Dissatisfaction Induced by Pairwise Swaps*

Discussion Paper

Alessandro Fabris[1,2], Gianmaria Silvello[2], Gian Antonio Susto[2] and Asia J. Biega[1]

[1]Max Planck Institute for Security and Privacy, Bochum, Germany
[2]University of Padova, Padova, Italy

Abstract

Fairness is increasingly recognized as an important property of information access systems. Pairwise fairness is a measure of equity in ranking whose normative grounding has not been clearly studied nor discussed in the literature. In this work, we target this gap by providing a clear interpretation for this family of measures, by demonstrating and remedying its key limitations, and by analysing its relationship to other measures of fair ranking.

1. Introduction

Information Access Systems (IAS) have become increasingly prominent in recent years as they help users interact with large amounts of content through the ranking and presentation of items based on their estimated relevance or merit [2, 3]. In the context of IAS, content producers are now seen as important stakeholders, whose economic and societal needs should be considered alongside consumers to promote a fair and productive information ecosystem [4, 5, 6]. Algorithmic fairness [7, 8] is a research field concerned with ensuring equitable algorithmic outcomes through specific measures [9, 10], algorithmic designs [11, 12], and auditing procedures [13, 14].

In this work, we explore the *pairwise fairness* family of ranking measures [15, 16, 17], providing a new interpretation based on browsing models and highlighting limitations of existing metrics. We propose a new metric that overcomes these limitations by modeling realistic browsing behaviors and individual provider perspectives. This new measure captures aspects of observed unfairness and dissatisfaction, specifically related to the perceived quality of IAS by content producers. Additionally, we characterize the relationship between pairwise and exposure-based fairness measures both analytically and empirically. Overall, we make significant contributions by offering a new interpretation of pairwise fairness, proposing a novel metric, and studying the relationship between pairwise and exposure-based fairness.

IIR2023: 13th Italian Information Retrieval Workshop, 8th - 9th June 2023, Pisa, Italy

*Full paper originally presented at WSDM 2023 [1].

alessandro.fabris@mpi-sp.org (A. Fabris); silvello@dei.unipd.it (G. Silvello); susto@dei.unipd.it (G. A. Susto); asia.biega@mpi-sp.org (A. J. Biega)
0000-0001-6108-9940 (A. Fabris); 0000-0003-4970-4554 (G. Silvello); 0000-0001-5739-9639 (G. A. Susto); 0000-0001-8083-0976 (A. J. Biega)

CEUR Workshop Proceedings (CEUR-WS.org)

2. Background and Related Work

Notation. We let $i \in \mathcal{I}$ denote an item in a set to be ranked, and r_i denote its relevance. We let $g \in \mathcal{G} = \{A, B\}$ indicate a (binary, for ease of exposition) sensitive attribute,[1] and $i \in g$ the membership of i in group g. We use σ_* for an "ideal" ranking, i.e., a permutation which orders items decreasingly by relevance: $\sigma_* = \text{argsort}(r_i)$. Finally, we let σ denote a ranking returned by the IAS in response to a query, and $\sigma(k)$ indicate the item ranked by σ in position k.

Discordant pairs. The notion of *discordant pair* is key to pairwise fairness. Two items $i, j \in \mathcal{I}$ form a discordant pair if their relative order in σ_* and σ is different. Let $\sigma^{-1}(i)$ indicate the position of item i in σ, i.e., $\sigma^{-1}(i) = k \iff \sigma(k) = i$. The indicator function for a discordant pair in rankings σ and σ_* is defined as

$$d(i,j) = \underbrace{\mathbb{1}(\sigma^{-1}(i) < \sigma^{-1}(j), \sigma_*^{-1}(i) > \sigma_*^{-1}(j))}_{d_F(i,j)} + \underbrace{\mathbb{1}(\sigma^{-1}(i) > \sigma^{-1}(j), \sigma_*^{-1}(i) < \sigma_*^{-1}(j))}_{d_U(i,j)}$$

In other words, i is in a discordant pair when ranking σ unfairly puts it at an advantage (d_F) or a disadvantage (d_U) on item j; subscripts F and U denote that the first item is in a Favorable Discordant Pair (FDP) or an Unfavorable Discordant Pair (UDP).

Pairwise Fairness. Inter-Group Inaccuracy (IGI) [15] and Rank Equality Error (REE) [16], are the two most popular measures of pairwise fairness, defined as

$$M_{AB} = \frac{1}{C_{AB}} \cdot \sum_{i \in A} \sum_{j \in B} d_U(i,j), \tag{1}$$

where C_{AB} is a normalizing constant. The literature lacks an explicit discussion of the normative reasoning behind these pairwise fairness metrics and the construct they capture. For instance, for IGI, Beutel et al. [15] "draw on the intuition of Hardt et al. [11] for equality of odds, where the fairness of a classifier is quantified by comparing its false positive rate and/or false negative rate.", while REE is based on the "postulate that there is value in considering error-based fairness criteria for rankings" [16].

3. What does Pairwise Fairness Actually Measure?

Browsing model. To provide an interpretation for pairwise fairness, we begin by demonstrating and deriving its implicit user browsing model. REE and IGI are related to Kendall's Tau [18], according to which the inaccuracy of a ranking can be written as

$$M = \frac{1}{C} \cdot \sum_{i} \sum_{j \neq i} d(i,j) = \frac{1}{C} \cdot \sum_{k=0}^{n-1} \sum_{k'=0}^{k-1} F(k') d_U(k, k'), \tag{2}$$

where we use $d(k, k') = d(\sigma(k), \sigma(k'))$ as shorthand notation for a discordant pair of items ranked by σ at positions (k, k'). Moreover, we let $F(k)$ denote the probability that users will

[1] We follow the literature on pairwise fairness and consider binary sensitive attributes.

visit the item $\sigma(k)$. The equality in Equation (2) holds under a trivial browsing model where users visit all items with the same probability $F(k) = 1 \ \forall k$.
Interpretation. At rank k, item producers evaluate ranking σ by focusing on the most visible cases of unfair treatment against their item $\sigma(k)$. Their dissatisfaction with σ grows each time they encounter a UDP for $\sigma(k)$, which is an item of lesser relevance ranked better than their own. The inner summation $\sum_{k'=0}^{k-1} F(k')d_U(k, k')$ represents a weighted counter of UDPs, with the weight proportional to the visibility of the unjustly favored item. Kendall's Tau is interpreted as operationalizing aggregate producer dissatisfaction with σ for unjustly favoring other items.

This interpretation also applies to pairwise fairness (Equation 1), by focusing on cross-group comparisons.

$$M_{AB} = \frac{1}{C_{AB}} \cdot \sum_{k=1}^{n-1} \sum_{k'=0}^{k-1} F(k')d_U(k, k') \cdot \mathbb{1}(\sigma(k) \in A, \sigma(k') \in B)$$

This formulation summarizes the dissatisfaction of items and their producers in one group due to being unfairly ranked lower than items of lesser relevance from another group. Pairwise fairness thus communicates observed injustice, which can affect perceptions of platform quality [19, 20], and influence the loyalty of item producers [21].

4. Current Limitations and Proposed Improvements

Fabris et al. [1] describe several limitations of pairwise fairness and overcome them with targeted reformulations, two of which are presented below.
Top-heaviness. Pairwise fairness metrics do not consider realistic browsing behaviors. In particular, they use a uniform visit probability for all ranks, which is not realistic in practice. The top ranking positions are more likely to be visited by searchers, and this should be accounted for in the metrics. As shown above, pairwise fairness measures can account for user browsing models $F(k)$:

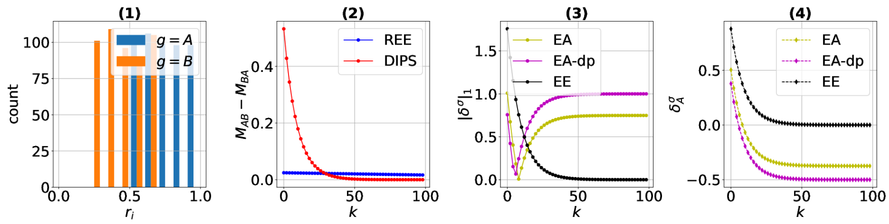
$$M_{AB} = \frac{1}{C_{AB}} \sum_{i \in A} \sum_{k=0}^{n-1} F(k)d_U(i, \sigma(k)) \cdot \mathbb{1}(\sigma(k) \in B). \tag{3}$$

The IAS literature has proposed and studied several top-heavy user models, including logarithmic ($F(k) \propto 1/\log(k)$ [22]) and exponential discount ($F(k) \propto \gamma^k$ [23]).
Tie handling. Pairwise fairness metrics such as IGI and REE do not consider ties in relevance scores. Ties are common in practical applications like recommender systems and information retrieval, where relevance judgments are often discrete or quantized. This means that IAS that favor a group by breaking ties in its favor are not flagged as problematic by IGI or REE.

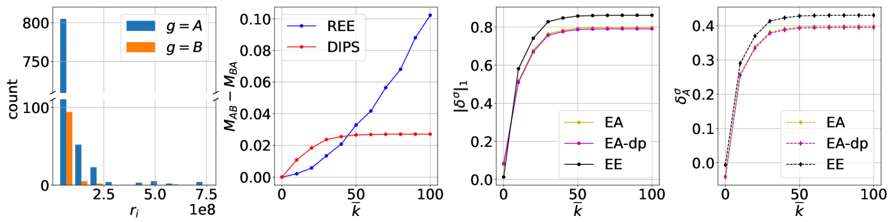
Since $\sigma_* = \text{argsort}(r_i)$, we rewrite the indicator function for UDPs as $d_U(i, j) = \mathbb{1}(\sigma^{-1}(i) > \sigma^{-1}(j), r_i > r_j)$. We generalize UDPs to handle ties as:

$$d_U(i, j) = \mathbb{1}(\sigma^{-1}(i) > \sigma^{-1}(j), r_i > r_j) + c_t \mathbb{1}(\sigma^{-1}(i) > \sigma^{-1}(j), r_i = r_j) \tag{4}$$

Here c_t indicates the dissatisfaction of an item ranked worse than another of same relevance. We call this *partial UDP*. Values for c_t range in $(0, 1)$, where $c_t = 0$ indicates indifference to ties, while $c_t = 1$ corresponds to partial UDPs leading to the same dissatisfaction as regular UDPs.



(a) Synthetic data.



(b) Real-world data.

Figure 1: Distribution of relevance r_i (1); comparison of pairwise fairness REE and DIPS (2) with exposure-based measures EE, EA, EA-dp: aggregate measure $|\delta^{\sigma}|_1$ (3) and individual component δ^{σ}_A (4).

5. Relation to Exposure-based Fairness

Based on the limitations and improvements discussed above, we propose Dissatisfaction Induced by Pairwise Swaps (DIPS), a new pairwise fairness measure defined as

$$M_{AB}^{\text{DIPS}} = \frac{1}{C_{AB}^{\text{DIPS}}} \sum_{i=0}^{n-1} \sum_{k=0}^{n-1} F(k) d_U(i, \sigma(k)) \cdot \mathbb{1}(i \in A, \sigma(k) \in B), \tag{5}$$

which can model top-heavy browsing models $F(k)$ and handle ties through parameter c_t in the definition of $d_U(\cdot)$. In Figure 1, we compare DIPS with Equity of Attention (EA) [24] and Expected Exposure (EE) [25] on both a synthetic and real-world dataset. These experiments, presented in more detail in [1], along with an analytical comparison between these measures, yield two key interpretations. On one hand, DIPS inherits a top-heavy behavior from $F(k)$ and is thus suited to highlight UDPs at highly visible ranks, similarly to EA, EE and in opposition to REE (Figure 1a). On the other hand, DIPS captures a different construct from EE and EA, enabling a desirable outcome: fairness interventions in favor of a group can have a sizeable impact on group equity, as measured by EA and EE, while maintaining dissatisfaction low for the privileged group, as measured by DIPS (Figure 1b).

6. Conclusion

Our work motivates and generalizes pairwise fairness in ranking by retrospectively mapping it to the construct of producer dissatisfaction, highlighting its current limitations and proposing specific improvements. We also compare it to other families of fair ranking measures. We add to the ongoing discussion about the normative reasoning of algorithmic fairness, supporting an informed and contextualized adoption of these measures.

References

[1] A. Fabris, G. Silvello, G. A. Susto, A. J. Biega, Pairwise fairness in ranking as a dissatisfaction measure, in: Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, WSDM '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 931–939. URL: https://doi.org/10.1145/3539597.3570459. doi:10.1145/3539597.3570459.

[2] R. Baeza-Yates, B. Ribeiro-Neto, et al., Modern information retrieval, volume 463, ACM press New York, 1999.

[3] T. Joachims, Optimizing search engines using clickthrough data, in: Proc. of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '02, ACM, 2002, p. 133–142. URL: https://doi.org/10.1145/775047.775067. doi:10.1145/775047.775067.

[4] M. D. Ekstrand, A. Das, R. Burke, F. Diaz, Fairness and discrimination in information access systems, arXiv preprint arXiv:2105.05779 (2021).

[5] E. Pitoura, K. Stefanidis, G. Koutrika, Fairness in rankings and recommendations: an overview, The VLDB Journal (2021) 1–28.

[6] M. Zehlike, K. Yang, J. Stoyanovich, Fairness in ranking, part i: Score-based ranking, ACM Comput. Surv. (2022). URL: https://doi.org/10.1145/3533379. doi:10.1145/3533379, just Accepted.

[7] S. Barocas, M. Hardt, A. Narayanan, Fairness and Machine Learning, fairmlbook.org, 2019. http://www.fairmlbook.org.

[8] A. Fabris, S. Messina, G. Silvello, G. A. Susto, Algorithmic fairness datasets: the story so far, Data Mining and Knowledge Discovery 36 (2022) 2074–2152.

[9] A. Castelnovo, R. Crupi, G. Greco, D. Regoli, I. G. Penco, A. C. Cosentini, A clarification of the nuances in the fairness metrics landscape, Scientific Reports 12 (2022) 4209.

[10] A. Raj, M. D. Ekstrand, Measuring fairness in ranked results: An analytical and empirical comparison, in: Proc. of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022.

[11] M. Hardt, E. Price, N. Srebro, Equality of opportunity in supervised learning, in: Proc. of the 29th Annual Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, ES, 2016, pp. 3323–3331.

[12] S. C. Geyik, S. Ambler, K. Kenthapadi, Fairness-aware ranking in search & recommendation systems with application to linkedin talent search, in: Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining, 2019, pp. 2221–2231.

[13] A. Fabris, A. Mishler, S. Gottardi, M. Carletti, M. Daicampi, G. A. Susto, G. Silvello, Algorithmic audit of italian car insurance: Evidence of unfairness in access and pricing, in: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, AIES '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 458–468. URL: https://doi.org/10.1145/3461702.3462569. doi:10.1145/3461702.3462569.

[14] A. Fabris, A. Esuli, A. Moreo, F. Sebastiani, Measuring fairness under unawareness of sensitive attributes: A quantification-based approach, Journal of Artificial Intelligence Research 76 (2023) 1117–1180.

[15] A. Beutel, J. Chen, T. Doshi, H. Qian, L. Wei, Y. Wu, L. Heldt, Z. Zhao, L. Hong, E. H. Chi,

C. Goodrow, Fairness in recommendation ranking through pairwise comparisons, in: Proc. of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, ACM, 2019, p. 2212–2220.

[16] C. Kuhlman, M. VanValkenburg, E. Rundensteiner, Fare: Diagnostics for fair ranking using pairwise error metrics, in: The World Wide Web Conference, WWW '19, ACM, 2019, p. 2936–2942. URL: https://doi.org/10.1145/3308558.3313443. doi:10.1145/3308558.3313443.

[17] H. Narasimhan, A. Cotter, M. Gupta, S. Wang, Proc. of the AAAI Conference on Artificial Intelligence 34 (2020) 5248–5255. URL: https://ojs.aaai.org/index.php/AAAI/article/view/5970. doi:10.1609/aaai.v34i04.5970.

[18] M. G. Kendall, A new measure of rank correlation, Biometrika 30 (1938) 81–93.

[19] R. Dudley, Amazon's new competitive advantage: Putting its own products first, https://www.propublica.org/article/amazons-new-competitive-advantage-putting-its-own-products-first, 2020.

[20] A. Jeffries, L. Yin, Amazon puts its own "brands" above better rated products, https://themarkup.org/amazons-advantage/2021/10/14/amazon-puts-its-own-brands-first-above-better-rated-products, 2021.

[21] J. Kim, Platform quality factors influencing content providers' loyalty, Journal of Retailing and Consumer Services 60 (2021) 102510. URL: https://www.sciencedirect.com/science/article/pii/S096969892100076X. doi:https://doi.org/10.1016/j.jretconser.2021.102510.

[22] K. Järvelin, J. Kekäläinen, Cumulated gain-based evaluation of ir techniques, ACM Trans. Inf. Syst. 20 (2002) 422–446. URL: https://doi.org/10.1145/582415.582418. doi:10.1145/582415.582418.

[23] A. Moffat, J. Zobel, Rank-biased precision for measurement of retrieval effectiveness, ACM Trans. Inf. Syst. 27 (2008). URL: https://doi.org/10.1145/1416950.1416952. doi:10.1145/1416950.1416952.

[24] A. J. Biega, K. P. Gummadi, G. Weikum, Equity of attention: Amortizing individual fairness in rankings, in: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR '18, ACM, 2018, p. 405–414.

[25] F. Diaz, B. Mitra, M. D. Ekstrand, A. J. Biega, B. Carterette, Evaluating stochastic rankings with expected exposure, in: Proc. of the 29th ACM International Conference on Information & Knowledge Management, CIKM '20, ACM, 2020, p. 275–284. URL: https://doi.org/10.1145/3340531.3411962. doi:10.1145/3340531.3411962.