IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified
Vol. 5, Issue 12, December 2017

Recommendation Model of Mathematical Expressions Based on Collaborative Filtering

Zhaoyu Wang[1], Xinfu Li[2], Fang Yang[3], Xiao Xu[1], Xuedong Tian[2*]

Master's Student, School of Cyber Security and Computer, Hebei University, Baoding, China [1]

Professor, School of Cyber Security and Computer, Hebei University, Baoding, China [2]

Associate Professor, School of Cyber Security and Computer, Hebei University, Baoding, China [3]

Abstract: The rapid development of the Internet has caused the problem of "information overload", especially in the field of information retrieval. Since mathematical retrieval as part of information retrieval, this paper proposed a model of interest recommendation for it. First, the public data sets of mathematical expressions were normalized for the mathematical formulas which have the same computing meaning but different parameters. Secondly, the concept of fuzzy set was employed for making the system to know the similarity of users according to the different habits of each user in the multi features fuzzy pattern through integrating these attributes to form a "user-formula" scoring matrix for calculating the closeness degree between users. And the nearest neighbor set of the target user is obtained. Finally, using the fuzzy elements of score and the corresponding neighbors to predict the score of the target user with non-scoring formulas to form the recommendation list of interested formulas through the further filtering of the threshold factor, and realize the personalized recommendation for the mathematical expressions at last. The feasibility and effectiveness of the proposed method are verified by the experiment and data.

Keywords: Mathematical expressions, collaborative filtering, fuzzy pattern with multi features, normalization, personalized recommendation

I. Introduction

With the popularity of the Internet and the development of technology, it becomes more and more difficult for people to find the message what they need in the vast ocean of information. Although the traditional search engines have brought great convenience for people to retrieve information and become the best way to find target messages quickly, they have many shortages as the pure search engines cannot meet the needs of users for discovering information. In many cases, users are not clear their own needs, or their needs are not easily described by simple content and formula symbols. With the emergence of the recommendation system, people's search habits have gradually developed from simple explicit methods to more advanced information discovery according to their habits. Today, with the development of recommendation technology, recommendation systems have been successfully applied on the field of e-commerce and some social networking sites. This also further explains that in the era of big data information, users are more interested in the information recommending mechanism which is more intelligent and can understand their needs.

Different from text, mathematical expressions have complex two-dimensional structure and nested format. So the process of editing, inputting, displaying and retrieving mathematical expressions in the front end or matching mechanism of formulas in the back end are restricted greatly. In additional to the various expressing modes of mathematical formulas, such as picture, Word, PDF, LaTeX, MathML and other formats [1], the retrieval technology of mathematical formula lags far behind the text retrieval.

The technology of mathematical expression retrieval has been developed in different degrees. Some prototypical systems of mathematical expression retrieval such as MathWebSearch [2], WikiMirs [3], EgoMath [4], MathDex [5,6], DLMFSearch [7], LeActiveMath [8,9], MIaS [10], MCAT [11] and MathSearch [12] were proposed. These model make a good foundation for realizing math retrieval. However, the interest model of mathematical expressions are rarely involved in the research of math retrieval. This paper introduced a recommendation algorithm of collaborative filtering and designed a recommendation model of user interest in the mathematical expression retrieval, and made a relevant evaluation to the recommendation results.

II. Overview of Collaborative Filtering

The recommender systems can be classified into four types called Association Rule-based Recommendation(called ARBR), Content-based Recommendation(called CBR), Collaborative Filtering Recommendation(called CFR) and Hybrid Recommendation(called HR) [13]. Among these methods, collaborative filtering algorithm is the most widely used recommendation algorithm.

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 12, December 2017

Recommendation engines could be divided into several categories according to the contentful correlation of different data sources called Demographic-based Recommendation [14] and Content-based Recommendation. As well as the Collaborative Filtering-based Recommendation, Association Rules-based Recommendation and Slope one recommendation algorithm [15].

Collaborative Filtering-based Recommendation is based on the preference to the information of all users to find the neighbours' similar to the current user on interest, and recommends the content which is interested but not touched before to the target user. The principle of this method is relatively simple and has higher recommendation accuracy. It has been widely used because it doesn't require strict modelling for users and projects, and the description of content is not limited. What's more, the recommendation calculated by this method is open, and can draw lessons from other people [16]. It can be divided into three sub categories: User-based Recommendation, Item-based Recommendation and Model-based Recommendation.

User-based Recommendation algorithm finds the collection of neighbours for the target user with the user's preferences and recommends the content to the target user according to the similarity degree between neighbours. Item-based Recommendation provides the recommendations according to the user's historical preferences through analysing the similarity between content. The former is based on the users' historical preferences, and the latter is according to the attribute characteristics of the content. User-based Recommendation is applicable for the condition of user's number is much less than the amount of data as the user data is relatively stable and the possibility of change is small. So the amount of calculation of user's similarity is small, but also needn't be updated frequently [17].

From above analysis on the collaborative filtering algorithm we can see that the amount of users of collaborative filtering algorithm based recommendation model for mathematical expressions is growing slowly. At the same time, the system also contains a large amount of mathematical expressions. In other words, the amount of recommendation contents are far more than the number of users. The user set is basically unchanged. On the other hand, the formula data set is updated frequently. Therefore, this paper employs User-based Recommendation algorithm for math recommendation which can reduce the amount of calculation and get a higher accuracy of formula recommendation.

III. RECOMMENDATION MECHANISM FOR MATHEMATICAL EXPRESSIONS

A. Recommendation Principle of Collaborative Filtering Algorithm

In the collaborative filtering recommendation model for mathematical expressions, a "user-formula" scoring matrix is generated according to the search records and behavioral preferences of users. The matrix usually consists three elements: the user set is defined as $User_i = \{U_1, U_2, \cdots, U_n\}$, the formula set is defined as $Formula_j = \{F_1, F_2, \cdots, F_m\}$ and the scoring matrix as $Point_{ij}$ with $n \times m$ ($i \in \{1,2,\cdots,n\}$, $j \in \{1,2,\cdots,m\}$) (P_{ij} represents the score of user U_i to formula F_j) as (1), and the form of the matrix is shown in TABLE I.

$$Point_{ij} = \begin{cases} 1, & \text{Preference} \\ null, & \text{No record} \end{cases} \qquad (1)$$

TABLE II
"USER-FORMULA" SCORING MATRIX

	$Formula_1$	$Formula_2$	$Formula_3$	$Formula_4$
$User_a$	1	*null*	1	(recommend)
$User_b$	*null*	1	*null*	*null*
$User_c$	1	*null*	1	1

In TABLE III, $Point = 1$ means that the formula has been searched by the corresponding user, that is, a preference record. Otherwise, it means no record. As can be seen from the table, $User_a$ has searched $Formula_1$ and $Formula_3$. $User_b$ has searched $Formula_2$. $User_c$ has searched $Formula_1$, $Formula_3$ and $Formula_4$. By analyzing the search behavior of these users, we can find that the search preferences of $User_a$ and $User_c$ are similar. So it is considered $User_a$ as the neighbor user to $User_c$. At the same time, the $Formula_4$ has searched by $User_c$ and didn't appear in the record of $User_a$. According to the recommendation thought of collaborative filtering algorithm, it can be predicted that $User_a$ may also be interested in $Formula_4$, and then the system recommends $Formula_4$ to $User_a$. The basic principle of the algorithm is illustrated in Fig. 1.

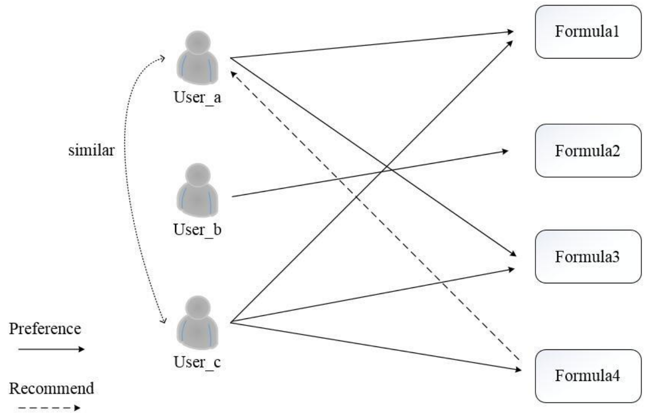

Fig. 1 Basic principle of User-based Recommendation

B. Mathematical Expressions Preprocessing

The premise of collaborative filtering recommendation is "user-formula" scoring matrix. But among the thousands of mathematical expressions. Different formulas might come from different documents related to different users who have different habits, and they use different names of two variables but represent the same meaning of the operation such as "$a+b=c$" and "$x+y=z$".

In this paper, the method of variable substitution is replacing all variables in each mathematical expression by a specific symbol "$\&$". At the same time, constants, operators, positional relations or the formula structure remain unchanged. For example, two formulas in the original dataset are $\sqrt{a^2+b^2}$ and $\sqrt{m^2+n^2}$. Their normalized form is $\sqrt{\&^2+\&^2}$. The corresponding LaTeX formula string is \[\sqrt{{&^2}+{&^2}}\].

In addition, in the process of expression processing, the variables which are replaced by the symbol "$\&$" are usually English letters. But some reserved keywords in LaTeX formula string are also usually composed of English letters. For example, the corresponding LaTeX description of formula $\sqrt{a^2+b^2}$ is \[\sqrt{{a^2}+{b^2}}\]. If the analyzing program regarded the radical sign "\sqrt" in LaTeX as variables, the processed LaTeX formula string would become \[\&&&&{{&^2}+{&^2}}\], which would results in errors and the losses of its original meaning. We designed a data dictionary of LaTeX keywords for storing LaTeX reserved keywords and the key values corresponding to each keyword. Just like α、$\sum$、$\prod$、$\sin\pi$ and many other symbols of LaTeX are included in the data dictionary. And it contains binary operators, Greek alphabets, trigonometric functions and so on.

The algorithm for preprocessing mathematical formula preprocessing is described as follows.

Input: LaTeX description of a mathematical formula;

Output: The mathematical expression string in LaTeX after the substitution of variables;

Step1: Read the data set which contains mathematical expressions in LaTeX format;

Step2: Traverse the string of mathematical expressions in LaTeX format;

Step3: If the next character is an end character, goto Step7, else goto Step4;

Step4: If the current stored character is an English letter, goto Step5, otherwise retain the current character;

Step5: Use the minimum matching method to determine whether the English word in the current location is the reserved keyword in the defined data dictionary in LaTeX. If it is, replace it with the code and goto Step3, otherwise goto Step6;

Step6: If the current character is an English letter variable instead of the LaTeX reserved keywords, replace it with the symbol "$\&$" and goto Step3;

Step7: Generate a new LaTeX string of the corresponding mathematical expression;

Step8: Complete the traversal of the LaTeX string, the algorithm terminates.

C. Multi-feature Fuzzy Pattern

It is too absolute for users to evaluate the contents from different aspects in the traditional collaborative filtering recommendation system, which can't describe the user's preference to the formulas in many ways. To solve this problem, we proposed a recommendation model of mathematical expressions based on collaborative filtering which employed multi-feature fuzzy pattern in users' similarity evaluation [18]. The fuzzy set method with multiple evaluation indexes can effectively avoid the traditional set theory with the logical truth values of $\{0,1\}$. It can availably depict the uncertainty in the evaluation index's decision.

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 12, December 2017

Usually, in the personalized recommendation system, users' behaviors are divided into the following characteristics, such as page browsing time(TIME), hit frequency(HIT), storage(STORAGE) and download(DOWNLOAD). We remember it as a fuzzy set of $X = \{x_t, x_h, x_s, x_d\}$, and $x_t, x_h, x_s, x_d \in [0,1]$. The multi-feature fuzzy score of each user is denoted as a fuzzy subset $\underset{\sim}{X}$ on the domain U. The weight vector of the four evaluation indexes is $w_s (s = 1,2,3,4)$, and it satisfies the conditions of $w_s \in [0,1]$ and $\sum_{s=1}^{4} w_s = 1$.

The inverse tangent is used in this paper to describe the membership functions of the four features. The membership function and characteristic weight vector of the above four features are defined as follows.

(1) User's behaviors: According to the user's browsing time of the web pages or the mouse's stay time to the scroll bar operation, and the user's behaviors to the different formulas of clicking, collecting and downloading. Their membership function is defined as follows.

$$p(x) = \begin{cases} \dfrac{2}{\pi} \arctan x, & 0 \le x < 10 \\ 1, & x \ge 10 \end{cases} \qquad (2)$$

Where x is the four feature elements in the fuzzy set $\underset{\sim}{X} = \{x_t, x_h, x_s, x_d\}$ with multi features. p is the corresponding score for each characteristic element of x corresponding.

(2) Feature weights: According to the importance of the four evaluation features in (x_t, x_h, x_s, x_d), the weight vector is set to $w_s = (w_1, w_2, w_3, w_4) = (0.1, 0.2, 0.3, 0.4)$.

Assume there are n users $U_i (i = 1,2,\cdots,n)$ and m formulas $F_j (j = 1,2,\cdots,m)$ in the recommendation system. The system records the scores of m mathematical formulas in the multi-feature fuzzy pattern according to the n users as shown in TABLE II.

TABLE II SCORING SET TO DIFFERENT FORMULAS BY USERS

	F_1	$\cdots$	F_m
U_1	$(p_{t_{11}}, p_{h_{11}}, p_{s_{11}}, p_{d_{11}})$	$\cdots$	$(p_{t_{1m}}, p_{h_{1m}}, p_{s_{1m}}, p_{d_{1m}})$
U_2	$(p_{t_{21}}, p_{h_{21}}, p_{s_{21}}, p_{d_{21}})$	$\cdots$	$(p_{t_{2m}}, p_{h_{2m}}, p_{s_{2m}}, p_{d_{2m}})$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
U_n	$(p_{t_{n1}}, p_{h_{n1}}, p_{s_{n1}}, p_{d_{n1}})$	$\cdots$	$(p_{t_{nm}}, p_{h_{nm}}, p_{s_{nm}}, p_{d_{nm}})$

In TABLE II, $\underset{\sim}{X}_{nm}$ represents the fuzzy subset formed by U_n to F_m in $(p_{t_{nm}}, p_{h_{nm}}, p_{s_{nm}}, p_{d_{nm}})$.

According to the fuzzy set $\underset{\sim}{X}_{ij}$ in TABLE II, we can calculate the multi-feature comprehensive score P_{ij} as shown in the (3). The "user-formula" scoring matrix can be obtained by combining the three elements of the user U_i, the formula F_j and the score $Point_{ij}$.

$$\begin{aligned} Point_{ij} = P_{ij} &= w_1 \times p_{t_{ij}} + w_2 \times p_{h_{ij}} + w_3 \times p_{s_{ij}} + w_4 \times p_{d_{ij}} \\ &= \sum_{s=1}^{4} (w_s \times \underset{\sim}{X}_{ij}) \end{aligned} \qquad (3)$$

D. User Similarity Measurement

The "user-formula" scoring matrix can be regarded as a fuzzy matrix, and the similarity between two users can be calculated by measuring the similarity between two corresponding fuzzy sets. Considering that the scoring matrix is a two-dimensional matrix containing n users and m mathematical expressions, this paper introduces the principle of selecting the nearest of multi-feature fuzzy recognition pattern. It is defined as follows [18].

Let $\underset{\sim}{A}_1, \underset{\sim}{A}_2, \cdots, \underset{\sim}{A}_n$ be the n known patterns, if $\underset{\sim}{B}$ is closest to $\underset{\sim}{A}_i$, then the $\underset{\sim}{B}$ is classified into the pattern of $\underset{\sim}{A}_i$. This principle is called the principle of fuzzy closeness [18].

In the "user-formula" scoring matrix, the multi-feature comprehensive scores from n users are regarded as n known patterns, and the m formulas are regarded as m characteristics. So there are $n \times m$ fuzzy sets which represent the user's multi-feature comprehensive scores for different formulas

$$[\underset{\sim}{A}_{ij}]_{n \times m} \quad (i = 1,2,\cdots,n; j = 1,2,\cdots,m) \qquad (4)$$

IJIREEICE
ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 12, December 2017

The fuzzy set of multi-feature comprehensive scores from the target user U_q to m formulas is

$$\{\underset{\sim}{B}_j\}\ (j=1,2,\cdots,m) \quad (5)$$

If exist

$$S_{i_0}=\bigvee_{1\le i\le n}\left[\bigwedge_{1\le j\le m}(\underset{\sim}{A}_{ij},\underset{\sim}{B}_j)\right] \quad (6)$$

it is determined that the user U_q is closest to the user U_{i_0} which ranked as the $i_0 th$. That is, the target user U_q and the user U_{i_0} are the nearest neighbors. So we can find the similar users whose closeness degrees [19] with target user are top num, and arrange num analogical users in descending order. The ranking near the top will have a higher degree of fuzzy closeness. The set of close neighbors of the target user U_q is described as the set H_{U_q}.

$$H_{U_q}=\{U_{i_0},U_{i_1},\cdots,U_{i_p}\} \quad (7)$$

The user U_{i_p} means the neighbor user in the set whose ranking is p, $p\in\{1,2,\cdots,num\}$.

There are some unrecorded $(null)$ formulas F_n in the user U_q, and we can combine the corresponding score records of neighbors in the set H_{U_q} for this formulas with the closeness degree weights of this nearest user. Then calculate the unrecorded mathematical formulas' prediction value of the target user U_q in the scoring matrix. The calculation method is

$$\begin{aligned}PV(U_q,F_n)&=S(U_q,U_{i_0})\times P(U_{i_0},F_n)+S(U_q,U_{i_1})\times\\ &P(U_{i_1},F_n)+\cdots+S(U_q,U_{i_p})\times P(U_{i_p},F_n) \quad (8)\\ &=\sum_{a=1}^{p}\left[S(U_q,U_{i_a})\times P(U_{i_a},F_n)\right]\end{aligned}$$

Where $PV(U_q,F_n)$ means prediction value of unrecorded formulas in U_q. $S(U_q,U_{i_p})$ is the closeness degree of the target user U_q and the neighbor user U_{i_p}. $P(U_{i_a},F_n)$ $a\in\{1,2,\cdots,p\}$ represents multi-feature comprehensive score of the neighbor U_{i_a} for the formula F_n in the scoring matrix.

E. Generation of Recommendation List

The interest recommendation list of mathematical expressions is generated according to the formulas of the user list in the neighbor user set. Here, we use an approach called Threshold-Interest List. Its main idea is all points located at the range which has the center of the current point and β as the radius are regarded as the similar interest targets of current point. As shown in Fig. 2. In this method, the number of recommended formulas is uncertain, but the error of similarity is not serious.

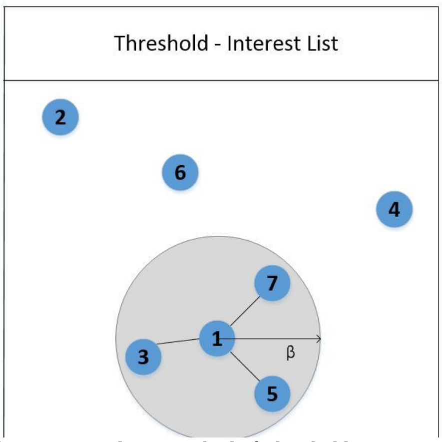

Fig. 2 The recommendation method of Threshold-Interest List

According to the (8), we can calculate the predicted scores of the unrecorded formulas F_n in the list of user U_q and rank scoring results. A threshold variable β is set to further filter the prediction score. In addition, we can obtain x target formulas $F_{\sigma(x)}$ which predictive value is greater than the threshold β and select the screening results as a

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified
Vol. 5, Issue 12, December 2017

mathematical formula interest recommendation list $List\{F_{\sigma(1)}, F_{\sigma(2)}, \cdots, F_{\sigma(x)}\}$. So that the x mathematical expressions with the higher prediction score are obtained in the nearest neighbors, that is, the formulas that the user most interested will be recommended to the target user U_q.

IV. EXPERIMENTAL RESULT AND EVALUATION

The experimental platform is based on the Windows7-64 bits operating system, integrated development environment with Eclipse 4.5.0, JDK version 1.8.0 and database management system MySQL. Hardware environment is Intel processor of Core i7-6700 3.40GHz, 8GB RAM and 1TB ROM.

A. Experimental Evaluation Method

The recommended quality standards of recommender systems usually include three estimated methods called the mean absolute error(MAE), the root mean square error(RMSE) and the normalized mean absolute error(NMAE). This paper uses the RMSE evaluation method because it is sensitive to a group of errors in the measurement performance. So it can reflect the precision of the measurement.

In the experiment, we selected 20% of the data in the "user-formula" scoring matrix as the testing set to compare multi-feature comprehensive score with the proposed prediction value method. The set of prediction scores is expressed as $PV_i \in \{PV_1, PV_2, \cdots, PV_n\}$, and the set of multi-feature comprehensive scores is represented as $P_i \in \{P_1, P_2, \cdots, P_n\}$. The calculation formula of RMSE evaluated method is shown in (9) [20].

$$\begin{aligned} RMSE &= \sqrt{\frac{(PV_1 - P_1)^2 + (PV_2 - P_2)^2 + \cdots + (PV_n - P_n)^2}{N}} \\ &= \sqrt{\frac{\sum_{i=1}^{n}(PV_i - P_i)^2}{N}} \end{aligned} \tag{9}$$

The root mean square error reflects the extent of the predicted data PV_i deviates from the real data P_i. The smaller the value of $RMSE$, the closer the predicted values to the actual values, and the higher the accuracy of the formula recommendation.

B. Experimental Result

The data set used in this experiment is from Wikipedia public data set, which contains several types' mathematical expressions in LaTeX format. The users' scoring table used in the test results contains more than 35 thousand scoring record data of 52 system users for 890 mathematical expressions. The experiment calculates the RMSE value by setting two parameters of the prediction scores' threshold β and the number *num* of nearest neighbours, then analyses and obtains the optimal result. By adjusting the two parameter variables, we can get the results in TABLE III.

TABLE III RMSE VALUES UNDER DIFFERENT PARAMETERS

	2.0	2.5	3.0	3.5
10	1.49474	1.49474	1.49474	1.98815
15	0.98643	0.98643	1.01482	2.26604
20	0.66693	0.66693	0.75515	2.55114
25	0.53887	0.53887	0.69116	3.30938
30	0.50309	0.50309	0.56802	NaN

Comparing the data in TABLE III with the number of neighbouring users and the threshold factor as the abscissa, we can draw the line chart that shown in Fig. 3 and Fig. 4 for comparison.

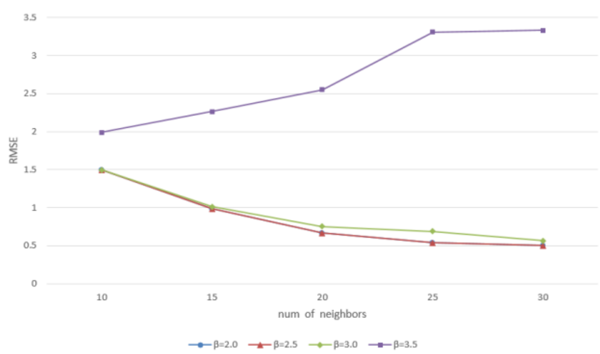

Fig. 3 RMSE values under different *num*

IJIREEICE
ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 12, December 2017

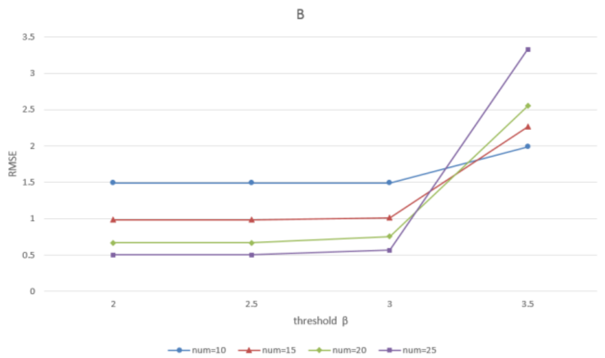

Fig. 4 RMSE values under different threshold

In Fig. 3, it can be observed that most of the broken lines tend to be stable with the increase of the nearest neighbours' number. From Fig. 4 we can know that the range of threshold β value is basically concentrated in $[2.5, 3.5]$, and the most of the fluctuation range is gathered around $[2.5, 3.5]$. This is because the system users' scoring interval is between 0 and 4, and the scores basically conforms to the normal distribution. When the threshold factor β is equal to 2.0 or 2.5 and the nearest neighbour set number num is 30, the system can obtain relatively small value of $RMSE$ and be able to approach the best interest recommended effect of mathematical expressions.

V. CONCLUSION

Collaborative filtering is a widely used technology in the field of interest recommender systems. In this paper, the collaborative filtering algorithm is incorporated into the field of mathematical formula retrieval. A mathematical expression interest recommendation model is proposed which can infer users' intentions from a large set of mathematical expressions according to the behaviour of each user. Hence to save the query time of searching formulas and improve the efficiency of query. However, the model also has some improved space. For example, when the amount of data in the recommendation system is too small, it will cause decline of recommended accuracy. The next step, on the one hand, is to expand the kinds and quantities of mathematical expressions. On the other hand, the users' behaviours in various sides and improve users' information. For the new users who have just joined the system, we can consider the appropriate filling of the matrix according to the personal information and other methods to alleviate the problem of users' cold start.

ACKNOWLEDGMENT

This work is supported by the **National Natural Science Foundation of China** (Grant No. 61375075). The Key Project of the Science and Technology Research Program in University of Hebei Province of China (Grant No. ZD2017208; ZD2017209). The Project of "One Province One University".

REFERENCES

[1] Hao, B.S. (2011) Research on Mathematical Expression Retrieval and Matching Technology. Popular Science & Technology, 5, 58-59.
[2] Math Web Search[EB/OL]. http://kware.eees.iu-bremen.de/, 2017.10.18.
[3] X.hu, L.C.Gao, X.Y.Li,et al. WikiMirs: a mathematical information retrieval system for wikipedia[C]. Proceedings of 13th ACM/IEEE-CS Joint Conference on Digital Libraries. Indiana, United States, Jul. 2013: 11-20.
[4] J.Misutka and L.Galambos. Mathematical extension of full text search engine indexer[C]. Proceedings of 3rd International Conference on Information and Communication Technologies: From Theory to Applications. Damascus,Syria,Apr.2008: 1-6.
[5] MathDex search tool[EB/OL]. http://www.mathdex.com/8080/mathfind/search, 2017.10.18.
[6] R.Miner and R.Munavalli. An approach to mathematical search through query formulation and data normalization[J]. Lecture Notes in Computer Science: Towards Mechanized Mathematical Assistants, 2007, 4573: 342-355.
[7] B.Miller and A.Youssef. Technical aspects of the digital library of mathematical functions[J]. Annals of Mathematics and Artificial Intelligence, 2003, 38(1): 121-136.
[8] P.Libbrecht and E.Melis. Methods to access and retrieve mathematical content in activemath[M]. Mathematical Software-ICMS 2006. Berlin: Springer,2006: 331-342.
[9] P.Libbrecht and E.Melis. Semantic search in leactivemath[C]. Proceedings if the Forst WebALT Conference and Exhibition. Eindhoven,Holland,jan. 2006: 97-109.
[10] Sojka P,Liška M. Indexing and Searching Mathematics in Digital Libraries[C]. Proceedings of International Conference on Intelligent Computer Mathematics. Berlin,Germany: Springer, 2011: 228-243.
[11] G.Y.Kristianto,M.Q.Nghiem and A Aizawa. The MCTA Math Retrieval System for ntcir-10 Math Track[C]. Proc.of the 10th NTCIR Conference. Tokyo,Japan,June 2013: 680-685.
[12] Liu, Z.W. (2011) Research on Mathematical Search Engine. Graduate Thesis, Lanzhou University, Lanzhou.
[13] Gong, L.L. (2013) Research on User Interests Modeling Techniques for Personalized Information Recommendation. Graduate Thesis, Shandong Normal University, Shandong.

[14] PAZZANI M J. A framework for collaborative, content-based and demographic filtering[J]. Artificial Intelligence Review, 1999, 13(5): 393-408.
[15] Yang, J. (2013) Research on User Interest Model and Real-time Algorithm of Personalization Recommendation. Graduate Thesis, Nanjing University Of Posts and Telecommunications, Nanjing.
[16] Xu, H.L. (2009) Comparison Study of Internet Recommendation System. Journal of Software, 20, 350-362.
[17] Guo, X. (2013) Research of New Collaborative Filtering Algorithm Based on Distributed Computing. Graduate Thesis, Xidian University, Xi'an.
[18] Li, S.Y. (2004) Engineering Fuzzy Mathematics with Applications. 130-134.
[19] Li, S.Y. (2004) Engineering Fuzzy Mathematics with Applications. 74-75.
[20] Li, G., Chen, S.H., Han, Z.Y., et al. (2014) Location-aware Recommendation Based on Collaborative Filtering. Computer Science, 41, 340-346.