

Algorithmic Bias and Information Diversity: Formation Mechanism and Countermeasure Path of Filter Bubbles in Social Media

Runhe Gu

*School of Culture and Media, Xi'an Eurasian College, Xi'an, China
grrrrhi@163.com*

Abstract. Driven by new media technology, the algorithmic recommendation mechanism of short video platforms has enhanced user stickiness through personalised content matching, but also triggered the systematic erosion of information diversity by filter bubbles. Drawing upon relevant academic literature, this study employs a literature review methodology to analyze the symbiotic relationship between algorithmic bias and the information cocoon phenomenon. It further explores the mechanisms through which filter bubbles influence users' cognitive perceptions and the content ecosystem, as well as corresponding countermeasure strategies. It is concluded that algorithmic recommendation reinforces user preferences through "data dependence + commercial logic", leading to information narrowing and value illusion; although users can break through part of the closure through active search, they need to rely on the synergy of technology optimisation, media literacy enhancement and policy regulation. Existing research still exhibits gaps in exploring cross-generational user differences, the impact of AIGC technology, and long-term social effects. Future studies ought to construct an algorithmic ethical framework that integrates technical rationality and value rationality.

Keywords: algorithmic bias, filter bubble, information cocoon, information diversity, platform governance

1. Introduction

By 2025, the scale of global short video users have exceeded 5 billion, and algorithmic recommendation technology have become the core engine of content distribution. Personalized recommendation systems based on collaborative filtering algorithms achieve an exponential improvement in information matching efficiency by analyzing user behavioral data—such as clicks, dwell time, and shares—to generate a pool of customized content [1]. However, this technological logic currently triggers a deep contradiction: the algorithm reinforces the user's existing preferences through a "feedback loop", forming a closed "filter bubble", leading to the marginalisation of high-quality content and the narrowing of user perceptions [2]. Academics have developed multiple research perspectives on this issue: early studies focused on the unidirectional correlation between technological efficiency and information cocoon [3], and in recent years, they have gradually

expanded to the framework of multi-role interaction between technology users and society, and explored the ethical risks of algorithmic bias [4,5].

In this paper, this study adopts the literature review method to systematically sort out the relevant studies on China Knowledge Network (CKN) from four perspectives: technological logic, user behaviour, social impact, and governance strategy. At the theoretical level, this study expands the interdisciplinary research perspective on algorithmic ethics; at the practical level, it offers doctrinal support for the design of platform anti-filter bubble mechanisms and policy formulation, while facilitating the balance between technical efficiency and information diversity.

2. Efficiency advantages and potential risks of algorithmic recommendation technology

2.1. The dialectical relationship between efficiency advantages and potential risks

Algorithmic recommendation technology shows significant advantages in improving information matching efficiency. Empirical research based on Jittery [1] shows that the personalised recommendation mechanism enhances users' perceived usefulness and positively increases their willingness to continue to use, reflecting the technological advantage of algorithms in meeting personalised needs. The platform leverages data mining and machine learning techniques to rapidly capture user preferences and mitigate information retrieval costs [4].

However, this efficiency-driven advantage is accompanied by multiple inherent risks. First of all, in the dimension of information access, the algorithm's data-dependent logic intertwined with the logic of platform traffic [2], not only exacerbate the narrowing of information, making quality content marginal, but also trapping the user in the "preference cage" constructed by their own preferences, a mechanism that is essentially an "efficiency match. This mechanism is essentially in the name of "efficiency matching", compressing the cognitive space of users to access diversified information. Following this, at the level of content ecology, ShakeEn's overlay recommendation mechanism upgrades the "preference cage" into an "echo chamber effect" by strengthening the dissemination of vulgar and homogeneous content [6]: homogeneous information efficiently pushed by algorithms not only leads to users' information predilection, but also intensifies polarisation through the resonance of views between groups. The resonance of opinions among groups exacerbates the polarization phenomenon, giving rise to the paradox of "greater efficiency, narrower cognition." At a deeper level, algorithms alienate the "efficiency advantage" into a cognitive erosion tool with the hidden nature of "legal deception" [3]: they take up users' time through the efficient filling of huge amounts of entertainment content, and systematically dissolve the ability to think in depth, while incidents such as minors imitating dangerous videos expose the algorithm's influence on the user's ability to think in depth, while minors imitating dangerous videos. Incidents such as minors imitating dangerous videos have exposed the deficiency of algorithms in value-oriented screening—the "efficiency-first" design logic ultimately impairs users' cognitive health at the level of content ethics.

2.2. Passive and active roles of users

Users in the algorithmic ecology present the dual characteristics of passive trapped and active resistance. Research has shown [5] that the stereotypical nature of users' categorical labels and interaction lag drives information avoidance behaviours (e.g., turning off the recommendation function or uninstalling the application), reflecting passive resistance. Drawing on Marcuse's theory [3], it is observed that users who rely on algorithms over an extended period may be alienated into

"technological giants," forfeiting their capacity for active information retrieval and succumbing to profound technological dependence.

However, users are not passive recipients: structural equation modelling [4] has shown that media literate "active audiences" can break through the filter bubble through active search and cross-platform information access, counteracting some of the effects of algorithmic closure. Users' awareness of perceived control and privacy concerns drive behavioural avoidance through the mediating effect of psychological resistance [5], which is manifested in active intervention behaviours such as setting privacy permissions and using the "Not Interested" function, demonstrating the users' ability to regulate the impact of algorithms in a proactive manner.

3. Existing collaborative governance paths

Based on the current literature, this study analyses that there have been three main governance frameworks

3.1. Technical optimisation

The "anti-inertia recommendation" algorithm advocates for the mandatory insertion of public issues and knowledge-oriented heterogeneous content into personalized content streams [2], thereby balancing user preferences and information diversity. Therefore, cross-domain recommendation technology [6] is introduced, which extends from food content to food culture, social issues and other related fields, extends the user interest boundary, and reduces label curing.

3.2. User empowerment

Several studies [4,5] have emphasised the importance of enhancing users' media literacy and advocated the development of critical information processing skills, including identifying algorithmic limitations and actively verifying the authenticity of information. In the context of platform function design, it is imperative to provide a "Recommendation Logic Transparency" interface [5] and a "Heterogeneous Content Ratio Setting" tool [2] to enhance users' control over algorithms.

3.3. Policy and ethics

Instrumental rationality has been proven to have a very important role [3], incorporating value rationality into algorithm design, establishing a content value audit mechanism, and prioritising the recommendation of content that is in the public interest. Although the "Regulations on the Administration of Algorithmic Recommendation of Internet Information Services" (2022) have required platforms to provide the option to turn off recommendations, echoing the claim about the "user's right to choose" [6], there is still a need to strengthen the ethical review of algorithms [3], and to build a whole-chain governance system that covers technological research and development, content auditing, and user feedback.

4. Innovative countermeasure strategies and research directions

4.1. Systematic construction of media literacy education

4.1.1. Algorithmic cognition course for teenagers

Include "algorithmic experience lessons" in secondary school IT classes, and let students understand the recommendation logic through interactive games -- for example, using different coloured cards to represent users' interest labels, and simulating how platforms push content based on click records. Referring to the "Critical Thinking Training" proposed by Bu Wei [7], this study sets up a practical session of "Short Video Content Discernment" to allow students to compare the content differences of similar accounts and learn to identify homogeneous information. At the same time, drawing on the "three-stage ability cultivation system" [8], students are trained to identify homogenised content through the "information screening lab" at the junior high school stage, and the "algorithmic ethics debate class" is added at the senior high school stage " to discuss the impact of algorithmic bias on diverse perspectives in the context of "the risk of cognitive solidification among adolescents" as pointed out by Huang Chuxin [9].

4.1.2. Information verification training for middle-aged and elderly people

A "mobile phone information counterfeiting class" can be held in the community to teach people over 50 years old to identify the authenticity of information, which can be done in three ways: "checking whether the publisher is an official account, checking whether the same content is said consistently across multiple platforms, and checking whether the information is outdated at the time it was published." Check. At the same time APP can be supporting the development of "a key to verify the reason for the design [10]. In addition by the community education centre in conjunction with the mainstream media to carry out the "three check learning workshop", to enhance the information discernment ability of the middle-aged and elderly.

4.1.3. Platform intelligent reminder function

The development team of APPs such as Jitterbug and Shutterbug can set up a "30-minute reminder": when the platform detects that the user has continuously watched the same kind of video for half an hour, a floating window automatically pops up to ask "Do you want to take a look at it?" applet, and the click can show the credibility rating of the information source, which can be Combined with "traditional media high trust" original other types?" , and recommend 3 pieces of tech or cultural content. Implement the "algorithmic transparency obligation" [11], safeguard the user's right to know through interpretable design, add a small question mark interaction icon in the lower right corner of the floating window recommendation card, click on it to pop up the layer to show the basis of recommendation in symbolic language, accompanied by a simple description, for example, "because you recently watched food videos. So food culture content is recommended". At the same time, with reference to the empirically supported "hybrid recommendation algorithm" [10], the intelligent mapping mechanism between the main domain of interest and the subdomain of value is constructed, a certain embedding strategy is adopted, and one cross-domain content is inserted after every six personalised contents, and the "expanded content" is prompted by a light grey horizontal bar at the bottom of the content " identity, balancing the intervention of mandatory and user experience. The platform jointly builds a "silver hair expansion content pool" with professional institutions, implements a double audit system of content scientific audit and ageing function

adaptation, and sets a user feedback slider in the recommended floating window to dynamically adjust the matching accuracy of cross-field content with real-time access to the algorithm model, so as to achieve the dynamic balance of "precise intervention without disturbing". " dynamic balance.

4.2. Mechanism innovation of cross-platform collaborative governance

4.2.1. Establishment of head platform content convention

Head platforms such as Jitterbug and Shutterbug can refer to the "bottom line of the proportion of public issues" [9] and jointly sign the "Convention on Content Diversity" to implement content ecological governance in three aspects: firstly, to clarify the bottom line of the supply of knowledge-based content, and to agree that the proportion of science, history and other knowledge-based content in the daily recommended content is not less than 20%, to ensure that the user's information balance of the consumption structure; then establishing a transparent regulatory reporting system, submitting weekly Content Variety Reports to the regulatory authorities, applying the "domain entropy method" [12] to calculate the content diversity index, and presenting the proportion distribution of entertainment and knowledge content in the form of visual charts to provide quantitative bases for regulatory assessment; and then empowering users to In addition, it gives users cross-platform preference management authority, and develops a one-key synchronisation function to "reduce certain recommendations" based on the three-angle process of filing, auditing and chastisement [11], so as to strengthen users' independent control over personalised recommendations. In addition, the convention supporting the establishment of non-compliance constraint mechanism, the platform does not meet the standard of content diversity, with reference to the "public welfare content bottom line" and "algorithmic audit" requirements [12], to reduce the home page recommendation and other penalties, forcing the platform to balance the commercial interests and public values.

4.2.2. Tracking the entire content process

To build a tracking mechanism that covers content production, dissemination, and feedback, platforms are required to generate an "electronic ID card" for each piece of content that contains multiple pieces of information, recording in detail the content's production body, the person responsible for auditing, the scope of the target audience, the frequency of pushing, and user interaction data (likes and reports). This information is stored in a system similar to a "network ledger", forming a traceable and untamperable content dissemination log. In the event of problematic content, the system allows regulators to quickly locate specific aspects of the algorithm's operation, determine whether there is over-recommendation of controversial topics, low-quality content, etc. and enforce accountability. This mechanism integrates the concepts of "data traceability" and "algorithm log uplinking" [12,11], realizes the transparent supervision of the whole process of content dissemination by technical means, provides a traceability tool for algorithm governance, and effectively improves the efficiency of disposal and the accuracy of responsibility determination of the offending content. Disposal efficiency and the accuracy of responsibility determination.

5. Conclusion

This study systematically analyzes the phenomenon of filtering bubbles caused by algorithmic recommendations on short video platforms, and finds that algorithms reinforce user preferences

through "data dependence + commercial logic", leading to narrowing of information and cognitive solidification. Although users can breach certain forms of closure through active searching, effective governance necessitates the synergy of technological optimization, media literacy education, and policy regulation.

The study expands the perspective of algorithmic ethics through the four dimensions of social governance of technology users, and verifies the dual roles of users' "passivity-activity". The proposed age-specific media literacy education (e.g., algorithm awareness courses for teenagers, information verification training for middle-aged and old-aged people) and cross-platform governance mechanisms (content conventions, full-process tracking) provide practical paths for balancing technological efficiency and information diversity.

There are gaps in the existing research on cross-generational user differences, the impact of AIGC technology and long-term social effects, etc. In the future, emphasis should be placed on the construction of algorithmic ethical frameworks, the assessment of information diversity in the AIGC era, and cross-cultural long-term tracking research, so as to provide prospective guidance for the governance of the digital information ecosystem.

References

- [1] Tan Yuelong. Research on the Impact of Short Video Personalised Recommendation Service on Users' Willingness to Continuously Use [J]. Commercial Exhibition Economy, 2023 (23): 118-121.
- [2] Shen Lin. Formation Mechanism and Governance Approach of Short Video Filtering Recommendation Algorithm in the Age of Digital Intelligence [J]. Journalism University, 2023 (5): 42-55.
- [3] Liu JX. Information cocoon and value illusion: the death of algorithmic recommendation of short video [J]. Changjiang Literary Review, 2020 (6): 45-49.
- [4] Guo Rong. Empirical study on information cocoon effect of algorithmic recommendation short video platform [D]. Shandong University, 2022.
- [5] Chen Liying. Study on the Influence Factors of Information Avoidance of Users on Personalised Recommendations of Short Video Platforms [D]. Northeast Normal University, 2023.
- [6] Xue Yaoyun. Research on short video "filter bubble" problem under algorithmic recommendation mechanism -- taking Jittery as an example [J]. New Media Research, 2019, 5 (14): 21-22.
- [7] Bu Wei. On the meaning, content and method of media education [J]. Modern Communication, 1994 (3): 1-9.
- [8] Zhang C.A., Shen G.L.. Media Literacy: An Urgent Subject for Education for All [J]. Journalist, 2004 (5): 31-33.
- [9] Huang Chuxin, Licence. The Development Logic and Governance Path of "Information Cocoon" in the Digital Era [J]. New Observations on Media Convergence, 2022 (3): 11-14.
- [10] Yu Guoming, Fang Ke Ren. Does Algorithmic Content Pushing Lead to Information Cocoon? [J]. Shandong Social Science, 2020 (11): 170-175.
- [11] Zhang Linghan. The construction of algorithmic accountability for online platform regulation [J]. Oriental Law, 2021 (3): 22-40.
- [12] Liu Youhua. Research on algorithmic bias and its regulatory path [J]. Journal of Law, 2019, 40 (6): 55-66.