IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

Recommendation Scheme for Research Studies via Graph Based Article Ranking

Dhinakaran D[1], Joe Prathap P.M[2]

Assistant Professor, Department of Computer Science & Engineering, Peri Institute of Technology, Mannivakkam, Kanchipuram, India[1]

Associate Professor, Department of Information Technology, RMD Engineering College, Kavaraipettai, Tiruvallur, India[2]

Abstract: Nowadays, the recommendation practice are becoming more popular as a result of expeditious growth in the information technology and the growing amount of scholarly data. The ultimate purpose of the proposed work is recommending the items of potential interest automatically to the users. In modern years, recommender systems are being employed in digital libraries to help researchers in finding relevant articles. This search tool gives the roll of relevant articles support on the keywords in the query. It is difficult for a beneficiary to describe his needs by a limited set of keywords. On the divergent, the results will be similar to all researchers if the keywords are same. The existing recommender scheme use the same algorithm to produce suggestion for users without considering their situation. To improve the recommendation, the proposed system incorporates common author relation between articles. The system suggests the articles for users who are having author based search pattern and also recommends conference venue for all researchers to assist them to gain more knowledge in their interested area. The principle notion of the project is to determine the users having author based search pattern. To serve this purpose, the system uses pair wise articles with common author relations and most frequently appeared author.

Keywords*:* Common Author Relation, Frequently Appeared Author, Article Recommendation, Recommends Conference Venue, Pair wise Articles.

I. INTRODUCTION

To serve the purpose of generating a new research idea and for writing a good article, researchers need to intensively read some scientific articles that are correlated to their comfort. With the rapid development in the volumes of scientific articles finding relevant articles is flattering further and more challenging for researchers. Academic recommender system aims to work out the problem of information overload in big learned data such as finding appropriate research paper, appropriate publication venue, etc. At present, web search tools employed by scientific digital libraries like IEEE Explore, and literature search engines like Google Scholar, can serve the purpose of retrieving a list of appropriate articles in different technological fields via keyword-based queries. However, these search tools have several weakness as follows: (i) The searchers' need cannot be described by limited keywords (ii) The obtained results are same for all researchers, if the keywords are identical. (iii)It is ambiguous to look for articles when a researcher has no information of what they are look for.

Article-researcher recommender system aims to repeatedly recommend modified articles of potential interest. intellectual recommender systems seek to solve the information surplus

crisis in big intellectual data such as finding applicable research paper, applicable journal location, etc. shows the corresponding recommendation tasks in above-mentioned

scenarios, including (1) User registration and profile creation- create profile for registered users (2) Article search and user's library creation - create new library for saved articles for each user (3) Target researcher selection –user selection (4) Graph construction-construct graph for most habitually become visible authors (5) Graph based article ranking-Find the most relevant articles by comparing extracted articles with other user's library (6) Publication preprocessing-Remove the special characters and 'journal' keywords from publications (7) Conference venue recommendation-for suggesting a topic relevant conference or journal to bring out a novel article.

In this paper, we spotlight on article recommendation and recommends conference venue i.e., common author relation between articles and also recommends conference venue for all researchers to help them to gain knowledge in their interested area. Recommender systems are being employed in the digital libraries to help the researchers in finding applicable articles. Online access services such as Google Scholar, ACM Portal, IEEE Xplore, Elsevier and Springer Link have been able to provide powerful web search tools to retrieve relevant articles in diverse research areas. But these search tools have the following disadvantages given below. First, only several keywords are not enough to represent researcher's need for articles of interest. The usage of incorrect keywords or those keywords involving in multiple research areas may generate inadequate search results.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

Second, the search outcome are not personalized. For any researcher, the results are the same if only the used keywords are also the same. Third, it is impossible to search articles when researchers, especially junior researchers, have no obvious idea of what they are looking for.
The rest of the paper is provide as pursue, Section 2 reviews associated work on article recommendation. Section 3 presents system design and our module descriptions. Section 4 introduces the problem declaration. Section 5 evaluates metrics and Section 6 finally concludes the paper.

II. RELATED WORK

A. Graph Based Article Ranking

Xia et al. [1] proposed a method that aims in recommending relevant articles for target researchers. Author information is used for building relations between articles. Then, the proposed scheme uses this article relation and researchers historical preference to make a graph for article ranking. Random-walk with regenerate algorithm is worn to rank the articles for making the recommendation. The idea of including familiar author associations is that the sustained growth of internet technology permits researchers to easily contribute to publications with others, which makes convenient for researchers who have search pattern based on authors to search articles published by same authors. This method defines features to find relevant target researchers by analyzing researcher historical preference.
To determine target users the system considers the following features. The ratio of the total number of pair wise articles with common author relations to the total number of all pair wise articles and the ratio of the occurrence number of the most frequently appeared author in articles to the total number of articles for a researcher. Experiments were conducted using a real-world dataset site you like to assess the influence of the two features and the proposed method produces more precise recommendations for relevant researchers when compared to a Baseline method.

B. Citation-Based Scientific Article Recommendation

Liu et al. [2] presented a novel article recommendation method named as Citation-based scientific Article Recommendation (CAR). The proposed method integrates the details of beneficiary historical fondness and citation associations among articles. The system considers that though researchers usually find articles of interest by searching citations, not all pair wise articles with citation relations are really relevant. So the projected method first filters the weak citations through an association mining technique based on researchers' historical preference. Then, these filtered citation relations are included into a graph-based article ranking method for improving recommendation quality.
This proposed method builds a graph where vertices are researchers and articles, and run a random walk with restart algorithm in the graph to compute rankings of relevant articles to generate top-N suggestion for target researchers. This method is evaluated using real-world dataset and the result shows that this method outperforms other existing baseline methods in provisions of precision, recall, and F1. The proposed method didn't consider author relations between articles.

C. Collaborative Topic Modeling

Wang, Chong, and David M. Blei [4] developed an algorithm for suggest scientific articles to the user. Each user saved his interesting article in his library, and this system goal is to find relevant articles for the recommendation. The system recommends older articles and new articles to users. Older articles are recommended for learning the details about the fields and to understand the foundations of their fields. New articles are recommended to the users to let them know about the trending concepts in their interested research area.
This method integrates the advantages of conventional mutual filtering and probabilistic subject modeling. It offer an illustratable concealed configuration for users and items and recommends both existing and new articles. This algorithm is evaluated against real-world data from site you Like and proved that it provides a more successful recommender scheme than conventional collaborative filtering. This method didn't consider author relation and citation relation between articles.

D. Context-Aware Citation Recommendation

He et al. [5] presented an proposal of building a substance-aware citation recommendation system. Suggests high-quality citations is a difficult one. The system developed a substance-aware approach to recommends a set of citations with high feature. A model system in CiteSeerX is implemented. CiteSeerX is a digital library experimented with many baselines methods and the result shows the efficiency and the flexibility of this approach.

E. Topic-Based Citation Recommendation

Tang, Jie, and Jing Zhang. [6] proposed a two-layer Restricted Boltzmann Mechanism (RBM) model, referred to as RBM-CS. The representation learns a combination of subject allocation over paper stuffing and citation associations by giving a collection of papers with citation relationship. A method is presented to rank papers for a given citation context, based on the learned topic model. The top-ranked papers are recommended. A method is described to find the

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

association between the suggested papers and the citation sentences. Experiment results show that RBM-CS can drastically outshine baseline scheme for citation suggestion.

F. Social Network Analysis

Luong, Hiep, et al [3] presented a new entrance that allows researchers to automatedly find suitable journal location for their research paper by exploring author's network of related co-authors. Then the proposed system bring in a new entrance using the social network study to explore author network's publication history in order to produce relevant conferences venues. Trial with a set of ACM conferences show that this new approach outshines the content-based approach and offers a better recommendation.

III. SYSTEM DESIGN

The system retrieves saved article from user's library and construct the graph using the methods mentioned in the architecture. Graph-based article ranking [1] is used to recommend the relevant articles to the users who are having author based explore prototype.

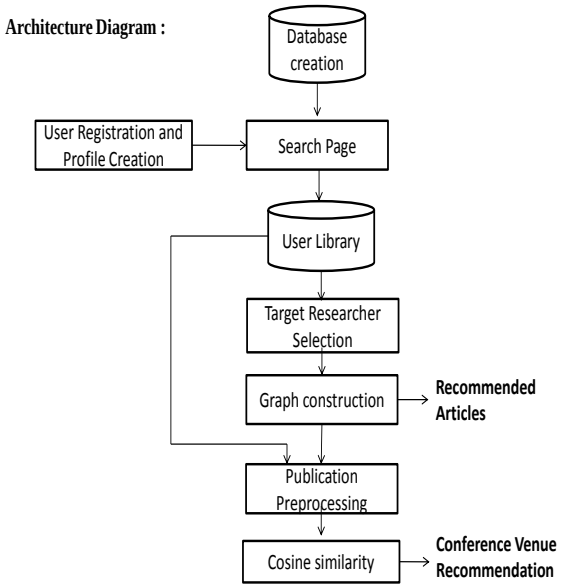

Fig.1 System architecture

A. User Registration And Profile Creation

The collaborative environment is created where the new user can record by providing the necessary details. These details are stock up in the user database and the profile is produced for the registered user. After the successful registration, the user can login using valid credentials.

Input: User details.
Output: User's successful login.
Algorithm:

```
Login page creation.
begin
    if already a member then
        login
    else
        register
 end.
Profile creation for registered users.
```

B. Article Search And User's Library Creation

The user can search the articles and can view the abstract of the article. If the user wants to read the full article then he can use the link provided which will navigate to the entire article. The user can save the article in his library by clicking the save button if he is interested in that article.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

Input: List of Articles
Output: Saved Articles
Algorithm:

```
for each researcher do
  for each keyword do
  obtain list article, author, publication through e-search
      for each  saved article do
   Create new file and store article, author, publication
      end for
    end for
  end for
```

C. Target Researcher Selection

The system dynamically validate whether the user having author based search pattern or not. To find that the system extracts the details from user library and calculates the following kind: The ratio of the total number of pair wise articles with common author relations to the entire number of all pair wise articles and the ratio of the occurrence number of the most frequently appeared author in articles to the entire number of articles. Then compare these features with a threshold value to find whether the user is having author based search pattern or not.

Input: Researcher
Output: Target researcher or Not
Algorithm:

```
To construct writing relation graph
  Extract details from researcher's library.
   if there exist a relation between article and author then
      draw edge between article and author.
    end if
To construct common author relation graph
   if there exist a common author between articles then
      draw edge between those articles.
    end if
```

D. Graph Construction

The user-article, author-article and article-article graph is constructed for recommending articles to the target user using graph-based article ranking algorithm. The user-article graph represents the reading relation between user and article. Author article graph represents the writing relation between author and article. Similarly, the article-article graph [6] represents common author relation between articles.

Input: Users {R1,R2,…RN}
Articles {A1,A2,…AM}
Authors {U1,U2,…UL}
*Output***:** Undirected Graph with users and articles.
Undirected Graph with authors and articles.
Undirected Graph with articles and articles.
*Algorithm***:**

W_{RA} ←Reading relations between users and articles.
W_{UA} ←Writing relations between authors and articles.
W_{AA} ← Common author relations between articles.
V_R ,V_A← Represents user vertex and article vertex.
E_{RA},E_{AA} ← Represents user article edge and article article edge.
W_{RR} ← Relation between researchers.

```
Construct Graph G
  begin
    W_RR =[0]
     if Ri expressed interest in Aj then
        W_RA(i,j) =1
    else
        W_RA(i,j) =0
if Ui  is the Author of Aj then
```

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

$W_{UA}(i,j) = 1$
else
$W_{UA}(i,j) = 0$
if there is/are common author(s) between articles Ai and Aj then
$W_{AA}(i,j) = 1$
else
$W_{AA}(i,j) = 0$
if $W_{RA}(i,j) = 1$ then
draw $E_{RA}(i,j)$
if $W_{AA}(i,j) = 1$ then
draw $E_{AA}(i,j)$
if $W_{UA}(i,j) = 1$ then
draw $E_{UA}(i,j)$
end

E. Graph Based Article Ranking

The graph-based article ranking method uses the graph to find the most frequently appeared author in the user's library and extracts the other articles written by that author from the database and recommend the relevant articles by comparing other user's library.

Input : Graph G Target Researcher vertex
Output: Relevant articles

Algorithm:
Step 1: Extract articles from user's library
Step 2: Find most frequently appeared author in the user's library
Step 3: Extract the articles written by most frequently appeared author from the database.
Step 4: Find the most relevant articles by comparing extracted articles with other user's library.
Step 5: Recommend relevant articles.

F. Publication Preprocessing.

Articles publication are extracted from user's library and pre-processed to find the keywords in the publication.
Input : Articles Publication
Output : Publication keywords

Algorithm:
Step 1: Extract articles publications from user's library.
Step 2: Extract 2nd half from the publication by splitting the Publication using 'on' or 'of' or 'in'.
Step 3: Remove the special characters and 'journal' keywords from publication.
Step 4: Store the Publication keywords.

G. Conference Venue Recommendation

Doc2vec algorithm is applied for the pre-processed publication keywords to find the frequency of words in the publication vector. Similarly, doc2vec is applied for the conference vector. Then the system applies cosine similarity for publication vector[7] and conference vector[11] to find the most relevant conference. Top 5 conferences are recommended based on cosine similarity. The user can view the details about the conference by clicking the given link to know more about the recommended conference.

Input : Conference dataset and publication keywords
Output: Relevant Conference details

Algorithm:
Step 1: Apply doc2vec algorithm to find the frequency of keywords in the publication
Step 2: Apply doc2vec algorithm to find the frequency of words in the given conference dataset
Step 3: Find the cosine similarity between publication vector and conference vector.

IARJSET
ISSN (Online) 2393-8021
ISSN (Print) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

Step 4: Sort the conferences based on cosine similarity
Step 5: Remove the old conferences by comparing conference date with system date
Step 6: Do step 1 to 5 for all the publications in the user's library
Step 7: Sort the conferences based on cosine similarity and recommend top 5 conference details

IV. PROBLEM STATEMENT

This project aims to develop a collaborative environment for the users to search and save articles of their interest. By using these saved articles, the system performs conference venue recommendation (conference name, location, date) for all researchers and article recommendation for the users having author based search pattern. Some authors may have expertise in the particular research area and have written a lot of articles in that research area. To improve the recommendation the system identifies whether the user has expressed interest in articles written by the particular author.
To determine target users we consider the following features:
1) The ratio of the total number of pair wise articles with common author relations to the total number of all pair wise articles.
2) The ratio of the occurrence number of the most frequently appeared author in articles to the total number of articles for a researcher.
The system will recommend articles written by the most frequently appeared author in the researcher's library to him. For conference venue recommendation, the system applies doc2vec algorithm to find the frequency of keywords in the user's saved article publication and conference dataset. Then the system performs cosine similarity between publication vector and conference vector to find the list of relevant conferences for the recommendation.

V. EVALUATION METRICS

A. Precision

Precision is the fraction of recommended articles that are relevant to the user's interest.
Precision = Relevant articles ∩ Recommended articles / Recommended articles

B. Recall

Recall is the fraction of articles that are relevant to the user's interests that are successfully recommended.
Recall =Relevant articles ∩ Recommended articles / Relevant articles

C. F-Measure

F-measure consider both precision and recall to compute score. It indicates an overall utility of the recommendation list. F-Measure = 2 * Precision * Recall / Precision + Recall

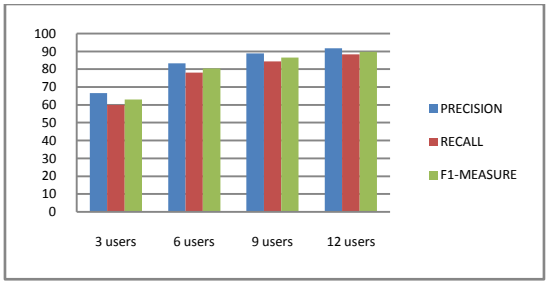

Fig. 2 Precision, recall and f1-measure of article recommendation

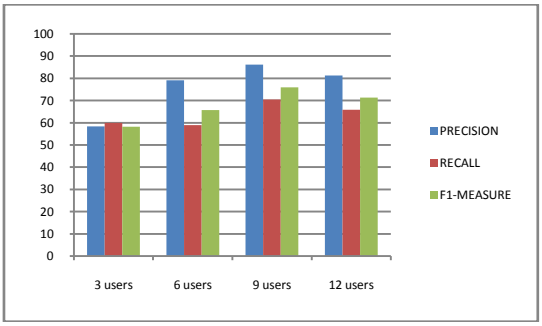

Fig. 3 Precision, recall and f1-measure of conference venue recommendation

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified
Vol. 5, Issue 3, March 2018

VI. CONCLUSION

A collaborative environment is created successfully in the proposed project and this environment helps the users to search and save the articles of interest. The saved articles of interest for analyzing the search pattern of the user and based on this it modifies the recommendation system accordingly. This also helps in recommending conference venue for the user.

As our future work will improvise on the limitation proposed above. It will aim to provide a recommendation system for the user not having author based search pattern by using citation relation between articles.

REFERENCES

[1] Feng X., Liu H., Lee I. and Cao L. (2016), 'Scientific article recommendation: Exploiting common author relations and historical preferences', IEEE Transactions on Big Data, Vol. 2, No. 2, pp. 101-112.

[2] Haifeng L., Yang Z., Lee I., Xu Z., Yu S. and Xia F. (2015), 'Car: Incorporating filtered citation relations for scientific article recommendation', IEEE International Conference on Smart City, pp. 513-518.

[3] Luong H., Huynh T., Gauch S. and Hoang K. (2012), 'Publication venue recommendation using author network's publication history', Asian Conference on Intelligent and Database Systems, pp. 426-435.

[4] Wang C. and Blei D.M. (2011), 'Collaborative topic modelling for recommending scientific articles', Proceedings of the 17th ACMSIGKDD International Conference on Knowledge Discovery and Data mining, pp. 448– 456.

[5] Qi H., Pei J., Kifer D., Mitra P. and Giles L. (2010), 'Context-aware citation recommendation', Proceedings of the 19th International Conference on World Wide Web on ACM, pp. 421-430.

[6] Tang J. and Zhang J. (2009), 'A discriminative approach to topic-based citation recommendation', Proceedings of the 13th Pacific-Asia Conference on Advances in Knowledge Discovery and Data mining, pp. 572–579.

[7] Menon, Varun G.; Joe Prathap P. M. (2013), ' Performance analysis of geographic routing protocols in highly mobile ad hoc network', Journal of Theoretical & Applied Information Technology, Vol. 54 Issue 1, p127-133.

[8] E. Medvet, A. Bartoli, and G. Piccinin, "Publication venue recommendation based on paper abstract," in Proceedings of the 26th IEEE International Conference on Tools with Artificial Intelligence, 2014, pp. 1004–1010.

[9] J. Tang, S. Wu, J. Sun, and H. Su, "Cross-domain collaboration recommendation," in Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2012, pp. 1285–1293.

[10] M. Qu, H. Zhu, J. Liu, G. Liu, and H. Xiong, "A cost-effective recommender system for taxi drivers," in Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 45–54.

[11] Menon, Varun G.; Joe Prathap P. M. (2016),' Analysing the Behaviour and Performance of Opportunistic Routing Protocols in Highly Mobile Wireless Ad Hoc Networks', International Journal of Engineering and Technology, Vol 8 No 5,p1916-1924.

[12] K. Emamy and R. Cameron, "Citeulike: a researcher's social bookmarking service," Ariadne, no. 51, p. 5, 2007.

[13] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," ACM Transactions on Information Systems, vol. 22, no. 1, pp. 5–53, 2004.