Cross-Platform Question Routing for Better Question Answering

Mossaab Bagdouri
Department of Computer Science
University of Maryland
College Park, MD, USA
mossaab@umd.edu

ABSTRACT

The last two decades have made the Internet a major source for knowledge seekers to fulfill their needs. Several tools and platforms were developed to find answers to one's questions such as search engines and online encyclopedias. The wide adoption of social media pushed the possibilities even further by giving people the opportunity to stimulate the generation of answers that are not already present on the Internet. Some of these are primarily community question answering (CQA) services, while the others have a more general audience but can also be used to ask and answer questions.

Assessing which platform is best suitable for certain information need of a particular user, and subject to some specific constraints, is a problem that has seen little attention (e.g. [3]). Hence, we consider *cross-platform question routing* to automatically suggest where an input question should be asked or redirected to. Several stages are required to address this problem. First, we need to analyze the differences between the questions asked on different platforms. In fact, on a search engine, a user would provide the least context to the query. On a CQA website, she would summarize the question in a title and provide more context in a description. On the other hand, the (online) identity of the asker in a microblogging service is revealed to her friends, leading to a higher exposure of context. For this stage, we gather questions from different platforms to train a classifier on the distribution of question types [4] over these platforms. Services that are not restricted to questions and answers (e.g., Twitter) need special mining to find content that contains questions. We will extend the work of [2] to find out which posts are answer seeking questions.

Second, we need to characterize the users, as not all demographics have the same needs. For instance, the types of information seeked by a journalist are arguably different from those of a teenager. Thus, we build some classifiers that differentiate users based on the questions they might ask. We apply some algorithms that we developed to retrieve the Twitter accounts of Arab journalists, and adapt the taxonomy of the more general question types to this group of professionals.

Routing questions needs new evaluation measures that take in consideration various aspects beyond the traditional dimension of relevance. For instance, for a user who is seeking opinions of "ordinary" people about non-factoid questions, informality of the answers is a constraint that might be added to the evaluation [1]. The cost of wrong answers is an additional dimension to the evaluation that is often traded-off with response time. For example, a journalist cannot accept an unverified report and might spend a long time verifying it. On another hand, someone who is trying to get recommendations for a nearby restaurant might trade the correctness for a timely answer. Thus, we need to model the likelihood of whether and when a question will get a response, if a response contains an answer to that question and to which extent that answer is correct and satisfies the expectations of the asker.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

Keywords

Question Answering; Question Routing; Social Media

REFERENCES

[1] M. Bagdouri, D. W. Oard, and V. Castelli. CLIR for informal content in Arabic forum posts. CIKM, 2014.

[2] M. Hasanain, T. Elsayed, and W. Magdy. Identification of answer-seeking questions in Arabic microblogs. CIKM, 2014.

[3] A. Oeldorf-Hirsch, B. Hecht, M. R. Morris, J. Teevan, and D. Gergle. To search or to ask: The routing of information needs between traditional search engines and social networks. CSCW, 2014.

[4] Z. Zhao and Q. Mei. Questions about questions: An empirical analysis of information needs on Twitter. WWW, 2013.