

Retraction

Retracted: Research on Book Recommendation Algorithm Based on Collaborative Filtering and Interest Degree

Wireless Communications and Mobile Computing

Received 28 November 2023; Accepted 28 November 2023; Published 29 November 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Z. H. Wang and D. Z. Hou, "Research on Book Recommendation Algorithm Based on Collaborative Filtering and Interest Degree," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 7036357, 7 pages, 2021.

Research Article

Research on Book Recommendation Algorithm Based on Collaborative Filtering and Interest Degree

Zhi Hui Wang¹ and De Zhi Hou²

¹Wuxi Vocational College of Science and Technology, Wuxi 214101, China

²R&D Department, Beijing Huapin Borui Network Technology Co., Ltd., Beijing 100020, China

Correspondence should be addressed to Zhi Hui Wang; wzh@wxsc.edu.cn

Received 26 May 2021; Revised 17 June 2021; Accepted 29 June 2021; Published 12 July 2021

Academic Editor: Balakrishnan Nagaraj

Copyright © 2021 Zhi Hui Wang and De Zhi Hou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the increase of library collections, it is difficult for readers to quickly find the books they want when choosing books. Book recommendation system is becoming more and more important. Based on the previous research, this paper proposes a book recommendation algorithm based on collaborative filtering and interest. Take the interest of the book itself as an important measurement index, including the number of searches, borrowing time, borrowing times, borrowing interval, and renewal times. Through the analysis of MAE and RMSE experiments, the results show that the method proposed in this paper converges faster than the traditional method.

1. Introduction

With the continuous deepening of the informatization process, all walks of life are carrying out informatization reforms. In this context, the process of digitalization of library management information is gradually improving [1]. Digital libraries are deeply loved by readers for their convenient and quick document retrieval methods, personalized recommendations, and other characteristic services [2]. It is difficult for readers to find books of interest in a short period of time in the face of various bibliographies. Therefore, the user experience of the traditional library borrowing method is poor. At present, many scholars have proposed book recommendation methods, including content-based recommendation algorithms, which recommend products similar to the products he liked in the past according to the products that the user liked in the past [3]. A recommendation algorithm based on association rules [4], the same user borrows different books can be considered as having an association relationship [5]. Searching for the collection of books with the highest degree of association from the borrowing information

can be used as an important reference for book recommendation [6]. A combination recommendation model based on tagging and association rule mining and user-based collaborative filtering has nothing to do with items [7]. Find the most similar neighboring users of the current user, and recommend the books borrowed by neighboring users to the current user [8]. A collaborative filtering recommendation algorithm is based on the popularity of the item scoring system combined with average preference weights [9]. A book recommendation algorithm is based on social networks and so on [10].

This paper proposes a book recommendation algorithm based on collaborative filtering and interest degree. Collaborative filtering uses cosine similarity to calculate; interest degree includes search times, borrowing time, borrowing times, borrowing interval, and renewing times according to book attributes. Finally, the average deviation and the root mean square error are used to measure.

The rest of this paper is organized as follows: Section 2 analyzes the factors that affect book recommendation from the two aspects of interest and similarity. The book recommendation algorithm based on collaborative filtering and

interest degree is discussed in Section 3. Section 4 shows the experimental analysis, and Section 5 concludes the paper with summary and future research directions.

2. Analysis of the Factors Affecting Book Recommendation

Recommending books to users can be analyzed in terms of interest and similarity, as shown in Figure 1.

In interest degree, interest is the characteristic of the book itself, and it is recommended from the perspective of attractiveness to the target user [11]. It includes the interest of the book itself and the interest of the user. Book interest refers to the attributes of the book itself, including search times, borrowing time, borrowing interval, borrowing times, and renewing times [12]. User interest refers to books that users like, books that have already been borrowed, and so on.

In similarity, the similarity is to recommend the target user from the perspective of relevance to the user. There are two situations for being associated with users [13]. Either they have common attributes, including major, grade, and gender, or they have a common borrowing group with the books borrowed by the user, and the books borrowed by these people can also be recommended to the user [14].

After comprehensively analyzing the interest and similarity, the relevant model is used to make predictions, and a recommendation list is generated [15]. Finally, comprehensive evaluation is carried out through evaluation indicators, and the prediction results are analyzed.

3. Book Recommendation Algorithm Based on Collaborative Filtering and Interest

3.1. Book Interest Model. Attributes related to books include search times, borrowing time, borrowing times, borrowing interval, and renewing times.

The proportion of search times of a certain book is the proportion of search times to the search times of all books. For normalization, the proportion is divided into five levels; the formula is as follows:

$$A = \frac{M}{N} = \left\{ \begin{array}{l} 5 \quad 0 - 20\% \\ 4 \quad 21 - 40\% \\ 3 \quad 41 - 60\% \\ 2 \quad 61 - 80\% \\ 1 \quad 81 - 100\% \end{array} \right\}. \quad (1)$$

In which, A is the grade score of the search proportion, M is the number of searches, and N is the total number of searches for all books. The higher the ranking means the higher the score.

The length of the borrowing time can basically reflect the popularity of a book [16]. Of course, the borrowing time may be too long because you forget to return the book, or you cannot return the book because of the holiday. This special case

is not considered here for the time being, and the formula is as follows:

$$B = \frac{\sum_1^K \left(\frac{t'(i) - t(i)}{t} \right)}{K} = \left\{ \begin{array}{l} 5 \quad 0 - 20\% \\ 4 \quad 21 - 40\% \\ 3 \quad 41 - 60\% \\ 2 \quad 61 - 80\% \\ 1 \quad 81 - 100\% \end{array} \right\}. \quad (2)$$

In which, B is the grade score of the borrowing time, K is the number of borrowers of the book, t is the longest borrowing time of the book, $t'(i)$ is the book return time, and $t(i)$ is the book borrowing time. The higher the ranking means the higher the score.

The number of borrowings can accurately reflect the popularity of the book [17]. The more borrowing times, the higher the popularity. The formula is as follows:

$$C = \frac{Q}{Q'} = \left\{ \begin{array}{l} 5 \quad 0 - 20\% \\ 4 \quad 21 - 40\% \\ 3 \quad 41 - 60\% \\ 2 \quad 61 - 80\% \\ 1 \quad 81 - 100\% \end{array} \right\}. \quad (3)$$

In which, C is the rating value of the number of borrowings, Q is the number of borrowings of the book, and Q' is the total number of borrowings of all books. Divide the number of borrowings into five levels, and rank according to the number of borrowings. The ranking is within 20%, the highest level. The ranking is 81%-100%, the level is the lowest, and the score is the lowest.

The borrowing interval refers to the time interval for a book to be borrowed after being returned [18]. If the time is shorter, the demand for the book is greater, or the popularity of the book is greater. On the contrary, if it is returned and is no longer borrowed, it means that the popularity of the book is very low. The formula is as follows:

$$D = \frac{\sum_1^K \left(\frac{t(i+1) - t'(i)}{t} \right)}{K} = \left\{ \begin{array}{l} 5 \quad 0 - 20\% \\ 4 \quad 21 - 40\% \\ 3 \quad 41 - 60\% \\ 2 \quad 61 - 80\% \\ 1 \quad 81 - 100\% \end{array} \right\}. \quad (4)$$

In which, D is the grade score of the borrowing interval, K is the number of borrowers of the book, t is the longest borrowing time of the book, $t'(i)$ is the return time, and $t(i+1)$ is the next person's borrowing time. \sum is the sum. The higher the ranking means the higher the score.

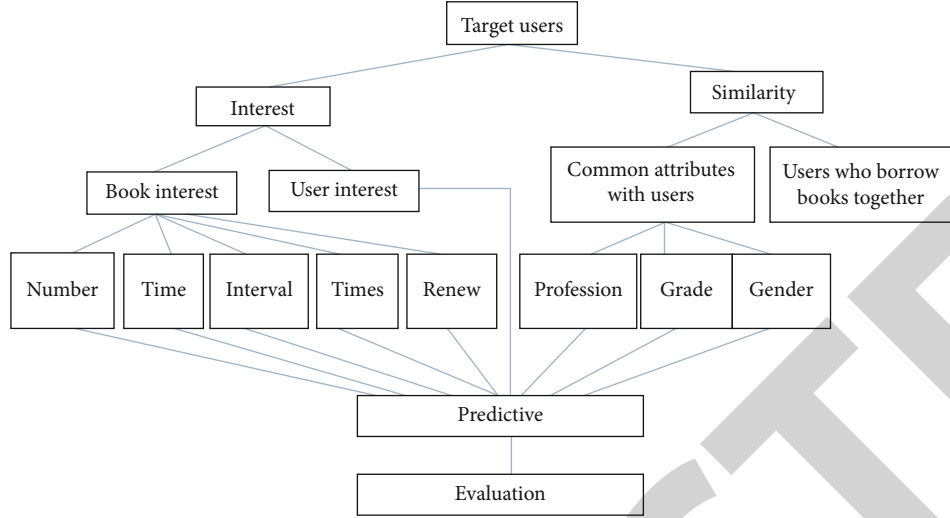


FIGURE 1: Recommended elements for target users.

The number of renewals of a book can also reflect the popularity of the book to a certain extent. The formula is as follows:

$$E = \frac{P}{P'} = \left\{ \begin{array}{l} 5 \quad 0 - 20\% \\ 4 \quad 21 - 40\% \\ 3 \quad 41 - 60\% \\ 2 \quad 61 - 80\% \\ 1 \quad 81 - 100\% \end{array} \right\}. \quad (5)$$

In which, E is the grade score of the proportion of the number of renewals, P is the number of renewals of the book, and P' is the total number of renewals of all books. The higher the ranking means the higher the score.

Finally, the average of the five indicators is used as a comprehensive indicator of book interest; the formula is as follows:

$$X = \frac{A + B + C + D + E}{5}. \quad (6)$$

3.2. Collaborative Filtering Recommendation Model. The basic idea of the collaborative filtering algorithm is to find similar users of the current user and predict the current user's score based on the similar user's score information to make recommendations. A recommendation system based on collaborative filtering does not analyze information from data but establishes an effective evaluation feedback mechanism to allow users to form a good feedback [19]. In other words, the recommendation users get may not be mined from the data at all but contributed by other users. There are three main steps: collecting scoring data, finding neighbors, and generating a recommendation list.

The scoring data is shown in Table 1.

The user-based collaborative filtering algorithm is to find neighbor users with high similarity for the current user and

TABLE 1: The scoring data.

	Profession	Grade	Gender	Same borrower
User 1	D_{11}	D_{12}	D_{13}	D_{14}
User 2	D_{21}	D_{22}	D_{23}	D_{24}
\vdots	\vdots	\vdots	\vdots	\vdots
User m	D_{m1}	D_{m2}	D_{m3}	D_{m4}

then recommend items that the neighbor users have rated, and the current user has not rated to the current user. The steps include calculating the similarity between the current user and other users, sorting according to the similarity from the highest to the bottom and the user with the highest ranking as the current user's neighbor, filtering the current user's rating items from the neighboring user's rating list, predicting the current user's rating of unrated item scoring, selecting the one with the highest score, and recommending it to the current user.

Cosine similarity can describe the linear correlation between two sets of data, and its value range is between -1 and 1. The cosine similarity is calculated based on the set of items jointly evaluated by two users [20]. When using this method to calculate, it is necessary to remove the average value of all commodities evaluated by the user. Generally, the following calculation formula is used to calculate the similarity.

$$\text{Sim}(u, v) = \frac{\sum_{i \in I_{uv}} (R_{ui} \times R_{vi})}{\sqrt{\sum_{i \in I_u} R_{ui}^2} \times \sqrt{\sum_{i \in I_v} R_{vi}^2}}. \quad (7)$$

In which, I_u is the set evaluated by user u , I_v is the set evaluated by user v , R_{ui} is the rating of user u on item i , and R_{vi} is the rating of user v on item i .

In this paper, the interest degree and collaborative filtering are averaged for a comprehensive analysis.

TABLE 2: 2020 classified borrowing statistics.

Class name	Class number	Jan.	Feb.	...	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Marxist	A	2	0	...	0	4	0	3	14	5	5
Philosophy	B	30	0	...	9	7	0	18	92	126	55
Social science	C	26	0	...	6	3	0	6	24	49	44
Politics	D	0	0	...	1	13	0	5	9	32	34
Military	E	0	0	...	0	0	0	0	4	5	4
Economic	F	18	0	...	17	10	0	34	51	145	102
Culture	G	17	0	...	13	0	0	31	42	27	43
Language	H	65	0	...	35	9	0	88	370	480	330
Literature	I	127	0	...	131	43	0	115	788	1528	765
Art	J	24	0	...	22	0	0	23	98	102	59
History	K	23	0	...	11	14	0	23	78	110	61
Natural	N	0	0	...	12	0	0	2	1	8	1
Mathematical	O	9	0	...	4	0	0	7	41	47	73
Astronomical	P	0	0	...	1	0	0	0	2	6	1
Biological	Q	1	0	...	2	0	0	0	2	8	1
Medicine	R	2	0	...	10	0	0	13	8	18	15
Agriculture	S	1	0	...	0	0	0	1	0	2	1
Industry	T	98	0	...	36	27	3	92	147	296	233
Traffic	U	0	0	...	1	0	0	0	3	1	3
Aviation	V	0	0	...	0	0	0	0	0	0	0
Surroundings	X	1	0	...	0	0	0	0	0	0	2
Comprehensive	Z	9	0	...	18	1	0	16	20	21	17

TABLE 3: Part of MAE data.

Number	MAE	Number	MAE	Number	MAE
1	0.89	160	0.6	320	0.58
10	0.88	170	0.6	330	0.57
20	0.83	180	0.61	340	0.57
30	0.74	190	0.6	350	0.56
40	0.64	200	0.59	360	0.56
50	0.64	210	0.59	370	0.56
60	0.64	220	0.59	380	0.56
70	0.64	230	0.59	390	0.55
80	0.62	240	0.59	400	0.55
90	0.62	250	0.57	410	0.56
100	0.62	260	0.57	420	0.54
110	0.62	270	0.58	430	0.54
120	0.62	280	0.57	440	0.54
130	0.61	290	0.57	450	0.54
140	0.61	300	0.57	460	0.54
150	0.61	310	0.57	470	0.54

3.3. *Evaluation Model.* Mean deviation and root mean square error are usually two standards to measure the accuracy of the recommended system.

Divide the attributes into five levels and rank them according to the data. The ranking is within 20%, the highest level; the ranking is 81%-100%, the lowest level, and the score is the lowest.

The formula for mean deviation is as follows:

$$MAE = \frac{\sum_{i=1}^n |p_i - q_i|}{n}. \quad (8)$$

In which, p_i is the predicted user rating, and q_i is the user's actual rating. The smaller the deviation of the average value, the closer the predicted score of the recommendation algorithm is to the actual score.

The formula for the root mean square error is as follows:

$$RMSE = \sqrt{\frac{\sum_{(u,i) \in T} (r_{ui} - r'_{ui})^2}{|T|}}. \quad (9)$$

In which, T represents the test data, $|T|$ represents the size of the test data set, u represents the user, i represents the book, r_{ui} represents the user's actual score for the book, and r'_{ui} represents the user's predicted score for the book.

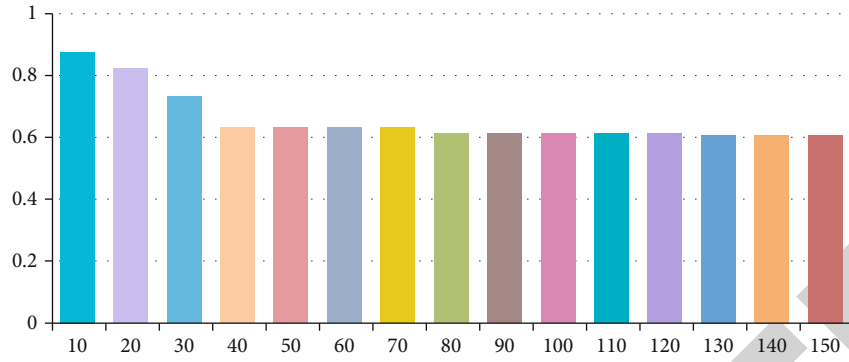


FIGURE 2: MAE comparison chart.

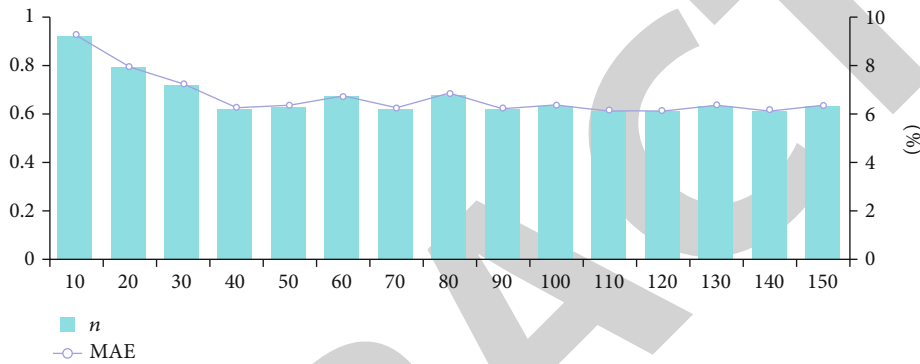


FIGURE 3: MAE comparison chart of science and engineering.

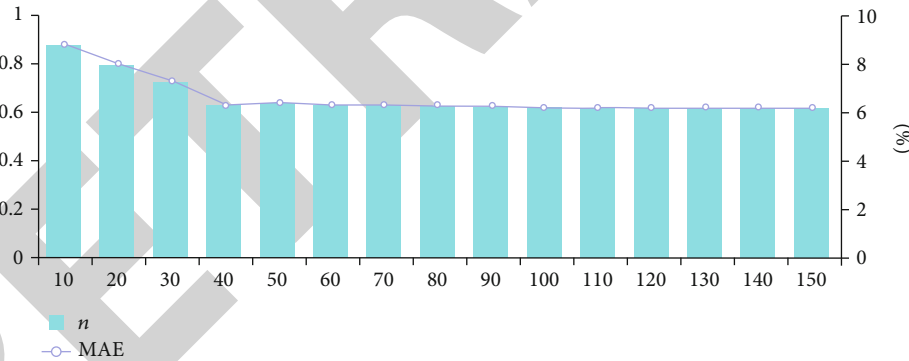


FIGURE 4: MAE comparison chart of literature and history.

4. Experimental Results and Analysis

The experimental data selected the borrowing data and user data of Wuxi Vocational College of Science and Technology from 2014 to 2020, and the data was copied and expanded to three times the original data volume. The borrowing data for 2020 is shown in Table 2. In the experiment, the data set is divided into training set and test set in an 8:2 manner.

The MAE is calculated according to formula (8), and the partial comprehensive data obtained is shown in Table 3, and the MAE of the first 150 neighbors is shown in Figure 2.

It can be seen from Figure 2 that in the comparison of the first 150 neighbor users, it gradually decreased at the begin-

ning, and the maximum value was 0.89. When it reaches about 40 neighbor users, it basically tends to be stable, and the stable value is around 0.6. The data is divided into literature and history and science and engineering for analysis. Figure 3 is the MAE value of science and engineering, and Figure 4 is the MAE value of literature and history. It can be seen that the literature and history category basically tends to be stable when it reaches about 40 neighboring users, similar to Figure 2. The science and engineering category also tends to be stable among 40 neighbor users, but there are some fluctuations. This shows that the data of literature and history accounts for a higher proportion of the whole data, and the frequency of borrowing is also higher.

TABLE 4: Part of RMSE data.

Number	Literature and history	Science and engineering	Number	Literature and history	Science and engineering
10	0.97	1.02	160	0.61	0.64
20	0.96	1	170	0.61	0.63
30	0.84	0.9	180	0.62	0.63
40	0.67	0.83	190	0.6	0.63
50	0.64	0.77	200	0.6	0.61
60	0.63	0.65	210	0.59	0.61
70	0.63	0.65	220	0.59	0.62
80	0.64	0.69	230	0.6	0.6
90	0.62	0.66	240	0.58	0.6
100	0.63	0.65	250	0.58	0.59
110	0.62	0.64	260	0.58	0.59
120	0.62	0.67	270	0.58	0.6
130	0.63	0.63	280	0.57	0.59
140	0.62	0.65	290	0.57	0.59
150	0.62	0.64	300	0.57	0.59

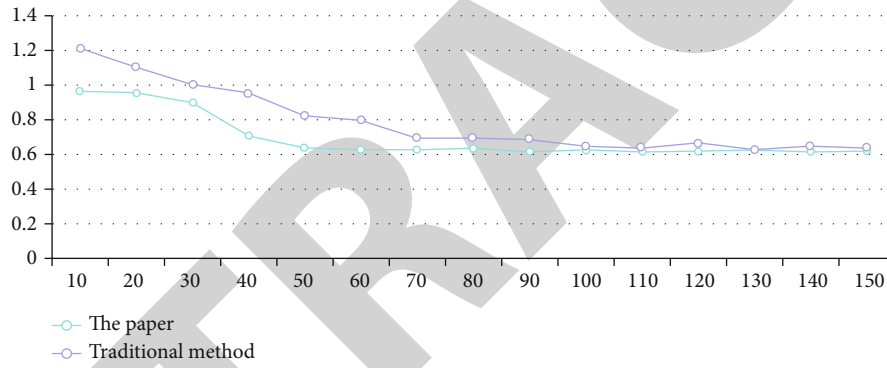


FIGURE 5: RMSE comparison chart of literature and science.

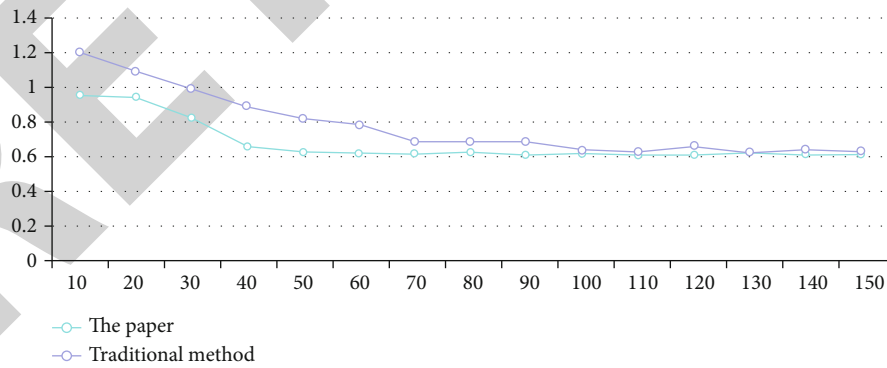


FIGURE 6: Comparison chart with traditional method RMSE.

The RMSE is calculated according to formula (9), and the partial comprehensive data obtained is shown in Table 4, and the RMSE of the first 150 neighbors is shown in Figure 5.

It can be seen from Figure 5 that the literature and history category also tends to be stable with 40 neighbor users. The science and engineering category tends to be stable with 60 users. This is the same as the previous article, indicating that

the data of literature and history account for a higher proportion of the entire data, and the frequency of borrowing is also higher.

This article analyzes from two perspectives of collaborative filtering and interest and compares it with the traditional single cosine similarity collaborative filtering, as shown in Figure 6. It can be seen that the method proposed in this

paper tends to be stable at about 50 times, and the fluctuation is small. The traditional method only stabilizes after about 70 times, and the fluctuations are still greater after stabilization. It shows that the method in this paper has better convergence and stability.

5. Conclusion

This paper proposes a book recommendation algorithm based on collaborative filtering and interest. Collaborative filtering uses cosine similarity for analysis, and the interest degree uses the basic attributes of the book as a measurement index. Through analyzing the statistical data of the library of Wuxi Vocational College of Science and Technology for many years, using MAE and RMSE two measurement indicators for experimental analysis, it is concluded that the method proposed in this paper has a good convergence result. The goal of the next step is to optimize the collaborative filtering algorithm and at the same time to optimize the measurement indicators, so as to have better convergence results.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] P. Li and Z. Hong, "Book recommendation algorithm based on the interest and type factor for university," *Journal of Zhejiang university of technology*, vol. 47, no. 4, pp. 425–429, 2019.
- [2] L. Sun, "Review and prospects of research on online personalized recommendation," *Foreign Economics and Management*, vol. 38, no. 6, pp. 82–99, 2016.
- [3] C. Zuo, Q. Chen, L. Tian, L. Waller, and A. Asundi, "Transport of intensity phase retrieval and computational imaging for partially coherent fields: the phase space perspective," *Optics and Lasers in Engineering*, vol. 71, no. 8, pp. 20–32, 2015.
- [4] F. Yin, X. Xue, C. Zhang et al., "Multifidelity genetic transfer: an efficient framework for production optimization," *SPE Journal*, vol. 23, no. 2, pp. 1–22, 2021.
- [5] X. Xue, K. Zhang, K. C. Tan et al., "A semantic-aware real-time correlation tracking framework for UAV videos," *IEEE Transactions on Cybernetics*, vol. 8, no. 2, pp. 1–12, 2020.
- [6] Q. Jiang, F. Shao, W. Lin, K. Gu, G. Jiang, and H. Sun, "Optimizing multistage discriminative dictionaries for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2035–2048, 2018.
- [7] K. Zhang, J. Zhang, X. Ma et al., "History matching of naturally fractured reservoirs using a deep sparse autoencoder," *SPE Journal*, vol. 21, no. 4, pp. 1022–1030, 2021.
- [8] S. Qu, W. Xu, J. Zhao, and H. Zhang, "Design and implementation of a fast sliding-mode speed controller with disturbance compensation for SPMSM system," *IEEE Transactions on Transportation Electrification*, vol. 12, no. 8, pp. 445–455, 2021.
- [9] Z. Wu, A. Song, J. Cao, J. Luo, and L. Zhang, "Efficiently translating complex SQL query to mapreduce jobflow on cloud," *IEEE transactions on cloud computing*, vol. 8, no. 2, pp. 508–517, 2020.
- [10] N. Xiao, R. Xinyi, Z. Xiong et al., "A diversity-based selfish node detection algorithm for socially aware networking," *Journal of Signal Processing Systems*, vol. 93, no. 7, pp. 811–825, 2021.
- [11] P. Wang, "Research and practice of book recommendation system based on SNS," *Computer Applications and Software*, vol. 29, no. 12, pp. 21–23, 2012.
- [12] G. Wang, "Recommendation algorithm for E-learning resources based on improved collaborative filtering," *Journal of Chinese Computer Systems*, vol. 42, no. 5, pp. 940–945, 2021.
- [13] C. Liu, "Collaborative filtering hybrid recommendation algorithm based on improved biased and cluster user nearest neighbor," *Computer Applications and Software*, vol. 38, no. 5, pp. 288–293, 2021.
- [14] Y. Leng and Q. Lu, "Overview of collaborative filtering recommendation technology," *Pattern Recognition and Artificial Intelligence*, vol. 27, no. 8, pp. 720–734, 2014.
- [15] M. Li, "Research on book combination recommendation system model based on tag and association rule mining," *Computer Application Research*, vol. 31, no. 8, pp. 2390–2393, 2014.
- [16] T. Huang, "Application research of collaborative filtering algorithm based on neighbor model," *Journal of Beijing Information Science and Technology University (Natural Science Edition)*, vol. 32, no. 4, pp. 90–94, 2017.
- [17] J. He, "Collaborative filtering personalized recommendation algorithm based on popularity division combined with average preference weight," *Computer Science*, vol. 45, no. 6, pp. 25–33, 2018.
- [18] X. Weng and Z. Wang, "Research progress of collaborative filtering recommendation algorithm," *Computer Engineering and Applications*, vol. 54, no. 1, pp. 25–31, 2018.
- [19] J. Zhang, "A fusion collaborative filtering algorithm based on rating information entropy," *Nanjing Universal of Posts and Telecommunications (Natural Science Edition)*, vol. 41, no. 2, pp. 76–81, 2021.
- [20] R. Niu, "The inverse methods based on SFEMs with an adaptive SVD regularization technique for solving Cauchy inverse heat transfer problems," *Engineering Analysis with Boundary Elements*, vol. 10, no. 7, pp. 79–95, 2019.