Toward a Fairer Information Retrieval System

Ruoyuan Gao
Rutgers University
gaoruoyuan@gmail.com

Abstract

With the increasing popularity and social influence of information retrieval (IR) systems, various studies have raised concerns on the presence of bias in IR and the social responsibilities of IR systems. Techniques for addressing these issues can be classified into *pre-processing*, *in-processing* and *post-processing*. Pre-processing reduces bias in the data that is fed into machine learning models. In-processing encodes fairness constraints as a part of the objective function or learning process. Post-processing operates as a top layer over the trained model to reduce the presentation bias exposed to users. This dissertation explored ways to bring the pre-processing and post-processing approaches, together with the fairness-aware evaluation metrics, into a unified framework as an attempt to break the vicious cycle of bias and improve fairness in IR.

We first investigated the existing bias presented in search engine results. Specifically, we focused on the top-k fairness ranking in terms of statistical parity fairness and disparate impact fairness definitions. With Google search and a general purposed text cluster as a lens, we explored several topical diversity fairness ranking strategies to understand the relationship between relevance and fairness in search results. Our experimental results showed that different fairness ranking strategies resulted in distinct utility scores and performed differently with distinct datasets. Second, to further investigate the relationship of data and fairness algorithms, we developed a statistical framework that was able to facilitate various analysis and decision making. Our framework could effectively and efficiently estimate the domain of data and solution space. We derived theoretical expressions to identify the fairness and relevance bounds for data of different distributions, and applied them to both synthetic datasets and real world datasets. We presented a series of use cases to demonstrate how our framework was applied to associate data and provide insights to fairness optimization problems. Third, we proposed an evaluation metric FAIR for the ranking results that encoded fairness, diversity, novelty and relevance. This metric offered a new perspective of evaluating fairness-aware ranking results. Based on this metric, we developed an effective ranking algorithm that jointly optimized for fairness and utility. Our experiments showed that our new metric was able to highlight results that achieved good user utility and fair information exposure at the same time. We showed how FAIR metric related to existing metrics through correlation analysis and case studies, and demonstrated the effectiveness of our FAIR-based algorithm.

Awarded by: Rutgers University, USA **on** 15 June 2020.
Supervised by: Chirag Shah.
Available at: https://doi.org/doi:10.7282/t3-ened-mx36.

Selected Publications

Zuohui Fu, Yikun Xian, Ruoyuan Gao, Jieyu Zhao, Qiaoying Huang, Yingqiang Ge, Shuyuan Xu, Shijie Geng, Chirag Shah, Yongfeng Zhang, and Gerard de Melo. Fairness-aware explainable recommendation over knowledge graphs. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, page 69–78, 2020.

Ruoyuan Gao and Chirag Shah. How fair can we go: Detecting the boundaries of fairness optimization in information retrieval. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*, pages 229–236, 2019.

Ruoyuan Gao and Chirag Shah. Toward creating a fairer ranking in search engine results. *Information Processing & Management*, 57:102–138, 2020a.

Ruoyuan Gao and Chirag Shah. Counteracting bias and increasing fairness in search and recommender systems. In *The 14th ACM Conference on Recommender Systems*, 2020b.

Yingqiang Ge, Shuchang Liu, Ruoyuan Gao, Yikun Xian, Yunqi Li, Xiangyu Zhao, Changhua Pei, Fei Sun, Junfeng Ge, Wenwu Ou, and Yongfeng Zhang. Towards long-term fairness in recommendation. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, page 445–453, 2021.