Media landscape in Twitter: A world of new conventions and political diversity

Jisun An[*]   **Meeyoung Cha**[†]   **Krishna Gummadi**[‡]   **Jon Crowcroft**[*]

[*]University of Cambridge   [†]Graduate School of Culture Technology, KAIST   [‡]Max Planck Institute for Software Systems

Abstract

We present a preliminary but groundbreaking study of the media landscape of Twitter. We use public data on whom follows who to uncover common behavior in media consumption, the relationship between various classes of media, and the diversity of media content transmitted via social links. Our analysis shows that there is a non-negligible amount of indirect media exposure, either through friends who follow particular media sources, or via retweeted messages. We show that the indirect media exposure expands the political diversity of news to which users are exposed to a surprising extent, increasing the range by between 60-98%. These results, which have not been readily available to traditional media, can help predict how we will read news, and how media publishers will interact with the audience in the future.

Introduction

In the era of realtime web, people are shifting from scanning traditional media such as newspapers and television to using the Internet and social media sites like Twitter to find news. Social media has often scooped traditional media in reporting current events, for instance, in the recent turmoil in Egypt. While most original reporting comes from traditional journalists, social publishing and syndication platforms make it increasingly possible for an attentive audience to tap into breaking news. It has even been said that *news no longer breaks, it tweets* (Solis 2010).

The paradigm shift in media journalism, also known as micro journalism, has received much attention in recent years. Newspapers and magazines have begun publishing on social networking sites like Facebook and Twitter. Once passive, users now filter news and discuss what media publish. Moreover, they propagate interesting stories further into the social network at unprecedented scale and frequency. With the evolution of these new technologies, some experts predict traditional print journalism will ultimately disappear, to be replaced by new complex socially-mediated channels.

This great excitement has led to a number of studies that seek to understand the new social media. Studies have provided insights into the patterns of user participation in social media, such as propagation of news, extraction of urgent updates, and evaluation of news comments (Diakopoulos and Naaman 2011; Yardi and danah boyd 2010). Others have looked into the patterns of new journalistic conventions and how they affect the newsroom (Matheson 2004).

Building upon these studies, our paper characterizes micro journalism in one of the most popular social media, Twitter. Using publicly visible data on media sources, their followers, and interactions among them, we conducted a detailed analysis of the media landscape in Twitter: from the evolving practices in media publishing and consumption, to the shared readership between different types of media, and to the diversity of opinion social contacts bring. Conducting a similar study on the traditional media would have been difficult, as it would have required extensive surveys to gather the required data. By contrast, since all interactions in social media are recorded online and are often made publicly accessible, gathering and aggregating data—processes that are largely automated—can yield a view of the media landscape.

For our study, we used information on the follow links and tweets of 80 popular media sources and their 14 million audience members in late 2009. The 80 media sources comprise a diverse group: news and television networks, magazines, and journalists. Since Twitter is a young platform founded in 2006 and has not had enough time to evolve and consolidate, our dataset exhibits demographic biases like the over-representation of technology-savvy and liberal users. Despite these limitations, analyses in this paper provide a valuable snapshot of the social media landscape at an early stage of micro journalism.

We make several key observations. First, there is much about the media landscape in Twitter that is 'old media'. Established media outlets retain the role of publishing news and stories without much interaction with readers. However, the features of the 'new media' age are reflected in the way journalists and audience engage in new communication patterns, communicating with each other directly, and tapping into breaking news.

Second, users show a strong tendency to receive information from multiple media sources, especially on similar topics. Users are more likely to subscribe to multiple media sources within a given topic (e.g., political or technology news) than to media sources across different topics, (The average co-subscription probability within a topic was 1.4

times higher than that across different topics.) We also observe that certain media sources, especially journalists, excel in connecting media from different topics, indicating that Twitter users who follow journalists tend to seek more diverse types of information. These findings can clearly be observed by visualization of a map of media landscape based on the directional distance measure in *Eq.(1)*.

Third, compared to their audience through direct subscription, media organizations reach a considerably larger audience through indirect exposure via social links, for instance, through a friend who follows a particular media source, or via a retweet from a friend. The exact benefits from indirect exposure vary from one source to another. Indirect exposure also increases the diversity of media sources from which an audience gets its information; we found that with indirect exposure, users receive information from six to ten times more media sources than from direct exposure alone.

Lastly, indirect media exposure increases the diversity of political opinions seen by users: between 60-98% of the users who directly followed media sources with only a single political leaning (left, right, or center) are indirectly exposed to media sources with a different political leaning. In order to reach this conclusion, we use public classification of news sources and infer the political preference of every audience member. One can only speculate about the effect of political diversity, because users do not necessarily read the complete Twitter timeline nor do they always prefer receiving diverse political opinions (Munson and Resnick 2010). Nonetheless our results show the power of social media, in that users are exposed to information they did not know they were interested in, serendipitously.

Methodology

The Twitter dataset We used the Twitter data gathered from our previous work (Cha et al. 2010), which comprises the following three types of information: profiles of 54M users, 1.9B directed follow links among these users, and all 1.7B public tweets that were ever posted by the collected users. For the analysis, we first identified a list of media sources by consulting Twitter's 'Find People' directory[1] and *http://wefollow.com*, a user powered directory service that lists popular Tweeters by genre. From those two lists, we searched for categories related to media, including news outlets and magazines, as well as individuals in those businesses, such as reporters, journalists, and magazine editors. We only considered sources having at least 10,000 followers in order to make sure that each media source had a large audience. This left us with a total of 80 media sources.

We collected all follow links to media sources and the tweets posted by them. Also for each user following the media source, we gathered her follow links and her tweets in order to study how users interact with media on Twitter. Users replying to a media source were identified based on the inclusion of '@medianame' in their tweeted text. In addition we identified users retweeting based on the inclusion of an 'RT' or 'Via' in the tweeted text.

Genre	Account	Followers	Tweets	Mentioned
News	cnnbrk	2,596,796	1,078	28,499
(40 sources)	nytimes	1,755,740	35,822	75,108
	TerryMoran	895,157	463	2,633
Technology	BBCClick	1,165,991	370	309,630
(13)	mashable	1,270,763	4,217	392,158
Sports	NBA	1,172,755	2,115	8,143
(7)	nfl	981,309	503	1,674
Music	MTV	294,971	1,769	10,569
(3)	iTunesTrailers	814,011	371	10,888
Politics	nprpolitics	1,145,170	3,630	17,676
(5)	jdickerson	953,993	3,470	11,438
Business	davos	750,523	553	1,210
(2)	alleyinsider	861,715	3,114	10,363
Fashion &	themoment	1,094,496	1,679	4,652
Gossip (4)	peoplemag	1,289,415	1,083	10,020
Others	trazzler	944,266	1,362	2,562
(6)	goodhealth	653,939	1,339	8,096

Table 1: Summary the 80 media sources studied

In our dataset, the 80 media sources had posted 189,083 tweets and had a total of 14,158,007 subscribers. Media tweets were generally popular, resulting altogether in 2,940,013 retweets and 623,630 replies. Including the retweets, we used information about the profiles of 35.7M Twitter users and 1.2B directed follow links. Table 1 displays the major categories by genre with a few representative sources in each category.

Media and audience characteristics Out of the 80 media sources, 20 of them were individual reporters and journals; a majority of the journalists belonged to the 'news' category, while some belonged to 'technology' and 'politics.' In terms of geography, the majority of the news sources were US-based although certain news sources like CNN, BBC and the Onion have international appeal.

A media tweet was 3.3 times more likely to contain a URL (73%) than an average tweet. A retweet indicates that the original post by the media source is shown to Twitter users who are not subscribing to the media source. Hence, a high rate of retweeting of a media source implies that there is a significant amount of *indirect* exposure or subscription to that media through the social network on Twitter. One of our goals in this paper is to quantify this exposure.

Since each of the 80 media sources we studied has a large number of followers, our findings in this paper are likely to apply to mass media sources, but not necessarily to smaller media sources. The set of users who followed at least one of the 80 media sources is estimated to contain nearly a quarter of all Twitter users, based on the counts of total registered users.

Social media landscape

Social media has fundamentally changed the way people interact with media. Open platforms like Twitter allow one to get a global picture of the media landscape—a landscape that was previously hidden and private. In this section, we
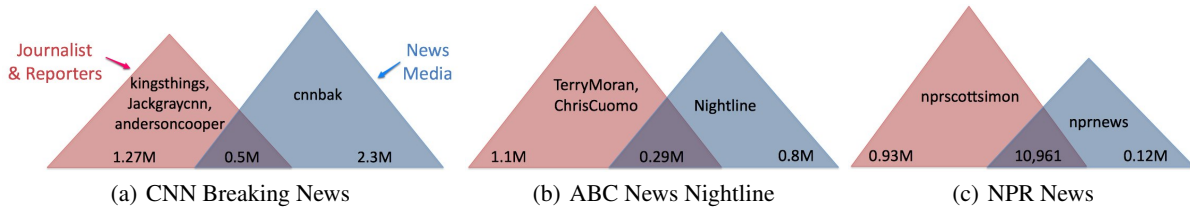
Figure 1: Comparison of the size of audience for the mainstream media and journalists

present the key features of social media and a map of media landscape based on the co-subscription pattern of users.

Key features of social media

We highlight two prominent features of social media: the role played by journalists and user participation in disseminating information.

Redefined role of media journalists Journalists in traditional media stayed out of the spotlight behind their publications. They had few means of interacting with (or "listening to") their audience. Social media journalists, in contrast, can reach the audience directly and build a more personal presence. The personal popularity of journalists is shown in Figure 1, which shows that the size of a journalist's audience often rivals the size of his media organization's audience and sometimes even exceeds it. For example, Scott Simon, a reporter at NPR, has nearly 940,000 followers—7 times more than NPR itself. Furthermore, contrary to our expectation that the audience of journalists would largely overlap with those of media organizations, the overlap in audience between the two was relatively small.

The reasons for why some users follow journalists instead of media organizations might be rooted in the different ways journalists and media organizations interact with their audience. First is the published content. We found that journalists' tweets are not always constrained to news stories. They sometimes covered personal updates and greetings. Second is the conversational nature of interaction between journalists and audience. We found that journalists are 6 times more likely than media organizations to receive a mention for each of their tweets. Journalists also converse reciprocally with audience; they are 5.4 times more likely than media organizations to reply to mentions received. These behaviors of journalists humanize the media source and make it easier for an audience to feel more connected with them.

On the other hand, journalists also tapped into news publication by posting the same URL that was posted by their media organizations. The purpose of co-tweeting was diverse. Usually it served an immediate need for sharing breaking information with a wider audience. For instance, CNN and its reporter Anderson Cooper both tweeted breaking news on the pardon of two U.S. journalists in North Korea and incurred 915 and 270 retweets, respectively. Co-tweeting was sometimes intended to encourage participation or raise anticipation for readers about the upcoming events, e.g., raising funds or sneak peek of upcoming shows.

User participation The empowerment of users is another prominent feature of social media. Twitter users retweet notable events and participate in the spread of realtime news.

While our goal is not to characterize retweeting behaviors, we examined what kinds of media tweets were most likely to spread deep and wide in the social network and whether users affect which news stories are covered in the media.

Mapping retweets to their original tweet is not easy, because users add short sentiments and modify the text as they retweet. Hence we used URLs that appear in tweets as a clean piece of information to track retweets from media sources and users, as follows. An edge from user A to user B is added to the retweet tree only when B follows A, B posted the URL after A posted the same URL, and B is not already a part of the tree. There were 138,432 media tweets with URLs, which spawned an average of 15.5 retweets each. Some media tweets propagated far; the longest retweet tree had 7 consecutive retweet hops. Some propagated to many other users in a single hop; the widest fan-out at any given retweet hop was 17,978.

Media tweets that generated a wide retweet pattern were on topics like breaking news, technology, politics and sports. These topics were not always those stories that appealed to the masses. Inspection of the tweets in the top 50 longest chains showed that topics that propagated far were often emotional such as comic videos or highly sarcastic news. For instance, a short video clip posted on the Today Show website titled 'dad surprises daughter with return from Iraq', which many Twitterers described as "very touching", was retweeted 84 times. Twitter in this case acted as an echo-chamber allowing various different communities to easily access the news.

Map of social media landscape

Social media serve as a useful platform for studying the media landscape since all interactions from both media sources and the audience members are recorded online. Here we examine which media sources are co-subscribed by users. We say that media sources with many subscribers in common are *closely related*.

As a measure of *closeness*, we calculate the fraction of common audience. Intuitively, the closer two media sources are the more their audiences overlap. Let A represent the media of interest and $\{B_1, B_2, \cdots, B_n\}$ be the set of n other media sources for which we would like to measure the distance from A. Then, the closeness value of A from B_i is defined as the probability that a random follower of B_i also follows A:

$$\text{closeness}(A|B_i) = \frac{|A \cap B_i|}{|B_i|} \tag{1}$$

For every media source, we calculated *Eq.(1)* [2] to all other

_____

[2] We also tested other variations such as $\frac{|A \cap B|}{|A|} \cdot \frac{|A \cap B|}{|B|}$ and

media sources and examined which ones appear the closest. Once the closeness values were calculated between all directed pairs of media sources, we used them to plot a map of the media landscape, as in Figure 2.
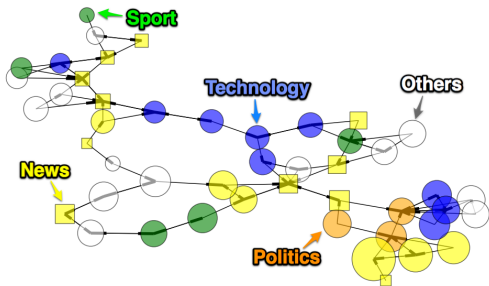

Figure 2: Pictorial map of the media landscape in Twitter

Nodes represent media sources and links represent co-subscription relationships. The node size is scaled to represent the log of audience share and its color represents topical categories. All 80 media sources are used in the analysis, but for visual clarity we show the map only for nodes that have more than 100,000 followers and their two strongest relationships per media, following the methodology in (Bollen et al. 2009) to unclutter the map. The distance between nodes is determined automatically by the Graphviz graph layout and visualization package, which tries to optimize the node positions based on the closeness values. Hence, a short distance in the map could be an artifact of visualization.

The map shows a macroscopic view of the social media landscape—a pictorial description that could not be easily obtained from traditional social media. Nodes of the same color associate and bond together (i.e., homophily), indicating a strong tendency of users to receive information from multiple sources, especially on similar topics. For example, a user following NBA had a high chance of also following NFL. The average co-subscription probability within a topic, or intra-category, was 1.4 times higher than that across different topics.

In the inter-category subscription, certain nodes connected different types of media. One non-obvious example is Larry King, shown in the bottom left corner of the map, who had a high closeness value to People magazine, a celebrity gossip magazine. Interestingly, journalists (indicated as rectangles in the map) often connected media of different categories. There may be several reasons for why certain nodes excel in connecting different types of media, such as subject matter, tone of voice, or audience with varied interests. We leave investigating this as future work.

In this section, we observed several key changes of journalistic conventions and cultures taking place in social media. However, the most significant characteristic of social media is the connection created among people. Since people share news with their friends, it is hard to say that media consumption is limited to direct media subscription. There is a non-negligible amount of indirect exposure to media

$\frac{|A \cap B|}{|A \cup B|}$, which are symmetric distance measures. We used *Eq.(1)* because it represents the *directional* distance measure.

through social links. In the following section, we examine how the map of the social media landscape changes when we incorporate the benefits social links bring to both media sources and their audience.

Impact of social links

The example in Figure 3 depicts the follow relationship (in the upward direction) between one media source, NBC, and five users. We consider the following three routes to media exposure, which we explain using the example.

- Direct subscription: when a user follows a media source (e.g., user A and user B in the toy example)
- Social network: when a user follows another user who follows a media source (e.g., users C, D, and E)
- Social interaction: when a user receives tweets from a media source via another Twitter user (e.g., user D and user E, where user B retweeted NBC's tweet)
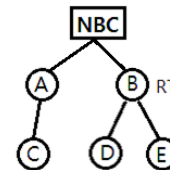

Figure 3: Example of a media subscription network

Unlike direct subscription, the latter two types represent what we call *indirect* exposure. Media exposure through the 'social network' is subtle and implicit—it represents merely a hint of media influence due to one's followees. By contrast, media exposure through 'social interaction' is explicit; the user is actively exposed to media because his followees have retweeted that media outlet's tweet. Since a media tweet receives on average 15.5 retweets, there is continual indirect exposure to media through social interaction.

In order to quantify the increase in media exposure due to social links, direct subscription is included to the other two cases. Hence, by 'social interaction', we refer to the case when a user either directly subscribes to a media source or receives a retweet from media through his followees.

Media perspective: impact on audience reached

We first investigated how social links impact the audience size for the media sources. Accounting for indirect media exposure increased the audience size for all 80 media sources. Figure 4 shows, for every media source sorted in decreasing order based on the number of direct followers, the ratio of the audience size under two types of indirect exposures against that of direct subscription.

The social network increases the reach of media sources to varying degrees. Through the social network, some media sources show a 2-10 fold increase in their reach and others a 100-fold increase. On average, media sources increased their audience by a factor of 92.8. Social interaction shows a slightly smaller increase of 28-fold on average. The ratio tends to increase rapidly for lower ranked media sources, shown as an upward trend in the y-axis that is in log-scale.

This is because the top ranked media sources have disproportionately large audiences to begin with, making them less susceptible to a larger-scale increase in audience after indirect media exposure.
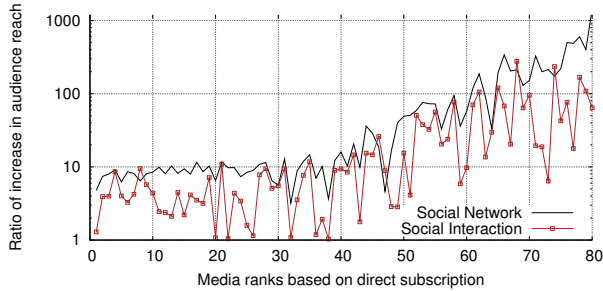

Figure 4: Increased readership due to indirect exposure

Most media sources reach a larger audience through their 2-hop social network than by social interaction. However, five media sources gained a larger fraction of audience through retweets. These media sources, including BBCClick, MLB and Jordan Times, are examples of media sources that benefit largely from social interaction.

Accounting for indirect media exposure results in differing amounts of increase in audience size for the 80 different media sources. As a result, certain media sources became more popular than others. Table 2 displays the top 10 media based on degree centrality, which is an indicator of the total subscription to the media across the three paths to media exposures.

Rank	Dir-Sub	Soc-Net	Soc-Int
1	cnnbrk	cnnbrk	BBCClick
2	nytimes	nytimes	mashable
3	TheOnion	mashable	timoreilly*
4	mashable	kingsthings*	todayshow
5	eonline	TheOnion	nerdist*
6	nprpolitics	andersoncooper*	TEDchris*
7	NBA	maddow*	MLB
8	BBCClick	timoreilly*	om*
9	peoplemag	nprpolitics	espn
10	GMA	davidgregory*	AnnCurry*

Table 2: Top ranked media based on degree centrality. Journalists are marked with the * sign.

Under direct subscription, established media sources like CNN and NYTimes made up the top list, each having several millions direct followers. Under social network exposure, the list changes significantly: it shows a 50% overlap with the list based on direct subscriptions, but the other half of the list was made up by media journalists like Rachel Maddow, Larry King and Anderson Cooper. The size of the audience has also increased by 10M-13M, reaching approximately a quarter of all Twitter users. The overlap in this top list is lower for social interaction (20%), in which case even the set of established media outlets changes and the Today Show, MLB, and ESPN turned out to be most effective in spawning retweets.

The increase in exposure of media journalists was unexpected, because they do not necessarily have the largest direct subscription. However, users could easily find these journalists through their social links. The median number of followers of journalist's audience was around 100, which was more than twice as high as the number of direct followers of established media sources. For instance, the NYTimes had a median of 6 followers. It implies that users who have a large following, called power users, tend to follow journalists. Since many users follow these power users, power users can play a role as a connector between journalists and those people who are not interested in news.

Audience perspective: impact on media diversity

Indirect media exposure increased the diversity of media sources for the audience members. The number of distinct media sources that an individual user subscribes to is not high, shown here as a solid line in the cumulative distribution function plot in Figure 5. 80% of users follow up to 10 media sources, typically from 2-3 different media types. The most popular media types were news, sports, and technology.
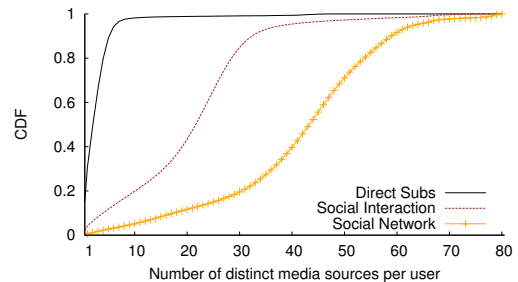

Figure 5: Diversity in the number of media sources per user

Under indirect exposure, 80% of the users hear from up to 55 and 27 media sources through social network and social interaction, indicating a 16 and 8 fold increase respectively. The number of media types accordingly showed an increase. A user's probability of subscribing to multiple types of media was around 30.1% for direct subscription, but 74.5% and 92.5% through social interaction and social network, respectively.

In summary, social links broaden the types of information users receive and allow less popular media to gain more presence. As a result, incorporating indirect media exposures changed the map of media landscape in Figure 2 significantly. We omit visualization due to space constraint. Smaller media outlets have a chance of being top ranked, which resulted in more similar node sizes in the map. Due to a high probability of inter-category exposure, grouping of the same color became less obvious.

Next we present a case study on the landscape of news, one of the most popular genres, and examine whether social links help users receive politically diverse news.

Does social media increase opinion diversity?

Media stations and newspapers are known to have some degree of political bias, liberal, conservative or other. For example, The New York Times and Chicago Tribune provide

different viewpoints in their coverage of stories on health care and national defense. In this case, it can be a challenge to determine the bias for readers who seek a diverse selection and viewpoints of stories.

In this section, we ask *to what extent do users receive politically diverse opinions through social links?* Answering this question is not easy, because one needs to know the political leanings of all the other connected users. We describe a methodology to infer the political preference of users and to measure political diversity in news subscriptions. We later compare this result with the level of political diversity found in direct media subscription alone. We also summarize the result of a small-scale survey that tries to identify the benefits of seeking diverse opinions in Twitter.

Method for inferring political leaning of users

We mapped the political leaning of media sources into three groups, left-wing (liberal), center, and right-wing (conservative), using a number of public data including a seminal paper by (Milyo and Groseclose 2005) and web resources such as *http://www.left-right.us/about.html*. In cases where public data conflicted on the political leaning of a specific media source, we used the most recent data. We also tried to find a fine-grained mapping, for instance ABC News is considered left-wing, but one of its outlets, Good Morning America, is considered center.

Not all 40 news media sources could be classified due to a lack of public data, but we could classify 34 of them. Only four of them, Fox News, Chicago Tribune, U.S. News & World Report, and Washington Times, were classified as right-wing and nine others including BBC and CNN were classified as center. The remaining 21 news outlets classified as left-wing included Huffington Post, NPR, and NYTimes.

Our inference method is based on examination of the media list a user subscribes to. As a rule of thumb, we assumed that a user follows media sources that match his political leaning if all the media sources he follows are of one political leaning. These users with clear political preference, whom we call *seeds*, accounted for 7 million or 50.8% of all audience. Most seeds were either left-wing (62%) or center (37%). Only 61,164 users (1%) were right-wing.

Measuring the impact of social links

Once we considered indirect media exposures, some seeds became more neutral in their political leaning and others retained their political leaning. Table 3 shows the fraction of users in each seed group whose political preference changed.

Many users connected to at least one friend who followed news media with a different political view (see Table 3A). Only 1.7% of the left-wing seeds remained unaffected by the social network at all. For the remaining users, 68.7% of them had at least one friend who followed right-wing media. The largest fraction of right-wing seeds (17.4%) remained unchanged in their political view, yet 75% of them received indirect exposures to left-wing media. Center-view seeds also received a high rate of indirect exposures to left-wing (93.6%) and right-wing (57.4%) media.

This high level of political diversity does not necessarily mean that Twitter users friend with those who have po-

Seed	Unchanged	Indirect media exposures to		
		Left	Center	Right
A. Social network				
Left	1.7%	-	88.1%	68.7%
Center	6.3%	93.6%	-	57.4%
Right	17.4%	75%	79.2%	-
B. Social interaction				
Left	20.4%	-	77.9%	17.8%
Center	17.9%	81.9%	-	15.2%
Right	40 %	57.2%	40%	-

Table 3: Political leaning of the known seeds, after incorporating indirect media exposures through social links

litically conflicting views. The trend may be due to users having a news junkie friend who follows many news outlets of diverse political views. In fact, 29.3% of news readers were directly following more than one political opinion from Twitter. Furthermore, a simple connection to a friend who gets different political opinions is after all a *subtle* influence.

Unexpectedly, we observed a high level of political diversity also in a more straightforward type of influence—social interaction (see Table 3B). Compared to the social network exposure, social interaction left more users unchanged in their views. However, a non-negligible portion of users were exposed to different political views; 17.8% of the left-wing seeds received retweets from right-wing media and 57.2%, vice versa. Our finding suggests that social media, whether the exposure level is subtle or straightforward, helped users get diverse opinions.

Why Twitter users follow multiple news sources

To better understand the cause of this high level of political diversity, we conducted a survey on Twitter and asked users what benefits they get from following multiple news outlets. We randomly picked 100 users following both left-wing and right-wing media and sent them questionnaires through Twitter. We received 56 responses. Many of them introduced themselves as 'independent news junkies' and had between 2,000 and 11,000 followers. The list below is not exhaustive, but attempts to capture representative replies.

- **On the purpose of following both sides of media:**
  - To be well balanced and well informed (e.g., *Following both sides somewhat stabilizes the spinning for me:-)*)
  - To know the enemy better (e.g., *I follow both because as a conservative you need to know your enemy, too.*)
  - For self-gain and entertainment (e.g., *I find politics and its coverage fascinating so I follow both.*)
  - Distrust of existing media (e.g., *Journalists are neither leftist nor rightists. They are typists.*)
- **On the usefulness of receiving diverse opinions:**
  - For accurate judgment (e.g., *I like to get all the facts from both sides and then make my choice on the truth.*)
  - To improve one's ability to refute what they are against (e.g., *Help me understand the opposing view point.*)
  - Diversity as valuable experience (e.g. *Understand both sides is a very valuable perspective. I wish more people tried to understand both sides.*)

thejoelstein L (66.3)	– TerryMoran L	– jdickerson L (66.6)	– Nightline L	– nytimes L (73.7)	– cnnbrk C (56.0)	– andersoncoopser C (56.0)	– bbcbreaking C	– foxnews R (39.7)	– usnews R	– washtimes R (35.4)

nprpolitics L (66.3)	– nytimes L (73.7)	– jdickerson L (66.6)	– Nightline L	– TheOnion L	– nprscottsimon L (66.3)	– GMA C (56.1)	– bbcbreaking C	– foxnews R (39.7)	– usnews R	– washtimes R (35.4)

washtimes R (35.4)	– washingtonpost L (66.6)	– foxnews R (39.7)	– usnews R	– bbcbreaking C	– earlyshow L (66.0)	– nytimes L (73.7)	– ariannahuff L	– ObamaNews L	– nprpolitics L (66.3)	– TheOnion L

Figure 6: Media landscape for three news sources marked in rectangles, based on the closeness measure in *Eq.(1)*. Political leaning of each media source is given under its name, L (left), C (center) and R (right), along with the ADA score when possible.

The survey result shows that users find it useful to receive diverse opinions, whether the purpose of multiple subscriptions was for information, entertainment or experience.

Political diversity in the map of media landscape

Finally revisiting the measure of distance between media sources, we investigated if the political dichotomy naturally arises when we only consider direct media subscription.

A well-known measure that quantifies such bias is the ADA (Americans for Democratic Action) score, which is calculated based on various quantities such as the number of times a media outlet cites various think tanks and other policy groups (Milyo and Groseclose 2005). The ADA score is scaled from 0 to 100, where 0 means strongly conservative and 100, strongly liberal. For instance, ADA scores are 39.7 for Fox News and 73.7 for NYTimes. The original reporting listed ADA scores of 18 news sources, out of which we found 10 media outlets in our dataset. They comprise 6 left-wing, 2 center, and 2 right-wing news sources.

For each one of these 10 news sources, we picked 4 nearest neighbors based on the absolute difference in ADA scores, then checked what fraction of them appear in the set of 5 nearest neighbors based on the closeness value defined by *Eq.(1)*. The two measures gave similar results with a high matching probability of 77.5% and a perfect match for every left-wing news source.[3]

A striking similarity in the two measures implies that, when direct subscription is considered alone, most Twitter users receive only biased political views they agree with. However, as Table 3 shows, the news media landscape changes dramatically under the influence of social links. The majority of users have access to politically diverse views.

Having shown that the closeness values could represent ADA scores relatively well, we further examined the map of the news media landscape for all of the 34 media outlets whose political leaning could be clearly identified. Figure 6 shows three examples for a journalist Joel Stein with left-wing media Time magazine, left-wing media NPR's political section (which had its own Twitter account), and right-wing media Washington Times. For each of them, the list shows the 10 most popular media sources based on the closeness measure, from left to right. Here, we assigned media journalists the ADA scores of their corresponding media organizations, unless their political leaning differed.

---

[3]The probability remained high (72%) when matched against the 5 nearest neighbors, instead of 4.

The ADA scores and the closeness values again matched very well. For instance, NPR Politics had the highest closeness values with NYTimes, indicating a strong liberal view. Center media like Good Morning America (GMA) and right-wing media like Fox News had low closeness values with NPR Politics, appearing towards the right hand side. It is noteworthy that the map of the media landscape, generated automatically from the co-subscription patterns of users without any consideration of tweet content, matched the ADA scores that require a tremendous amount of sophisticated text classification.

There were, however, a few exceptions. Left-wing media Washington Post is closest to Washington Times, which is a right-wing media. Despite the opposing political views, the two may have a high closeness value because both are popular newspapers from the same politically active region, i.e., Washington, D.C., and people in the region may follow both to learn more about happenings in the area like local events or sports.

Discussion

Our observations about the media landscape in Twitter help researchers understand how people read news and media updates in social media. Our findings also offer useful insights for future studies on media journalism. For example:

Distance measure The map of the media landscape, which is based on gathering online data and aggregating it via a distance measure, is useful because it did not require any complex procedures or private data as in the past. Hence, our methodology is suitable for a large-scale, repeated study. In fact, media landscape continues to evolve over time and the relationship among media sources changes. For example, U.S. News & World Report was once classified as left-leaning, but now is classified as right-leaning. In such case, our distance model can be used to track the transition phase over time.

Social filters Media tweets that had long and wide retweet chains were not always on topics that appeared as top stories in the media outlets' websites. Often the top stories in social media were urgent, emotional, critical, sarcastic, or humorous. This means that widely repeated media stories in Twitter are chosen with an authentic voice, and reading media updates ordinary users share, so called *social filters*, provides a unique experience for social media users. In this case, users who filtered content did more than a mere transmission of 'data bits' of information—a trend that has

been also shown in the blogosphere (Matheson 2004). In fact, some users are strong advocates of social filters and claim they no longer go to mainstream media directly to read news.[4] While we do not argue that social filters like Twitter can entirely replace news subscription because the two emphasize different sets of popular topics, our study showed that social filters could increase the audience reach of media sources significantly, by a factor of 28.

Political diversity Indirect media exposure expanded the political diversity of news users obtained by a significant amount. Being exposed to different viewpoint is known to bring positive social consequences: it helps people find common grounds on important issues and improve group decision-making skills. Nonetheless, we deemphasize the potential benefit of such political diversity because not everyone prefers to receive diverse political opinions, for instance in the case of challenge-averse individuals who only seek opinions that support their own views (Munson and Resnick 2010). Hence different strategies are required to assist heterogeneous individuals when news aggregators plan to increase opinion diversity.

The high political diversity also indicates that audience members reading different political views are connected to each other. This observation is particularly surprising because other studies have found a stronger tendency of homophily; blogs of different political views rarely linked to each other (Adamic and Glance 2005). However, social links in Twitter were less dichotomous in political views. One possible reason is that Twitter network encompasses several different relationships, from shared interest, to familial ties, friends, and acquaintances, so that political similarity doesn't necessarily exist in all such ties. We wish to understand this phenomenon better by conducting surveys in the future.

Related work

Several studies exist on the journalistic aspect of social media. In (Hermida 2010), authors referred to Twitter as a system of ambient journalism and addressed the need to study such phenomenon. Towards understanding news propagation, (Yardi and danah boyd 2010) looked into characterized the use of Twitter as a medium of local news sharing. Expanding on these work, we examined the macroscopic view of the media landscape in Twitter.

Pertinent to our study of the news media landscape are (Milyo and Groseclose 2005) and (Gentzkow and Shapiro 2010) which studied the political orientation of mainstream news media. The distance model in this paper could be further developed as an alternative way to identify media bias with much less effort. Orthogonal to our study, several interesting systems have been proposed that visualize the news in novel ways in order to help readers receive diverse opinions. For example, NewsCube (Park et al. 2009) and BLEWS (Gamon et al. 2008) detect various political aspects of news articles, classify them, and suggest a grouping with opposite views.

---

[4]*http://www.salon.com/news/feature/2009/07/28/ wired*

Conclusion

This paper presented a first-of-its-kind, although preliminary study of the media landscape of Twitter. While there is much about the media landscape that is as it was for traditional media, we identified several key changes in journalistic conventions and cultures taking place in this new world, including active participation of media journalists and audience members and increased diversity of information distributed via social contacts. The indirect media exposure through social links also broadened the opportunity for users to receive updates from politically diverse media outlets.

We hope that this paper represents a step towards understanding how people will read news and how publishers will interact with them in the future. We hope to understand the dynamics of this media landscape by repeating the study in the future. Our study has implications for various inference problems in social media; for instance, in determining the political leanings of influential individuals (e.g., politicians, celebrity) and their role in the media landscape.

Acknowledgement

References

Adamic, L. A., and Glance, N. 2005. The political blogosphere and the 2004 U.S. election: Divided they blog. In *ACM SIGKDD International Workshop on Link Discovery*.

Bollen, J.; de Sompel, H. V.; Hagberg, A.; Bettencourt, L.; Chute, R.; Rodriguez, M. A.; and Balakireva, L. 2009. Clickstream data yields high-resolution maps of science. *PLoS*.

Cha, M.; Haddadi, H.; Benevenuto, F.; and Gummadi, K. 2010. Measuring user influence in Twitter: The million follower fallacy. In *ICWSM*.

Diakopoulos, N., and Naaman, M. 2011. Towards quality discourse in online news comments. In *CSCW*.

Gamon, M.; Basu, S.; Belenko, D.; Fisher, D.; Hurst, M.; and Konig, A. C. 2008. Blews: Using blogs to provide context for news articles. In *ICWSM*.

Gentzkow, M., and Shapiro, J. M. 2010. What drives media slant? evidence from U.S. daily newspapers. *Econometrica, Econometric Society* 78(1):35–71.

Hermida, A. 2010. Twittering the news: The emergence of ambient journalism. *Journalism Practice*.

Matheson, D. 2004. Weblogs and the epistemology of the news: Some trends in online journalism. *New Media & Society*.

Milyo, J., and Groseclose, T. 2005. A measure of media bias. *The Quarterly Journal of Economics* 120(4):1191–1237.

Munson, S., and Resnick, P. 2010. Presenting diverse political opinions: How and how much. In *ACM CHI*.

Park, S.; Kang, S.; Chung, S.; and Song, J. 2009. Newscube: Delivering multiple aspects of news to mitigate media bias. In *ACM CHI*.

Solis, B. 2010. The information divide between traditional and new media, *http://tinyurl.com/ya8etcf*, Internet Draft.

Yardi, S., and danah boyd. 2010. Tweeting from the town square: Measuring geographic local networks. In *ICWSM*.