

Improving Search and Recommendation by Asking Clarifying Questions

Jie Zou

University of Amsterdam

j.zou@uva.nl

Abstract

Recently, there has been renewed interest in interactive search and recommender systems through the development of systems that ask clarifying questions (CQs) to users to better understand their information needs. Although the significance and effectiveness of CQs have been recognized, research on modelling CQs to improve search and recommender systems is limited, compared to traditional search and recommender systems. Also, understanding users' willingness or ability to answer these CQs, and the extent to which CQs impact users' behavior and their ability to identify relevant information remains relatively unexplored.

The work in this thesis provides a series of models and online user studies for search and recommender systems that construct and use CQs. In the first part of the thesis, we focus on CQ-based search and recommendation algorithms. We first propose a novel CQ-based document search algorithm, Sequential Bayesian Search based method for Technology-Assisted Review (SBSTAR) and its extension, to efficiently retrieve the last few, but significant, relevant documents [Zou et al., 2018; Zou and Kanoulas, 2020]. The algorithm sequentially selects and directly asks CQs to users about the presence or absence of an entity in the missing relevant documents. Next, we propose a novel CQ-based product search algorithm, Question-based Sequential Bayesian Product Search (QSBPS), to effectively locate products that users are looking for [Zou and Kanoulas, 2019]. The model is based on a duet learning framework that learns product relevance and the reward of the potential CQ to be asked to the user. Furthermore, we propose a novel CQ-based recommendation algorithm, Question-based recommendation (Qrec), to assist users to find items interactively [Zou et al., 2020a]. The model is first trained offline by a novel matrix factorization algorithm, and then iteratively updates the user and item latent factors online based on the user's answers. We experimentally test the performance of the proposed models and demonstrate that they outperform state-of-the-art baselines.

In the second part of the thesis, we focus on the evaluation of CQ-based search and recommendation. We first conduct an online user study by deploying a CQ-based system in the domain of online retail, to understand to what extent users can answer CQs. We explore the user willingness, ability, and user perception to provide answers to CQs [Zou et al., 2020b]. Next, we conduct a large user study on web search to understand the impact of CQ interactions on user search behavior and satisfaction, and how users interact with different levels of quality of CQs under different user perceptions and conditions.

Awarded by: University of Amsterdam, Amsterdam, The Netherlands **on** 6 September 2021.

Supervised by: Evangelos Kanoulas, Maarten de Rijke.

Available at: <https://hdl.handle.net/11245.1/3c46aa78-98af-4997-aef7-44db7ab2fa3d>.

Selected Publications

Jie Zou and Evangelos Kanoulas. Learning to ask: Question-based sequential bayesian product search. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, pages 369–378, 2019.

Jie Zou and Evangelos Kanoulas. Towards question-based high-recall information retrieval: Locating the last few relevant documents for technology-assisted reviews. *ACM Trans. Inf. Syst.*, 38(3), May 2020.

Jie Zou, Dan Li, and Evangelos Kanoulas. Technology assisted reviews: Finding the last few relevant documents by asking yes/no questions to reviewers. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '18, pages 949–952, 2018.

Jie Zou, Yifan Chen, and Evangelos Kanoulas. Towards question-based recommender systems. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '20, pages 881–890, 2020a.

Jie Zou, Evangelos Kanoulas, and Yiqun Liu. An empirical study on clarifying question-based systems. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 2361–2364, 2020b.