

# Balancing Exposure and Relevance in Academic Search

Andres Ferraro  
Universitat Pompeu Fabra  
Barcelona, Spain  
andres.ferraro@upf.edu

Lorenzo Porcaro  
Universitat Pompeu Fabra  
Barcelona, Spain  
lorenzo.porcaro@upf.edu

Xavier Serra  
Universitat Pompeu Fabra  
Barcelona, Spain

## ABSTRACT

The TREC 2020 Fair Ranking Track focuses on the evaluation of retrieval systems according to how well they fairly rank academic papers. The evaluation metric considered estimates how much the ranked papers are relevant and fairly represent different groups of authors, groups unknown to the track’s participants. In this paper, we present the three different solutions proposed by our team to the given problem. The first solution is built on a learning-to-rank model to predict how much the documents are relevant for a given query and modify the ranking based on this relevance and a randomization strategy. The second approach is also based on the relevance predicted by a learning-to-rank model, but it additionally selects the authors using categories defined by analyzing collaborations between authors. The third approach uses the DELTR framework, and it considers different categories of authors based on the corresponding H-class. The results show that the first approach gives the best performance, with the additional advantage that it does not require extra information about the authors.

## KEYWORDS

Information retrieval, TREC 2020 fair ranking track, exposure, fairness

## 1 INTRODUCTION

In the field of information retrieval, the area of fair ranking focuses on minimizing the disparity in items’ exposure while at the same time giving relevant results. However, fairness is not a concept generalizable and universally accepted in a univocal way. In the context of the Text Retrieval Conference (TREC) 2020, the specific goal of the Fair Ranking Track was to provide fair exposure to different groups of authors in an academic search task. For accomplishing it, we explored different ways to consider the trade-off between the relevance of the results and group exposure leading to diverse outcomes. First, we used a learning-to-rank model combined with randomization (§2.1), to ensure that results with the same relevance are likely to be ranked similarly. Second, on top of the previous logic, we build a group classification of authors based on their collaborations, making use of the authors’ network. We eventually used such classification for re-ranking the results obtained through the relevance-based model (§2.2). Third, we consider the DELTR framework proposed in [5], using the authors’ H-class as the main feature for training the learning-to-rank model (§2.3).

## 2 PROPOSED MODELS

We proposed three different alternatives for the reranking problem in TREC 2020 Fair Ranking Track. In this Section, we describe the

proposed solutions. The evaluation metric has been provided by the organizers of the track<sup>1</sup>, and it is expressed by the formula [4]:

$$e_a^\pi = \sum_{i=1}^n \left[ \gamma^{i-1} \prod_{j=1}^{i-1} (1 - p(s|\pi_j)) \right] I(\pi_i \in \mathcal{D}_a) \quad (1)$$

where

- $n$  : number of documents in ranking  $\pi$
- $\mathcal{D}_a$  : documents including  $a$  as an author
- $\pi_i$  : document a position  $i$
- $\gamma$  : continuation probability (fixed to 0 for the final position in the ranking)
- $p(s|d)$  : probability of stopping given user examined  $d$

The final expected exposure for  $a$  is the sum over all the impressions for a query,  $\Pi_q$ :

$$e_a = \sum_{\pi \in \Pi_q} e_a^\pi \quad (2)$$

## 2.1 Relevance

Our first approach is based on the solution proposed by Bonant<sup>2</sup> [2] to the 2019 Fair TREC. We first train a ranking model [3] using features extracted from the provided data, including the length of the query and the following information for each publication: title, abstract, entities, venue, journal, authors’ names, out-citations and in-citations.

Once the ranking model is trained, for a given query of length  $K$ , we compute the relevance score predicted by the model for each item ( $\hat{r}_i$ ). Then, we sort the items according to the predicted relevance and compute the mean difference ( $d$ ) following the Equation 3.

$$d = \frac{1}{K} \cdot \sum_{i=2}^K \hat{r}_i - r_{i-1} \quad (3)$$

Since the goal of this approach is that items with the same relevance will have the same probability to be in a position, we apply a randomization to the results based on the predicted relevance. The new relevance ( $\tilde{r}$ ) is computed following the Equation 4.

$$\tilde{r}_i = \text{random}(0, d) + \hat{r}_i \quad (4)$$

We then obtain the final ranking by sorting the items using the new relevance  $\tilde{r}_i$ .

<sup>1</sup><https://fair-trec.github.io/2020/doc/guidelines-2020.pdf>

<sup>2</sup>Code available at <https://github.com/irgroup/fair-trec>

## 2.2 Groups exposure

The main goal of the second solution was to give equal exposure to authors of different groups. In order to define the groups, we make use of a graph of authors connected by the collaborations between them in the papers. Afterwards, we use the method proposed by Blondel et al. [1] to create groups of authors, depending on the distance that they presented in the graph. Using this method, each author in the dataset was assigned to a single group. Therefore, authors in the same group should be co-authors in a paper, or should have some co-author in common.

In order to generate the reranking, for a given query we first obtain a relevance value following the procedure explain the Section 2.1, and then we assign the corresponding group to each candidate paper. Finally, we select the paper of each group with higher relevance and we repeat this process until all the papers are ranked.

## 2.3 Protected groups

The third solution proposed is based on the framework created by Zehlike and Castillo [5]. DELTR is a learning-to-rank framework that allows to take into account the different categories of authors to addresses potential issues of discrimination and unequal opportunity. One of the parameters that needs to be defined in DELTR is the 'gamma', which indicates the importance of the provided class. Gamma equals zero gives more importance to relevance over fairness, and gamma value of one gives more importance to fairness over relevance. In our case, we used the H-class provided by the organizers to train DELTR. In order to obtain a trade-off between relevance and fairness we train two models of DELTR, one with gamma=1 and another with gamma=0, then for each query we used a weighted average of the score obtain by each model.

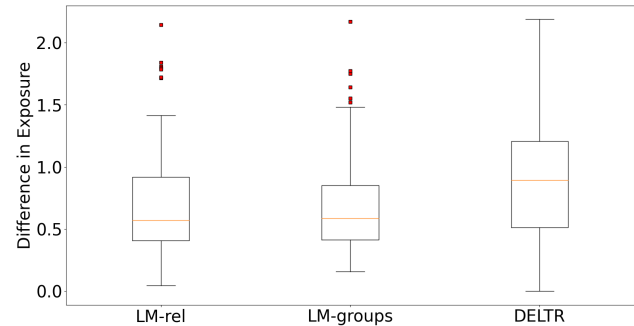
## 3 DISCUSSION AND FUTURE WORK

In Figure 1, we provide details about the performance of the three approaches from the information published by the organizers, showing the distributions of the difference in exposure compared with the expected exposure on each query. These results are also summarized in Table 1, where the average expected exposure for each model is reported.

We can see that the performance of the randomized ranking based on relevance gives a lower difference with the expected exposure. Therefore, the proposed alternative for the first solution based on authors' groups, in this case, does not improve the results. However, including the groups' information seems to mitigate the wrong behaviors of the first model, especially in the cases where it shows worst performances (see the outliers in Figure 1). This mitigation is also proven by the smaller variance of the results of the second model, in comparison to the first one, indicating that this solution could be effective for joining group constraints and relevance. Randomizing the order of the groups when selected in the reranking step could also be helpful in reducing the variance of the results. Nonetheless, the groups created with our model are not symmetrically distributed, having few groups with a lot of authors, and a series of small groups with only two authors. Because of how academic networks are built, where few big communities collaborate while many small research labs are not connected, our grouping represents a real-world scenario, which however implies

**Table 1: Results of the proposed solutions**

Model	Difference in exposure – $avg(e_a)$
LM-rel	0.6733
LM-groups	0.6755
Deltr	0.9097



**Figure 1: Difference in exposure summary statistics – box-plots.**

a structural unfairness which partly has been embedded in our model. Using a group definition which does not take into account only collaboration, but also other socio-political differentials should lead to improved results.

We also see that the proposed solution based balancing fairness and relevance using DELTR gives a higher difference with the expected exposure compared to the other approaches.

In summary, we proposed three different alternatives to balance exposure and relevance of the documents. Basing on the results, we see that the more simple solution is the one that performed better. Nevertheless, is important to mention that the proposed alternatives followed an exploratory approach and is possible that further studies could lead to improvements in the results of each solution.

## ACKNOWLEDGMENTS

This work is partially supported by the European Commission under the TROMPA project (H2020 770376). This work is also partially supported by Kakao Corp.

## REFERENCES

- [1] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.
- [2] Malte Bonart. [n.d.]. Fair ranking in academic search. ([n. d.]).
- [3] Christopher JC Burges. [n.d.]. From ranknet to lambdarank to lambdamart: An overview. In *Technical Report MSR-TR-2010-82*.
- [4] Fernando Diaz, Bhaskar Mitra, Michael D Ekstrand, Asia J Biega, and Ben Carterette. 2020. Evaluating Stochastic Rankings with Expected Exposure. *arXiv preprint arXiv:2004.13157* (2020).
- [5] Meike Zehlike and Carlos Castillo. 2020. Reducing disparate exposure in ranking: A learning to rank approach. In *Proceedings of The Web Conference 2020*. 2849–2855.