Adaptive Support for Educational Question Answering

Ivan Srba and Mária Bieliková

Institute of Informatics and Software Engineering
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 2, 842 16 Bratislava, Slovakia

{srba,bielik}@fiit.stuba.sk

Abstract. Nowadays, it is possible to access almost unlimited sources of information by ubiquitous information and communication technologies. However, sometimes it is difficult to find required information by standard web search engines. In these situations, Internet users have a possibility to ask their questions in popular community question answering systems (CQA) such as Yahoo! Answers or Stack Overflow. We are interested in an idea to provide similar opportunity also for users in intra-organizational context, and more specifically for students in educational environment. On the basis of analyses of existing approaches in standard CQA systems, we concentrate in our research on the open problem how to adapt these approaches to match the specifics of intra-organizational context and particularly how to recommend newly posted questions to students who are most likely to provide the appropriate answer.

Keywords: Community Question Answering, Knowledge Sharing, Collaborative Learning

1 Community Question Answering

With the development of Web 2.0, popularity of systems based on user generated content such as Wikipedia, YouTube or Flickr is continuously increasing. One type of these systems is Community Question Answering (CQA). CQA is a web service where people can seek information by asking a question and share knowledge by providing answer on the particular question [13]. One kind of CQA systems provides users with possibility to ask any general question without any topic restriction (e.g. Yahoo! Answers or Wiki Answers). On the other hand, there are topic-focused CQA systems dedicated to some specific domain too (i.e. Stack Overflow where users concern with questions related to computer programming).

Compared with the traditional information retrieval systems, CQA systems are based on communities of different Internet users and thus they can apply the best of users' collective wisdom to satisfy knowledge seekers with the most accurate answers. This kind of systems for knowledge sharing is very effective and successful especially when the answer cannot be found easily by standard web search engines

(e.g. when users cannot describe their information needs by keywords or when the information is distributed in different sources).

The significant part of state-of-the-art research in the domain of CQA systems falls into *knowledge sharing* perspective. According to knowledge management theory, knowledge sharing is a process in which knowledge is exchanged among members of particular community. Furthermore, besides knowledge sharing, there is also another interesting possibility how CQA systems can be analyzed. Searching for the answer to the question we are asking is actually informal learning. And thus, we can consider CQA systems also as an innovative kind of *collaborative learning* environments.

Various adaptive approaches have been proposed to support collaboration among users in CQA systems so far. In spite of huge variance of these methods in their purpose, we are not aware of any systematic categorization of approaches employed in CQA systems. Therefore, on the basis of extensive analyses and evaluation of more than 70 research studies performed in the domain of CQA systems, we have proposed views for a complex categorization of approaches supporting CQA process. In the proposed categorization, we divided the approaches according five views, which are related to different phases of question's lifecycle:

1. **Domain entities analyses** concern with classification and evaluation of users, questions and corresponding answers (e.g. identification of users' intent [3] or estimation of question/answer quality [10]). These analyses are independent on question's lifecycle and serve as an important input for following groups of approaches.
2. **Question creation** group of approaches concerns with bridging the gap between web search and CQA, and with the assignment of questions to topics. Existing approaches solve the problems such as the maintenance of topic hierarchy [7] or how to automatically identify question's topic [11].
3. **Question routing** refers to routing newly posted questions to potential answerers. This process is essential for successful CQA system and thus many various approaches have been proposed for this purpose so far. They are based on different estimations of user expertise (e.g. [3], [5]) and user activity (e.g. user authority or availability [3]).
4. **Question answering** is supported by approaches concerned with voting calibration [2] or selection of the best answer. The best answer can be selected by the asker, by the community of answerers or by CQA system itself.
5. **Question search** refers to the retrieval of the best archived question-answer pairs which provide a user with the same information as is required for answering his/her original question. Different methods based on semantic similarity between the searched question and archived questions are proposed (e.g. [8]).

2 Intra-organizational Community Question Answering

Millions of successfully answered questions in CQA systems prove that the popularity of CQA systems rapidly increased in several last years. The rising popularity and growing number of CQA users cause that new opportunities for supporting collaboration constantly emerge. Therefore, CQA systems became the subject of many research

studies. However, in spite of increasing number of studies on CQA systems in recent years, the beneficial effect of CQA systems has not been fully discovered in *intra-organizational context* yet, such as in educational, business or research environments.

Employing the successful and verified concepts of CQA systems inside organizations is a complex task because organization-wide CQA differs from traditional CQA systems in several aspects. Some of them make the knowledge sharing more complicated, e.g. the number of users is significantly lower and thus it is more difficult to route questions to appropriate answerers. On the other hand, there are differences which provide new opportunities for collaboration support, such as the great amount of accessible information about users or a possibility to manage knowledge sharing process by a supervisor, an instructor or a team leader. As the result of these differences, informal learning in standard CQA systems becomes non-formal or even formal in intra-organizational environments. All these aspects make the transition of CQA systems from the Web to intra-organizational context challenging.

However, each type of organizational environment is specific. We focus on educational organizations where students are often struggling with many problems related to their learning process. Therefore, the idea of providing students with the possibility to ask their questions in faculty or university CQA system seems quite promising.

2.1 Research Questions

According to the motivation stated above we formulate the research question of our dissertation project as follows: How an intra-organizational educational CQA system should be designed to take specific organizational conditions into consideration while preserving well established aspects of CQA? Answering of our main research question leads to a set of derived sub-questions:

1. What are the specific conditions in information systems employed in intra-organizational educational environment? Especially how different students' characteristic influence and motivate collaboration during question-answering process?
2. How can we adapt existing approaches to match specifics of intra-organizational environment?
3. Which additional/modified functionalities should be provided by an educational intra-organizational CQA system in comparison with standard CQA systems?

2.2 Research Methodology

We divided our work on the dissertation project into three phases:

1. In the first preliminary phase, we analyzed state of the art in research of CQA systems from the perspective of knowledge sharing as well as collaborative learning. Moreover, we performed extensive study of existing approaches which have been proposed to support collaboration during question answering process so far. In addition, on the basis of the results of our previous work, we examined the students' collaboration while answering well-defined questions prepared by a teacher. The

main aim of this analysis was to determine how students' characteristics influence collaboration and motivate for participation in question answering process.

2. In the current second phase, we build on outputs from the first phase. From the proposed categorization of approaches employed in CQA systems, we recognized question routing as significantly influenced by the transition to intra-organizational environment and thus the proposal of new method for question routing represents the main focus of our research. Besides, the method proposal, we plan to design and implement the prototype of educational CQA system.
3. In the next phase, we plan to evaluate the proposed method by employing the created prototype of CQA system. We plan to start the experiment with the limited number of students who enroll the course Principles of software engineering in bachelor study programme Informatics. Afterwards, in a long-term experiment, we plan to involve wider group of students moving to faculty-wide environment with possibility to ask questions related to various topics across several courses. Finally, we will conclude achieved results and derive implications at generic level, i.e. suggestions which can be applied for intra-organizational environments in general.

2.3 Preliminary Results

In our previous work [12], we proposed a group formation method for automatic creation of short-term dynamic study groups for collaborative solving of various questions. The proposed method is based on the optimization approach named Group Technology [9]. It takes into consideration different students' personal and collaborative characteristics. We implemented the proposed method as a part of collaborative environment PopCorm (Popular Collaborative Platform) which is integrated with the system dedicated to individual learning named Adaptive Learning Framework ALEF [1]. PopCorm provides for task solving four collaborative tools: a text editor, a graphical editor, a categorizer, and a semi-structured discussion (see Figure 1).

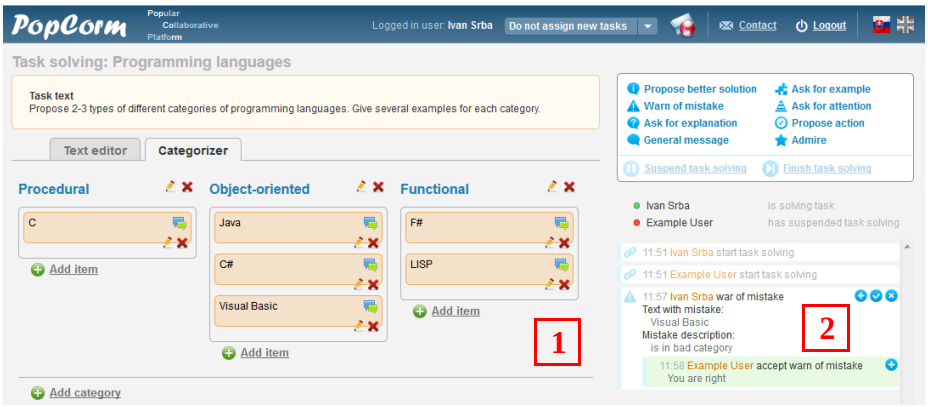

Fig. 1. Screenshot from the collaboration platform PopCorm; the categorizer tool (1) and semi-structured discussion (2) is displayed. The categorizer is a special tool developed for solving tasks the solution of which consists of one or more lists. Semi-structured discussion provides 18 different types of messages which allow us to automatically identify student's activities.

PopCorm served for evaluation of the proposed method in a long-term experiment in which students solved the collaborative questions prepared by a teacher. The data acquired during the experiment were used in analyses how students' characteristics influence their collaboration. More specifically, we focused on students' study results, personal characteristics determined from Big Five questionnaires and collaborative characteristics proposed by McManus and Aiken [6]. In our analysis we consider data of totally 129 students who actively participated in ALEF and PopCorm during one semester of the course (12 weeks). Students were repeatedly assigned to 254 groups in PopCorm in which totally 3,763 activities were recorded. In addition, more than 55,400 interactions with learning objects were recorded in ALEF. The results of analyses provide us with important findings how students' characteristics influence overall activity, correctness of created solutions, task and time management, students' self-regulation and motivation, and, finally, evaluation and providing feedback. These findings represent important input for following phases of our research work.

2.4 Adaptive Question Routing

In our dissertation project, we focus on question routing which is probably the most important part of the proposed educational CQA system. It refers to recommendation of potential answerers who are most likely to provide appropriate answer on the newly posted question. We propose a method for question routing on the basis of existing methods for question routing while taking the specifics of intra-organizational educational domain into consideration. The proposed method considers three groups of answerer's characteristics which are important to provide an appropriate answer.

User expertise. Students' expertise can be derived from previous activities (e.g. asking the question, providing the answer, voting for the best answer). In addition, we can take advantage of information about students which are available specifically in educational environment (e.g. study results or enrolled subjects).

User activity. Besides the overall students' activity in a CQA system, we propose to consider also students' availability (estimation that a potential answerer will login to the system in the time dedicated to answering the question) and authority.

User motivation. On the basis of performed analyses, we found out that students perceive reciprocity as an important motivation factor. Therefore, we suggest recommending questions with considering the symmetry in knowledge each student provides and receives from the CQA system.

3 Conclusion

Community question answering systems has become recently the subject of many research studies. However, we suppose that their potential for supporting of organizational knowledge sharing and collaborative learning is only to be discovered. In our dissertation project, we concern with the transition of CQA from the Web to educational environment. More specifically, we recognized question routing to potential answerers as unique opportunity how information technologies can support students'

collaboration. The proposed method, which will be employed in innovative educational CQA system, will provide students new opportunities how to solve their own questions related to learning process at our faculty.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0675/11 and by the Slovak Research and Development Agency under the contract No. APVV-0208-10.

References

1. Bieliková, M., Šimko, M., Barla, M., Tvarožek, J., Labaj, M., Móro, R., Srba, I., Ševcech, J.: ALEF: from Application to Platform for Adaptive Collaborative Learning. In Recommender Systems for Technology Enhanced Learning: Research Trends & Applications, Springer, Heidelberg, (to appear, 2013)
2. Chen, B.-C., Dasgupta, A., Wang, X., Yang, J.: Vote calibration in community question-answering systems. In Proc. of the 35th int. ACM conf. on Research and development in information retrieval - SIGIR'12. New York, USA: ACM Press, pp. 781-790 (2012)
3. Chen, L., Zhang, D., Mark, L.: Understanding user intent in community question answering. In Proceedings of the 21st international conference companion on World Wide Web - WWW'12 Companion, pp. 823–828 (2012)
4. Li, B., King, I.: Routing questions to appropriate answerers in community question answering services. In Proc. of the 19th ACM int. conf. on Information and knowledge management - CIKM'10. New York, USA: ACM Press, pp. 1585-1588 (2010)
5. Li, B., King, I., Lyu, M.R.: Question routing in community question answering. In Proc. of the 20th ACM int. conf. on Information and knowledge management - CIKM'11. New York, USA: ACM Press, pp. 2041-2044 (2011)
6. McManus, M.M., Aiken, R.M.: Monitoring computer-based collaborative problem solving. Journal of Artificial Intelligence in Education 6 (4), 307–336, (1995)
7. Miao, Y., Li, C., Tang, J.,Zhao, L.: Identifying new categories in community question answering archives. In Proc. of the 19th ACM int. conf. on Information and knowledge management - CIKM'10. New York, USA: ACM Press, pp. 1673-1676 (2010)
8. Pera, M.S., Ng, Y.-K.: A community question-answering refinement system. In Proc. of the 22nd ACM conf. on Hypertext and hypermedia - HT'11. New York, USA: ACM Press, pp. 251-260 (2011)
9. Selim, H.M. et al.: Cell formation in group tech.: review, evaluation and directions for future research. Comput. Ind. Eng. 34, 1 (January 1998), pp. 3-20 (1998)
10. Shah, C., Pomerantz, J.: Evaluating and predicting answer quality in community QA. In Proc. of the 33rd int. ACM conf. on Research and development in information retrieval - SIGIR'10. New York, USA: ACM Press, pp. 411-418 (2010)
11. Singh, A., Visweswariah, K: CQC: classifying questions in CQA websites. In Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM'11, pp. 2033–2036 (2011)
12. Srba, I., Bieliková, M.: Encouragement of Collaborative Learning Based on Dynamic Groups. In Proc. of the 7th European Conf. of Technology Enhanced Learning - EC-TEL'12. LNCS, vol. 7563, 432–437. Springer, Heidelberg, (2012)
13. Xu, F., Ji, Z.,Wang, B.: Dual role model for question recommendation in community question answering. In Proc. of the 35th int. ACM conf. on Research and development in information retrieval - SIGIR'12. New York, USA: ACM Press, pp. 771–779 (2012)