Dissemination Biases of Social Media Channels:
On The Topical Coverage of Socially Shared News

Abhijnan Chakraborty[*][#]**, Saptarshi Ghosh**[o]**, Niloy Ganguly**[*]**, Krishna P. Gummadi**[#]

[#]Max Planck Institute for Software Systems, Germany
[*]Indian Institute of Technology Kharagpur, India
[o]Indian Institute of Engineering Science and Technology Shibpur, India

Abstract

In a marked departure from traditional offline media, where all subscribers of a particular news media source (e.g., New York Times) used to get the same news stories through printed newspapers, online news media presents multiple options for the readers to consume news. For example, the subscribers of a media source can get news directly from the news website, or from what their peers share over social media sites like Facebook and Twitter. It is, however, unclear whether there are any differences in the news disseminated on these different online channels. In this work, we analyze data from a popular online news media site (`nytimes.com`), and show that each of these different channels tends to highlight some types of stories more than other stories. We believe that consumers of online news as well as media organizations need to be aware of such differences in various online news dissemination channels.

Introduction

As the number of users receiving news via traditional offline methods (e.g., via print newspapers or weeklies) is in steep decline, *online news media sites* like `nytimes.com` and `cnn.com` are emerging as the primary sources of news for people world-wide. A recent survey by the Pew Research Center (Pew 2012) found that the proportion of Americans reading news on a printed newspaper halved to 23% in 2012 compared to 47% in 2000. On the other hand, 55% of the regular readers of The New York Times declared that they read the news stories online, similar to 48% of regular USA Today and 44% of Wall Street Journal readers.

The wide-spread adoption of social media sites like Facebook and Twitter is fueling the growth of online news consumption further. In a separate survey (Mitchell et al. 2014), Pew Research Center found that around 48% of American Internet users got politics news via social media sites like Facebook, almost as many as those that got such news from local television channels.

In a marked departure from traditional offline media, where all subscribers of a particular news media source (e.g., New York Times) used to get the same news stories through printed newspapers, online news media presents multiple options for the readers to consume news. For example, a

user can directly visit the website (e.g., `nytimes.com`) and read the stories published therein. Whereas, another user may only read the news stories shared by her social contacts on Facebook or Twitter. Effectively these different news dissemination mediums become news channels for the users who can consume news from one or multiple such channels.

Given the complex landscape of online news consumption, it is important to understand the differences in the news consumption process contributed by each of these different news dissemination channels. In this work, we investigate whether there are systemic differences in coverage of news stories over different channels. Specifically, we attempt to understand whether certain types of news stories are covered more in certain channels than others.

For our study, we gathered extensive data from one of the world's most popular online news media sites, `nytimes.com` (henceforth referred to as NYTimes). We collected all news stories published at NYTimes over a period of 6 months. We further gathered the stories which became most popular across different dissemination mediums (e.g., most viewed on website, most emailed, most shared on Facebook, most shared on Twitter) during this 6 month period.

We analyzed the differences in the *topical distributions* of stories covered by different mediums. Our analysis demonstrates that there are significant differences in the topical coverage of news stories that gain popularity over different dissemination channels. For example, opinions and local stories (related to New York and US regions) tend to be more widely shared on Facebook, while business and world news stories tend to be shared more on Twitter.

It is unclear whether the online news consumers are aware of such differences in the topical coverages of the different mediums from which they consume news stories. Additionally, designers of various news recommendation systems also use the popularity of news stories in different mediums as signals to rank / recommend news stories. Through this study, we want to spread awareness both among the news consumers and the designers of news recommendation systems, about the differences in news disseminated on various mediums. Finally, our work is an early attempt, and much future work still remains to be done on understanding the effects of the differences across the different mediums on the news consumers.

Background and Related Work

Comparing online and offline news media: There have been prior research on the coverage of news stories on the offline and online editions of media sources. For instance, (Althaus and Tewksbury 2002) investigated whether readers of printed newspaper and website of NYTimes get different perceptions of political news. (Quandt 2008) compared the distribution of articles per news sections in online news websites with printed newspaper and TV news channels, and found prevalence of news on national politics and economy consistently in all three mediums. Complementary to the above works, we compare different *online* mediums of news consumption, namely news websites, email, Facebook and Twitter, which was not explored before.

Social media and propagation of news: As more and more people are relying on online sources for news, there have been several attempts to understand the flow of news stories on social media (Bhattacharya and Ram 2012; Jisun et al. 2011). There have also been studies on how different factors affect the coverage of news consumed by users. For instance, (Jisun et al. 2011) examined how indirect media exposure in Twitter expands the political diversity of news stories consumed by the users. Our prior work (Chakraborty et al. 2015) investigated whether coverage of trending stories can differ depending on the browsing habits of users. In this paper, we show that the coverage of news stories consumed by users can also vary with the medium which a user chooses to consume news from.

Ideological segregation and filter bubbles: Researchers have investigated the impact of personalized search / recommendations on social media, where individual users get content based on their profiles (e.g., locations), social media neighborhood, past click behaviors, search histories, and so. The concern is whether such exposures increase ideological segregation (Flaxman, Goel, and Rao 2013) and filter bubbles (Pariser 2011). Our study here raises the concern that the social media channels selected by users to receive news may be implicitly filtering stories on certain topics.

Dataset Used

In this work, our objective is to understand how different news stories are covered across different dissemination mediums. We attempt to investigate this question in the context of one of the most popular news media site – NYTimes. Using the NYTimes developer API[1], we collected all news-stories appearing on NYTimes during a period of 6 months, July – December, 2015. Overall, we collected 120,231 distinct news-stories.

Additionally, NYTimes API also returns sets of daily 'Most Viewed', 'Most Emailed', 'Most Shared on Facebook', and 'Most Tweeted' stories, all of which contain 20 news stories at a time.[2] We collected all such stories returned by NYTimes API at 5-minute intervals. Table 1 shows the number of distinct stories that appeared in different sets during this period.

---

¹ developer.nytimes.com/docs/

² For instance, the most emailed stories can be accessed at www.nytimes.com/most-popular-emailed.

Type	No. of distinct stories
All stories published on site	120,231
Most viewed stories	3,008
Most emailed stories	2,667
Most tweeted stories	2,756
Most shared stories on Facebook	2,472

Table 1: **Number of distinct NYTimes news stories that became popular in different dissemination channels during July – December, 2015**

Each NYTimes news-story is published under a *topical category* assigned by the NYTimes site itself. Examples of some topical categories are *Arts*, *Education*, *Politics*, *Sports*, *Science*, and so on. We also gathered the topical annotations for every story using the NYTimes API. In this work, we compute the topical coverage of a set of stories (which have become popular on a particular medium) as the distribution of stories over these topical categories.

Topical Coverage of Socially Shared News

In this section, we compare the topical coverage of stories which are (i) *most popular on the NYTimes website*, i.e., most viewed stories, and (ii) *most socially shared*, which includes most emailed, most shared on Facebook, and most tweeted stories. To better understand the differences in the coverage between these two groups of stories, we also consider a baseline – the overall coverage of *all stories* published online at NYTimes.

Figure 1(a) shows the topical coverages of most viewed, most socially shared and all published stories. Figure 1(b) shows a Venn diagram that represents the (non-)overlap between these three sets of stories. We observe that there are significant fractions of most viewed stories which are not most socially shared, and vice-versa.

To further characterize the non-overlapping stories in Figure 1(b), we look at the topical coverages of stories which are either most viewed or socially most shared, but not both (shown in Figure 1(c)). From Figure 1(a) and Figure 1(c), we observe two interesting trends.

First, the topical coverage of most viewed and most socially shared stories differs significantly from the topical coverage of all published stories, suggesting that online news consumers are expressing a preference for certain topical categories of NYTimes stories over other topical categories, both when viewing and sharing stories online.

Stories on topics of broad interest like 'Sports', 'Business', 'U.S.' and 'World' are less viewed and socially shared, compared to their relative coverage in all published stories. At the same time, stories on topics of narrower (niche) interest like 'Health', 'Fashion', 'Science', and 'Opinion' are more viewed and shared socially than their share of all published stories.

One possible explanation for the differences in the coverage is that users' are more likely to view or socially share topical stories that are uniquely found on NYTimes compared to topical stories that can also be found on other online media sites.
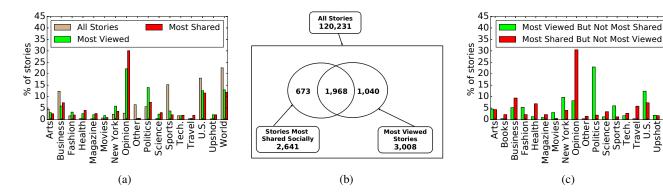
Figure 1: **(a) Topical coverage of most viewed, most socially shared and all published stories at NYTimes, (b) Overlap between stories most viewed, most socially shared and all published stories, (c) Topical coverage of stories that are either most viewed or most socially shared, but not both.**
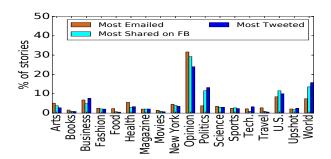

Figure 2: **Topical coverage of stories shared on different channels.**

Figure 3: **Overlap between most shared stories via (i) email, (ii) Facebook, and (iii) Twitter.**

Second, topics like 'Politics' are more viewed on website than shared. On the other hand, 'Opinion' pieces and topics like 'Travel', 'Business' and 'Health' are more shared than viewed. The differences may be explained by the observation that users have a greater incentive to socially share stories on topics of narrower interest that their friends are less likely to find elsewhere.

Dissemination Bias of Social Media Channels

We now drill down to the most socially shared stories, and compare the set of stories that are most shared over three social media channels – (i) most shared over email, (ii) most shared over Facebook, and (iii) most shared over Twitter.

Figure 2 shows the topical coverages of these different sets of stories. We see that stories on certain topics like 'Health', 'Travel' and 'Opinion' are shared more via email than via social media like Facebook or Twitter. On the other hand, stories on 'World' or 'Politics' are more shared on social media than over email.

Figure 3 shows the overlap among the three sets of stories, and we find significant differences in the sets of stories most shared via different sharing channels. To further characterize the difference in the coverages, for every pair of channels, we look at the stories covered by *only one* of them (as shown in Figure 4(a), Figure 4(b) and Figure 4(c)).
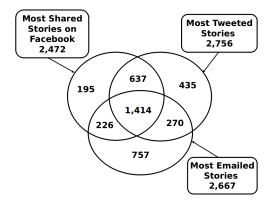
For example, Figure 4(c) shows the topical coverages of stories most shared either on Facebook or Twitter, but not on the other. The comparison between Facebook and Twitter as media channels highlights that the popular stories on Facebook are more 'Opinion' pieces, 'Movies' and 'Arts' stories. On the other hand, higher fraction of news stories about 'Business', 'Technology', 'Politics', and 'World' get shared on Twitter. Interestingly, very similar trends are observed, when comparing coverage of stories disseminated over Email and Facebook. Table 2 shows some example stories which are most shared on one channel but not on another.

These differences in the topical coverage of news stories on different channels are probably due to the differences in the *level of personalization* of the various channels. Email (mostly one-to-one private communications) can be considered a more personal dissemination channel than Facebook (mostly conversations with a group of friends) which in turn is a more personal dissemination channel than Twitter (mostly public communications to everyone). As the medium becomes more public, less of opinionated stories (on narrower topics) and more of factual stories (on topics of broader interest) are shared. For instance, 'Opinion' stories
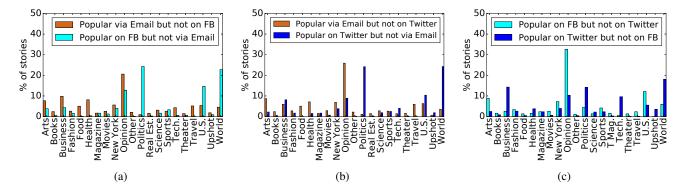
Figure 4: **(a) Topical Coverage of stories popular either via Email or on Facebook, (b) Topical Coverage of stories popular either via Email or on Twitter, (c) Topical Coverage of stories popular either on Facebook or Twitter.**

Type of news story	Title of news story	Topic
Most shared on FB, but not most tweeted	Two Ways of Dealing With Guns	Opinion
Most tweeted, but not most shared on FB	Japan's Parliament Approves Overseas Combat Role for Military	World
Most shared on FB, but not most emailed	Hillary Clinton Says She Cannot Explain Why Previously Undisclosed Emails Turned Up	Politics
Most emailed, but not most shared on FB	Early Detection of Ovarian Cancer May Become Possible	Health
Most tweeted, but not most emailed	14 Years After U.S. Invasion, the Taliban Are Back in Control of Large Parts of Afghanistan	World
Most emailed, but not most tweeted	36 Hours in Cape Town	Travel

Table 2: **Examples of news stories that are popular on one channel but not on other.**

are most shared over email, next over Facebook, and least over Twitter. On the contrary, 'Politics' and 'World' stories are most shared over Twitter and least over email.

Conclusion

Our study here shows that there are substantial differences in the topical coverage of news stories shared over different social media channels. It is unclear whether the online news consumers are aware of such differences in the topical coverages of the different channels from which they consume news. Additionally, designers of various news recommendation systems also use the popularity of news stories in different channels to recommend news stories. Through this study, we want to spread awareness both among the news consumers and the designers of news recommendation systems, about the differences in news disseminated on various channels. In future work, we plan to investigate the extent to which these differences in coverage across the different channels affect the news consumers.

References

Althaus, S. L., and Tewksbury, D. 2002. Agenda setting and the 'new' news patterns of issue importance among readers of the paper and online versions of the new york times. *Communication Research*.

Bhattacharya, D., and Ram, S. 2012. Sharing news articles using 140 characters: A diffusion analysis on twitter. In *ASONAM*.

Chakraborty, A.; Ghosh, S.; Ganguly, N.; and Gummadi, K. 2015. Can trending news stories create coverage bias? on the impact of high content churn in online news media. In *Computation and Journalism Symposium*.

Flaxman, S.; Goel, S.; and Rao, J. M. 2013. Ideological segregation and the effects of social media on news consumption.

Jisun, A.; Cha, M.; Krishna, G.; and Jon, C. 2011. Media landscape in twitter: A world of new conventions and political diversity. In *AAAI ICWSM*.

Mitchell, A.; Gottfried, J.; Kiley, J.; and Matsa, K. E. 2014. Social Media, Political News and Ideology — Pew Research Center. http://www.journalism.org/2014/10/21/section-2-social-media-political-news-and-ideology/.

Pariser, E. 2011. *The filter bubble : what the Internet is hiding from you*. Penguin Press.

Pew, R. C. 2012. In Changing News Landscape, Even Television Is Vulnerable. Trends in News Consumption: 1991 – 2012. http://www.people-press.org/2012/09/27/in-changing-news-landscape-even-television-is-vulnerable/.

Quandt, T. 2008. News on the world wide web? a comparative content analysis of online news in europe and the united states. *Journalism Studies* 9.