International Journal of Advanced Research in
Computer Science Engineering and Information Technology
Volume: 4 Issue: 1 19-Dec-2014,ISSN_NO: 2321-3337

SURVEY ON PERSONALIZED SOCIAL SEARCH WITH CLOUD BASED Q&A SYSTEM

[1]R.JAYA SHREE, [2]M.RAMESH KUMAR

[1]M.E Student, [2]Assistant Professor

[1,2]Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, Tamil Nadu, India.

ABSTRACT- *Question and Answer (Q&A) system on social search engines provide a way for retrieving an information to motivate the people to answer the questions. These Questions can get High-Quality answers but it does not able to balance the needs of the question askers because some askers may want answers urgently but instead they waiting long time for the friends reply without getting answers for their questions and by these it also occur high server Bandwidth. In this paper, we propose personalized social search with cloud based Q&A system that provide quick response for the question askers. These Q&A services can share non-factual queries and also colloquial languages. Here the askers sharing questions is feasible and we can also use large resources to store the information. We also analyze the appropriate settings for the Time-To-Live (TTL) value, for TTL-Limited flooding, that provides a satisfactory success ratio and it avoids redundant message overhead.*

Key Terms- Question and answer systems, Web search, online social networks.

1, INTRODUCTION

Traditional search engines such as Google and Bing are the primary way for information retrieval on the Internet. To improve the performance of search engines, social search engines[8] have been proposed to determine the results searched by keywords that are more relevant to the searchers. These social search engines[8] group people with similar interests and refer to the historical selected results of a person's group members to decide the relevant results for the person.

Although the search engines perform well in answering factual queries for information already in a database, they are not suitable for non-factual queries that are more subjective, relative and multi-dimensional (e.g., can anyone recommend a professor in advising research on social-based question and answer (Q&A) systems?), especially when the information is not in the database (e.g., suggestions, recommendations, advices). One method to solve this problem is to forward the non-factual queries to humans, which are the most "intelligent machines" that are capable of parsing, interpreting and answering the queries, provided they are familiar with the queries. Accordingly, a number of expertise location systems[12] have been proposed to search experts in social networks or Internet aided by a centralized search engine. Also, web Q&A sites[2] such as Yahoo! Answers and Ask.com provide high-quality answers[17] and have been increasingly popular.

To enhance the asker satisfaction on the Q&A sites, recently, emerging research efforts have been focused on social network based Q&A systems[12],[5],[3],[9],[2] in which users post and answer questions through social network maintained in a centralized server. As the answerers in the social network know the backgrounds and preference of the askers, they are willing and able to provide more tailored and personalized answers to the askers. The social-based Q&A systems can be classified into two categories: broadcasting-based[12],[5],[3] and centralized[9],[2]. The broadcasting-based systems broadcast the questions of a user to all of the user's friends. In the centralized systems[9],[2], since the centralized server[2] constructs and maintains the social network of each user, it searches the potential answerers for a given question from the asker's friends, friends of friends and so on.

To tackle the problems in the previous social-based Q&A systems and to realize a mobile Q&A system, a key hurdle to overcome is: How can a node identify friends most likely to answer questions in a distributed fashion? To this problem, in this paper, we propose a distributed Social-based mobile Q&A System (SOS) with low node overhead and system cost as well as quick response to question askers. SOS is novel in that it achieves lightweight distributed answerer search, while still enabling a node to accurately identify its friends that can answer a question. We have also deployed a pilot version of SOS for use in a small group in Clemson University. The analytical results of the data from the real application show the highly satisfying Q&A service and high performance of SOS.

SOS leverages the lightweight knowledge engineering techniques to transform users' social information and closeness, as well as questions to IDs, respectively, so that a node can locally and accurately identify its friends capable of answering a given question by mapping the question's ID with the social IDs. The node then forwards the question to the identified friends in a decentralized manner. After receiving a question, the users answer the questions if they can or forward the question to their friends. The question is forwarded along friend social links for a number of hops, and then to the server. The cornerstone of SOS is that a person usually issues a question that is closely related to his/her social life. As people sharing similar interests are likely to be clustered in the social network the social network[6] can be regarded as social interest clusters intersecting with each other. By locally choosing the most potential answerers in a node's friend list, the queries can be finally forwarded to the social clusters that have answers for the question. As the answerers are socially close to the askers, they are more willing to answer the questions compared to strangers in the Q&A websites. In addition, their answers are also more personalized and trustable.[7]

In a nutshell, SOS is featured by three advantages:

(1) Decentralized. Rather than relying on a centralized server, each node identifies the potential answerers from its friends, thus avoiding the query congestion and high server bandwidth and maintenance cost problem.

(2) Low cost. Rather than broadcasting a question to all of its friends, an asker identifies the potential answerers who
are very likely to answer this question, thus reducing the node overhead, traffic and mobile Internet access.

(3) Quick response. An asker identifies potential answerers from his/her friends based on their past answer quality and answering activeness to his/her questions.

1.1 Objective:

The main aim is to develop a personalized social search Q&A system for quick response to question askers.

2, SYSTEM ANALYSIS

2.1 EXISTING SYSTEM

A Distributed Social-based mobile Q&A System (SOS) which provide a searches with an additional mechanism which cannot be easily resolved by web search engines[8]. SOS can achieve a high query precision and recall rate. The feedback from the users shows that SOS can provide high-quality answers[3]. Here the First Order Logic technique can be used to calculate their interest and speed of the answers. NLP Technique has also been used to divide the question in to group of related words, 2-words phrases, and the wh-type(when, where and what).

The drawback of this Existing System that has limited resources to store the information and it also have high server bandwidth and waiting long time for friends reply and it accept only non-factual queries and not colloquial languages.

2.2 PROPOSED SYSTEM

In this paper, we propose a personalized social search[8] on cloud based Q&A system that provide the large resources to store the information. It also provide quick response to the question answers and sharing questions is feasible. We also analyze the appropriate settings for the Time-To-Live (TTL) value, that provides a satisfactory success ratio and it avoids redundant message overhead and waiting time is reduce because of using TTL. Here we also share non-factual queries and colloquial languages for asking questions. The First Order Logic Technique also been used to calculate their interest and speed of the answers.

2. ARCHITECTURAL DIAGRAM

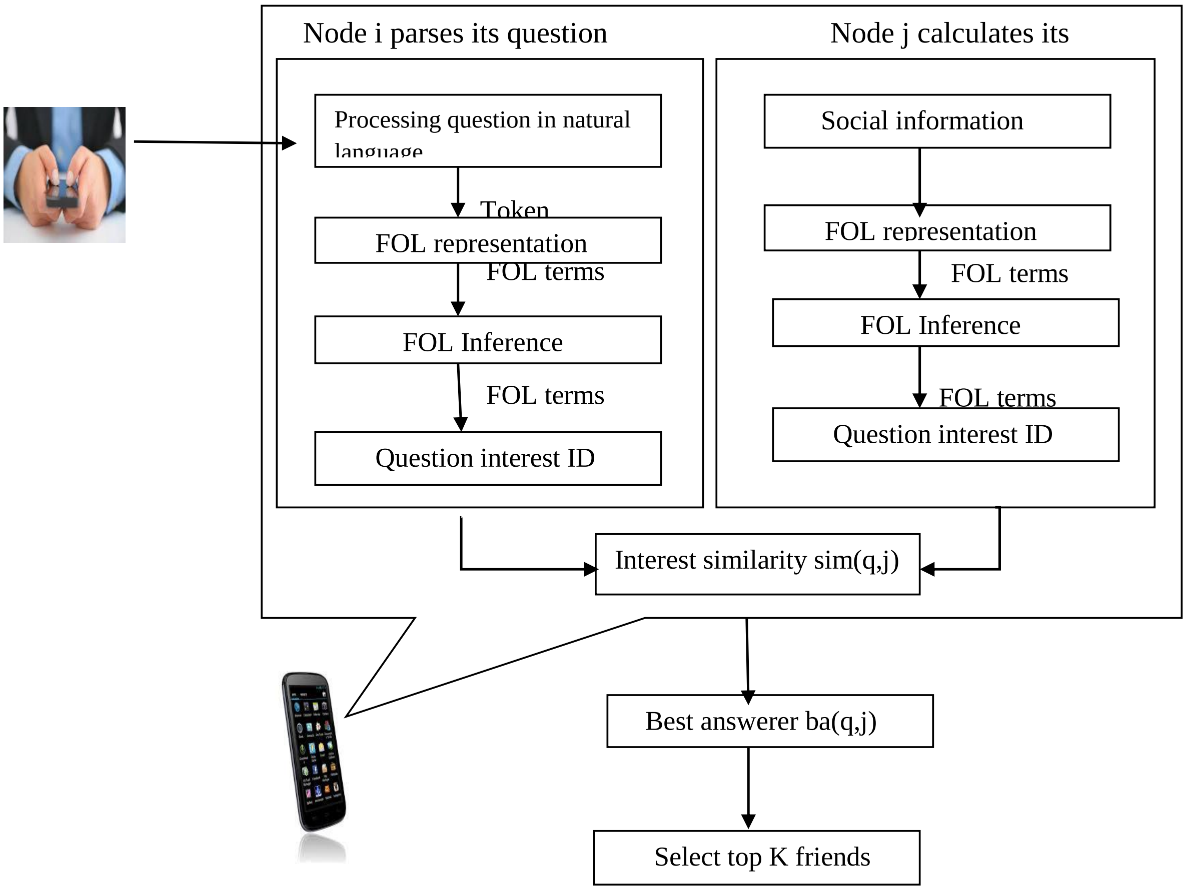

Fig:1 Architecture diagram

These architectural diagram shows that the local answerer selection process for forwarding a question in one mobile node in the SOS system. To parse a question, the node first processes the question using natural language processing (NLP), and then represents the question in the FOL format and uses the FOL inference to infer the question's interests. Finally, it transforms the question to a question ID in the form of a numerical string. After node i parses its initiated question qi to a question ID, it calculates interest similarity S(qi;j) for each of its friends j 2 Fi, where Fi denotes the set of node i's friends. It then calculates the best answerer value (BA(qi; j)) for each friend j by combining S(qi;j) and answer quality from friend j (Q(i;j)).

3, LITERATURE SURVEY

[1] C. Lampe, J. Vitak, R. Gray, and N.B. Ellison have looked at how Social network sites(SNS) are used for information seeking, both in workplace setting and to fill common information needs. So they analyzed how the Facebook connects a people to communicate, exchange messages, post the status and update their profiles. But they cannot able to predict the facebook as an information seeking behavior because it contain multiple channels for information seeking, including dyadic and group messaging, as well as network-wide interaction via status updates.

[2] M. Richardson and R.W. White have investigates by predicting the variety of outcomes during the question life cycle. Here they used Synchronous social Q&A system which provide the searches with an additional mechanism beyond the general-purpose search engines through which to get questions answered. They also used the data from an instant messaging for automatically identifies candidates answerers by ranking all users by representation of their interest and routes the questions only to those available and most able to answer, and mediates the dialog between the asker and answerer. Using this we found that they only gathered the interaction data from the real users.

[3] J. Teevan, M.R. Morris, and K. Panovich have explored what factors might influence how people respond to other online. So they analyzed the quality, quantity and speed of the responses to the question askers. Here the users use their status messages to ask questions of their networks, these includes to create social awareness, encouraging the asker to reflect on a current need, building social ties and finding answers. Here they used ANOVA tests for measuring the response quantity and response quality. If the questions are phrased as the statement that look like a regular status updates, and it not be responded as a question.

[4] D. Horowitz and S.D. Kamvar are presented Aardvark which is used to find the right person to satisfy the user information need. They also find the vast majority of questions get answered promptly and satisfactorily so the users also surprisingly active for both in asking and answering. The questions include the background details but it is not obvious how to translate these information needs in to keyword searches.

[5] M.R. Morris, J. Teevan, and K. Panovich have presented the detailed data from a survey of facebook and Twitter users on the topic of asking and answering questions via social status messages and those responses are very valuable. This approach would be most effective if the speed of receiving social answer could be optimized. Survey participants appreciated the context inherent in interactions with their social network. This finding suggests that enhancing the effectiveness of personalized search algorithms is an important area of further study for improving the usability of search engine.

[6] J. Raacke and J. Bonds-Raacke have studied to identify the dimensions of use for friend networking sites that including the information dimension, the friendship dimension, and the connection dimension. The reasons in the information dimension were related to using friend networking sites to gather and share information. The reason in the friendship dimensions were conceptually related to sustaining friendship. Then the reason in the connection dimensions that were related to making the connection with others and this dimension included to make new friends, to find a significant other, and to feel connected in general.

[7] M.R. Morris, J. Teevan, and K. Panovich have conducted a study to compare social and non-social search for complex, self-motivated information seeking tasks. In this paper they focused on the aspects of social search where the searchers ask a question to the group of people they personally by means of a social network status message update. They compared this experienced to the searching for the same information with a Web search engine. They found an average of 1.4 responses waiting, and a maximum of 5 responses and 5 participants received no responses compared to the social search.

[8] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N. Har'el, I.Ronen, E. Uziel, S. Yogev, and S. Chernov have investigated the personalized social search based on the user's social relations which provide the effectiveness of Familiarity, similarity and overall based network. Their assumption behind this personalization approach is that the preferences to other people, who are expected to have "similar" interests as the searcher and assist the search results that might subjectively satisfied the searcher needs.

4, BEST ANSWERER IDENTIFICATION ALGORITHM

The pseudo code of the best answerer identification executed by node i.

Input: ID_i, ID_j, $Q_{(i,j)}$ (j£Fi)
Output: top-K best answerers
For each friend j in friend list Fi **do**
  Update $Q_{(i,j)}$
end for
If create a question or receive a question it cannot answer **then**
  If TTL>0 **then**
    For each friend j in friend list F_i **do**
      Calculate $s_{(qi,j)}$ using ID_{qi} and ID_j
      Calculate $BA_{(i,j)}$ using $Q_{(i,j)}$ and $s_{(q,j)}$
      Add $BA_{(i,j)}$ to a list List

end for
Quick Sort partition around the Kth largest element in List
Find the top-K friends having the highest $BA_{(i,j)}$
TTL-=1
Send the question to the identified K friends
end if
end if
if does not receive answers for its created question during the time corresponding to TTL **then**
Resort to the centralized server for the answers
end if

These Algorithm shows the pseudo code of the process for the best answerer metric calculation and best answerer selection conducted by node i. If node i does not receive answers for its created question during the time corresponding to TTL, it resorts to the centralized server for the answers, where all users conduct Q&A activities as in online Q&A sites. Lines 3-5 are used to periodically update answer quality of each of its friends. Lines 7-12 calculate each friend's best answerer metric and generate a list including all metric values.

Lines 13-17 identify the top-K friends with the highest best answerer metric values and send the question to them. Answer quality $Q_{(i;j)}$ is pre-processed, and only interest similarity $S_{(qi;j)}$ needs to be calculated at run time. The $S_{(qi;j)}$ calculation has a time complexity of $O(|F_i|)$. As the number of keywords in a question is generally very small, the calculation of $S_{(qi;j)}$ should take a short time and costs little computation resources of the mobile devices. This top-K friend selection algorithm has a time complexity of $O(|F_i|)$.

5, RELATED WORK

6.1 PERSONALIZED SEARCH

In recent years many researchers utilize query log and click-through analysis for web search personalization. Joachims et al.[4] study clicks applicability as implicit relevance judgments. They show that users' clicks provide a reasonably accurate evidence of their preferences. Tan et al.[3] propose a language modeling approach for query history mining. Their small-scale study demonstrates significant improvement of personalized web search with a history-based language model over regular search. The user modeling approach described in [30] is based on a decision-theoretic framework to convert implicit feedback into a user profile that is used to re-rank search results. Agichtein et al. [2] introduce an alternative user modeling method, in which a set of rules is applied to a query log. While user models are usually targeted at search personalization, they could also be applied for personalized information filtering, as was shown in [7] who analyze click history for the identification of regular users' interests. Recent work of

Teevan et al. [11] on “groupization” shows that combining implicit user profiles from several related users has a positive impact on personalization effectiveness.

In addition to regular web log data, several works consider personalization using desktop data and external resources. For example, in [12] the authors index desktop information and experiment with different representations of users, documents and queries for personalized web search. Chirita et al. [9] explore personalized query expansion based on users’ desktop information. Several approaches for personalized Web search are based on global interests using the Open Directory Project (ODP) categories. In the authors map previously visited pages to ODP categories and use this mapping to build a user profile. Another work [10] proposes a personalized version of Page Rank, in which a hand-picked set of preferred users’ categories are applied for result re-ranking.

Recently, new approaches for adaptive personalization focus on the user task and the current activity context . There are several approaches trying to predict applicability of personalization while considering the current context of the user’s task on query submission.

CONCLUSION

In this paper, we present how these Q&A system can accurately identify the answerers that are able answer the questions. To analyze the problem getting only factual queries, we examine the non-factual queries that are more subjective, relative and multi- dimensional. Here we also proposed that the answerer will also ask questions by colloquial languages for better understanding of their friends. We can also get the data from any location by accessing the internet by means of cloud server. Here we also considers both its friend’s parsed interests and answer quality in determining the friends similarity value, which measures both the capability and willingness of the friend to answer / forward the questions.

REFERENCES

[1] C. Lampe, J. Vitak, R. Gray, and N.B. Ellison, “Perceptions of Facebook’s Value as an Information Source,” Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI ’12), 2012.
[2] M. Richardson and R.W. White, “Supporting Synchronous Social Q&A Throughout the Question Lifecycle,” Proc. 20th Int’l Conf. World Wide Web (WWW ’11), 2011.
[3] J. Teevan, M.R. Morris, and K. Panovich, “Factors Affecting Response Quantity, Quality, and Speed for Questions Asked via Social Network Status Messages,” Proc. ICWSM, 2011.
[4] 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.
[5] M.R. Morris, J. Teevan, and K. Panovich, “What Do People Ask Their Social Networks, and Why?: A Survey Study of Status Message Q&A Behavior,” Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI ’10), 2010.

[6] J. Raacke and J. Bonds-Raacke, "MySpace and Facebook: Applying the Uses and Gratifications Theory to Exploring Friend-Networking Sites," Cyber Psychology & Behavior, vol. 11, p. 169, 2008.
[7] M.R. Morris, J. Teevan, and K. Panovich, "A Comparison of Information Seeking Using Search Engines and Social Networks," Proc. Fourth Int'l AAAI Conf. Weblogs and Social Media (ICSWM '10), 2010.
[8] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N. Har'el, I.Ronen, E. Uziel, S. Yogev, and S. Chernov, "Personalized Social Search Based on the User's Social Network," Proc. 18th ACM Conf. Information and Knowledge Management (CIKM '09), 2009.
[9] R.W. White, M. Richardson, and Y. Liu, "Effects of Community Size and Contact Rate in Synchronous Social Q&A," Proc. SIGCHI Conf. Human Factors in Computing Systems, 2011.
[10] E. Amitay, D. Carmel, N. Har'El, S. Ofek Koifman, A. Soffer, S. Yogev, and N. Golbandi, "Social Search and Discovery Using a Unified Approach," Proc. 20th ACM Conf. Hypertext and Hypermedia (HT '09), 2009.
[11] Z. Li and H. Shen, "SOAP: A Social Network Aided Personalized and Effective Spam Filter to Clean Your E-mail box," Proc. IEEE INFOCOM, 2011.

[12] F. Harper, D. Raban, S. Rafaeli, and J. Konstan, "Predictors of Answer Quality in Online Q&A Sites," Proc. SIGCHI Conf. Human Factors in Computing Systems (SIGCHI '08), 2008.