AHITS-UPT: A High Quality Academic Resources Recommendation Method

Lu Meilian*, Wei Xudan, Gao Jie, Shi Yan
State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, China
*E-mail: mllu@bupt.edu.cn

***Abstract*—Personalized recommendation technology provides the possibility for users to obtain academic resources quickly and accurately. However, the existing recommendation methods based on user's historical behaviors and paper contents are limited in terms of expanding user perspectives. The existing methods that evaluate the authority and quality of academic resources based on academic network ignore paper information or consider some unreasonable information, leading to various quality levels of recommendation results. In order to recommend high quality academic resources to users and expand the horizons of users, we propose an Advanced Hyperlink Induced Topic Search (AHITS) algorithm to evaluate the quality and authority of academic resources, propose a user research interest model based on constructing a tripartite graph, namely User-Paper-Topic (UPT), and propose an academic resource recommendation method based on AHITS and UPT. Experimental results show that the methods presented in this paper can effectively remedy the problem that content based recommendation method is not conducive to expand the horizons of users, recommend authoritative authors and high quality papers to users, improve the accuracy of the recommendation results, and effectively reduce the time complexity of algorithm.**

Keywords—academic recommend; graph model; authority value; quality value

I. INTRODUCTION

With the rapid development of Internet technology, the number of academic resources on the Internet is rapidly growing. Academic search engines can help users quickly find the required information from the massive academic resources, but cannot meet users' personalized demands, and unable to provide satisfactory service in case that users don't have explicit search targets. Personalized academic recommender can mine user preferences according to user's historical behaviors and the operated academic resources, and recommend user with the academic resources meeting user's personalized requirements quickly and accurately.

Recommend authoritative authors and high quality papers are the core requirements of personalized academic information service, and how to evaluate the quality of the recommended academic resources is a key issue. Some of the existing schemes evaluate the authority of authors and the quality of papers through constructing academic network and using the original or improved HITS algorithm or PageRank algorithm. Wu Harris [1] uses HITS algorithm to calculate the user's hub and the authority of document by constructing user-label-document tripartite graph. Antonietta Grasso [2] constructs user-document network and user-author network, and uses the HITS algorithm to respectively compute the document quality and the author authority based on the above both networks. Men Rui [3] builds citation network according to the citation relationship, and proposes a method to evaluate the initial quality value of paper using journal impact factor, title of author and fund of paper, then uses the PageRank algorithm to measure the citation weight, finally obtains the final paper quality according to the initial paper quality and the citation weight. However the above methods do not consider content information and time information, some of the literatures do not consider the level of journals and authors, meanwhile, the time complexity of the algorithm is high, and the recommendation effect is not satisfied. Yunhong Xu [4] proposes a recommendation method based on bi-layer concept network, which combines the semantic and social network information together, and measures the weights of edges in the network through semantic relevance and social networks. However, the concept layer only considers the keywords information, without considering other paper information. Content-based personalized recommendation method [5, 6] recommends academic resources by constructing user interest model. This kind of methods can recommend the most similar academic resources to user, but they have limitations in broadening user's research perspectives. Wang Yanran [7] introduces time parameter and literature importance into the traditional content-based recommendation method to remedy the defects in terms of distinguishing literature quality under the same theme. However, it does not take into account the topic information in building literature model.

In order to remedy the defects of existing quality and authority evaluation methods, utilizing the mutually reinforcing relationship of author authority and paper quality, and considering author level, citations, publish time, and the level of journal/conference, an improved HITS algorithm, named as Advanced Hyperlink-Induced Topic Search (AHITS), is proposed to calculate the author authority and the paper quality and filter out the authors and papers with lower authority and quality from the recommendation results. Furthermore, in order to remedy the defects of traditional

This work is partially funded by the National Science Foundation Projects (Grant No. 61300183 and Grant No. 61471060).

978-1-5090-1893-2/15 $31.00 © 2015 IEEE
DOI 10.1109/SmartCity.2015.120

content-based academic resource recommender method in broadening user perspectives, a tripartite graph model named as User-Paper-Topic (UPT) is constructed to mine user interest based on the relationships between users and papers and the relationships between papers and topics. Based on AHITS algorithm and UPT model, the authoritative authors and the high quality papers can be recommended.

II. AHITS: EVALUATE AUTHOR AUTHORITY AND PAPER QUALITY

In this part, the HITS algorithm is improved, and a new method referred as AHITS (Advanced hyperlink induced topic search) is proposed. AHITS iteratively calculates the author authority and paper quality by constructing author-paper relation network and combining with author level, author sequence number in the paper, citations, journal/conference level and time information.

In the traditional HITS algorithm, the initial value of authority and hub are set to 1, and the algorithm's time complexity is relatively high. In order to make the HITS algorithm converge fast and reduce the time complexity, in AHITS algorithm, the authority value is initialized according to author level, and the quality value is initialized based on paper citations, publish time and the level of journal/conference.

The author authority is initialized as a_{i_0}, which indicates the level of author r_i. In this paper, the author level is categorized as professor, associate professor, lecturer, assistant professor, Ph.D student and graduate student, which are respectively assigned values as 15, 10, 8, 5, 3 and 2 [3].

The paper quality is initialized as $q_{j_0} = \zeta_j + H_j e^{-\varphi(t-\tau_j)}$. Where, ζ_j indicates the level of the journal/conference where paper p_j published, and which is indicated with the influence factors of the journal/conference. H_j is the citations of the paper. φ is time decay factor, the value is 0.62 [8]. t is the current time, and τ_j is the publication time of the paper.

Based on the idea of HITS, the authority of author r_i is calculated as follows:

$$a_i = (1-\lambda)\sum_{j\in V(i)} (\omega_{ij} q_j) + \lambda a_{i_0} \tag{1}$$

Where, λ is the weight of author authority, $V(i)$ is the paper collection published by author r_i, q_j is the quality value of paper p_j. $\omega_{ij} = g_{ij} / \sum_{z=1}^{M_j} z$ is the weight of author r_i in paper p_j, which is determined by the author sequence number in the paper. M_j is the number of authors in paper p_j, g_{ij} is the reverse sequence number of author r_i in paper p_j.

The quality value of paper p_j is calculated as:

$$q_j = (1-\gamma)\sum_{i\in E(j)} (\omega_{ij} a_i) + \gamma q_{j_0} \tag{2}$$

Where, γ is the weight of quality value, $E(j)$ is the author collection of paper p_j.

According to (1) and (2), we can iteratively calculate the author authority value and the paper quality value.

III. UPT: USER INTEREST MODELING METHOD

In this paper, a tripartite graph model is proposed which combines user behaviors and paper contents, namely User-Paper-Topic (UPT). Based on the tripartite graph model, we can dig out the user's preferences for papers and construct user interest model, and further explore the user's potential research interest.

A. UPT model

UPT is a directed graph, which is comprised of user layer, paper layer, topic layer, and the edges between adjacent layers, as shown in Fig. 1. The user layer comprises all the users who have operated papers, the paper layer consists of all the papers operated by the users in user layer, and the topic layer is composed of all the topics the papers in the paper layer belong to. Where, the topics can be extracted from the title, abstract, and key words of all papers in the data set through LDA model. The topic distribution of paper P_j is denoted as $(T_{j1}, T_{j2}, ..., T_{jk}, ..., T_{jK})$. Where, T_{jk} is the weight of paper p_j in topic t_k, and $\sum_{k=1}^{K} T_{jk} = 1$. K is the specified topic number.

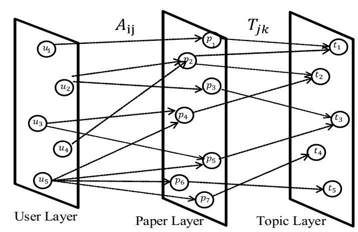

Fig. 1. Hierarchy User Interest Model

In Fig.1, A_{ij} represents user u_i have operated paper p_j and indicates the preferences of user u_i for paper p_j. T_{jk} represents that the weight of paper p_j in topic t_k is greater than a specified threshold and indicates the topic distribution of paper p_j in topic t_k.

The user's behaviors on paper include reading, collecting, sharing, downloading and scoring. We define the weight of the edge between user u_i and paper p_j as $A_{ij} = f_{ij} \times e^{-\mu(t-t_{ij})}$. Where, $e^{-\mu(t-t_{ij})}$ is time attenuation coefficient, which means that, with the passing of time, the user preference for paper will reduce. Considering that the interests of users generally change after a longer period, the value of the attenuation factor μ is set as 0.8. t indicates the current year, and t_{ij} indicates the year user u_i operated paper p_j. f_{ij} is the score user u_i for paper p_j, which is calculated as:

(a) If u_i has an explicit score on p_j, then the value of f_{ij} is equal to the score;

(b) If u_i hasn't an explicit score on p_j, then $f_{ij} = \begin{cases} 5 & downloading \\ 4 & sharing \\ 3 & collecting \end{cases}$.

B. Modeling user interest

Based on the content and topic characteristics of the papers user have operated, we define the interest model of user u_i as:

$$M_{ui} = \{U_{Ti}, U_{Bi}, U_{Ci}, U_{Di}\} \quad (3)$$

Where, U_{Ti} is the collection of user's current research topics, U_{Bi}, U_{Ci} and U_{Di} are the user's title feature space vector, abstract feature space vector, and keywords feature space vector. U_{Ti} reflects the research topics user interested, and U_{Bi}, U_{Ci}, U_{Di} reflect the paper content user interested.

According to the UPT model, the preference of user u_i for topic t_k is defined as:

$$P_{u_i t_k} = \sum_{j=1}^{J} A_{ij} T_{jk} \quad (4)$$

Sort all the interested topics of user u_i by the size of preference value, and select the previous W topics to construct the user's current research topic collection U_{Ti}.

Assuming that the papers user u_i operated is R_i, extract the title, abstract and keywords of these papers, and construct the user's title feature space vector U_{Bi}, abstract feature space vector U_{Ci} and keywords feature space vector U_{Di}.

$$\begin{aligned} U_{Bi} &= \{(B_{i1}, b_{i1}), (B_{i2}, b_{i2}), ..., (B_{ix_i}, b_{ix_i}), ..., (B_{iX_i}, b_{iX_i})\} \\ U_{Ci} &= \{(C_{i1}, c_{i1}), (C_{i2}, c_{i2}), ..., (C_{iy_i}, c_{iy_i}), ..., (C_{iY_i}, c_{iY_i})\} \\ U_{Di} &= \{(D_{i1}, d_{i1}), (D_{i2}, d_{i2}), ..., (D_{iz_i}, d_{iz_i}), ..., (D_{iZ_i}, d_{iZ_i})\} \end{aligned} \quad (5)$$

In (5), X_i, Y_i, Z_i are separately the total numbers of title feature words, abstract feature words, and keywords, x_i, y_i, z_i are the sequence number of the feature words. B_{ix_i} and b_{ix_i}, C_{iy_i} and c_{iy_i}, D_{iz_i} and d_{iz_i} are separately $x_i th$ title feature word, $y_i th$ abstract feature word, $z_i th$ keywords feature word and their TF-IDF values.

C. Predicting potential user interest

According to (4), the user's topic preference vector is constructed as $Tu_i = (P_{u_i t_1}, P_{u_i t_2}, ..., P_{u_i t_k}, ... P_{u_i t_K})$.

Where, $P_{u_i t_k}$ indicates the preference of user u_i for topic t_k.

The interest similarity between user u_i and u_s is calculated by cosine similarity:

$$sim(u_i, u_s) = \cos(Tu_i, Tu_s) = \frac{\sum_{k=1}^{K} P_{u_i t_k} \times P_{u_s t_k}}{\sqrt{\sum_{k=1}^{K} (P_{u_i t_k})^2} \times \sqrt{\sum_{k=1}^{K} (P_{u_s t_k})^2}} \quad (6)$$

Then, the similar users of user u_i are sorted in descending order according to $sim(u_i, u_s)$, and the front Q users are selected as the final similar users of user u_i, which is denoted as a collection $v(i,Q)$.

Finally, the preferences of user u_i for the topics which the users in $v(i,Q)$ are interested can be calculated:

$$P'_{u_i t_{k'}} = \overline{P_i} + \frac{\sum_{s \in v(i,Q)} sim(u_i, u_s)(P_{u_s t_{k'}} - \overline{P_s})}{\sum_{s \in v(i,Q)} sim(u_i, u_s)} \quad (7)$$

Where, $P'_{u_i t_{k'}}$ is the predicted preference of user u_i for topic $t_{k'}$, $P_{u_s t_{k'}}$ is the preference of user u_s for topic $t_{k'}$, $\overline{P_i}$ and $\overline{P_s}$ are respectively the average preference values of user u_i and user u_s for the topics related with the papers that have been operated by user u_i and user u_s.

According to the predicted preference value in (7), we select F topics with highest preference values as user's potential interest topics U'_{Ti}.

Based on the above prediction scheme, we experimented and obtained the user's interest topics and potential interest topics as shown in Table 1. In Table 1, the research direction of user 1 is about peer to peer network performance, the recommended topics of user 1 include network security and routing protocol, both are related with the user interest topics and also expand the research fields of the user 1. The interest topics of user 2 are related with the semantic and decision tree, and the recommended results include Support Vector Machine, which is a branch of decision tree, and contains the contents of Database Query. The research direction of user 3 is about recommender system, the recommended results include formal concept analysis, which can be used for the construction of ontology. Therefore, predicting potential interest can effectively expand the user perspective.

TABLE I. INTEREST TOPICS AND PREDICTED INTEREST TOPICS

	User 1	*User 2*	*User 3*
User Interest Topics	Peer to Peer Network	Semantic Knowledge	Cluster Algorithm
	Quality of Service	Matrix Compute	Search Engine
	Network Protocol	Decision tree	Similarity Measure
	Network Performance	Random Smpling	Author Recommender
Potential User Interest Topics	Routing Protocol	Database Query	Formal Concept Analysis
	Network Security	Support Vector Machine	Prediction Theory

IV. RECOMMEND AUTHORITATIVE AUTHORS AND HIGH QUALITY PAPERS

Based on the AHITS algorithm proposed in section II and the UPT model proposed in section III, we can recommend authoritative authors and high quality papers for users. The recommendation algorithm is as follows:

Step 1: According to user interest modeling method and the potential user interest predicting method in section III, we can get the target user's current research topic collection and potential interest topic collection, and compose both to construct the final user interest topics.

Step 2: Extract the papers and the related authors from each interest topic of the target user u_i, and construct the author-paper network for each topic.

Step 3: According to the AHITS algorithm proposed in section II, calculate the authority value of each author and the quality value of each paper in each interest topic, and sort them in descending order.

Step 4: For each interest topic t_k of the target user u_i, select the front E author with the highest authority values to construct the target user's initial candidate authoritative author recommendation list in this topic, merge all the lists of all topics and delete the replicated authors, then the final candidate authoritative author recommendation list is generated.

Step 5: Calculate the authority value $a'_{r_j u_i}$ of each author r_j in the final candidate authoritative author recommendation list for the target user u_i:

$$a'_{r_j u_i} = P_{u_i t_k} * a_{r_j t_k} \tag{8}$$

Where, $a_{r_j t_k}$ is the authority value of author r_j in topic t_k.

According to the values calculated in (8), resort the authors in the final candidate authoritative author recommendation list, and select the *topN* authors to form the target user's final authoritative author recommendation list.

Step 6: For each interest topic t_k of the target user u_i, calculate the similarity between each paper p_j in the paper list sorted in step 3 and the target user u_i:

$$\begin{aligned} sim(u_i, p_j) = \theta sim(U_{Bi}, P_{Bj}) + \sigma sim(U_{Ci}, P_{Cj}) \\ + \psi sim(U_{Di}, P_{Dj}) \end{aligned} \tag{9}$$

Where, U_{Bi}, U_{Ci}, U_{Di} respectively are the title feature space vector, abstract feature space vector and keywords feature space vector of target user u_i. P_{Bj}, P_{Cj}, P_{Dj} respectively are the title feature space vector, abstract feature space vector and keywords feature space vector of paper p_j. θ, σ, ψ are respectively set as 0.4, 0.2, 0.4[5].

As an example, here give the calculation method of $sim(U_{Bi}, P_{Bj})$ in (9):

$$sim(U_{Bi}, P_{Bj}) = \frac{\overrightarrow{U^*_{Bi}} \bullet \overrightarrow{P^*_{Bj}}}{\left|\overrightarrow{U^*_{Bi}}\right| \times \left|\overrightarrow{P^*_{Bj}}\right|} \tag{10}$$

Where, U_{Bi} and P_{Bj} are set as two-tuples (key, value), and $\overrightarrow{U^*_{Bi}}$, $\overrightarrow{P^*_{Bj}}$ are the corresponding feature vector.

$$U_{B_i} = \{(word_{i1}, w_{i1}), (word_{i1}, w_{i1}), ..., (word_{in}, w_{in})\}$$

$$P_{B_j} = \{(word_{j1}, w_{j1}), (word_{j1}, w_{j1}), ..., (word_{jm}, w_{jm})\}$$

Set $Word_i = \{word_{i1}, word_{i2}, \cdots word_{in}\}$,

$Word_j = \{word_{j1}, word_{j2}, \cdots word_{jm}\}$,

$S = Word_i \bigcup Word_j = \{s_1, s_2 \cdots s_k \cdots s_z\}$, $Z = |S|$.

$$\forall s_k \in S \qquad w_{ik} = \begin{cases} w_{ix} & s_k \in Word_i \\ 0 & s_k \notin Word_i \end{cases}$$

$$w_{jk} = \begin{cases} w_{jx} & s_k \in Word_j \\ 0 & s_k \notin Word_j \end{cases}$$

$$\overrightarrow{U^*_{Bi}} = (w_{i1}, w_{i2} \cdots w_{iZ}), \ \overrightarrow{P^*_{Bj}} = (w_{j1}, w_{j2} \cdots w_{jZ}) \tag{11}$$

Using the similar method, we can get:

$$sim(U_{Ci}, P_{Cj}) = \frac{\overrightarrow{U^*_{Ci}} \bullet \overrightarrow{P^*_{Cj}}}{\left|\overrightarrow{U^*_{Ci}}\right| \times \left|\overrightarrow{P^*_{Cj}}\right|}, \quad sim(U_{Di}, P_{Dj}) = \frac{\overrightarrow{U^*_{Di}} \bullet \overrightarrow{P^*_{Dj}}}{\left|\overrightarrow{U^*_{Di}}\right| \times \left|\overrightarrow{P^*_{Dj}}\right|} \tag{12}$$

According to the value of $sim(u_i, p_j)$ in (9), delete the papers the similarity is less than a specified threshold, in this paper we specify the threshold as 0.2 [6], delete the papers the target user have operated, and select the former N_k paper in topic t_k as the candidate paper recommendation list for target user u_i in current topic. Merge all the candidate paper recommendation lists of all topics, and delete the repeated papers, then we can get the final high quality paper recommendation list of target user u_i.

V. EXPERIMENTS AND RESULTS

A. Experimental scheme

We verify the methods proposed in this paper based on an academic recommendation website, which is currently in the trial operation stage. Users of the website can search, download, read, score, or share papers and authors. The dataset in our experiments include 160000 papers in the fields of computer and 185279 authors obtained from Microsoft Academic Search (http://academic.research.microsoft.com/). We select 10 registered users, 856 user behaviors to evaluate our scheme. The topic number of LDA model is set as $K = 100$. Time complexity of AHITS algorithm is measured by the iteration number of the author authority and the paper quality, and effectiveness of the algorithm is evaluated through the recommendation accuracy of papers and authors.

Three comparison schemes are selected, which are HITS-based recommendation method [1], MHITS-based algorithm [2] and traditional content-based filtering recommendation method [7]. Our proposed method is referred as AHITS-based. Through following experiments, related parameters are determined to ensure the optimal recommendation effect, and the advantages of our proposed model and algorithm are demonstrated in this paper.

Experiment1: Determine the optimal value of the weight parameters λ and γ in (1) and (2) according to the iteration number of AHITS, the authoritative author rate and the high quality paper rate. First, several better pairs of λ and γ is determined based on the iteration number, then calculate the authoritative author rate and the high quality rate in the recommendation result in the condition of the better parameter pairs, and determine the optimal λ and γ .

Experiment 2: Compare the iteration number of AHITS, HITS [1] and MHITS [2], and verify the method proposed in this paper has lower time complexity.

Experiment 3: Compare the authoritative author rate and the high quality paper rate of AHITS-Based, HITS-Based and MHITS-Based algorithm, and verify the advantages of the proposed method in this paper.

Experiment 4: Further verify the performance of the AHITS-Based recommendation method in terms of accuracy rate, in which the comparison scheme is the traditional Content-Based recommendation method [7].

B. Experimental results analysis

1) Determine the optimal value of λ and γ

Fig. 2 gives the influence of weight parameters λ in formula (1) and γ in formula (2) on the iteration number of AHITS algorithm. From Fig. 2 we can see, when $\lambda = 0.9$ and $\gamma = 0.9$, the iteration number is the least. However, we determine the final λ , γ values by obtaining the authoritative author rate and high quality paper rate based on six pairs of λ , γ values with lower iteration number.

The effects of λ and γ on authoritative author rate and high quality paper rate are respectively presented in Fig.3 and Fig. 4. From Fig.3 and Fig.4, we can see, when $\lambda = 0.9$ and $\gamma = 0.7$, the high quality paper rates of all users are relatively high, while the authoritative author rate is hardly affected. Therefore, taken together the iteration number, high quality paper rate and authoritative author rate, the optimal pair of λ and γ is determined as $\lambda = 0.9$ and $\gamma = 0.7$.

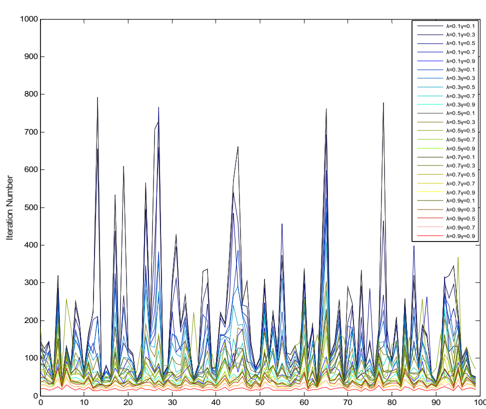

Fig. 2. Influence of λ , γ on Iteration Number of AHITS Algorithm

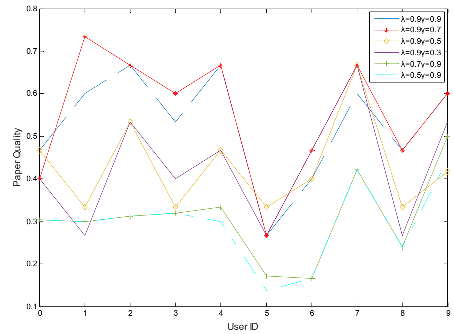

Fig. 3. Influence of λ and γ on AHITS-based Paper Recommendation

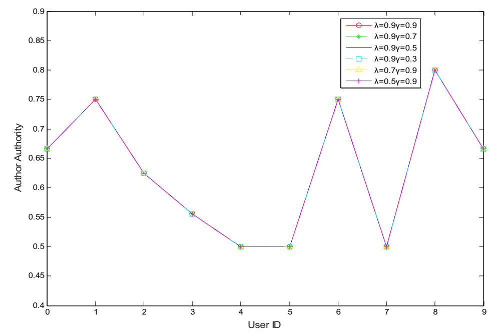

Fig. 4. Influence of λ and γ on AHITS-based Author Recommendation

2) Comparison of time complexity of different algorithms

In case of $K = 100$, the iteration number of AHITS, HITS [2] and MHITS [3] are compared in Fig.5. It can be seen, the iteration number of the proposed AHITS in this paper is between 4-40 times for majority topics. While the iteration number of HITS and MHITS are mostly between 100-600, even more than 1000 times for some topics. So in terms of iteration number, our method has lower time complexity. This is mainly because that the method and algorithm proposed in this paper take into account multiple factors which can reflect the author authority and paper quality, use the author level to initialize the authority value, and use the citation number, publishing time and the level of journal/conference to initialize the paper quality.

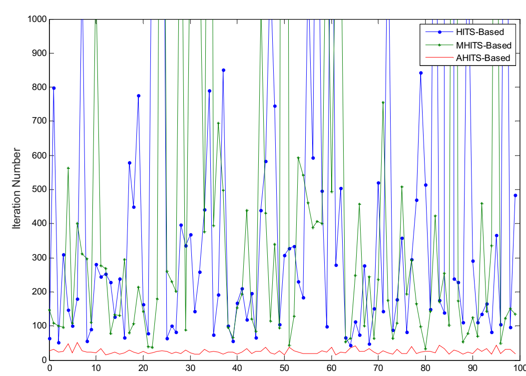

Fig. 5. Iteration Number of Different Author Authority/Paper Quality Algorithm

3) Comparison of the authoritative author rate and high quality paper rate of different algorithms

Fig. 6 and Fig. 7 respectively show the authoritative author rate and high quality paper rate of the three kinds of algorithms. From Fig. 6 and Fig. 7, we can see that the AHITS algorithm proposed in this paper is obviously superior to the comparison schemes not only in terms of the authoritative author rate but also the high quality paper rate.

4) Recommendation effect of AHITS algorithm

Fig. 8 compares the proposed scheme with the Content-Based [7] scheme in terms of the accuracy of recommended paper. In Fig. 8, the horizontal coordinates are the user ID, and the vertical coordinates are accuracy rate of recommended paper. For considering the characteristics of topic information and user interest in the potential direction, the user interest model constructed in this paper more accurately characterize the user's interest, and ensure the authority of recommended authors and the quality of recommended papers. So, the recommendation accuracy of this paper's method is obviously superior to the compared schemes.

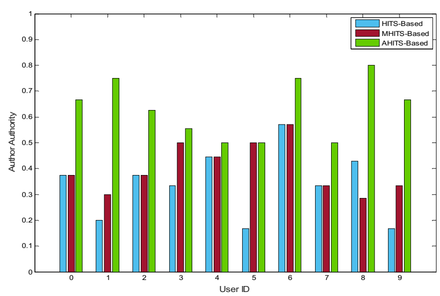

Fig. 6. Author Authority Rate of Different Author Authority/Paper Quality Algorithm

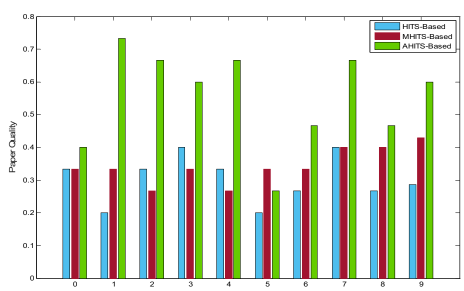

Fig. 7. Paper High Quality Rate of Different Author Authority/Paper Quality Algorithm

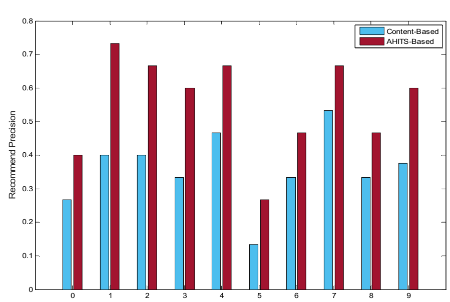

Fig. 8. Recommendation accuracy of different recommendation method

VI. CONCLUSIONS

Aiming at the defects of existing academic recommendation methods in terms that they are not conducive to expand the visions of researchers and the quality of the recommended academic resources is not high, a new user interest modeling method based on tripartite graph model and a new algorithm for calculating academic authority value and quality values are proposed. The tripartite graph model characterizes the user preference for related topics through combining user behaviors, paper content and paper topics. The authority and quality algorithm comprehensively considers factors which influence the author authority value and the paper quality value, uses the author level to initialize the authority values, and uses publish time, cited number and journal/conference level to initialize the quality value. Experimental results verify that the proposed method can effectively solve the problem of uneven quality of academic resources with lower time complexity, and can expand the user perspectives. However, because the tripartite graph model and user interest model rely on user behaviors, the cold start problem still exists. The future work will try to solve the problem.

References

[1] Harris Wu, Mohammad Zubair, Kurt Maly. Harvesting social knowledge from folksinomies[C]. HYPERTEXT'06: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, August 22-25, 2006, Odense, Denmark, ACM, pp111-114.

[2] Antonietta Grasso, Andre Bergholz. Method and system for expertise mapping based on user activity in recommender systems: US7240055B2. 2007.

[3] Men Rui. Study on high-quality personalized paper recommender system[D]. Tian Jin University. 2011

[4] Yunhong Xu, Xitong Guo, Jinxing Hao, Jian Ma, Raymond Y.K. Lau, Wei Xu. Combining social network and semantic concept analysis for personalized academic researcher recommendation[J]. Decision Support Systems. 2012, 54(1), pp 564-573.

[5] Yueheng Sun,WeijieNi, Rui Men. A personalized paper recommendation approach based on web paper mining and reviewer's interest modeling[C]. 2009 International Conference on Research Challenges in Computer Science,December, 28-29, 2009, Shanghai, China, IEEE, pp 49-52.

[6] Kazunari S, Kan M Y. Scholarly paper recommendation via user's recent research interests[C]. Proceedings of the 10thAnnual Joint Conference on Digital Libraries, June 21-25, 2010, Gold Coast, Queensland, Australia, ACM, pp 29-38.

[7] Wang Yanran, Chen Mei, Wang Hanhu, Zhang Xin. A content-based filtering algorithm for scientific literature recommendation[J]. Computer Technology and Develpoment, 2011, 21(2): 66-69.

[8] Yujing Wang, Yunhai Tong, Ming Zeng. Ranking scientific articles by exploiting citations, authors, journals, and time information[C]. Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14–18, 2013, Bellevue, Washington, USA, AAAI Press, pp 933-939.