

# ADVISOR RECOMMENDATION SYSTEM FOR KNOWLEDGE SHARING

Nihar Khandekar  
Department of Information Technology  
Atharva College of Engineering,  
Malad (West), Mumbai, India.

Akash Gupta  
Department of Information Technology  
Atharva College of Engineering,  
Malad (West), Mumbai, India

Abhishek Mishra  
Department of Information Technology  
Atharva College of Engineering,  
Malad (West), Mumbai, India.

Chanda Chouhan  
Assistant Professor, Department of Information Technology  
Atharva College of Engineering,  
Malad (West), Mumbai, India.

**Abstract:** In this paper, the problem of knowledge hoarding has been overcome by sharing knowledge in collaborative environments. The proposed system enables users searching on the web to get connected and share learned knowledge. The web surfing activities of users are recorded and they are mined for extracting keywords using RAKE algorithm. The keywords associated with a particular user are fed into the database. The user enters a topic to search and gets suggested an advisor based on user-keyword database. The user can then directly contact advisor for acquiring knowledge.

**Keywords - Collaborative environments, RAKE algorithm, Advisor**

## I. INTRODUCTION

[1] Collaborative environment means people work together with others to achieve a common goal. Interaction with people and/or colleagues to acquire information is the most common way adopted by the members of collaborative environment. Other ways for people to gather knowledge or information and solve their problems are:

- 1) By using machines, such as google, yahoo, bing kind of search engines.
- 2) Asking for help directly from subject/domain experts.

We're now into an era of hyper-specialized corporations where every single job requires a strong level of expertise. Knowledge dissemination cannot be controlled by a certain elite anymore. Companies should therefore leverage their employees' knowledge by helping them to **store, share, categorize and disseminate** it to the largest possible number. The aim is that companies can maintain a "common minimum knowledge level" so everyone can carry on with the same amount of information. Also it will help the employees to understand a particular topic from the knowledgeable person of that field rather than wasting precious time searching on the internet.

## II. RELATED WORK

### 2.1 Expert Search

[7] Advisor search is different from traditional expert search. Advisor search is dedicated to retrieving people who are most likely possessing the desired piece of fine-grained knowledge. It is important to recognize the semantic structures and summarize the session data into micro-aspects so that we can find the desired advisor accurately. Nonparametric generative models have been created to mine micro-aspects to show the superiority of search scheme over the simple idea of applying traditional expert search methods on session data directly.

### 2.2 Keyword Extraction Methods

Research has been done on various keyword extraction methods including TF-IDF, Text Rank, RAKE etc. [6] Early approaches to automatically extract keywords focus on evaluating corpus-oriented statistics of individual words. Jones (1972) and Salton et al. (1975) describe positive results of selecting for an index vocabulary the statistically discriminating words across a corpus. Later keyword extraction research applies these metrics to select discriminating words as keywords for individual documents. For example, Andrade and Valencia (1998) base their approach on comparison of word frequency distributions within a text against distributions from a reference corpus.

### III. PROBLEM STATEMENT

The problem of knowledge hoarding exists in many organizations and they fall flat in knowledge management. To overcome this problem, knowledge is shared in collaborative environments using various models.

### IV. AIMS AND OBJECTIVES

#### 4.1 Aim

The project aims to recommend an 'advisor' based on his web-surfing activities.

#### 4.2 Objectives

The objective is to save repeating efforts of acquiring same knowledge which is already acquired by other members of collaborative environment.

Another objective is to achieve time efficient quick response for the query.

### V. EXISTING SYSTEM

There exists a system which finds most relevant information by analyzing knowledge (documents) which is already acquired by other members of collaborative environment. In this system, documents uploaded by members of collaborative environment are analyzed and indexing of these uploaded documents is done. After that, documents are retrieved based on search keyword. Clustering of documents is done with cluster label to each cluster for identifying documents where a search term is used in the intended sense and expert search method is applied to get most relevant document. For this system, input is various text documents uploaded by members of collaborative environment.

### VI. PROPOSED SYSTEM

We propose a system in which advisor is recommended by analyzing knowledge gained by other members of collaborative environment. The project is carried out in two phases.

In the first phase, users are asked to search for their topics on the internet. This web-surfing data is recorded using *SoftActivityMonitor* software installed on users' computers. The data is then treated through RAKE algorithm for extracting keywords and key phrases from it. The keywords extracted are then stored in the database corresponding to the user. Also the time duration of searching a topic recorded by *SoftActivityMonitor* software is stored in the database.

In the second phase, users are asked to search for a particular topic or keyword and he/she will be suggested a list of advisors ranked from high to low based on the time spent by advisors on a particular topic.

#### 6.1 Rake Algorithm

RAKE (Rapid Automatic Keyword Extraction) is based on observation that keywords frequently contain multiple words but rarely contain standard punctuation or stop words, such as the function words *and*, *the*, and *of*, or other words with minimal lexical meaning.

The input parameters for RAKE comprise a list of stop words (or stop list), a set of phrase delimiters, and a set of word delimiters. RAKE uses stop words and phrase delimiters to partition the document text into candidate keywords, which are sequences of content words as they occur in the text. Co-occurrences of words within these candidate keywords are meaningful and allow us to identify word co-occurrence without the application of an arbitrarily sized sliding window. Word associations are thus measured in a manner that automatically adapts to the style and content of the text, enabling adaptive and fine-grained measurement of word co-occurrences that will be used to score candidate keywords.

#### 6.2 Algorithm for proposed system

- Step 1: Start
- Step 2: Members search for topics on the internet
- Step 3: Web surfing data is recorded
- Step 4: RAKE algorithm is used for keyword extraction
- Step 5: Keywords along with time duration are stored in the database
- Step 6: Users enter a topic or keyword in search field
- Step 7: Displays a list of advisors
- Step 8: Stop

### 6.3 Proposed System Architecture

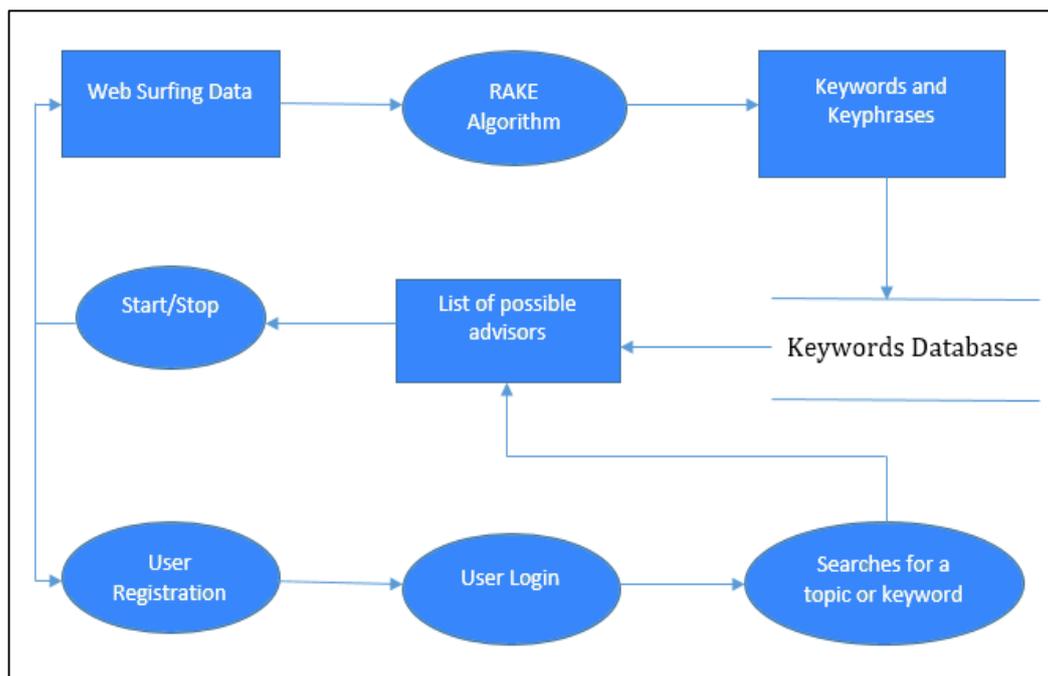


Figure -1: Architecture of proposed system

### VII. CONCLUSION AND FUTURE WORK

The proposed system is going to save repeating efforts of the members in collaborative environment by finding the expert, who can provide the relevant required knowledge. We identified grained knowledge reflected by people's interactions with the outside world. The two-step framework used for system is very efficient in terms of finding information.

In future scope, the system can be improved by applying it to whole organization with geographically distant offices and instead of considering only text documents, audio and video can also be included.

### VIII. REFERENCES

- [1]Kolte, C. and Shrivasa, A. (2017). Intelligent knowledge sharing for agricultural information. *2017 International Conference on Trends in Electronics and Informatics (ICEI)*.
- [2]Soor, P. and Ghorpade-Aher, J. (2017). Expert search based knowledge sharing system for collaborative environment. *2017 International Conference on Emerging Trends & Innovation in ICT (ICEI)*.
- [3]M. J. Beal, Z. Ghahramani, and C. E. Rasmussen. The infinite hidden markov model. In *Advances in neural information processing systems*, pages 577–584, 2001.
- [4] D. M. Blei, T. L. Griffiths, M. I. Jordan, and J. B. Tenenbaum. Hierarchical topic models and the nested chinese restaurant process. In *NIPS*, 2003.
- [5] 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.
- [6]Rose, Stuart & Engel, Dave & Cramer, Nick & Cowley, Wendy. (2010). Automatic Keyword Extraction from Individual Documents. 10.1002/9780470689646.ch1.
- [7]Z. Guan, S. Yang, H. Sun, M. Srivasta and X. Yan, "Fine-Grained Knowledge Sharing in Collaborative Environments - IEEE Journals & Magazine", *Ieeexplore.ieee.org*, 2019. [Online].Available:<https://ieeexplore.ieee.org/document/70/56511>.