

Review of the Research on the Relationship Between Algorithmic News Recommendation and Information Cocoons

Zhiyu Kong^{1*} Xiaoru Zhang² Ruilin Wang³

¹School of Languages and Communication Studies, Beijing Jiaotong University, Haidian, Beijing 10044, China

²School of Language and Communication Studies, Beijing Jiaotong University, Haidian, Beijing 10044, China

³School of Computer and Information Technology, Beijing Jiaotong University, Haidian, Beijing 10044, China

*Corresponding author. Email: 19322045@bjtu.edu.cn

ABSTRACT

In recent years, the rapid growth of algorithmic news recommendation has brought huge commercial value. Considering the current development and application of algorithmic news recommendation, many experts and scholars worry that it will lead to information cocoons, which will destroy the value system and social consensus of society, influence sound social development, even enfeeble users' self-consideration capacity and be reduced to a tool for algorithms and capital. The relationship between algorithmic news recommendation and information cocoon has been hotly debated as it may do harm. In this paper, the representative literature on the relationship between algorithmic news recommendation technology and information cocoons is reviewed and straightened out, making the relationship between algorithmic news recommendation and information cocoons clear gradually. Finally, research on the relationship between algorithmic news recommendation and information cocoons is summed up, considered and predicted.

Keywords: Information cocoons (IC), Algorithmic news recommendation (ANR), Social consensus, News, Society

1. INTRODUCTION

The news and information industry is improving rapidly and now it has a large number of users, influencing every user to obtain information and establish values. ANR is just the driver of the rise of the news and information industry as businesses use ANR to reap substantial commercial benefits and an increasing number of users become dependent on the ANR distribution mechanism and less able to do independent information retrieval and thinking as well. During the time businesses develop the ANR mechanism to reap greater commercial benefits, they are very likely to harm users' interests and depart from social values, which has caused concern among academics and industry at home and abroad from the perspective of social values and journalism ethics. They suggest that the current application of the ANR mechanism has hidden dangers and risks among which the most worrying one is that ANR may lead to IC. It can further cause a decline in the quality and diversity

of public discussions, polarization of opinions, fragmentation of the public sphere, the absence of values, etc, causing unnecessary harm to human society. In this paper, representative research on topics of "ANR" and "IC" is straightened out and reviewed to clarify the current status of research on the relationship between ANR and IC, and to propose thoughts and prospects for future related research.

2. CONCEPT SOURCE AND REVIEW

2.1. Concept Source

The concept of IC was first put forward in *Infotopia: how many minds produce knowledge* written by American Professor Sunstein and published in 2008, and whereby Sunstein criticizes ANR-induced negative effects on society. That's to say, according to Sunstein, in the Web 2.0 network environment internet businesses develop personalized information recommendation for users so that they will be exposed only to information of

their own choice as well as satisfying and pleasurable information, and lose their ability of profound thinking [1].

However, Sunstein also considers that real-life information doesn't confine to Internet information filtered by ANR, and that users aren't Web-users living in the Web world completely. He once admitted that "In the next 10 years or more, my personal argument should not be regarded as an empirical argument about the personal characteristics of information selection." Thus, the concept of IC should be classified as a hypothesis rather than a proven theory.

With the popularity of the Internet, Internet information recommended by ANR has become the main information source for many netizens. Owing to the limited controllability and predictability of ANR, the gradual complication of the algorithm as well as the purpose of commercial confidentiality, the ANR mechanism of businesses remains a "black box" for the public and scholars now [2]. Besides, considering mixed Internet information and limited ability of value judgement of ANR, many experts and scholars worry that social fragmentation and conflicts would intensify, and considerable attention has been paid to the relationship between ANR and IC.

2.2. Review Description

Reviewed literature is obtained from CNKI, CMCC, and Web of science with Information cocoons, recommender systems, echo chamber, algorithm news recommendation and filter bubbles as subject terms. After de-duplicating, screening and combing collected literature, this paper divides the research review on the relationship between ANR and IC accordingly into two parts that ANR leads to IC and ANR doesn't lead to IC, totally 42 typical papers reviewed.

3. ANR LEAD TO IC

Current research in favor of ANR leading to IC is based on two arguments: ANR leads to information narrowing, ANR reinforces selective-exposure behavior in users. The difference between 2 arguments is that, the former one claims ANR as the fundamental factor of IC while the latter one claims selective-exposure behavior in users as the fundamental factor of IC, and ANR the cause factor at play.

3.1. ANR Lead to Narrowing Information

Currently, the ANR principles of content-based filtering and collaborative filtering are commonly used [3]. With human assistance, ANR keeps learning, optimizing recommendation strategies or changing the algorithm structure, finally being used in production [4]. Andrejevic finds that for the ANR mechanism of

businesses facial relevance is the main criteria for judging. That's because current ANR can't understand the hidden information of context, pictures and so on, and also can't offer overall information recommendation according to users' real-life requirements, such as their living environment [5]. Unlike in the traditional news era, current ANR actively meets user's information needs to achieve the situation that information recommendation varies from person to person as nowadays it's what users like to read that counts.

In the long run, users may gradually receive fixed types of news, causing information narrowing and finally IC. Pariser points out vividly that under the influence of ANR, users are immersed in the environment created by ANR and have difficulties in seeing the whole information environment. Influenced by ANR which filters information, users often can but get narrowed information and have difficulties in reaching different ideas when browsing information, like politics, ethics, social hot spots, easily getting into the opinion-homogenized IC [6]. After entering such IC, users will constantly receive narrowed information, unwilling to take other opinions into account, which in turn turns discussions of opinions into a zero-sum game, and even causes serious social fragmentation.

However, the above inferences are likely to be challenged by subjective inferences. Through quantitative research, Chaney and Lunardi verify the ability of ANR to filter information and find that it is difficult for ANR to push diverse information [7]. To make the experimental results more accurate, Chaney and Lunardi develop the ANR mechanism for specific experiments, and add multidimensional metrics, such as diversified push, novel push, and causal push. As a result, ANR is still found to lead to IC [8].

3.2. ANR Technique Reinforce Selective Exposure

The earliest theory related to information narrowing is the selective exposure theory proposed by Lazarsfeld in 1940. Many scholars believe that IC should be regarded as the result of the reinforcement of selective-exposure mind in users, while ANR only reinforces selective-exposure behavior in users, and is not a key factor in the formation of IC [9].

According to users' behaviors and purposes of obtaining information with the Internet, many scholars have argued that ANR reinforces selective-exposure behavior, and leads to IC. Bozdag Passe, et al. argue that the public is more inclined to browse content they identify with and avoid information that holds opinions different to theirs, willing to remain in their comfort zone, and thus keep such behaviors in the Internet environment [10]. From the political context of the

United States, Goldie et al. (2014) argue that people constantly reinforce their preconceived notions when browsing information recommended through news algorithms on Twitter [11]. With bibliometrics and social media analysis, they also find that users' selective exposure may lead to IC. Peng Lan analyzes the filtering mechanisms of platforms, social networks, algorithms and individual self-filtering in users' access to information. He concludes that individual self-filtering plays a decisive role among these four filtering mechanisms while the filtering effect of ANR is based on reinforcing users' selective minds, thus completing the argumentation of users' selective exposure leading to IC. Besides, Goldie et al. analyze the data of users' remarks in Twitter and in-depth interviews, and find that selective exposure of Internet users is the key factor in the formation of IC.

4. ANR NOT LEAD TO IC

In Journalism and Communication, research on the topic of ANR not leading to IC performs its argumentation generally from two points, including the complex and variable actual information environment, the inherent existence of IC.

4.1. Complex and Variable Actual Information Environment

Early in 2009, Garrett suggested that in the actual information environment people are influenced by the social circles they are in and the social groups they belong to receive information, but this hasn't been verified by empirical research [12]. To complete the argument, Yang Guang suggests that specific factors that affect information obtainment can be judged out by comprehensive analysis of the multiple ways people obtain information and how they filter them [13].

In order to cope with the impact of the complex and variable actual information environment on arguing, many scholars have taken an alternative approach and finished argument of ANR not leading to IC from the theoretical perspective of psychology, social network analysis and information ecology and others.

Barber et al. analyze the relationship between ANR and IC from the perspective of psychology. They consider the diverse accesses for Internet users to information, their specific preferences and habits, and find that ANR leads to IC with finished psychological analysis [14]. LI Weidong and Peng Jing find in recent research that social networks are possible important factors in the formation of IC. After verifying the key factors and effects, they propose that ANR and IC are not strongly connected and that IC effects can be artificially controlled [15]. Based on Information Ecology theory and from the perspective of human, information, information environment and information technological interaction, Duan Hui constructs a

research model of the formation mechanism of IC in Internet users, and proposes that the value perceptions of Internet users themselves are the main factor in the formation process of IC [16].

4.2. Inherent Existence of IC

Another possibility of IC is that IC may still exist as Internet users leave the Internet, and only by comparing the formation mechanism of the IC in real life with that of the Internet world can ANR be confirmed to lead to ANR [17]. For humans, algorithms create a "black box world", and it is difficult for them to understand and anticipate the behavior mechanism of complex algorithms. Many engineers admit that when maintaining the existing algorithms of businesses, they often don't know what it means to implement a certain function or meet certain needs, so today's research on algorithms can easily fall into the trap of sci-fi utopian thinking [18]. From a philosophical point of view and in the context of real life, IC is found to be the destiny of individuals, but it is only because of the popularity of algorithms that IC has become a public issue, amplified by the progress of media technology [19]. As ANR shapes users' news-using habits, their behaviors also influence algorithms in learning. Their relationship emerges as a two-way interactive relationship and users' media literacy has become the main factor in changing the ANR mechanism. Through the combination of machine learning and social network analysis, Elanor (2019) et al. find that the position of political information received by users in Twitter mainly depends on users' pre-existing beliefs, media behavior, and media literacy, etc [20].

5. CONCLUSION

This proactive research is inextricably linked to current media ecology, social harmonious and healthy development. Current research should remain aware of simplifying the relationship between ANR and IC, and dig into it through the commercial benefits of ANR and user's attitudes toward using it. The relationship question cannot be explored by simplification under the technological determinism sight.

The present concept of IC is still a hypothesis rather than a proven and recognized theory. Insights from academia to the question of whether ANR leads to IC are diverse, and academia tends to explore the question after understanding the operation principle and mechanism of ANR. A minority of research tends to explore the question through logical arguments.

In addition, the exploration of the relationship between ANR and IC is actually also the exploration of the game between algorithms and human society. Continuing vigilance against algorithms gradually gaining subjective power from human society to control

humans to some extent is necessary. However, as a new media technology, ANR can be more viewed as part of a new kind of social relations, or new spiritual ethos and development opportunities. Thus, algorithms should be gradually dissected as research gradually revolves and fully discusses the position of algorithms in the current or future social context so as to establish correct algorithmic values. Human society must master subjectivity in the game with algorithms, and get a boost with the aid of new technologies like algorithms.

ACKNOWLEDGMENTS

Thanks very much to Professor Chen Jie and his support for this research, and also to The School of Languages and Communication Studies of Beijing Jiaotong University for their attention to this research. Without the support of Professor Chen Jie and the encouragement of the school, this research would have been difficult to complete.

REFERENCES

- [1] Lan Peng. The multiple factors leading to the information cocoon and the path of "breaking the cocoon"[J]. *Journalism and Mass Communication Monthly*, 2020(01): 30-38+73.
- [2] Nechushtai E, Lewis S C. What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations[J]. *Computers in Human Behavior*, 2019, 90: 298-307. DOI: <https://doi.org/10.1016/j.chb.2018.07.043>
- [3] Schedl M, Knees P, McFee B, et al. *Recommender systems handbook[M]*//*Music Recommender Systems*. Springer, 2015.
- [4] Ge M, Gedikli F, Jannach D. Placing High-Diversity Items in Top-N Recommendation Lists[C]//*ITWP@ IJCAI*. 2011.
- [5] Andrejevic M. *Infoglut: How too much information is changing the way we think and know[M]*. Routledge, 2013. DOI: https://doi.org/10.4324/9780203075319://doi.org/10.1007/978-3-540-30494-4_16
- [6] Liu P, Shivaram K, Culotta A, et al. The Interaction between Political Typology and Filter Bubbles in News Recommendation Algorithms [C]//*Proceedings of the Web Conference 2021*. 2021: 3791-3801. DOI: <https://dl.acm.org/doi/abs/10.1145/3442381.3450113>
- [7] J.B. Chaney, B.M. Stewart, B.E. Engelhardt, How algorithmic confounding in recommendation systems increases homogeneity and decreases utility, in: *Proceedings of the 12th ACM Conference on Recommender Systems*, in: *RecSys '18*, ACM, New York, NY, USA, 2018, pp. 224–232. DOI: <https://doi.org/10.1145/3240323.3240370>
- [8] Li Lunardi G M, Machado G M, Maran V, et al. Ametric for Filter Bubble measurement in recommender algorithms considering the news domain[J]. *Applied Soft Computing*, 2020, 97: 106771. DOI: <https://doi.org/10.1016/j.asoc.2020.106771>
- [9] Cho J, Ahmed S, Hilbert M, et al. Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization[J]. *Journal of Broadcasting & Electronic Media*, 2020, 64(2): 150-172. DOI: <https://doi.org/10.1080/08838151.2020.1757365>
- [10] Bozdag E, Gao Q, Houben G J, et al. Does offline political segregation affect the filter bubble? An empirical analysis of information diversity for Dutch and Turkish Twitter users[J]. *Computers in human behavior*, 2014, 41: 405-415. DOI: <http://dx.doi.org/10.1016/j.chb.2014.05.028>
- [11] Goldie D, Linick M, Jabbar H, et al. Using bibliometric and social media analyses to explore the “echo chamber” hypothesis[J]. *Educational Policy*, 2014, 28(2): 281-305. DOI: <https://doi.org/10.1016/j.chb.2018.07.043>
- [12] Del Vicario M, Bessi A, Zollo F, et al. The spreading of misinformation online[J]. *Proceedings of the National Academy of Sciences*, 2016, 113(3): 554-559. DOI: <https://doi.org/10.1073/pnas.1517441113>
- [13] Guang Yang, Jialing Yu. Information Visibility, user initiative and Information Cocoon effect of News Algorithm Recommendation: A Perspective of algorithm and User interaction[J]. *Journalism Research*, 2020(02): 102-118+123. DOI: CNKI: SUN: XWDX.0.2020-02-012.
- [14] Barberá P, Jost J T, Nagler J, et al. Tweeting from left to right: Is online political communication more than an echo chamber?[J]. *Psychological science*, 2015, 26(10): 1531-1542. DOI: <https://doi.org/10.1177/0956797615594620>
- [15] Weidong Li, Jing Peng. Simulation analysis of echo chamber effect of information transmission on social network platform[J]. *Modern Communication (Journal of Communication University of China)*, 2019, 41(04): 142-149. DOI: CNKI: SUN: XDCB.0.2019-04-024.

- [16] Nguyen T T, Hui P M, Harper F M, et al. Exploring the filter bubble: the effect of using recommender systems on content diversity [C]//Proceedings of the 23rd international conference on World wide web. 2014: 677-686.DOI:<https://doi.org/10.1145/2566486.2568012>
- [17] Savolainen R. Filtering and withdrawing: strategies for coping with information overload in everyday contexts[J]. *Journal of information Science*, 2007, 33(5): 611-621.DOI:<https://doi.org/10.1177%2F0165551506077418>
- [18] Rubel A. *The Black Box Society: The Secret Algorithms that Control Money and Information*, by Frank Pasquale. Cambridge: Harvard University Press, 2015. 320 pp. ISBN 978-0674368279[J]. *Business Ethics Quarterly*, 2016, 26(4): 568-571.DOI:<https://philpapers.org/go.pl?id=RUBTBBS-2&proxyId=&u=https%3A%2F%2Fdx.doi.org%2F10.1017%2Fbeq.2016.50>
- [19] Bao T, Liang B, Riyanto Y E. Unpacking the negative welfare effect of social media: Evidence from a large scale nationally representative time-use survey in China[J]. *China Economic Review*, 2021: 101650.D <https://doi.org/10.1016/j.chieco.2021.101650>
- [20] Passe J, Drake C, Mayger L. Homophily, echochambers, & selective exposure in social networks: What should civic educators do? [J]. *The Journal of Social Studies Research*, 2018, 42(3): 261-271.DOI:<https://doi.org/10.1016/j.jssr.2017.08.001>