logistic curve is maximized at the center point, one can take the logistic regression coefficient β_x and divide it by 4 to get an upper bound on how much a unit difference in x (e.g., whether article is about politics or not) would change the outcome variable (e.g., probability of retweeting the article). If β_x is, for example, 0.8, then articles about politics are likely to be retweeted with a probability 20% ($\frac{0.8}{4} = 0.2$) more than articles of any other subject.

6.1 General news sharing

We first investigate the generic news sharing pattern. We consider retweeting cases not only of political news but also of other kinds of news for comparison. Table 4 reports the results of the logistic regression: the 'original model' column fits the original dataset, while the 'revised model' column includes only the significant predictors whose sign remain unchanged compared to those of the original model. Both models fit the data better than the null model and the prediction error rate of our model is only 0.19, while that of the null model is 0.5. Below we discuss the findings.

Gratification: The F1 feature is statistically significant, while F2 is not. The number of repeated exposures to the same article (F1) is positively correlated with retweeting the news, emphasizing the importance of a news article being informative to be retweeted. The positive coefficient of 0.93 indicates that one extra exposure to the article increases one's retweeting probability by 23% (0.93/4 = 0.23). On the other hand, what a user generally likes is not correlated to what he shares (F2). This finding counters what had been found in more traditional settings: one major motivation for consuming and sharing news is entertainment [29].

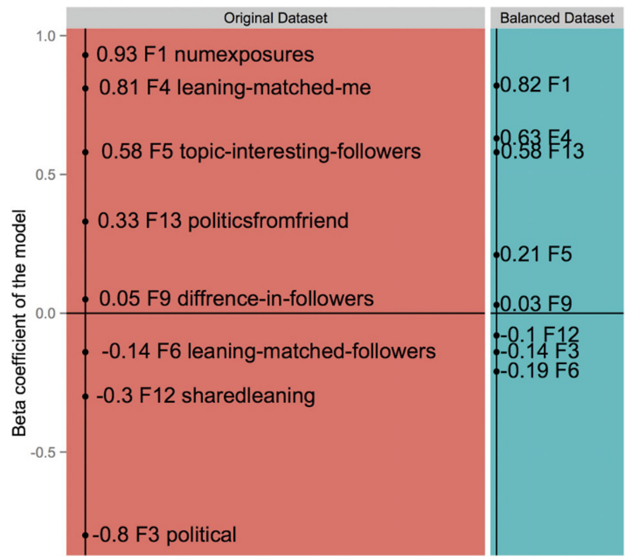

Figure 5 Factors ranked based on their predictive power (i.e., beta coefficient value of the model) to retweeting behavior. Each column reports the results from the original dataset and the balanced dataset. Even the scale of impacts may vary, the list of significant factors and their signs are matching in two models, which gives a confidence in the original dataset.

Selective exposure: Both F3 and F4 are statistically significant variables. People tend to retweet news articles in subject areas other than politics. The negative correlation for F3 indicates that a user is 20% less likely to retweet political articles as opposed to other types of news (–0.8). When articles about politics are concerned, one retweets them more with a high positive correlation, if they express political views one agrees with (F4, 0.72). This suggests that although Twitter allows the flow of politically diverse news articles, people have a strong tendency to retweet only what matches their views.

Socialization: We find that what one's followers are interested in (F5) is positively related to what one chooses to share (0.57). This finding is in line with findings from other work [31] in that social interaction is a key factor that encourages information sharing in the online world. Trying to please one's friends may be particularly important in Twitter.

Trust and intimacy: The results show that all the variables except for F10 are statistically significant, and only few are mildly correlated. The significance of source credibility (F7) shows a negative correlation (–2.09). This indicates that a user is 52% more likely to retweet news articles that come from media sources than from friends. However, news from a friend who has a mutual relationship (F8) have a 3% higher probability of being retweeted (0.13). To a limited extent, one is also likely to preferentially retweet news coming from popular friends (F9). Finally, political news is unlikely to be shared, yet a user is 8% more likely to share a political article given that it was shared by a friend (F13, 0.31). This peer pressure effect was even true for friends who had opposing political views (F12, –0.37).

The regression analysis can determine the relative importance of the 13 predictors (in the following order): trust & intimacy, gratification, selective exposure, and socialization. Significant factors are ranked based on how much they increase the retweeting probability and are summarized in Figure 5. Each column reports the results for the original dataset

Table 5 Predictors for retweeting political news articles

Predictive variables	Agree with article	Disagree with article
F1 numexposures	0.74***	0.83***
F2 topic-interesting-me	0.16	1.23*
F5 topic-interesting-followers	0.22	–0.36
F6 leaning-matched-followers	–0.44*	1.55**
F8 mutualfriend	0.14*	0.52.
F9 difference-in-followers	0.06***	–0.13**
F10 sharedfollowers	0.00	0.00*
F11 sharedfollowees	0.00	–0.01*
F12 sharedleaning	–0.27***	–1.60**

Both models pass the Goodness-of-fit tests: Agree ($\chi^2 = 4.5227$, $p = 0.81$) and Disagree ($\chi^2 = 1.6879$, $p = 0.97$). Signif. codes: 0 (***) 0.001 (**) 0.01 (*) 0.05 (.).

and the politically balanced dataset, respectively. For the two datasets, the impact of each factor varies in scale, but their signs (positive or negative) do not, speaking for the validity of the results.

To sum up, the credibility of a news outlet (trust & intimacy) and the informativeness or the enjoyment of the articles themselves (gratification) are the two strongest factors that motivate people to share news. Socialization plays a role in choosing news topics to a certain extent - what a user shares depends on what his friends like. In sharing political news, we see that people share political news less frequently than other types of news; however, when they do so, the political stances of articles are likely to match those of the users (selective exposure) or of their friends. As one might expect, one's taste is a strong motivation to encourage to share a news article. However the above results also suggest that social relationships do affect media consumption in notable ways.

6.2 Political news sharing

Next, we focus on the specific question of whether users retweet articles differently depending on the article's political views. We consider two situations: one in which a news article matches the retweeter's political views - that incorporates 3,379 positive retweeting cases, and the other in which it does not match (701 negative retweeting cases). We run a logistic regression for these two cases separately, and report the results in Table 5. For the two regressions, the likelihood ratio test were significant at the 5% level. In both cases, the strongest predictor is *numexposures*, which is the number of times the retweeter has been exposed to the article. If the article agrees with the retweeter's political views, then the article does not necessarily agree with the followers' political views (–0.44) and is likely to come from reciprocal friends (0.14), who might happen to have diverse political views (–0.27).

In contrast, if the article disagrees with the retweeter's political views, then the article is likely to be of retweeter's interest (1.23) but not necessarily of followers' interest (–0.36), match followers' political preferences (1.546), come from friends who have different political views (–1.60), and come from friends with whom one has a mutual relationship (0.52). This means that, when people decide to retweet political articles, they do care about their online social relationships (e.g., who shared, who is the audience). When it is an article contrasting their views, then social context becomes more significant. As such, contextualizing the news reading experience could offer ways of nudging people to accept a variety of political views.

7 Limitations and implications

7.1 Limitations

This work has some main limitations. First, our dataset shows biases, which might inherently come from the biases of the Twitter population. For example, we have an overrepresentation of liberal users, but that is because the number of left-wing media outlets is higher than that of right-wing ones. To check whether this would impact our results, we have also considered a 'balanced' (sample) dataset that includes the same number of left and right media outlets, and we found the results to be consistent in both.

The second limitation of our work is the method we used for classifying the topic of news articles - Alchemy. We find that *F2 topic-interesting-me* factor is not strongly related to news sharing behavior, which is counter-intuitive. Had only Alchemy been used, we might have been unsure whether our results hold true in general, or whether they are the product of classification artifacts. To validate Alchemy's categorization of news articles, we compared the classifications for the New York Times articles returned by Alchemy and the official classifications offered on the New York Times site. For example, a url http://www.nytimes.com/2009/04/13/us/politics can be categorized as 'Culture Politics' based on the URL itself. Showing 82% mcathing probability with New York Times' categorization, we believe that is it acceptable to use Alchemy. However, there is still a room to examine whether the lack of correlation of *F2* and retweeting probability is produced by the Alchemy or by its inherent absence of relationship.

Third, in building our PoNS model, we generate an artificial 50-50 positive-negative retweeting cases by taking random negative retweeting cases. Given a large number of tweets individuals receive, 50% of negative retweeting cases may not reflect the reality - in fact, only a few news articles are shared. However, such sample creation helps us to understand which of four factors in our PoNS model is the strongest one in relating to news sharing behavior. Yet, further investigation on the effect of samples can be conducted. For example one could test the model by changing the proportion of negative cases in generating samples.

Last but not least, we do not consider the sentiment of a user when he shares an article. If a user shares a news article of an hostile media outlet, it does not necessarily mean that he is vouching for it - he might simply make fun of it. Yet, what we observed from our analysis is that when an individual shares news articles that conflict with his own political view, it is about his friends' interests rather than his own, and this stays valid even thought we do not consider the sentiment of tweets. However, recent studies have emphasized the role of the sentiment of a tweet in its virality, especially when it is news content. Negative sentiment tends to be a strong promoter of news sharing [46, 47] and the stronger the emotion of a tweet is, the higher the chance it is retweeted [48, 49]. Thus investigating on how the sentiment of tweet come across with the factors we considered seems like an interesting follow up work. We leave this as future work.

7.2 Theoretical implications

This work has important implications for theories on information consumption, information sharing, and opinion diversity. Our results suggest that news sharing depends on four factors: (1) gratification; (2) selective exposure; (3) socialization; and (4) trust and intimacy. These factors have been studied before [28–31], but only separately, mainly because of lack of data. Here we have studied them together.

In terms of opinion spreading, it is tempting to think that Twitter allows us to connect to thousands of individuals who collectively hold diverse political views. The reality is that homophily limits who connects to whom (who one follows or is followed by, in Twitter parlance) - users are likely to connect to and exchange news articles with other like-minded users. As such, it is hard for ideas to pass between groups who are separated. Both online and offline, one important dimension separating groups of people is politics [50]. When people are separated by political views, they perceive each other as far apart and are unlikely to share opinions and offer any kind of support. This results into the creation of echo-chambers where like-minded individuals talk with each other and, as a result, reinforce each other's views. In our work, we found that Twitter users segregate themselves into echo chambers by sharing like-minded opinions even though they are exposed to different opinions.

In terms of opinion diversity, it is known that exposures to balanced information brings positive social consequences; it helps people set common grounds on important issues and improve group decision-making [51]. On the other hand, previous studies have also shown that exposure to balanced information does not change people's minds but, in contrast, increases commitment to original perceptions [52–54]. This effect is called *cognitive dissonance* [55], i.e., people tend to deny claims that contradict their beliefs. For example, exposing people to balanced political news generally leads them to hold more intense beliefs than they held beforehand. So the simple approach of exposing people to diverse political opinions might not work, and more sophisticated approaches should be used. Our study suggests that social context (e.g., one's followers) is associated with low levels of cognitive dissonance. Challenge-averse individuals were prepared to lose their reticence and retweet some articles with views different from their own - these articles generally came from friends.

In terms of information diffusion, there are a few studies on the relation of 'impact of number of exposures' to different outputs (e.g., hashtag adoption of Twitter [33] and Facebook fan page creation [56]). These studies all concluded that the more an individual is exposed to some piece of information, the more likely the individual will be persuaded by it. For example, [56] reports that '*after controlling for News Feed exposure variables, neither demographic characteristics nor number of Facebook friends seems to play an important role in the prediction of maximum diffusion chain length*'. Our study shows a similar trend and also finds that, after controlling for *numexposures*, other variables become important, and their importance changes across individuals: some users may like the popular stories (hence larger *numExposures*), while others value stories coming from close friends (hence *mutualfriend*).

7.3 Practical implications

We have found that users are more likely to retweet articles that are shared by their popular friends. This means that news aggregators might want to rank news depending on how popular or socially central the individual sources are. In general, offering personalized news articles on politics is more challenging than offering other types of articles. However, not all users find such exposure challenging. Support-seeking or diversity-seeking users are expected to be open minded and be willing to receive political news that do not necessarily reflect their own views. However, challenge-averse users may not appreciate such exposure. Our findings suggest that offering news through social-networking friends could be a reasonable way to 'scratch' challenge-averse users' echo chambers.

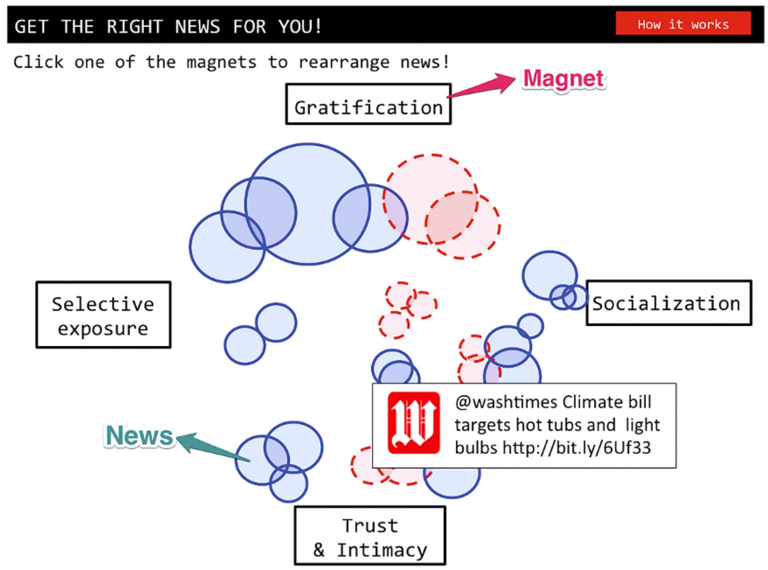

Figure 6 An example visualization of PoNS model based news presenting application.

Our findings have practical implications for the design of news aggregators. Twitter users strongly care about their followers' interests, including their political views. Traditional news aggregators return news a user might like based on the user's interest only. Our findings suggest that aggregators might also return news that not only are of interest to the user but also encourage interactions with friends.

Based on these findings, we introduce a new visualization for presenting news articles that gives users control over the PoNS' four dimensions (gratification, selective exposure, socialization, and trust & intimacy). This visualization is based on the Dust & Magnet visualization technique [57] that uses a magnet metaphor in which the individual data cases are represented as particles of iron dust, and magnets represent the different variables of the dataset. Users can interactively manipulate the magnets and then the dust moves appropriately.

One can develop an application that collects news articles a user receives in Twitter and present them with PoNS' four factors, providing more context to what they read. Figure 6 shows an example design of such application, which is run on a left-leaning user' news articles. The rectangles are magnets (i.e., PoNS' dimensions), and the circles are dust (i.e., news articles). The blue circles having solid line are news articles from left-wing media and the red ones having dotted line are from right-wing media. The size of each circle reflects the popularity of the news article based on the number of retweeted. By clicking any dust object, a user can see the detailed pop-up of the corresponding tweet.

A user can also click on any magnet to adjust the magnitude of attraction of magnet. When a magnet is clicked, dust particles are attracted to the magnet based on the value of the dimension corresponding to the magnet. For example, if the magnet represents socialization, a piece of dust with a higher value for socialization attracts more than a piece of dust with a lower value for it. As a result, users receives a sorted list of news articles. By allowing users to explore news articles along psychological dimensions, one could encourage them to expand their normal news reading patterns.

8 Conclusion

To counter information overload, people increasingly turn to their friends to receive filtered information as a proxy for relevance. If one hears about a story from a friend, then that story suddenly becomes relevant and salient even when its political orientation is different [58]. This established pattern of social behavior guides our actions not only offline but also online.

In Twitter, some of those who tend to be diversity-averse in their consumption of political news still promote stories they disagree with, and they do so because these stories are relevant to their online friends. This finding suggests that social ties are a proxy for relevance online. This striking resemblance with what happens offline happens likely because human behavior, which took thousands of years to evolve, changes much more slowly than the Web, which is only about 20 years old. As a result, it is easier for our online world to align itself with our offline world [59].

The media landscape continues to evolve over time and how people use certain medium also changes. This work has offered only one snapshot of the Twitter political landscape. To extend this, we will conduct a longitudinal study using the same 4-factor model, and see how the contributions of those factors changes over time and during large-scale events (e.g., elections).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
JA collected the data. JA, DQ, and MC performed the analysis. JA, DQ, MC, KG, and JC conceived the experiments and wrote the manuscript.

Author details
[1]Qatar Computing Research Institute, Majlis Al Taawon Street, Doha, Qatar. [2]Yahoo Labs, Avinguda Diagonal 177, Barcelona, Spain. [3]Graduate School of Culture Technology, KAIST, 291 Daehak-ro, Daejeon, Republic of Korea. [4]Max Plank Institute for Software Systems, Campus E1 5, Saarbrücken, Germany. [5]Computer Laboratory, University of Cambridge, 15 JJ Thomson Avenue, Cambridge, UK.

Acknowledgements
JA was supported in part by the Google European Doctoral Fellowship in Social Computing. MC was supported by the BK21 Plus Postgraduate Organization for Content Science in Korea.

Received: 1 March 2014 Accepted: 31 August 2014 Published online: 25 September 2014

References

1. Milyo J, Groseclose T (2005) A measure of media bias. Q J Econ 120(4):1191-1237
2. Gentzkow M, Shapiro JM (2010) What drives media slant? Evidence from U.S. daily newspapers. Econometrica 78(1):35-71
3. Della Vigna S, Kaplan E (2007) The Fox News effect: media bias and voting. Q J Econ 122:1187-1234
4. Zaller JR (1992) The nature and origins of mass opinion. Cambridge University Press, Cambridge
5. Glynn CJ, Herbs S, OKeefe GJ, Shapiro RY (1999) Public opinion. Westview Press, Boulder
6. Quirk PW (2009) Iran's Twitter revolution. http://www.fpif.org/articles/irans_twitter_revolution. Accessed 28 Jul 2014
7. Tumasjan A, Sprenger TO, Sandne PG, Welpe IM (2010) Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Proceedings of the 4th international AAAI conference on weblogs and social media (ICWSM'10)
8. Lotan G, Graeff E, Ananny M, Gaffney D, Pearce I, et al (2011) The Arab Spring - the revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions. Int J Commun 5:1375-1405
9. Diakopoulos N, Naaman M (2011) Towards quality discourse in online news comments. In: Proceedings of ACM conference on computer supported cooperative work (CSCW'11)
10. An J, Cha M, Gummadi K, Crowcroft J (2011) Media landscape in Twitter: a world of new conventions and political diversity. In: Proceedings of the 5th international AAAI conference on weblogs and social media (ICWSM'11)
11. Pariser E (2012) The filter bubble: how the new personalized web is changing what we read and how we think. Penguin Books, London
12. Severin WJ, Tankard JW (2000) Communication theories: origins, methods and uses in the mass media. Addison-Wesley, Boston
13. Sears DO, Freedman JL (1967) Selective exposure to information: a critical review. Public Opin Q 31(2):194-213

14. Christakis NA, Fowler JH (2009) Connected: the surprising power of our social networks and how they shape our lives. Little Brown and Company, New York
15. Krebs V (2008) Political polarization in Amazon book purchases. http://www.orgnet.com/divided.html USA. Accessed 27 Jul 2014
16. Bishop B (2008) The big sort: why the clustering of likeminded America is tearing us apart. Houghton Mifflin Company, New York
17. Mutz DC (1996) Political persuasion and attitude change. University of Michigan Press, Ann Arbor
18. Adamic LA, Glance N (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)
19. Gentzkow M, Shapiro JM (2011) Ideological segregation online and offline. Q J Econ 126(4):1799-1839
20. Livne A, Simmons MP, Adar E, Adamic L (2011) The party is over here: structure and content in the 2010 election. In: Proceedings of the 5th international AAAI conference on weblogs and social media (ICWSM'11)
21. Conover MD, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A (2011) Political polarization on Twitter. In: Proceedings of the 5th international AAAI conference on weblogs and social media (ICWSM'11)
22. Yardi S, Boyd D (2010) Dynamic debates: an analysis of group polarization over time on Twitter. Bull Sci Technol Soc 30(5):316-327
23. Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? In: Proceedings of the 19th international world wide web conference (WWW'10)
24. Java A, Song X, Finin T, Tseng B (2007) Why we Twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on web mining and social network analysis (WebKDD/SNA-KDD'07)
25. Yardi S, Boyd D (2010) Tweeting from the town square: measuring geographic local networks. In: Proceedings of the 4th international AAAI conference on weblogs and social media (ICWSM'10)
26. Boyd D, Golder S, Lotan G (2010) Tweet, tweet, retweet: conversational aspects of retweeting on Twitter. In: Proceedings of the 43rd Hawaii international conference on system sciences (HICSS'10)
27. Lin N (1999) Social networks and status attainment. Annu Rev Sociol 25:467-487
28. Lin C, Salwen MB, Abdulla RA (2005) Uses and gratifications of online and offline news: new wine in an old bottle. In: Online news and the public, pp 221-236
29. Diddi A, LaRose R (2006) Getting hooked on news: uses and gratifications and the formation of news habits among college students in an Internet environment. J Broadcast Electron Media 50:193-210
30. Dunne A, Lawlor M, Rowley J (2010) Young people's use of online social networking sites - a uses and gratifications perspective. J Res Interact Mark 4:46-58
31. Chiu C, Hsu M, Wang E (2006) Understanding knowledge sharing in virtual communities: an integration of social capital and social cognitive theories. Decis Support Syst 42:1872-1888
32. Latane B (1981) The psychology of social impact. Am Psychol 36(4):343-356
33. Romero DM, Meeder B, Kleinberg J (2011) Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter. In: Proceedings of the 20th international world wide web conference (WWW'11)
34. Lin YR, Margolin D, Keegan B, Baronchelli A (2013) # Bigbirds never die: understanding social dynamics of emergent hashtag. In: Proceedings of the 7th international AAAI conference on weblogs and social media (ICWSM'13)
35. Freeman S, Walker MR, Borden R, Latane B (1975) Diffusion of responsibility and restaurant tipping: cheaper by the bunch. Pers Soc Psychol Bull 1(4):584-587
36. Darley JM, Latane B (1968) Bystander intervention in emergencies: diffusion of responsibility. J Pers Soc Psychol 8(4p1):377
37. Stroud N (2011) Niche news. Westview Press, Boulder
38. http://thenextweb.com/twitter/2014/04/29/twitter-passes-255m-monthly-active-users-198m-mobile-users-sees-80-advertising-revenue-mobile/
39. http://twitter.com/#!/who_to_follow/interests/news
40. Quercia D, Askham H, Crowcroft J (2012) TweetLDA: supervised topic classification and link prediction in Twitter. In: Proceedings of the 4th annual ACM web science conference (WebSci'12)
41. Efron M (2004) The liberal media and right-wing conspiracies: using cocitation information to estimate political orientation in web documents. In: Proceedings of the 13th ACM international conference on information and knowledge management (CIKM'04)
42. Golbeck J, Hansen D (2011) Computing political preference among Twitter followers. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI'11)
43. Miller GA (1955) Note on the bias of information estimates. In: Information theory in psychology: problems and methods, pp 95-100
44. Munson S, Resnick P (2010) Presenting diverse political opinions: how and how much. In: Proceedings of the 28th ACM conference on human factors in computing systems (CHI'10)
45. Gelman A, Hill J (2006) Data analysis using regression and Multilevel/Hierarchical models. Cambridge University Press, Cambridge
46. Thelwall M, Buckley K, Paltoglou G (2011) Sentiment in Twitter events. J Am Soc Inf Sci Technol 62(2):1532-2882
47. Hansen LK, Arvidsson A, Nielsen FA, Colleoni E, Etter M (2011) Good friends, bad news - affect and virality in Twitter. In: Future information technology
48. Pfitzner R, Garas A, Schweitzer F (2012) Emotional divergence influences information spreading in Twitter. In: Proceedings of the 6th international AAAI conference on weblogs and social media (ICWSM'12)
49. Jenders M, Kasneci G, Naumann F (2013) Analyzing and predicting viral tweets. In: Proceedings of the 22nd international conference on world wide web companion (WWW'13 companion)
50. Sunstein CR (2001) Republic.com. Princeton University Press, Princeton
51. Nemeth CJ, Rogers J (1996) Dissent and the search for information. Br J Soc Psychol 35:67-76
52. Ross L, Lepper MR, Hubbard M (1975) Perseverance in self-perception and social perception: biased attributional processes in the debriefing paradigm. J Pers Soc Psychol 32(5):80-92

53. Sunstein CR (2009) On rumors: how falsehoods spread, why we believe them, what can be done. Farrar, Straus and Giroux, New York
54. Nyhan B, Reifler J (2010) When corrections fail: the persistence of political misperceptions. Polit Behav 32(2):303-330
55. Festinger L (1957) A theory of cognitive dissonance. Stanford University Press, Stanford
56. Sun E, Rosenn I, Marlow C, Lento TM (2009) Gesundheit! Modeling contagion through Facebook news feed. In: Proceedings of the 3rd international AAAI conference on weblogs and social media (ICWSM'09)
57. Yi JS, Melton R, Stasko J, Jacko JA (2005) Dust magnet: multivariate information visualization using a magnet metaphor. Inf Vis 4(4):239-256.
58. Hogan K (2004) The science of influence: how to get anyone to say "Yes" in 8 minutes or less! Wiley, Hoboken
59. Adams P (2011) Grouped: how small groups of friends are the key to influence on the social web. Pearson Education, Upper Saddle River

doi:10.1140/epjds/s13688-014-0012-2
Cite this article as: An et al.: **Sharing political news: the balancing act of intimacy and socialization in selective exposure.** *EPJ Data Science* 2014 **2014**:12.