OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

RANKING RESULTS IN A WEB BASED COMMUNITY QA SYSTEM

Ms. Monika Sonmale[1], Mrs. Sushma Nandgaonkar[2]
P. G. Student, *Department of Computer Engineering, VPKBIET, Baramati, India[1]*
Assistant Professor, *Department of Computer Engineering, VPKBIET, Baramati, India[2]*
monikasonmale@gmail.com[1] , *sushma.nandgaonkar@gmail.com[2]*

Abstract: **Web forums like Stack overflow, Yahoo Answer and Quora comes under community question answering system (CQA) are gaining popularity. CQA archives of previously asked questions and their answers, this means that one can freely ask any question on web forums and expect some good, honest answers. On other side, it takes effort to go through all possible answers of search question and to make sense of them but this process take too much time like popular question to have hundreds of answers, and it is very time-consuming for a user to inspect them all. To resolve this problem thus system mainly works on ranking of QA pairs and online forum. This QA pairs are ranked by fine grained and QA forwarding technique. This technique ranks based on rating given by common user. This technique explicitly captures relationships between questions and their best answers by modeling topical dependencies. In some of condition unsatisfied answers on system resolved by online forum system with direct communication. In QA system not only user can post question and answer, but also it can correct spell at the time of posting question and answer, this system used fuzzy spell check API.**

Keywords: *CQA, Fuzzy spell check API, fine grained and QA forwarding, online forum*

.----------∴∴∴----------

I INTRODUCTION

Community Question Answering (CQA) on web forums such as Quora, Yahoo answer and Stack Overflow are more popular. In which user can post question and get answers freely. This has been seen on two sides a) a user can freely ask any question and can expect a best variety of answers based on the answers rating. b) It takes efforts to go through the provided answers of varying quality and to make sense of them. It is not a better option for a popular question to have hundreds of answers, and it is very time-consuming for a user to inspect them all.

The main approach is to propose a system which may help to automate the process of finding best answers of newly posed questions. So going forward to this paper, it introduced rank based QA pair and online forum support technique.

The Community Question and Answer (CQA) system have huge number of users where they have different types of questions. User needs to login with interest in specific community. User asks questions in his community and rates the answers so that answerer can act like expert. User acknowledges the answerer when he gets best answer by rating same answer and likewise high rated answers act like experts and also helps to and near to best answer [1].

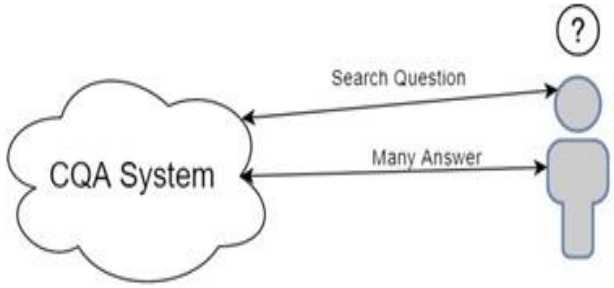

Figure 1 Introduction of CQA

It improves response latency and answer quality in both the social community and global community. Here, interest coefficient based uncategorized question forwarding algorithm and weak tie assisted social based potential answerer location algorithm will be used [1].

Earlier QASs provided answers to the user by matching keywords and concept, but because of that sometimes the user didn't get proper answers to the

questions. To solve this challenge, CQA system was developed (shown in Fig. 1). CQA websites like Stack Overflow, Quora, Answers, Yahoo!, Cross etc to get answers to their questions. These websites allow individuals to post their questions online, and multiple experts across the globe answer them.

A. Problem Statement

When user try to find answers from QA systems he/she get many answers and to find appropriate answer is very tough work because it consumes lots of time and user have to read each and every answers manually. To address this following system is proposed.

B. Overview of Our System

In existing system when user ask new question in a search engine results will be retrieved which may not be appropriate. Getting an appropriate answer is difficult. Also it can take long time to get answers. And there can be confusions because of lots of answers from different user having different view. So there are some drawbacks of existing system and to solve this problem a new system called CQA is implemented.

When user ask question it is first forwarded to that community. Users in that community answers question if it is not found there then it are forwarded to the global community. Then asker will rate users with the quality of answers. And thus they will act like experts to askers. Their answer will have more preference than others. And likewise best answers can be decided. Fig. 2 shows the block diagram of proposed system.

Here every user will first register themselves with all information including their interests, education and knowledge. And then users will be grouped in community so that in a community all related users will be present with their matching interests, education and knowledge. The user of the system has the authority to post new questions into the search box. When user post or ask question then system will try to find the best answers of the asked questions.

To find best answers, the system will be using the fine-grained technology. It identifies the top K users with the highest rating as the appropriate answerer candidates and sends the question to them. If there is no better answer after a time out, the virtual server posts the question on the forum, where each user can see and answer the question asked by the user [1].

Figure shows the system architecture of proposed System, in this user need to do registration providing all information including his interest. This will be reflected in their particular community, this will be related to users interest, education and knowledge, it will be used for matching profile.

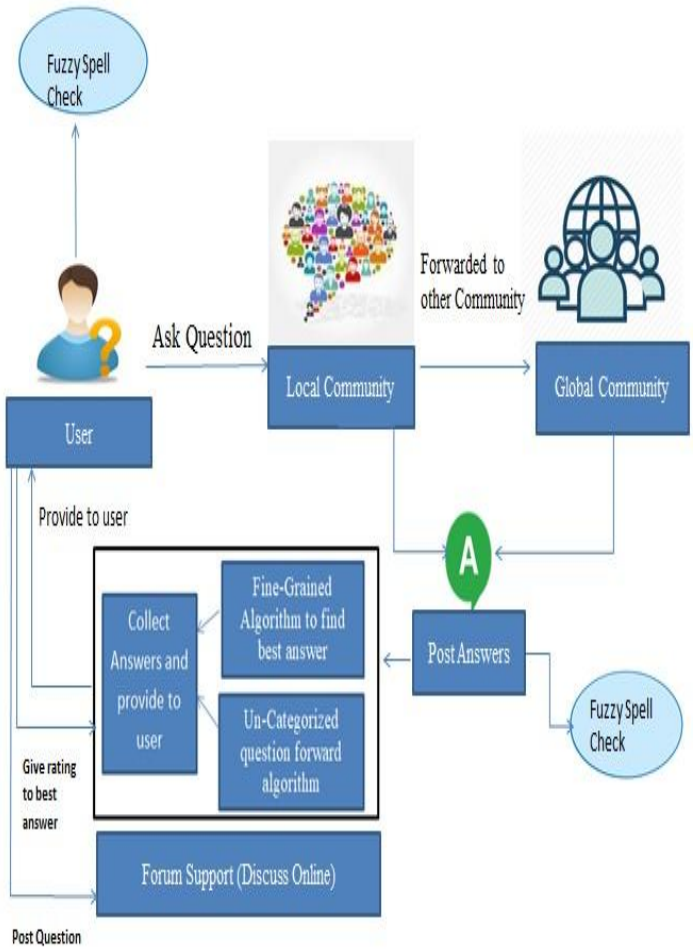

Figure 2 System Architecture

II RELATED WORK

Recently, research in social-based QA systems increased rapidly [2], [3], [4], [5], [6], [8], [9], [10], [11]. Most of social-based QA systems use distributed approach to identify credible answerer. Ze Li and Haiying Shen [2] proposed a distributed Social-based mobile QA System (SOS), which enables mobile users to send questions to potential answerer in their friend lists in a decentralized manner. It depends on lightweight knowledge engineering techniques to accurately find out friends who are able to and willing to answer questions, thus reducing the search and computation costs of mobile nodes [2].

Even though recent commercial search engines are mostly based on information related with QA, but it is still difficult to collect an appropriate related content from numerous user specified answers in QA websites. And in order to get the users which could help people to find relevant answers, GunWoo [3] proposed a ranking algorithm called Influence Rank, which works as the base for analyzing the relationship between user's activities and their mutual understanding. QA system in [4] helps to find best answer based on up votes and down votes using users rating on answers. But, there is very much less knowledge about

properties of experts and non-experts and on what basis experts need to be decided in general topics or a specific topic.

Social networks have been used for efficient and cooperative file sharing and distribution in peer-to-peer (P2P) Networks [5], [6]. Cheng et al. [5] suggested NetTube, a good associated video forwarding framework that looks for the clustering in social networks for short video sharing. NetTube solved a long queue of key-design issues to understand the system, which included bi-layer overlay, an effective indexing method, and a pre-fetching methodology using social networks. The works in [6] focuses on locating experts and authoritative users as potential answers for QA systems. To recognize reliable users and content in social media, Jiang Bain [6] developed a semi-supervised coupled common boosting framework. This framework concurrently calculates elements quality and user position. This framework needed relatively less marked examples to start the training process of the system.

David et al [8] re-ranked the search results by calculating their relevance with individuals in the requesters social network. S. Bao [9] optimizes a web search by using social annotations from the following two aspects: similarity ranking and Static ranking.

Guangyou Zhou et al [10] Question retrieval in CQA will mechanically notice the foremost relevant and up to date questions that are resolved by alternative users. However, the word ambiguity and word match issues bring on new challenges for question retrieval in CQA. State-of-the-art approaches address these problems by implicitly increasing the queried questions with extra words or phrases mistreatment monolingual translation models. Rui Zhao et al [11] BoW-based vector illustration of a document, every part denotes the normalized variety of prevalence of a basis term within the document. To count the amount of prevalence of a basis term, BoW conducts actual word matching, which may be thought to be a tough mapping from words to the basis term. BoW illustration suffers from its intrinsic extreme sparsity, high dimensionality, and inability to capture high-level linguistics meanings behind text information. To handle the higher than problems, planned a new document illustration technique named Fuzzy Bag-of-Words (FBoW) is used. Additionally, it uses word clusters rather than individual words as basis terms and develop Fuzzy Bag-of-WordClusters (FBoWC) models. Document representations learned by the planned FBoW and FBoWC area unit dense and able to encode high-level semantic. The results on seven real word document classification datasets in comparison with six document representation learning methods have shown that our methods FBoW and FBoWC achieve the highest classification accuracies.

TABLE 1
SUMMARY OF LITERATURE SURVEY

Sr. No.	Authors	Techniques/Algo/Methods used
1	ZeLi ,Haiying[2012]	ZeLi ,Haiying[2] works on non- factual questions posted by users by searching friends close to them and interested.
2	P. GunWoo [2011]	P. GunWoo [3] uses ranking algorithm for CQA users. It can be used for answering questions of other users.
3	Sumanth P, Kyumin Lee [2015]	This work focuses on finding experts from sites like Quora where they can act like answer system to users.
4	Cheng et al. [2009]	In this research work peer to peer video forwarding system is made so that there will be fewer loads on servers of YouTube and there will be less cost.
5	Jiang Bian[2009]	In this work, Jiang Bian [6] found out the quality of content and the users who answer the questions in CQA systems.

III BASIC TERMINOLOGY

Question and answer Posting and fuzzy spell check API- When question posted on CQA system it will be forwarded to system using fuzzy spell check for spell correction [11].

Centralized data management- Data is stored in centralized location where data from all communities is stored so that it can be shared to other communities and it will help use to get answer.

Fine grained system- Here, based on rating and reputation best answer is found and then based on his reputation we can find expert so that he will help user to get best answer [1].

Question forwarding algorithm- This algorithm is used to forward question to communities where the question matches with askers interest and community interest so that it will get forwarded to community where user will get best answer in less time [1].

Reward system- When user with good reputation and rating is obtained. They are given rewards in terms of recognition [1].

Forum site- A Web forum is a website or section of a website that allows visitors to communicate with each other by posting messages. Most forums allow anonymous visitors to view forum postings, but require you to create an account

in order to post messages in the forum. In a forum, you can create new topics or post replies within existing message [10].

IV TECHNICAL REPRESENTATION

We use following technique:

1. Fine grained technique: it finds the best answerer candidates with the help of user rating.
2. Social based potential answers: Find posted question Answers in local community.
3. Global Based Potential answers: Find same question's answers.
4. Reward System: Best answers get a reward for best answer. When the user is not getting good answers, he/she enter that question in forum site Fig. 2 shows the flow of the whole system of CQA, in that main entity belongs to the fined grained answers selection techniques. This form the best answers related to the previous system.

ALGORITHM

Forum Support-

1. User U= u1, u2, u3.....un; complete the registration(User)
2. getUniqueId()< -0;
3. message< -null
4. subject< -null1
5. language< -L= (l1, l2,l3....ln);
6. foreach(i in U) do
7. if (User register)then // generate unique userId
8. getUniqueId< -i;
9. Else // not register
10. registration(User)
11. End if
12. IncludeCurrentDiscussionTopic()
13. //previous user are talking on particular subject
14. getSubject()
15. getFamiliarLanguage(language l)
16. DiscussForum() // send message in the group
17. message< -sent message;
18. Return message;

V EVALUATIONS

In this section, we evaluate the Effectiveness and efficiency of different QA systems. In that describes Fuzzy Dictionary, Question and Answer, Users with Gmail Accounts, Performance metrics.

Experimental Data and Settings

Fuzzy Dictionary: When user post new question and answer our API this will work fine with correcting spell. This dictionary has more than 50,000 words. These words compared with entered keywords and given correct output. Sample data set of fuzzy spell is (mydictionary.txt):

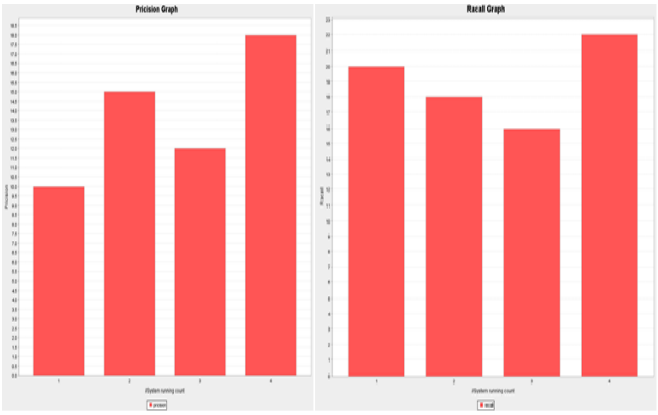

Figure 3 Effectiveness and efficiency of different QA systems.

TABLE II

PRECISION AND RECALL VALUE

System Running Count	Precision	Recall
1	15	18
2	18	22
3	10	20

Users with email Accounts: User want to enter in CQA system login with help of email id, at the time of posting question, the posted question sent via mail using email to respective people present in community. Currently we have more than 40 Gmail users.

Performance metrics: Recall, Precision was performance metrics which used throughout the experiments. Recall quantifies the percentage of the true question with his respective answer that is correctly recognized. Precision quantifies the percentage of the extracted all question with low rating.

Precision = TP / TP + FP + 0.1;

Recall = TP / TP + FN + 0.1;

Where, TP: True Positive (all High rated answer)

FP: False Positive (Initial rated answer but present in search result) and 0.1 is constant value to avoiding exception at the time of developenemt.

VI PERFORMANCE ANALYSIS

This section investigates the impact of different proposed CQA system components on its performance.

A. Performance of request sending and response time at single request

Fig shows that request time of searching question and calculating total time. Calculating time graph:

StartTime=System.getCurrentTime();

CQA System(best answerer)

EndTime=System.getCurrentTime();

Total time=EndTime StartTime

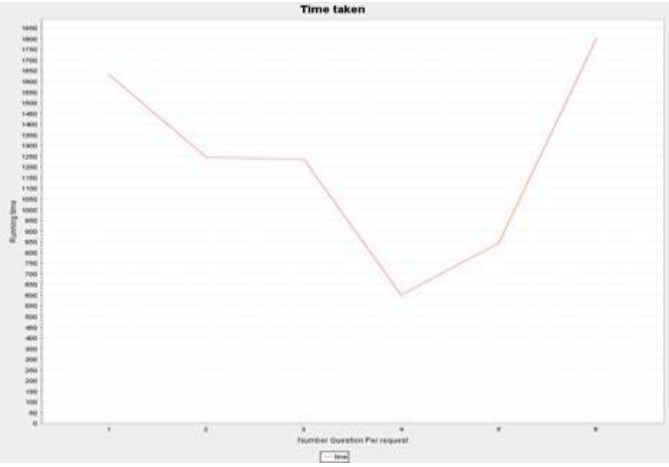

Figure 4 Time graph.

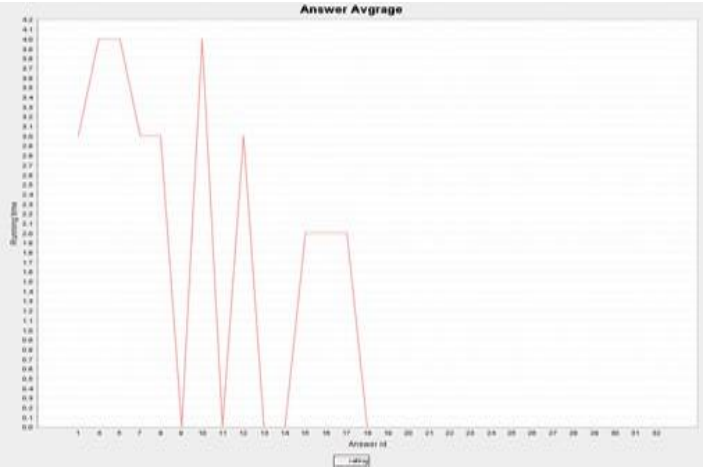

Figure 5 Average graph

B. Average rating of answer**.**

In this section we are calculating average rating of answer which are been posted by multiple user as following AvgRating = number of rating of answer / number of user given rating

VII APPLICATION

- Community forum
- Best answer finding
- On the spot solution finding

VIII CONCLUSION AND FUTURE WORK

This system uses concept called CQA which will use two communities to answer the question asked by user. To find good answerer candidates in a users social network, CQA uses a question forwarding which consider multiple factors in evaluating the answer QoS of the users friends. If a answers is not obtained in social then question is forwarded to global community [1]. CQA builds central servers where information is stored to efficiently locate answerer candidates in the interest of the question. CQA has a fine grained reputation system to find experts, and which depends on a reputation-based reward system that adaptively rewards question answerers based on their reputations, in order to provide some value in answering questions we also use weak tie assisted social based potential answerer location algorithm and the interest coefficient based uncategorized question forwarding algorithm to further improve its performance. So, this system helps user to find best answer [1].

In future, we will study the fault tolerance after CQA system failure and some recommendation related to the eBook with detail information and with appropriate links can be added.

ACKNOWLEDGMENT

This paper would not have been written without the valuable advices and encouragement of Mrs. S. S. Nandgaonkar, guide of ME Dissertation work. Authors special thanks go to all the professors of computer engineering department of VPKBIET, Baramati, for their support and for giving an opportunity to work on ranking results in a web based community QA system.

REFERENCES

[1] G. Liu and H. Shen, "iASK: A Distributed QA System Incorporat- ing Social Community and Global Collective Intelligence,"*In Proc. of the IEEE Transactions on Parallel and Distributed Systems,* 2017.

[2] Ze Li; Haiying Shen; Guoxin Liu; Jin Li. SOS: A Distributed Mobile Qamp;A System Based on Social Networks, IEEE 32nd Interna- tional Conference on Distributed Computing Systems,Pages: 627 636, Year: 2012.

[3] P. GunWoo, Y. Soung Woung, L. SooJin, and L. SangHoon Credible user identification using social network analysis in a qa site, 2011.

[4] Sumanth Patil1, Kyumin Lee, "Detecting experts on Quora: by their activity, quality of answers, linguistic characteristics and temporal behaviors,"2015.

[5] X. Cheng and J. Liu. NetTube: Exploring Social Networks for Peer- to- Peer Short Video Sharing. In *Proc. of INFOCOM,* 2009.

[6] J. Bian, Y. Liu, D. Zhou, E. Agichtein, and H. Zha. Learning to Recognize Reliable Users and Content in Social Media with Coupled Mutual Reinforcement. In *Proc. of WWW,* Volume:11, Pages: 485 489, 2009.

[7] Markus Heckner, Christian Wolff. Towards Social Information Seeking and Interaction on the Web. *In: Kuhlen, Rainer, (ed.) Information: Droge, Ware oder Commons,Pages: 235-241,* 2009.

[8] E. Amitay, D. Carmel, N. HarEl, S. A. Ofek-Koiman. Golbandi, "Social search and discovery using a unified approach," In *Proc. of HT,* 2009.

[9] Z. Wang, L. Sun, X. Chen, W. Zhu, J. Liu, M. Chen, and S. Yang. Propagation-based Social-Aware Replication for Social Video Contents. In *Proc. of ACM Multimedia,* 2012.

[10] Guangyou Zhou, Zhiwen Xie, Tingting He, Jun Zhao and Xiaohua Tony Hu, "Learning the Multilingual Translation Representations for Question Retrieval in Community Question Answering via Non-negative Matrix Factorization," 2016.

[11] Rui Zhao and Kezhi Mao, "Fuzzy Bag-of-Words Model for Doc- ument Representation," 2017.

[12] E. Bakshy, I. Rosenn, C. Marlow, and L. A. Adamic. The Role of Social Networks in Information Diffusion. CoRR, 2012.

[13] M. Granovetter. The Strength Of Weak Ties. American Journal of Sociology 78, 1360-80, 1973.

[14] Ask, http://www.ask.com, [Accessed in May 2015].

[15] Answers, http://www.answers.com, [Accessed in May 2015].

[16] Yahoo! Answers, http://answers.yahoo.com, [Accessed in May 2015].

[17] Stackoverflow, http://stackoverflow.com/, [Accessed in May 2015].

[18] Quora, http://www.quora.com, [Accessed in May 2015].

BIOGRAPHY

Ms. Monika S. Sonmale has received the bachelors degree in computer engineering in 2016 from SCOEM, Shivaji University, Satara, Maharashtra, India. Currently pursing Master of Com-puter Engineering in VPKBIET,Baramati,Pune.

Mrs. Sushma S. Nandgaonkar has received his Bachelor degree from WIT, Solapur(Shivaji University) and Master degree from COEP, Pune(Pune University). She is currently working as Assistant Professor, Department of Com- puter Engineering, Vidya Pratishthans Kamal- nayan Bajaj Institute of Engineering and Technology, Baramati, Pune, Maharashtra, India. His research areas include Uncertain Data Mining, Application of Data Mining in Business and Intelligent predictions.