

# Information Consumption on Social Media: Efficiency, Trust, and Divisiveness

Mahmoudreza Babaei

Max Planck Institute for Software Systems (MPI-SWS), Germany

## ABSTRACT

A growing number of people rely on social media platforms, such as Twitter and Facebook, for their news and information needs [15, 24], where users themselves play a role in selecting the sources from which they consume information, overthrowing traditional journalistic gatekeeping [23]. Since users can just select their information sources, they don't have full control on the content they receive. Moreover, it is very hard to ascertain the quality, relevance, and credibility of information produced by social media users [1, 6, 7], raising interesting questions like: (i) how efficient are users at selecting their information sources?, (ii) how do users perceive the truthfulness of information?, and (iii) how does the consumed information impact users and the society?

## THESIS RESEARCH

In this thesis, we analyse social media users' information consumption along the three dimensions of efficiency [3, 4], trust [2], and divisiveness [5], which we briefly describe next.

### Measuring Efficiency of Users for Selecting Information Sources [3,4]

Task of selecting information sources from potentially tens to hundreds of millions of users poses serious challenges and raises important questions. For example, recent studies have observed that out of fear of missing out on important information users tend to follow too many other users [10]. In the process, they receive a lot of redundant information [3], become overloaded, and effectively miss the information they are interested in [9, 14].

These observations raise questions about how efficient users are at choosing their information sources. In our recent work [3,4], we identified three notions of efficiency for a user: (i) link (i.e., number of sources the user follows), (ii) in-flow (i.e., the amount of (redundant) information she acquires), and (iii) delay (i.e., the delay with which she receives the information). We introduced a computational framework to quantify a user's efficiency, which estimated the optimal set of users from whom the user could have acquired the same pieces of information more efficiently, and compared this optimal set with the user's set of followees.

We observed that Twitter users exhibit sub-optimal efficiency across our three notions of efficiency. Moreover, we showed that this lack of efficiency is a consequence of the triadic closure mechanism by which users typically discover and follow other users on social media platforms. Finally, we developed a heuristic algorithm that enables users to be significantly more efficient at acquiring the same unique pieces of information.

### Utilizing Users' Truth Perceptions to Fact-check News Stories [2]

Having characterized how efficient users are at selecting information sources, we next turn towards examining how do users perceive the truthfulness of the information they consume from these sources. Recently social media sites have been severely criticized for allowing *fake news* stories to spread unchecked on their platforms [8, 13, 12]. To counter the proliferation of fake news, social media sites are relying on their users' perceptions of the truthfulness of news stories to select stories to fact check. However, to date, few studies have focused on understanding how users perceive truth in news stories, or how biases in their perceptions might affect current strategies to detect and label

fake news stories. Moreover, the goal of fact checking is also not clear. In the second part of the thesis, we describe three important goals for fact checking news stories based on users' perceptions of truthfulness of stories:

- **Goal 1: Removing false news stories from circulation:** By using users' truth perception values as a proxy, stories which are flagged by more users as false could be selected for fact checking with a higher probability than true stories.
- **Goal 2: Correcting the misperception of users:** Based on the biases in users' perceptions, stories where users' perceived truth levels differ significantly from ground truth levels (which we define as the Total Perception Bias) could be prioritized for fact checking. Using our metric of Total Perception Bias and combining it with user demographics, we designed an automated method for identifying news stories with high and low perception bias with an accuracy of 82%.
- **Goal 3: Decreasing the disagreement amongst users' perceptions of truth:** For the society to have fruitful debates in the public sphere, it is essential for there to exist a common ground for the different, possibly disagreeing sections of the society. It is desirable to decrease the disagreement amongst users' truth perceptions by prioritizing those stories for fact checking where people disagree most about the truth value of the stories. For achieving the goal, we prioritize stories with high values of disputability (i.e., high variance in users' truth perceptions of the story).

## Identifying and Removing Divisive Content from Social Media [5]

Finally, in the last part of the thesis, we focus on impact of the consumed information on the users and the society. Within our societies, there are many divisive topics for which different subgroups hold opposing ideological positions (e.g., Republicans and Democrats in the US). Social media platforms provide a wide variety of news sources covering this ideological spectrum, yet many users largely limit themselves to news stories, which reinforce their preexisting views. This selective exposure and consumption of divisive information may lead to a more politically fragmented, less cohesive society [16].

To minimize the possibility of social media users getting trapped in 'echo chambers' or 'filter bubbles', as well as consuming just divisive information, prior works have proposed to introduce diversity in the news that users are consuming [20, 22, 11]. Often, such approaches, which highlight the most belief challenging news, increase the chances of users rejecting them, thereby defeating the original purpose [21, 17, 18, 19, 25]. Here, we propose a complementary approach to inject diversity in users' news consumption by highlighting and identifying non-divisive news posts which evoke similar reactions from different readers, irrespective of their own political leanings.

To identify (non) divisive news automatically, we propose a novel class of features of social media posts on Twitter, which we term as audience leaning based features. These features describe the distribution of the political leanings of audience subgroups interacting with a post - namely the retweeters and repliers of a post. Intuitively, most of the retweeters are likely to be supportive of it, while repliers have a higher likelihood of opposing it. We show that our proposed features are well suited to help identify (non) divisive tweets automatically with high accuracy, leading to significantly better performance than can be achieved using previously proposed methods.

## BIOGRAPHY:

I am a fourth year Ph.D. candidate at Max Planck Institute for Software Systems (MPI-SWS) in the Social Computing Research group under the supervision of Prof. Krishna P. Gummadi. Prior to joining MPI-SWS, I did a Masters in Computer engineering from Sharif University in Iran.

My research focuses on information consumption in social media particularly along three dimensions of efficiency, trust, and divisiveness. My research interests lie at the intersection of data mining, social computing and natural language processing.

## OPEN QUESTIONS:

I am eager to participate in the Doctoral Colloquium at CSS Symposium because unlike most other venues, the symposium truly has a large spectrum of experts from multiple disciplines and I would love to take this opportunity to

get feedback from people with different perspectives. Since my thesis goals are interdisciplinary in nature, I believe such varied feedback would help enrich my thesis research and help create an informed storyline.

## REFERENCES

- [1] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne. Finding high-quality content in social media. In *Proceedings of the 1st International Conference on Web Search and Data Mining*, pages 183–194, 2008.
- [2] **M. Babaei**, A. Chakraborty, J. Kulshrestha, E. M. Redmiles, M. Cha, and K. P. Gummadi. Analysing biases in perception of truth in news stories and their implications for fact checking. 2018.
- [3] **M. Babaei**, P. Grabowicz, I. Valera, and M. Gomez-Rodriguez. On the users' efficiency in the twitter information network. In *Proceedings of the 9th International AAAI Conference on Weblogs and Social Media*, 2015.
- [4] **M. Babaei**, P. Grabowicz, I. Valera, K. P. Gummadi, and M. Gomez-Rodriguez. On the efficiency of the information networks in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 83–92. ACM, 2016.
- [5] **M. Babaei**, J. Kulshrestha, A. Chakraborty, F. Benevenuto, K. P. Gummadi, and A. Weller. Purple feed: Identifying high consensus news posts on social media. In *Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics & Society (AIES)*, 2018.
- [6] C. Castillo, M. Mendoza, and B. Poblete. Information credibility on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, pages 675–684, 2011.
- [7] M. Farajtabar, M. Gomez-Rodriguez, N. Du, M. Zamani, H. Zha, and L. Song. Back to the past: Source identification in diffusion networks from partially observed cascades. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*, 2015.
- [8] A. Friggeri, L. A. Adamic, D. Eckles, and J. Cheng. Rumor cascades. In *ICWSM*, 2014.
- [9] M. Gomez-Rodriguez, K. Gummadi, and B. Schoelkopf. Quantifying Information Overload in Social Media and its Impact on Social Contagions. In *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*, pages 170–179, 2014.
- [10] N. Hodas and K. Lerman. How visibility and divided attention constrain social contagion. In *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing*, pages 249–257, 2012.
- [11] J. Keegan. Blue feed, red feed - see liberal facebook and conservative facebook, side by side. <http://graphics.wsj.com/blue-feed-red-feed/>, 2017.
- [12] S. Kwon, M. Cha, and K. Jung. Rumor detection over varying time windows. *PLoS one*, 12(1):e0168344, 2017.
- [13] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang. Aspects of rumor spreading on a microblog network. In *Proceedings of the SocInfo*, 2013.
- [14] K. Lerman and T. Hogg. Leveraging position bias to improve peer recommendation. *PLoS One*, 9(6), 2014.
- [15] J. Lichterman. New Pew data: More Americans are getting news on Facebook and Twitter, 2010. <http://tinyurl.com/News-on-Social-Media>.
- [16] Z. Liu and I. Weber. Is twitter a public sphere for online conflicts? a cross-ideological and cross-hierarchical look. In *International Conference on Social Informatics*, pages 336–347. Springer, 2014.
- [17] C. G. Lord and L. Ross. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, pages 2098–2109, 1979.

- [18] A. G. Miller, J. W. McHoskey, C. M. Bane, and T. G. Dowd. The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *Journal of Personality and Social Psychology*, 64(4):561, 1993.
- [19] G. D. Munro and P. H. Ditto. Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, 23(6):636–653, 1997.
- [20] S. Munson, S. Lee, and P. Resnick. Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, pages 419–428, Boston, USA, 2013. AAAI press.
- [21] S. A. Munson and P. Resnick. Presenting diverse political opinions: How and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10*, pages 1457–1466, New York, NY, USA, 2010. ACM.
- [22] S. Park, S. Kang, S. Chung, and J. Song. Newscube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 443–452. ACM, 2009.
- [23] P. J. Shoemaker, T. P. Vos, and S. D. Reese. Journalists as gatekeepers. *The handbook of journalism studies*, 73, 2009.
- [24] J. Teevan, D. Ramage, and M. R. Morris. #twittersearch: a comparison of microblog search and web search. In *WSDM '11 Proceedings of the fourth ACM international conference on Web search and data mining*, pages 35–44, Hong Kong, China, 2011. ACM.
- [25] C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 2018.