Public Sentiment Toward Personalized Recommendation System: Evidence in Chinese Social Media

Lewis Bingwen Huang
University of Macau
Joseph Sy-Changco
University of Macau

Cite as:
Huang Lewis Bingwen, Sy-Changco Joseph (2025), Public Sentiment Toward Personalized Recommendation System: Evidence in Chinese Social Media. *Proceedings of the European Marketing Academy*, 54th, (125697)

Paper from the 54th Annual EMAC Conference, Madrid, Spain, May 25-30, 2025

Public Sentiment Toward Personalized Recommendation System: Evidence in Chinese Social Media

Personalized Recommendation Systems (PRS) are algorithmic tools designed to deliver tailored content based on users' digital footprints, gaining attention in recent research. This study examines Chinese public sentiment toward PRS using user-generated content from Weibo, employing sentiment analysis and topic modeling. Findings reveal a neutral to positive sentiment overall but indicate declining satisfaction over time. PRS enhances fanbase growth, social engagement, and user experience but raises concerns about filter bubbles, annoying ads, and privacy. Recommendations include diversifying content and increasing algorithm transparency to mitigate negative sentiment. Retailers should address declining PRS effectiveness by reassessing strategies and resource allocation.

Keywords: *Personalized recommendation systems, social media, sentiment analysis*
Track: *Digital Marketing & Social Media*

1. Introduction

Consider the following scenarios: after liking a video on social media, you are presented with a stream of similar video recommendations; or upon purchasing a chair on an e-commerce platform, you are promptly shown suggestions for complementary furniture. Such occurrences, which many experience regularly, are not mere coincidences but the result of sophisticated recommendation systems at work. Recommendation systems, or recommendation algorithms, have become integral to modern digital activities. By predicting user preferences and delivering tailored content based on digital footprints, they mitigate information overload and enhance online experiences. Research highlights their significant potential in business, improving marketing effectiveness and boosting sales for retailers. Moreover, social platforms benefit from the technology by offering superior marketing outcomes to business clients, increasing their pricing leverage. While recommendation systems appear advantageous for all stakeholders, ongoing research continues to refine and improve these algorithms.

In fact, the algorithm technology is still far from perfection, it may arouse certain drawbacks. From the user perspective, the filtered content may cause filter bubbles and echo chambers, which limit users' perceptions, leading to perception bias. From the business perspective, as recommendation systems tend to recommend popular items to users, products or services with less popularity would not gain significant benefits from algorithms, which in turn may not be helpful in driving business growth (Elahi et al., 2021). Nevertheless, it seems that there is a hidden factor that determines the success of recommendation systems, the user's attitude. If the general public are negative toward such technologies, the aforementioned business potential will then no longer exist. Hence, it is important to explore the public sentiment toward the recommendation system. This research focuses on exploring the sentiment of Chinese online users, as China is known for its immense purchasing power.

According to official data from the National Bureau of Statistics, China's online retail sales in 2023 exceeded 15,426.4 billion yuan, marking an 11.0% growth from the previous year—the highest globally (National Bureau of Statistics, 2024). This underscores the importance for companies entering the Chinese market to conduct in-depth investigations to gain insights and drive business growth. Moreover, China's Cybersecurity Law mandates online users to submit verified identities and disclose their IP territories on social platforms, providing marketers with richer demographic data for analysis.

Despite limited research on public sentiment toward personalized recommendation systems, this study seeks to address this gap. By analyzing user-generated content on Sina Weibo, one of China's largest social networking platforms, the research examines public attitudes toward recommendation systems. The findings aim to offer actionable insights for marketers and companies to develop more effective marketing strategies tailored to the Chinese market.

2. Exploring the Digital Landscape

China's internet development stands out globally, driven by advancements in digital technology and the economy. As of 2023, the Chinese internet penetration rate reached 76.4% of the population, with 1.079 billion users—over 20% of the global internet user base (Statista, 2024). This growth is mirrored by the 1.03 billion social media users reported in January 2023, a testament to the significant role of social media in China's internet ecosystem (Statista, 2024). Nearly all internet users (99.8%) access it through mobile devices, spending an average of 29.1 hours online weekly (CNNIC, 2023). E-commerce also thrives, with 884 million users—82% of the online population—engaging in online shopping (CNNIC, 2023). China's internet advertising market, valued at 223 billion yuan, heavily invests in social media and short video platforms (Statista, 2024). These advancements are facilitated by stringent regulations under the Chinese Cybersecurity Law, requiring users to verify identities and display Internet Protocol (IP) information on platforms. Although foreign sites like Facebook face restrictions through the Great Firewall, this has allowed domestic platforms such as WeChat and Sina Weibo to flourish. For example, Sina Weibo reported 260 million daily active users in September 2023 (Weibo Corporation, 2023). This regulatory and technological ecosystem provides fertile ground for personalized recommendation technologies.

2.1 User-Generated Content (E-WOM)

The rise of social media has transformed users from passive information consumers to active content creators, with user-generated content (UGC) forming a cornerstone of digital interaction. Hennig-Thurau et al. (2004) define electronic Word of Mouth (eWOM) as any online consumer statement, positive or negative, about a product or company. This content is instrumental in fostering community engagement and drives the growth of platforms like Facebook, Instagram, and Weibo (Buhalis & Law, 2008; Filieri, 2015). UGC offers diverse,

authentic perspectives, often perceived as more trustworthy than brand-generated content (Yamagishi et al., 2023). It allows businesses to glean insights into customer preferences, refine products, and conduct cost-effective marketing (Li et al., 2022). Moreover, influencers—a byproduct of UGC—extend its impact by acting as brand advocates (Duffy, 2020). Platforms such as Little Red Book demonstrate the effectiveness of UGC, where positive reviews by foodies boost visibility for restaurants.

2.2 Text mining and natural language processing

Analyzing UGC presents challenges due to its unstructured nature. Text mining and Natural Language Processing (NLP) techniques address this by extracting meaningful patterns from vast textual data (Feldman & Sanger, 2013). Text mining integrates data mining, machine learning, and NLP to process unstructured data efficiently, uncovering trends within UGC (Barnaghi et al., 2016). However, human language's complexity and context sensitivity can complicate machine-based analyses. NLP, utilizing machine learning and large language datasets, bridges this gap by enabling precise comprehension of textual content (Chowdhary, 2020). By processing public opinions on social media, NLP allows businesses to predict consumer needs, enhancing decision-making in product development and service offerings (Afriliana et al., 2022).

2.3 Personalized recommendation system

The digital era has led to information overload, making personalized recommendation systems (PRS) indispensable (Liu, 2017). These systems filter vast amounts of data to deliver tailored content, enhancing user experience and enabling efficient navigation (Sun et al., 2019). Collaborative filtering (CF), a widely used technique, predicts user preferences based on historical data (Zhang et al., 2020). PRS first gained traction in e-commerce, where platforms like Amazon and Taobao use them to suggest products, while Netflix and Google apply them to recommend entertainment and ads (Linden et al., 2003; Das et al., 2007). These systems are now integral to apps, demonstrating significant commercial value. For instance, Koren et al. (2009) found that personalized recommendations increase user engagement, driving higher satisfaction and purchase likelihood.

Despite their advantages, PRS face limitations. Over-personalization can create filter bubbles and echo chambers, limiting exposure to diverse viewpoints and reinforcing biases (Elahi et al., 2021). In addition, privacy concerns arise from the collection and use of personal

data, with public apprehension about potential misuse growing in recent years (Meng et al., 2022).

While PRS improves user experience and marketing efficiency, its flaws necessitate a deeper understanding of public sentiment. Key questions addressed in this study include:

1. Do users predominantly embrace or reject PRS?
2. What are the main reasons for their attitudes?
3. How do demographic factors influence perceptions?
4. Have these perceptions evolved over time?

3. Methodology

This study employs a systematic approach to evaluate public sentiment toward personalized recommendation systems (PRS) and identify the factors influencing attitudes. The methodology involves data collection, pre-processing, topic modeling, and sentiment analysis, leveraging advanced computational tools tailored for Chinese-language text.

Given China's unique cyber environment, where platforms like Facebook and Twitter are restricted, Weibo was selected as the data source due to its extensive and diverse user base. Posts from 2019 to 2023 containing keywords such as “个性化推荐” (personalized recommendation), “算法推荐” (algorithm recommendation), and “大数据推荐” (big data recommendation) were collected using the Selenium Python library. Demographic information, such as users’ IP territory and gender, was included to enrich the dataset, which was archived in CSV format for further analysis.

To uncover key themes in the dataset, the study applied Latent Dirichlet Allocation (LDA), a probabilistic model widely used in topic modeling. LDA assumes that each document contains a mix of topics, characterized by distributions of words (Blei & Jordan, 2003). Unlike simple frequency-based methods, LDA analyzes relationships between words to reveal hidden patterns, providing deeper insights into user concerns about PRS. This enabled the identification of major discussion points, such as valued features and recurring issues.

Sentiment analysis was conducted using SnowNLP, a Python library specifically designed for Chinese text, to determine the emotional polarity of Weibo posts. Each post was

assigned a polarity score ranging from 0 (negative) to 1 (positive). SnowNLP was chosen for its superior performance in processing Chinese text, including part-of-speech tagging and word segmentation. The analysis also explored sentiment variations across demographic groups to provide a comprehensive understanding of public attitudes toward PRS.

By integrating advanced tools such as LDA and SnowNLP, this methodology effectively processed large volumes of unstructured text, uncovering meaningful insights into user sentiment. The study's approach highlights the complexity of analyzing public perspectives and provides actionable findings for enhancing algorithmic transparency and user experience.

4. Public Sentiment and Insights on Personalized Recommendation Systems

After analyzing Weibo posts from January 2019 to December 2023, 11,334 valid records were retained from the initial 12,501 collected. Gender distribution showed 47% of posts from female users and 53% from male users, indicating no significant gender bias in concerns about personalized recommendation systems (PRS). The volume of posts grew steadily over time, increasing from 1,232 in 2019 to 3,538 in 2023, reflecting heightened public interest in PRS.

The average sentiment score was 0.625, with 50% of posts scoring above 0.65, signifying a generally neutral to positive sentiment. However, 25% of posts scored below 0.5, highlighting some concerns. Male users exhibited slightly higher sentiment scores than females, suggesting greater acceptance of PRS, whereas women appeared more cautious. Regionally, users from economically developed cities such as Beijing, Shanghai, and Shenzhen demonstrated marginally higher sentiment scores than those from less developed areas, indicating greater openness in urban regions. Over time, sentiment scores remained largely within the range of 0.5 to 0.75, reflecting consistent neutral to positive attitudes. A slight downward trend emerged, potentially linked to increasing public awareness of data privacy concerns and skepticism about PRS. Notable fluctuations in sentiment, such as a drop in late 2020, may be attributed to data availability during specific periods.

Keyword analysis revealed distinct themes in posts with positive and negative sentiments. Positive posts highlighted PRS's role in enhancing engagement and expanding fanbases, with terms like "fans," "interaction," and "strengthen" frequently appearing. In contrast, negative posts emphasized privacy risks, intrusive ads, and algorithm overuse, with

keywords such as "privacy," "ads," and "turn off" being prominent. Latent Dirichlet Allocation (LDA) identified seven topics related to PRS, spanning its applications in social networking, e-commerce, and technical domains. Posts with positive sentiment focused on benefits like media engagement and leisure activities, while negative posts centered on concerns over data misuse, privacy violations, and excessive advertising.

These findings demonstrate the dual perception of PRS. While the technology offers clear benefits, including enhanced user experience and engagement, concerns about privacy and over-personalization persist. This underscores the need for platforms to prioritize algorithmic transparency and balance to address public apprehensions while leveraging PRS's potential.

5. Insights, Implications, and Limitations of the Study

Recommendation systems, from e-commerce product suggestions to streaming media recommendations, have become integral to digital activities. By predicting preferences through users' digital footprints, they enhance online experiences while offering significant commercial potential for businesses and marketers. However, concerns such as echo chambers, perception bias, and privacy risks persist, drawing attention to public attitudes toward these technologies. This study addresses this gap by analyzing Chinese public sentiment toward personalized recommendation systems (PRS) using user-generated content (UGC) on Weibo. Key questions explored include overall attitudes, demographic variations, primary concerns, and temporal changes in sentiment.

Analysis revealed a neutral to positive overall sentiment, with male users and those from economically developed cities exhibiting higher acceptance of PRS. Sentiment trends showed a slight decline over time, suggesting increasing public apprehension about data privacy and algorithmic transparency. Keyword and topic analyses highlighted both benefits and concerns, with users praising PRS for enhancing engagement and expanding fanbases while criticizing intrusive ads and potential privacy violations.

These findings offer valuable managerial insights. For online retailers, PRS remains an effective marketing tool, particularly for male users and those in urban areas. However, the downward trend in sentiment indicates that businesses should monitor public attitudes and reassess their resource allocation to PRS. For platform developers, enhancing transparency about how recommendations are generated and data is used could alleviate privacy concerns

and correct misconceptions. Additionally, diversifying content recommendations can reduce the risks of echo chambers and filter bubbles, fostering a more positive user experience. These measures could sustain PRS's business value over time despite growing user skepticism.

This study provides insights into public attitudes toward PRS, including key factors influencing perceptions and demographic differences. However, limitations exist. The narrow range of keywords used for data collection may have excluded relevant posts, leading to an incomplete representation of public sentiment. Expanding the keyword list in future research could provide a more comprehensive understanding. Weibo's daily-life focus may limit the diversity of opinions compared to more academically oriented platforms like Baidu Tieba or Zhihu. Including multiple platforms in future studies would enhance data representativeness.

Moreover, this research did not analyze comments, likes, or reposts, which could provide valuable indicators of sentiment agreement. Cultural nuances, such as encoded slang or filtered words on Chinese social media, may also affect sentiment accuracy. Advanced methodologies, including large language models like ChatGPT or Llama, could improve analysis precision. Finally, future research could combine data mining with surveys to apply theoretical frameworks such as the Technology Acceptance Model (TAM) or the Theory of Planned Behavior (TPB), offering deeper insights into user acceptance and behavior toward PRS.

References

Afriliana, N., Made, N., & Iswari, S. (2022). Sentiment Analysis of User-Generated Content: A Bibliometric analysis. *Journal of System and Management Sciences*. https://doi.org/10.33168/jsms.2022.0634

Barnaghi, P., Ghaffari, P., & Breslin, J. G. (n.d.). *Opinion mining and sentiment polarity on Twitter and correlation between events and sentiment*. 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, United Kingdom of Great Britain and Northern Ireland. https://doi.org/10.1109/bigdataservice.2016.36

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.5555/944919.944937

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management*, *29*(4), 609–623. https://doi.org/10.1016/j.tourman.2008.01.005

Chowdhary, K. R. (2020). Natural language processing. In *Springer eBooks* (pp. 603–649). https://doi.org/10.1007/978-81-322-3972-7_19

Das, A. S., Datar, M., Garg, A., & Rajaram, S. (2007). Google News Personalization: Scalable online Collaborative filtering. https://doi.org/10.1145/1242572.1242610

Duffy, B. E. (2020). Social media influencers. *The International Encyclopedia of Gender, Media, and Communication*, 1–4. https://doi.org/10.1002/9781119429128.iegmc219

Elahi, M., Jannach, D., Skjærven, L., Knudsen, E., Sjøvaag, H., Tolonen, K., Holmstad, Ø., Pipkin, I., Throndsen, E., Stenbom, A., Fiskerud, E., Oesch, A., Vredenberg, L., & Trattner, C. (2021). Towards responsible media recommendation. *AI And Ethics*, *2*(1), 103–114. https://doi.org/10.1007/s43681-021-00107-7

Feldman, R., & Sanger, J. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. *Choice Reviews Online*, *44*(10), 44–5684. https://doi.org/10.5860/choice.44-5684

Filieri, R. (2014). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. Journal of Business Research, 68(6), 1261–1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52. https://doi.org/10.1002/dir.10073

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender Systems. *Computer*, *42*(8), 30–37. https://doi.org/10.1109/mc.2009.263

Li, S., Liu, F., Zhang, Y., Zhu, B., Zhu, H., & Yu, Z. (2022). Text Mining of User-Generated Content (UGC) for Business Applications in E-Commerce: A Systematic Review. *Mathematics*, 10(19), 3554. https://doi.org/10.3390/math10193554

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, *7*(1), 76–80. https://doi.org/10.1109/mic.2003.1167344

Liu, H., Xia, F., Chen, Z., Asabere, N. Y., Ma, J., & Huang, R. (2015). TrUCom: Exploiting Domain-Specific Trust Networks for multicategory item recommendation. IEEE Systems Journal, 11(1), 295–304. https://doi.org/10.1109/jsyst.2015.2427193

Meng, S., Fan, S., Li, Q., Wang, X., Zhang, J., Xu, X., Qi, L., & Bhuiyan, M. Z. A. (2022). Privacy-Aware Factorization-Based hybrid recommendation method for healthcare services. *IEEE Transactions on Industrial Informatics*, *18*(8), 5637–5647. https://doi.org/10.1109/tii.2022.3143103

Statista. (2024, February 18). *Number of social media user accounts APAC 2024, by country*. https://www.statista.com/statistics/295606/social-media-mau-%20asia-pacific-countries/

Sun, X., Xu, X., & Xia, F. (2019). CROA: A Content-Based Recommendation Optimization Algorithm for Personalized Knowledge Services. IEEE Systems Journal, 29, 804–810. https://doi.org/10.1109/hpcc/smartcity/dss.2019.00118

Yamagishi, K., Canayong, D., Domingo, M., Maneja, K. N., Montolo, A., & Siton, A. (2023). User-generated content on Gen Z tourist visit intention: a stimulus-organism-response approach. *Journal of Hospitality and Tourism Insights*, *7*(4), 1949–1973. https://doi.org/10.1108/jhti-02-2023-0091

Zhang, J., Adomavicius, G., Gupta, A., & Ketter, W. (2020). Consumption and Performance: Understanding longitudinal dynamics of recommender systems via an Agent-Based simulation framework. *Information Systems Research*, *31*(1), 76–101. https://doi.org/10.1287/isre.2019.0876