Quantifying User Agency vs. Algorithmic Influence in Movie Recommender Systems

Marco Conti
University of Illinois
Chicago, Illinois, USA

Sai Shridhar Balamurali
University of Illinois
Chicago, Illinois, USA

Nitheesh Mannava
University of Illinois
Chicago, Illinois, USA

ABSTRACT

This study examines the impact of user choices and algorithmic recommendations on the formation of filter bubbles in movie recommendation systems. We conducted an extensive literature review, presenting the formal definition of filter bubbles and the most critical studies on the topic. We then perform a simulation study using a collaborative filtering recommendation system trained on the MovieLens dataset to evaluate the impact of algorithmic recommendations and user choices on narrowing content exposure. We also propose a method for computing the recommendation system's average treatment effect on the recommendation set's homogeneity.

1 INTRODUCTION

Recommender systems have become integral to our digital experiences. These systems leverage user data and algorithms to provide personalized suggestions to enhance the user experience. While they offer convenience, concerns have been raised about "filter bubbles" where users' choices are isolated from diverse perspectives, potentially reinforcing biases and limiting exposure to new ideas [11]. As the influence of recommender systems grows, a critical question emerges: Are the recommendation algorithms truly responsible for these filter bubbles, or are the people to blame?

A filter bubble has ethical and economic impacts, as a user may get bored of the lack of diversity, leaving the platform. Our research focuses on movie streaming platforms and aims to quantify user agency versus algorithmic influence in filter bubble formation to provide a strategy that promotes user satisfaction, content diversity, and economic returns.

2 RELATED WORK

Eli Pariser first introduced the term "filter bubble" in his book *The Filter Bubble: What the Internet Is Hiding from You* [**?**]. He described how popular social media platforms and search engines employ algorithms that display content based on users' existing behaviours and beliefs, effectively creating a "bubble" around each individual. These filter bubbles have a feedback loop that reinforces preexisting views, leading to isolation, radicalization, and a loss of creativity. One major problem with filter bubbles is their invisibility, which makes users unaware of their isolated condition.

One of the consequences of these filter bubbles is their potential to give rise to or sustain an echo chamber. According to the literature review by Argueda et al. [1], an echo chamber is a situation of limited exposure or confinement to new ideas or content, influenced either by media supply or by the user's own preferences. By contrast, a filter bubble is an echo chamber created by a ranking algorithm without any active choice by the user. People themselves may be blamed for filter bubble formation: they seek information that aligns with their views — whether or not it is factually valid — getting trapped in an echo chamber where opposing views are excluded or minimized.

Filter bubbles are a complex phenomenon explored by numerous researchers, sometimes with contradictory results. Nguyen et al. [10] were among the first to study the presence and effects of filter bubbles in movie recommendation systems. By analyzing the MovieLens dataset, they found that the narrowing effect of the recommendation system is statistically significant but small in value and that the population that follows the recommendations reports an overall more positive experience. They developed a metric to assess content diversity in the movie domain based on the Tag Genome, a system of tags and relevance scores, designed by Vig et al. [12] to facilitate user exploration and integrated into MovieLens.

This is one of several studies suggesting that the effect of filter bubbles is not as severe as initially thought. In "Eli Pariser is wrong," Linden et al. [7] argue that recommender systems are essential for improving serendipitous encounters with new items, as users cannot discover an item they are unaware of. Another study by Bakshy et al. [3] demonstrated that individual choice plays a more vital role than Facebook's ranking algorithm in limiting exposure to cross-cutting content. However, these studies have limitations. For instance, Bakshy's study examines an already polarized audience, typically less susceptible to algorithmic influences. This is why further research is necessary to validate or challenge these findings across different scenarios and with varying assumptions.

As an example of an opposing result, we reference the research by Chaney et al. [2]. They ran simulations and measured, through the Jaccard index, the amount of homogeneity in collaborative filtering recommendation systems, showing that they homogenize user behaviour beyond what is necessary to increase utility. Additionally, they draw parallels with the explore/exploit paradigm and warn against using confounded data when selecting recommender systems. For example, the MovieLens dataset is biased toward collaborative filtering algorithms.

Another study is by Lunardi et al. [9], who introduce a metric for homogeneity, the homogeneity level (HL), study filter bubble formation, and propose diversification to reduce this effect. However, their study claims that the reduction of homogeneity due to diversification is not statistically significant. Liu et al. [8] developed an innovative news recommendation system with a transparent approach to making users aware of the presence of a filter bubble. Jiang et al. [6] introduced a novel approach to distinguish between an echo chamber and filter bubble effects. They approximate user interests, showing they tend to degenerate under mild conditions. They also conducted simulation experiments on degeneration in various recommender systems and proposed continuous exploration and an expanding candidate pool as the only solutions to reduce this effect.

In their work, "Causal Inference in Recommender Systems: A Survey and Future Directions," Gao et al. [4] provide a survey of causal investigations within recommender systems. Notable related applications of causal inference to filter bubbles include the work of Wang et al. [13], who apply the backdoor adjustment method to alleviate the filter bubble effect, and Xu et al. [14], who employ counterfactual reasoning to mitigate the echo chamber effect. To the best of our knowledge, no significant studies have analyzed the role of algorithms in filter bubble formation using causal inference.

3 FORMAL PROBLEM DESCRIPTION

A filter bubble occurs when algorithms automatically narrow the range of recommended content, creating self-reinforcing patterns or feedback loops. The homogeneity of the algorithm-recommended items can measure a filter bubble. We aim to quantify the individual effects of the extent to which the user's choices and the algorithm's recommendations influence the creation of a filter bubble. The effect of the recommendation system on filter bubble formation can be viewed as the average causal effect of the algorithm on the homogeneity of the results. For this task, we identified the structural causal model shown in Figure 1.

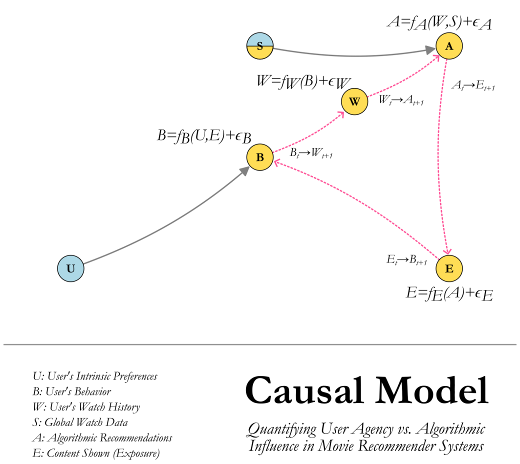

Figure 1: Structural Causal Model of User Agency and Algorithmic Influence in Recommender Systems.

4 DATA DESCRIPTION AND SIMULATION

This project is based on the MovieLens 32M dataset [5]. MovieLens 32M is a widely used benchmark dataset in recommender systems research, containing 32 million ratings and two million tag applications for 87,585 movies by 200,948 users. It also includes the tag genome, where each movie is rated based on attributes that describe the film. The dataset is gathered from MovieLens, a collaborative filtering-based website for movie recommendations.

Additionally, we use a simulation to gather supplementary data. Specifically, we developed a collaborative filtering recommender system, trained it on the MovieLens dataset, and ran two simulations with two different types of users:

- A random user without a specific preference.
- A polarized user who prefers a specific type of film, in this case, crime/thriller.

Both users pick randomly from the recommended items, but the polarized user only positively rates the movies that align with their taste. For simplicity, we assume both users' preferences are static over time. This approach allows us to observe the effect of the recommendation system on the variation of the proposed items, with an intervention on user preferences. In the future, we plan to conduct further simulations with other types of users, random items introduced into the recommended set, or by relaxing the static preference constraint.

5 DESCRIPTION OF INITIAL SOLUTIONS

To analyze the simulation results, we must formally define a filter bubble. Following the work of Lunardi et al. [9], a filter bubble can be described as the amount of homogeneity in the recommended set, which indicates the narrowing of the proposed results. The homogeneity between the two recommended sets can be considered the average pairwise distance between the elements of the two sets. We propose the following metric, which computes the distance between movies using actor, director, and genre differences.

$$
\begin{aligned}
H &= \text{mean}(G + A + D), \\
G &= \frac{1}{N(N-1)} \sum_{i \neq j} S_{ij}, \\
A &= \frac{1}{N(N-1)} \sum_{i \neq j} T_{ij}, \\
D &= 1 - \frac{U}{N}
\end{aligned}
\tag{1}
$$

- G is the genre homogeneity, which is the average pairwise Jaccard index between the genres of the recommended movies (S_{ij}).
- A is the actor homogeneity, which is the average pairwise Jaccard index between the actor lists of the recommended movies (T_{ij}).
- D is the director homogeneity, where U is the number of unique directors, and N is the total number of movies.

The use of the Jaccard index for measuring filter bubbles was first proposed by Chaney et al. [2]. The generic definition of the Jaccard Index is:

$$
J_{ij} = \frac{|\text{features of movie}_i \cap \text{features of movie}_j|}{|\text{features of movie}_i \cup \text{features of movie}_j|}
$$

The proposed homogeneity metric is used to evaluate the degree of homogeneity in the recommended items generated during the simulation. Other metrics, such as the one proposed by Nguyen et al. [10], may be used based on the Tag Genome of MovieLens.

6 EXPERIMENTAL RESULTS

We have developed a collaborative filtering recommender system and trained on the MovieLens data. We performed two simulations with 50 agents over 100 timestamps in two independent runs:

- In the first simulation, the agents were not polarized and selected and rated a movie randomly from the recommendation set at each step.
- In the second simulation, agents were polarized towards crime/thriller movies. They selected items randomly from the recommendations but rated only crime/thriller movies positively.

Our purpose is to analyze the recommendation system's effect on non-polarized and polarized users and evaluate the evolution of the homogeneity of the recommended set. Figure 2 shows the homogeneity trend for 4 random agents in the first simulation. Figure 3 shows the homogeneity trend for 4 polarized agents in the second simulation.

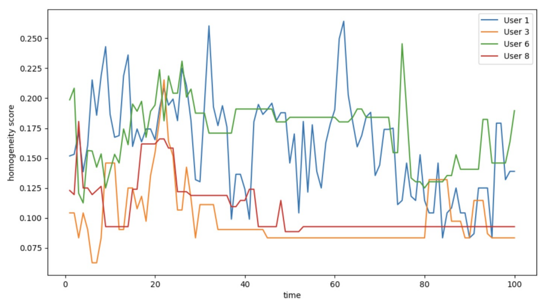

Figure 2: Homogeneity score for random users.

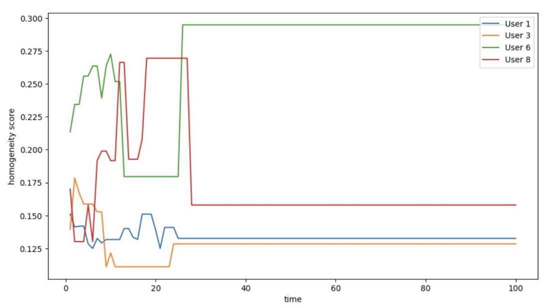

Figure 3: Homogeneity score for polarized users.

Figure 2, representing a scenario with random selection and random ratings, exhibits dynamic fluctuations in homogeneity scores, ranging from approximately 0.075 to 0.25. The patterns appear volatile and unpredictable, with no clear convergence to a single value and different users showing varying trends over time. This suggests that the system actively explores diverse recommendations.

In contrast, Figure 3, which depicts a scenario with random selection but biased towards the crime/thriller genre, the graph contains homogeneity scores that rapidly homogenise after initial fluctuations. All users converge to relatively stable scores around timestamp 30, with User 6 stabilizing at a notably higher level of approximately 0.3. In contrast, other users stabilize between 0.125 and 0.16, exhibiting significantly less variation than Figure 2. Additionally, filter bubbles can vary in homogeneity levels, with some being stronger than others.

7 IDEAS FOR NEXT STEPS

Our plan for the next steps is to estimate the average treatment effect (ATE) of the recommendation system on the formation of filter bubbles, which are characterized by recommendations with a high level of homogeneity. Our simulation provides data about the probability:

$$P(\text{homogeneity of watched items} \mid \text{do(recommendation)}),$$

but we lack information about the probability:

$$P(\text{homogeneity of watched items} \mid \text{do(!recommendation)}),$$

which prevents us from computing the ATE directly.

To address this, we plan to use the MovieLens dataset. Specifically, we can extract a subgroup of users, use our simulation to compute their recommended items at each step, and compare this set with the movies they choose. If a chosen film is among the suggested ones, we can conclude that the recommendation system influences the user. Suppose the subgroups are relatively small (since, in reality, we don't know the exact state of the recommendation system at the moment a user chooses to watch a film). In that case, this approach allows us to define two subgroups of users: those who follow the recommendations and those who do not, enabling us to estimate the ATE.

Finally, we will explore the potential for developing tools to automatically monitor and optimize the balance between user agency and algorithmic influence in movie recommender systems.

REFERENCES

[1] Firstname Argueda and Firstname Others. 2023. Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review. *Journal Name* X, Y (2023), Z–ZZ. https://example.com
[2] A.J.B. Chaney, B.M. Stewart, and B.E. Engelhardt. 2018. How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *Proceedings of the 12th ACM Conference on Recommender Systems*. 224–232.
[3] S. Messing E. Bakshy and L. A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
[4] C. Gao, Y. Zheng, W. Wang, F. Feng, X. He, and Y. Li. 2024. Causal inference in recommender systems: A survey and future directions. *ACM Transactions on Information Systems* 42, 4 (2024), 1–32.
[5] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015), 1–19.
[6] R. Jiang, S. Chiappa, T. Lattimore, A. György, and P. Kohli. 2019. Degenerate feedback loops in recommender systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 383–390.
[7] Greg Linden. 2011. Eli Pariser is Wrong. http://glinden.blogspot.com/2011/05/eli-pariser-is-wrong.html
[8] P. Liu, K. Shivaram, A. Culotta, M.A. Shapiro, and M. Bilgic. 2021. The interaction between political typology and filter bubbles in news recommendation algorithms. In *Proceedings of the Web Conference 2021*. 3791–3801.
[9] G.M. Lunardi, G.M. Machado, V. Maran, and J.P.M. de Oliveira. 2020. A metric for filter bubble measurement in recommender algorithms considering the news domain. *Applied Soft Computing* 97 (2020), 106771.
[10] T.T. Nguyen, P.-M. Hui, F.M. Harper, L. Terveen, and J.A. Konstan. 2014. Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web*. 677–686.
[11] Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding from You.* Penguin UK. 10 pages.
[12] J. Vig, S. Sen, and J. Riedl. 2012. The tag genome: Encoding community knowledge to support novel interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 3 (2012), 1–44.
[13] W. Wang, F. Feng, L. Nie, and T.-S. Chua. 2022. User-controllable recommendation against filter bubbles. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1251–1261.
[14] D. Xu, C. Ruan, E. Korpeoglu, S. Kumar, and K. Achan. 2020. Adversarial counterfactual learning and evaluation for recommender system. In *Advances in Neural Information Processing Systems*, Vol. 33. 13515–13526.