ISSN: 2321-7782 (Online)
Volume 3, Issue 10, October 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Review on Knowledge Sharing In Collaborative Environment

Nilofar N. Pathan[1]
Computer Science and Engineering
H.V.P.M's C.O.E.T
Amravati, India

Ranjit R. Keole[2]
Professor, Information Technology
H.V.P.M's C.O.E.T
Amravati, India

Abstract: ***In working area where it is common that members try to acquire similar information on the web in order to gain specific knowledge in one domain sharing environment is needed. Like in collaborative environments, members may try to acquire similar information on the web in order to gain knowledge in one domain. For example, in a company several departments may successively need to buy same software and employees from these departments may have studied online about different tools and their features independently. It will be productive to get them connected and share learned knowledge. In this dissertation work investigation is done on fine-grained knowledge sharing in collaborative environments. In this work a methodology is proposed to analyze member's web surfing data to summarize the fine-grained knowledge acquired by them. A two-step framework is proposed for mining fine-grained knowledge: (1) web surfing data is clustered into tasks by a nonparametric generative model; (2) a new infinite Hidden Markov Model is developed to mine fine-grained aspects in each task. Finally, the classic expert search method is applied to the mined results to find proper advisor for knowledge sharing.***

Keywords: Fine-grained knowledge sharing, Advisor Search, Infinite Hidden Markov Model, Collaborative environment, nonparametric generative model.

I. INTRODUCTION

In a collaborative environment, it could be common that members try to acquire similar information on the web in order to gain specific knowledge in one domain. This dissertation presents a new method to identify, how to enable such knowledge sharing mechanism by analyzing user data. For example, Alice starts to surf the web and wants to learn how to develop a Java multithreading program, which has already been studied by Bob. In this case, it might be a good idea to consult Bob, rather than studying by herself. Such recommendations are provided with this methodology by analyzing surfing activities automatically. In this example, not necessarily Bob is an expert in every aspect of Java programming; however, due to his significant surfing activities in Java multithreading, it is reasonable to assume that he has gained enough knowledge in this area so that he can help Alice.

This method departs from the traditional expert search problem in that expert search aims to find domain experts based on their associated documents in an enterprise repository, while the goal of this proposed work is to find proper "advisors" who are most likely possessing the desired piece of fine-grained knowledge based on their web surfing activities. In order to analyze the knowledge acquired by web users, new method is proposed to log and analyze user's web surfing data. User's interactions with the web can be segmented into different "tasks", e.g., "learning Java" and "shopping". Textual contents of a task are usually cohesive. This dissertation defines a session as an aggregation of consecutively browsed web contents of a user that belong to the same task. Sessions are atomic units in our analysis. A task can be further decomposed into fine-grained aspects (called micro-aspects). A micro-aspect could be roughly defined as a significantly more cohesive subset of sessions in a task. For example, the task "learning Java" might contain "Java IO" and "Java multithreading". To this end, a novel infinite Hidden

Markov Model (iHMM) is proposed to mine micro-aspects in each task. Finally, a language model based expert search method is applied over the mined micro aspects for advisor search.

II. Literature Review and Related Work

In year 2003 D. M. Blei, A. Y. Ng, and M. I. Jordan [2] had analyzed topic modeling. Topic modeling is a popular tool for analyzing topics in a document collection. The most prevalent topic modeling method is Latent Dirichlet Allocation (LDA). It is a generative probabilistic model for collections of discrete data. Topic modeling decomposes a document into topics. But it doesn't recover the semantic structures of people's online learning activities from their web surfing data, i.e. identifying groups of sessions representing tasks (e.g. learning "Java") and micro-aspects (e.g. learning "Java multithreading"). After applying topic modeling methods on session data, it is still difficult to find the right advisor by using the mined topics.

In year 2005 X. Liu, W. B. Croft, and M. Koll [3] has also been studied expert retrieval in other scenarios, e.g. online question answering communities. People using such services are like a community – anyone can ask, anyone can answer, and everyone can share, since all of the questions and answers are public and searchable immediately. But there are hundreds of questions asked each day but some portion of them may not be answered or there may be a lag between the time when a question is asked and when it is answered. Also the answers may not be satisfactory.

In year 2006 K. Balog, L. Azzopardi, and M. de Rijke [4] proposed a language model framework for expert search. Expert search aims at retrieving people who have expertise on the given query topic. Their Model 2 is a document-centric approach which first computes the relevance of documents to a query and then accumulates for each candidate the relevance scores of the documents that are associated with the candidate. It locates documents on topic, and then finds the associated expert. Balog showed that Model 2 performed better. But the nature of these methods is still accumulating relevance scores of associated documents to candidates. Traditional expert search does not explicitly retrieving people who are most likely possessing the desired piece of fine-grained knowledge it focused on finding experts only rather than to mine fine-grained aspects for each task.

In year 2008 R. Jones and K. Klinkner [5] found that search tasks are interleaved and used classifiers to segment the sequence of user queries into tasks. They studied real sessions manually labeled into hierarchical task. They proposed and evaluated a method for the automated segmentation of users' query streams into hierarchical units. But it considers search engine query logs only, rather than general web surfing contents (including search). Query logs do not record the subsequent surfing activity after the user clicked a relevant search result. Also it dint try to address advisor search by exploiting the data generated from users' past online behaviors.

In year 2011 A. Kotov, P. Bennett, R. White, S. Dumais, and J. Teevan [6] designed classifiers to identify same-task queries for a given query and to predict whether a user will resume a task. They introduced and addressed the two problems in the context of analysis of cross-session search tasks: (i) identifying queries from earlier sessions on the same task, and (ii) predicting whether a user will return to the same task during a later session. But it doesn't provided richer prediction models and alternative feature sets, exploring new prediction and classification problems in the context of cross session information needs. It also didn't tried to mine fine-grained aspects for each task. Summarizing fine-grained aspects can provide a fine-grained description of the knowledge gained by a person.

In year 2015 Ziyu Guan, Shengqi Yang, Huan Sun, Mudhakar Srivatsa, and Xifeng Yan [1] suggested a fine-grained knowledge sharing in collaborative environments. They proposed a method to find proper "advisors" who are most likely possessing the desired piece of fine-grained knowledge based on their web surfing activities. But the fine-grained knowledge could have a hierarchical structure. And how to search over this hierarchy is not a trivial problem. Also this work creates an issue of privacy.

III. AIM AND OBJECTIVES

» To retrieve people who are most likely possessing the desired piece of fine-grained knowledge.

» To address advisor search by exploiting the data generated from user's past online behaviors.

» To summarize fine-grained aspects which can provide a fine grained description of the knowledge gained by a person.

» To provide a healthy sharing environment.

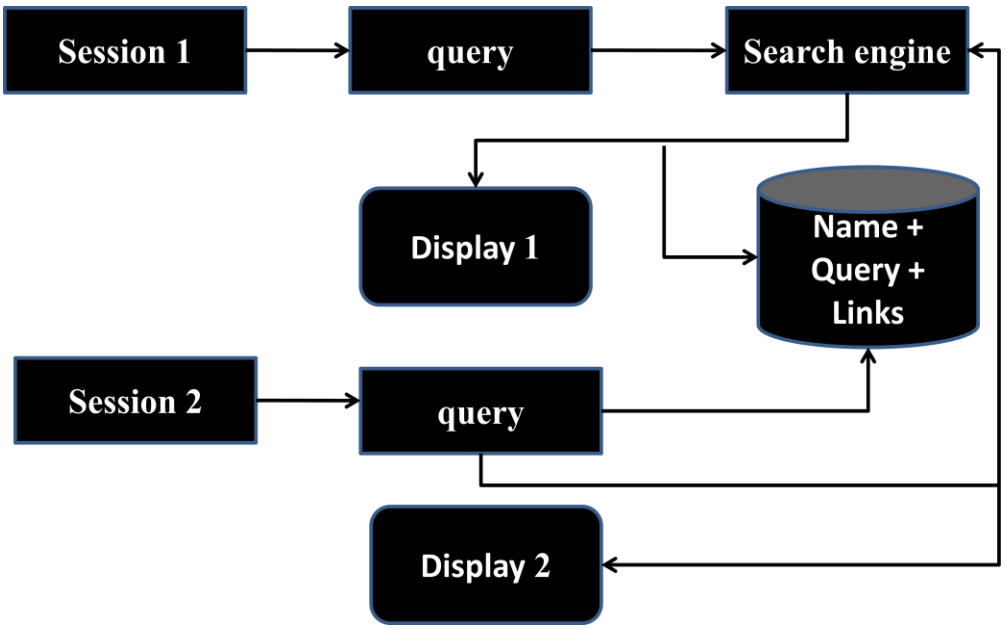
» To analyse user online behaviours to mine the tasks and provide advisor search.

» To provide ease of access of desired information and save time of repetitive efforts.

IV. PROPOSED METHODOLOGY

The goal of this method is not finding domain experts but a person who has the desired piece of knowledge. The proposed methodology provides technique to find proper "advisors" who are most likely possessing the desired piece of fine-grained knowledge based on their web surfing activities. This work proposes the fine-grained knowledge sharing in collaborative environments. This method is proposed to solve the problems by first summarizing web surfing data into fine grained aspects, and then search over these aspects.First the user entered web surfing data including queries and name is analyzed and extracted. This web surfing data is clustered into tasks by a nonparametric generative model. These tasks can be further decomposed into fine-grained aspects (called micro-aspects). Then infinite Hidden Markov Model is developed to mine fine-grained aspects in each task and to employ comparison among same searches. Finally, a language model based expert search method is applied over the mined micro aspects for advisor search.

Representing the flow of work diagrammatically :



1 User will start the session 1 with his /her name and enters the query 2 search engine.

2 Search engine provides the display of information and as a background process database is created with name, query entered and related searched links.

3 When new user comes with new session 2 the query is first compared with queries in our database.

4 If match is found then the stored data is recommended and if not then new display is provided with required topic.

V. Conclusion

This method provides an easy way to retrieve people who are most likely possessing the desired piece of fine-grained knowledge by addressing advisor search by exploiting the data generated from user's past online behaviours. It provides ease of access of desired information and save time of repetitive efforts. Also identified digging out fine-grained knowledge reflected by people's interactions with the outside world as the key to solving the problems. This method proposed a two-step framework to mine fine-grained knowledge and integrated it with the classic expert search method for finding right advisors.

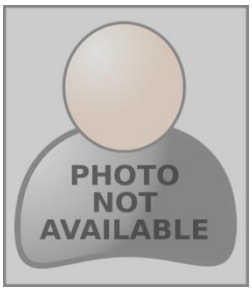
Acknowledgment

My thanks to the Guide, Prof. R.R.Keole and Principal Dr.A.B.Marathe, who provided me constructive and positive feedback during the preparation of this paper.

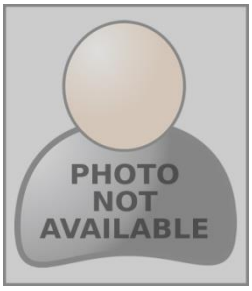
References

1. Ziyu Guan; Shengqi Yang; Huan Sun; Srivatsa, M.; Xifeng Yan, "Fine-Grained Knowledge Sharing in Collaborative Environments," in Knowledge and Data Engineering, IEEE Transactions on , vol.27, no.8, pp.2163-2174, Aug. 1 2015
2. D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," J. Mach. Learn. Res., vol. 3, pp. 993–1022, 2003.
3. X. Liu, W. B. Croft, and M. Koll, "Finding experts in communitybased question-answering services," in Proc. 14th ACM Int. Conf. Inf. Knowl. Manage., 2005, pp. 315–316.
4. K. Balog, L. Azzopardi, and M. de Rijke, "Formal models for expert finding in enterprise corpora," in Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2006, pp. 43 50.
5. R. Jones and K. Klinkner, "Beyond the session timeout: Automatic hierarchical segmentation of search topics in query logs," in Proc. 17th ACM Conf. Inf. Knowl. Manage., 2008, pp. 699–708.
6. A. Kotov, P. Bennett, R. White, S. Dumais, and J. Teevan, "Modeling and analysis of cross-session search tasks," in Proc. 34th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2011, pp. 5–14.

Author(s) Profile



Nilofar N. Pathan, received the B.E.degree in Information Technology from H.V.P.M's College Of Engineering And Technology, Amravati in 2014. She is currently persuing Master's Degree in Computer Science and Engineering from H.V.P.M's College of Engineering And Technology, Amravati.



Prof. Ranjit R.Keole, received the B.E.and M.E degree in Computer Science from Prof. Ram Meghe Institute of Technology, Badnera in 1992 and 2008, respectively. His field of specialisation is web Mining. He is currently working as Associate Professor at H.V.P.M's college of Engineering and Technology,Amravati.

