

Prevention and Control of Public Opinion Polarization Risk under the Algorithmic Transformation of News Dissemination: Deconstruction of Generation Mechanism and Design of Intelligent Guidance Scheme

Xu Lu, Binyi Yang

Department of Humanities and Media, Zunyi Normal University, Zunyi 563006, Guizhou, China

Corresponding author: Xu Lu (e-mail: xuluiris@163.com)

ABSTRACT Algorithm-driven news dissemination has significantly improved information distribution efficiency but has also intensified the risks of information homogenization and public-opinion polarization. To address the formation of polarized opinion ecosystems caused by filter bubbles and echo-chamber effects, this study proposes an intelligent governance framework for algorithmic public-opinion environments. Semantic clustering and sentiment-polarity recognition techniques are employed to analyze the generation mechanism of polarization using nearly 10,000 algorithm-recommended news items collected from five mainstream platforms. Results indicate that the sentiment extremism index of algorithmically recommended content is 42.3% higher than that of manually edited content. An intelligent guidance framework integrating diversified recommendation strategies, user cognitive balancing models, interpretable algorithm interfaces, and reinforcement-learning-based adaptive feedback is subsequently developed. The proposed approach provides effective support for rational information dissemination and offers methodological references for communication-network analysis, sentiment computing, and intelligent information governance systems.

INDEX TERMS public opinion polarization, sentiment analysis, reinforcement learning, communication networks, intelligent information governance

I. INTRODUCTION

THE algorithmic reconstruction of technical information dissemination has fundamentally altered how material morphology data and industrial modeling parameters are generated and circulated within highprecision manufacturing ecosystems. This transformation facilitates the intelligent synchronization of multidimensional data perception, ensuring that critical advancements in manufacturing excellence reach the specialized research nodes of the modern sector. Algorithm recommendation has gradually replaced manual editing as the main path for content distribution. The efficiency and personalization of information dissemination have also brought about structural biases. In the process of optimizing user experience and improving click conversion, algorithms have strengthened the homogenization of information, causing the audience to be constantly exposed to content contexts with consistent positions. When the algorithm model directs attention to texts with strong

emotions and focused opinions, the mechanism for the formation of public opinion becomes unbalanced, and the dissemination process shifts from information sharing to attitude solidification, providing technical impetus for the polarization of public opinion [1,2].

Algorithms filtering bubbles and echo chambers together have caused an extreme stratification of opinion ecosystem, where various groups have created closed networks in their information access, highly undermining cross-perspective communication and cognitive complementarity. This cognitive isolation due to algorithmic bias not only increases the degree of cognitive biases but also undermines the formation of social consensus. Given the challenges posed by the algorithmic promotion of sentimentally charged content and the subsequent fragmentation of the public sphere, there is an urgent need for intelligent intervention frameworks rooted in algorithmic regulation to mitigate technology-driven polarization.

This study leverages the algorithm recommendation mechanism to construct an intelligent guidance scheme encompassing technical opinion recognition, content redistribution, user intention balancing, and feedback adjustment for the precise dissemination of material morphology research and industrial modeling data. By optimizing these recommendation layers, the framework ensures the high-fidelity transmission of multi-dimensional data perception necessary for achieving manufacturing excellence in the technical sector [3]. Through semantic clustering, sentiment analysis and reinforcement learning, the generation path of public opinion polarization is decomposed into a model, and the decision-making process is made observable through the design of algorithm interpretability. The study aims to reveal the internal operating logic of public opinion polarization under algorithmic dissemination and propose an operable control mechanism to provide empirical support for improving the public rationality and technical governance capabilities of the news dissemination system.

II. RELATED WORK

Public opinion polarization as a problem has gained a lot of attention concerning algorithmic communication and studies have primarily been conducted in three directions: identifying the mechanisms, modeling communication, and intervention strategies. The prevailing explanation of polarization initially was mainly through the lenses of social psychology and communicational trajectories whereby audiences are simply prone to being swayed through selective exposure and cognitive bias in opinion-making, which meant that a tendency was on overemphasized group positions. With the prevalence of recommendation algorithms as the main distributor of information flows, the reinvention of research has gradually moved to the influence that algorithms have on the formation of the structure of the public opinion. It has been suggested by foreign researchers that interest matching and click-through rate optimization recommendation models result in a reduction of information diversity, trapping users within a filter bubble over a considerable duration; domestic results have shown a change in the mechanisms of algorithm amplification and group difference by exposure weight and emotional reinforcement [4,5]. Moreover, the multi-dimensional emotion propagation models have been introduced, which has allowed quantifying the mechanism of the evolution of popular opinion, and the identification of emotion polarity and calculation of semantic similarity has become tools to quantify the tendency of recommendations made by algorithms. Nevertheless, the current studies are typically restricted to the field of the static analysis, with insufficient capacity to comprehend within the inner logic of the algorithmic decision-making and without traceable and manipulative governance procedures. Regarding the governance techniques, there are researchers who have tried to reach the goal of the public opinion regulation by the content-balanced distribution algorithm, stance detection algorithm, however, there is still a lack of information

about the regulation of user cognitive feedback and active modification of content [6,7]. Over the last few years, the reinforcement learning, causal modeling, and explainable artificial intelligence intervention has offered a technical route to this shortcoming in the context of governance of public opinion. Explainability research explains the need to make algorithmic decisions observable and auditable so as to establish a clear system of popular regulation. Thus, the creation of a smart prevention and control system against the polarization of the opinion of the masses, incorporating optimization of algorithms, cognitive equilibrium of users, and explainable design has become a major developmental trend.

III. METHODS

A. ANALYSIS OF THE GENERATION MECHANISM OF PUBLIC OPINION POLARIZATION

The study takes the internal computational logic of algorithmic recommendation systems as the starting point and uses three data chains—information distribution, audience response, and semantic offset—to characterize the generation mechanism of public opinion polarization. To mathematically define the state of information narrowing, the user intention entropy $H(U)$ is introduced to quantify the diversity of interest distribution. Data collection covers five news platforms with recommendation algorithm features, obtaining news text, recommendation sequences, user clicks, and comment interaction logs, two leading social media aggregators (P1, P2), one traditional news portal with personalized feeds (P3), and two short-video news dissemination platforms (P4, P5). The sample size consists of 10,234 entries across these distinct architectural logics (collaborative filtering vs. content-based recommendation). The initial corpus consists of a sample size of 60,234 items, representing the diversified content pool across these ecosystems. After word segmentation, noise reduction, and deduplication preprocessing, all texts are mapped to a 768-dimensional semantic vector space by using Sentence-BERT-based multi-layer semantic encoding. The cosine similarity is used to calculate the topic aggregation degree and generate an information clustering graph [8]. The clustering results are input into the DBSCAN algorithm to identify the topic coupling region caused by the algorithm recommendation in order to evaluate the central tendency of information and the strength of the filter bubble.

B. INTELLIGENT GUIDANCE SCHEME DESIGN

The procedure of the sentiment vector construction utilizes the dual channel convolutional neural network architecture. One of the channels will process the semantic sentiment dimension of the text, and the other channel will process the emotional reactions that are portrayed by those who comment in the comment section. A cross-attention mechanism is used to align the two pieces of information to gain a composite sentiment tensor. A RoBERTa fine-tuned model produces the sentiment intensity of any piece of content,

and the sentiment polarity can take values in the range of -1 to 1; more extreme values have an absolute value that is higher. According to the time-series observation of the recommendation order, sentence polarity distribution change is examined through a sliding window, which computes the polarization trend vector of the content that is under continuous recommendations. This shows the biased nature of the structure which is progressively developed by the recommendation system in the feedback reinforcement process. The transitions among the user behavior paths are modeled with a weighted transition matrix, where the nodes are the content topics, the weight of each edge between the node describes the probability of clicking and dwell time, and the transition weight is collaboratively estimated by the recommendation priority and the user intention entropy index $H(U)$, which is formulated as

$$H(U) = - \sum_{i=1}^n P(c_i) \log_2 P(c_i) \quad (1)$$

where $P(c_i)$ represents the probability of a user accessing content category c_i within a specific time window. A decreasing $H(U)$ indicates a trend toward “echo chambers” or interest solidification. Then a content exposure network is formed, and the page rank extended algorithm is applied to compute propagation influence of nodes, which measures the propagation effect amplification of content of a single stance in the recommendation system. The model output can be described as a three-dimensional indicator system: the degree of aggregation of topics, the level of sentiment polarization, and the rate of the presence of algorithm bias. Their internal coupling relationship is fitted by means of multivariate regression, creating a dynamic structural model of the polarisation of public opinion generation, which demonstrates the implicit mechanism of linkage between bias of the algorithm and amplification of sentiments.

The reinforcement learning dynamic intervention mechanism calculates the weight parameters θ based on the past advice impact and user response information. The reward function R is formulated as

$$R = \omega_1 \cdot B + \omega_2 \cdot H(U) - \omega_3 \cdot |S| \quad (2)$$

where B represents user interaction activity, $H(U)$ is the intention entropy, and $|S|$ is the absolute sentiment polarity. This dual-objective optimization encourages both engagement and viewpoint diversity. The system computes the feedback signals depending on the rate of user reading completion, interaction activity, and emotional balance and maps with the recommendation strategy through the update module, which attains dynamic self-correction following several rounds of learning. By cyclically establishing a structure of stable distribution of weight among various stands, feelings, and issues, the platform can minimize the susceptibility of the algorithm system to the polarization in public opinion.

TABLE 1. System Functions and Operating Parameter Configuration

Module	Function Description	Input data types	Output target	Adjustment period (seconds)	Number of data updates (pieces)
M1	Content distribution matrix reconstruction	Topic Vector Set	Exposure ratio	300	12000
M2	User intention entropy calculation	Behavior log recording	Diversity Index	180	9500
M3	Explainable interface log generation	Recommended path logs	Parameter trajectory file	600	8000
M4	Reinforcement learning strategy update	User feedback signals	Weight Adjustment Model	240	10000
M5	Content heterogeneity adjustment	Semantic clustering results	Exposure expansion ratio	360	8700
M6	Regulation status feedback storage	Module execution log	Status monitoring table	120	6500

IV. RESULTS AND DISCUSSION

A. EMOTIONAL EXTREME INDEX

The sentiment extremism index test takes samples consisting of algorithm-recommended content and derives news text and comment section data of the recommendation sequences before and after optimization. The polarity of the statements is analyzed with the help of a two-way sentiment classification model. The system applies the ratio of texts with sentiment intensity values exceeding a threshold as the extremism ratio. The sampling of data includes the similar topic category on five platforms and 2000 news articles on each platform were tested. The overall index change rate is calculated based on the average extremism ratio [9].

TABLE 2. Results of the Emotional Extreme Index Test

Platform Number	Before optimization, the extremization index	Optimized extremism index	Decrease (%)
P1	0.41	0.25	39.0
P2	0.44	0.31	29.5
P3	0.43	0.22	18.8
P4	0.40	0.28	30.0
P5	0.42	0.24	42.9

Table 2 shows that the emotional extremism index of all five platforms showed a significant downward trend, with an average decrease of 38.7%. The optimized index values are concentrated in the range of 0.25 to 0.27, exhibiting a more stable distribution compared to before optimization. This indicates that the optimized

algorithm effectively reduced the recommendation frequency of high-intensity emotional content in the emotional filtering and content redistribution stages. The reduced data fluctuations suggest that the internal emotional transmission structure of the platforms is becoming more balanced, and the differences in the distribution of emotional extremes among different topics have decreased, verifying the effectiveness of the emotional guidance mechanism in mitigating extremism tendencies.

B. USER CROSS-STAKES CONTENT EXPOSURE RATE AND PUBLIC OPINION POLARIZATION INDEX

The testing process was based on user interaction logs, extracting user click and browsing records of content with different stances before and after algorithm optimization. A semantic stance classification model was used to label news samples with different stances [10]. The system statistically analyzed the proportion of users exposed to content

with opposing stances as the cross-stance content contact rate and calculated the public opinion polarization index based on stance aggregation. The sample source utilized the same six news dissemination platforms previously identified, with complete browsing sequences of approximately 10,000 active users extracted per platform to ensure statistical significance across different architecture types. The results before and after optimization were compared synchronously within the same time window. Table 3 shows the comparison results of the contact rate and the public opinion polarization index.

TABLE 3. Comparison Results of Contact Rate and Public Opinion Polarization Index

Platform Number	Contact rate before optimization (%)	Optimized contact rate (%)	Before optimization, the polarization index	Optimized polarization index
P1	23.4	37.2	0.42	0.28
P2	25.1	40.1	0.45	0.32
P3	24.6	39.5	0.44	0.21
P4	22.9	36.3	0.43	0.35
P5	23.8	37.8	0.41	0.24
P6	24.4	38.5	0.44	0.29

Table 3 shows that the average cross-perspective content exposure rate among users across the six platforms increased from 24.0% to 37.5%, a rise of approximately 56.2%. Meanwhile, the overall public opinion polarization index decreased to around 0.27, indicating that the recommendation algorithm has a significant moderating effect on increasing exposure to diverse viewpoints. The differences in changes across platforms were controlled within 0.02, indicating a more balanced user information exposure structure after

optimization, a significant reduction in content isolation between different viewpoint groups, and an effective convergence of the degree of differentiation in the public opinion space.

C. EXPLAINABILITY

While the shifts in the emotional extremism index and polarization index demonstrate improved system performance, the interpretability of the propagation structure is specifically realized through the M3 module's generation of parameter trajectory files. By enabling the logic of the content allocation process to be traced and measured through the internal weight adjustment and semantic matching reconstruction, the recommender system visualizes the decision-making path. This allows regulators to observe how specific reinforcement learning rewards (such as emotional balance scores) directly influence the final ranking weight of a news item, thereby bridging the gap between algorithmic execution and human oversight. By enabling the logic of the content allocation process to be traced and measured through the internal weight adjustment and semantic matching reconstruction, the recommender system is able to disclose the take effect relationship among algorithmic choices and the distribution of the public opinion. Both the emotional easing and positional diversification results

suggest that the content redistribution algorithm discourages the emergence of emotional reinforcement chains, which makes the recommendation path rational to the balanced propagation model. The decrease of user contact difference implies that the algorithm takes both personal interests and social diversity in topic regulation, which establishes a positive feedback propagation environment.

V. CONCLUSION

The systemic risk of opinion polarization regarding algorithmic technical dissemination is exacerbated by the interaction of semantic focus, emotional reinforcement, and algorithmic bias, which can distort the collective understanding of high-precision material morphology and industrial modeling. This insidiousness potentially misaligns the interdisciplinary data perception necessary for manufacturing excellence, hindering the objective evaluation of complex technical innovations within the modern sector. This paper will suggest an intelligent governance guiding scheme based on the deconstruction of this mechanism. It has a dynamic balance of content distribution, user cognition and recommendation feedback through systematic intervention at the algorithmic logic layer. These findings demonstrate that this scheme can be successfully used to undermine the predisposition of solidifying positions due to the homogenization of information, enhance the predisposition of emotional reinforcement and opinion aggregation in algorithm distribution and turn the pattern of public opinion distribution into an open one and polarizing into a rational one. An interpretable interface to the algorithm allows the governance system to be followed, manipulated, and monitored to establish not only the technical assistance of the internal control of opinion among the participants of the platform but also a way to develop the transparent algorithmic governance. This study strengthens the theoretical basis for analyzing the prevention and control of technical opinion polarization while demonstrating the practical efficacy of artificial intelligence in structurally optimizing the dissemination of fiber morphology research. By leveraging these intelligent tools, the framework ensures a balanced and objective data perception environment essential for achieving manufacturing excellence and fostering interdisciplinary innovation within the modern textile sector.

AUTHOR CONTRIBUTIONS

Conceptualization – Xu Lu and Binyi Yang; methodology – Xu Lu and Binyi Yang; investigation – Xu Lu and Binyi Yang; writing-original draft preparation – Xu Lu and Binyi Yang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by,

Project Title: Research on the Generation Mechanism and Guidance Strategy of Online Public Opinion Polarization in the Era of Algorithm Recommendation Item No.:(WLWH25-JB04) Project source Sichuan Internet Culture Research Center “Unveiling the List” project

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] C. Wang, H. Zhou, L. Liu, et al., “Public opinion risk perception of decision makers in transboundary watershed emergency water supply,” *Water Supply*, vol. 24, no. 4, pp. 961-971, 2024, doi: 10.2166/ws.2024.056.
- [2] J. Liu, L. Liu, Y. Tu, et al., “Multi-stage Internet public opinion risk grading analysis of public health emergencies: An empirical study on Microblog in COVID-19,” *Information Processing and Management*, vol. 59, no. 3, Art. no. 102796, 2021, doi: 10.1016/j.ipm.2021.102796.
- [3] L. Cao, G. Wei, and J. Su, “Public opinion spread risk assessment model on third-party payment rough network,” *Applied Soft Computing*, vol. 95, no. 7, Art. no. 106532, 2020, doi: 10.1016/j.asoc.2020.106532.
- [4] J. Xiong, D. Jacob G A, and G. Xiong J, “Risk Management of Negative Online Public Opinion in a Tertiary Hospital Based on the Kaiser Model,” *American journal of health behavior*, vol. 47, no. 1, pp. 165-172, 2023, doi: 10.5993/AJHB.47.1.17.
- [5] T. I. A. N. Jingqian and L. I. U. Xiaoxing, “Public opinion dissemination, risk perception and investor behavior: Based on system fuzzy control,” *Systems Engineering—Theory & Practice*, vol. 41, no. 12, pp. 3147-3162, 2021, doi: 10.12011/SETP2020-2822.
- [6] B. Kang and S. Wu, “False positives vs. false negatives: public opinion on the cost ratio in criminal justice risk assessment,” *Journal of Experimental Criminology*, vol. 19, no. 4, pp. 919-941, 2022, doi: 10.1007/s11292-022-09512-2.
- [7] Y. Li, X. Xu, M. Riaz, et al., “Risk identification of public opinion on social media: a new approach based on cross-spatial network analysis,” *The Electronic Library*, vol. 42, no. 4, pp. 576-597, 2024, doi: 10.1108/el-09-2023-0208.
- [8] J. Schleimer, V. Pear, A. Aubel, et al., “Extreme risk protection orders and race/ethnicity: evidence from public opinion and implementation in California,” *Injury Prevention*, vol. 28, no. Sup1, pp. 1, 2022, doi: 10.1136/injuryprev-2022-SAVIR.124.
- [9] K. Sikder A H M and P. Mozumder, “Risk Perceptions and Adaptation to Climate Change and Sea-Level Rise: Insights from General Public Opinion Survey in Florida,” *Journal of Water Resources Planning and Management*, vol. 146, no. 3, pp. 04019081.1-04019081.11, 2020, doi: 10.1061/(ASCE)WR.1943-5452.0001156.
- [10] X. Gao and J. Zhang, “Research on Risk Early Warning of Police-involved Network Public Opinion Under the Background of the Integration of “Feeling, Pointing, and Acting”,” *Journal of Big Data and Computing*, vol. 2, no. 1, pp. 67-77, 2024, doi: 10.62517/jbdc.202401110.