Content Recommendation from Archives of Questions Answered in Communities

Viktória LOVASOVÁ*

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

Community Question Answering (CQA) sites, such as Yahoo! Answers or Stack Overflow, have become valuable platforms to create, share, and seek a massive volume of human knowledge. By posting questions for other participants to answer, information seekers can obtain specific answers to their questions.

Archives of questions, provide a potential which has not yet been fully discovered – to recommend solved questions that may be useful for users – (e.g to expand their current knowledge or to provide topics they are interested in). In our work, we focus on this area.

Knowledge sharing in community systems can be divided into two groups: from community and from archive of questions. The both approaches are based on recommendations. In the community-based approach, questions are recommended to the users (experts), who might know the answer – Question Routing [1]. In archive-based approaches, similar questions are searched in the archives and then they are recommended to users – Question Retrieval (Figure 1).

The main objective of CQA systems is to encourage users to answer questions which are not resolved yet. A huge number of solved issues are stored in the archives while these questions remain unused. CQA systems can search in archives for questions using different methods but they do not try to actively get (recommend) the data to the users.

Recommendation of unanswered questions in this area is largely explored. On the other side, it has not been studied how to recommend already resolved questions to users in order to extend their current knowledge or gain insight into problems and solutions in the area. The aim of our work is therefore to join existing approaches: searching questions from archives and routing questions to users and thus recommend answered questions based on users interests derived from their previous activity in the system. Wherein the input would by the user and the archive of issues and output would be an ordered list of questions that he might be interested in.

---

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

Spring 2015 PeWe Workshop, April 11, 2015, pp. 51–52.

More specifically, we propose a method that would recommend questions with their answers to users from archive of questions to improve their knowledge. It takes into account the interests of user: what questions he frequently views, answers, asks, comments and rates. This method should improve the knowledge of the user for the area he is interested in.

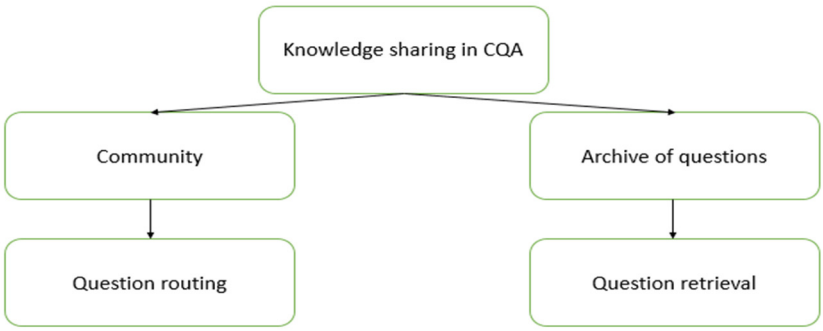

Figure 1. Knowledge sharing in CQA.

In the proposal of the methods, we plan to use Latent Dirichlet allocation (LDA) [2], which is a topic modelling approach that is widely used in CQA to model the topics of questions (e.g. [3]). Then we can find a match between latent topics of questions in the archive (only with right answers) and questions a particular user previously work with to perform the recommendation.

We plan to implement and evaluate the method by the data from the system Stack Overflow - CQA system for professional programmers or in system Askalot, which is used by students at Faculty of Informatics and Information Technologies in Bratislava, which will allow us to perform not only an off-line verification but also a live experiment with students.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2012). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993–1022.

[2] Yang, L., Qiu, M., Gottipati, S., Zhu, F., Jiang, J., Sun, H., & Chen, Z. (2013). CQARank: Jointly Model Topics and Expertise in Community Question Answering.

[3] Zhou, T. C., Lyu, M. R., & King, I. (2012). A classification-based approach to question routing in community question answering. Proceedings of the 21st International Conference Companion on World Wide Web - WWW '12 Companion, 783.