Exposure to the Invisible: Algorithm Awareness from the Individual to the Collective

Motahhare Eslami
University of Illinois at Urbana-Champaign
Computer Science Department & Coordinated Science Laboratory

INTRODUCTION

We should interrogate the architecture of cyberspace as we interrogate the code of Congress. —Lawrence Lessig

Today, algorithms exert great power in the curation of everyday online content by prioritizing, classifying, associating, and filtering information. In doing so, they can exert power to shape the users' experience and even their perception of the world [1]. While such powerful algorithms are omnipresent online, they are rarely highlighted in the interface, leaving users unaware of their presence. Although the lack of users' awareness about these hidden processes can sometimes indicate a successful design, in some cases this invisibility can cause problems. A clear example is in Morris's study of social network use by new mothers. She questioned the common complaint that new mothers primarily posted photos of their babies. She found that Facebook News Feed created this misperception because it prioritizes posts that receive likes and comments – photos of babies often received attention from a large audience. Because users lack knowledge about the News Feed algorithm, they may have an inaccurate picture of how their and others' actions influence their personal feeds, resulted in actions such as blocking new mothers on Facebook [2]. Feed curation algorithms are some of the many algorithms that can influence people's experience in the online world. The increasing prevalence of these invisible algorithms coupled with such substantial influences raises many questions:

RQ1-Algorithm Awareness: How knowledgeable are users about these algorithms, and how aware should they be of their existence and functionality?

RQ2-Awareness Effects: If we can provide insight to users about an invisible algorithm's existence or functionality, how will this insight affect their interaction experiences?

RQ3-Algorithm Bias: How can we detect whether algorithms have biases affecting users' experiences?

Here, I present innovative approaches that my colleagues (from Computer Science department and Coordinated Science Laboratory) and I have done to answer the above questions. These approaches, while being unique, lead us to promising outcomes.

ALGORITHM AWARENESS

To answer RQ1, we explore users' awareness and perception of the Facebook News Feed curation algorithm as an example in our recent work [3]. We interviewed 40 Facebook users and asked them to imagine they had a 'friend,' Sarah, and she shared a public story visible on her wall to all her friends. We then asked them whether this story would appear in their own News Feed. Surprisingly, we discovered that more than half (62.5%) were not aware that News Feed hid stories. They believed every single story from their followed pages and friends appeared in their News feed.

To assist us in the interviews, we then presented *FeedVis* to the participants, a tool I have developed to disclose what we call "**the algorithm outputs**": the differences in users' News Feeds when they have been curated by the algorithm and when they have not. Figure 1 shows the first view of FeedVis, *Content View*, consisting of two columns: The right column, 'Shown stories,' displayed only the stories shown on the user's News Feed. The left column, 'All stories,' showed every story from all the users' friends and all the pages the user follows. We distinguished these stories based on being shown or not in the user's News Feed by different colors. To create this column, we extracted all viewable stories from the page of each of the participant's friends and checked the information provided by Facebook API to determine their color. Many of the participants who were unaware of News Feed algorithmic curation were initially surprised by how long the all stories' column was in comparison to the 'shown stories' column.

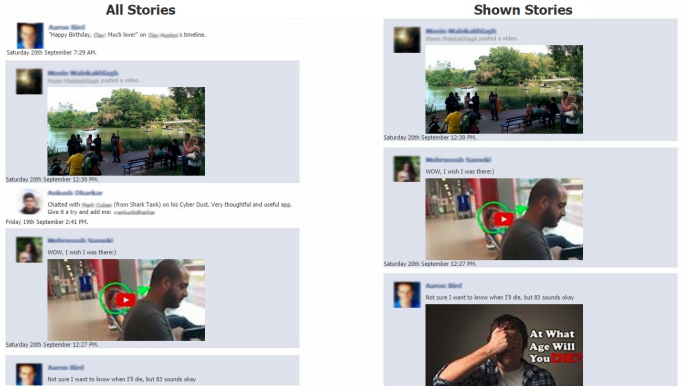

Figure 1: Content View. Shown stories are in blue, while the hidden stories are in white.

We provided another view named *Friend View* (Figure 2), to reveal social patterns by helping the users understand whose stories appeared and whose were hidden in their News Feed. This view divided the users' friends into three categories based on the proportion of each friend's stories that had appeared in the users News Feed during the previous week: 'rarely shown', 'sometimes shown', and 'mostly shown' friends. Observing the algorithm outputs also surprised many participants by revealing misperceptions about their friends whose stories were rarely shown. Importantly, some participants disclosed that they had previously made inferences about their personal relationships based on the algorithm outputs in Facebook News Feed. For instance, some participants mistakenly believed that their friends intentionally chose not to show them stories because they were not close enough. They were surprised to learn via FeedVis that those hidden stories were removed by Facebook and not their friends: "*I have never seen her post anything!, and I always assumed that I wasn't really that close to that person!*". Along with surprise, many participants expressed dissatisfaction when missing stories were revealed to them on FeedVis particularly when close friends and family were not shown. Despite these initial reactions, more than half of the participants became more satisfied with the algorithm over the course of the study. Even as they first scrolled down the Content View, many mentioned that they began to understand why Facebook hid some stories from them.

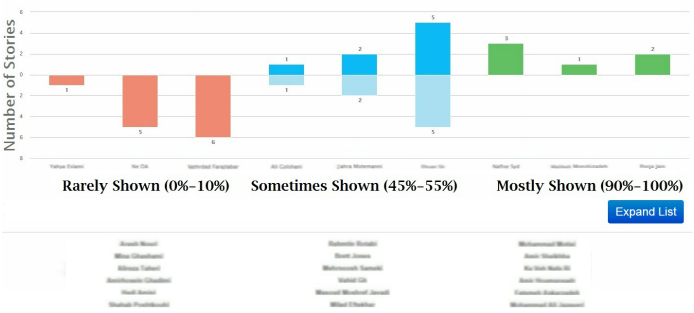

Figure 2: Friend View. The number of the shown and hidden stories are displayed above and below the x-axis, respectively.

After exploring the algorithm outputs, we wanted to gauge participants' desire to change those outputs. Therefore, we created two new views that invited participants to 'tweak' the algorithm. One allowed for adjustment based on authorship of stories, and the other invited manual filtering based on the content of stories. Participants moved their friends to another category around three times more than the stories. This high rate of change demonstrates that the algorithm is not effectively capturing the strong feelings participants had about which friends should appear in their News Feed. These findings suggest that while filtering is both generally needed and appreciated, a lack of awareness of the existence of this process results in concern and dissatisfaction.

AWARENESS EFFECTS

To address RQ2 that how the gained insight into an algorithm would affect a user's interaction experience, we contacted participants via e-mail two to six months after conducting the study. We asked them whether participation in the study resulted in any changes in their Facebook usage behavior. Interestingly, most of the participants (83%) reported changes in their behavior due to participation in our study. Most of them started to manipulate what appeared on their News Feed by using the News Feed adjustment options Facebook provided such as 'most recent' and 'top stories' options. In addition, some changed their interaction behavior with some of their friends, assuming that this interaction change would result in a change to their News Feed. For example, some started to be "*more selective about clicking 'like' because it will have consequences on what [they] see/don't see in the future.*". There were participants who began to play around with Facebook and discuss with their friends on ways to streamline what they were receiving in their News Feed. A few reported that they used Facebook less than they had before the study since the algorithm broke their trust between them and Facebook.

We also asked the participants whether participation in our study affected their satisfaction with News Feed. The majority (80%) reported a similar or higher satisfaction level with News Feed after the study. Many discussed how they felt more knowledgeable about the algorithm. For instance, one previously unaware participant of the algorithm's presence stated that becoming aware of the algorithm's existence resulted in less dissatisfaction when her stories did not receive enough attention from others: *"Because I know now that not everything I post everyone else will see, I feel less snubbed when I make posts that get minimal or no response. It feels less personal."*. Finding such substantial effects suggests that algorithms must be considered more than simply a way to manage information – but rather a way to offer users agency, control, and a deeper relationship with the platform itself.

ALGORITHM BIAS

To find out whether an algorithm might have biases that influence users' experiences (RQ3), we designed a study in which we presented a user with different algorithms' results for the same process. Then, we asked the user to interact with those results to see how differences in results might bias user's behavior. To this end, we considered friend grouping in social networks. In our paper [4], we proposed a tool that applies three clustering algorithms on a Facebook friendship network. These algorithms used the same input (network structure) to create groups of friends automatically. In a series of interviews, we asked 18 Facebook users to evaluate and modify the groupings generated by each algorithm, and create their final desired grouping. Surprisingly, we discovered a 14% difference on average between the final desired groupings the same user created after modifying the groupings.

To understand possible causes of this difference, we correlated each desired group resulting from the modification of an algorithm's results with other desired groups from other algorithms. By talking with participants, we found that this difference was caused mainly by 'following what algorithms create'. Some participants stated that if an algorithm did not find a specific group, they would not create it themselves. For example, one of the participants mentioned that one algorithm put her 'church' friends in a separate group. If she had manually created groups, she would not have considered a 'church' group, but the group made sense, she liked it, and kept it. Such examples demonstrate that automating the friend grouping process influences users to follow what algorithms seed. This is a case where the algorithm biased the user in a positive way, however, negative biases may also be possible. For example, since these groupings are often used for selectively sharing in social media, this bias might cause privacy issues.

Our next major challenge is to generalize this approach in other domains by giving users different versions of an algorithm (including no algorithm), and compare the results to see whether there is a bias or not. However, how can we characterize the bias or power of an algorithm in larger scales, specifically when most of the algorithms are hidden behind walls of intellectual property? Our plan to tackle this challenge is *'collective audit'*; an approach where a large number of regular users help researchers collaboratively to find misleading biases created by invisible algorithms. We note that a crowdsourcing platform is an appropriate candidate to provide enough data and analyze the possible biases created by algorithms. Other alternatives may include recruiting volunteers who want to become more aware of algorithms, algorithm functionality and algorithm biases. We know that such users exist due to number of requests we have received to make FeedVis public. This interest suggests the importance of understanding algorithm awareness in good system design.

REFERENCES

1. Diakopoulos, N. Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Tow Center for Digital Journalism (2014).
2. Morris, M. R. Social networking site use by mothers of young children. CSCW 2014.
3. Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Voung, A., Karahalios, K. , Hamilton, K. and Sandvig, C. I always assumed that I wasn't really that close to [her]: Reasoning about invisible algorithms in the news feed, CHI 2015.
4. Eslami, M., Aleyasen, A., Zilouchian Moghadam, R., and Karahalios, K. Friend Grouping Algorithms for Online Social Networks: preference, bias, and implications, SocInfo 2014.