Sherpa - Expert's Social Network

Sangramsingh M Ghatage[1], Siddharth K Gaikwad[2]

[1]Department of Computer Engineering and Technology, College of Engineering, Pune, India

Abstract: *Nowadays, lots of people with different goals are registered in social networks and do wide range of activities. Social networks provide a structure that allows users to interact and share resources (e.g. Documents, Pictures and Videos) on the basis of a trustworthy relationship (e.g. Friendship). One of the most important feature of social networks is knowledge sharing. Social networks and online communities have become important places for people to seek and share expertise. One major problem on social networks and online communities is a wide range of shared knowledge and experience is available but there is no mechanism to determine their validity. This Problem can be viewed as finding the influential experts who posts topic-specific high-quality content. By finding experts in social networks and determining their level of knowledge, the validity of their posts could be determined. Finding expertise in a research area also helps researchers to know who are the experts working on particular research area and seek help from them. Researchers can use this platform to work with Experts of concerned area for their Research Topic. Considering the problem of Expert Finding and available technology: Social Networking; a novel solution of Expert's Social Network is proposed. This will be a Social Network System where Experts Can Share their knowledge and other members of network can use it. This System will help to A User to get topic-specific high-quality content, Trending Topics, Share data and work with peer members of Network. Our approach demonstrates that for finding an Expert and to get topic-specific high-quality content, Social Networking is one of the ways.*

Keywords: Expert Search, Expert's social network, Expert Advice, research help, research guidance

1. Introduction

Human communication and interaction way has experienced changed a lot during last few years. The way we communicate and interact with each other in day-to-day life is changed by email, chats, discussion forums, blogs, etc. For common man it was never so easy to reach and share the voice to a large and growing audience. Social networks popularity is increasing constantly. New ways of using technology of social networking with some twists are being discovered. For instance, many science communities are now using social networks. Biotechnology companies share their scientific knowledge with each other using networking sites. Collaboration has always helped to grow and flourish science and various forums are best places to promote sharing of information and knowledge. A business focused virtual community can be considered as Professional Network rather than social network which emphasis on the relationships and interactions between businesses. So basics features of Social networking are same, but how they are tailored for various purposes differ. Along with basic features special features are added. These extra features help to establish a social network of special interest. E.g. Linked In for professionals, Flicker for Photo Sharing, etc.

Over last several years Search engines have made a great progress. However, in many cases, people would prefer individual help and supporting. They would like to search for individuals instead of documents. e.g. a researcher would be interested in consulting a expert from his/her research area rather than referring large amount of documents. It is important issue to get support from Experts in our daily life. Expert finding is difficult task. Finding expertise in a research area helps researchers to know whom are the experts working on the research area. This is very useful to get the trend of your research area or related research papers from that expert can be analyzed. Academic communities are the most interested and active in research of finding expertise for a topic.

Online sharing of knowledge and learning systems are supported by systems like Ask to Expert and Online communities. We aim to use such systems with Social network and adding some focused features related to Expert's advice and researcher coworking. This will create Support of social interactions in educational spaces. We named this Expert's Social Network as 'Sherpa'. Sherpa are the people who live in Himalayas and work there as Guides for Trekkers. In Tough times and track of Himalayan Mountains they show the path with all their dedication and efforts to reach Trekkers to ultimate destination. We visualize this Expert's Social Network in same way; that will help people, researchers to connect and get guidance from Experts to reach to their ultimate goal.

2. Social Networking & Features

The top social networking sites have become the most frequently visited pages on the internet. It's important that you find the ones that best represent your personality and allow you to share your thoughts and feelings with the widest selection of friends, even if you're sharing nothing more than an adorable cat video. SNSs as web-based services allowing individuals to do three basic tasks i.e. construct a profile, articulate a lists of users with whom they share connection and view and traverse the lists. They do talk about other elements that SNSs may have: avatars, privacy settings, customization of relation names, posting resources to users wall, private messages and any type of content sharing (photos, videos, etc.) However, they do not provide a formal measure on how these features are adopted by current social network platforms. The presence of information about relations is a main characteristic of a social network. The concepts behind a social network are actors and their relations. These are the basic elements that need to be present to start talking about a social network.

Paper ID: 02014512

1384

Features:

2.1 Social actors (Users)

Social entities are referred to as actors. They can be people of course, but also collective social units such as corporations, organizations, institutes, cities or even events. This feature should be implemented by every site that claims to be a SNS.

2.2 User accounts

Users must be allowed to register in the site. They should fill a form, provide some basic data, typically name, email and password, though this could vary from one site to the other. Later, in the log-in process, users provide credentials in order to be authenticated. Nowadays, leading websites are providing authentication services to other websites. SNS may delegate authentication to them by using these services.

2.3 Social relations

In SNS, actors must be able to establish some kind of relations with other actors. Relations can be unidirectional o bi-directional. Unidirectional relations are established by one of the parties and do not need confirmation by the other party. Follow is a popular case of unidirectional relation. The paradigm of this relation is Twitter. Bidirectional relations mean that both parts need to approve the relation to be established. Facebook is the paradigm of this relation type.

2.4 Content

Content-oriented SNS are built around some specific type of content. These features take the influence from the work in web content management systems (CMS). Content types managed by SNS may be very different. Most popular content types include text, pictures and video. Examples of SNS built around content are Flickr for photographs, YouTube for videos, Github for code repositories. Specific type of content help to focus the website on some specific type of user,

2.5 Communication tools

SNS sites may also support other communication tools. One example is system notifications, by which the system alerts users of some types of events, for example the posting of new content that may be of their interest. Other examples include private messages between users, chat and video-conferences. These communication tools are often connected to email.

2.6 Privacy and content visibility

Access control mechanisms allow content to be shared only with a given audience. There are different levels of information disclosure. Existence of different levels of content visibility may facilitate users to share more content. There are some sensitive pieces of information that would not be shared if the only option is publishing them to the public, but they might be shared with a restricted audience.

2.7 Ratings

In many social network sites, people are allowed to rate content and activities from other users. Rating systems take several forms, from like, binary like and dislike, 5 starts, etc. This functionality is the base for popularity rankings. They can be used for building recommendations to users and promoting featured content.

2.8 Activities timeline

An activities timeline consists on a more-recent ordered list of actions performed by one or several actors. It may refer to posting some content, creating new friendships, rating content, etc. The activities timeline help users to be aware of the activity of others, and have a sense of what is happening in the social network.

2.9 Wall

We define a wall as the posting box that allows actors adding or creating any type of content to his own profile or to other actor in the SNS. The box is usually besides the activity timeline, so after the user posts a new content, a new activity appears in the timeline, a way of providing users with instant gratification. The box usually supports posting several types of contents, such as text and files. It can also appear in several sections of the SNS.

2.10 Home

As users are authenticated, SNS may have a special page called home or dashboard. This is the first page offered to users after signing into the site. This page is used to present users a summary of the most important information in the user's social network.

2.11 Profile

The heart and soul of social networking sites are people's personal profiles. Each one is its own online sanctuary – a place where people can express their thoughts and feelings, post their photographs and interact with their network of friends. The most popular social network websites put a strong emphasis on the user's profile, making it easy to use yet still reflective of each user's personality. They represent the external image that social actors project, they are the representation of users to the rest of the social network.

2.12 Search

In SNS search functionality is a must – simply because of the vast amount of available information. The search, however, has multiple dimensions: apart from the traditional content search, social applications also provide an advanced search of connections in the social graph – be it groups, communities or interests. A common feature in social applications is the use of **live search results and filtering**. When you type into the box, the results are filtered out in a drop-down style.

Paper ID: 02014512

1385

3. Current Expert Search Approaches and New Approach

3.1 Web Mining Approach Using Clustering

It consists of the following components: Indexing Agents, Author Clustering, Document Clustering, Multi-Clustering and Retrieval Engine [3].

a) **Indexing Agents:** Indexing Agents search for research web sites, especially those from academic institutions, download scientific publications from the web sites and convert the PDF or PostScript files into text data. They parse the bibliographic section to extract the citation information, which is then stored in the Web Citation Database.

b) **Document clustering:** Document clustering mines the Web Citation Database. Paper titles of all the citation records in the Web Citation Database are selected as feature factors to represent the document vectors. Then, the titles of citations are preprocessed to extract keywords. The extracted keywords are then converted into document vectors. Next, KSOM and Fuzzy ART are used to generate document clusters from document vectors.

c) **Author Clustering:** Author clustering mines the Web Citation Database based on author co-citation analysis. Author co-citation analysis is based on the premise that two authors are correlated if the frequency that they are cited together by later works is high.

d) **Multi-Clustering Technique:** The Multi-Clustering technique consists of the following steps.

1) Vectorization - It represents the multiple clustering data from document clusters and author clusters as vectors. The multiple clustering data is obtained from the Document Clusters database and Author Clusters Database.

2) Distance Evaluation - It calculates the distances between the generated vectors.

3) Linkage Generation - It performs the AHC technique to cluster the vectors and generates linkage information between the vectors within each cluster. The vector clusters and linkage information are stored in the Multi-Clusters Database.

e) **Retrieval Engine:** Retrieval Engine supports the finding of expertise by sending database queries on records stored in Multi-Cluster Database.

3.2 Link Analysis Approach

In order to identify the importance of people in a social network, link analysis approach is used. In fact, it is modified version of PageRank algorithm to make it possible to be used in social networks for expert finding. In PageRank algorithm, different pages are navigated and the probability of each page selection is calculated using the Marcov chain. Finally, importance of each page is calculated. In this algorithm, instead of web pages, there are people in social networks and the connections between them are used as hyper links.

In the Fig. the connections between people are illustrated by weighted links. The weight of each link is the number of posts which two people sent to each other. First of all, the adjacency matrix of social network is created. The existence of relation between two people is shown by 1, otherwise by 0. Then any row in the table which has no cell with 1 value will be filled by 1/N in which N is the number of people in social network (e.g. N is 4 here).

Then for other rows, any cell which has the value 1 is calculated using equation

$$\frac{R_{ij}}{\sum_{z=0}^{z=n} R_{iz}}$$

In this relation, R_{ij} represents the weight of the link between person i and person j which is the number of posts between them. Now some transformations like multiplying by probability of teleport action are carried out on this matrix. Then operation of PageRank values calculations will be started. For this reason, the Marcov chain is used.

- **Crawler:** First of all, the relevant user's data from the social network are extracted using the crawler and then they're stored into text format files.

- **Data Formatting:** Before the data could be used, they should be stored in a well formed structure in order to make queries possible on the data. In this phase, the proper data are separated and after doing pre-processing and cleansing operations are stored in a structured way into the social network data warehouse.

- **Experts Finding:** In this section, we find experts by applying the variation of PageRank algorithm on the data warehouse. Output is then converted in to displayable format.

3.3 Social Networking Approach: New Perspective

When we consider above approaches there are some pros and cons. These methods have certain limitations. Both of them heavily depend on Data mining techniques. In first Web mining approach there is implicit requirement of downloading documents, research papers, and publications from academic websites, Conferences websites, Journals or scientific publications websites. It is not possible to have free access to most of such websites. These websites charge people to get such data from their site. They are usually not free. Even with paying charges they provide limited access not allow to download all their data for single subscription.

Second method of Link Analysis needs heavy crawlers to crawl all networks data and format it. If Social Network has huge data then crawlers needs to be of that potential to span such large data. Also determining the person's importance depends on no of interactions rather than quality of interaction which may sometimes mislead.

Considering above limitations and other available technologies, we thought problem of Expert Finding and Quality content Search can be viewed from a different perspective. Problem can be approached in different way than above approaches. This sparked a thought that we can create a complete social network of experts where people will decide expert and content quality than machines and algorithms. A person's expertise and Content quality will be

decided by peer reviews. Using this approach Expert's of a topic or domain can share their knowledge; connect with other Users, Seek advice or help from other Experts, share resources, etc. Users of Network will get Topic Specific High Quality content, Current topics on which most of people are working or topics most being discussed. Users can work with peer researchers for their proposed work or recommend others for research work.

4. System Design and Implementation

4.1 CMS: WordPress with Buddypress & bbpress

Today in Market various Content Management Systems are available. Examples of such systems are Wordpress, Drupal, Joomla[12]. After doing analysis of these content management systems and our problem statement, we thought that its better choice to use one of the content management systems instead of building a social network from scratch. Choosing a CMS has some advantages rather building a site from scratch. Some of them are: Scalable Architecture, Time Saving, and Support Available. Considering above features of Content Management Systems, we decided to Use Wordpress along with its Extensions Buddypress and bbpress.

Wordpress: WordPress is a powerful semantic publishing platform, and it comes with a great set of features Its freely distributed, standards-compliant, fast, light and free content management system with an customizable core[9]. It is a free and open source blogging tool and a content management system (CMS). It's an elegant, well-architectured personal publishing system built on PHP and MySQL and licensed under the GPLv2. It is one of the world's most popular blog publishing applications.

Features of Wordpress: User Management, Full Standards Compliance, Extend with Plugins, Freedom, Community Support, Admin Dashboard, Proven, Dynamic page generation.

Buddypress: BuddyPress is a powerful plugin that takes your WordPress.org powered site beyond the blog with social-network features like user profiles, activity streams, user groups, and more. Users can create a focused social network for new product. BuddyPress makes it easy for you to build your community or social network and establish relationships among members with default components such as Extended Profiles, Activity Streams, Friend Connections, Private Messaging, User Groups and Notifications.

Features of Buddypress: Profile, Private Messaging, Friends, Groups, Activity Stream, Notifications.

BBpress: An internet forum is a discussion area on a website where members submit posts, and then others discuss them. A forum can deal with various topics or can be dedicated to a single issue. These forums are also called as Message Boards, Bulletin Boards and Discussion Groups.

4.2 System Architecture

With Chosen content management system, overall system architecture can be viewed as in following figure. Wordpress is at the Core of the system. Buddypress and bbpress are built on top of the Wordpress. The functionalities that we will be extending will be built on top of these and will make use of APIs and functions of Wordpress, Buddypress and bbpress. Structure can be viewed as similar to OS kernel which is at Core and other functionalities are built using system calls.
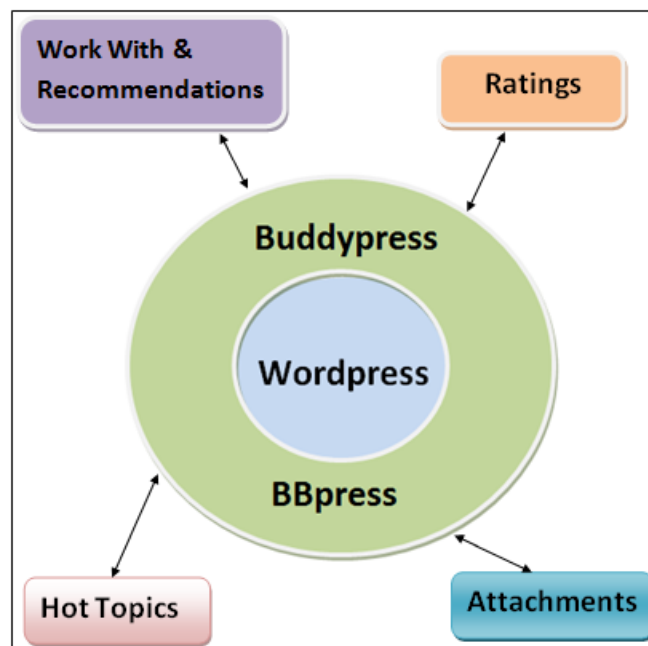

Figure 1: System Architecture

4.3 System Implementation

Creating Environment
We are adding new features on top of existing wordpress, buddypress and bbpress combination, we need to these to be configured. Wordpress, buddypress and bbpress are written in PHP and uses MySQL as database. To run the PHP scripts, we need an Apache web server. 'XAMPP' is the packaged freeware available which comes with these applications bundled together. 'XAMPP' is the most preferred tool for 'PHP and MySQL' based web application development. Now we will be creating additional functionalities by developing new plugins with help of wordpress APIs which are called as Hooks. We also need to modify some core files or make use of Shortcodes.

Developing work with plug-in
This plugin is intended to provide functionality that a User can apply to work with an Expert for research purpose. Other people can recommend that person and Expert can see who has applied to work with him. We are implementing this functionality with help of Shortcodes and editing one Core file[8].
- **Shortcodes:** Shortcodes in wordpress are like handlers. We register a function with help of shortcode. Wherever we call the shortcode registered for that function in Wordpress, in turn that function is executed.

- **Work With:** We add a 'Work-with' button on Profile of every member. When a researcher wants to work under guidance of an Expert, researcher will go to expert's profile and request for 'work-with'. This will create a database entry for the request.

Recommendations

For recommendations we need to create a new Page from Admin for Main Navigation Menu. This page will call shortcode for Recommendations and execute that. The database entries created during work with requests are displayed on recommendation page in three sections. One for requests that you have made. another for requests that you have received and last for Other request that you can recommend.

Deciding Content Quality - Rating Plug-in

We are developing this functionality so that User can rate the quality of the post. To implement this we use APIs called actions and filters in wordpress terminology. We also use some predefined functions of wordpress and bbpress. To display the Rating icons and total ratings, we need to edit the HTMLarea of the Post. We make use of two predefined tables 'sh_usermeta' and 'sh_postmeta' and update these tables with posted, userID and ratings whenever a logged in User updates the rating for the post.

Topics Most Discussed – Hot Topic

This functionality intends to display list of most active topics on network. This Functionality is implemented as a Widget. Wordpress Core class of Widget is extended for that. We first set default values for parameters that show No. of Topics, User's Name, Forum Name and Order to Display. Then we read these values from Admin dashboard set during activation of the plugin. Now we prepare query to retrieve data from database tables of posts using predefined bbpress functions as per options configured in dashboard. Post this we display the content retrieved using query.

Implementing Attachments for Forums

Attachments functionality is implemented using Wordpress media library and its supportive functions. Files that are attached in Topic post or replies are stored in 'wp-content\uploads' directory of the project. Meta data related to attachments like file name is stored in database table of post meta. We make use of 'add_action' hooks to manage topic posts, manage reply posts. We define file formats that are allowed to upload and error messages for different conditions. For uploading files on server we use wordpress media functions wp_insert_attachment(), wp_generate_attachment_metadata(), wp_update_attachment_metadata(). These functions upload the file in Uploads directory.

5. Testing and Results

5.1 Testing Basic Features

A Basic set of features comes with Wordpress combined with Buddypress and bbpress package. These features give most of the functionalities that any Social Networking Site will need. Here are highlights of how these basic features

work which are provided by Wordpress along with Buddypress and bbpress.

5.1.1 User Registration and Login

User can create own account using registration process and login to network using login credentials created using registration process.

5.1.2 Members

Every registered user of the network is a member. Every member has certain other components attached to him/her. Each member has Profile which contains data filled during registration which can be altered. With every member's page; certain associated components are available which include User's Activity, User Profile, Notifications, Messages, Friends, Groups, Forums and Settings.

5.1.3 Groups and Forums

User Can Create own group. Group admin can alter group details like Name, Description, Privacy Setting, Avatar, etc. Most important feature of a Group is that Forum associated with it. Under the hood of Forum, Members can create Topics. Users can create multiple topics which enable Topic bases discussion and organize it better rather than having a common group discussion board.

5.1.4 Activity

Activity is displayed at component level or system-wide. There are certain places where Activity is displayed like home page, groups. Activities range from New User Registration, Profile updates, New Friends connection created. Topics created and replied.

5.1.5 Site Admin

Site Admin can manage the site, its content and features using the dashboard. Site Admin Can View, modify, add, delete various contents like Users, Groups, Topics, replies, Forums, Pages, etc. Site Admin can configure different options available under Dashboard to decide how site will appear to End Users and what features will be accessible for them.

5.2 Testing Extended Features

Following sections cover functionalities and their results that are implemented as part of extension to basic features and would be focusing on special features to users of this site.

5.2.1 Work-With and Recommendations

Purpose of the Functionality is that automating the system of Applying for research and seeking recommendations from others. User can send work with requests to Experts and other users can recommend that pair. Below is screen how Recommendation page will look and User can recommend.

5.2.2 Rate Content

If User likes or agrees with the Content of the Topic or reply posted on Topic or Question in forum; then user can rate that content. If user disagrees the content or reply given in forum then user can rate it negatively. User can rate positively or negatively for content. Total Cumulative rating is displayed for each post.

5.2.3 Hot Topics

For topics that are discussed most on network, a widget is displayed on left panel of the site. The widget is displayed throughout the site in left panel for easy and quick access. User can be on any page of site and can access most discussed topics in one click. Widget displays the five most discussed topics in panel. If user wishes to see more details about the topic then User can click on the topic and will be directed to the topic in the forum.

Figure 2: Recommendations

5.2.4 Attachments

When User posts in Forum under a topic or gives reply to a post, User can add attachment as part of the post. Attachments are helpful to better explain the topic details. If a question is asked in forum then, answer can be better explained with help of some diagrams or reference files. User can attach files of type image, word, pdf, etc.
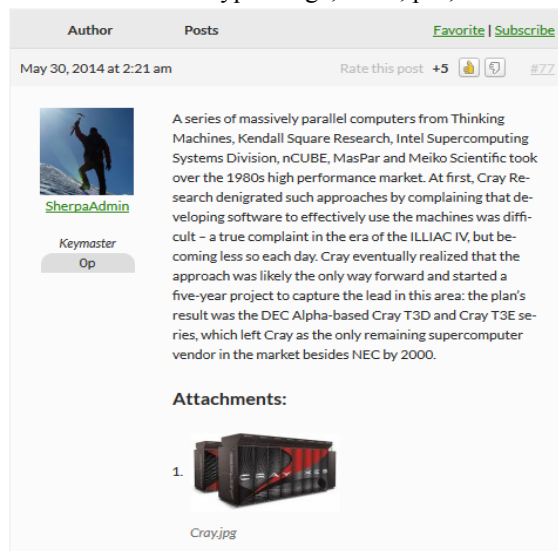

Figure 3: Attachments and ratings in Forum

6. Conclusion and Future Scope

6.1 Conclusion

We discussed various approaches that are used for expert search. We proposed that Expert search problem can be tried to solve using a different technique. Using Social Networking; we proposed that an Expert's Social Network can be one of the answers to solve problem of Expert search and get help from Expert. We used a framework of Social network and added useful functionalities on top of the basic features, so that Researchers can take help of these and get benefited from Expert's Advice. User of this network can make friends, send messages, create groups, receive notifications, etc. User can share Topic based articles or ask questions in Forums for knowledge sharing. Apart from this User can rate the content quality, See current trending topics that are most discussed and Share files of various formats for better communication. User of the network can apply for work with an Expert and get recommendations from other members of network; simplifying the process of getting recommendations. Expert's Social Network can be considered as alternative to other Expert search techniques.

6.2 Future Work

Features added could be 'connecting with other social networks or websites to search an Expert'. For creating trust factor Expert Verification could be provided. Expert verification could be provided by verifying author published documents in Journals and conferences. Feature for Expert suggestions could be included where for each Topic few experts will be suggested by Network itself.

References

[1] Guojun Liu, Ming Zhang, Fei Yan - Large-Scale Social Network Analysis based on MapReduce, 2010 International Conference on Computational Aspects of Social Networks

[2] Data Infrastructure at LinkedIn, 2012 IEEE 28th International Conference on Data Engineering Quan Thanh Tho, Siu Cheung Hui and A.C.M. Fong - A Web Mining Approach for Finding Expertise in Research Areas.

[3] Ahmad Kardan, Salar Mohtaj - A Model for Expert Finding based on Social Network Structure and Underlying Information Diffusion Network, 2013 5th Conference on Information and Knowledge Technology

[4] Yanyan Li, Shaoqian Ma, Yonghe Zhang, Ronghuai Huang - Expertise Network Discovery via Topic and Link Analysis in Online Communities, 2012 12th IEEE International Conference on Advanced Learning Technologies

[5] Anongnart Srivihok - Building FExpert: System for Searching Experts in Research University Using K-Means Algorithms, 2012 IEEE Symposium on Computers & Informatics

[6]   Ahmad Kardan, Amin Omidvar, Farzad Farahmandnia - Expert Finding on Social Network with Link Analysis Approach

[7]   Antonio Tapiador and Diego Carrera - A survey on social network sites' functional features, Proceedings of IADIS WWW/Internet 2012. Madrid

[8]   WordPress Plugin Development Cookbook – by Yannick Lefebvre

[9]   www.wordpress.com

[10]   www.blogboostup.com

[11]   www.webolia.com

[12]   Satyen Parikh, V.R. Rathod, Savan K Patel - Joomla. Drupal and WordPress - A Statistical Comparison of Open Source CMS

[13]   Jian Chen, Jin Huang - Expert Finding in Virtual Learning Community

[14]   Lan Zhang_, Xiang-Yang Li_,z, Yunhao Liu_,y, QiuYuan Huangx, Shaojie Tang - Mechanism Design for Finding Experts Using Locally Constructed Social Referral Web,

[15]   Tang Rui, Simon Fong, Sababady Sarasvady - Expert Finding and Query Answering for Collaborative Inter-Organizational System by Using Rule Responder

Author Profile

Mr Sangramsingh M Ghatage received a B.E(Comp) from Shivaji University, Kolhapur. He has two years of Professional Experience of IT Industry. He is currently M.Tech student at College of Engineering, Pune, India.

Mr Siddharth K Gaikwad completed M.Tech in Computer Science. He is currently working as assistant professor at college of engineering,Pune,India. He is life Member of Indian Society for Technical Education