AI Facilitated Isolations? The Impact of Recommendation-based Influence Diffusion in Human Society

Yuxuan Hu[1], **Shiqing Wu**[1], **Chenting Jiang**[1,3], **Weihua Li**[2], **Quan Bai**[1] and **Erin Roehrer**[1]
[1]School of Information and Communication Technology, University of Tasmania
[2]School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology
[3]Tasmanian Institute of Agriculture, University of Tasmania
{yuxuan.hu, shiqing.wu, chenting.jiang, quan.bai, erin.roehrer}@utas.edu.au, weihua.li@aut.ac.nz

Abstract

AI recommendation techniques provide users with personalized services, feeding them the information they may be interested in. The increasing personalization raises the hypotheses of the "filter bubble" and "echo chamber" effects. To investigate these hypotheses, in this paper, we inspect the impact of recommendation algorithms on forming two types of ideological isolation, i.e., the individual isolation and the topological isolation, in terms of the filter bubble and echo chamber effects, respectively. Simulation results show that AI recommendation strategies severely facilitate the evolution of the filter bubble effect, leading users to become ideologically isolated at an individual level. Whereas, at a topological level, recommendation algorithms show eligibility in connecting individuals with dissimilar users or recommending diverse topics to receive more diverse viewpoints. This research sheds light on the ability of AI recommendation strategies to temper ideological isolation at a topological level.

1 Introduction

Social media enables users to exchange information and interact with others. AI techniques like Recommendation Systems have been widely applied to online social media, providing users with more personalized services and bringing fundamental changes to the influence diffusion style. Recommendation-based social medias, such as YouTube and Facebook, feed user information (influence messages) posted by others based on recommendation strategies. Traditional influence diffusion models capture the interpersonal influence between each pair of users based on network structure [Liu *et al.*, 2017], i.e., arcs denoting physical relationships. Whereas, in recommendation-based social medias, influence flows to a user from the poster of a recommended influence message even though they have no connection. In other words, AI matches two users according to recommendation strategies and establishes a virtual link between them, which can spread influence. The influence diffusion on a recommendation-based social media is denoted as recommendation-based influence diffusion, in which influence diffuses through both network structure and AI-established virtual links.

Along with changing fundamental diffusion styles, AI recommendation techniques also change how users perceive information. Rather than receiving information from adjacent neighbors or searching for required information, individuals are fed by AI recommendation techniques that estimate the potential preferences of users [Adomavicius and Tuzhilin, 2005]. In offering personalized information based on user preference, recommendation algorithms were expected to reduce information diversity, leading users to ideological isolation [Haim *et al.*, 2018]. Such a hypothesis is termed the "filter bubble" effect [Pariser, 2011], which describes a unique and personal information universe that results from recommendation techniques, leading users to become individually isolated from diverse information. A similar concept focuses on the increased chance of interacting with like-minded peers, termed the "echo chamber" effect [Flaxman *et al.*, 2016], suggesting topological isolation in which individuals' opinions or beliefs towards a topic get reinforced due to the repeated interaction with these peers. Several studies have explored the "filter bubble" and "echo chamber" hypothesis [Nguyen *et al.*, 2014; Cinelli *et al.*, 2021; Jiang *et al.*, 2019]. However, most of these works ignore the effect of influence diffusion triggered by AI recommendation algorithms, in which individuals not only perceive information from their neighbors but are also passively fed by the AI recommendation platform.

In this paper, we investigate the impact of AI on leading two types of isolation in human society, i.e., individual isolation and topological isolation, by exploring the role of AI in the evolution of filter bubble and echo chamber, respectively. We first propose a Recommendation-based Influence Diffusion Model (RIDM) for computing influence diffused through both network structure and AI-established virtual links. We explore the role of AI in the evolution of two types of ideological isolation from two dimensions, i.e., the information diversity and the preference similarity. We conduct simulations to explore the impact of AI in forming the filter bubble effect and the echo chamber effect with the comparison of the traditional influence diffusion model. We also investigate different AI recommendation strategies' impact on forming filter bubbles and echo chambers. The experimental results reveal that AI in recommendation-based influence diffusion significantly impacts the evolution of the filter bubble effect compared with traditional peer-to-peer influence diffusion and leads users to ideological information isolation. Regardless of the attendance of recommendation algorithms, the echo chamber effect evolves in the network over time. However, AI recommendation algorithms are able to recommends diverse information by connecting dissimilar users or recommending diverse topics to temper the echo chamber effect. To summarize, the main contributions of this paper are summarized as follows:

- To the best of our knowledge, RIDM is the first model which considers the impact of AI recommendation in the influence diffusion process. It computes both peer-to-peer influence and influence from AI recommendations, which is suitable for real-world social media with recommendation services.
- We propose a novel quantification approach to explore the impact of AI in recommendation-based influence diffusion on individuals' ideological isolation in terms of the filter bubble effect and echo chamber effect.
- Simulation results reveal that AI recommendation has an impact on the evolution of the filter bubble effect and causes individuals' ideological isolation. However, recommendation algorithms, i.e., user-based filtering and content-based filtering, are eligible to connect individuals with others who can spread diverse information. This sheds a light on tempering the echo chamber effect with AI recommending diverse information.

The rest of this paper is organized as follows. In Section 2, we review literatures related to this research work. In Section 3, we introduce the preliminaries of this work, including the framework of RIDM and formal definitions. Section 4 describes the proposed RIDM diffusion model in detail. Section 5 elaborates the proposed quantification approaches for measuring ideological isolations. The simulation-based experimental results are demonstrated and analyzed in Section 6. Finally, the paper is concluded in Section 7.

2 Related Works

In recent years, AI recommendation techniques have raised concerns about the filter bubble effect [Pariser, 2011]. Nguyen et al. measure the filter bubble effect brought by a collaborative filtering recommendation strategy in terms of content diversity [Nguyen *et al.*, 2014]. They suggest the recommendation strategy exposed users to narrowing items over time. Cinelli et al. quantify the echo chamber effect in social networks from two dimensions, i.e., homophily and bias. They find social networks with news feed algorithms that consider users preferences foster the formation of the echo chamber effect [Cinelli *et al.*, 2021]. In contrast, researchers found that individuals who use social networks and search engines to browse news are more ideologically distant from each other as the vast majority of users simply visit the home pages of their favorite, tempering the consequence of AI recommendation techniques [Flaxman *et al.*, 2016]. Dubois et al. conduct a national survey of 2,000 internet users and suggest that individuals have high-choice media sources to get more diverse information, thus avoiding the echo chamber effect [Dubois and Blank, 2018].

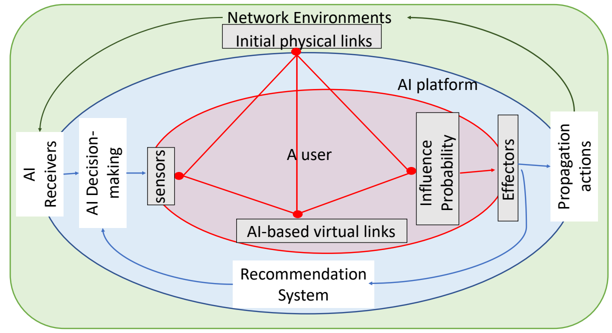


Figure 1: The Framework of RIDM

However, these works neglect the impact of influence diffusion triggered by AI algorithms, in which individuals passively feed information and experience the information propagated by their neighbors. To fill this gap, in this work, we aim to investigate the impact of recommendation algorithms to embed with an influence diffusion process.

Independent Cascade (IC) model [Kempe *et al.*, 2003] is recognized as one of the fundamental influence diffusion models. In the IC model, influences are propagated from active nodes to inactive neighbors with an influence probability. Each inactive neighbor has a single chance to be activated. The influence diffusion process stops until no node can be activated. To capture the individual behavior in influence diffusion, Agent-based Modelling (ABM) [Li *et al.*, 2016] is adopted, where the influence diffusion process is modelled as an evolutionary process driven by individual agents' behaviors [Li *et al.*, 2019].

3 Preliminaries

In this section, we introduce the preliminaries of RIDM, including framework and formal definitions.

3.1 Framework

The framework of RIDM, shown in Figure 1, describes the influence process of an individual user in a social network. This model is composed of three layers and two stages. Firstly, the outermost layer (green color) represents a network environment filled with all kinds of messages based on different topics. Secondly, the middle layer (blue color) denotes the AI platforms, as a transfer channel with some intelligent functions, which consider the user's behaviors and a recommendation system to make decisions in the initial stage. Thirdly, the innermost layer (red color) presents a user that could get a type of influence through a message transmission due to AI-based virtual links and initial physical relationships.

In the first stage, the platform selects some messages from all received messages and then transfers the selected messages to the user. This stage is regarded as the AI-based message transmission, reflecting the impact of AI platforms in the diffusion process. The decision of AI selection derives from different recommendation strategies. Subsequently, the sensor of a user receives messages from an AI platform, moving on to the second stage. This stage is the influenced phase of a user. Whether the user is impacted by a received message depends on physical relationships between the sender and the receiver, and virtual linkage through the AI system.

3.2 Formal Definitions

A social network refers to a directed graph $G = (V, E, A, T)$, which consists of a set of users $V = \{v_1, v_2, ..., v_n\}$, a set of edges $E = \{e_{ij} | 1 \leq i, j \leq n\}$, an AI platform A, and a set of topics $T = \{T_1, T_2, ..., T_n\}$.

Definition 1. ***A User*** *$v_i \in V$ is defined as a vertex in the social network graph, representing an interactive and autonomous user in an AI platform. Each user is active in an AI platform to interact with its physical neighbors and virtual neighbors. That is to say, a user is able to handle multiple topics influence. Every user has his or her dynamic preference for diverse topics, which will be explained in Definition 6.*

Definition 2. ***An Edge*** *$e_{ij} = (v_i, v_j)$ denotes the influence relationship between v_i and v_j, $e_{ij} \neq e_{ji}$, where $E = \{e_{ij} | 1 \leq i, j \leq n\}$. Different from traditional influence diffusion models, the edge in RIDM is expressed by online social platform, including the direct and indirect linkages. The direct (or physical) edge means two users have an offline physical relationship, such as friends, families, and colleagues. While the indirect (or virtual) edge presents two users who have an online virtual relationship on a social media AI platform.*

In the real-world applications, each pair of users connected by a direct edge can share messages directly. However, if there is no direct edge existing between two users, the message sharing can be only realized by the AI platform's recommendations, which would build an indirect edge between them. Therefore, both types of edges coexist in RIDM, and also a pair of users could have both relationships simultaneously.

The neighbors of a user are indicated as $\mathcal{N}_i = \mathcal{N}_i^{di} \cup \mathcal{N}_i^{in}$. $\mathcal{N}_i^{in} = \mathcal{N}_i^{in^s} \cup \mathcal{N}_i^{in^r}$ is the neighbors which have virtual links with v_i, where $\mathcal{N}_i^{in^s} = \{v_j | v_j \in V, v_j \neq v_i, e_{ji} \in E\}$ is a sender set of $\mathcal{N}_i^{in}$, and $\mathcal{N}_i^{in^r} = \{v_j | v_j \in V, v_j \neq v_i, e_{ij} \in E\}$ is a receiver set of $\mathcal{N}_i^{in}$. $\mathcal{N}_i^{di} = \mathcal{N}_i^{di^s} \cup \mathcal{N}_i^{di^r}$ is the neighbor set contains the users who have physical relationships with v_i, where $\mathcal{N}_i^{di^s}$ and $\mathcal{N}_i^{di^r}$ are the sender-neighbor set and the receiver-neighbor set of the physical neighbors $\mathcal{N}_i^{di}$, respectively.

Definition 3. ***An Influence Message*** *$m_{ij}^{x_k}$ in the social networks represents a piece of information $k \in \mathbb{N}^+$ about a topic x transferred through an AI-based platform, and this message is diffused from v_i to v_j. A set of messages $M^x = \{m^{x_1}, m^{x_2}, ..., m^{x_k}\}$ includes all messages about topic x.*

Definition 4. ***A Topic*** *x is considered as a label of messages, where $x \in T$ and $T = \{T_1, T_2, ..., T_n\}$. Messages with the same topic label describe the same major content, so that each message only has one topic label. In RIDM, a user may have different preferences toward distinct topics. Meanwhile, we define s^x represents the number of messages about a topic x.*

Definition 5. ***Topic Correlation*** *$\rho(x, x') \in [0, 1]$ denotes the relevancy between topic x and x' in an AI-based platform. The value of topic correlation explains the influence possibility between two topics, where $\rho(x, x') = \rho(x', x)$. That means when transferring a message with topic x in a platform, the user has a probability of being influenced by a relevant topic x', and vice versa. Topic correlation reveals the complexity of influence diffusion in a multi-topic social platform. If $\rho(x, x') = 0$, topic x and x' do not have any relation, and they cannot impact each other. Whereas if $\rho(x, x')$ infinitely closes to 1, x and x' display a very strong correlation in potential influence diffusion. Normally, the value of $\rho(x, x')$ more, the topic correlation stronger.*

Definition 6. ***User Preference*** *r_i^x represents a personal emotional attitude toward a topic on messages. $r_i^x \in (0, 1)$ expresses the degree of user preference on topic x. When $r_i^x \to 0$, the user v_i dislike topic x, meaning a negative attitude. When $r_i^x \to 1$, v_i is in favor of topic x, displaying a positive attitude. $r_i^x = 0.5$ is regarded as a neutralizing attitude. The user preference is a dynamic value.*

We assume the dynamic values of user preference are calculated by:

$$r_{i,t}^x = \frac{\sum_1^t s_{i,t}^x}{\sum_1^t s_{i,t}}, \tag{1}$$

in which $s_{i,t}$ is the total number of messages v_i has sent from time step 0 to time step t, and $s_{i,t}^x$ is the number of messages related to topic x that v_i has sent from time step 0 to time step t.

4 Recommendation-based Influence Diffusion Model Derivation

In RIDM, an influence message $m_{ij}^{x_k}$ with a specific topic x is diffused from a user v_i to another user v_j. From a macro perspective, the recommendation-based influence diffusion can be composed of two phases: the AI recommending message transmission and the influence diffusion. At the first stage, AI recommends a fixed number of influence messages to v_i, establishing the virtual links between v_i and the sender of the recommended messages. Influence diffusion starts at the second stage. Besides peer-to-peer influence diffusion on physical links, influences are diffused through AI-established virtual links. In the following subsection, we will introduce these two phases in details.

4.1 AI Recommending Message Transmission

The message transmission phase focuses on sending a message from one user to another through an AI-based platform. Two recommendation strategies are designed based on two traditional filtering methods, i.e., User-based Collaborative Filtering (UC) and Content-based Filtering (CB) [Adomavicius and Tuzhilin, 2005]:

User-based collaborative filtering predicts the utility of messages to a particular user v_i based on the historical messages sent by other users in the social network. The utility between a message $m_{ji}^{x_k}$ and a user v_i is calculated by the similarity of two users and the user preference in Equation 2.

$$score_{v_i,m_{ji}^{xk}} = sim(v_i, v_j) \cdot r_i^x, \tag{2}$$

where m_{ji}^{xk} represents the k-th influence message with a particular topic x sent by user v_j and transmitted to v_i. r_i^x denotes the preference of v_i to a topic x. $sim(\cdot)$ represents a function, calculating the similarity between two users based on their historical behaviors, which is formulated as:

$$sim(v_{i,t}, v_{j,t}) = \frac{\mathbf{s}_{i,t}\mathbf{s}_{j,t}^{\top}}{||\mathbf{s}_{i,t}|| \cdot ||\mathbf{s}_{j,t}||}, \tag{3}$$

where $s_{v_{i,t}}$ and $s_{v_{j,t}}$ denote the number of sending messages of v_i and v_j from time step 0 to time step t.

Content-based filtering is a method that compares a message m^x with a user v_i according to topic correlation in Definition 5 and v_i's historical behavior at time step t:

$$score_{v_i,m^x} = \sum_{x' \in T} \rho(x, x') s_{i,t}^x. \tag{4}$$

In Equation 4, $\rho(x, x')$ is the topic correlation between a topic and any of a topic in T. $s_{i,t}^x$ is the number of messages sent by v_i from time step 0 to time step t.

We assume the AI recommendation is conducted over each user in the social network with a particular recommendation strategy. The AI platform collects all the influence messages that users posted to the public space at each time step. Then, the AI platform calculates the score between each message and a targeted user. By ranking the score of each message, the AI platform selects Top-k messages to transmit to the targeted user and establishes virtual links between the senders of the recommended messages and the targeted.

4.2 Influence Diffusion

After message transmission, a topic's influence diffusion probability consists of physical influence possibility and virtual influence possibility. These two probabilities are derived from two types of edge, i.e. physicals and virtual relationships. To be specific, in an AI platform, a user receives an influence message in two possible ways, i.e., peer-to-peer propagation from its neighbors and the feeding by AI. When AI feeds an influence message to a user, AI establishes a virtual relationship between the users and the sender of the message. The Physical Influence Probability (i.e. PIP) depends on the number of messages of a pair of users at time $t-1$, as follow:

$$PIP_{j,t}^x = \sum_{x' \in T} \sum_{i \in \mathcal{N}_j^{dis}} \frac{1}{|\mathcal{N}_j^{dis}||T|} \rho(x, x') \frac{s_{i,t-1}^x + s_{j,t-1}^x}{s_{i,t-1} + s_{j,t-1}} r_{j,t-1}^x, \tag{5}$$

where $|\mathcal{N}_j^{dis}|$ represents the number of users in $\mathcal{N}_j^{dis}$. r_i^x and r_j^x denote the preferences of v_i and v_j to the topic x.

Meanwhile, the Virtual Influence Probability (i.e. VIP) is determined by the sets of messages of senders and receivers at time step $t-1$.

$$VIP_{j,t}^x = \sum_{x' \in T} \sum_{i \in V} \frac{1}{|V||T|} \rho(x, x') \frac{|\hat{M}_{i,t-1}^x \cap \check{M}_{j,t-1}^x|}{|M_{i,t-1}^x \cup M_{j,t-1}^x|}, \tag{6}$$

In Equation 6, $\hat{M}_{i,t-1}^x$ denotes the set of sent messages of user v_i at $t-1$. Similarly, $\check{M}_{j,t-1}^x$ denotes the set of received messages of user v_j at $t-1$. $|M|$ represents the cardinality of M.

Combining Equations 5 and 6, the probability of influence diffusion in the influence diffusion phase is demonstrated, as follows:

$$P_{j,t}^x = \lambda PIP_{j,t}^x + (1-\lambda) VIP_{j,t}^x, \tag{7}$$

where λ denotes a weight between $PIP_{j,t}^x$ and $VIP_{j,t}^x$. Suppose there is no physical link between v_i and v_j, $\lambda = 0$. Otherwise, $0 < \lambda < 1$, meaning e_{ij} involves physical and virtual link simultaneously. Notably, if there is no virtual link between v_i and v_j or no AI platform on the social network, RIDM can be returned to the traditional influence model only based on physical influence probability (i.e. $P_{j,t}^x = PIP_{j,t}^x$).

5 Quantification of Ideological Isolations

AI algorithms feed users with influence messages based on various recommendation strategies, forming personal, unique information universes for users. This raises a risk of self-reinforcement and reduced diversity for one's consumed information as AI algorithms continuously recommend influence messages that match a user. This phenomenon is so-called the "filter bubble" effect [Pariser, 2011]. A similar concept focuses on increasing the chance of like-minded neighbors [McPherson *et al.*, 2001], forming a situation in which an individual is more likely to interact with peers who have similar points of view, which is called the "echo chamber" effect [Gillani *et al.*, 2018]. In this section, the impact of the recommendation-based influence diffusion on a user is measured with two indices. At an individual level, we measure the filter bubble effect in terms of the diversity of the influence messages that a user consumes. At a topological level, we measure the echo chamber effect in terms of preference similarity between a user and its social relationships.

5.1 Measuring Filter Bubble Effect

We first measure the filter bubble effect which is formed by the recommendation algorithms in the recommendation-based influence diffusion in terms of information diversity. At an individual level, recommendation algorithms feed a user v_i influence messages towards a particular topic.

Specifically, v_i holds an influence message repository $\check{M}_{i(t)}$, a collection of influence messages that v_i received at time step t. A user v_i can receive any influence message from the AI platform via a recommendation strategy. Meanwhile, v_i also possibly receives influence messages from its physical sender-neighbors $v_j \in \mathcal{N}_i^{dis}$. $|\check{M}_{i,t}| = s_{ir,t}$ is the total number of influence messages that v_i received at time step t, where $s_{ir} = \sum_{1}^{n} s_{ir}^x$. The effect of an individual filter bubble on an AI platform. $Q_f[\check{M}_{i,t}]$ is the received influence message repository $\check{M}_{i,t}$ of a user v_i with a quantification function $Q_f(\cdot)$.

We adopt the idea of information theory to quantify the diversity of a user's influence repository with entropy [Jaynes, 1957]. Specifically, a higher $Q_f[\check{M}_{i,t}]$ implies a more disorder influence repository a user v_i holds, suggesting v_i is in a more diverse information universe and less likely to stay in a filter bubble. In this sense, the quantification equation to measure the filter bubble effect on an AI platform of a user v_i can be formulated as:

$$Q_f[\check{M}_{i,t}] = -\sum_{x=1}^{n} \frac{s_{ir}^x}{s_{ir}} \ln(\frac{s_{ir}^x}{s_{ir}}), \tag{8}$$

in which $\frac{s_{ir}^x}{s_{ir}}$ describes the ratio of the number of influence messages v_i received about a specific topic x to the total number of influence messages v_i received on the AI platform. A higher $Q_f[\check{M}_{i,t}]$ indicates the influence messages in $\check{M}_{i,t}$ are more diverse. $Q_f[\check{M}_{i,t}] = 0$ suggests that all the influence messages in $\check{M}_{i,t}$ are related to only one specific topic.

5.2 Measuring Echo Chamber Effect

We also consider how likely a user v_i is to exist in an echo chamber, considering its social relationships in terms of the similarity between v_i and its neighbors $\mathcal{N}_{i,t}$. A higher similarity suggests a higher chance of interacting with like-minded peers and encountering information that supports its preference. In this sense, the user is more likely to be trapped in an echo chamber. Thus, the echo chamber indice of a user v_i as a quantification function $Q_e[\mathcal{N}_{i,t}]$ on a user v_i's sender-neighbor set $\mathcal{N}_{i,t}^s = \mathcal{N}_{i,t}^{in^s} \cup \mathcal{N}_{i,t}^{di^s}$ is denoted as:

$$Q_e[\mathcal{N}_{i,t}^s] = \frac{\sum sim[r_{i,t}, r_{j,t}]}{|\mathcal{N}_{i,t}^s|}. \tag{9}$$

$r_{i,t}$ is the preference of a user v_i at time step t. $sim[r_{i,t}, r_{j,t]}$ is the preference similarity between a user v_i and any of its neighbor v_j illustrated in the following equation:

$$sim[r_{i,t}, r_{j,t}] = \frac{r_{i,t} r_{j,t}^\top}{||r_{i,t}|| \cdot ||r_{i,t}||} \tag{10}$$

$Q_e[\mathcal{N}_{i,t}] \in [0,1]$ quantifies the similarity between a user v_i's preference and the preference of its neighbors. A higher $Q_e[\mathcal{N}_{i,t}]$ implies a higher similarity between v_i and its neighbors and a higher chance for v_i to stay in an echo chamber on the AI platform. $Q_e[\mathcal{N}_i(t)] = 0$ suggests user v_i has no similarity with all of its neighbors, while $Q_e[\mathcal{N}_i(t)] = 1$ implies the preference of user v_i complete overlap its neighbors.

We use preference similarity in Equation 10 to quantify the echo chamber effect because it provides evidence of homophily, which is a main mechanism behind the echo chamber effect [Cinelli *et al.*, 2021; McPherson *et al.*, 2001]. Homophily describes a principle wherein contact between similar people occurs higher than in dissimilar people [McPherson *et al.*, 2001].

6 Experiments

6.1 Datasets and Recommendation Strategies

The experiments are conducted by using two real-world datasets, ego-Facebook [1] and ego-Twitter [2] [Leskovec and Mcauley, 2012]. The ego-Facebook dataset has 4,039 nodes and 88,234 edges. To solve the computational cost, we select a sub-graph in the ego-Twitter dataset with 236 nodes and 2,478 edges.

We implement our influence diffusion model under two recommendation strategies derived in Sub-section 4.1:

- Content-based: AI recommends a user with influence messages based on the topic similarity between the influence messages and their historical behaviors.
- User-collaborative: AI recommends a user with influence messages based on the similarity between a user and other users in the network.

We also adopt the traditional IC model as a benchmark, having no recommendation algorithm involved in the process of influence diffusion.

6.2 Experimental Settings

Assumptions are given as follows:

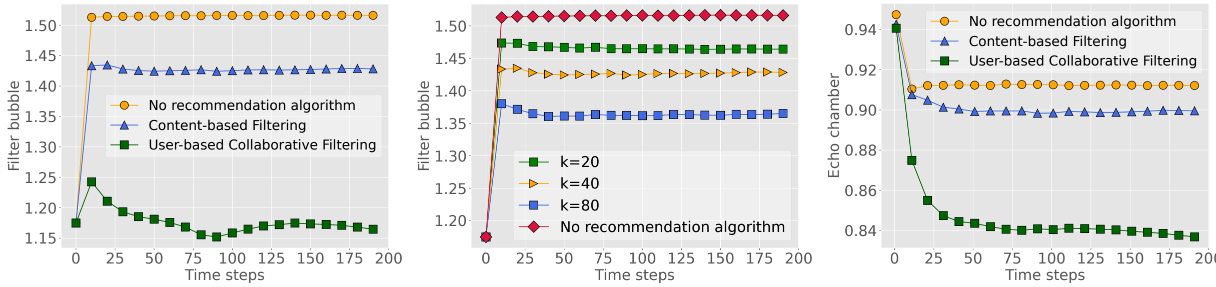
- Starts from time step 0, every user in the network would send a message with their favorite topic at each time step. A user's preference is updated after its sending behavior based on its historical behavior at each time step according to Equation 1.
- The AI recommendation-based influence diffusion process starts from time step 1. At each time step $t > 0$, AI feeds users with recommended messages that other users sent to AI platform at previous time steps. If a user is influence by a topic x at time step t, it will take action, i.e., send a message to the AI platform with the same topic. A user can be influenced by multiple messages at the same time.
- We assume each user randomly selects a topic x to send an influence message before step 0. In other words, $s_{i,0}^x = 1$.

6.3 Experiment 1: Quantifying filter bubble effect

To quantify the filter bubble effect in a social network, we adopt Equation 8 to estimate the individual information diversity over all the users in the network and plot the average $\bar{Q}_f[\check{M}_{i,t}] = \frac{\sum_{v_i \in V} Q_f[\check{M}_{i,t}]}{|V|}$ at each time step. The results of three strategies in two real-world networks are demonstrated in Figures 2(a) and 3(a). As can be observed from both figures, the filter bubble effect evolves over time under three methods. Compared with influence diffusion without any recommendation algorithms, the filter bubble index $\bar{Q}_f[\check{M}_{i,t}]$ shows a rapid decrease with the attendance of recommendation algorithms. In particular, UC has the lowest $\bar{Q}_f[\check{M}_{i,t}]$ in both datasets, suggesting its ability to form the most monotonous information universes of users.

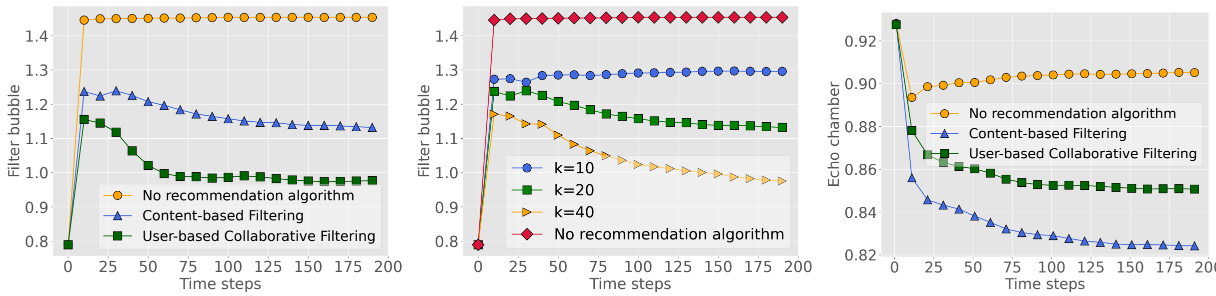
We further analyzed the relationship between the filter bubble effect and the k number of recommending messages. By deploying different k on one particular recommendation strategy, i.e., content-based, the result of the average $\bar{Q}_f[\check{M}_{i,t}]$ of different k at each time step are shown in Figures 2(b) and 3(b). With the increasing number of k, the filter bubble effect

[1] https://snap.stanford.edu/data/ego-Facebook.html
[2] https://snap.stanford.edu/data/ego-Twitter.html



(a) ego-Facebook, k=40 (b) ego-Facebook, CB (c) ego-Facebook, k=40

Figure 2: Development of Filter Bubble and Echo Chamber under different recommendation approaches (the ego-Facebook dataset)



(a) ego-Twitter, k=20 (b) ego-Twitter, CB (c) ego-Twitter, k=20

Figure 3: Development of Filter Bubble and Echo Chamber under different recommendation approaches (the ego-Twitter dataset)

is more severe in the two networks. The results reveal that individuals are high-likely in individual information isolation if they consumed more information from AI recommendations (i.e., spend more time on AI-based social platforms).

6.4 Experiment 2: Quantifying echo chamber effect

To explore how RS contributes to topological connections, we adopt $Q_e[\mathcal{N}_{i,t}]$ in Equation 9 to quantify the echo chamber degree of a group in which RS establishes the connections in a social network. After several rounds of influence diffusion, we recorded the average $\bar{Q}_e[\mathcal{N}_{i,t}] = \frac{\sum_{v_i \in V} Q_e[\mathcal{N}_{i,t}]}{|V|}$ of all the user in the network. The results in two datasets are presented in Figures 2(c) and 3(c).

The average $\bar{Q}_e[\mathcal{N}_{i,t}]$ converges with both three methods in two networks, suggesting the echo chamber effect formation in the networks. In other words, no matter the interactions between users or the interactions between users and AI platforms unavoidably bring the echo chamber effect with different degrees. The results reveal that both UC and CB have a lower $\bar{Q}_e[\mathcal{N}_{i,t}]$ than no recommendation algorithm. This is because UC is a recommendation strategy that finds the most similar users according to the similarity of their historical behaviors, rather than focusing on similar preferences. While CB considers the topic correlation between two topics. In this sense, the topics related to a user's favorite topic could be recommended, making it receive diverse viewpoints.

7 Conclusion

In this paper, we investigated the impact of recommendation algorithms on causing ideological isolations. We used filter bubble indices and echo chamber indices to quantify the corresponding "filter bubble" and "echo chamber" effects evolved in a recommendation-based influence diffusion process on a single AI platform. The filter bubble indice refers to the information diversity of a user's consumed information. The echo chamber indice denotes the preference similarity between a user and its neighbors. The simulation results suggest that AI recommendation algorithms impact accelerating the filter bubble effect. The echo chamber effect is evolved in the influence diffusion process, regardless of the attendance of AI. Whereas, AI recommendation methods can connect users with dissimilar ones, tempering the user's ideological isolation at a topological level. In the future, we aim to further investigate these effects on multiple AI platforms, to fit the real-world situation that individuals involved in multiple media environments.

References

[Adomavicius and Tuzhilin, 2005] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.

[Cinelli *et al.*, 2021] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

[Dubois and Blank, 2018] Elizabeth Dubois and Grant Blank. The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5):729–745, 2018.

[Flaxman *et al.*, 2016] Seth Flaxman, Sharad Goel, and Justin M Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320, 2016.

[Gillani *et al.*, 2018] Nabeel Gillani, Ann Yuan, Martin Saveski, Soroush Vosoughi, and Deb Roy. Me, my echo chamber, and i: Introspection on social media polarization. In *Proceedings of the 2018 World Wide Web Conference*, page 823–831, 2018.

[Haim *et al.*, 2018] Mario Haim, Andreas Graefe, and Hans-Bernd Brosius. Burst of the filter bubble? effects of personalization on the diversity of google news. *Digital journalism*, 6(3):330–343, 2018.

[Jaynes, 1957] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.

[Jiang *et al.*, 2019] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, and Pushmeet Kohli. Degenerate feedback loops in recommender systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, page 383–390, 2019.

[Kempe *et al.*, 2003] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 2003 ACM International Conference on Knowledge Discovery and Data Mining*, page 137–146, 2003.

[Leskovec and Mcauley, 2012] Jure Leskovec and Julian Mcauley. Learning to discover social circles in ego networks. In *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[Li *et al.*, 2016] Weihua Li, Quan Bai, and Minjie Zhang. Agent-based influence propagation in social networks. In *Proceeding of the 2016 IEEE International Conference on Agents*, pages 51–56, 2016.

[Li *et al.*, 2019] Weihua Li, Quan Bai, and Minjie Zhang. A multi-agent system for modelling preference-based complex influence diffusion in social networks. *The Computer Journal*, 62(3):430–447, 2019.

[Liu *et al.*, 2017] Shenghua Liu, Houdong Zheng, Huawei Shen, Xueqi Cheng, and Xiangwen Liao. Learning concise representations of users' influences through online behaviors. In *Proceedings of the 2017 International Joint Conference on Artificial Intelligence*, pages 2351–2357, 2017.

[McPherson *et al.*, 2001] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.

[Nguyen *et al.*, 2014] Tien T. Nguyen, Pik-Mai Hui, F. Maxwell Harper, Loren Terveen, and Joseph A. Konstan. Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 2014 International Conference on World Wide Web*, page 677–686, 2014.

[Pariser, 2011] Eli Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.