



A DISTRIBUTED MOBILE SYSTEM BASED ON SOCIAL NETWORKS

N. Hanumantha Rao

Associate Professor,

Newton's Institute of Engineering, Macherla

ABSTRACT

Recently, emerging research efforts have been focused on question and answer (Q&A) systems based on social networks. The social-based Q&A systems can answer non-factual questions, which cannot be easily resolved by web search engines. These systems either rely on a centralized server for identifying friends based on social information or broadcast a user's questions to all of its friends. Mobile Q&A systems, where mobile nodes access the Q&A systems through Internet, are very promising considering the rapid increase of mobile users and the convenience of practical use. However, such systems cannot directly use the previous centralized methods or broadcasting methods, which generate high cost of mobile Internet access, node overload, and high server bandwidth cost with the tremendous number of mobile users. We propose a distributed Social-based mobile Q&A System (SOS) with low overhead and system cost as well as quick response to question askers. SOS enables mobile users to forward questions to potential answerers in their friend lists in a decentralized manner for a number of hops before resorting to the server.

I. 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 have been proposed to determine the results searched by keywords that are more relevant to the searchers. These social search engines 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 multidimensional (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

have been proposed to search experts in social networks or Internet aided by a centralized search engine. Also, web Q&A sites such

as Yahoo! Answers and Ask.com provide high-quality answers 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 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 and centralized. The broadcasting based systems broadcast the questions of a user to all of the user's friends. In the centralized systems since the centralized server constructs and The number of mobile users who access Twitter Increased 182 percent from 14.28 million in January 2010 to 26 million in January 2011. It was estimated that Internet browser-equipped phones will surpass 1.82 billion units by 2013, eclipsing the total of 1.78 billion PCs by then. The mobile Q&A systems enable users to ask and answer questions anytime and anywhere at their fingertips. Further, broadcasting to a large number of friends cannot guarantee the quality of the answers. The centralized methods, by serving a social network consisting of hundreds of millions of mobile users (which are also rapidly increasing), suffer from high cost of mobile Internet access for clients, high query congestion, and high server bandwidth and maintenance costs. It was reported that Facebook spent more than \$15 million per year for server bandwidth costs and data center rental in addition to \$100 million for purchasing 50,000 servers to relieve the high burden of traffic.

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 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.

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. The contributions of this work are summarized as follows:

1. As far as we know, it is the first work to design a distributed Q&A mobile system based on social networks, which can be extended to low-end mobile devices. The system can tackle the formidable challenge facing distributed systems: precise answerer identification.

2. We propose a method that leverages lightweight knowledge engineering techniques for accurate answerer identification.

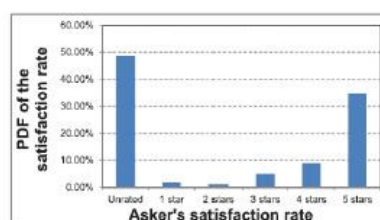
3. We use answer quality to represent both the willingness of a node to answer another node's questions and the quality of its answers. We propose a method that considers both interest similarity and answer quality based on past experience in question forwarder selection in order to increase the likelihood of the receiver to answer/forward the question.

4. We have studied our crawled data from Yahoo! Answer and Twitter with regards to node interactions in online Q&A systems and online social networks. We then conducted extensive trace-driven simulations based on the crawled data. Experimental results show the high answerer identification accuracy, low cost and short response delay of SOS. 5. We have deployed a pilot version of SOS for use in a small group in Clemson University and revealed interesting findings in the mobile social-based Q&A system. Though Google earns a little higher user satisfaction degree than SOS on factual questions, SOS gains much higher satisfaction degree for nonfactual questions than Google.

This journal version presents more comprehensive experimental results compared to its conference version. The remainder of the paper is organized as follows. Section 2 and Section 3 present the trace data and the design of SOS. Section 4 presents the trace-driven simulation results. We conclude this paper with remarks on future work in Section 5. The supplemental material, available online, presents an overview of related work and additional experimental results for the effectiveness of SOS's feedback mechanism, and for our implemented SOS prototype.

II. DATA STUDY FOR POTENTIAL ADVANTAGES OF SOCIAL-BASED Q&A

In order to study the features of people interactions in online Q&A sites and social networks, we crawled 9;419

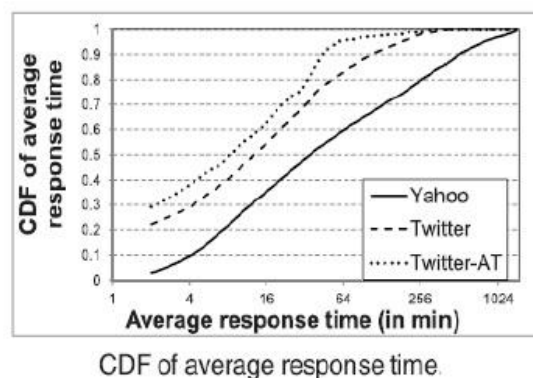
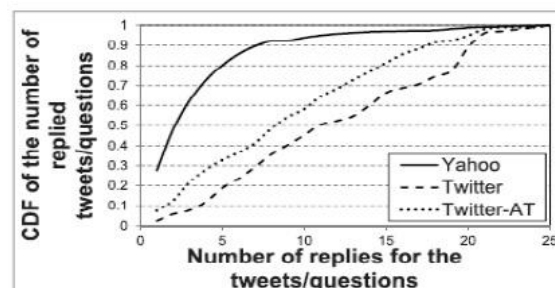


questions posted in the “Entertainment & Music– Movies” section in Yahoo! Answer and 7;559 tweets from Twitter. Since Twitter is not designed to ask questions, we chose user Read Write Web that has a large number of followers in order to find more tweets and user tweeting/reply interactions.

Specifically, we crawled 2;559 tweets posted by user

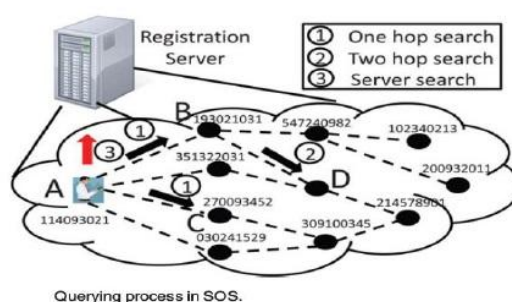
Read Write Web and all his/her followers and followers’ followers that do not have @username and

5;000 tweets that have @username between October 5, 2010 and October 26, 2010. In Twitter, a user A can specifically notify another user B about a tweet by adding B’s @username in the tweet. User B will then receive a notification message from the system about the tweet. In the figures below, we use Twitter- AT to denote the tweets with @username and use Twitter to denote tweets without @username. First, we analyze the satisfaction of Yahoo!Answers users based on their feedbacks. In Yahoo!Answers, an asker is allowed to rate an answer with rating stars 1-5. Fig. 1 shows the histogram of all the askers’ satisfaction with all answers in the data set. We can see that 32 percent of the answerers receive five stars and 8 percent of the answerers receive four stars. These answers take up 80 percent of the answers that receive ratings. The result conforms to the observations in the previous research that the answers provided in yahoo!Answers are quite satisfying. However, nearly 50 percent of answers did not receive ratings. We suspect that one reason is because the users in Yahoo!Answers are not closely tied, and they may visit Yahoo!Answers only when they need to ask questions. Also, users may not take the questions or answers very seriously. Some askers may even forget to check the answers. If a question is not answered when it is in the first few pages, it may never be answered later. Fig. 2 further shows the cumulative distribution function (CDF) of the number of replies for the replied questions in Yahoo! Answers and all crawled tweets in Twitter. In Yahoo! Answers, nearly 73 percent of all the questions receive more than one response, in contrast to Twitter,



where more than 95 percent tweets receive more than one response. Also, less than 20 percent of all the questions receive more than five answers in Yahoo! Answers, in contrast to Twitter, where more than 80 percent of the tweets receive more than five responses. We suspect that one reason is that the users in Yahoo! Answer do not have social relationship, thus they may not take others' questions as seriously as in social networks. Also, too many questions are posted on the forum every day, so it is not easy for a person who may be able to answer the question to find the questions in the first place. In Twitter, the tweets of users are only pushed to their friends that are connected by their interests and social relationships. Therefore, it is more likely for them to interact with each other. Common interest, close social relationship and frequent contact motivate a user's friends to answer his/her questions. We can also see from the figure that about 30 percent of the tweets with @ receive responses less than five, while 20 percent of the tweets without @ receive responses less than five. This is because the tweets that target to a specific user will not attract discussions from other users, leading to a decreased number of responses. However, as the figure shows the users with @in Twitter are more likely to respond to each other than the Yahoo! Answer. In Yahoo! Answer, nearly 73 percent of all the questions receive more than one response while in Twitter with @, more than 90 percent of the users receive more than one response. This is because if a node A sends a tweet specifically to another node B, node B is very likely to reply the tweet to node A. In Yahoo! Answer, users do not know each other and may not have the same incentives to reply questions as in Twitter.

Fig. 3 shows the CDF of the average response time for the questions that are rated in Yahoo! Answers and tweets in Twitter. We can see that less than 30 percent of the questions rated in Yahoo! Answers receive answers in less than 15 minutes. In contrast, in Twitter, more than 50 percent of the questions are responded within 15 minutes. This result shows the advantages of shorter response time in social-based Q&A compared to Yahoo!Answers. In social networks, as users are connected by their interests and social relationships, they are more willing to interact with each other, resulting in a low response delay. We can also see from the figure that in Twitter-AT, more than 60 percent of the questions are responded within 15 minutes, the percentage of which is higher than tweets without @. This is because when a user is mentioned in a tweet, s(he) is more likely to respond as s(he) knows that the sender is expecting the reply from him/her. Considering the social tie between them, the receiver is very likely to respond the tweet in a short time.



Summary. The figures show that questions posted in online Q&A sites are likely to receive few responses with long delay, though they are a good channel to inquire information. Similar result is also found in [17], which shows that the latency for receiving a satisfying answer in an online Q&A site is high with the average equals 2:52:30 (hh:mm:ss) even when the number of the registered users is very large (290,000). This is because

anonymous users in a Q&A site do not have social relationship between each other, so they may not have incentives to answer others' questions. By leveraging the close social relationship and interest similarity properties among friends in social networks, social-based Q&A systems can help to overcome the inherent problems in online Q&A sites with high response rate and low response delay, since people with similar interests or close social relationship are likely to interact with each other, especially when a user specifically sends a tweet to another user.

III. SYSTEM DESIGN

Question Routing

SOS incorporates an online social network, where nodes connect each other by their social links. As shown in Fig. 4, a registration server is responsible for user registration. Each user has an interest ID, which represents his/her interest. Users sharing more common interests with an asker's question are more likely to be able to answer the question. Also, users who have been willing to answer questions and provided high-quality answers (measured by answer quality) to node i's questions previously are more likely to be willing to answer node i's questions and provide high-quality answers.

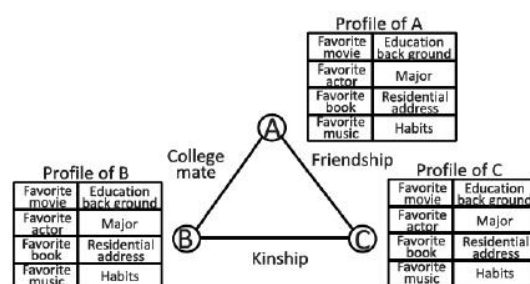


Fig. 5. An example of a node's social network.

IV. QUESTION/USER INTEREST REPRESENTATION

When a user first uses the SOS system, s(he) is required to complete his/her social profile such as interests, professional background and so on. Based on the social information, the registration server recommends friends to the user, and the user then adds friends into his/her friend list. Fig. 5 shows a simple example of social network, where users A, B and C are connected with each other by their social relationships. Each user locally stores her/his own profile and interest ID, and her/his friend list and their interest IDs and answer quality values. Each user calculates his/her own interest ID based on his/her social information and sends it to his/her friends. To calculate interest ID, as shown on the right part of Fig. 6, a node

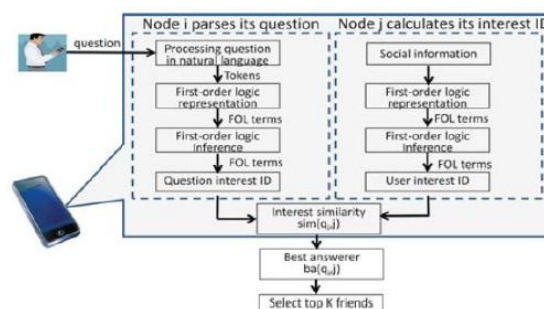


Fig. 6. Answerer selection for forwarding a question in one node.

first derives the first-order logic representation (FOL) [31] from its social information, then conducts firstorder logic inference to infer its interests, from which it decides the interest ID. For instance, an asker may ask a question “Where is the best place to watch the movie Avatar in lemson?”. The corresponding keyword list of this question is resolved to the FOL format [where, place, movie, Avatar, Clemson] after natural language processing. After the FOL inference, the FOL format is changed to [movie(sci-fi), director(James Cameron), place(Clemson)], which is subsequently encoded as a numerical string such as 3200001000.

Similarly, a student in Clemson University who likes to watch sci-fi movie is represented as [movie(sci-fi), career(student), place(clemson)] after the FOL inference and be further encoded as interest ID 3202001001. By comparing the similarity between a question’s ID and its friend’s interest ID, a node can identify its friends that are able to answer questions. More details of the parsing process for a question or for a user is demonstrated in Figs. 7 and 8, respectively. The figures list the three steps in the process: FOL representation, FOL inference, and ID transformation. Below, we introduce the details of the three steps.

V. PRELIMINARY OF THE FIRST-ORDER LOGIC

FOL is a powerful tool to describe objects and their relationships in real life. FOL has basic rules or axioms, which serve as the base of the inference. For example, the FOL for an

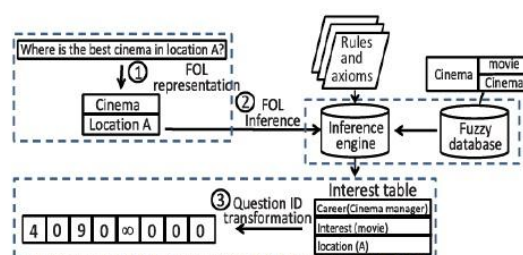


Fig. 7. An example of FOL inference for a question.

axiom in natural language “All computer science (CS) male students who like reading like sci-fi movies” is $\forall x (CS(x) \wedge male(x) \wedge Activity(Reading) \rightarrow like(y))$; where “CS(x)”, “male(x)”, “Activity(Reading)”, and “like(Sci- Fi)” are predicate symbols, and $\wedge$ is connectives symbol. In an FOL representation, connectives symbols (e.g., $\wedge$, $\rightarrow$) and quantifiers logically connect constant symbols, predicate symbols which map from individuals to truth values (e.g., green (Grass)) and function symbols which map individuals to individuals (e.g.,

father-of(Mary)=John). These symbols represent objects (e.g., people, houses, numbers), relations (e.g. red, is inside) and functions (e.g., father of, best friend), respectively.

VI. FIRST-ORDER LOGIC REPRESENTATION

A question or user profile information is always expressed in the natural language. To convert a question or profile information into a format that a computer can understand, we can use part-of-speech tagging or modern NLP techniques to divide the question into a group of related words expressed by words, 2-word phrases, the wh-type (e.g., “what”, “where” or “when”). Then we transform questions into the FOL representation. First, we parse the natural language into token keywords, which are the constant symbols in the FOL representations. The step 1 in Fig. 7 shows an example of FOL representation of the query. The keywords of the question “Where is the best cinema in location A?” are “cinema” and “location A”.

Each node also transforms its social information into the FOL representation. Specifically, a node first represents its profile in the form of name:value pairs such as “movies: Avatar, The Social Network”, “music: Hey, Fig. 6. Answerer selection for forwarding a question in one node. Fig. 7. An example of FOL inference for a question.

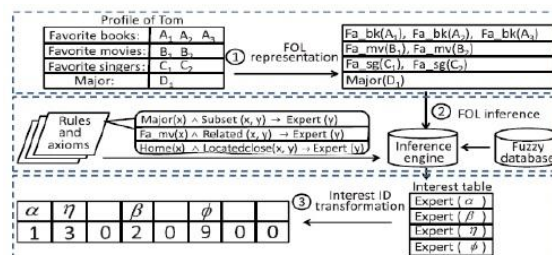


Fig. 8. An example of FOL inference for a user's social information.

VII. FIRST-ORDER LOGIC INFERENCE

As shown in the step 2 in Figs. 7 and 8, the FOL inference component consists of three parts: (1) fuzzy database, (2) rules and axioms, (3) inference engine. The goal of the inference is to identify node interests represented by a numerical string that can accurately represent the capability of a node to answer questions. The fuzzy database is used to store words that have relationships, including subset, alias(x), related, with the information in profiles. For example, $Related(cinema) \frac{1}{4} movie$, $Subset(computer\ science, algorithm)$, $Alias(USA) \frac{1}{4} US$. The rule and axioms provide basic formulas for the inference.

VIII. SIMILARITY VALUE CALCULATION

After users' social information and questions are transformed into numerical strings, the similarity between a user and a question can be calculated based on two parts: interest similarity between the user and question, and answer quality between the question sender and receiver.

IX. INTEREST SIMILARITY CALCULATION

To evaluate the interest similarity of a question of user $i(q_i)$ and a user j , we use a method proposed in. We use ID_{q_i} and ID_j to denote the interest strings of question q_i and user j , respectively. We use $n_{q_i,j}$ to denote the number of interests owned by ID_{q_i} but not by ID_j ; use $l_{q_i,j}$ to denote the number of categories of interest elements owned both by ID_{q_i} and ID_j , and $m_{q_i,j}$ the number of categories of interest elements owned by ID_j but not by ID_{q_i} . Then the interest similarity of question q_i and user j is defined as

$$S_{(q_i,j)} = \frac{l_{(q_i,j)} + 1}{2} \left(\frac{1}{l_{(q_i,j)} + n_{(q_i,j)} + 2} + \frac{1}{l_{(q_i,j)} + m_{(q_i,j)} + 2} \right).$$

The value of $S_{q_i,j}$ ranges in the classical spectrum $[0; 1]$, and it represents the level of likelihood that two strings under comparison are actually similar. If two strings have complete overlapping $n = m = 0$, $S_{q_i,j}$ approaches 1 as the number of common features grows.

X. ANSWER QUALITY CALCULATION

Social closeness directly affects the willingness of people to answer or forward questions. Several recent works have studied how to effectively calculate the social closeness between two users. However, these social closeness value calculation mechanisms are based on the whole social network topology, which are energy consuming. It is even worse when the social network dynamically changes. Therefore, the topology based social closeness calculation methods are not suitable for energy stringent mobile devices in SOS. To reduce the load on the mobile devices, each user in SOS locally manages its first hand information on the answer quality of each of his/her friends. As the performance of the SOS largely depends on the activeness and the knowledge base of the users, user i considers the number of received answers from user j and their associated quality ratings when calculating the answer quality of user j . We call it as the feedback mechanism. Specifically, SOS initially lets users indicate the answer quality value of a newly added friend. For each received answer, an asker can rate the quality of the answer within rating scale $R \in [1; 5]$. The answer quality value is periodically updated based on the number of answers received from friend j during each period T and the associated quality rating ($r \in [1; 5]$). For the k th question sent from node i to node j , if node i receives an answer from node j during T , $x_k \in [1; 5]$; otherwise, $x_k = 0$.

XI. BEST ANSWERER METRIC CALCULATION

Based on Sections 3.4.1 and 3.4.2, for its generated or received question q_i that it cannot answer, node i calculates the best answerer metric of each of its friends. That is,

$$BA_{(q_i,j)} = \beta S_{(q_i,j)} + (1 - \beta) Q_{(i,j)},$$

where $\beta \in [0; 1]$ is a parameter used to adjust the weight of the interest similarity and answer quality. Node i then selects the top K friends that have the highest $BA_{q_i,j}$ values and forwards the question to them. We confine the question forwarding TTL to three since the social trust between two nodes decrease exponentially with distance.



XII. PERFORMANCE EVALUATION

We evaluated the SOS system using our crawled questions from Yahoo! Answers. Since Yahoo! Answers does not have user profile information, we crawled 1;000 users from Facebook to form a social network. We used one user as a seed and used breadth-first search to crawl their personal profile information. We ignored users that did not fill out their profiles. The crawling stopped when 1,000 users were crawled. The users are highly clustered due to the high clustering feature of the social networks. Users' profiles contain their current locations, education backgrounds, hobbies and interests, such as books, movies, music and television programs. This information was parsed and converted to FOL and finally encoded as strings using the method introduced previously. In the experiment, we focused on evaluating the questions related to movies, because most of the Facebook users filled out a large amount of information in the movie section in their profiles.

XIII. CONCLUSION

In this paper, we present the design and implementation of a distributed Social-based mobile Q&A System (SOS). SOS is novel in that it achieves lightweight distributed answerer search, while still enables a node to accurately identify its friends that can answer a question. SOS uses the FOL representation and inference engine to derive the interests of questions, and interests of users based on user social information.

A node considers both its friend's parsed interests and answer quality in determining the friend's similarity value, which measures both the capability and willingness of the friend to answer/forward a question. Compared to the centralized social network based Q&A systems that suffer from traffic congestions and high server bandwidth cost, SOS is a fully distributed system in which each node makes local decision on question forwarding. Compared to broadcasting, SOS generates much less overhead with its limited question forwarding hops. Since each user belongs to several social clusters, by locally selecting most potential answerers, the question is very likely to be forwarded to answerers that can provide answers. The low computation cost makes the system suitable for low-end mobile devices. We conducted extensive trace-driven simulations and implemented the system on iPod Touch/iPhone mobile devices. The results show that SOS can accurately identify answerers that are able to answer questions. Also, SOS earns high user satisfaction ratings on answering both factual and non-factual questions. In the future, we will study the combination of SOS and cloud-based Q&A system. We will also release the application in the App Store and study the Q&A behaviors of users in a larger scale social network.

REFERENCES

- [1.] Google, <http://www.google.com>, 2013.
- [2.] B.M. Evans and E.H. Chi, "An Elaborated Model of Social Search," J. Information Processing and Management, vol. 46, pp. 656-678, 2009.
- [3.] L. Terveen, W. Hill, B. Amento, D. McDonald, and J. Creter, "PHOAKS: A System for Sharing Recommendations," Comm.ACM, vol. 40, pp. 59-62, 1997.



- [4.] L.G. Terveen, P.G. Selfridge, and M.D. Long, "Living Design Memory: Framework, Implementation, Lessons Learned," Human-Computer Interaction, vol. 10, pp. 1-37, 1995.
- [5.] 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.
- [6.] D. Carmel, N. Zwerdling, I. Guy, S. Ofek-Koifman, N. Har'el, 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.
- [7.] H. Kautz, B. Selman, and M. Shah, "Referral Web: Combining Social Networks and Collaborative Filtering," Comm. ACM, vol. 40, pp. 63-65, 1997.
- [8.] D.W. McDonald and M.S. Ackerman, "Expertise Recommender: A Flexible Recommendation System and Architecture," Proc. ACM Conf. Computer Supported Cooperative Work (CSCW '00), 2000.
- [9.] Yahoo answer, <http://answers.yahoo.com>, 2013.
- [10.] Ask.com., <http://www.ask.com>, 2013.
- [11.] 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.
- [12.] 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.
- [13.] 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.
- [14.] 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.
- [15.] 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.
- [16.] 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.
- [17.] J. Raacke and J. Bonds-Raacke, "MySpace and Facebook: Applying the Uses and Gratifications Theory to Exploring Friend-Networking Sites," CyberPsychology & Behavior, vol. 11, p. 169, 2008.
- [18.] 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.
- [19.] 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.



- [24.] C. Manning and H. Schuetze, Foundations of Statistical Natural Language Processing. The MIT Press, June 1999.
- [25.] N. Lao, T. Mitchell, and W.W. Cohen, "Random Walk Inference and Learning in a Large Scale Knowledge Base," Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP '11), 2011.