P-MMF: Provider Max-min Fairness Re-ranking in Recommender System

Chen Xu
Gaoling School of Artificial
Intelligence Renmin University of
China
xc_chen@ruc.edu.cn

Sirui Chen
School of Information Renmin
University of China
csr16@ruc.edu.cn

Jun Xu*
Gaoling School of Artificial
Intelligence Renmin University of
China
junxu@ruc.edu.cn

Weiran Shen
Gaoling School of Artificial
Intelligence Renmin University of
China
shenweiran@ruc.edu.cn

Xiao Zhang
Gaoling School of Artificial
Intelligence Renmin University of
China
zhangx89@ruc.edu.cn

Gang Wang
Huawei Noah's Ark Lab
wanggang110@huawei.com

Zhenghua Dong
Huawei Noah's Ark Lab
dongzhenhua@huawei.com

ABSTRACT

In this paper, we address the issue of recommending fairly from the aspect of providers, which has become increasingly essential in multistakeholder recommender systems. Existing studies on provider fairness usually focused on designing proportion fairness (PF) metrics that first consider systematic fairness. However, sociological researches show that to make the market more stable, max-min fairness (MMF) is a better metric. The main reason is that MMF aims to improve the utility of the worst ones preferentially, guiding the system to support the providers in weak market positions. When applying MMF to recommender systems, how to balance user preferences and provider fairness in an online recommendation scenario is still a challenging problem. In this paper, we proposed an online re-ranking model named Provider Max-min Fairness Re-ranking (P-MMF) to tackle the problem. Specifically, P-MMF formulates provider fair recommendation as a resource allocation problem, where the exposure slots are considered the resources to be allocated to providers and the max-min fairness is used as the regularizer during the process. We show that the problem can be further represented as a regularized online optimizing problem and solved efficiently in its dual space. During the online re-ranking phase, a momentum gradient descent method is designed to conduct the dynamic re-ranking. Theoretical analysis showed that the regret of P-MMF can be bounded. Experimental results on four public recommender datasets demonstrated that P-MMF can outperformed the state-of-the-art baselines. Experimental results also show that P-MMF can retain small computationally costs on a corpus with the large number of items.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

Max-min Fairness, Provider Fairness, Recommender System

1 INTRODUCTION

Raised out of social, ethical, and economic considerations, the fairness problem becomes non-negligible in recommendation [1, 2]. In multi-stakeholder recommender systems (RS), there are several different participants, including users, items, providers, etc [1]. Various models have been proposed for recommending fairly from the viewpoints of users [29], providers [8, 35], or both [45]. In this paper, we are concerned about the problem of ensuring provider fairness in multi-stakeholder recommender systems.

There are many cultural variations regarding fairness [43, 44] in sociological researches. One practical fairness definition is based on two sociological concepts: equality and equity [31]. According to Matsumoto and Juang [31], equality can be defined as: *everyone is treated the same and provided the same resources to succeed*, while equity can be defined as: *ensuring that resources are equally distributed based on needs*. Figure 1 gives an intuitive example of the two types of fairness under the recommendation scenario. Suppose we have
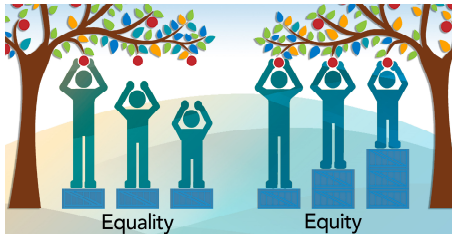
Figure 1: An intuitive example of equality and equity.[1]

some resources (i.e., exposure slots) to ensure that providers can get the apples (i.e., survival in the market). As shown in Figure 1, equality ensures that the RS evenly gives each provider the same support, while equity (known as distributive justice) emphasizes that RS will assign resources to providers as different ratios [26].

Based on the concepts of equity [25], the metrics of proportion fairness (PF) and max-min fairness (MMF) [33] have been respectively proposed. PF and MMF have been widely used for computation network and transportation [9, 17, 36]. Their formal formulations in the provider fair recommendation scenario will be explained in Section 3. Intuitively, PF and MMF tries to assign resources to the specified ratios [7]. PF is based on welfare-based principles (known as Aristotle's Principle and Nash solution [33]), which maximizes the sum of welfare of providers. MMF (known as Rawls' Principle and Kalai-Smorodinsky solution [23]) aims to improve the worst-off providers' utilities. MMF has proven to be a better metric for the provider fairness problem, because worse-off providers, who occupy the majority of the platform, can survive with these supports. Taking care of these weak providers will increase the stability of the recommender market [35]. For example, according to the report by Erickson [18], small sellers in Amazon have difficulty facing the challenges of "lower profitability" and "inability to personalize" on their own. The unfair system without ensuring MMF may result in a broken relationship between the providers and RS and, finally, force the providers to leave.

Existing provider fair recommendation models either consider the PF of providers [8, 32, 35, 47] or use heuristics to guarantee the utilities of the worse-off providers [8, 35]. However, these heuristic methods do not directly consider the MMF and lack theoretical guarantees when adapting to online scenarios. Moreover, the heuristic methods lack the flexibility to trade-off with user utility, which inevitably hurts the users' experience.

This paper aims to develop a practical re-ranking model that considers MMF from the providers' perspectives. We formulate the provider fair recommendation as a process of resource allocation [10]. In such a process, resources can be regarded as limited ranking slots, and providers are viewed as the demanders. The cost of allocation is defined as the preference of the users. Moreover, an MMF regularizer is imposed on the allocation to maximize the minimum allocation to a specific provider.

To adapt the method to the online recommendation scenarios, we proposed an online re-ranking algorithm called P-MMF that focuses on provider max-min fairness by viewing the provider fair online recommendation as a regularized online optimizing problem. However, the optimization problem contains many integral variables, making it notoriously difficult to solve. We then derive

its dual problem and propose an efficient algorithm to optimize the problem in the dual space. In the online setting, a momentum gradient descent method was developed to make an effective and efficient online recommendation. Our theoretical analysis shows that the regret of the P-MMF can be well-bounded. Furthermore, the P-MMF is also computationally efficient due to its insensitivity to the number of items.

We summarize the major contributions of this paper as follows:

(1) We analyzed the importance of ensuring equity for providers in multi-stakeholder recommender systems, and propose to use the max-min fairness metric in provider fair recommendation.

(2) We formulated the provider fair recommendation as a resource allocation problem regularized by the max-min fairness, and proposed a re-ranking model called P-MMF that balances the provider fairness and the user preference. Our theoretical analysis showed that the regret of P-MMF can be bounded.

(3) Simulations on a small dataset showed the superiority of MMF compared to PF in recommending fairly to providers. Extensive experiments on four publicly available datasets demonstrated that P-MMF outperformed state-of-the-art baselines, including the PF-based and MMF-based methods.

2 RELATED WORK

Fairness has become a hot research topic in multi-stakeholder recommender systems [1, 2]. Researchers have proposed several user-oriented and item/provider-oriented fairness re-ranking models. For user-oriented fairness, Abdollahpouri et al. [3] proposed a user-centered evaluation that measures users' different interest levels in popular items. Li et al. [29], Sacharidis [40], Serbos et al. [41] addressed the user-centered fairness from a group fairness perspective. For provider-oriented fairness, Gómez et al. [21] proposed a rule-based algorithm to ensure that the exposures of items should be equally distributed. Ge et al. [20] aimed to enhance dynamic fairness when item popularity changes over time. However, these provider-oriented methods did not consider the user's perspective.

Recently, there are also studies that jointly consider the trade-off between user preference and provider fairness. Chakraborty et al. [11] claimed that all providers should receive the amount of exposures proportional to their relevance in economy platforms. Rahmani et al. [38] studied the trade-offs between the user and producer fairness in Point-of-Interest recommendations. TFROM [47] and CPFair [32] formulated the trade-off as a knapsack problem and a relaxed linear programming problem, respectively. However, they used greedy-based algorithms in online scenarios, which only improves the proportion fairness of providers. Some studies noticed that the utilities of worse-off providers should also be guaranteed. For example, FairRec [35] and its extension FairRec+ [8] proposed hard constraints to ensure that every provider should have the lowest exposures. Welf [16] proposed a Frank-Wolfe algorithm to maximize the welfare functions of worse-off items. However, they were all designed for offline scenarios and suffered from high computational costs, which prevented them from being applied to real online recommendation systems [50, 51].

In this paper, we formulate the re-ranking task as the resource allocation problem [10], which is crucial in communications and transportation. In online resource allocation, most studies [14, 28] focused on designing the time-separable reward functions, which is the sum of rewards over periods. Li et al. [27] proposed a local-based sub-gradient algorithm for the linear reward. Cheung et al.

[13] designed a dual-based online algorithm with learning from the distribution of requests. Balseiro et al. [5] proposed a mirror-descent method to solve the fairness-regularized online allocation problem.

3 PROVIDER FAIR RE-RANKING

In this section, we first define the notations in a multi-stakeholders recommender system. Then we give the formal definitions of the proportion and max-min fairness.

In a multi-stakeholders recommender system, multiple participants exist, including users, item providers, and other stakeholders. Let $\mathcal{U}$, $\mathcal{I}$, and $\mathcal{P}$ be the set of users, items, and providers, respectively. Each item $i \in \mathcal{I}$ is associated to a unique provider $p \in \mathcal{P}$. The set of items associated with a specific provider p is denoted as $\mathcal{I}_p$. When a specific user $u \in \mathcal{U}$ accesses the recommender system, a list of K items, denoted by $L_K(u) \in \mathcal{I}^K$, is provided to the user. For each user-item pair (u, i), the recommender model estimates a preference score $s_{u,i} \in \mathbb{R}$. These items are ranked according to their preference scores. In this paper, we define the user-side utility of exposing item list $L_K(u)$ to u as the summation of the preference scores in the list, denoted by $f\left(L_K(u)\right) = \sum_{i \in L_K(u)} s_{u,i}$. We follow the literature convention [32, 35] and define the fairness vector of providers as $\mathbf{e} \in \mathbb{R}^{|\mathcal{P}|}$, where for a specific provider p, $\mathbf{e}_p \in \mathbb{R}^+$ denotes the number of exposed items of provider p. The goal of provider fair re-ranking is to compute a new fair list $L_K^F(u) \in \mathcal{I}^K$ which well balances the user utilities $f\left(L_K^F(u)\right)$ and a provider fairness metric defined over $\mathbf{e}$.

In real-world applications, the users arrive at the recommender system sequentially. Assume that at time t user u_t arrives. The recommender system needs to consider long-term provider exposure during the entire time horizon from $t = 0$ to T. Our task can be formulated as a resource allocation problem [5] with time-separable fairness. Specifically, the optimal utility of the recommender system can be defined as a time-separable utility function, which is the accumulated reward [5] over periods from 0 to T. In this case, $\mathbf{e}_p$ can be seen as the total number of exposed items of provider p, accumulated over the period 0 to T. Formally, when trading-off the user preference and provider fairness, we have the following mathematical program:

$$\max_{L_K^F} \quad \frac{1}{T} \sum_{t=1}^{T} f\left(L_K^F(u_t)\right) + \lambda r(\mathbf{e}),$$
$$\text{s.t.} \quad \mathbf{e} \le \boldsymbol{\gamma} \tag{1}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^{|\mathcal{P}|}$ denotes the weights of different providers, e.g., weighted PF or MMF [7], and $r(\mathbf{e}) \in \mathbb{R}$ is a provider fairness metric that serves as a fairness regularizer. Note that the constraint $\mathbf{e} \le \boldsymbol{\gamma}$ can also be viewed as the maximum resources allocated to the providers. Following the definitions of PF and MMF [7], $r(\mathbf{e})$ also has two different forms:

Proportion Fairness (PF): $r(\mathbf{e}) = \sum_{p \in \mathcal{P}} \log \left[1 + \mathbf{e}_p/\boldsymbol{\gamma}_p\right]$.

Max-min Fairness (MMF): $r(\mathbf{e}) = \min_{p \in \mathcal{P}} \left[\mathbf{e}_p/\boldsymbol{\gamma}_p\right]$.

Using the proportion fairness as the regularizer, we can reduce the difference between $\mathbf{e}$ and $\boldsymbol{\gamma}$, while using the max-min fairness we can improve the relative exposure (w.r.t. weights $\boldsymbol{\gamma}$) of the least exposed provider.
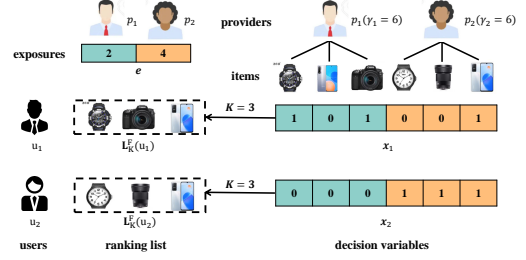

Figure 2: A toy example for time-separable fairness in recommendation

4 OUR APPROACH: P-MMF

In this section, we first formulate provider fairness in the recommendation as a resource allocation problem. Then, we propose an algorithm called P-MMF for the online recommendation.

4.1 Resource allocation with provider fairness

We formulate the providers' fair recommendation problem as a resource allocation process [10]. In such a process, the resources can be regarded as limited ranking slots that are allocated to providers. The allocation cost is defined based on the preference of the users and the max-min fairness.

Formally, based on Equation (1), the fair recommendation problem can be written as a linear programming:

$$\max_{\mathbf{x}_t} \quad \frac{1}{T} \sum_{t=1}^{T} g(\mathbf{x}_t) + \lambda r(\mathbf{e})$$
$$\text{s.t.} \quad \sum_{i \in \mathcal{I}} \mathbf{x}_{ti} = K, \quad \forall t \in [1, 2, \dots, T]$$
$$\mathbf{e}_p = \sum_{t=1}^{T} \sum_{i \in \mathcal{I}_p} \mathbf{x}_{ti}, \quad \forall p \in \mathcal{P} \tag{2}$$
$$g(\mathbf{x}_t) = \sum_{i \in \mathcal{I}} \mathbf{x}_{ti} s_{u_t,i}, \quad \forall t \in [1, 2, \dots, T]$$
$$\mathbf{e} \le \boldsymbol{\gamma}, \mathbf{x}_{ti} \in \{0, 1\}, \quad \forall i \in \mathcal{I}, \forall t \in [1, 2, \dots, T]$$

where $\mathbf{x}_t \in \{0, 1\}^{|\mathcal{I}|}$ is the decision vector for user u_t, and $g(\cdot)$ is the user-side utility function. Specifically, for each item i, $\mathbf{x}_{ti} = 1$ if it is added to the fair ranking list $L_K^F(u_t)$, otherwise, $\mathbf{x}_{ti} = 0$. Note that $g(\cdot)$ is equivalent to $f(\cdot)$ in the sense that they produce the same result, while $g(\cdot)$ takes the binary decision vector as input. The first constraint in Equation (2) ensures that the recommended lists are of size K. The second constraint in Equation (2) suggests that the exposures of each provider p are the accumulated exposures of the corresponding items over all periods. In general, we think time-separable fairness would be preferred under the scenarios with weak timeliness. For example, recommending the items with long service life (e.g., games and clothes etc.).

Figure 2 gives a toy example of the linear programming problem (2). In the example, we set $T = 2$. Suppose there are two users, u_1 and u_2, arriving at the system one by one. At each time step, the system recommends a list of $K = 3$ items. Therefore, the system has overall $2 \times 3 = 6$ slots to expose. Suppose the system has two providers p_1 and p_2, each owning three items. Let's set identical

weights for p_1 and p_2 (i.e., $\gamma = [\gamma_1 = 6, \gamma_2 = 6]$). The model solves problem (2) and the solution is two binary vectors: $\mathbf{x}_1$ for u_1 and $\mathbf{x}_2$ for u_2. Finally, we count the exposures over the time $T = 2$: $\mathbf{e} = [e_{p_1} = 2, e_{p_2} = 4]$, which means by recommending the ranking lists $L_K^F(u_1)$ (created based on $\mathbf{x}_1$) and $L_K^F(u_2)$ (created based on $\mathbf{x}_2$) to u_1 and u_2, provider p_1 and $_2$ get 2 and 4 exposures, respectively.

Although here we have already given a linear programming solution Eq.(2) to the problem, it can only be solved small-scale problems in an offline way. In online recommendation systems, for each user u_t access, the model needs to generate a fair ranking list $L_K^F(u_t)$ from large-scale item corpus immediately. This means we have no idea about the information after t. Next, we will discuss how to use MMF in the online recommendation problem.

4.2 P-MMF for online applications

In this section, firstly, we formulate the MMF in online scenarios. Next, we consider the dual of the original problem. Focusing on the dual problem has several advantages: the dual problem has significantly fewer variables, and the variables no longer need to be integers as in the original problem. Finally, we proposed a momentum gradient descent algorithm for efficient online learning.

4.2.1 Max-min fairness in online scenarios. At time step t, the recommender system receives a request from user u_t. An online algorithm h produces a real-time decision vector $\mathbf{x}_t \in \{0, 1\}^{|\mathcal{I}|}$ based on the current user u_t and the previous history $\mathcal{H}_{t-1} = \{u_s, \mathbf{x}_s\}_{s=1}^{t-1}$: $\mathbf{x}_t = h(u_t \mid \mathcal{H}_{t-1})$. We define the online reward of max-min fairness to be the summation of rewards over all time steps:

$$W = \frac{1}{T}\sum_{t=1}^{T} g(\mathbf{x}_t) + \lambda \min\left[\mathbf{e}/\gamma\right], \tag{3}$$

where $\min\left[\mathbf{e}/\gamma\right]$ corresponds to the max-min fairness regularizer.

Our goal is to design an algorithm h that attains low regret. Denote by W_{OPT} the optimal value, we measure the regret of an algorithm as the expectation difference between the optimal performance W_{OPT} of the offline problem and that of the online algorithm W over user distributions $\mathcal{U}$:

$$\text{Regret}(h) = \mathbb{E}_{u_t \sim \mathcal{U}}\left[W_{OPT} - W\right]. \tag{4}$$

4.2.2 Dual problem. From the original problem in Equation (1), we know that the integer decision variable $\mathbf{x}_t$ is of size $|\mathcal{I}|$ for each time t, which is hard to solve. However, considering its dual problem, we can significantly reduce its computational cost.

THEOREM 1 (DUAL PROBLEM). *The dual problem of Equation* (1) *can be written as:*

$$W_{OPT} \le W_{Dual} = \min_{\boldsymbol{\mu} \in \mathcal{D}}\left[g^*(\mathbf{A}\boldsymbol{\mu}) + \lambda r^*(-\boldsymbol{\mu})\right], \tag{5}$$

where $\mathbf{A} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{P}|}$ *is the item-provider adjacent matrix, and* $A_{ip} = 1$ *indicates item* $i \in \mathcal{I}_p$, *and 0 otherwise. Letting* $\mathcal{X} = \{\mathbf{x}_t | \mathbf{x}_t \in \{0, 1\} \land \sum_{i \in \mathcal{I}} \mathbf{x}_{ti} = K\}$, $g^*(\cdot), r^*(\cdot)$ *are the conjugate functions:*

$$g^*(c) = \max_{\mathbf{x}_t \in \mathcal{X}}\sum_{t=1}^{T}\left[g(\mathbf{x}_t)/T - c^\top \mathbf{x}_t\right], \, r^*(-\boldsymbol{\mu}) = \max_{\mathbf{e} \le \gamma}\left[r(\mathbf{e}) + \boldsymbol{\mu}^\top \mathbf{e}/\lambda\right],$$

and $\mathcal{D} = \{\boldsymbol{\mu} | r^*(-\boldsymbol{\mu}) < \infty\}$ *is the feasible region of dual variable* $\boldsymbol{\mu}$.

Proof of Theorem 1 can be found in Appendix A.1. From Theorem 1, we can have a new non-integral decision variable $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{P}|}$. In practice, the provider size $|\mathcal{P}| \ll |\mathcal{I}|$. Besides, due to $\mathbf{A}$'s sparsity, it is very efficient to compute $\mathbf{A}\boldsymbol{\mu}$, which aims to project the variable $\boldsymbol{\mu}$ from provider space into item spaces.

In our online algorithm discussed in Section 4.2.3, we can have a closed form of the conjugate function $g^*(\cdot)$ in constant time. As for the feasible region $\mathcal{D}$ and the conjugate function $r^*(\cdot)$, we have Theorem 2 and Lemma 1.

THEOREM 2 (DUAL FEASIBLE REGION). *In the MMF, the feasible region of the dual problem*

$$\mathcal{D} = \left\{\boldsymbol{\mu} \, \middle| \, \sum_{p \in \mathcal{S}} \gamma_p \mu_p \ge -\lambda, \forall \mathcal{S} \in \mathcal{P}_s\right\},$$

where $\mathcal{P}_s$ *is the power set of* $\mathcal{P}$, *i.e., the set of all the subsets of* $\mathcal{P}$.

The proof of Theorem 2 can be found in Appendix A.1, which implies the following Lemma 1:

LEMMA 1. *The conjugate function* $r^*(\cdot)$ *has a closed form:* $\max_{\boldsymbol{\mu} \le \gamma} r^*(-\boldsymbol{\mu}) = \gamma^T \boldsymbol{\mu}/\lambda + 1$, *and the optimal dual variable is:* $\arg\max_{\boldsymbol{\mu} \le \gamma} r^*(-\boldsymbol{\mu}) = \gamma/\lambda$.

4.2.3 The P-MMF algorithm. Algorithm 1 illustrates P-MMF algorithm. Following Balseiro et al. [5], P-MMF keeps a dual variable $\boldsymbol{\mu}_t$, the remaining resources $\boldsymbol{\beta}_t$ and the gradient $\mathbf{g}_t$ for each time t.

Whenever a user arrives, the algorithm computes the recommended variable $\mathbf{x}_t$ based on the remaining resources and the dual variable $\boldsymbol{\mu}_t$ (line 7). The final dual variable is estimated as the average dual variable for each time t: $\boldsymbol{\mu} = \sum_{t=1}^{T} \boldsymbol{\mu}_t/T$. Intuitively, for $\boldsymbol{\mu}_t$, when the values of dual variables are higher, the algorithm naturally recommends fewer items related to the corresponding provider. The remaining resources $\boldsymbol{\beta}_t$ ensure that the algorithm only recommends items from providers with remaining resources. Note that in line 7, the formulation is linear with respect to $\mathbf{x}_t$. Therefore, it is efficient to compute $\mathbf{x}_t$ through a top-K sort algorithm in constant time.

The online learning process is as follows. Firstly, we get the closed form of the conjugate function of max-min regularizer $r^*(-\boldsymbol{\mu}_t)$ according to Lemma 1. Then we can get the subgradient of the dual function $g^*(\mathbf{A}\boldsymbol{\mu}) + \lambda r^*(-\boldsymbol{\mu})$:

$$-\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t \in \partial\left(g^*(\mathbf{A}\boldsymbol{\mu}_t) + \lambda r^*(-\boldsymbol{\mu}_t)\right).$$

We also add the last time momentum [37] to the updated gradient $\mathbf{g}_t$. Finally, we utilize $\mathbf{g}_t$ to update the dual variable by performing the online descent in line 14, where we used weighted $\ell_2 - norm$: $\|\boldsymbol{\mu}\|_{\gamma^2}^2 = \sum_{j=1}^{|\mathcal{P}|} \gamma_j^2 \mu_j^2$. Therefore, the dual variable will move towards the directions of the providers with fewer exposures, and the primal variable $\mathbf{x}_t$ will move to a better solution.

Note that the projection step in line 14 can be efficiently solved using convex optimization solvers [5] since $\mathcal{D}$ is coordinate-wisely symmetric.

$$\min_{\boldsymbol{\mu} \in \mathcal{D}} \|\boldsymbol{\mu} - \boldsymbol{\mu}_t\|_{\gamma^2}^2 = \min_{\sum_{p \in \mathcal{P}}} (\mu_p \gamma_p - \widetilde{\mu}_p \gamma_p)^2$$

$$\text{s.t.} \sum_{j=1}^{m} \gamma_j \widetilde{\mu}_j + \lambda \ge 0, \, \forall m = 1, 2, \dots, |\mathcal{P}|,$$

where $\widetilde{\boldsymbol{\mu}}$ satisfies: $\gamma_1 \widetilde{\mu}_1 \le \gamma_2 \widetilde{\mu}_2 \le \cdots, \le \gamma_{|\mathcal{P}|} \widetilde{\mu}_{|\mathcal{P}|}$.

Algorithm 1: Online learning of P-MMF

Input: User arriving order $\{u_i\}_{i=1}^N$, time-separate size T, ranking size K, user-item preference score $s_{u,i}, \forall u, i$, item-provider adjacent matrix $\mathbf{A}$, maximum resources $\boldsymbol{\gamma}$ and the trade-off coefficient λ.

Output: The decision variables $\{\mathbf{x}_i, i = 1, 2, \ldots, N\}$

1: **for** $n = 1, \cdots, N/T$ **do**
2:    Initialize dual solution $\boldsymbol{\mu}_1 = 0$, remaining resources $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}$, and momentum gradient $\mathbf{g}_0 = 0$.
3:    **for** $t = 1, \cdots, T$ **do**
4:       User u_{nT+t} arrives
5:       $\mathbf{m}_p = \begin{cases} 0, & \boldsymbol{\beta}_{tp} > 0 \\ \infty, & \text{otherwise} \end{cases}$
6:       `// Make the recommendation:`
7:       $\mathbf{x}_t = \arg\max_{\mathbf{x}_t \in \mathcal{X}} \left[g(\mathbf{x}_t)/T - (\mathbf{A}(\boldsymbol{\mu}_t + \mathbf{m}))^\top \mathbf{x}_t \right]$
8:       $\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t - \mathbf{A}^\top \mathbf{x}_t$
9:       $\mathbf{e}_t = \arg\max_{\mathbf{e}_t \le \boldsymbol{\beta}_t} r^*(-\boldsymbol{\mu}_t)$
10:      $\widetilde{\mathbf{g}}_t = -\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t$
11:      $\mathbf{g}_t = \alpha\widetilde{\mathbf{g}}_t + (1 - \alpha)\mathbf{g}_{t-1}$
12:      $\boldsymbol{\mu}_{t+1} = \arg\min_{\boldsymbol{\mu} \in \mathcal{D}} \left[\langle \mathbf{g}_t, \boldsymbol{\mu} \rangle + \eta \|\boldsymbol{\mu} - \boldsymbol{\mu}_t\|_{\boldsymbol{\gamma}^2}^2 \right]$
13:    **end for**
14: **end for**

We provide a regret bound on Regret(h) in Theorem 3 but defer the proof to Appendix A.3 due to space limit. Intuitively, larger ranking size K and time T will lead to larger biases in P-MMF.

THEOREM 3 (REGRET BOUND). *Assume that the function $\|\cdot\|_{\boldsymbol{\gamma}^2}^2$ is σ-strong convex and there exists a constant $G \in \mathbb{R}^+$ and $H > 0$ such that $\|\widetilde{g}_t\| < G$, $\|\boldsymbol{\mu}_t - \boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2 \le H$.. Then, the regret can be bounded as follows:*

$$Regret(h) \le \frac{K(1 + \lambda\bar{r} + \bar{r})}{\min_p \gamma_p} + \frac{H}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(T - 1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}, \tag{6}$$

where $\bar{r}$ is the upper bound of MMF regularzier, and in practice, $\bar{r} \le 1$.

Setting the learning rate as $\eta = O(T^{-1/2})$, we can obtain a sub-linear regret upper bound of order $O(T^{1/2})$ of magnitude.

5 EXPERIMENT

We conducted experiments to show the effectiveness of the proposed P-MMF for provider-fair recommendations. The source code and experiments have been shared at github [2].

5.1 Experimental settings

5.1.1 Datasets. The experiments were conducted on four large-scale, publicly available recommendation datasets, including:

Yelp[3]: a large-scale businesses recommendation dataset. We only utilized the clicked data, which is simulated as the 4-5 star rating samples. The cities of the items are considered as providers. It has 154543 samples, which contains 17034 users, 11821 items and 23 providers.

Amazon-Beauty/Amazon-Baby: Two subsets (beauty and digital music domains) of Amazon Product dataset[4]. We only utilized the clicked data, which is simulated as the 4-5 star rating samples. Also, the brands are considered as providers. They has 49217/59836 samples, which contains 9625/11680 users, 2756/2687 items and 1024/112 providers.

Steam[5] [24]: We used the data for gamed played for more than 10 hours in our experiments. The publishers of games are considered as providers. It has 29530 samples, which contains 5902 users, 591 items and 81 providers.

As a pre-processing step, the users, items, and providers who interacted with less than 5 items/users were removed from all dataset to avoid the extremely sparse cases. We also removed the providers associated with less than 5 items.

5.1.2 Evaluation. We sorted all the interactions according to the time and used the first 80% of the interactions as the training data to train the base model (i.e., BPR [39]). The remaining 20% of interactions were used as the test data for evaluation. Based on the trained base model, we can obtain a preference score $s_{u,i}$ for each user-item pair (u, i). The chronological interactions in the test data were split into interaction sequences where the horizon length was set to T. We calculated the metrics separately for each sequence, and the averaged results are reported as the final performances.

As for the evaluation metrics, the performances of the models were evaluated from three aspects: user-side preference, provider-side fairness, and the trade-off between them. As for the user-side preference, following the practices in [47], we utilized the NDCG@K, which is defined as the ratio between the sum of position-based user-item scores [47] in the original ranking list $\mathbf{L}_K(u_t)$ and that in the re-ranked list $\mathbf{L}_K^F(u_t)$:

$$\text{NDCG@K} = \frac{1}{T} \sum_{t=1}^T \frac{\sum_{i \in \mathbf{L}_K(u_t)} s_{u_t,i}/\log(\text{rank}_i + 1)}{\sum_{i \in \mathbf{L}_K^F(u_t)} s_{u_t,i}/\log(\text{rank}_i^F + 1)},$$

where rank_i and rank_i^F are the ranking positions of the item i in $\mathbf{L}_K(u_t)$ and $\mathbf{L}_K^F(u_t)$, respectively.

As for the provider fairness, we directly utilized the definition of MMF in Section 3 as the metric:

$$\text{MMF@K} = \min_{p \in \mathcal{P}} \left\{ \sum_{t=1}^T \sum_{i \in \mathbf{L}_K^F(u_t)} \mathbb{I}(i \in \mathcal{I}_p)/\gamma_p \right\},$$

where $\mathbb{I}(\cdot)$ is the indicator function.

As for the trade-off performance, we used the online objective traded-off with Equation (3):

$$W_\lambda @K = \frac{1}{T} \sum_{t=1}^T \left(\sum_{i \in \mathbf{L}_K^F(u_t)} s_{u_t,i} \right) + \lambda \cdot \text{MMF@K},$$

where $\lambda \ge 0$ is the trade-off coefficient.

5.1.3 Baselines. The following representative provider fair re-ranking models were chosen as the baselines: **FairRec** [35] and **FairRec+** [8] aimed to guarantee at least Maximin Share (MMS) of the provider exposures. **CPFair** [32] formulated the trade-off problem as a knapsack problem and proposed a greedy solution.

We also chose the following MMF models: **Welf** [16] use the Frank-Wolfe algorithm to maximize the Welfare functions of worst-off items. However, it is developed under off-line settings; **RAOP** [5] is a state-of-the-art online resource allocation method. We applied it to the recommendation by regarding the items as the resources and users as the demanders.

We also compared the proposed P-MMF with two heuristic MMF baselines: **K-neighbor**: at each time step t, only the items associated to the top-K providers with the least cumulative exposure are recommended; **min-regularizer**: at each time step t, we add a regularizer that measures the exposure gaps between the target provider and the worst-providers. Appendix A.4 shows the algorithm details.

5.1.4 Implementation details. As for the hyper-parameters in all models, the learning rate was tuned among $[1e-2, 1e-4]/T^{1/2}$ and the momentum coefficient α was tuned among $[0.2, 0.6]$. For the maximum resources (i.e., the weights) γ, following the practices in [47], we set γ based on the number of items provided by the providers: $\gamma_p = KT\eta|\mathcal{I}_p|/|\mathcal{I}|$, where η is the factor controlling the richness of resources. In all the experiments, we set $\eta = 1 + 1/|\mathcal{P}|$. We implemented P-MMF with both CPU and GPU versions based on cvxpy [15] and its PyTorch version [4], respectively.

5.2 Simulation on a small dataset

Note that Problem (2) can use MMF or PF as its regularizer, though directly solving it on large-scale datasets is difficult. To verify the correctness of the formulation and to investigate different impacts of MMF-based and PF-based regularizers on provider fair recommendation, we first conducted a numerical simulation using 5% of the Yelp data, which consists of 844 users, 813 items, 10 providers, and with the length of the horizon $T = 256$.

Specifically, we solved the Problem (2) with PF based or MMF based regularizers (i.e., setting their provider weights as even distribution $\gamma_p = KT, \forall p \in \mathcal{P}$), using the solver cvxpy [15]. At each time t, after receiving the recommendation decision variable $\mathbf{x}_t$, we calculated the overall provider exposure $\mathbf{e}$ on the dataset. Then the Lorenz curves [19] of provider exposures were drawn and shown in Figure 3. The Lorenz curve is often used to represent exposure distribution. Here it shows the proportion of overall exposure percentage assumed by the bottom $x\%$ providers. In other words, for the bottom $x\%$ providers, what percentage ($y\%$) of the total exposures they have. Note that the diagonal line to the upper right is known as the absolute fair line.

From the curves shown in Figure 3, we can observe that when the $\lambda \to \infty$ (i.e., only consider provider's fairness), both PF and MMF can achieve the expected proportion γ (i.e., even exposures here). However, after considering the user preference, the MMF-based regularizer tends to consider the worst providers' exposure first, while PF does not. For example, when the trade-off coefficient λ changes among $[0.01, 0.05, 0.1]$, 60% of the least exposure providers will increase $[18.81\%, 18.96\%, 12.58\%]$ exposures by MMF. If using PF-based regularizer, then the increased exposure ratios become
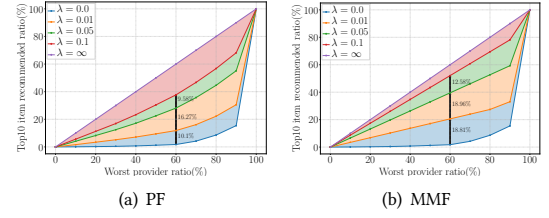

Figure 3: The Lorenz curves of provider exposure with PF regularizer (a) and MMF regularizer (b), respectively. The experiments were conducted based on 5% of the Yelp dataset.

$[10.1\%, 16.27\%, 9.58\%]$. The results verified that formulating recommendations as a resource allocation problem regularized by MMF leads to better provider-fair recommendations than PF.

5.3 Experimental results on full datasets

In this section, we conducted the experiments with the online algorithm developed in Section 4.2. In all of the experiments, BPR [39] was chosen as the base ranking model for generating preference scores. We set the length of the horizon $T = 256$.

Table 1 reports the experimental results of P-MMF and the baselines on all datasets in terms of the metric $W_1@K$. Underlined numbers mean the best-performed baseline. To make fair comparisons, all the baselines were tuned and used $W_1@K$ as the evaluation metric. Note that similar experiment phenomena has also been observed on other λ values.

From the reported results, we found that P-MMF outperformed all of the PF-based baselines in terms of $W_1@K$ ($K = 5, 10, 20$), verified that P-MMF can give supports to the poor-conditioned providers. We also observed that P-MMF outperformed all the MMF-based baselines, indicating P-MMF's effectiveness in enhancing provider fairness while keeping high user preference.

Figure 4 shows the Pareto frontiers [30] of NDCG@K and MMF@K on four datasets with different ranking size K. The Pareto frontiers were drawn by tuning different parameters of the models and choosing the (NDCG@K, MMF@K) points with the best performances. In the experiment, we selected the baselines of CPFair, min-regularizer, Welf, and ROAP, which achieved relatively good performances among all baselines.

From the Pareto frontiers, we can see that the proposed P-MMF Pareto dominated the baselines (i.e., the P-MMF curves are at the upper right corner). Pareto dominance means that under the same NDCG@K level, P-MMF achieved better MMF@K; Under the same MMF@K level, P-MMF achieved better NDCG@K. The results demonstrate that P-MMF splendidly improves the utilities of poor-conditioned providers without sacrificing users' utilities too much.

5.4 Experiment analysis

We also conducted experiments to analyze P-MMF on Yelp.

5.4.1 Ablation study on different base models. P-MMF and other baselines are re-ranking models which re-rank the results outputted by a base recommender model. To verify the effectiveness of P-MMF with different base ranking models, we choose three other base models to generate user-item preference scores $s_{u,i}$, including

Table 1: Performance comparisons between ours and the baselines on Yelp, Beauty, Baby, and Steam. We choose trade-off co-efficient $\lambda = 1$ on three different ranking sizes K to investigate the effectiveness of ours.

Dataset	Metric	PF-based baselines			MMF-based baselines				Our approach	
		FairRec	FairRec+	CPFair	k-neighbor	Welf	min-regularizer	ROAP	P-MMF(ours)	Improv.
Yelp	$W_1@5$	2.659	2.955	3.366	2.929	2.986	3.296	3.395	**3.455**	1.8%
	$W_1@10$	5.046	5.335	6.124	5.345	5.639	6.125	6.126	**6.247**	2.0%
	$W_1@20$	10.256	10.860	11.370	10.829	11.648	11.600	11.640	**11.797**	1.3%
Amazon-Beauty	$W_1@5$	2.617	3.783	4.959	3.958	4.820	4.953	4.863	**5.034**	1.5%
	$W_1@10$	5.197	7.244	9.195	7.312	9.166	9.309	9.147	**9.434**	1.3%
	$W_1@20$	10.312	13.492	17.686	13.474	17.650	17.880	17.637	**17.983**	0.6%
Amazon-Baby	$W_1@5$	2.696	3.377	4.129	3.438	4.344	4.556	4.303	**4.583**	0.6%
	$W_1@10$	5.388	6.499	8.080	6.384	8.187	8.345	8.137	**8.440**	1.1%
	$W_1@20$	10.591	12.193	15.527	11.871	15.480	15.713	15.544	**15.873**	1.1%
Steam	$W_1@5$	2.953	3.173	3.994	2.992	3.947	4.005	3.915	**4.443**	10.9%
	$W_1@10$	5.486	5.716	7.736	5.479	7.678	8.033	7.893	**8.148**	1.4%
	$W_1@20$	9.915	12.699	14.864	9.724	14.733	15.039	15.081	**15.235**	1.0%

(a) Amazon Beauty $K = 5$ (b) Amazon Baby $K = 5$ (c) Yelp $K = 5$ (d) Steam $K = 5$

(e) Amazon Beauty $K = 10$ (f) Amazon Baby $K = 10$ (g) Yelp $K = 10$ (h) Steam $K = 10$

(i) Amazon Beauty $K = 20$ (j) Amazon Baby $K = 20$ (k) Yelp $K = 20$ (l) Steam $K = 20$
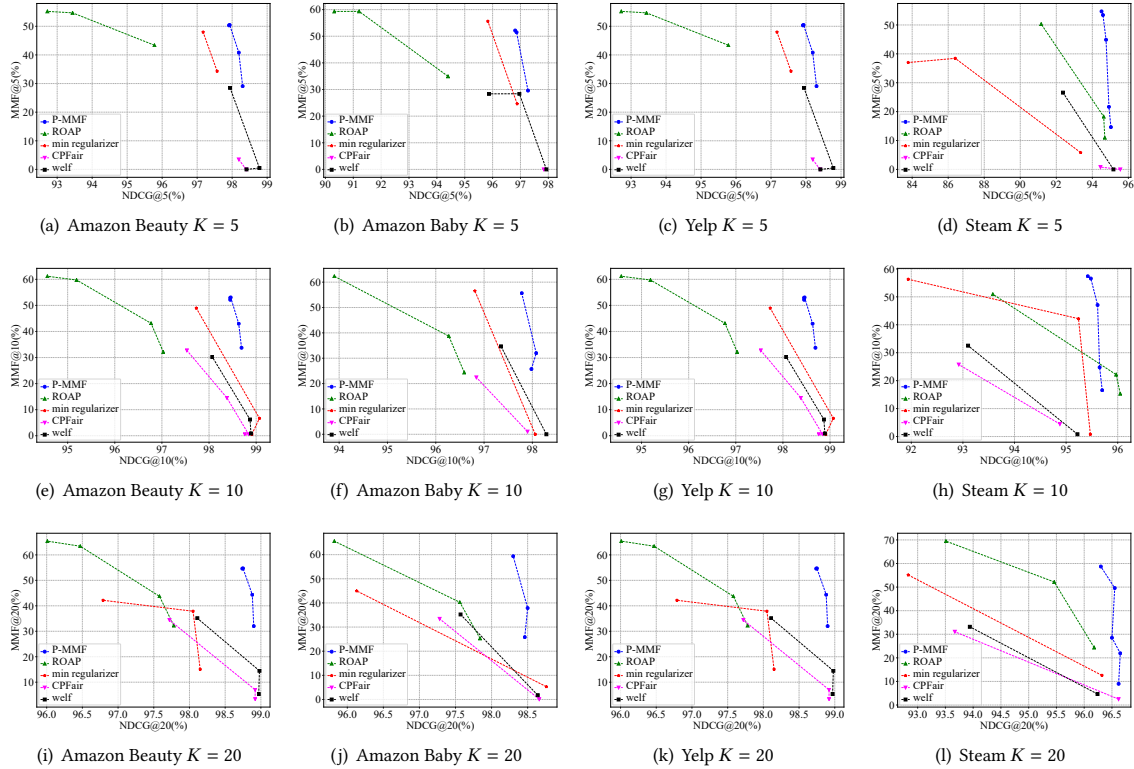
Figure 4: Pareto frontier of four different datasets with different top-K ranking.

DMF [48] which optimizes the matrix factorization with the deep neuarl networks; **LINE** [42] is a matrix factorization model based on graph embeddings; and **LightGCN** [22] which builts a user-item interaction graph and adapts Graph Convolutional Network [49] to conduct recommendation. All the experiments were also conducted on the full Yelp dataset with $T = 256$.
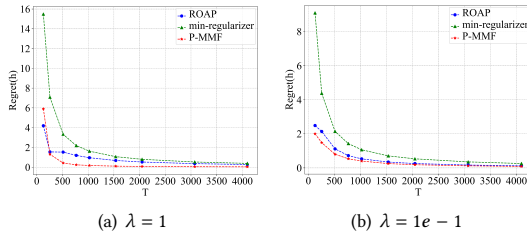
Table 2 shows the experimental results of P-MMF with different base models. We observed that P-MMF consistently outperformed the PF-based and MMF-based baselines. The results also verified

that P-MMF is more effective than the baselines in re-ranking the results outputted by different base models.

5.4.2 Regret bound experiments. To further show the effectiveness of the P-MMF, we directly compute the regret $\text{Regret}(h) = W_{OPT} - W$, where W_{OPT} is the oracle performance defined in Equation (1), and W is the online performance defined in Equation (3). Due to the huge computational cost of obtaining W_{OPT} on large-scale datasets, we conducted the experiments on the 5% of the Yelp data

Table 2: Performance comparisons between ours and the baselines on Yelp with different base recommender models. We choose trade-off co-efficient $\lambda = 1$ on three different ranking sizes K to investigate the effectiveness of ours.

Base Model	Metric	PF-based baselines			MMF-based baselines				Our approach	
		FairRec	FairRec+	CPFair	k-neighbor	Welf	min-regularizer	ROAP	**P-MMF(ours)**	Improv.
DMF [48]	W_1@5	3.1982	4.087	5.070	4.114	4.981	<u>5.198</u>	5.190	**5.272**	1.4%
	W_1@10	6.364	7.829	9.929	7.957	9.867	<u>9.968</u>	9.878	**10.003**	0.3%
	W_1@20	12.643	15.148	19.372	15.208	19.292	<u>19.410</u>	19.324	**19.448**	0.2%
LINE [42]	W_1@5	2.323	4.074	5.237	4.269	5.153	<u>5.239</u>	5.159	**5.407**	3.2%
	W_1@10	4.614	7.866	10.087	8.099	10.064	<u>10.124</u>	10.042	**10.280**	1.5%
	W_1@20	9.164	15.041	19.814	15.829	<u>19.826</u>	19.450	19.723	**19.973**	0.7%
LightGCN [22]	W_1@5	2.729	3.714	4.531	3.963	4.633	<u>4.782</u>	4.718	**4.837**	1.2%
	W_1@10	5.468	7.2368	8.982	7.584	8.975	9.016	<u>9.144</u>	**9.283**	1.5%
	W_1@20	10.927	13.995	17.799	14.856	17.774	17.812	<u>17.871</u>	**18.035**	0.9%

(a) $\lambda = 1$ (b) $\lambda = 1e - 1$

Figure 5: Regret versus time-separate length T.

(a) CPU version   (b) GPU version

Figure 6: Online inference time per user access under CPU and GPU implementations w.r.t. the number of total items.

created in Section 5.2. We compared P-MMF with the state-of-the-art provider-fair online MMF baselines: min-regularizer and ROAP. The experiments were conducted on ranking size $K = 10$. Note that in the experiment, we fixed the user arriving size N, and the regret is computed through the summation over the N/T samples where T is the length of the horizon. According the Theorem 3, the summation Regret(h) is comparable with $O(T^{1/2}N/T) = O(N/T^{1/2})$.

Figure 5 shows the summation Regret(h) curves w.r.t. T. Figure (a) and (b) show the curves when the trade-off co-efficient λ was set as 1 and 0.1, respectively. From Figure 5, we can see that P-MMF has lower regret bound than other online models, especially when λ is large. Moreover, we can see that the regret bound of P-MMF is decreasing with the increase of T, which verified the theoretical analysis results. Similar results have also been observed for other λ values.

5.4.3 Online inference time. We experimented with investigating the online inference time of P-MMF and the most practical RS model DMF [48]. In real recommender systems, the number of providers is relatively small and steady. However, the number of users and items is usually huge and grows rapidly. Therefore, we tested the inference time under CPU and GPU implementations of P-MMF and DMF w.r.t. the different number of items while keeping the number of users and providers unchanged. The GPU implementation is based on PyTorch [34].

Figure 6 reports the curves of inference time (ms) per user access w.r.t. item size. We can see that P-MMF with CPU and GPU versions need only about 20-40ms and 17-18ms for online inference, respectively. Moreover, we can see that the inference time of P-MMF did not increase much with the increasing number of

items. For example, by increasing the item size from 0 to 200000, the P-MMF CPU version only needs a little bit more time (19ms) for online inference. The inference time for the GPU version almost kept unchanged. As for comparisons, DMF's inference time increased rapidly: both CPU and GPU versions need much more time (about 65ms). The phenomenon can be easily explained with the dual problem analysis in Theorem 1. We see the parameter size of P-MMF is provider size, which is far less than the item size $|\mathcal{P}| \ll |\mathcal{I}|$. Therefore, the online inference time is not sensitive to item numbers. We conclude that P-MMF can be adapted to the real online recommendation scenarios efficiently because of its low and robust online computational cost, even when the number of items grows rapidly.

6 CONCLUSION

This paper proposes a novel re-ranking model called P-MMF that aims to balance provider fairness and user preference. Firstly, We formalize the task as a resource allocation problem regularized by a max-min fairness metric. Then, to adapt the online scenarios in the recommendation, we proposed a momentum gradient descent method to conduct online learning for resource allocation in the dual space. Theoretical analysis showed that the regret of P-MMF is bounded. Extensive experimental results on four large-scale datasets demonstrated that P-MMF outperformed the baselines and Pareto dominated the state-of-the-art provider fair baselines.

ACKNOWLEDGMENTS

REFERENCES

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. 2020. Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction* 30, 1 (2020), 127–158.

[2] Himan Abdollahpouri and Robin Burke. 2019. Multi-stakeholder recommendation and its connection to multi-sided fairness. *arXiv preprint arXiv:1907.13158* (2019).

[3] Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. 2019. The unfairness of popularity bias in recommendation. *arXiv preprint arXiv:1907.13286* (2019).

[4] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J Zico Kolter. 2019. Differentiable convex optimization layers. *Advances in neural information processing systems* 32 (2019).

[5] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. 2021. Regularized online allocation problems: Fairness and beyond. In *International Conference on Machine Learning*. PMLR, 630–639.

[6] Santiago R. Balseiro, Haihao Lu, Vahab Mirrokni, and Balasubramanian Sivan. 2022. From Online Optimization to PID Controllers: Mirror Descent with Momentum. https://doi.org/10.48550/ARXIV.2202.06152

[7] Dimitris Bertsimas, Vivek F Farias, and Nikolaos Trichakis. 2011. The price of fairness. *Operations research* 59, 1 (2011), 17–31.

[8] Arpita Biswas, Gourab K Patro, Niloy Ganguly, Krishna P Gummadi, and Abhijnan Chakraborty. 2021. Toward Fair Recommendation in Two-sided Platforms. *ACM Transactions on the Web (TWEB)* 16, 2 (2021), 1–34.

[9] Thomas Bonald, Laurent Massoulié, Alexandre Proutiere, and Jorma Virtamo. 2006. A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing systems* 53, 1 (2006), 65–84.

[10] Joseph L Bower. 1972. *Managing the resource allocation process: A study of corporate planning and investment.* Irwin Homewood.

[11] Abhijnan Chakraborty, Aniko Hannak, Asia J Biega, and Krishna Gummadi. 2017. Fair sharing for sharing economy platforms. In *Fairness, Accountability and Transparency in Recommender Systems-Workshop on Responsible Recommendation*.

[12] Lei Chen. 2021. *Deep Learning and Practice with MindSpore.* Springer Nature.

[13] Wang Chi Cheung, Guodong Lyu, Chung-Piaw Teo, and Hai Wang. 2020. Online planning with offline simulation. *Available at SSRN 3709882* (2020).

[14] Nikhil R Devanur, Kamal Jain, Balasubramanian Sivan, and Christopher A Wilkens. 2011. Near optimal online algorithms and fast approximation algorithms for resource allocation problems. In *Proceedings of the 12th ACM conference on Electronic commerce*. 29–38.

[15] Steven Diamond and Stephen Boyd. 2016. CVXPY: A Python-embedded modeling language for convex optimization. *The Journal of Machine Learning Research* 17, 1 (2016), 2909–2913.

[16] Virginie Do, Sam Corbett-Davies, Jamal Atif, and Nicolas Usunier. 2021. Two-sided fairness in rankings via Lorenz dominance. *Advances in Neural Information Processing Systems* 34 (2021), 8596–8608.

[17] VR Dondeti and H Emmons. 1996. Max-Min matching problems with multiple assignments. *Journal of optimization theory and applications* 91, 2 (1996), 491–511.

[18] KJ Erickson. 2018. In Their Own Words: Why Sellers Are Fed Up With Amazon. https://medium.com/public-market/in-their-own-words-why-sellers-are-fed-up-with-amazon-e97da44f7f18

[19] Joseph L Gastwirth. 1971. A general definition of the Lorenz curve. *Econometrica: Journal of the Econometric Society* (1971), 1037–1039.

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

[21] Elizabeth Gómez, Ludovico Boratto, and Maria Salamó. 2022. Provider fairness across continents in collaborative recommender systems. *Information Processing & Management* 59, 1 (2022), 102719.

[22] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*. 639–648.

[23] Ehud Kalai and Meir Smorodinsky. 1975. Other solutions to Nash's bargaining problem. *Econometrica: Journal of the Econometric Society* (1975), 513–518.

[24] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*. IEEE, 197–206.

[25] Julian Lamont. 2017. *Distributive justice.* Routledge.

[26] Johannes Langemeyer and James JT Connolly. 2020. Weaving notions of justice into urban ecosystem services research and practice. *Environmental science & policy* 109 (2020), 1–14.

[27] Xiaocheng Li, Chunlin Sun, and Yinyu Ye. 2020. Simple and fast algorithm for binary integer and online linear programming. *Advances in Neural Information Processing Systems* 33 (2020), 9412–9421.

[28] Xiaocheng Li and Yinyu Ye. 2021. Online linear programming: Dual convergence, new algorithms, and regret bounds. *Operations Research* (2021).

[29] Yunqi Li, Hanxiong Chen, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2021. User-oriented fairness in recommendation. In *Proceedings of the Web Conference 2021*. 624–632.

[30] Alexander V Lotov and Kaisa Miettinen. 2008. Visualizing the Pareto frontier. In *Multiobjective optimization*. Springer, 213–243.

[31] David Matsumoto and Linda Juang. 2016. *Culture and psychology.* Cengage Learning.

[32] Mohammadmehdi Naghiaei, Hossein A Rahmani, and Yashar Deldjoo. 2022. Cpfair: Personalized consumer and producer fairness re-ranking for recommender systems. *arXiv preprint arXiv:2204.08085* (2022).

[33] John F Nash Jr. 1950. The bargaining problem. *Econometrica: Journal of the econometric society* (1950), 155–162.

[34] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in PyTorch. (2017).

[35] Gourab K Patro, Arpita Biswas, Niloy Ganguly, Krishna P Gummadi, and Abhijnan Chakraborty. 2020. Fairrec: Two-sided fairness for personalized recommendations in two-sided platforms. In *Proceedings of The Web Conference 2020*. 1194–1204.

[36] Larry L Peterson and Bruce S Davie. 2007. *Computer networks: a systems approach.* Elsevier.

[37] Ning Qian. 1999. On the momentum term in gradient descent learning algorithms. *Neural networks* 12, 1 (1999), 145–151.

[38] Hossein A Rahmani, Yashar Deldjoo, Ali Tourani, and Mohammadmehdi Naghiaei. 2022. The unfairness of active users and popularity bias in point-of-interest recommendation. *arXiv preprint arXiv:2202.13307* (2022).

[39] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2012. BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618* (2012).

[40] Dimitris Sacharidis. 2019. Top-n group recommendations with fairness. In *Proceedings of the 34th ACM/SIGAPP symposium on applied computing*. 1663–1670.

[41] Dimitris Serbos, Shuyao Qi, Nikos Mamoulis, Evaggelia Pitoura, and Panayiotis Tsaparas. 2017. Fairness in package-to-group recommendations. In *Proceedings of the 26th international conference on world wide web*. 371–379.

[42] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015. Line: Large-scale information network embedding. In *Proceedings of the 24th international conference on world wide web*. 1067–1077.

[43] Tom R Tyler and E Allan Lind. 2002. Procedural justice. In *Handbook of justice research in law*. Springer, 65–92.

[44] Tom R Tyler and Heather J Smith. 1995. Social justice and social movements. (1995).

[45] Lequn Wang and Thorsten Joachims. 2021. User Fairness, Item Fairness, and Diversity for Rankings in Two-Sided Markets. In *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*. 23–41.

[46] David Williams. 1991. *Probability with martingales.* Cambridge university press.

[47] Yao Wu, Jian Cao, Guandong Xu, and Yudong Tan. 2021. Tfrom: A two-sided fairness-aware recommendation model for both customers and providers. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1013–1022.

[48] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep matrix factorization models for recommender systems.. In *IJCAI*, Vol. 17. Melbourne, Australia, 3203–3209.

[49] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. 2019. Graph convolutional networks: a comprehensive review. *Computational Social Networks* 6, 1 (2019), 1–23.

[50] Xiao Zhang, Sunhao Dai, Jun Xu, Zhenhua Dong, Quanyu Dai, and Ji-Rong Wen. 2022. Counteracting user attention bias in music streaming recommendation via reward modification. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2504–2514.

[51] Xiao Zhang, Haonan Jia, Hanjing Su, Wenhan Wang, Jun Xu, and Ji-Rong Wen. 2021. Counterfactual reward modification for streaming recommendation with delayed feedback. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 41–50.

A APPENDIX

A.1 Proof of Theorem 1

PROOF. For max-min fairness, we have the regularizer as $r(\mathbf{e}) = \min_{p \in \mathcal{P}} (\mathbf{e}_p / \boldsymbol{\gamma}_p)$, we can easily proof that the exposure vector $\mathbf{e}$ can be represented as the dot-product between decision varible $\mathbf{x}_t$ and the item-provider adjacent matrix $\mathbf{A}$: $\mathbf{e} = \sum_{t=1}^{T} (\mathbf{A}^\top \mathbf{x}_t)$. Then we treat the $\mathbf{e}$ as the auxiliary variable, and the Equation (1) can be written as:

$$W_{OPT} = \max_{\mathbf{x}_t \in \mathcal{X}, \mathbf{e} \leq \boldsymbol{\gamma}} \left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}) \right]$$

$$s.t. \mathbf{e} = \sum_{t=1}^{T} (\mathbf{A}^\top \mathbf{x}_t),$$

where $\mathcal{X} = \{ \mathbf{x}_t | \mathbf{x}_t \in {0,1} \wedge \sum_{i \in I} \mathbf{x}_{ti} = K \}$. Then we move the constraints to the objective using a vector of Lagrange multipliers $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{P}|}$:

$$W_{OPT} = \max_{\mathbf{x}_t \in \mathcal{X}, \mathbf{e} \leq \boldsymbol{\gamma}} \min_{\boldsymbol{\mu} \in \mathcal{D}} \left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}) - \boldsymbol{\mu}^\top \left(-\mathbf{e} + \sum_{t=1}^{T} \mathbf{A}^\top \mathbf{x}_t \right) \right]$$

$$\leq \min_{\boldsymbol{\mu} \in \mathcal{D}} \left[\max_{\mathbf{x}_t \in \mathcal{X}} \left[\sum_{t=1}^{T} g(\mathbf{x}_t)/T - \boldsymbol{\mu}^\top \sum_{t=1}^{T} \mathbf{A}^\top \mathbf{x}_t \right] + \max_{\mathbf{e} \leq \boldsymbol{\gamma}} (\lambda r(\mathbf{e}) - \boldsymbol{\mu}^\top \mathbf{e}) \right]$$

$$= \min_{\boldsymbol{\mu} \in \mathcal{D}} \left[f^*(\mathbf{A}\boldsymbol{\mu}) + \lambda r^*(-\boldsymbol{\mu}) \right] = W_{Dual},$$

where $\mathcal{D} = {\boldsymbol{\mu} | r^*(-\boldsymbol{\mu}) < \infty}$ is the feasible region of dual variable $\boldsymbol{\mu}$. According to the Lemma 1 in the Balseiro et al. [5], we have $\mathcal{D}$ is convex and positive orthant is inside the recession cone of $\mathcal{D}$.

□

A.2 Proof of Theorem 2

PROOF. We let the variable $\mathbf{z}_p = (\mathbf{e}_p / \boldsymbol{\gamma}_p - 1)$, we have:

$$r^*(\boldsymbol{\mu}) = \max_{\mathbf{e} \leq \boldsymbol{\gamma}} \left[\min (\mathbf{e}_p / \boldsymbol{\gamma}_p) + \boldsymbol{\mu}^\top \mathbf{e}/\lambda \right]$$

$$= \boldsymbol{\mu}^\top \boldsymbol{\gamma}/\lambda + 1 + \max_{\mathbf{z}_p \leq 0} \left[\min(\mathbf{z}_p) + 1/\lambda \sum_{p \in \mathcal{P}} \mu_p \gamma_p z_p \right]$$

Let $s(\mathbf{z}) = \min_p \mathbf{z}_p$ and $\mathbf{v} = (\boldsymbol{\mu} \odot \boldsymbol{\gamma})/\lambda$, $\odot$ is the hadamard product. Then we define $s^*(\mathbf{v}) = \max_{\mathbf{z} \leq 0} (s(\mathbf{z}) + \mathbf{z}^T \mathbf{v})$. We firstly show that if $\sum_{p \in \mathcal{S}} \mathbf{v}_p \geq -1, \forall \mathcal{S} \in \mathcal{P}_s$, then $s^*(\mathbf{v}) = 0$ and $\mathbf{z} = 0$ is the optimal solution, otherwise $s^*(\mathbf{v}) = \infty$.

We can equivalently write $\mathcal{D} = {\mathbf{v} | \sum_{p \in \mathcal{S}} \mathbf{v}_p \geq -1, \forall \mathcal{S} \in \mathcal{P}_s}$. We firstly show that $s^*(\mathbf{v}) = \infty$ for $\mathbf{v} \notin \mathcal{D}$. Suppose that there exists a subset $\mathcal{S} \in \mathcal{P}_s$ such that $\sum_{p \in \mathcal{S}} \mathbf{v}_p < -1$. For any $b > 1$, we can get a feasible solution:

$$\mathbf{v}_p = \begin{cases} -b, & p \in \mathcal{S} \\ 0, & otherwise. \end{cases}$$

Then, because such solution is feasible and $s(\mathbf{z}) = -b$, we obtain that

$$s^*(\mathbf{v}) \geq s(\mathbf{z}) - b(\sum_{p \in \mathcal{S}} \mathbf{v}_p) = -b(\sum_{p \in \mathcal{S}} \mathbf{v}_p + 1).$$

Let $b \to \infty$, we have $s^*(\mathbf{v}) \to \infty$.

Then we show that $s^*(\boldsymbol{\mu}) = 0$ for $\mathbf{v} \in \mathcal{D}$. Note that $\mathbf{z} = 0$ is feasible. Therefore, we have

$$s^*(\mathbf{v}) \geq s^*(0) = 0.$$

Then we have $\mathbf{z} \leq 0$ and without loss of generality, that the vector $\mathbf{z}$ is sorted in increasing order, i.e., $\mathbf{z}_1 \leq \mathbf{z}_2, \cdots, \leq \mathbf{z}_{|\mathcal{P}|}$. The objective value is

$$s^*(\mathbf{v}) = \mathbf{z}_1 + \sum_{j \in |\mathcal{P}|} \mathbf{z}_j \mathbf{v}_j$$

$$= \sum_{j=1}^{|\mathcal{P}|} (\mathbf{z}_j - \mathbf{z}_{j+1}) \left(1 + \sum_{i=1}^{j} \mathbf{v}_i \right) \leq 0.$$

Thus we can have $s^*(\boldsymbol{\mu}) = 0$ for $\mathbf{v} \in \mathcal{D}$ and we have

$$r^*(-\boldsymbol{\mu}) = \boldsymbol{\mu}^\top \boldsymbol{\gamma}/\lambda + 1.$$

□

A.3 Proof of Theorem 3

PROOF. Firstly, in practice, we normalize the user-item preference score $s_{u,i}$ to $[0,1]$. Therefore, $\sum_{t=1}^{T} g(\mathbf{x}_t)/T \leq K$. In max-min regularizer $r(\mathbf{e})$. Let's abbreviate its upper bound to $\bar{r}$. In practice, $\bar{r} \leq 1$ We have

$$W_{OPT} \leq K + \lambda \bar{r}. \tag{7}$$

We consider the stopping time τ of Algorithm 1 as the first time the provider will have the maximum exposures, i.e.

$$\sum_{t=1}^{\tau} \mathbf{A}^\top \mathbf{x}_t \geq \boldsymbol{\gamma}.$$

Note that is τ a random variable.

Similarly, followed the prove idea of Balseiro et al. [5], First, we analysis the primal performance of the objective function. Second, we bound the complementary slackness term by the momentum gradient descent. Finally, We conclude by putting it to achieve the final regret bound.

Primal performance proof: Consider a time $t < \tau$, the recommender action will not violate the resource constraint. Therefore, we have:

$$g(\mathbf{x}_t)/T = g^*(\mathbf{A}\boldsymbol{\mu}_t) + \lambda \boldsymbol{\mu}_t^T \mathbf{A}^\top \mathbf{x}_t,$$

and we have $\mathbf{e}_t = \arg\max_{\mathbf{e} \leq \boldsymbol{\gamma}} {r(\mathbf{e}) + \boldsymbol{\mu}^\top \mathbf{e}/\lambda}$

$$r(\mathbf{e}_t) = r^*(-\boldsymbol{\mu}) - \boldsymbol{\mu}_t^T \mathbf{e}_t/\lambda.$$

We make the expectations for the current time step t for the primal functions:

$$\mathbb{E} [g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}_t)] = \mathbb{E} \left[g^*(\mathbf{A}\boldsymbol{\mu}_t) + \boldsymbol{\mu}_t^T \mathbf{A}^\top \mathbf{x}_t + \lambda r^*(-\boldsymbol{\mu}) - \boldsymbol{\mu}_t^T \mathbf{e}_t \right]$$

$$= W(\boldsymbol{\mu}_t) - \mathbb{E} \left[\boldsymbol{\mu}_t^T (-\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t) \right].$$

Consider the process $Z_t = \sum_{j=1}^{T} \boldsymbol{\mu}_j^t (-\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t) - \mathbb{E} \left[\boldsymbol{\mu}_t^T (-\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t) \right]$ is a martingale process. The Optional Stopping Theorem in martingale process [46] implies that $\mathbb{E} [Z_\tau] = 0$. Consider the variable $w_t(\boldsymbol{\mu}_t) = \boldsymbol{\mu}_t^T (-\mathbf{A}^\top \mathbf{x}_t + \mathbf{e}_t)$, we have

$$\mathbb{E} \left[\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) \right] = \mathbb{E} \left[\sum_{t=1}^{\tau} \mathbb{E} [w_t(\boldsymbol{\mu}_t)] \right]$$

Algorithm 2: Online learning of Min-Regularizer

Input: User arriving set $\{u_i\}_{i=1}^N$, time-separate size T, ranking size K, user-item preference score $s_{u,i}, \forall u, i$, item-provider adjacent matrix $\mathbf{A}$, maximum resources $\boldsymbol{\gamma}$ and the trade-off coefficient λ.

Output: The decision variables $\{\mathbf{x}_i, i = 1, 2, \cdots, N\}$

1: **for** $n = 1, \cdots, N/T$ **do**
2:    Initial dual solution $\boldsymbol{\mu}_1 = \mathbf{0}$. remain resources $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}$ and momentum gradient $\mathbf{g}_0 = \mathbf{0}$.
3:    **for** $t = 1, \cdots, T$ **do**
4:      Receive u_{nT+t}
5:

$$\mathbf{m}_p = \begin{cases} 0, & \boldsymbol{\beta}_{tp} > 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

6:      // Make the recommendation:
7:

$$\mathbf{x}_t = \underset{\mathbf{x}_t \in \mathcal{X}}{\arg\max} \left[g(\mathbf{x}_t) - \lambda \left(\mathbf{A}(\mathbf{e} - (\min_p \mathbf{e}_p [1, 1, \ldots, 1]^\top) + \mathbf{m}) \right)^\top \mathbf{x}_t \right]$$

8:      // Update the remaining resources:
9:

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t - \mathbf{A}^\top \mathbf{x}_t$$

10:    **end for**
11: **end for**

Moreover, in MMF, the dual function W_{Dual} is convex proofed in Theorem 1, we have

$$\mathbb{E}\left[\sum_{t=1}^{\tau} g(\mathbf{x}_t)/T + \lambda r(\mathbf{e}_t)\right] = \mathbb{E}\left[\sum_{t=1}^{\tau} W_{Dual}(\boldsymbol{\mu}_t)\right] - \mathbb{E}\left[\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t)\right]$$

$$\leq \mathbb{E}\left[\tau W_{Dual}(\widetilde{\boldsymbol{\mu}_\tau})\right] - \mathbb{E}\left[\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t)\right], \tag{8}$$

where $\widetilde{\boldsymbol{\mu}_\tau} = \sum_{t=1}^{\tau} \boldsymbol{\mu}_t/\tau$.

Complementary slackness proof Then we aim to proof the complementary slackness $\sum_{t=1}^{T} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu})$ is bounded. Suppose there exists $G, s.t.$ the gradient norm is bounded $\|\widetilde{\mathbf{g}}_t\| \leq G$. Then we have:

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu}) \leq \frac{\lambda^2}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(\tau - 1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}, \tag{9}$$

where the project function $\|\boldsymbol{\mu} - \boldsymbol{\mu}_t\|_{\boldsymbol{\gamma}}^2$ is $\sigma-$strongly convex.

Next we prove the inequality in Equation (9). According to the Theorem 1 in [6], we have

$$\|\mathbf{g}_t\|_2^2 = \|(1-\alpha)\sum_{s=1}^{t} \alpha^{t-s}(\widetilde{\mathbf{g}}_s)\|_2^2 \leq G^2,$$

and

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) - w_t(\boldsymbol{\mu}) \leq \frac{\|\boldsymbol{\mu}_t - \boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(\tau-1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}, \forall \boldsymbol{\mu}.$$

Assuming there exists $H > 0$, s.t. $\|\boldsymbol{\mu}_t - \boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2 \leq H$. According to the Cauchy-Schwarz' inequality. The results follows. Let $M = \frac{H}{\eta} + \frac{G^2}{(1-\alpha)\sigma}\eta(T-1) + \frac{G^2}{2(1-\alpha)^2\sigma\eta}$. We now choose a proper $\boldsymbol{\mu}$, s.t. the complementary stackness can be further bounded.

For $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}} + \theta$, where $\theta \in \mathbb{R}^{|P|}$ is non-negative to be determined later and $\hat{\boldsymbol{\mu}} = \arg\max_{\boldsymbol{\mu}} -\boldsymbol{\mu}^\top (\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t)/\lambda$. According to the constraint $\mathbf{e} = \sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t$, we have that

$$\sum_{t=1}^{T} (r(\mathbf{e}_t) + \boldsymbol{\mu}^\top \mathbf{e}_t \lambda) \leq r^*(-\hat{\boldsymbol{\mu}}) = r(\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t) + \hat{\boldsymbol{\mu}}^T (\sum_{i=1}^{T} \mathbf{A}^\top \mathbf{x}_t)/\lambda.$$

Note that in proof of Theorem 2, the feasible region $\mathcal{D}$ is recession cone, therefore, $\boldsymbol{\mu} \in \mathcal{D}$.

Therefore, we have

$$\sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) = \sum_{t=1}^{T} w_t(\hat{\boldsymbol{\mu}}) - \sum_{t=\tau+1}^{T} w_t(\hat{\boldsymbol{\mu}}) + \sum_{t=1}^{\tau} w_t(\theta) + M. \tag{10}$$

Put them together: Under the max-min fair, we obtain that

$$W_{OPT} = \frac{\tau}{T}W_{OPT} + \frac{T-\tau}{T}W_{OPT} \leq \tau W_{Dual}(\widetilde{\boldsymbol{\mu}_\tau}) + (T-\tau)(K + \lambda \bar{r}). \tag{11}$$

Therefore, combining Eq. (10,12,13) the regret $\text{Regret}(h)$ can be bounded as:
$\text{Regret}(h) = \mathbb{E}[W_{OPT} - W]$

$$\leq \mathbb{E}\left[W_{OPT} - \sum_{t=1}^{\tau} \left(g(\mathbf{x}_t)/T - \lambda r(\mathbf{A}^\top \mathbf{x}_t/\boldsymbol{\gamma})\right)\right]$$

$$\leq \mathbb{E}\left[W_{OPT} - \tau W_{Dual}(\widetilde{\boldsymbol{\mu}_t}) + \sum_{t=1}^{\tau} w_t(\boldsymbol{\mu}_t) + \sum_{t=1}^{T}(\mathbf{e}_t - \mathbf{A}^\top \mathbf{x}_t)\right]$$

$$\leq \mathbb{E}\left[(T-\tau)(K + \lambda\bar{r}) + \sum_{t=1}^{T} w_t(\hat{\boldsymbol{\mu}}) + \sum_{t=1}^{\tau} w_t(\theta)\right] + M$$

$$\leq (T-\tau)(K + \lambda\bar{r} + \lambda K) + \sum_{t=1}^{\tau} w_t(\theta) + M. \tag{12}$$

Let $C = K + \lambda\bar{r} + \lambda K$, then setting the $\theta = C\min_p \boldsymbol{\gamma}_p \mathbf{u}_p$, where $\mathbf{u}_p$ is the p-th unit vector. We have

$$\sum_{t=1}^{\tau} w_t(\theta) = C/(\min_p \boldsymbol{\gamma}_p) - C(T - \tau).$$

Then the $\text{Regret}(h) \leq M + C/(\min_p \boldsymbol{\gamma}_p)$, when we set $\eta = O(T^{-1/2})$, the Regret(h) is comparable with $O(T^{1/2})$. The result follows.

□

A.4 Algorithm for Min-reguarlizer

In this section, we proposed a heuristic method for provider MMF online application, named Min-reguarlizer. It has a regularizer that measures the exposure gaps between the target provider and the worst-providers. The detailed algorithm is shown in Algorithm 2. The notations are the same as P-MMF in Algorithm 1.