

## **STUDY ON RESEARCH PAPER RECOMMENDER SYSTEM**

Pranali Manekar<sup>1</sup>, Anjali Raut<sup>2</sup>

<sup>1</sup>Computer Science And Engineering, SGBAU, pranalimanekar28@gmail.com

<sup>2</sup>Computer Science And Engineering, SGBAU, anjali\_dahake@rediffmail.com

**Abstract—** Research article plays a vital role in developing a knowledge which is helpful in the analysis. There are many reference papers already made by researchers. To search an appropriate research paper is a tedious task. The traditional approach is focused on keyword-based searches to obtain a list of papers. And it requires researchers to review papers for selection as a reference paper. This is a time-consuming approach; this new technique is efficient and has a great demand in the processing of papers. To make it possible citation recommendation is used with mutual reinforcement on three layers. Each paper is considered as a vertex of the paper layer, author layer, and venue layer. To improve citation recommendation clustering is used with three strategies.

**Keywords—**Clustering, Mutually reinforced model, personalized citation recommendation, three-layered interactive clustering, Collaborative filtering, Tokenization.

### **I. INTRODUCTION**

World is changing day by day. Every day some new inventions and experiments are found in technologies trying to find an easy solution for the existing system. Likewise, scientific literature also trending, therefore it is very difficult for researchers to go through all the available literature. The regular approach to search the paper is depending on keyword-based searching. In this, it fetches the list of related papers and requires the analyst to review them and then choose the correct paper as a reference paper. But this required a large time and especially it is challenging for young researchers. Therefore the next technique is found which is effective for processing. Citation recommendation is a new technique that recommends a list of the reference paper which is related to the researcher's information needs and it is the required technique to solve the problem. There are various forms of citation recommendation available that are Elsevier, Pub Med, and Spring Link which recommend the paper which is related to personalized interest of research by email alerts but to recommend, researchers must have to provide the information of their categories of interest others assume that researchers have provided some information which may be citation context, a small part of the reference list. But this assumption is impractical therefore here the main focus on a genuine personalized citation recommendation task. There are three strategies for citation recommendation 1) Graph-based approaches 2) Content-based filtering (CBF) 3) Collaborative filtering (CF). Therefore here for personalized citation recommendation needs to integrate the identity of researchers and query text information that gives the list of the related reference paper.

### **II. LITERATURE REVIEW**

A literature review is an overview of all the knowledge available on a specific topic till date. When it comes to a research topic the first step is in the direction of conducting research is learn more about the previous research published on the topic, and this eventually translates into literature review.

- **Citation Recommendation:-** Searching related papers on site is very hard work for researchers, but citation recommendation can overcome this drawback.

Yang *et al.* [1] expand a ranking oriented CF approach, based on user's access logs to recommend papers. Kang *et al.* [2] provide missing elements of rating matrix using low-rank assumption and then construct recommendations based on the recovered matrix. Chandrasekaran *et al.*[3] generate user profiles based on user authorship of the already published paper. The related paper is recommended by measuring the similarities between the profiles of the papers in the collection and the user's profile. CF is domain independent because it collects user feedback on items and recommends this item based on similarities displayed between the profiles of users. M. Balabanovi'c and Y. Shoham [4] CBF has much quality like the ability to generate recommendations over all items in the domain. Nascimento *et al.* [5] evaluate the content of full paper as input and to construct queries, and then request to the recommend candidate papers that are related to the input paper. Torres *et al.* [6] expand hybrid approaches for merging CF and CBF to recommend research papers. CF used KNN algorithm to output an ordered list of input citation as recommendation and CBF recommended papers by merging similarities between the current paper and the papers in the collection. McNee [7] created a paper recommendation by a hybrid approach by combining CF CBF techniques. Zhou *et al.* [8] measured the similarity between author paper graph, paper citation graph, and venue paper graph. Gori and Pucci [9] developed a random walk based approach to recommend research papers. Meng *et al.* [10] presented a personalized citation recommendation

approach which contains a different kind of information like the content of papers, citation into unified graph model and authorship.

- **Mutual Reinforcement:** - Mutual reinforcement integrate interrelationship between author-to-paper, author-to-venue, venue-to-paper relationship and intrarelationship between author-to-author, venue-to-venue, and paper-to-paper relationship. Into a three-layered graph and develop citation recommendation with the help of a mutually reinforcing manner.

X. Xu, Z. Huang, L. Zuo, and H. He [11] Mutual reinforcement is using for summarization research. Zha [12] developed a principal of mutual reinforcement to extract sentences and key phrases. Wei et al. [13] applied mutual reinforcement principal to the term, document chain and sentences and proposed a ranking approach for query oriented summarization.

- **Multilayered Clustering:**- In multilayered clustering two types of objects are simultaneously clustered so the much attention is drawn in multilayered clustering.

L. Labiod and M. Nadif [14] some methods are applied for a two-layered clustering algorithm to solve the higher order clustering algorithm. Bekkerman et al. [15] applied different types of objects clustered by distributed clustering. Cheng et al. [16] nonnegative matrix factorization one by one clustered terms, documents and authors but they only used interrelationship. Han, R. Quan, D. Zhang, and F. Nie [17] multilayered classification and multilayered clustering are different things. Cai and Li [18] proposed two frameworks i.e. integrated clustering framework and interactive clustering framework to cluster different types of text objects.

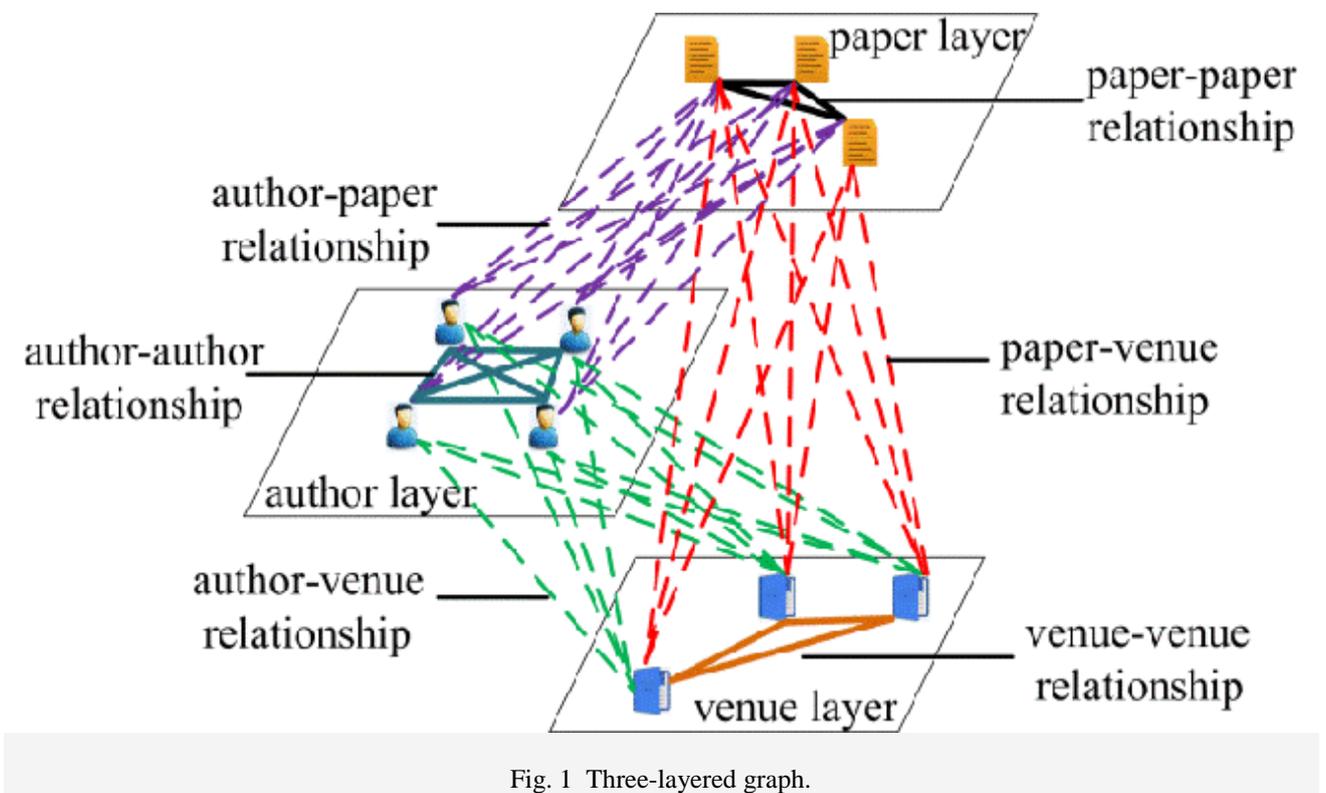


Fig. 1 Three-layered graph.

### III. PRAPOSED WORK

The proposed work includes a citation search engine and recommendation system in which any student/researcher can search for a research paper and get recommended reference papers along with matching papers. To improve the performance of the searching the proposed work includes the k-means clustering algorithm to form documents clusters. It also considers reference paper recommendations by using a user-defined algorithm. Along with papers searching and reference papers recommendation, the paper will be recommended in the model in which collaborative, profile wise and preferences wise citation recommendation techniques will be considered. The proposed work will include the following modules

- A) **Admin panel:-** Admin will view researchers detail and search papers
- B) **User management:-** User will upload a research paper with details and search research paper
- C) **Document searching:-** document searching will includes

- 1) *Specify search query*
  - 2) *Tokenization*
  - 3) *Keywords extraction*
  - 4) *Keywords re-formulation with keywords synonyms*
  - 5) *Search keywords in the database*
  - 6) *Display matching results in clusters using k-means clustering.*
- D) *Document clustering*:- Clustering algorithm will form clusters of search documents.
- E) *Reference Document Recommendation*:- when the user selects any paper for more information system will automatically find out reference papers related to the selected paper and recommend it to the user.

#### **IV. CONCLUSION**

Three-layered mutual reinforcement are useful in citation recommendation. A three-layered mutual reinforcement model is recommended by personalized query oriented reference papers that perform experiments on datasets and proposed approaches in comparison with the other citation recommendation approaches. This is a time-consuming approach; this new technique is efficient and has a great demand in the processing of papers.

#### **REFERENCES**

- [1] C. Yang, B. Wei, J. Wu, Y. Zhang, and L. Zhang, "CARES: A ranking-oriented CADAL recommender system," in Proc. 9th Joint Conf. Dig. Libraries, Austin, TX, USA, 2009, pp. 203–212.
- [2] Z. Kang, C. Peng, and Q. Cheng, "Top-N recommender system via matrix completion," in Proc. 30th AAAI Conf., Phoenix, AZ, USA, 2016, pp. 179–185.
- [3] K. Chandrasekaran, S. Gauch, P. Lakkaraju, and H. Luong, "Concept-based document recommendations for CiteSeer authors," in Proc. Int. Conf. Adap. Hypermedia Adap. Web Syst., Hannover, Germany, 2008, pp. 83–92.
- [4] M. Balabanović and Y. Shoham, "Fab: Content-based, collaborative recommendation," Commun. ACM, vol. 40, no. 3, pp. 66–72, 1997.
- [5] C. Nascimento, A. Laender, A. Silva, and M. G. Alves, "A source independent framework for research paper recommendation," in Proc. 11th Joint Conf. Dig. Libraries, 2011, pp. 297–306.
- [6] R. Torres, S. M. McNee, M. Abel, J. A. Konstan, and J. Riedl, "Enhancing digital libraries with TechLens+," in Proc. 4th Joint Conf. Dig. Libraries, Tucson, AZ, USA, 2004, pp. 228–236.
- [7] S. McNee, "Meeting user information needs in recommender systems," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Minnesota, Minneapolis, MN, USA, 2006.
- [8] D. Zhou et al., "Learning multiple graphs for document recommendations," in Proc. 17th Int. Conf. World Wide Web, Beijing, China, 2008, pp. 141–150.
- [9] M. Gori and A. Pucci, "Research paper recommender systems: A random-walk based approach," in Proc. IEEE/WIC/ACM Int. Conf. Web Intell., Hong Kong, Dec. 2006, pp. 778–781.
- [10] F. Meng, D. Gao, W. Li, X. Sun, and Y. Hou, "A unified graph model for personalized query-oriented reference paper recommendation," in Proc. 22nd ACM Conf. Inf. Knowl. Manag., San Francisco, CA, USA, 2013, pp. 1509–1512.
- [11] X. Xu, Z. Huang, L. Zuo, and H. He, "Manifold-based reinforcement learning via locally linear reconstruction," IEEE Trans. Neural Netw. Learn. Syst., vol. 28, no. 4, pp. 934–947, Apr. 2017.
- [12] H. Zha, "Generic summarization and key phrase extraction using mutual reinforcement principle and sentence clustering," in Proc. 25th SIGIR Conf., Tampere, Finland, 2002, pp. 113–120.
- [13] F. Wei, W. Li, Q. Lu, and Y. He, "Query-sensitive mutual reinforcement chain and its application in query-oriented multi-document summarization," in Proc. 31st SIGIR Conf., Boston, MA, USA, 2009, pp. 283–290.
- [14] L. Labiod and M. Nadif, "A unified framework for data visualization and coclustering," IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 9, pp. 2194–2199, Sep. 2015.
- [15] R. Bekkerman, R. Yaniv, and A. McCallum, "Multi-way distributional clustering via pairwise interactions," in Proc. 22nd Int. Conf. Mach. Learn., Bonn, Germany, 2005, pp. 41–48.
- [16] H. Cheng, P. Tan, J. Sticklen, and W. Punch, "Recommendation via query centered random walk on k-partite graph," in Proc. 7th IEEE Int. Conf. Data Min, Omaha, NE, USA, Oct. 2007, pp. 457–462.

- [17] *J. Han, R. Quan, D. Zhang, and F. Nie*, “Robust object co-segmentation using background prior,” *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1639–1651, Apr. 2018.
- [18] *X. Cai and W. Li*, “Enhancing sentence-level clustering with integrated and interactive frameworks for theme-based summarization,” *J. Assoc. Inf. Sci. Tech.*, vol. 62, no. 10, pp. 2067–2082, 2011.
- [19] *Z. Kang, C. Peng, and Q. Cheng*, “Kernel-driven similarity learning,” *Neurocomputing*, vol. 267, pp. 210–219, Dec. 2017.
- [20] *Z. Kang, C. Peng, and Q. Cheng*, “Twin learning for similarity and clustering: A unified kernel approach,” in *Proc. 31st AAAI Conf.*, San Francisco, CA, USA, 2017, pp. 2080–2086.