Formulating Effective Questions for Community-based Question Answering

Saori Suzuki  
College of Knowledge and Library Sciences  
University of Tsukuba  
s0711607@u.tsukuba.ac.jp

Shin’ichi Nakayama  
Graduate School of Library, Information and Media Studies  
University of Tsukuba  
nakayama@slis.tsukuba.ac.jp

Hideo Joho  
Graduate School of Library, Information and Media Studies  
University of Tsukuba  
hideo@slis.tsukuba.ac.jp

ABSTRACT

Community-based Question Answering (CQA) services have become a major venue for people’s information seeking on the Web. However, many studies on CQA have focused on the prediction of the best answers for a given question. This paper looks into the formulation of effective questions in the context of CQA. In particular, we looked at effect of contextual factors appended to a basic question on the performance of submitted answers. This study analysed a total of 930 answers returned in response to 266 questions that were formulated by 46 participants. The results show that adding a questionnaire’s personal and social attribute to the question helped improve the perceptions of answers both in information seeking questions and opinion seeking questions.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Query formulation; H.3.5 [On-line Information Services]: Web-based services

General Terms

Experimentation, Human Factors

Keywords

Question formulation, context, CQA

1. INTRODUCTION

Taylor’s 1968 model is perhaps one of the earliest work to point out that people’s expression of their information needs can be compromised by a given environment [9]. Short queries frequently submitted to search engines show that Taylor’s model is also evident on the Web. One goal of Information Retrieval (IR) research is to retrieve relevant information in response to such compromised expressions of people’s information needs. In this regard, community-based Question Answering (CQA) sites are an interesting venue for information seeking on the Web, since the expression of information needs does not have to be compromised as much as typical search engine queries. People use natural language to formulate a query and it is common to use multiple sentences to elaborate the information needs in CQA services. Meanwhile, Yahoo! Answers alone are reported to have 200 million visitors worldwide in 2009\(^1\), making CQA an important research area in IR. However, existing studies on CQA have been focused on the prediction of the best answers for a given question [1, 2, 8]. Investigations on how to formulate effective questions have been limited.

To address the problem, this paper examined effect of contextual factors in question formulation in the context of CQA. Appending a contextual factor has been shown to be effective in query formulation [7] in IR. Therefore, we hypothesised that appending a contextual factor to a basic question could improve the performance of submitted answers. To examine the hypothesis, we asked participants to formulate genuine questions with different contextual factors, submitted them to a CQA site, and assessed the answers responded to the questions.

2. METHOD

2.1 Contextual Factors

We were also interested in how effect of contextual factors differed across the types of questions. We decided to look

at two types of questions based on the investigations by [1, 5, 4]. One type was called information seeking questions while another type was called opinion seeking questions. As the name suggests, the former type was generally looking for information such as facts and methods, while the latter type was looking for opinions and suggestions.

2.3 User Study

46 participants (27 females and 19 males) were recruited for the study. The majority of participants were the undergraduate and postgraduate students of our institution, with varied backgrounds (e.g., Management, Physics, Mathematics, Linguistics, Engineering, Medicine, Cultural Study, Media Study, Computer Science, Information Science, Chemistry, Art, Robotics, Biology, etc).

Participants were asked to formulate six genuine questions. Three of them were information seeking questions while others were opinion seeking questions. Participants were asked to include one of the five contextual factors in their question formulation. The remaining question was formulated without any contextual factor. We rotated the order of contextual factors to append across participants.

The formulated questions were then submitted to one of the major CQA sites in Japan by the experimenter, and recorded any answers given to the questions for 7 days. Participants were called again and asked to assess the answers from eight perspectives. They were Overall satisfaction, Topicality, Novelty, Situational relevance, Attitude, Trustiness, Promptness, and Length. A 5-point Likert scale was used to capture participants’ assessment of answers.

266 questions were formulated and 930 answers were collected and assessed by participants.

3. MAIN RESULTS

We examined effect of five contextual factors used in question formulation on the perception of answers, by comparing to the questions formulated without using contextual factors. This section reports only on a subset of analysis we conducted based on the perception of answers, by comparing to the questions formulated without using contextual factors. This section reports only on a subset of analysis we conducted based on the perception of answers.

First, we examined the number of answers submitted to the questions. Although opinion seeking questions tended to receive more answers than information seeking questions (3.9 vs. 2.9, \( p \leq .005 \) by T-test), there was no significant effect of contextual factors on the number of answers.

Second, we examined participants’ perceptions on the answers. The result pattern differed across the question types. In information seeking questions, the questions with Attribute and Thought had a higher average score of Attitude of answers (Attribute: \( p \leq .001 \), Thought: \( p \leq .003 \), by T-test). In opinion seeking questions, the questions with Attribute had a higher average score of Novelty of answers (\( p \leq .003 \), by T-test). On the other hand, the factors such as Task and Situation had a negative impact on the perceptions of their answers.

Overall, we gained some evidence suggesting that appending a contextual factor is a promising way to formulate effective questions in a CQA site. At the same time, the results suggest that some contextual factors should be avoided in a CQA site. Also, our results suggest that effective contextual factors can differ across the question types. A comprehensive description of this study and analysis of the results will be reported elsewhere.

4. CONCLUSION

This paper explored effect of contextual factors as a means of formulating effective questions in the context of community-based question answering services. Genuine questions were formulated by our participants and some questions were presented with a contextual factor and others without. Based on the analysis of over 900 answers submitted to participants’ questions, we gained evidence to support the following findings. First, including a contextual factor to the question can improve a questioner’s assessment of the answers. Second, affective contextual factors can differ across the question type (information seeking and opinion seeking type). These findings can be a good starting point in guiding how beginners use CQA services to formulate their questions.

5. ACKNOWLEDGEMENTS

The authors thank to Yahoo JAPAN Corporation and National Institute of Informatics who gave us access to Yahoo! Chiebukuro Data (Version 1).

6. REFERENCES


