ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)
Vol. 3, Issue 3 (July - Sept. 2015)
International Journal of Advanced Research in Computer Science & Technology (IJARCST 2015)

Social Based Q&A System in Cloud Environment

[I]G.Taruni, [II]A.P.Siva Kumar
[I]M.Tech Scholar, Dept of CSE, Jawaharlal Nehru Technological University, Anantapuram, India
[II]Assistant Professor, Dept of CSE, Jawaharlal Nehru Technological University, Anantapuram, India

Abstract
Q&A systems on social search engines provide a way to get information from the people who are socially connected. But they are not providing satisfactory results as the friends reply with unnecessary answers to the questions which lead to high server bandwidth. In this paper, we propose a Cloud based Social Q&A system which gives quick response for the question askers with appropriate answers. The system forwards the question to the accurate and willing person and we also analyze the appropriate TTL (Time to Live) value. So that it will avoid the unnecessary message overload.

Keywords
Question and answer system, On-line social network, non-factual questions, first order logic.

I. Introduction

Traditional search engines like Google, Bing, and Yahoo provide information for factual queries. These do not provide any answers for nonfactual queries which are specific, imaginary, subjective and multi-dimentional. So inorder to enhance the performance of search engines, social search engines are proposed. The social search engines gather and group the people of similar interest and initiate the search query to the relevant person in the group.

Although the search engines answer factual queries that are already stored in centralized server hence this technique is not suitable for answering non factual queries. For example, "can anyone recommend me a Doctorate professor for doing my project in social network...?" If the valid information is not found in database then we need to forward these queries to the human, which are the most "intelligent machines". The persons who are expertise in that particular topic can give perfect answer.

The advent of Social Web is clearly a game-changer, on numerous fronts. The rush of social web to implement and focus on marketing specially Versus the business is more universal, "social media marketing" is just the latest technique in marketing, but the truth is it is a new way that can be used to market and promote your services and products. It also concentrates on interaction with colleagues, latent customers as well as current clients and helps to spread message in a serene and conversational way. It also results in many costly disturbances to users by sending queries that cannot be answered and hence increasing the workload by looking on the queries through the pool of queries. It leads to high server bandwidth and high query congestion and maintenance costs.

To enhance the askers satisfaction, researchers focused on systems in which the users post and answer in social network. SOS (Social based Q&A system) is such system in which the answerers in the network are connected socially. SOS also leverages a lightweight technique to transform the users closeness and their interests. So that the question is forwarded to the answerer who are close and have similar interests. By which, the node overload is reduced and a quick response is received within low cost. As the answerers have similar interest, the accuracy of answer is also increased.

SOS transforms the question into ID's, so that the node can easily compare it with the social ID. Then the node forwards the question to the particular person with that social ID. After receiving the questions the answerer can answer if he knows the answer or if he doesn't, he can forward the question to his friends. As the answerers are socially related to askers, the willingness to answer the question is more compared to the strangers. By choosing the potential answerer from the friends list, the question is finally answered.SOS uses NLP and FOL techniques to calculate questions ID. NLP technique will divide the question into groups of related words (e.g., wh-type). First order logic will parse the question into tokens.

In SOS, the question is forwarded to friends in the friend list. The question can be forwarded to TTL number of hops, so that if the askers' friends don't know the answer they can forward to their friends hop. By this the accuracy of getting an answer is improved.

II. Related work

The paper that helped in this work is,[2]Damon Horowitz's Aardvark, a social search engine. With Aardvark users can ask queries through emails, text messages or either by voice or instant messaging. Aardvark routes the question to the users in the social network who are more likely to answer the question. The main components of Aardvark are:

1. Crawler and Indexer: finds the appropriate user.
2. Query Analyzer: To extract the topic of the query.
3. Ranking Function: To select the best user that provides the information.
4. UI: To present the information to the user in an understandable manner.

In any case, it is difficult for the asker to check the content that is given back is trustworthy or right for them. In these cases, askers are looking for personal opinions, recommendations, or advice, from someone they feel a connection with and trust.

[3] The main objective of this research is to enhance the quality of answers and to reduce wait time by forwarding questions to users who are interested or expert in the area to which the question belongs. This paper proposed SocialQ&A which is a social network based Q&A system that identify and notify the users who are more interested to answer a question. SocialQ&A includes three major components:

User Interest Analyzer: *It* analyzes the interest of the user.
Question Categorizer: identify the category of the question.
Question-User Mapper: identifies a list of potential answer providers for each question.

A protypical testing was also done on the system to evaluate the performance, which was satisfactory.

[4] In this paper PeopleRank approach was developed in which the nodes are ranked according to the social information inorder to reduce the retransmission. PeopleRank is a social distributed algorithm which measures the importance of a node in a social graph based on the social interaction contact frequency between

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

the nodes .It achieved a success rate by reducing the number of retransmission by 50%.

III. Proposed Work

In this paper, we propose social search on cloud based Q&A system that provides the large resources to store the information. It also provides quick response to the question and sharing questions is also feasible. We also analyze the appropriate value for the Time-To-Live (TTL) , that provides a satisfactory success ratio, it avoids redundant message overhead and reduces the waiting time . The First Order Logic Technique also been used to calculate interest ID's and speed of the answers.

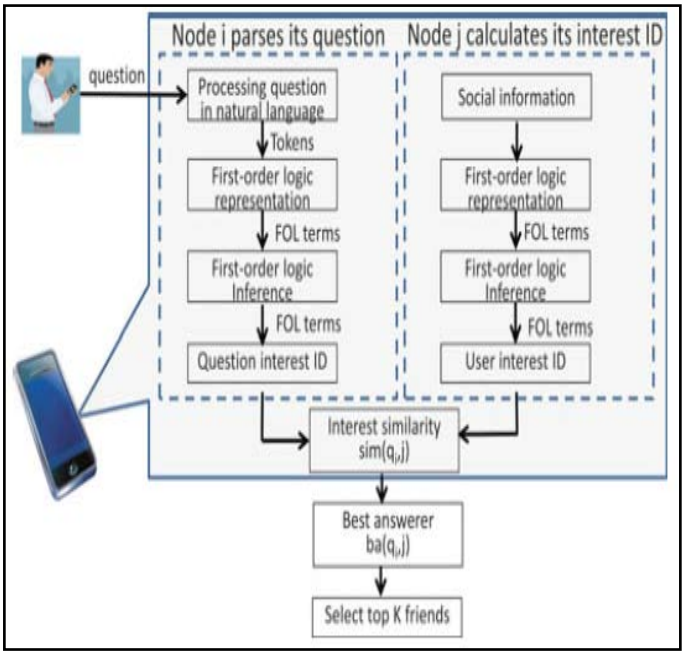

Fig 1: System Architecture

SOS associates with an online social network, where nodes are socially connected. Every user has an interest ID which is created based on the profile of the user which represents the interest of the user. The users who give answers are considered as best answerers only if the asker is satisfied with the answer to that question. The architectural diagram in above figure helps to find the answerers.

When a question is posted by the user, the node processes the question using NLP (Natural Language Processing) and then represent the in first order logic format by dividing question into tokens keywords then apply inference rules on the tokens to infer the questions interest. Finally a question id is created based on the interest. This question id is compared with the friends Interest id and if the id's match, question is forwarded to those friends.

When a question is posted by the user, the node processes the question using NLP (Natural Language Processing) and then represent the in first order logic format by dividing question into tokens keywords then apply inference rules on the tokens to infer the questions interest. Finally a question id is created based on the interest. This question id is compared with the friends Interest id and if the id's match, question is forwarded to those friends.

After finding the friends with similar interest, it will select k number of best answerers among those friends. It will select the k answerers based on the feedback or performance of the users. If they have a good response feedback then they are selected among k friends.

An algorithm helps to find the best answerers and forwards the question to those best answerers. When a question is received or created, a question id is created and similarity is calculated for all friends. The node will give a list of similar friends and based on the quality of answers they give, they are sorted and the question is forwarded to k number of top friends

Algorithm:

Algorithm 1 Pseudocode of the best answerer identification executed by node i.

1: **Input:** ID_i, ID_j, $Q_{(i,j)}$ $(j \in \mathcal{F}_i)$
2: **Output:** top-K best answerers
3: *//Periodically update* $Q_{(i,j)}$ $(j \in \mathcal{F}_i)$
4: **for** each friend j in friend list $\mathcal{F}_i$ **do**
5: Update $Q_{(i,j)}$ based on Equation (2)
6: **end for**
7: **if** create a question or receive a question it cannot answer **then**
8: **if** TTL>0 **then**
9: **for** each friend j in friend list $\mathcal{F}_i$ **do**
10: Calculate $S_{(q_i,j)}$ using ID_{q_i} and ID_j based on Equation (1)
11: Calculate $BA_{(i,j)}$ using $Q_{(i,j)}$ and $S_{(q_i,j)}$ based on Equation (3)
12: Add $BA_{(i,j)}$ to a list $List$
13: **end for**
14: QuickSort partition around the K^{th} largest element in $List$
15: Find the top-K friends having the highest $BA_{(i,j)}$
16: TTL-=1
17: Send the question to the identified K friends
18: **end if**
19: **end if**
20: **if** does not receive answers for its created question during the time corresponding to TTL **then**
21: Resort to the centralized server for the answers
22: **end if**

If the answer is not answered by the k friends then for how much time the question can reside inside the system, that time is represented as TTL (Time to Live). There is a facility for the friends to forward the question to their friends. Whenever the question is forwarded TTL increases. If they go on forward the question to more hops then there is a risk of question overload. So inorder to solve the problem we limit the number of TTL.We will set value for TTL, whenever a question is forwarded to a hop the value is decreased by 1. If the value reaches to 0 then we will stop forwarding.

If the question is not answered, even though the TTL value tends to 0, the cloud server will share the question to all the members in the system that belongs to the interest of the question. So that the answer is answered finally. The cloud server will store all the unsolved questions.

IV. Conclusion

In this paper, we present how these Q&A system can accurately identify the best answerers who are expertise in that area. SOS leverages lightweight techniques to identify the related friends. This system also provides answers to non-factual queries which are specific and multi-dimentional. It uses FOL for the retrieval of interest of the user and the question. It also earns high user satisfaction ratings for the accurate answers. SOS generates very less overhead with limited question forwarding. Since each user is connected to several social groups, it selects most probable answerers and forward to an answerer that can provide an answer.

We can also get response from any location by accessing through internet by means of cloud server which provides storage of large resources. All the question and answers are stored in the cloud. The future of Q&A system in cloud is demandable and scope full.

References

[1] *Haiying Shen, Ze Li, Guoxin Liu, and Jin "SOS: A Distributed Mobile Q&A System Based on Social Networks", IEEE Transactions on Parallel and Distributed Systems Vol. 25, No. 4, April 014.*

[2] *D. Horowitz and S.D. Kamvar, "The Anatomy of a Large-Scale Social Search Engine," Proc. 19th Int'l Conf. World Wide Web(WWW '10), 2010*

[3] *Nikhil Vithlani "SOCIALQ&A: A NOVEL APPROACH TO NOTIFIYING THE CORRECT USERS IN QUESTION AND ANSWERING SYSTEMS", (2012). All Theses. Paper 1423.*

[4] *A. Mtibaa, M. May, C. Diot, and M. Ammar, "Peoplerank: Social Opportunistic Forwarding," Proc. IEEE INFOCOM, 2010.*

[5] *Z. Li, H. Shen, and K. Sapra, "Leveraging Social Networks to Combat Collusion in Reputation Systems for Peer-to-Peer Networks," Proc. IEEE Int'l Parallel and Distributed Processing Symp. (IPDPS '11), 2011.*

[6] *K. Toutanova and C.D. Manning, "Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger," Proc. Joint SIGDAT Conf. Empirical Methods in Natural Language Processing and Very Large Corpora (SIGDAT '00), 2000.*

[7] *H. Shen and K. Hwang, "Locality-Preserving Clustering and Discovery of Resources in Wide-Area Distributed Computational Grids," IEEE Trans. Computers, vol. 61, no. 4, pp. 458-473, Apr. 2012.*

[8] *H. Shen and C. Xu, "Locality-Aware and Churn-Resilient Load Balancing Algorithms in Structured Peer-to-Peer Networks," IEEE Transactions on Parallel and Distributed Systems, Vol.18, No. 6, June 2007.*

[9] *Google, http://www.google.com, 2013.*

[10] *Bing, http://www.bing.com, 2013.*

Author's Profile

G.Taruni received B.Tech degree in Computer Science Engineering from Rajeev Gandhi Memorial College of Engineering and Technology Nandyal, affiliated to JNTUA College of Engineering, Ananthapuramu, A.P, India, during 2009 to2013. Currently pursuing M.Tech in Software Engineering at JNTUA College of Engineering, Ananthapuramu, A.P, India, during 2013 to 2015.Her area of interest is Cloud Computing.

Dr. A.P.Siva Kumar is an Assistant Professor of Computer Science and Engineering at Jawaharlal Nehru Technological University College of Engineering, Ananthapuramu. He obtained his Bachelor degree in Computer Science & Information Technology from Jawaharlal Nehru Technological University Hyderabad, Master of Technology in Computer Science from Jawaharlal Nehru Technological University Hyderabad and Ph.D. in cross lingual information Retrieval from Jawaharlal Nehru Technological University Ananthapuramu. He has published several Research papers in National \ International Conferences and Journals. His research interests includes Information Retrieval and Natural Language Processing.