FACEBOOK ALGRORITHMS AND USER POLARIZATION

A Research Paper submitted to the Department of Computer Science
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By

Zach White

April 29, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISOR
Raymond S. Pettit, Department of Computer Science
Yixin Sun, Department of Computer Science

Facebook algorithms and user polarization

A literature review of user interaction with News Feed algorithms

Zachary White
Computer science
University of Virginia
Charlottesville, VA 22903
zsw2wf@virginia.edu

ABSTRACT

In social media applications like Facebook, algorithms control the vast majority of what a user sees on their main content page. Facebook has enlisted the help of artificial intelligence to help manifest user engagement and community on their platform through personalization and networking. These AI have proven immensely valuable to advertisers who can better target groups of people through Social Networking Marketing [7]. The same algorithms that recommend products accurately have also been shown to lead users into more polarized mindsets [3], leading to tangible tensions and effects in reality.

This technical report will be an analysis of current research into how Facebook users and the platform's news feed algorithms interact and can lead to polarization. Despite the fact that bias has existed and fostered socially in all of human civilization, research in this area is particularly important due to the novel role of algorithms in the information filtration process. Current research has studied various aspects of how echo chambers are created, grown, and mitigated [3], as well as potential methods of altering the algorithm's role in platform content distribution. It is important for users and Facebook itself to be aware of how an account's feed is affected by user actions, in order to maintain an understanding of the role of artificial intelligence in the online community.

Facebook and Algorithms

The Facebook business model consists of a few major goals: optimizing individual user engagement, expanding the networks, and targeted advertising. *Algorithms*, used interchangeable in this paper with *artificial intelligence*, are used by the platform to achieve these goals. These algorithms are behind News Feed content selection, friend and group recommendations, and the advertisements that eventually seem to read your mind. Most of the ways a user can interact with the application are interpreted into a user response that is evaluated by the algorithm before the next iteration of content display. This distinguishes these algorithms from being just simple recipes; these programs learn about the user and change approach to try to best maximize engagement, expand networking, or land an advertisement into the right brain.

Social Networking Marketing. Facebook creates most of the company revenue through advertising, which allows all users to consistently use the application for free. As television and newspapers before it, this new media distribution platform created a new and effective method for companies and products to reach broad audiences. But unlike the older forms of media, the *social* theme of the application allows advertisers access to user data and friend networks. This heightened interactivity model through which advertising is now flourishing is defined as Social Networking Marketing [7]. This approach takes advantage of a user's network and ability to share information to reach wider and more relevant audiences, including Facebook Friends and Groups. This alongside instant communication brought a massive media advertisement market to supplement the growth of online platforms.

Perhaps the largest advantage of Social Networking Marketing is the ability for platforms to collect data from user behavior. This data includes specific user interactions with media content, as well as Friends, groups, interests, and more. On the macro level, this allows for the platform to more effectively target similar users with relevant advertisements. This ability to effectively match the advertisement with the consumer is a crucial point of competition among social media companies, and advertisers

will invest more into a certain application's ads if that application has a higher success rate in customer feedback.

This same data is used by News Feed and recommendation algorithms to decide what next content or suggestion to deploy to the user. As further laid out in one of the articles under review [1], the algorithms predict which content will provide the most positive user response, and greedily approach this goal at every step. However, the problem begins to arise when these algorithms are unable to mitigate human tendencies, which manifest in heightened engagement with sources and content that affirm and agree with the user's existing beliefs and viewpoints [2].

Bias Propagation

The more an AI learns about a user, recommender algorithms have been shown to decrease the variety of what is shown, as they figure out main topics and characteristics of a user. The intensity of this effect varies through platforms, but a study at the McIntire School at UVA [3] observing the effects of these "echo chambers" across multiple social media platforms showed Facebook users in particular to be driven to more partisan sources. While this may seem harmless on the individual level, these affirmative and polarized bubbles have culminated in real-life, extreme global effects, like Trump's 2016 victory [4], the genocide in Myanmar against Muslim Rohingya citizens, and a surge in conspiracy theories and their contingent communities.

Related work to this literature review observes the power that these echo chambers have over even fact-checking mechanisms that are in place to try and combat misinformation [5]. Regardless of political viewpoint or orientation, users had experienced echo chamber effects from iterative algorithmic interactions that were a more significant consideration than a simple fact-check flag.

Because these algorithms are the trade secret of these applications, the best research can do from the outside is a black-boxed approach. The articles in this literature review center around the user interaction interface input to this black-box artificial intelligence to study and mitigate the polarizing effect that pure algorithmic control can bring over time. The first article, authored by Lorian Leong, explores how new users in Myanmar understood and navigated their interaction with the Facebook algorithm. The second, from L. Elisa Celis et. al, proposes a straightforward boundary methodology to mitigate the unrestrained polarizing tendencies of an iterative algorithm approach.

Facebook Users in Myanmar

Lorian Leong writes about a study conducted in Myanmar around the year 2016 [6], consisting of individual interviews of various Myanmarese citizens and their personal navigation of the application and News Feed.

In 2011 the long-standing military regime in Myanmar fell, and it was followed by a surge of Myanmarese citizens gaining access to the internet and its information. Smartphones and Facebook were the main point of access through which everyone saw their news and friend activity. The stark contrast to the restrictive, expensive internet of the previous era caused Facebook to be widely accepted with very minimal criticism. Networks and usage grew quickly, and this provided an opportunity for these researchers to observe how users interact with a new algorithm presence in their everyday lives.

Specifically, Leong discusses the idea of "domestication" of algorithm technology through Facebook; that is, how does this phenomenon get accepted into the weave of the existing structures of everyday life? The algorithm determines what users see and provides suggestions – how users react and provide data also influences how this algorithmic presence will settle. The author provides a framework for the audience to use for this social media technology, centering around the double meaning of technologies in different cultures and societies. "Double articulation" of this Facebook application recognizes it as two things: a tangible application that provided access to information and communication, and a symbolic representation of press freedom and transparency in Myanmar.

This is important because while Facebook centers its application around Friends and Groups, this symbolic freedom of information for Myanmarese internet users changes the dynamic of how Facebook is primarily used in the country. In one of the first findings of this exploratory study featuring Myanmarese informants [6], it was found to be common practice to add new accounts and connections specifically for their information sharing, despite not knowing the people in real life. This was sometimes layered onto social motives as well, as more Friend connections makes an account appear more connected, informed, and social to others. This interpretive flexibility by Myanmarese users shifts Facebook's use much more heavily toward the information realm, despite Facebook's intentions of personal connections.

The author connects these interactions to interfacing with the News Feed algorithm: most users understand that adding connections for information will cause the algorithm to display this information on their Feed, *because* of the user's

actions. But a significant disconnect between reality and the algorithm's lens, as Leong points out, is that connections are not distinguished by "type" of connection, being a real-life friend or an anonymous informant. These connections are only prioritized through the iterative actions in the user feedback loop that the algorithm sustains. More interactions with news connections will "teach" the algorithm that these connections are more valuable to the user, and will therefore appear more often.

While this is an example of the power users and culture can have over the adaptation and domestication of algorithm technology, users and culture are also susceptible to the power of the algorithm. Notoriously beginning soon after this study was conducted, over 700,000 Rohingya Muslims in Myanmar were forced from their home country in a mass exodus following the surge in anti-Islam propaganda on Facebook [8]. Military and government officials realized the power that could be manifested through Facebook's platform and the social spread of information, and they posed spreading hate messages into the newly free access of information to Myanmarese citizens. Tension and violence caused the mass exodus and awoke the world to the dangers of this new and unregulated social platform.

This unconstrained algorithm behavior signals the need for mitigating the extremes and negatives of personalization online. The next article under review [1], authored by researchers from Yale, IIT Kanpur in India, and EPFL in Switzerland, explores "a dashboard for controlling polarization in personalization".

Controlling Algorithmic Polarization

As the authors to this second article [1] acknowledge, there exists a need for "regulatory mechanisms" to control the output of a News Feed. One proposed way of approaching this is the "constrained algorithm" approach, which this demo represents on a website[1] with percentage limits and a news search bar. This idea limits the effects of otherwise "greedy" algorithms, which attempt to maximize the short-term reward rate at every individual step.

This graphic included from the article itself (Figure 1) illustrates the extra step included in each iteration of a News Feed display. The "constraints" implemented on the search results are shown atop in the user-accessible percentage graph, where limits can be chosen for conservative and liberal balancing. These represent how much of each viewpoint is allowed into the feed, and are satisfied by every News Feed display within the "balanced" output.

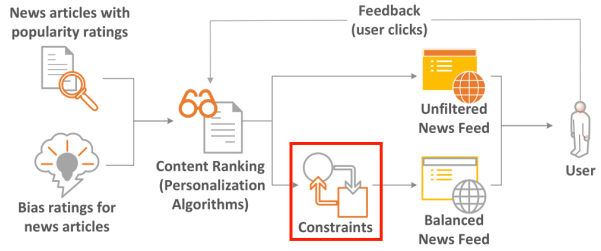

Figure 1: From the article under review [1], this flowchart illustrates the insertion of constraints into the previously unfiltered feedback loop.

In this demo, each click by the user takes them to an external news story, but also records as a "reward" for that instance of content. The algorithm updates its information on the user: specifically, the predicted success rate of a certain topic or viewpoint. Therefore, in the next iteration, the greedy algorithm displays more content aligned with the viewpoint of that article.

This happens for both the balanced and unconstrained feeds on this demo website, as the constraints are not meant to directly oppose personalization. Most articles displayed still align with the political views of the user, but the percentage limits offer just enough contrasting viewpoint to theoretically disrupt the "echo chambers" otherwise fostered online [3]. These limits are considered by the constrained algorithm at every step, and any inclusion that pushes the quantified polarization of the News Feed above a limit is terminated until the constraints are satisfied. Otherwise, much of the behavior of the algorithms is the same.

In a graph that the authors include (Figure 2 below), they demonstrate how over thousands of iterations, their "constrained-greedy" algorithm best balances personalization with these constraints, maintaining a changing weighted probability for content but underneath the specified maximum. A measurement coined "regret" determines how much potential reward for the algorithm was lost in choosing a particular piece of content, as compared to the maximum possible. This would inevitably increase as fewer polarizing articles are selected to display, but the underlying "greedy" operations of the algorithm at every step would still manage to optimize the reward within these new constraints.

[1] https://theory.epfl.ch/bias/personalized_news/fast/diversity.php

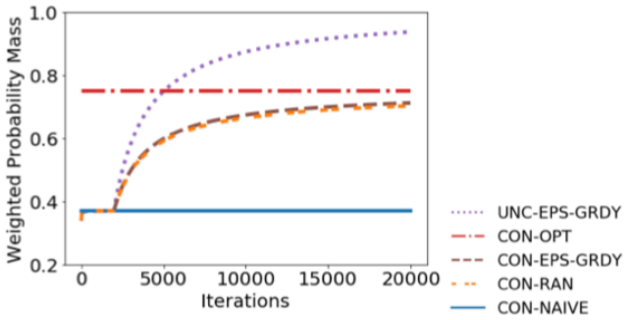

Figure 2: the authors of [1] display how the constrained-greedy algorithm best aligns with the goal constraints over time, as compared to unconstrained algorithms and other, more naïve implementations.

In summary, this demonstration of constraining algorithms offers a direct visualization of how News Feeds would be affected by regulatory measures. This simple measure is hugely instrumental in the overall fight against the development of opinion bubbles online.

Conclusion

Facebook has opened a new, incredibly efficient avenue for everyday humans with internet to access networks and information. This hugely significant role of the application puts their algorithms at the center of user interaction and content moderation, which are two important facets of understanding needed to effectively reduce bias propagation on these social platforms. Lorain Leong in Myanmar gives us an understanding of the client-side perception, and illustrates through double articulation of technology how the symbolic meaning of a social media can drastically affect its position and acceptance into a culture. Elisa Celis et al. create an interface to explore and model how News Feeds can be altered at the algorithm level with content distribution constraints.

Leong, having written this research months after the informant interviews were conducted, acknowledges the dynamic landscape and how Facebook may have changed since the information was collected, including major events like the extended Rohingya Muslim genocide in the country. Observing from the client-side also brings memory and bias limitations within the response and data themselves, and she suggests future research with actual user data to mitigate this.

The five authors of the constrained algorithm paper also summarize the potential complications of incorporation into real social platforms. The model uses data from *AllSides* to judge polarization weight, and the authors hint at further progress in the language processing realm that could advance how polarization is analyzed and weighted on applications. Different methods of evaluating media's polarization would have to be modeled as well. It is important to maintain an understanding of content and algorithms on Facebook, as our societies will only become more dependent on artificial intelligence to run the world around us. Exploring how to interact with and constrain these forces will remain a largely relevant field of research for the foreseeable future.

REFERENCES

[1] Celis, L. E., Kapoor, S., Salehi, F., Keswani, V., Vishnoi, N. K., & Weng, P. (2019). A dashboard for controlling polarization in personalization. AI Communications, 32(1), 77–89. https://doi.org/10.3233/AIC-180606

[2] Usher-Layser, N. (2016). Newsfeed: Facebook, filtering, and news consumption. *Phi Kappa Phi Forum, 96(3), 18-21.*

[3] Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption. MIS Quarterly, 44(4), 1619–1649. https://doi.org/10.25300/MISQ/2020/16371

[4] Baer, D. (2016, November 09). The 'filter bubble' explains why Trump won and you didn't see it coming. New York Magazine. Retrieved April 25, 2021, from https://www.thecut.com/2016/11/how-facebook-and-the-filter-bubble-pushed-trump-to-victory.html

[5] Ardévol-Abreu, Alberto; Delponti, Patricia; Rodriguez-Wangüemert, Carmen (2020). "Intentional or inadvertent fake news sharing? Fact-checking warnings and users' interaction with social media content". Profesional de la información, v. 29, n. 5, e290507. http://doi.org/10.3145/epi.2020.sep07

[6] Leong, L. (2020). Domesticating algorithms: An exploratory study of Facebook users in Myanmar. Information Society, 36(2), 97–108. https://doi.org/10.1080/01972243.2019.1709930

[7] Liao, S.-H., Hsian, P.-Y., & Wu, G.-L. (2014). Mining User Knowledge for Investigating the Facebook Business Model: The Case of Taiwan Users. Applied Artificial Intelligence, 28(7), 712–736. https://doi.org/10.1080/08839514.2014.927695

[8] Yue, N. (2020). The "weaponization" of facebook in myanmar: case for corporate criminal liability. Hastings Law Journal, 71(3), 813-844.