

Adaptive Collaboration Support in Community Question Answering

Marek GRZNÁR*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
marek.grznar@gmail.com*

The current web systems use a variety of search engines to offer theirs users the ability to effectively identify and obtain valuable information. Despite the fact that search engines in the last decade have improved their efficiency and effectiveness, there are still some aspects that search engines cannot meet. Often it happens that search engines fail because the provided results are not sufficient and necessary information is fragmented among several sources and therefore obtaining the full answers can be very challenging.

Web 2.0 offers an opportunity that helps to solve the mentioned problem. This option is to obtain necessary information by asking a community [1]. In the recent years, Community Questions Answering (CQA) more and more come to the fore. Popular CQA systems include Stack Overflow, Yahoo! Answers or Quora, to name a few.

CQA systems have recently become an increasingly popular place not only for knowledge seeking, but also knowledge sharing. The user can insert a question into system and other users by answering this question can share their knowledge (e.g. [4]). Searching for information using wisdom of a crowd can provide the most accurate answers. Thereby CQA systems also use many social features (e.g. following users).

The existing CQA systems, despite of their increasing popularity, fail to answer significant number of questions in required time. In some of current popular CQA systems, only 17,6% of the questions are answered sufficiently [2]. This critical problem was confirmed also in other studies, such as authors in [3] found out that only 11,95% from questions were answered in one day and in two days, there were just 19,95% of answered questions.

In recent years, we can observe an increasing amount of research studies concern with various aspects of the collaboration support in the CQA systems. We focus on question routing. It refers to a recommendations of questions to a user which is a suitable candidate for providing the correct answer. We think that question routing is

* Supervisor: Ivan Srba, Institute of Informatics and Software Engineering

probably the most important part of each CQA system and therefore we believe that question routing has the best chance to solve the problem of low number of answered questions.

Based on results achieved in [5] we decided to focus on non-QA data in our work. By taking into account the results of these researches, we believe that the use of non-QA data can help with the question routing issues (i.e. cold start problem, lurkers).

One of the goal of our method is routing question to users which are not so actively involved in interaction with a system (asking or answering questions). Using of different sources of non-QA data may help to engage users who are new or totally passive in the system (lurkers). It means that we can reduce a cold start problem because users spread a lot of information about themselves. Except of the user non-QA context, we use also historical data from the previous question answering process. We use one of the state-of-the-art method based on LDA to create user profiles from obtained information. Each potential answerer in the system is represented by two profiles (non-QA and QA). For a newly posted question, we also create a question profile. Finally when all profiles are created, we rank all the answerers based on the similarity of their user profiles with the question profile of the question to be routed.

We verify our proposed method on the dataset from existing CQA systems (e.g. Android Enthusiasts). We compare our method to a gold standard which consist of users who actually provided answers on questions in the system.

Extended version was published in Proc. of the 11th Student Research Conference in Informatics and Information Technologies (IIT.SRC 2015), STU Bratislava, 157-162.

Acknowledgement. This work was partially supported by the Cultural and Educational Grant Agency of the Slovak Republic, grant No. 009STU-4/2014.

References

- [1] Liu, Q., Agichtein, E., Dror, G., Maarek, Y., Szpektor, I.: When web search fails, searchers become askers. In: 35th int. ACM SIGIR conf. on Research and development in information retrieval - SIGIR '12, ACM, (2012), pp. 801-810.
- [2] Li, B., King, I.: Routing questions to appropriate answerers in community question answering services. In: Proceedings of the 19th ACM international conference on Information and knowledge management - CIKM '10, ACM Press, (2010), pp. 1585-1588.
- [3] Zhou, T. C., Lyu, M. R., King, I.: A classification-based approach to question routing in community question answering. In: Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion, ACM Press, (2012), pp. 783-790.
- [4] Xu, F., Ji, Z., Wang, B.: Dual role model for question recommendation in community question answering. In: 35th int. ACM SIGIR conf. on Research and development in information retrieval - SIGIR '12, ACM, (2012), pp. 771-779.
- [5] Pan, Y., Luo, L., Chi, C., Liao, Q.: To answer or not. In: Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13, ACM Press, (2013), pp 1253-1262.