Fair Ranking as Fair Division

Poch Laohrenu*
pl2839@columbia.edu

Jirat Suchato*
js5808@columbia.edu

Leo Hu*
ch3729@columbia.edu

Richard Uzeel*
ru2155@columbia.edu

1 Why Fair Ranking?

Fair ranking has applications to online marketplace settings. Allocating ranks to items should satisfy both users and item creators, providing all items with the chance to gain exposure, leading to economic opportunity, as well as users with utility. Traditional ranking systems rank items based on their relevance in descending order. However, this discourages new content creators, as they often struggle to gain initial exposure due to popularity bias. Not only does this hurt new creators, it hurts the platform in the long term. The objective of fair ranking is to balance the trade-off between providing exposure to new items with the users' utility from having the most relevant items appear first.

2 Fair Ranking as Fair Division

In this problem, the content can be considered as agents, and the ranking slots can be considered as items. Each content is assigned a probability distribution where $P_i(X = K)$ denotes the probability that item i is placed into slot K. Thus, the fair ranking as fair division problem can be represented as a bipartite graph with 2K nodes, of which K nodes represent items, and K nodes represent slots. For each node representing an item, K edges connect that node to each node representing an item.

2.1 Previous Study: Exposure-Based Fair Ranking

Previous studies explore how to explicitly link merit to the exposure that is allocated to individual items or groups of items. For example, the seminal papers by Singh and Joachims [5], as well as Biega et al [1], established concrete definitions of exposure-based fairness. These works introduced constraints to ensure that exposure is allocated in proportion to relevance, either across different groups in expectation or cumulatively across a series of rankings. However, there is no justification on the functional form that links merit and exposure. Formulating fair ranking as fair division aims to address this problem.

2.2 Impact-Based Fair Ranking

In the paper "Fair Ranking as Fair Division: Impact-Based Individual Fairness in Ranking," Saito and Joachims developed a fair ranking method that utilizes fair division principles.[4] The authors claim to be the first to successfully implement *impact-based individual item fairness*. Unlike in other works in this field, Saito and Joachims's formulation accounts for position bias in ranking. This adjustment makes the ideas more applicable to the real world where people are more likely to be persuaded by the items that they see first rather than last. Also, unlike the previous research [3], they proposed axioms of fairness based on the impact (i.e., clicks and revenue) among individual items, rather than on the item's exposure which produced substantial envy and did not differentiate between being shown to

*These authors contributed equally to this work

Preprint. Under review.

target and non-target audiences. To illustrate the relevance of impact, one can consider showing a puffer jacket to two different groups, those who live in warm and cold climates. *Ceteris paribus*, the two groups have the same exposure but a very different impact because a warm jacket might not be needed in areas with high average temperatures. Previous studies would have considered these two cases to be equivalent, but there is a clear difference.

Saito and Joachims build an item-centric matrix $X^{\pi}_{*,i,*}$ whose (u,k) element is: $X^{\pi}_{u,i,k} = \mathbb{P}(\sigma(i) = k \mid \pi, u)$, where $\sigma(i)$ is the rank of item i in the ranking σ, k is the position in a ranking for user u, and π is a ranking policy. In words, $X^{\pi}_{u,i,k}$ is the probability that item i is ranked at position k for user u under allocation π. The matrix, $X^{\pi}_{*,i,*}$, characterizes the allocation of positions, i.e., how big a fraction of the $k-th$ position in a ranking for user u goes to item i. Saito and Joachims further define an $impact$ metrics: $\mathrm{Imp}_i\left(X^{\pi}_{*,i,*}\right) := \sum_{u \in \mathcal{U}} \sum_{k=1}^{n} v_i(u,k) X^{\pi}_{u,i,k}$ where $v_i(u,k)$ is an application-dependent impact function, which defines how much impact (e.g., expected clicks, bookings, revenue) item i receives when it is ranked at the k-th position for user u. Using this, one can define utility to the users as

$$U(\pi) = \sum_{i \in \mathcal{I}} \mathrm{Imp}_i\left(X^{\pi}_{*,i,*}\right)$$

2.3 Fairness Axioms

Given a ranking policy π, their three axioms are as follows:

- Envy-Freeness: No item would have strictly more impact if it received another item's position allocation.
  - Formally: $Imp_i\left(X^{\pi}_{*,i,*}\right) \geq Imp_i\left(X^{\pi}_{*,j,*}\right), \forall i,j \in \mathcal{I}$.
- Dominance Over Uniform Ranking: Each item receives a weakly better allocation than under a uniform random policy, π_{unif}, where each permutation is equally likely. Additionally, (at least) one item is strictly better off than under π_{unif}.
  - Formally: $Imp_i\left(X^{\pi}_{*,i,*}\right) \geq Imp_i\left(X^{\pi_{unif}}_{*,i,*}\right), \forall i \in \mathcal{I}$.
- Pareto Optimality: No item can be made strictly better off without making another one strictly worse off.

2.4 Standard NSW Program

Saito and Joachims proposed an algorithm (NSW) to compute an allocation that is approximately envy-free and Pareto optimal at the cost of user utility. They find that the NSW algorithm maintains low envy and Pareto optimality as popularity bias increases and that utility decreases for NSW and increases in most cases for uniform and exposure-based algorithms. The NSW formulation is as follows:

$$\begin{aligned}
\pi_{NSW} = \underset{\left\{X^{\pi}_{*,i,*}\right\}_{i \in I}}{\arg\max} & \prod_{i \in \mathcal{I}} Imp_i\left(X^{\pi}_{*,i,*}\right), \\
\text{s.t. } & \sum_{k=1}^{n} X^{\pi}_{u,i,k} = 1, \forall (u,i) \\
& \sum_{i \in \mathcal{I}} X^{\pi}_{u,i,k} = 1, \forall (u,k) \\
& 0 \leq X^{\pi}_{u,i,k} \leq 1, \forall (u,i,k)
\end{aligned} \tag{1}$$

The first constraint expresses that each item needs to have probability 1 to be placed in some position, and the second one expresses that each position needs to have probability 1 of receiving an item. Here, the second constraint corresponds to the supply constraint in the standard MNW formulation which can be easily solved by CEEI method to achieve Pareto optimality, envy-free, and dominance over a uniform random allocation. However, with the first constraint being added, the result from CEEI does not necessarily hold [2].

2.5 α-NSW Program

Saito and Joachims also propose α-NSW algorithm, which allows for controlling the trade-off between maximizing impact-based fairness (low α) and user utility (high α), by choosing the hyperparameter α. When α = 0, this becomes the Standard NSW case from the previous section. Due to the trade-off, this new policy can lead to a higher user utility than in the Standard NSW.

$$\begin{aligned}
\pi_{\alpha-NSW} = &\underset{\left\{X^{\pi}_{*,i,*}\right\}_{i\in\mathcal{I}}}{\arg\max} \prod_{i\in\mathcal{I}} \mathit{Imp}_i\left(X^{\pi}_{*,i,*}\right)^{\text{Merit}_i^{\alpha}}, \\
\text{s.t.} &\sum_{k=1}^{n} X^{\pi}_{u,i,k} = 1, \quad \forall(u,i), \\
&\sum_{i\in\mathcal{I}} X^{\pi}_{u,i,k} = 1, \quad \forall(u,k), \\
&0 \le X^{\pi}_{u,i,k} \le 1, \quad \forall(u,i,k).
\end{aligned} \tag{2}$$

Here, $Merit_i$ is item i's relevance after being amortized over all of the users, found through $Merit_i$:= $\sum_{u\in\mathcal{U}} r(u,i)$. Additionally, $r(u,i)$ is a positive real number quantifying the relevance between user u and item i. The α here is an exponent for $Merit_i$.

3 Problem with the Extra Constraint

In the Scalable fair division for 'at most one' preferences paper from Kroer and Peysakhovich [2], the authors studied the allocation setting where individuals want at most one of any given item. Therefore, apart from the supply constraint, the NSW program needs to have an additional constraint that limits the allocation of each user-item pair to be at most 1, which they refer this as **at most one** (AMO) setting. This AMO setting is closely aligned with our fair division problem since an item would only need, at maximum, a probability of 1 to be placed in some position for a user. Kroer and Peysakhovich showed, by derivation, that the presence of this AMO constraint prevents the solution of the optimal NSW program from achieving market equilibrium and thus, the fairness axioms are violated. In the following sections, we will apply the same derivation to our setting to show theoretically that the presence of our extra constraints leads to the violation of the fairness axioms. We will also show numerically the severity of the violation by running experiments on synthetic data with different parameters.

4 Theoretical Derivation

Starting from the original NSW program from (1), we replaced the product of impacts in objective with the sum of their logarithms to linearize the objective function whilst maintaining convexity. We also substitute the expression of the impact into the objective function and add the budget of each item B_i in for the purpose of derivation. Since $B_i = 1$, it doesn't change the objective function.

$$\begin{aligned}
\max &\sum_{i\in\mathcal{I}} B_i \log(\sum_{u\in\mathcal{U}} \sum_{k=1}^{K} v_{u,i,k} \cdot X^{\pi}_{u,i,k})), \\
\text{s.t.} &\sum_{k=1}^{K} X^{\pi}_{u,i,k} \le 1, \forall(u,i) \\
&\sum_{i\in\mathcal{I}} X^{\pi}_{u,i,k} \le 1, \forall(u,k) \\
&0 \le X^{\pi}_{u,i,k} \le 1, \forall(u,i,k)
\end{aligned} \tag{3}$$

We then convert this optimization problem into the Eisenberg-Gale convex program by applying the Lagrange multiplier method to both of the constraints, which yields

$$\max \sum_{i\in\mathcal{I}} B_i \log(\sum_{u\in\mathcal{U}}\sum_{k=1}^{K} v_{u,i,k}\cdot X^{\pi}_{u,i,k})) - \sum_{u\in\mathcal{U}}\sum_{k=1}^{K} p_{u,k}(\sum_{i\in\mathcal{I}} X^{\pi}_{u,i,k} - 1) - \sum_{u\in\mathcal{U}}\sum_{i\in\mathcal{I}} y_{u,i}(\sum_{k=1}^{K} X^{\pi}_{u,i,k} - 1) \tag{4}$$

where $p_{u,k}$ is the Lagrangian multiplier of the supply constraint, representing shadow price of each ranking slot for user u and rank k. $y_{u,i}$ is the Lagrange multiplier of the additional constraint.

Let $\beta_i = \frac{B_i}{\sum_{u,k} v_{u,i,k} X^{\pi}_{u,i,k}}$, then it follows that the first-order optimally condition with respect to $X^{\pi}_{u,i,k}$ yields

$$v_{u,i,k}\beta_i - p_{u,k} - y_{u,i} \leq 0 \Rightarrow v_{u,i,k}\beta_i \leq p_{u,k} + y_{u,i}$$

If $X^{\pi}_{u,i,k} > 0$, then

$$v_{u,i,k}\beta_i = p_{u,k} + y_{u,i} \tag{5}$$

Multiplying (5) by $X^{\pi}_{u,i,k} > 0$ and summing over u, k yields

$$\sum_{u,k} X^{\pi}_{u,i,k}(p_{u,k} + y_{u,i}) = \sum_{u,k} X^{\pi}_{u,i,k} \frac{B_i}{\sum_{u,k} v_{u,i,k} X^{\pi}_{u,i,k}} = B_i \tag{6}$$

According to (6), we can see that when $y_{u,i} > 0$ for some u, i, we no longer achieve a market equilibrium under prices $p_{u,k}$ because the buyer (the item), is spending less than the budget B_i. Therefore, the fairness axioms are not necessarily satisfied.

Furthermore, notice that for $K = 1$, the extra constraint becomes trivial and thus $y_{u,i} = 0$. Therefore, the fairness axioms are only violated when $K > 1$.

5 Experiments: NSW (α = 0)

In this section, we ran simulations to test the impact of K on fairness. The code for all experiments can be found in the GitHub repository (https://github.com/pochl/AGM2023-FairRanking).

First, we define the metric we use to measure envy, which is **Mean Max Envy**.

$$\frac{1}{|\mathcal{I}|}\left\{\sum_{i\in\mathcal{I}} \max_{j\in\mathcal{I}} Imp_i\left(X^{\pi}_{*,j,*}\right) - Imp_i\left(X^{\pi}_{*,i,*}\right)\right\} \tag{7}$$

To generate synthetic data, the ground-truth relevance between user u and item i is defined as

$$r_{true}(u,i) = (1-\lambda)\cdot r_{unif}(u,i) + \lambda \cdot r_{pop}(u,i)$$

where $r_{unif}(u,i)$ is generated uniformly and independently within [0,1] range. $r_{pop}(u,i)$ is the popularity bias term, which is defined as

$$r_{pop}(u,i) := \begin{cases} \frac{n-i+1}{n}\cdot\frac{m-u+1}{m} & \text{(randomly sampled 70\% of the items)} \\ \frac{n-i+1}{n}\cdot\frac{u}{m} & \text{(rest of the items)} \end{cases} \tag{8}$$

We can control the degree of popularity bias by adjusting $\lambda \in [0,1]$, where the larger the value, the more popularity bias we have.

In this experiment, we use the following parameters, $|\mathcal{U}| = 100, |\mathcal{I}| = 50, \lambda = 0.5$, then we varied the values of K and observe the Mean Max Envy. For each value of K, we run the experiment 10 times and measure the average Mean Max Envy to reduce variability.

5.1 Result

According to Figure 1, it is clear that as K gets larger, the Mean Max Envy increases. This means that as more ranking slots are introduced to users, the further the solution is from being envy-free.

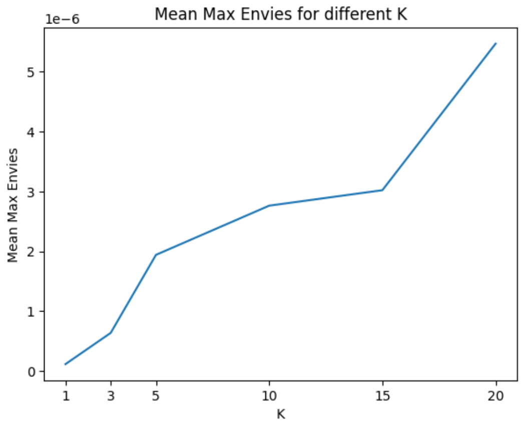

Figure 1: Mean Max Envy for different values of K (α = 0)

5.2 Approximate Equilibrium in Large Market

Even though the theoretical derivation shows that the envy-freeness property is violated, this doesn't necessitate that the algorithm cannot achieve envy-freeness in practice. Kroer and Peysakhovich came up with the concept of "twins" in their paper [2], where twins are buyers with approximately the same valuation of the products. Saito and Joachims [4] then applied this concept to the fair ranking problem and derived the theorem that says if every item has at least K+1 twins, then the solution approximately satisfies market equilibrium and thus is approximately envy-free.

5.3 Experimental Result

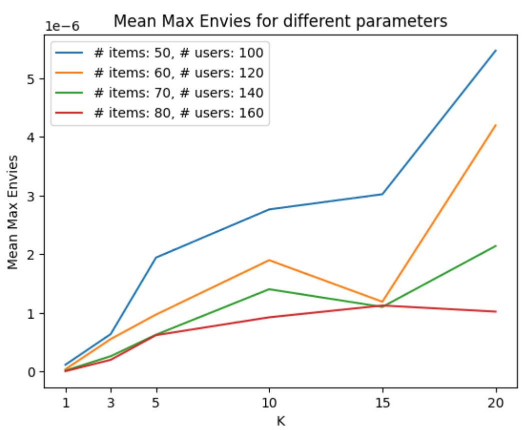

Figure 2: Mean Max Envy for different values of K and market size (α = 0)

To explore the impact of market size experimentally, we ran the same experiment as the previous section, but also with varying the number of users $|\mathcal{U}|$ and items $|\mathcal{I}|$, though we keep the ratio between users and items $|\mathcal{U}|/|\mathcal{I}|$ the same at 2. Although the Mean Max envy is generally increasing with K, it can be seen in Figure 2, that as the size of the market grows, the Mean Max Envy reduces. Since it is very common to have thousands of items and tens of thousands of users in practice, the Mean Max Envy will approximately converge to zero in a sufficiently large market.

6 Experiments: NSW with $\alpha > 0$

Next, we explored if the α-*NSW* problem can yield envy freeness in a large market for values of 0.5, 1, and 2. Again, we kept the ratio $|\mathcal{U}|/|\mathcal{I}|$ the same at 2. The results observed differ from the $\alpha = 0$ case because small α values guarantee impact-based fairness, so as the α value increases, there will be less fairness (but the user utility will be maximized, due to the trade-off). Consequently, in Figures 3,4,and 5, we no longer see that the Mean Max Envy will approximately converge to zero in a large market. In fact, we see that as the fairness decreases, the Mean Max Envy increases for the same number of items and users, as expected.

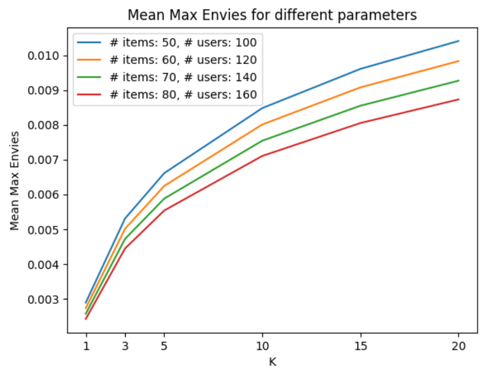

Figure 3: Mean Max Envy for different values of K and market size ($\alpha = 0.5$)

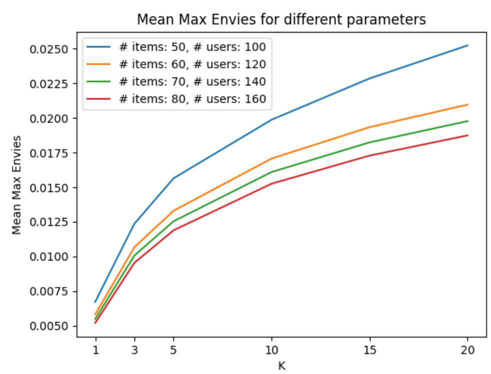

Figure 4: Mean Max Envy for different values of K and market size ($\alpha = 1$)

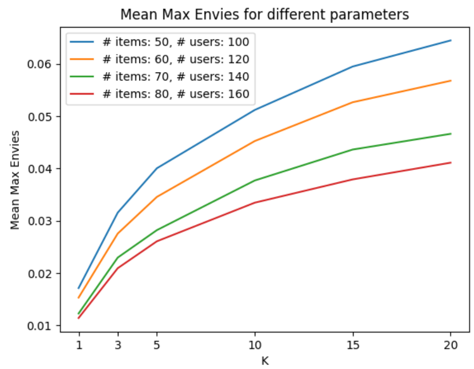

Figure 5: Mean Max Envy for different values of K and market size ($\alpha = 2$)

7 Conclusion

Although envy freeness does not hold theoretically, we find in practice solutions can be nearly envy free in large markets where the algorithms focuses on imapct based fariness. However, if we were to adapt the algorithm to favor user utility, envy freeness no longer holds, even in the large market case.

8 Further Work

We want to come up with 2 metrics, one for each of dominance over uniform ranking and Pareto optimality.

Further exploration can be done on fine tuning the alpha value to optimize the trade-off between utility and fairness where we could define a metric where custom weights are given to both of these components.

References

[1] Asia J. Biega, Krishna P. Gummadi, and Gerhard Weikum. Equity of attention: Amortizing individual fairness in rankings. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 2018.

[2] Christian Kroer and Alexander Peysakhovich. Scalable fair division for 'at most one' preferences, 2019.

[3] Gourab K Patro, Arpita Biswas, Niloy Ganguly, Krishna P. Gummadi, and Abhijnan Chakraborty. FairRec: Two-sided fairness for personalized recommendations in two-sided platforms. In *Proceedings of The Web Conference 2020*. ACM, apr 2020.

[4] Yuta Saito and Thorsten Joachims. Fair ranking as fair division: Impact-based individual fairness in ranking. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, aug 2022.

[5] Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018.