2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Analyzing the News Coverage of Personalized Newspapers

Abhijnan Chakraborty
Max Planck Institute for Software Systems, Germany
Indian Institute of Technology Kharagpur, India

Niloy Ganguly
Indian Institute of Technology Kharagpur, India

***Abstract*—Traditionally, news media organizations used to publish only a few editions of the printed newspapers, and all subscribers of a particular edition used to receive the same information broadcasted by the media organization. The advent of personalized news recommendations has completely changed this simpler news landscape. Such recommendations effectively produce numerous personalized editions of a single newspaper, consisting of only the stories recommended to a particular reader. Although prior works have considered news coverage of different newspapers, due to the difficulty of knowing what news is recommended to whom, there has been no prior study to look into the coverage of information in different personalized news editions. Moreover, the evolution of the effects of personalization on recommended news stories is also not explored. In this work, we make the first attempt to investigate these issues. By collecting extensive data from New York Times personalized recommendations, we compare the information coverage in different personalized editions and investigate how they evolve over time. We observe that the coverage of news stories recommended to different readers are considerably different, and these differences further change with time. We believe that our work will be an important addition to the growing literature on algorithmic auditing and transparency.**

***Index Terms*—Personalized News Recommendation, New York Times, Coverage Bias, Algorithmic Auditing**

I. Introduction

Traditionally, media organizations used to publish only a handful of editions of the printed newspaper. For instance, The New York Times (NYTimes) has four print editions: (i) New York, (ii) National, (iii) International Europe, and (iv) International Asia. Different readers can subscribe to one of these print editions depending on either their geographical locations, or their explicit choice for a particular edition. All subscribers of a chosen print edition would receive the same information broadcasted by the media organization.

However, the growth of online news consumption has completely changed this simple news landscape. Today, due to the large number of stories published online, the readers need to rely on news recommendations to find important news stories. To attract and retain the news readers to their websites, media organizations have gradually incorporated personalization in their recommended contents [1]. Such personalized recommendations effectively give rise to numerous personalized editions of a single newspaper, where each edition consists of the set of news stories recommended to a particular reader.

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain

Prior works have analyzed the news coverage of different print newspapers (potentially from different organizations) [2], stories shared on different social channels [3] or featured stories on website frontpages [4]. With the widespread adoption of personalization, researchers have raised the concern of *Filter Bubbles* arising out of the algorithmic efforts to match user interests [5]. Subsequently, Hannak *et al.* [6] and Nguyen *et al.* [7] attempted to measure the effect of personalization on web search, and movie recommendations respectively. Due to the hardness of gathering personalized recommendation data, there has been no prior study to look into the coverage of information in different personalized news editions. Moreover, the evolution of the effect of personalization on recommended news stories is also not considered either.

In this paper, we make the first attempt to answer two unexplored questions in the context of personalized news recommendations:
RQ1. How different are the stories recommended to different readers (i.e., how different are their personalized newspapers)?
RQ2. How are the personalized editions evolving over time?

To answer these questions, we consider the personalized recommendation deployed at the NYTimes website (`nytimes.com`). We simulate several news readers with different reading behaviors, and then automatically collect the news stories recommended by NYTimes to them. We find that the set of stories recommended to different readers are significantly different from each other, where we clearly observe the effect of feedback loop in the recommendations. We also observe how the initial news preferences made by particular readers can take them to very divergent news discourse. We believe that the news readers should be made aware of how their reading habits impact what they get recommended, and works such as ours are important steps to bring in more transparency into the news recommendation scenario.

II. Simulating NYTimes Readers

We simulate different NYTimes readers by using Selenium (`seleniumhq.org`) to automate the Mozilla Firefox browser. For every reader, we create a separate Firefox instance which stores all incoming cookies during execution. On start, each Firefox instance logs in to NYTimes using a particular NYTimes account, and downloads the contents of the 'Recommendations for You' webpage (available at nytimes.com/recommendations). This webpage includes 20 news stories recommended to a particular reader at a particular time. Then, depending on the reading habit of the corresponding reader, the Firefox instance navigates to other article pages. After all the required pages have been navigated, each Firefox instance closes, but the cookies are retained to build the reading history.

To analyze the effects of personalization in the stories being recommended, we simulate two types of readers:
(i) readers who randomly pick stories to read from the stories recommended to them, and
(ii) readers who mimic real Twitter users who share a lot of NYTimes stories on Twitter.

Baseline topical distribution: We further simulate a reader who visits NYTimes every hour, collects the stories recommended to her, but *does not read anything* (i.e., does not click on any of the URLs of the recommended stories). We consider the topical distribution of stories recommended to this reader as *the baseline*, because this set of stories is not influenced by the reading history of the reader.

All browser instances were run at machines in the same $/24$ subnet having the same Linux distribution, and they were started at the same time. All the NYTimes accounts were manually created using the default user profile. We collected the data of random readers for 40 days during June – July, 2016, and for the readers mimicking Twitter users, we collected 12 days of data during July, 2016.

III. Analysis

As mentioned earlier, we consider two types of simulated readers, and compare the set of stories that appeared in their personalized newspapers (i.e., stories recommended to them). We characterize the overlap and differences in the stories recommended to individual readers, and in the topical distribution of the recommended stories.

A. Readers Reading Randomly

These readers visit NYTimes once every hour and read five stories randomly chosen from the set of stories recommended to them. Figure 1(a) shows the overlap in the number of stories recommended to three such readers. Out of all news stories received by any reader during the 40 days period, about 65% stories are recommended to all three readers. However, the rest 35% stories are recommended to one or two of the readers. Thus, these stories will differentiate the information coverage of these readers.

To characterize this difference, Figure 1(b) compares the topical distribution of stories recommended to the three random readers along with the baseline. We can see in Figure 1(b) that the topical distribution of recommended stories are substantially different from the baseline. The baseline has a lot of 'U.S.' stories followed by stories on 'World' and 'Opinion' pieces. Whereas, different random readers received way more 'Sports' or 'Business' stories compared to the baseline.

Among the different readers with random reading habits, all of them get equivalent coverage of stories on topics such as 'U.S.', 'World' or 'Opinion'. However, there are considerable differences in coverage of topics like 'Sports', 'Fashion' or 'Business'. For example, Reader 1 covers more 'Sports' or 'Business' stories than Reader 2, who gets more stories on 'Fashion', 'Food' or 'New York'.

Feedback Loop in Recommendations

The random readers, as described above, exhibit the interesting effects of feedback loops in personalized recommendations. The readers are choosing five stories randomly from the set of recommended stories. The recommendation algorithm is inferring the interests of the readers from the set of stories read by them, and updating the recommended stories in the next iteration. The readers again choose five stories randomly, which influence the next recommendation and so on.

Figure 1(c) shows the topical distribution of stories read by and recommended to one such reader. As evident from Figure 1(c), such random reading implies that these two distributions are similar. As the set of recommended stories depends on the stories recommended earlier, the feedback loop has the potential of trapping a reader into reading stories similar to what she has already read, a notion described as *filter bubbles* [5].

Recall that Figure 1(b) showed that different readers reading random five stories every hour are being directed towards different topics. For example, Reader 1 is directed towards (gets a lot of recommendations on) 'Sports'; whereas, Reader 2 is directed towards 'Fashion'. Hence, it is important to track how the topical distribution of the recommended stories evolve over time.

How the Personalized Newspapers Evolve over Time

To check how the topical distribution of the stories recommended to a reader evolves over time, we compute how the information coverage of the recommended stories are deviating from the baseline. We measure the difference in terms of *Jensen Shannon Divergence* [8] from the baseline distribution, which is explained below.

For two discrete probability distributions P and Q, the Kullback Leibler (KL) divergence [9] from Q to P is defined as the amount of information lost when Q is used to approximate P. More formally,

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

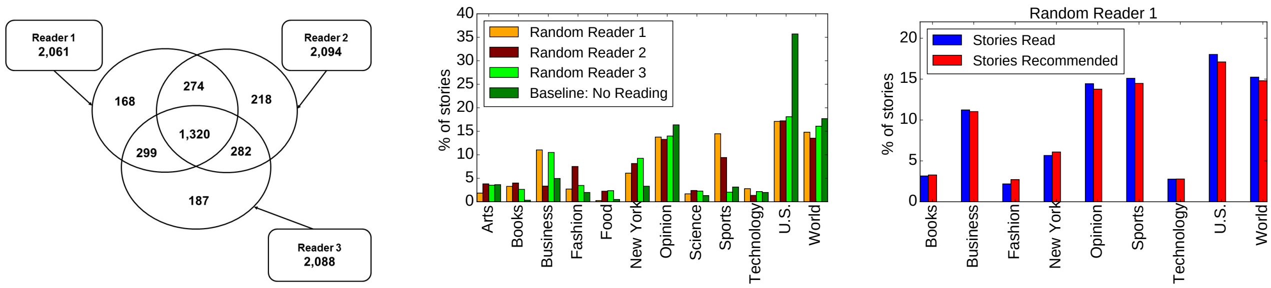
Jensen Shannon (JS) Divergence is the symmetric version of the KL Divergence between P and Q, and is defined as:

$$D_{\mathrm{JS}}(P \parallel Q) = \frac{1}{2} D_{\mathrm{KL}}(P \parallel M) + \frac{1}{2} D_{\mathrm{KL}}(Q \parallel M)$$

where $M = \frac{1}{2}(P + Q)$.

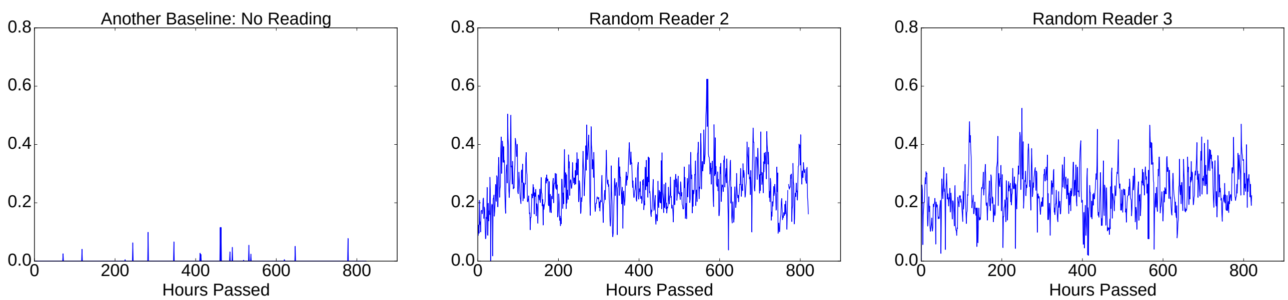
$D_{\mathrm{JS}}(P \parallel Q)$ ranges from 0 to 1, where 0 indicates that the two distributions P and Q are exactly same. Higher the score, two distributions are more different.

In the present context, we compute the topical distribution of stories recommended to a particular reader at a particular hour, and the topical distribution of stories recommended to the reader who does not read anything (i.e., the baseline) during the same hour. Then, the JS Divergence between



(a) (b) (c)

Fig. 1: **(a) Overlap between the news stories recommended to different random readers, (b) topical distributions of news stories recommended to them, (c) topical distributions of news stories read by and recommended to one of these readers. All of these readers visited NYTimes every hour and read 5 stories chosen randomly from the stories recommended at that time.**



(a) (b) (c)

Fig. 2: **How difference from baseline recommendation changed over time for (a) another baseline reader, (b) and (c) two readers with random reading habit.**

these two distributions represents the difference between the recommended stories at that hour.

We first check whether there is any difference between the stories recommended to two readers, both of whom do not read anything. We can see in Figure 2(a) that except a handful of hours, the stories recommended to them at most of the hours are the same. Thus, we can infer that the recommendation process is largely deterministic, and the differences reported for other readers are the *manifestation of personalization based on the readers' reading history*, rather than an effect of some random noise.

Figure 2(b) and Figure 2(c) show the differences between the distribution of stories recommended to two readers and the baseline distribution over time. We can see that the distributions start deviating from the baseline immediately after the readers start reading. The differences between the distributions keep on increasing for next few hours, and then start oscillating around a particular level.

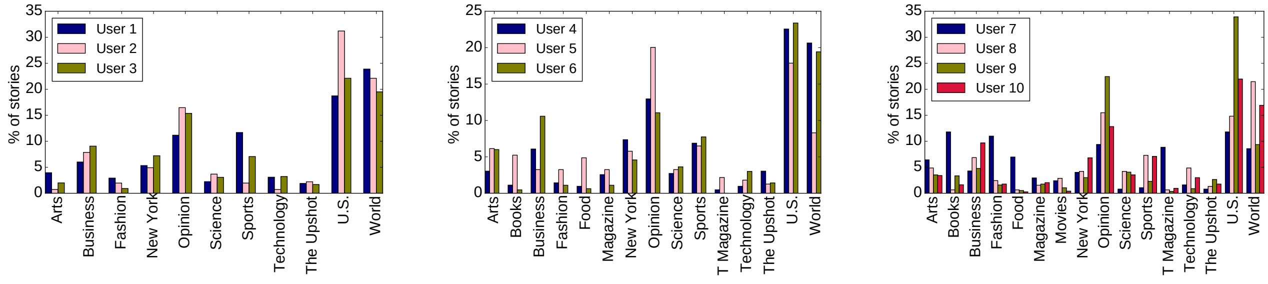
Although the topical differences between the baseline and the stories recommended to the random readers do not increase continuously, the initial random selections determine the level at which the differences stabilize. For example, for Random Reader 2 in Figure 2(b), the average JS Divergence from the baseline is 0.25, whereas for Random Reader 3 in Figure 2(c), the average JS Divergence is slightly lower at 0.2.

B. Readers mimicking the news sharing of Twitter users

The readers considered earlier simulate the random reading habits, which may not reflect the actual reading patterns of NYTimes readers. To analyze the stories recommended to the actual readers of NYTimes, we considered the sharing of a news story by a reader on Twitter as a proxy for her reading the article.

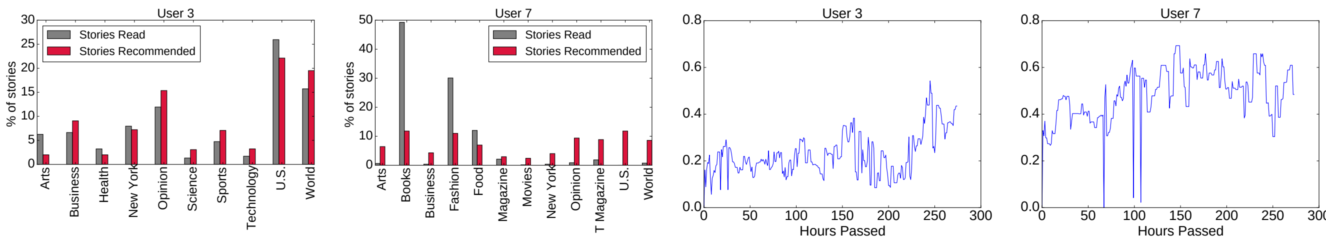
We considered 10 Twitter users who had shared more than 150 NYTimes articles on Twitter during the month of June 2016. Then, we simulated readers who read the stories shared by the corresponding Twitter users. We also maintained the sequence of stories as shared by them. Figure 3 shows the topical distribution of stories recommended to these readers. We can see substantial differences between the coverage of topics by the recommendations to different readers. For example in Figure 3(b), User 6 gets more stories on 'World' or 'Business' compared to User 5, who in turn gets more stories on 'Opinion', 'Food' or 'Fashion'.

To check how the reading habit influences the stories being recommended, we compare the topical distribution of stories read by a reader and the stories recommended to her. Figure 4(a) and Figure 4(b) show the comparisons for two such readers. We observe that for some readers (e.g., User 3 in Figure 4(a)), the recommendation algorithm select stories to closely match the topics of stories read by the readers. This



(a) (b) (c)

Fig. 3: **Topical distribution of stories recommended to different readers mimicking the sharing of NYTimes news stories on Twitter by real NYTimes followers.**



(a) (b) (c) (d)

Fig. 4: **(a) - (b) Topical distributions of news stories read by and recommended to two readers, who mimic the sharing of NYTimes news stories on Twitter. (c) - (d) How the difference from baseline recommendation changed over time for these readers.**

is the main reason for the differences between the topical distributions of stories recommended to different readers as observed in Figure 3.

However, for User 7 in Figure 4(b), it is interesting to observe that there is a significant difference between the distribution of stories read and recommended. There may be two reasons for this divergence:
(i) there may not be many stories on niche topics 'Books' or 'Food' which can be considered for recommendation at a particular time, and
(ii) there may be other stories on some different topics which share common keywords with the stories the user has shared on Twitter.

Finally, Figure 4(c) and Figure 4(d) track how the recommendations are evolving over time for User 3 and User 7. Similar to the earlier two types of readers, the distributions of recommended stories start differing from the baseline in the beginning, and gradually stabilize with time. However, depending on the topical preference exhibited by the stories read by different readers, they would end up with recommendations whose topical coverage may be closer or further from the baseline distribution.

IV. Conclusion

In this paper, we make the first attempt to compare the news coverage of different personalized editions. Our analysis reveals the effect of feedback loop in personalized recommendations where the initial news choices can take individual readers to very different news discourse. We believe that the news readers and recommendation designers should be aware about the impact of personalization and our work is a first step towards that direction.

Acknowledgments

A. Chakraborty is a recipient of Google India PhD Fellowship and Prime Minister's Fellowship Scheme for Doctoral Research, a public-private partnership between Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India and Confederation of Indian Industry (CII).

References

[1] N. Thurman and S. Schifferes, "The future of personalization at news websites: lessons from a longitudinal study," *Journalism Studies*, 2012.
[2] J. R. Lott Jr and K. A. Hassett, "Is newspaper coverage of economic events politically biased?" *Public Choice*, 2014.
[3] A. Chakraborty, S. Ghosh, N. Ganguly, and K. P. Gummadi, "Dissemination biases of social media channels: On the topical coverage of socially shared news." in *AAAI ICWSM*, 2016.
[4] A. Chakraborty, S. Ghosh, N. Ganguly, and K. Gummadi, "Can trending news stories create coverage bias? on the impact of high content churn in online news media," in *Computation and Journalism Symposium*, 2015.
[5] E. Pariser, *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.
[6] A. Hannak, G. Soeller, D. Lazer, A. Mislove, and C. Wilson, "Measuring price discrimination & steering on e-commerce web sites," in *IMC*, 2014.
[7] T. T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, and J. A. Konstan, "Exploring the filter bubble: The effect of using recommender systems on content diversity," in *WWW*, 2014.
[8] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information theory*, vol. 37, 1991.
[9] S. Kullback, *Information theory and statistics*. Courier, 1997.