

## Social Query: A Query Routing System for Twitter

Cleyton Souza, Jonathas Magalhães, Evandro Costa, Joseana Fechine

Laboratory of Artificial Intelligence - LIA

Federal University of Campina Grande

Campina Grande, Brazil

cleyton.caetano.souza@gmail.com, jonathas@copin.ufcg.edu.br, ebc.academico@gmail.com, joseana@dsc.ufcg.edu.br

**Abstract**—Social Query is a new and efficient way to get answers on the social networks. However, the popular method of sharing public questions could be optimized by directing the question to an expert, a process called query routing. In this work, we propose a Social Query System for query routing on Twitter, currently, one of the most popular social networks. The Social Query Systems analyzes the information about the questioner's followers and recommends the most suitable users to answer the questions. The use of the system changes the usual process, working apart of Twitter and allowing questioner and responder exceed the limit of 140 characters. Through a qualitative evaluation, we showed promising results and ideas for improving the system and the recommendation algorithm.

**Keywords**- query routing; social query; expertise finding systems; community question and answering sites; social network; twitter

### I. INTRODUCTION

Social query consists in sharing a problem (in the form of a question) with contacts in social networks and waiting for responses. It is an alternative to search engines and Community Question and Answering sites. Supported by popularity of social networks like Twitter [1] and Facebook [2], social query is a new and efficient way to find information on Web 2.0. The common strategy is to share a public question. However, this way is inefficient because after sharing a public question, there are several disappointing outcomes: receiving many answers (including wrong answers); keep receiving answers after having the problem solved; and never receiving an answer because people able to respond the question did not see it, since social network prioritizes visualization of most recent posts in Timeline, or did not feel an obligation to help [3, 4].

Horowitz and Kamvar [5] associate social query to the process of searching for answers in a village: when an individual in a village has a problem, before he goes to the library, he asks the most capable person that he knows; he does not ask everybody (the village paradigm). The same idea can be applied in the context of social networks. If the question is previously directed to someone (an expert), the social network will ensure that the expert will see the question through its notification system. However, choosing the right person is complicated and by choosing the wrong person, the author of the question may have to wait long for answers, may receive a wrong answer or may never receive

an answer if the expert ignores the question [6]. The addition of Expertise Finding Systems to social networks would optimize the process and consequently enable quick and right answers [5]. The process of identifying an expert and directing questions to that expert is called query routing.

In this work, we present the Social Query System, a tool for query routing on Twitter. The system analyzes the information made available by the questioner's followers and ranks them based on three criteria: knowledge, trust and activity. Then, top users are recommended to the questioner who chooses to whom to direct the question. Details about the model [7] and the algorithm [6] could be obtained in our previous work. Our goal in this paper is to present how the system works and to show promising results of a qualitative evaluation performed with the first version of the Social Query System.

This remainder of this paper is organized as follows: Section II presents related work, Section III presents the usual process of sharing questions on the social networks, Section IV shows or proposal and how it works, Section V describes the results of our qualitative evaluation and Section VI presents our conclusions and future work.

### II. RELATED WORK

Search engines are not always the best way to find information on the Web. Some problems are better solved by people (e.g., high contextualized questions, recommendations request, opinions request, advices request, and social connection request) [5]. An alternative to these types of problems is Community Question and Answering sites like Yahoo! Answers, which consists of online communities where users publish and answer questions voluntarily. After publishing a question, the user waits for answers from other users, who usually are unknown. However, people prefer to pose questions to their close friends in social networks rather than to unknown persons in Community Question and Answering sites [3, 5].

Regarding sharing questions on social networks, Morris et al. [8] presents statistics confirming that sharing questions is a viable method to obtain answers online. In their case study, 93.5% of users had their questions answered. In 90.1% of the cases, responses were provided within one day. However, that case study was conducted with Microsoft employees only, who possibly know each other and also use status messages from chat to ask questions. Paul et al. [4] conducted a similar study, but using only

Twitter users. They conclude that, in this specific context, only a only a small percentage of questions received answers (18.7%) and receiving an answer or not was strongly connected to the number of followers the questioner had. However, questions posted on Twitter are usually answered quickly. In their study, 67% of the responses come within the range of 30 minutes and 95% within the range of ten hours. These facts are due some features of Twitter: when a user posts questions to all followers, only a portion of these followers will view it and a smaller portion will respond. Thus, users with more followers are more likely to get answers, because there is a larger viewing of their messages. And, with respect to agility in receiving a response, this is mainly due to the nature of Twitter as a real-time social network. Actually, these statistics also show that, even as a regular strategy to obtain answers online, the social query process could be improved. We believe these results could be improved by applying query routing: after identifying an expert on the topic of the question and directing the question to him, the answer could come faster and with higher quality. Horowitz and Kamvar [5] establish a correlation between social query and the village paradigm: when people in a village are looking for information, before consulting the libraries, they first ask the most intelligent people they know.

In fact, the query routing problem could be understood as Expertise Finding problem. The query routing consists of a recommendation algorithm (or a technique) that finds an expert present in a group and directs a question to that expert [7]. In [9], it is presented a probabilistic and decentralized model for the question routing problem. This means that there is not an entity that makes all decisions and the routing algorithm works based on the repetition of actions taken in past. Probably, it was the first work about the query routing problem, but no system is proposed and the model is validated using simulated networks.

Davitz et al. [10] present a centralized model implemented in a tool called iLink. In this system, there is a global entity that monitors social network and decides who will receive questions, named super-node. Sometimes, the super-node is also able to offer answers. Other examples of systems are Aardvark [5], a social network that belonged to Google, and Q-Sabe [11], an academic tool for exchange of information focused on education. Both systems consist of Community Question and Answering sites where users could publish questions (questioners) that were routed to other users (responders) and these choose either answering or ignoring the question. Another example of query routing system is AskWho [12], a Facebook plugin that helps in the addition of mentions in the question. Silva et al. [13] proposed SWEETS, an Expertise Finding System, to AMIGOS, an academic online social network of Federal University of Pernambuco. Their recommender system monitors users and suggests experts based on a reading and writing profile.

The studies of Andrade et al. [11] and Horowitz and Kamvar [5] proposed query routing techniques and developed environments where they will work. Our research follows the reverse path. Our system works in a pre-existent context: Twitter, one of most popular social networks currently and that, apparently, will benefit of our technique [4]. In [12], it is presented a Facebook plugin, but AskWho does not use any special technique to match friends, consisting only in a search engine comparing keywords of the question with the profile of friends. Davitz et al. [10] and Silva et al. [13] also propose a system for pre-existent context, but iLink is only available to small communities due the computational effort to monitor the entire social network and SWEETS was not considered useful by AMIGOS users, being rejected by more than half of users [13]. Table I presents a comparison between the cited works.

TABLE I. WORKS COMPARISON.

Reference	Kind of Software	Recommendation Context	Features of Recommendation	Limitation	It uses an Activity criterion	It uses Relationship criterion
Andrade et al. [11]	Desktop.	Everyone who downloads the software.	Information Retrieval method.	Looking for people is not like looking for documents.	Yes (more active users are prioritized).	No.
Davitz et al. [10]	Web.	Users from Small Communities (e.g., forums, blogs).	Probabilistic method.	Only available to small communities.	No.	No.
Horowitz and Kamvar [5]	Web.	Users from Aardvark (a CQA site that belonged to Google)	Probabilistic method.	The project was closed in 2010.	Yes (more active users are prioritized).	Yes (users can optionally maintain a personal network).
Liu [12]	Web.	Friends on Facebook.	Keywords Matching.	Very simple.	No.	No.
Silva et al. [13]	Web.	AMIGOS Users.	Information Retrieval Method.	It was bad evaluated by the users.	No.	No.
Social Query System	Web.	Followers from Twitter.	Multi-criteria decision making method.	The current version only is available to Twitter users previously registered.	Yes (more active users are prioritized).	Yes (based on talks and similarities).

Our system uses three criteria during the recommendation process: knowledge (captured through the vocabulary used by followers), trust (captured through tweets exchanged and friends and followers in common) and activity (consisting in the mean latency time among messages). Besides, the followers are ranked using a multi-criteria decision making process called Weight Product Model (WPM). The use of multi-criteria is the main characteristic of the new generation of recommender systems and the WPM is the better method for the amount of criteria considered by us [14]. While most previous systems consider only the expertise of the candidates we use additional criteria related to the candidate’s availability and the relationship between questioner and expert candidate. The next section describes the usual process of sharing questions on Twitter

### III. THE USUAL WAY OF QUESTION AND ANSWERING ON TWITTER

Twitter is a microblog, a type of blog with some limitation of the content, where users can tweet (post a message) about any topic using 140 characters. In less than three years, Twitter reached such popularity that became the microblog with the largest number of users. Currently, five years after its creation, there are more than 200 million of users and daily registrations are about 460 thousand new users [15]. Another impressive data is the amount of posted messages: in January 2009 two million of tweets (messages) were sent per day; in January 2010 were 65 million and in 2011 were sent 200 million daily tweets [15]. According with the last numbers released by Twitter, currently, there are more than half billion of Twitter accounts and 340 millions of tweets being published per day [16]. This rapid growth has increased the interest of the scientific community on this online social network [1].

On Twitter, users can follow and be followed by other users. In this context, to follow a user means exposing interest in the content published by that user. The account of a user may be public or not, and in order to follow a protected account it is necessary to get the permission of its owner. Initially, the tweets (posts) are visible only to followers, they see in their timelines the tweets of those who they are following. Among the reasons that lead a user to follow another are admiration, friendship and reciprocity. In addition, a user may want to follow another one who posts content which may be considered relevant. Any user is allowed to reference others within a tweet and users can filter their mentions. Because of these features, many users use the microblog as a public chat [1].

When a user publishes a public question on Twitter, if it is not answered quickly, the chances of being visualized and answered in future decrease because the question will fall down in the followers’ timeline. In Fig. 1, it is illustrated the process of publishing a question on Twitter.

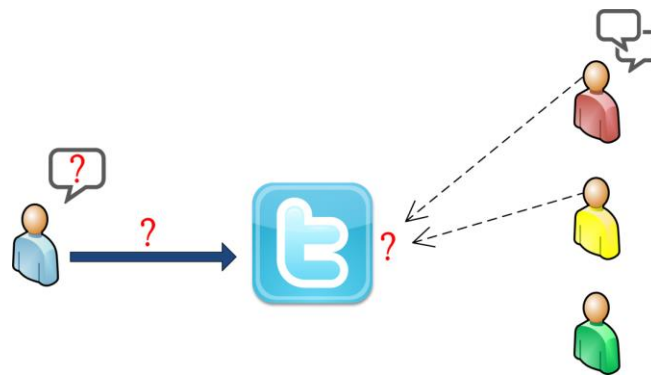


Figure 1. User Tweeting a Question

After publishing a question, probably not all users who follow the questioner (the blue user) will see the question. Among the users that visualize (the yellow user and the red user), only a few will answer (the red user) and there is no guarantee that any of these answers will satisfy the questioner. Some users will not provide an answer because, as the question was posted for all followers, they do not feel an obligation to help. As the time passes by, the chances of the question being viewed and consequently answered in the future decreases, because it will fall positions on the timeline of the questioner’s followers.

When a tweet (question) is previously directed to someone, the probability of it being visualized is much larger, because the mentioned user can filter the messages which mention him/her. In Fig. 2, it is illustrated the same process, but directing the question to a specific user (the green user).

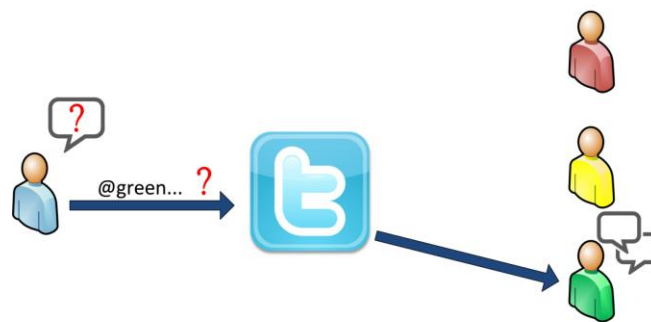


Figure 2. User Tweeting a Direct Question to Someone

When a user (the blue user) mentions another user (the green user), the mentioned user immediately receives an email informing about the message. The mentioned user may wish to disable these notifications, but any user can filter their mentions, as already commented. Given these facts, we believe that direct the question to someone, in practice, guarantees that the message will be visualized by this person, but there is no guarantee that the message will be answered, neither about the quality of the response.

It seems evident that directing a question increases the probability of it being visualized, while the probability of receiving a good answer depends on whom it will be directed to. A query routing model consists in a recommendation

algorithm that examines the information available on the social network to decide who is able to respond the question. In Fig. 3, it is illustrated the query routing process working.

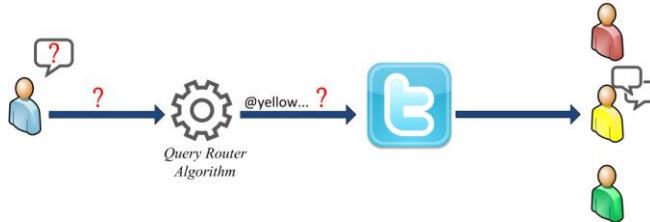


Figure 3. Using Query Routing to Tweet a Question

The question is formulated without a mention. The query routing algorithm (or routing algorithm) analyzes the information about the questioner’s followers and ranks them, according to their ability to offer a right and quick answer. Then, the algorithm adds a mention in the question. In Fig. 3, the question is being directed to only one follower (the yellow user) but, as the algorithm ranks all candidates, it would be possible send the question to the  $n$  followers in best positions.

#### IV. THE SOCIAL QUERY SYSTEM

Our system works outside Twitter. Users access our system to ask and answer. Then, they tweet about their questions and answers. To exemplify how it works, we are considering two basic entities: the questioner (the author of the question) and the responder (a user recommended by our system and chosen by the questioner to receive the question).

Basically, in the query process, the questioner accesses the Social Query System and informs his Twitter account and the question (he also could add keywords to improve the recommendation process). The system analyzes information about the questioner’s followers and recommends the top 5 users. Then, the questioner chooses to whom direct the question and tweets a message informing the chosen users (the responder) about the question. After clicking on the link informed in the tweet, the responder is directed to a page where the question can be answered, or someone else can be recommended to respond. Then, another message is tweeted informing about what was done. Working this way, the system has the advantage that questioners and responders could exceed the limit of 140 characters. Next, we will present in more details how the system works.

Fig. 4 presents the Homepage of our system. In this page, it is explained to the user how the system works and there is a link to the New Question page. While our system is on trial stage, it is only available for users previously registered by us. To require access it is needed to contact one of the authors. The Social Query System is available in two languages: English and Portuguese.

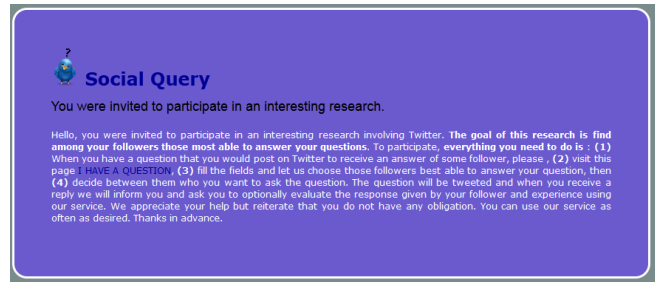


Figure 4. Homepage of the Social Query System

When the user clicks the “I HAVE A QUESTION” link, he is directed to the New Question page. This page is presented in Fig. 5. As already commented, there are three text fields: the questioner’s account name on Twitter, the question and keywords (optionally).

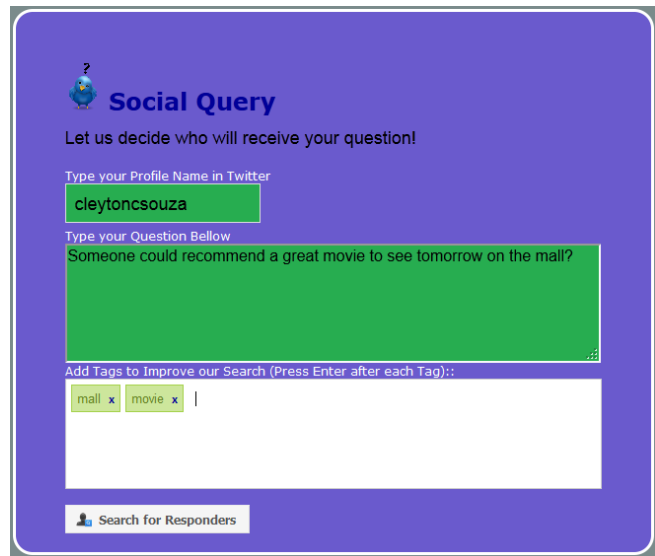


Figure 5. New Question Page of the Social Query System

When the questioner clicks the “Search for Responders” button, the system checks if the required text fields were filled and if the questioner is authorized to use the system. Then, if it is all right, the system analyzes information about the questioner’s followers and ranks them based on three criteria: knowledge (an attribute that relates follower and question), trust (an attribute that relates follower and questioner) and activity (an attribute that relates followers and the social network). The top five candidates are recommended as responders on the Show Recommended List page, as presented on the left side of Fig. 6. If no user is a good recommendation, we suggest the questioner to reformulate the question or change the keywords.

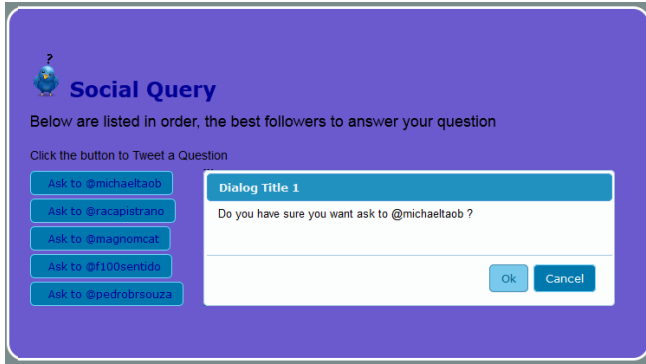


Figure 6. Show Recommendation List Page

Fig. 6 presents the recommendations. When the user clicks someone’s name, a modal appears to confirm the questioner’s intention to send the question to that person. When the questioner confirms, pressing the “OK” button, a new window is opened asking him to tweet the message “@questioner has a question for u @responder. Access <http://short.link> to answer. Thank u”. The tweet has the following information: mention to the questioner, mention to whom was chosen by the questioner and a link to answer the question. After tweet the question, the questioner can repeat the process with the other users that were recommended or just close the page and wait for the answer. Someone who clicks the short link in the tweet will be directed to the New Answer page of the Social Query System, presented in Figure 7. In this page, the responder can give an answer (left button), inform that he/she does not know the answer (mid button) or recommend someone else to answer the question (right button).

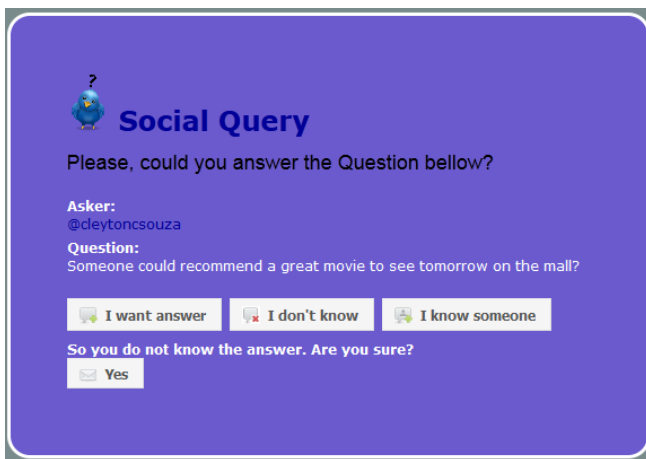


Figure 7. New Answer Page of the Social Query System

If the responder prefers to *suggest someone* else to answer the question, he clicks the “I know someone” button and informs his recommendation in a text field, and then tweets a message alerting the recommended user and the questioner about what this. The person recommended by the responder will click on a link in the tweet and will be

directed to the New Answer Page. A responder who *does not know the answer* will press the “I don’t know” button, and then will be directed to a page to confirm that he/she does not know the answer. The responder will be asked to tweet a message to the questioner informing about it. A responder who *knows the answer* clicks the “I Want Answer” button and is directed to the page presented in Fig. 8.

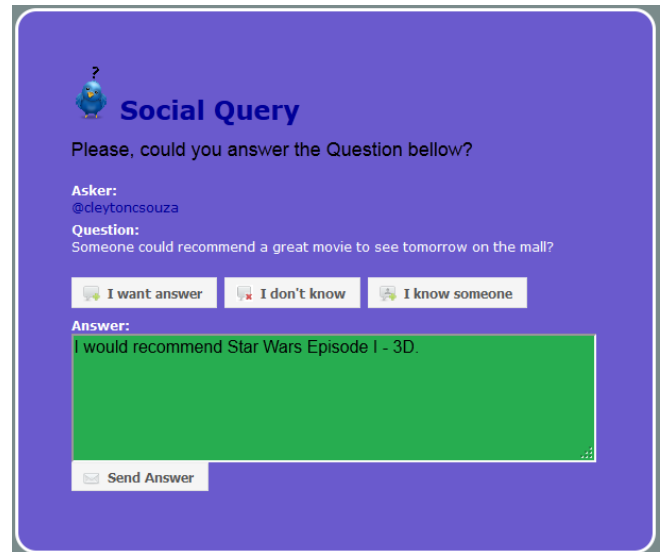


Figure 8. New Answer Page – Answering the Question

The responder writes the answer and clicks the “Send Answer” button. Then, a new window is open asking him/her to tweet the following message informing the questioner that the question was answered: “@questioner, @responder answered your question. Access <http://short.link> to see his answer. Thank u”. After clicking the short link, the questioner is directed to the “New Evaluation” page, where the answer can be seen and have its quality evaluated. This page is presented in Fig. 9.

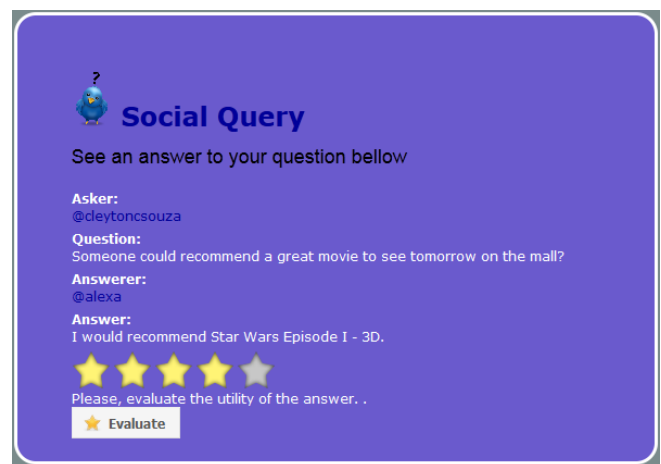


Figure 9. New Evaluation Page

The questioner clicks the “Evaluate” button and the evaluation is saved, ending the Question and Answering process. If the questioner sent the question to other responders, each responder will give an independent answer and the questioner will evaluate them individually.

V. EVALUATION

To validate our purpose, a qualitative evaluation was conducted. The goal of the evaluation is to analyze the opinion of volunteers about the recommendations made by the Social Query System. We used the following questions to perform the recommendations:

- a) *Looking for a new band to listen during weekend, does anyone have an indication?*
- b) *Going to the movie theater after years LOL. What is the best movie in theaters?*

Then, nine volunteers evaluate ten recommendations for each question. Each recommendation was labeled as *good*, *bad* or *neutral*. A recommendation was *good*, if the questioner believes that a relevant answer will be received. A recommendation was *bad*, if the questioner believes that an answer will not be received or an irrelevant answer will be received. A recommendation was *neutral*, if the questioner has no clue about the recommended person. For each question and each volunteer we calculate the percentile of good recommendations and compute the Normalized Discounted Cumulative Gain (nDCG), by considering *good* recommendation with a relevance of “1” and *neutral* and *bad* with “0”.

In Table II, we present the main results of the evaluation.

TABLE II. SUMMARY OF RESULTS

Cases	Amount of Followers	% of good	%of bad	nDCG
Best case for Question “a”	192	50%	10%	0.63
Worst case for Question “a”	129	30%	60%	0.25
Best case for Question “b”	121	60%	0%	0.74
Worst case for Question “b”	68	30%	0%	0.18
<b>Average for Question “a”</b>	<b>110</b>	<b>41%</b>	<b>28%</b>	<b>0.41</b>
<b>Average for Question “b”</b>	<b>110</b>	<b>50%</b>	<b>20%</b>	<b>0.51</b>

According to Table II, the recommendations for Question “b” were better evaluated by volunteers. That happens due the way as the Question “b” was formulated: while Question “a” is composed by few words that express its subject (e.g., Band), Question “b” has more significant words and they are repeated (e.g., Movie and Theaters). This fact affects the knowledge model.

With respect to the worst cases, both users who classified most part of recommendation as *bad* or *neutral* informed that they are followed by many unknowns and

they do not use Twitter to interact with friends but, mostly, to read and share news. With respect to the best cases they informed that they interact often with their followers and that follow back when they are real world friends. All these comments let us to add a new filter to recommendations related to the following back condition. Anyway, the mean performance of the Social Query System was considered very promising, almost half of recommendations were labeled as *good* and the *good* recommendations are better positioned in recommendation list.

VI. CONCLUSION AND FUTURE WORK

The main contribution of this work was the development of a tool to help people finding answers on social networks by improving the conventional social query process. This is a new and interesting research field with many limitations, mainly related to evaluation methods. Datasets provided by Community Question and Answering sites do not have all information available on social networks (e.g., relationship between users, messages directed to specific users, etc.) what hinders the use of a quantitative evaluation method. Qualitative methods have been used to evaluate recommender systems for a long time, but their results are very subjective and hard to compare. To evaluate our tool we used a qualitative evaluation method and we achieved promising results: more than half of recommendation list was considered useful by the volunteers. In addition, based on the opinion of volunteers, we added new features to the Social Query System, for instance, a filter to recommend only followers who are followed by the questioner and the inclusion of a new temporal and non-uniform criterion in recommendation that we call Availability.

As a future work, we are planning a quantitative evaluation of our model. We are creating a dataset and a quantitative method for evaluate our approach and compare with some previous work. In addition, another future work is a new qualitative study about real data collected with the mobile version of the Social Query System. Currently, we still projecting this app, but, probably, such application will be used by hundreds or thousands of users.

ACKNOWLEDGMENT

We want to thank all the volunteers who agreed to participate in this study.

REFERENCES

- [1] B. Huberman, D. Romero and F. Wu, “Social networks that matter: Twitter under the microscope,” *First Monday*, vol. 14, 2009, pp. 1-8.
- [2] Y. Mui and P. Whoriskey, “Facebook passes Google as most popular site on the Internet, two measures show,” *The Washington Post*, 2010.
- [3] M. Morris, J. Teevan and K. Panovich, “Comparison of information seeking using search engines and social networks,” *Proc. 4th International AAAI International Conference on Weblogs and Social Media (ICWSM)*, AAAI Press, 2010, pp. 291-294.
- [4] S. Paul, L. Hong and E. Chi, “Is twitter a good place for asking questions? a characterization study,” *Proc. Fifth AAAI International*

- Conference on Weblogs and Social Media (ICWSM), 2011, pp. 578-581.
- [5] D. Horowitz and S. Kamvar, "The anatomy of a large-scale social search engine," Proc. of the 19th International Conference on World Wide Web (WWW), ACM Press, 2010, pp. 431-440.
- [6] C. Souza, J. Magalhães, E. Costa and J. Fechine, "Predicting potential responders in twitter: a query routing algorithm," Proc. 12th International Conference on Computational Science and Its Applications (ICCSA), Springer, 2012, pp. 714-729.
- [7] C. Souza, J. Magalhães and E. Costa, "A formal model to the routing questions problem in the context of twitter," Proc. IADIS WWW/Internet (ICWI), IADIS Press, 2011, pp. 153-160.
- [8] M. 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. 28th International Conference on Human Factors in Computing Systems (CHI), ACM Press, 2010, pp. 1739-1748.
- [9] A. Banerjee and S. Basu, "A social query model for decentralized search," Proc. 2nd Workshop on Social Network Mining and Analysis, ACM Press, 2008.
- [10] J. Davitz, J. Yu, S. Basu, D. Gutelius and A. Harris, "iLink: search and routing in social networks," Proc. 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM Press, 2007, pp. 931-940.
- [11] J. Andrade, J. Nardi, J. Pessoa and C. Menezes, "Qsabe-um ambiente inteligente para endereçamento de perguntas em uma comunidade virtual de esclarecimento," Proc. First Latin American Web Congress (LA-WEB), IEEE press, 2003, pp.1-7.
- [12] C. Liu, "AskWho: finding potential answerers for status message questions in social networks," unpublished, 2010.
- [13] E. Silva, R. Costa, L. Schmitz and S. Meira, "SWEETS: um sistema de recomendação de especialistas aplicado a uma plataforma de gestão de conhecimento a introdução," Revista de Informática Teórica e Aplicada (RITA), vol. 18, 2011, pp. 153-160.
- [14] E. Triantaphyllou and S. Mann, "An examination of the effectiveness of multi-dimensional decision-making methods: A decision-making paradox," Decision Support Systems, vol. 5, 1989, pp. 303-312.
- [15] Twitter Blog. "#numbers", 2011. Available at: <http://blog.twitter.com/2011/03/numbers.html>. [retrieved: May, 2013]
- [16] Twitter Blog. "Twitter turns six", 2012. Available at: <http://blog.twitter.com/2012/03/twitter-turns-six.html>. [retrieved: May, 2013]