Expert Finding in Web Based Collaborative Knowledge Sharing Environment

Prof. Jayshree Ghorpade-Aher [#1], Prajyot Soor[#2]

[#] *Department of Computer Engineering, MIT College of Engineering, Pune, India*

Abstract-- *The biggest challenge for knowledge management is how to make sure that right knowledge could be available to the right people on the right time for any purposes. Traditional search engines return relevant Web documents to keyword queries, where users have to read tedious documents to find the exact information they need. In collaborative environment where people work together with others to achieve a common goal, knowledge sharing plays a very important role. With the help of knowledge sharing one can get faster access to the information. But to find a person who can give the right knowledge is a challenging part due to variety of information needs. This proposed system finds an expert who provides a most relevant knowledge acquired by them to the user in collaborative environment. When the knowledge mining is integrated with expert search, the accuracy improves as compared to traditional methods. This knowledge sharing system helps to get the exact knowledge in less time and to save a lot of repeating efforts for the same task. Thus, it increases the performance of the system by using efficient techniques.*

Keywords — *Expert Finding, collaborative learning, knowledge sharing, document centric expert finding, profile centric expert finding.*

I. INTRODUCTION

Knowledge management enables individuals to create, share and use knowledge in a systematic way. Knowledge sharing is important part of knowledge management, as effective use of knowledge would be more productive. The goal of Knowledge Sharing is to distribute the right knowledge to the right people at right time. Collaborative environment means people work together with others to achieve a common goal. In collaborative environment, people share knowledge, gives suggestions to other people for a specific problem in such a way to get desired knowledge. But to find a person who can give the right knowledge is a challenging part due to variety of information needs. The goal is to find the right person i.e. expert who possesses the desired piece of knowledge based on their relevant uploaded documents.

In collaborative environments, members are assigned a job of acquiring similar information on the web or on computer system in order to gain specific knowledge in one domain. It will be more beneficial to get them connected and share learned knowledge. Traditional search engines like Google, Yahoo return relevant web documents to queries entered by user, where users have to read lengthy documents to find the exact information they need. More recently, intelligent systems allow users to ask questions directly and aim at returning exact answers as per their requirement. But, the intelligence current machines have is still limited in many aspects. Human intelligence should be combined with machine to complement machine-aided knowledge discovery process.

In a classic collaborative environment, upon receiving a task, an expert first tries to solve the task but if he fails, the expert will route the task to another expert. The task will not complete till it reaches an expert who can provide a solution to that task.

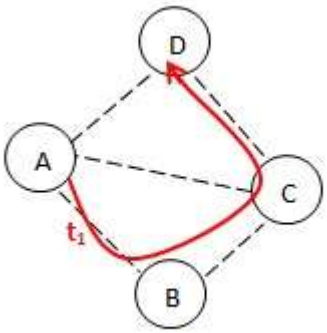

Fig.1. Collaborative network

Figure 1 shows a collaborative network with task solving examples. Task t1 starts at expert A and is resolved by expert D. The sequences A → B → C → D is called solving sequences of task t1.

It is observed that an expert tends to transfer a task or job to other members in collaborative environment whose expertise is neither much similar to nor much different from his own. This phenomenon can be explained as follows: An expert is less likely to transfer a task to another member in collaborative environment whose expertise is very similar, that means the current expert or member is not able to resolve the task. On the other hand, if the expertise of two members is very dissimilar, they might actually specialize in quite different domains; therefore, an member might not be knowing about the other's specialty and few tasks would be transferred between them. We conjecture that it is like the situation when an electronics professor gets an electromagnetics related task, he might consult an

electronics engineering professor who is working on electromagnetics, rather than directly ask a physics scientists who is expert in magnetism, though the latter might be a better candidate.

Information Retrieval (IR) is an activity of retrieving data from databases based on the query received by the system. The process of information retrieval is as shown in Fig. 2.

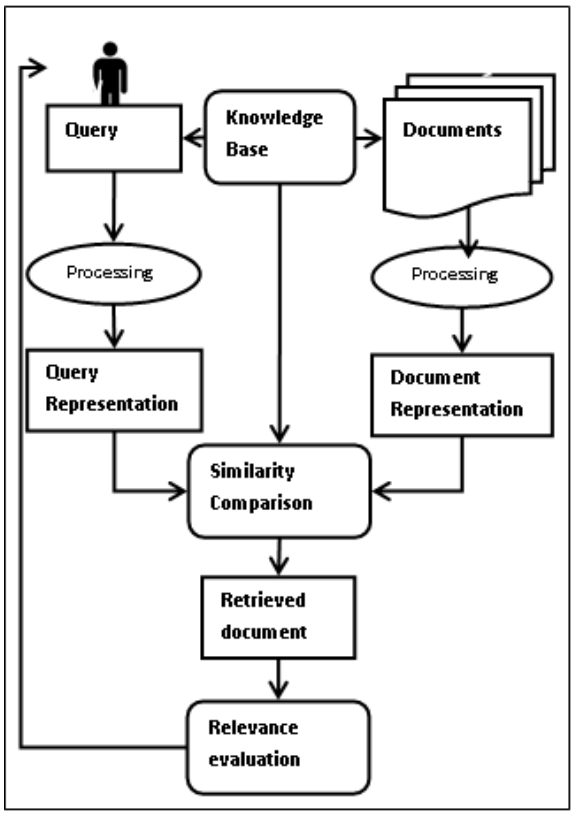

Fig.2. Information Retrieval Process

II. EXPERT FINDING

In information retrieval (IR) community, searching the human expertise has recently attracted much attention. Expert finding is one of the challenging types of search, which concerns itself with ranking peoples who are knowledgeable in a given topic.

The main approaches of expert finding can be divided into two types: profile-centric and document-centric approaches. In profile-centric method, associated documents with a person are used to decide expertise of person. In document-centric method, documents which are related to the query are found out and then experts are derived from them.

Guan, Yang, Sun, Shrivatsa, Yan[1] proposed to find proper advisor or expert who is most likely possessing the desired piece of relevant knowledge based on their web surfing activities. Author proposes to first summarize people's knowledge reflected in their web surfing activities by recognizing the semantic structures, and then searches for the mined pieces of knowledge. Finally, the classic expert search method is applied to the mined results to find proper members for knowledge sharing. Here author evaluated experiments on web surfing data collected from UCSB and IBM lab which shows that the fine-grained aspect mining framework works as expected and outperforms baselines. When information retrieval is integrated with expert search, the search accuracy increases, in comparison with applying the classic expert search method directly on web surfing data.

Proposed framework is divided into two phases (1) clustering of web surfing data into tasks; (2) discriminative infinite Hidden Markov Model is used to mine fine-grained aspects in each task. A framework for mining fine-grained knowledge (micro-aspects): (1) Identify tasks from sessions. (2) Find micro-aspects from sessions in each task. And (3) Apply expert search to find an expert [1].

The problem of finding an expert has been introduced in [2]. In this paper, the goal is to find an expert in bibliographic network where there is a large number of documents and expert candidates. The proposed approach is called topic dominance leaning.

Usually, the authors of a paper have different expertise levels in the domain of the paper. For example, while the supervisor of a Ph.D. project has a broader view of the research problem, other co-authors might be more involved in the detail of the project. Author proposes algorithm assigns no equal expertise scores to authors of a document so that author with more expertise can be given a more importance. These models assign higher scores to authors who are more dominant in the topic of the paper. This model is based on document centric expert ranking approach.

In paper [2], a goal is to predict the topic dominance of each author in a multiauthor document learning algorithms is proposed. The predicted value for the topic dominance of an author can be calculated to check the relative association of each expert candidate on a topic. Two main approaches are used to construct the ranking model using training data such that the model can sort authors according to their importance in a multi author document.

In first approach, algorithm treats the relevant experts in a multiauthor document as positive data, and non-relevant authors as negative data. More precisely, in this approach, each author in a multiauthor document is represented using a feature vector and the algorithm assigns a score for each author of a given multiauthor document. This approach is known as pointwise learning to rank method in the IR literature.

In our second approach, authors follow the idea of pairwise learning to estimate the topic dominance of authors. In this approach, for each pair of authors of a given document, they determine the preference of ranking according to given relevance judgement. Using these preferences, we learn a ranking function and then for a given topic, authors can be ranked according to their topic dominance on a given topic [2].

Gao, Shengxiang et al [3] aimed to determine review expert's rating by using the historical rating records. By constructing rating matrix for project and expert, the relevance of project and domain of expertise of expert is calculated. Author evaluated experiments on real data set which shows that the proposed method could predict the review expert rating more effectively.

In proposed method, algorithm obtains topics of project and experts based on latent Dirichlet allocation (LDA) model and then build the topic relationship network of projects and experts. After that through the topic similarity between projects and experts, algorithm finds a neighbour collection which have the largest similarity with target project and expert. Then integrate the collection into the collaborative filtering recommendation algorithm based on matrix factorization.

Finally, by using the rating matrix algorithm can predict the ratings that a target project will give candidate review experts, and achieves the review expert recommendation [3].

Lan Zhang, Xiang-Yang Li, Jingsheng Lei, Jiaguang Sun, and Yunhao Liu[4] addresses the problem of expert finding in a distributed expert scenarios using chains of social referrals and profile matching with only local information in social media. In this work authors proposes a mechanism in which expert finding query will be put forwarded by intermediate users in the network. Authors address the local social referral problem in a large scale social network by taking users' self-interests into consideration and designed an efficient mechanism that reduces the length of social referral chain, reduces the cost of social referrals, and improves the success rate, compared with previous efforts.

Author performed some extensive experiments on several data sets of online social networks. The evaluated result shows that the success rate of proposed mechanism is about 90 percent in finding closely matched experts within budget or cost using local search.

Various methods are used to improve the performance of distributed search i.e. profile matching, score function for selection of relay neighbours, and budget estimation [4].

Wang, Rong, Jie Deng, and Fei Men [5] proposed architecture of a social collaborative learning environment system. This system is divided into three layers i.e. (1) user layer, (2) system layer and (3) storage layer. In the user layer, individuals can construct their cognitions. The system layer used to aggregate individual cognitions into group cognition and discover relations among the knowledge, users and the resources. In this way, one user or members collaborates with other members in the system through the environment. This method analyses networks made of users, knowledge, resources and relations between them.

Author performed some experiments to see how the aggregation model and algorithm work and concluded that more experiments need to be done to decide suitable parameters in the model.

There are four main parameters i.e. learners, knowledge, resources and relations between two of the pervious elements. Here author analyses the network with its elements, explains the generation of cognition maps and grouping of four elements [5].

III. CONCLUSIONS

Knowledge sharing is a solution for improving knowledge access and enhancing the efficiency of knowledge sharing environment. By finding expert who can provide the required knowledge, lots of time and repeating efforts of the members in collaborative environment can be saved. Profile centric and document centric approaches can be used to find expert among the crowd. But among all, expert finding by document centric approach yield better results. The document centric approach for expert finding is more effective way of finding expert in a collaborative learning environment as compared to profile centric approach. And current research is also going on document centric expert finding.

REFERENCES

[1] Guan, Z., Yang, S., Sun, H., Srivatsa, M., Yan, X. (2014). "Fine-Grained Knowledge Sharing in Collaborative Environments."Knowledge and Data Engineering, IEEE Transactions on Year: 2015.

[2] Neshati, Mahmood, Seyyed Hadi Hashemi, and Hamid Beigy. "Expertise Finding in Bibliographic Network: Topic Dominance Learning Approach."Cybernetics, IEEE Transactions on 44.12 (2014): 2646-2657.

[3] Gao, Shengxiang, Zhengtao Yu, Linbin Shi, Xin Yan, and Haixia Song. "Review expert collaborative recommendation algorithm based on topic relationship." Automatica Sinica, IEEE/CAA Journal of 2, no. 4 (2015): 403-411.

[4] Lan Zhang, Xiang-Yang Li, Jingsheng Lei, Jiaguang Sun, and Yunhao Liu "Mechanism Design for Finding Experts Using Locally Constructed Social Referral Web", IEEE Transaction S On Parallel And Distribut Ed Systems, Vol. 26, No. 8, August 2015.

[5] Wang, Rong, Jie Deng, and Fei Men. "Social Group Collaborative Learning and Knowledge Sharing Environment." Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on. IEEE, 2009.

[6] Balog, Krisztian, Leif Azzopardi, and Maarten De Rijke. "Formal models for expert finding in enterprise corpora." Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2006.

[7] Abdullah, R., Jaafar, A., Ibrahim, H., Selamat, M. H., Napis, S., & Atan, R. (2008, August). Knowledge Management system architecture in agricultural industries environment. In Information Technology, 2008. ITSim 2008. International Symposium on (Vol. 2, pp. 1-9). IEEE.

[8] Zhang, Chenghong, Cheng Zhang, and Zheng Yan. "Knowledge Sharing in Collaborative Commerce: Its Characteristics, Research Framework and Process." PACIS 2004 Proceedings (2004): 108.

[9] Deng, Hongbo, Irwin King, and Michael R. Lyu. "Formal models for expert finding on dblp bibliography data." Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on. IEEE, 2008.

[10] Mookerjee, Vijay S. "Debiasing training data for inductive expert system construction." Knowledge and Data Engineering, IEEE Transactions on 13.3 (2001): 497-512.

[11] Huang, Changqin, et al. "Enabling Novel Knowledge Sharing with Activity Ontology Model in Collaborative Learning Environments." Information Technologies and Applications in Education, 2007. ISITAE'07. First IEEE International Symposium on. IEEE, 2007.

[12] Fischer, Gerhard, Kumiyo Nakakoji, and Yunwen Ye. "Metadesign: Guidelines for supporting domain experts in software development."Software, IEEE 26.5 (2009): 37-44.

[13] Nor, Mohd Zali Mohd, et al. "Knowledge sharing interactions in collaborative software maintenance environment." Computer Technology and Development, 2009. ICCTD'09. International Conference on. Vol. 2. IEEE, 2009.

[14] Maybury, Mark, Ray D'Amore, and David House. "Expert finding for collaborative virtual environments." Communications of the ACM 44.12 (2001): 55-56.