© 2018 IJSRSET | Volume 4 | Issue 7 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Conferencing Terming to End of Discussions as Query

M. Ramu[1], P. Harishwar Reddy[2], P. Sreejani[2], T. Sai Tarun[2]

[1]Assistant Professor, Department of Information Technology in Teegala Krisha Reddy Engineering College, Telangana, India

[2]UG Scholar, Department of Information Technology in Teegala Krisha Reddy Engineering college, Telangana, India

ABSTRACT

Question and Answer (Q&A) systems play a vital role in our daily life for information and knowledge sharing. Users post questions and pick questions to answer in the system. Due to the rapidly growing user population and the number of questions, it is unlikely for a user to stumble upon a question by chance that (s)he can answer. Also, altruism does not encourage all users to provide answers, not to mention high quality answers with a short answer wait time. The primary objective of this paper is to improve the performance of Q&A systems by actively forwarding questions to users who are capable and willing to answer the questions. To this end, we have designed and implemented SocialQ&A, an online social network based Q&A system. SocialQ&A leverages the social network properties of common-interest and mutual-trust friend relationship to identify an asker through friendship who are most likely to answer the question, and enhance the user security. We also improve SocialQ&A with security and efficiency enhancements by protecting user privacy and identifies, and retrieving answers automatically for recurrent questions. We describe the architecture and algorithms, and conducted comprehensive large-scale simulation to evaluate SocialQ&A in comparison with other methods. Our results suggest that social networks can be leveraged to improve the answer quality and asker's waiting time. We also implemented a real prototype of SocialQ&A, and analyze the Q&A behavior of real users and questions from a small-scale real-world SocialQ&A system.

Keywords: Question And Answer Systems, Social Networks, Information Search

I. INTRODUCTION

The Internet is an important source of information, where the amount of data is vast and constantly growing. Users rely on search engines to find specific information in this knowledge base. Search engines such as *Google* and *Bing* use keywords provided by the users to perform searches. Re- cently, industrial research and development activities, such as Microsoft and Facebook's social-featured Bing search en- deavor, try to combine search engines and online social net- works for higher search performance. As previous research has indicated , search engines perform well in indexing web pages and providing users with relevant content to their search but are not suited for non-factual questions such as "Which is the best local auto shop?". To address this particular class of non-factual questions, many Question and Answer (Q&A) systems such as *Yahoo! Answers*, *Baidu Zhidao*, *StackExchange*, *Quora* and *Ask* have been developed. Since their inception, Q&A systems have proved to be a valuable resource for sharing expertise and consequently are used by a large number of Internet users. They are not only important for sharing technical knowledge, but also as a source for receiving advice and satisfying one's curiosity about a wide variety of subjects.

To meet this need, we propose SocialQ&A, an online social network based Q&A system, that actively forwards questions to those users with the highest likelihood (capabil- ity and willingness) of answering them with expertise and interest in the questions' subjects. The design of SocialQ&A is based on two social network properties. First, social friends tend to share similar interests (e.g., lab members majoring in computer systems). Second, social friends tend to be trustworthy and altruistic due to the property of "friendship fosters cooperation" . Accordingly, SocialQ&A favors routing queries among friends and iden- tifies a question's potential answerers by considering two metrics: the interest of the friend towards the question and the social closeness of the friend to the asker/forwarder. Thus, the answer receivers have high probability of provid- ing high-quality answers in a short time. Different from the existing Q&A systems, due to the importance of users privacy, we future introduce security and efficiency enhancement to protect users privacy while users using social network answering questions.

Q&A:
Comparative trace-driven experiments.
The development of a real-world SocialQ&A.
The analysis of the data from real SocialQ&A.

•

A user's profile, questions and answers
Output: The user's interest vector VUj =< Ii, WIi >

1. Parse the "interests" field to generate a token stream TI
2. Parse the "activities" field to generate a token stream Ta
3. Use the inputs from the user's selection from the Music, Movie, Television and Book fields to generate token streams Tmu, Tmo, Tt and Tb
4. for each token stream Tx (Tx=TI , Ta, Tmu, Tmo, Tt, Tb) do
5. Check each token in the Synset
6. if a matching interest category Ii exists then
7. Update interest weight: WIi ++ end if
8. end for
9. Keep updating WIi based on questions asked and answered and profile update
10. Periodically update WIi using WIi = α ∗ WIiold

II. OBJECTIVE

Like all online social networks, the one in SocialQ&A has user profiles that record users' interests, education, hobbies and etc. Like *Yahoo! Answers*, SocialQ&A also predefines interest categories and subcategories. A total of 4 categories (music, movies, television, and books) and 32 subcategories (e.g., books: novel, drama) derived from *Yahoo! Answers* were used to implement SocialQ&A. We used these 4 cat- egories as an example and will add more categories in our future work.

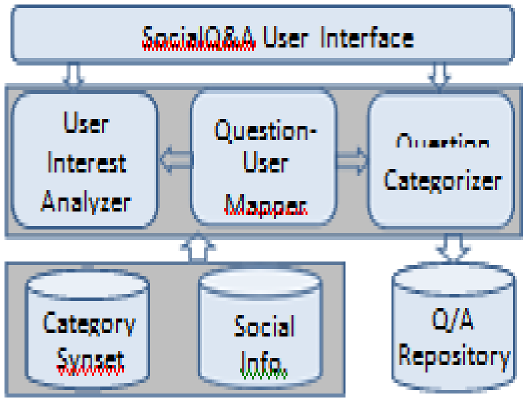

Figure 1 : architecture of interpretation on query

Figure shows the high-level architecture of SocialQ&A and the interaction between the core components: *User Inter- est Analyzer*, *Question Categorizer*, and *Question-User Mapper*. *User Interest Analyzer* analyzes data associated with each user in the social network to derive user interests. *Ques- tion Categorizer* categorizes the user questions into interest categories based on the *Category Synsets*, which stores the synonyms of all categories' keywords from WordNet [41]. *Question-User Mapper* connects these two components by identifying potential answerers who are most likely to be willing to and be able to provide satisfactory answers. The data from user questions and answers is stored on

Q/A Repository to serve subsequent similar questions. Below, we present each component and user interface.

User Interest Analyzer utilizes each user's profile information in the social network and user interactions (answers provided and questions asked) to determine the interests of the user in the predefined interest categories.

The primary task of *Question Categorizer* is to categorize a question into predefined interest categories based on the topic(s) of the question. We also allow users to input self- defined tags associate with questions, which are analyzed in question parsing. *Question Categorizer* generates a vector of question Q_i's interests, denoted by V_{Q_i} , using a similar algorithm as Algorithm 1. While processing a question, SocialQ&A uses WordNet to examine the tags and text of the question and generates a token string. The tokens are given as questions.

A successful response to a question includes answering or forwarding the question, in which if a question receiver has an answer to the received question, he replies to it; otherwise, he forwards the question. Intuitively, each potential answerer willing to answer the question should have at least one very high score for user questions. Thus, we give equal weights to all factors as stated as pairs of questions. The social closeness between friends ranges in data questions if an asker's friend has social closeness larger than 0.1 (s)he is willing to respond to the asker's question. If we set this willingness threshold to be larger, there will be fewer successful responses in both our method and comparison methods, and vice versa. The probability that other friends respond to the question was randomly chosen from 10%, 20%, 30% . The question query rate was set to one question per minute. These parameters are adjustable parameters and their changes will not affect the relative per- formance differences between the systems in comparison. The distribution of response time to a question follows the trace .

We use he Best Answer set of a question existing in the system, and use *RA* to represent theWe calculated the response rate as the number of all suc- cessful responses divided by the total number of question receivers. Questions shows that the response rate follows Social>SocialQ&A>SOS>Interest Random Flooding. In SocialQ&A and Social, users choose friends with higher social closeness who are most willing to answer questions, so they have a higher response rate than others. SOS does not consider the potential willingness of friends with many common interests when calculating social closeness. Thus, its response rate is lower than SocialQ&A and Social, but higher than the other three methods without social closeness consideration.

In SocialQ&A, users may choose friends with high interest similarity but lower social closeness. Thus, it generates a lower response rate than Social. We also see that the response rate of SocialQ&A, Social and SOS decreases as the number of selected answerers increases, since friends with lower social closeness are more likely to drop ques- tions. This result implies that SocialQ&A's incentive works well when the set of answerers selected is small.

III. IMPLEMENTATION

We then measure the performance of the onion routing based answer forwarding method in protecting the identi- ties of answerers and askers and its system overhead. In this experiment, we assume that 50% of users are malicious users, and the identities of answerers or askers are exposed if all relay users in the whole forwarding path are malicious users. This is because the malicious users form all relay users of this onion routing path and the path length of the onion routing is constant, so that they can know the whole path and the first (last) relay user knows that its predecessor (successor) is the answerer (asker). We define the exposure rate as the number of total identified askers and answerers by malicious users over the total number of asker and answerer. Figure 10 shows the exposure rate versus the path length of the onion routing. It shows that the expose rate decreases as the path length of the onion routing increases. This is because for a longer path length, the probability that all relay users are malicious users is lower. It indicates that the onion routing based answer forwarding can better protect the asker/answerer identities when the path length is longer. Figure 10 also shows the total computing time of all users in an onion path for answer forwarding per question. It shows that the computing time increases as the path length increases. This is because each secure communication between two relay users in the path involves data encryption and decryption. Thus, the figure is a showcase to help determine an appropriate path length in reality, that is, we should consider both the identity protection requirement and the system overhead for each user.

Next, we measure the performance of the answer re- trieval method for recurrent questions. In this experiment, each user's bloom filter result of the successfully answered questions is broadcasted through social links within three hops, and top 2 users with the highest scores of the bloom filter result matching are selected to send the recurrent ques- tion searching request. In reality, the newly asked question may not be the same as the former one. Thus, to generate a newly asked question, instead of using the exact recurrent question, we replaced *m* number of words in the question with *m* randomly selected other words among all words of all questions, where *m* is increased from 0 to 3 with a step size of 1. We measure the success rate as the percentage of questions, which are resolved by satisfying answers fetched by the answer retrieval method. We used *Question* to denote the method using the whole newly posted question to match former questions for answer retrieval, and use *n*-gram to denote the *n*-gram based answer retrieval. Figure 11 shows the success rate of answer retrieval methods with different *n*-grams versus the number of changed words in each question. It shows that *n*-grams can successfully retrieve the answers for most of the recurrent questions. Among them, the 1-gram has a better performance than the others when the asked question is not exactly same as its former one. This is because when matching a new question to a former question, 1-gram produces more the same grams than other *n*-grams, which makes the former question have a higher.

SocialQ&A allows users to register and modify user in- formation, add/remove friends, ask/answer/forward ques- tions and check question notifications. Consider a hypothet- ical user named Mike. When Mike registers, he is required to provide essential information about himself, such as his personal information, area of study/expertise, his current interests, and his involvement in other activities. Users are also encouraged to describe their interests in terms of a few predefined categories, such as movies, books, television, music. *User Interest Analyzer* uses the registration informa- tion to determine Mike's interests.

We used the number of questions and answers posted to characterize user activity. Out of 124 users, 75 unique users posted at least one question; 81 unique users provided at least one answer; 26 users (approximately 20%) did not post or answer any questions. The remaining 80% contributed actively to SocialQ&A.

In the test, a total of 24 out of 163 questions (around 15%) remain unanswered, while all other questions have at least one response. As SocialQ&A identifies potential answer providers who have more common interests, close social relationships with the asker, and have interest in the question's category, those question receivers are more likely to answer the question. Thus, SocialQ&A is able to achieve an improvement even with a very limited number of users. We expect that the number of unanswered questions tends to reduce with user growth, because with more users, the range of expertise also becomes broader, a user has more friends to ask questions and more users are willing to answer questions. Practically, we were not able to test SocialQ&A with millions of users. However, current results indicate the promises of SocialQ&A in improving current Q&A systems.

IV. CONCLUSION

Q&A systems are used by many people for purposes such as information retrieval, academic assistance, and discussion. To increase the quality of answers received and decrease the wait time for answers, we have developed and proto- typed an online social network based Q&A system, called SocialQ&A. It utilizes the properties of a social network to forward a question to potential answer providers, ensuring that a given question receives a high-quality answer in a short period of time. It removes the burden from answer providers by directly delivering them the questions they might be interested in, as opposed to requiring answer providers to search through a large collection of questions as in *Yahoo! Answers* or flooding a question to all of an asker's friends in an online social network. The bloom filter based enhancement methods encrypt the interest and friendship information exchanged between users to protect user privacy, and record all *n*-grams of answered questions to automatically retrieve answers for recurrent question. The onion routing based answer forwarding protects the identities of askers and answers. Our comprehensive trace- driven experiments and analysis results on the real-world Q&A activities from the SocialQ&A prototype show the promises of SocialQ&A to enhance answer quality and re- duce answer wait time in current Q&A systems, and demon- strate the secure and efficiency improvement achieved by the enhancements. Since same questions may be presented very differently and the same question may be answered differently in different situation. In the future, we will co- operate with other techniques (e.g. topic modeling [48] and word embedding [49]) into SocialQ&A to find the redundant question with a large scale user set. Due to the dynamic of user behavior, SocialQ&A can cooperate a machine learning method to adjust three parameters appropriately, which needs a large user base and much more usage. We will con- duct tests on a large user base in the real-world experiment.

V. REFERENCES

[1]. M. R. Morris, J. Teevan, and K. Panovich. A Comparison of Information Seeking Using Search Engines and Social Networks. In In Proc. of ICWSM, 2010.

[2]. 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. In Proc. of CHI, 2010.

[3]. Z. Gyongyi, G. Koutrika, J. Pedersen, and H. Garcia-Molina. Questioning Yahoo! Answers. In Proc. of QAWeb, 2008.

[4]. Yahoo!Answers Team. Yahoo! Answers BLOG. http://yahooanswers.tumblr.com/, Accessed on 10/20/2014.

[5]. B. Li and I. King. Routing Questions to Appropriate Answerers in Community Question Answering Services. In Proc. of CIKM, 2010.

[6]. L. A. Adamic, J. Zhang, E. Bakshy, and M. S. Ackerman. Knowl- edge Sharing and Yahoo Answers: Everyone Knows Something. In Proc. of WWW, 2008.

[7]. G. Drosatos, P. Efraimidis, A. Arampatzis, G. Stamatelatos, and Athanasiadis. Pythia: A privacy-enhanced personalized contex- tual suggestion system for tourism. In COMPSAC, 2015.

[8]. S. Li, Q. Jin, X. Jiang, and J. Park. Frontier and Future Development of Information Technology in Medicine and Education: ITME 2013. Springer Science & Business Media, 2013.

[9]. A. Mtibaa, M. May, C. Diot, and M. Ammar. Peoplerank: Social Opportunistic Forwarding. In Proc. of Infocom, 2010. E. Pennisi. How Did

Cooperative Behavior Evolve? Science, 2005H. Shen, Z. Li, G. Liu, and J. Li. Sos: A distributed mobile q&a systembased on social networks. TPDS, 2014.
[10]. A. Mtibaa, M. May, C. Diot, and M. Ammar. Peoplerank: Social Opportunistic Forwarding. In Proc. of Infocom, 2010.
[11]. E. Pennisi. How Did Cooperative Behavior Evolve? Science, 2005.
[12]. H. Shen, Z. Li, G. Liu, and J. Li. Sos: A distributed mobile q&a systembased on social networks. TPDS, 2014.
[13]. A. Spagnolli and L. Gamberini. Interacting via sms: Practices of social closeness and reciprocation. British Journal of Social Psychology, 2007.
[14]. M. L. Radford, C. Shah, L. Mon, and R. Gazan. Stepping Stones to Synergy: Social Q&A and Virtual Reference. Proceedings of the AmericanSociety for Information Science and Technology, 2011.
[15]. M. Richardson and R. White. Supporting Synchronous Social Q&A Throughout the Question Lifecycle. In Proc. of WWW, 2011. W. White, M. Richardson, and Y Liu. Effects of Community Size and Contact Rate in Synchronous Social Q&A. In Proc. of SIGCHI, 2011.
[16]. J. Teevan, M.R. Morris, and K Panovich. Factors Affecting Re- sponse Quantity, Quality, and Speed for Questions Asked via Social Network Status Messages. In Proc. of ICWSM, 2011.
[17]. Z. Li and H. Shen. Collective Intelligence in the Online Social Network of Yahoo!Answers and Its Implications. In Proc. of CIKM, 2012.
[18]. J. Bian, Y. Yang, and T. Chua. Predicting trending messages and diffusion participants in microblogging network. In Proc. of SIGIR, 2014.