

E-WISE: An Expertise-Driven Recommendation Platform for Web Question Answering Systems

Jie Yang^(✉), Alessandro Bozzon, and Geert-Jan Houben

Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
{j.yang-3,a.bozzon,g.j.p.m.houben}@tudelft.nl

Abstract. This demo presents E-WISE, an expertise-driven recommendation platform built upon Web Question Answering (QA) systems to assist askers in question-answering process. Despite that crowdsourcing knowledge (e.g., on-line question-answering) is becoming increasingly important, it remains a big challenge to accelerate its process. E-WISE blends the recently developed methods for knowledge crowdsourcing acceleration, including 1) an edit suggestion component to improve question quality; 2) a question routing component that suggests a list of ranked answerers. Both components are automatic, and meanwhile enable a human controlling part: askers can make their decisions in selecting the right edits/answerers among the suggested ones, which guarantees the effectiveness of the suggesting components and provides feedback to the suggesting methods. E-WISE will be demonstrated through a case study on Stack Overflow – a popular QA systems, to exemplify its functions and potential in on-line knowledge creation.

Keywords: Crowdsourcing · Knowledge creation · Question answering · Edit suggestion · Question routing

1 Introduction

Crowdsourcing knowledge has recently been proposed as a promising topic, which solicits the rich set of knowledge-related features of humans (e.g. expertise, skills, experience and opinions), to solve complex and cognitive intensive tasks. The growing popularity of knowledge crowdsourcing systems, such as on-line QA platform (e.g. Stack Exchange¹ and Yahoo! Answers²), content curation systems (e.g. Reddit³), diverse on-line forums and wiki's, have demonstrated the great potential of knowledge crowdsourcing in on-line knowledge creation.

Although crowdsourcing has drawn much attention from diverse domains, knowledge crowdsourcing has been much less studied until the most recent [1]. In the case of on-line knowledge creation, typical tasks such as building wiki's or QA systems are performed on an ad-hoc basis. The non-existence of a comprehensive

¹ <http://stackexchange.com/>

² <http://answers.yahoo.com/>

³ <http://www.reddit.com>

theory of knowledge crowdsourcing and a corresponding systematic engineering approach have induced big challenges: knowledge is usually created inefficiently. For instance in Stack Overflow⁴, which is one of the most active QA systems on the web, there are more than 30% questions do not have any up-voted answer; furthermore, the average time for a question to receive an accepted answer is at a magnitude of days.

Recent advances in the research community have illuminated the possibility of developing methods and tools to support accelerating knowledge creation process. Question routing, for example, that actively routes the newly posted question to potential answerers, has been proposed as a promising method to accelerate QA process. Another direction in this research is to improve question quality. Poorly formulated questions (around 40% in Stack Overflow) needs to be edited after they were initially posted, as responses to comments or answers. Since the interactions between askers and other users lengthens the time for a question to obtain a good answer, it has been proposed to detect poorly formulated question and automatically suggest edit to improve question quality such that the QA process can be accelerated [2,3].

This demo introduces E-WISE, an expertise-driven recommendation platform that provides a continuous and complete line of assistance to askers in web QA systems. E-WISE supports related question retrieval, automatic edit suggestion, user modeling and question routing. The platform implements and integrates the state-of-the-art methods [3,4], thus contributing to the creation of a comprehensive theory of knowledge crowdsourcing acceleration. Another important feature of E-WISE is that askers are allowed to make their own decisions, with the assistance of the automatic recommending methods, i.e. edit and answerer suggestion. Therefore, E-WISE acts as a social machine [5] that fuses the complementary power of humans and computers for knowledge creation.

E-WISE will be demonstrated with an instantiation integrated with Stack Overflow. The attendees will be given the opportunity to interact lively with Stack Overflow through E-WISE, such as searching and posting questions on Stack Overflow, with the assistance of edit and answerer suggestion by E-WISE.

2 The E-WISE Architecture

Fig. 1 depicts the architecture of E-WISE. It comprises two phases, namely, off-line training and on-line searching and recommendation. Due to efficiency reason, the training phase trains recommenders off-line based on historical QA data imported from web QA systems; while on-line phase recommends edits and answerers to askers in real-time using the trained recommenders. Askers can interact lively with the QA systems through E-WISE, such that the actions taken by the askers can be written back to the QA systems in real-time.

E-WISE consists of three components: related question searcher, edit suggester and question router. Standard information retrieving method is adopted for related question retrieval. While for edit suggestion and question routing, E-WISE implements our novel methods described in [3,4].

⁴ <http://stackoverflow.com/>

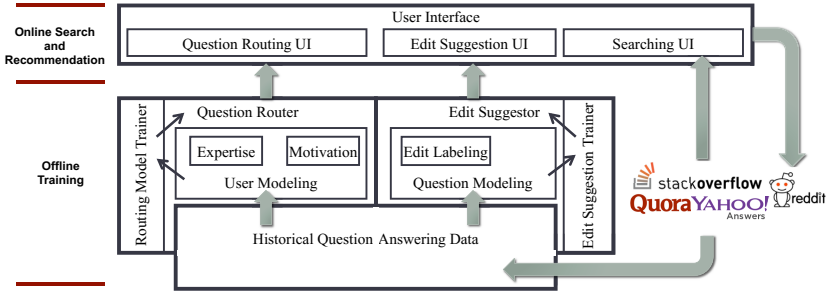


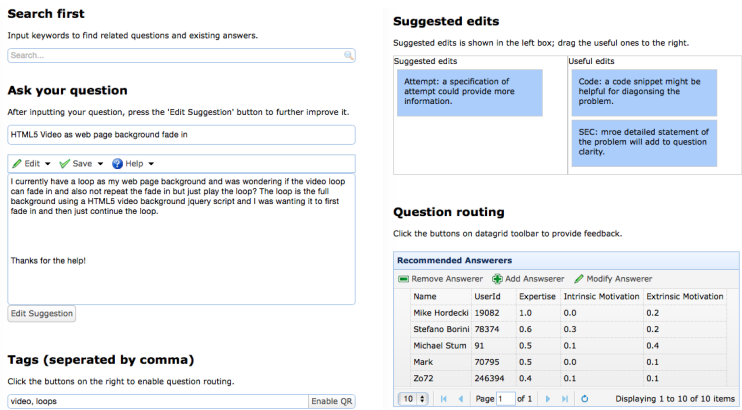
Fig. 1. System Architecture of E-WISE

Edit Suggester detects poorly formulated questions and suggests edits to improve question quality. Essentially, this component consists of two type of classifiers, one for edit/non-edit classification, the other one for edit type classification. The trainer takes (compressed) TFIDF representation of labeled questions as input, then trains the classifiers off-line for on-line edit suggestion.

After a question is classified as one that needs an edit, it will then be classified into some of pre-defined edit types. For instance the edit suggester of E-WISE for Stack Overflow defines the following edit types: 1) Code, which indicates that the question needs to have a code snippet; 2) SEC, indicating that the question needs clarification of either problem *S*tatement, *E*xample or *C*ontext; 3) Attempt, which indicates that the asker needs to specify attempts that he has tried. The asker will receive a list of suggestions, and decide to reformulate their questions correspondingly. The asker is further provided the option to provide feedback to the system, by indicating whether or not the suggested edits are helpful.

Question Router. After the asker input a question, he can enable the question router to actively route his question to potential answerers. Question router contains a user modeling part that models the topic-wise expertise and motivation of answerers, which are considered as the most important answering properties in QA systems [4]. The basic assumption is that questions of different topics require different combinations of user roles: questions relating to general concepts may require motivated answerers engaged in discussion to ultimately generate a best answer, while some other questions relating to specific topic only need one expert user to directly provide the right answer.

Given a newly posted question, question router first matches the question to answerers according to the topics, then ranks the answerers according to expertise and motivation. Question router trains a learning-to-rank model off-line for on-line answerer ranking. It takes as input the user models (expertise, motivation) and historical QA data, and output the model for answerers suggestion. The asker will be shown a ranked list of answerers, together with their user profiles as the interpretation to help the asker in making decision of whom to route the question to. The asker is further provided with the option of removing or adding answerers.



(a) User input interface. (b) Result and feedback interface.

Fig. 2. User interface

3 Demo Scenario: Stack Overflow

We demonstrate E-WISE with a case study on Stack Overflow. For performance reason, E-WISE takes the latest historical data for off-line training edit suggestor and question router; with the trained recommenders, E-WISE is seamlessly integrated with the live QA system, such that the askers are allowed to directly interact with Stack Overflow through E-WISE, e.g. searching related questions, post questions and inspect the profile of the suggested users.

Fig. 2a is the user interface in which askers can search, post questions, call edit suggestor and question router. Fig. 2b is the user interface that displays suggested edits and answerers. Askers can interact with the results to provide feedback for both edit suggestion and question routing: suggested edits that are useful on the left box can be dragged to the right box to indicate the usefulness; answerers can be removed from or added to the routing list.

References

1. Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. In: CSCW 2013, pp. 1301–1318. ACM, New York (2013)
2. Ravi, S., Pang, B., Rastogi, V., Kumar, R.: Great question! question quality in community Q&A. In: ICWSM 2014, pp. 426–435. AAAI, Palo Alto (2014)
3. Yang, J., Hauff, C., Bozzon, A., Houben, G.J.: Asking the right question in collaborative Q&A systems. In: Hypertext 2014, pp. 179–189. ACM, New York (2014)
4. Yang, J., Tao, K., Bozzon, A., Houben, G.-J.: Sparrows and owls: characterisation of expert behaviour in stackoverflow. In: Dimitrova, V., Kuflik, T., Chin, D., Ricci, F., Dolog, P., Houben, G.-J. (eds.) UMAP 2014. LNCS, vol. 8538, pp. 266–277. Springer, Heidelberg (2014)
5. Hendler, J., Berners-Lee, T.: From the Semantic Web to Social Machines: a Research Challenge for AI on the World Wide Web. Artificial Intelligence 174(2), 156–161 (2010)