P-MMF：推荐系统中的提供者最大最小公平性重排序

Chen Xu 高瓴人工智能学院 中国人民大学 xc_chen@ruc.edu.cn

Sirui Chen 信息学院 中国人民大学 csr16@ruc.edu.cn

Jun Xu*高瓴人工智能学院 中国人民大学 junxu@ruc.edu.cn

Weiran Shen 高瓴人工智能学院 智能 中国人民大学 shenweiran@ruc.edu.cn

Xiao Zhang 高瓴人工智能学院 智能 中国人民大学 zhangx89@ruc.edu.cn

Gang Wang Zhenghua Dong 华为 诺亚方舟实验室 wanggang110@huawei.com dongzhenhua@huawei.com

arXiv:2303.06660v1 [cs.IR] 12 Mar 2023

摘要

本文从提供方的角度探讨公平推荐问题，这一问题在多利益相关方推荐系统中变得日益重要。现有关于提供方公平性的研究通常侧重于设计比例公平性（PF）指标，首先考虑系统性公平。然而，社会学研究表明，为使市场更加稳定，最大最小公平性（MMF）是更好的指标。其主要原因在于，MMF旨在优先提升处于最不利地位者的效用，引导系统支持市场地位较弱的提供方。在将MMF应用于推荐系统时，如何在在线推荐场景中平衡用户偏好与提供方公平性，仍然是一个具有挑战性的问题。本文提出了一种名为提供者最大最小公平性重排序（P-MMF）的在线重排序模型，以解决这一问题。具体而言，P-MMF将提供方公平推荐建模为资源分配问题，其中曝光位被视为分配给提供方的资源，并在该过程中将最大最小公平性用作正则化项。我们表明，该问题可以进一步表示为一个正则化在线优化问题，并在其对偶空间中高效求解。在在线重排序阶段，我们设计了一种动量梯度下降方法来执行动态重排序。理论分析表明，P-MMF的遗憾值是有界的。在四个公开推荐系统数据集上的实验结果表明，P-MMF的性能优于最先进基线方法。实验结果还表明，在包含大量物品的数据集上，P-MMF仍能保持较低的计算成本。

ACM 概念分类

• **信息系统** →**推荐系统**。

关键词

最大最小公平性、提供方公平性、推荐系统

ACM 参考文献格式：

Chen Xu、Sirui Chen、Jun Xu、Weiran Shen、Xiao Zhang、Gang Wang 和 Zhenghua Dong。2023 年。P-MMF：推荐系统中的提供者最大最小公平性重排序。载于《*2023* 年 *ACM Web Conference* 论文集》（*WWW '23*），*2023* 年 *5* 月 *1* 日至 *5* 日，美国得克萨斯州奥斯汀。ACM，美国纽约州纽约，11 页。https://doi.org/10.1145/3543507.3583296

1 引言

出于社会、伦理和经济方面的考虑，推荐中的公平性问题变得不容忽视 [1, 2]。在多利益相关者推荐系统（RS）中，存在多个不同的参与者，包括用户、物品、提供方等 [1]。人们已从用户 [29], 、提供方 [8, 35], 或双方 [45]的角度提出了多种公平推荐模型。本文关注的是如何在多利益相关者推荐系统中确保提供方公平性的问题。

关于公平性，社会学研究中存在许多文化差异 [43, 44] 。一种实用的公平性定义建立在两个社会学概念之上：平等与公平 [31]。根据 Matsumoto 和 Juang 的观点 [31], ，平等可以定义为：每个人都受到同样的对待，并获得相同的资源以取得成功；而公平可以定义为：确保根据需求平等地分配资源。图 1 直观地展示了推荐场景下这两种公平性的示例。假设我们拥有一些资源（即曝光位），以确保提供方能够获得苹果（即在市场中生存所需的机会）。如图 1 所示，平等确保推荐系统向每个提供方平均给予相同的支持，而公平（即分配正义）则强调推荐系统应以不同的比例向提供方分配资源 [26]。

基于公平性概念 [25], ，比例公平（PF）和最大最小公平性（MMF）这两项指标 [33] 分别被提出。PF 和 MMF 已广泛应用于计算机网络和交通运输 [9, 17, 36]。它们在提供方公平推荐场景中的形式化

*Jun Xu 是通讯作者。部分工作在北京市大数据管理与分析方法重点实验室完成。

允许为个人或课堂使用而免费制作本作品的全部或部分数字副本或纸质副本，但前提是不得出于营利或商业利益而制作或分发副本，且副本首页必须附有本声明及完整引文。对于本作品中由作者以外的其他方拥有版权的组件，必须尊重其版权。注明出处后允许进行摘要。除上述情况外，如需复制、再版、发布到服务器或向列表重新分发，必须事先获得具体许可和/或支付费用。请通过 permissions@acm.org 申请许可。*WWW '23*，*2023*年*5*月*1*日至*5*日，美国得克萨斯州奥斯汀 © 2023 版权所有归所有者/作者所有。出版权已授权给 ACM。ACM
ISBN 978-1-4503-9416-1/23/04...$15.00
https://doi.org/10.1145/3543507.3583296

[1]图片来源：Interaction Institute for Social Change | 艺术家：Angus Maguire。（interac- tioninstitute.org 和 madewithangus.com）。萨斯卡通卫生区。

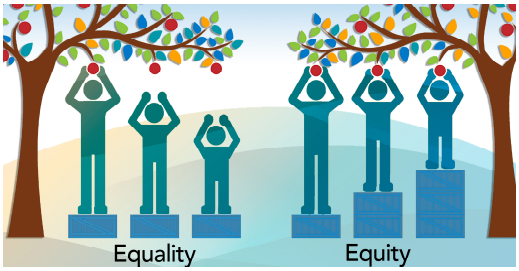


图 1：平等与公平的直观示例。[1]

表述将在第 3 节中说明。直观而言，PF 和 MMF 试图按照指定比例分配资源 [7]。PF 基于福利原则（即亚里士多德原则和纳什解 [33]），旨在最大化提供方福利之和。MMF（即罗尔斯原则和卡莱-斯莫罗金斯基解 [23]）旨在改善处境最差的提供方的效用。事实证明，对于提供方公平问题，MMF 是更好的指标，因为处境较差的提供方占平台的大多数，而这些支持能够帮助他们生存。关注这些弱势提供方将提高推荐市场的稳定性 [35]。例如，根据 Erickson 的报告 [18],，亚马逊的小型卖家难以独自应对“盈利能力较低”和“无法实现个性化”等挑战。未确保 MMF 的不公平系统可能导致提供方与 RS 之间的关系破裂，最终迫使提供方离开。

现有的提供方公平推荐模型要么考虑提供方的比例公平 [8, 32, 35, 47]，要么采用启发式方法来保障处境较差的提供方的效用 [8, 35]。然而，这些启发式方法并未直接考虑最大最小公平性（MMF），在适用于在线场景时也缺乏理论保证。此外，这些启发式方法缺乏与用户效用进行权衡的灵活性，这不可避免地会损害用户体验。

本文旨在开发一种从提供方视角考虑最大最小公平性（MMF）的实用重排序模型。我们将提供方公平推荐构建为资源分配过程 [10]。在这一过程中，资源可视为有限的排序槽位，而提供方则视为需求方。分配成本定义为用户的偏好。此外，我们在分配中加入 MMF正则化器，以最大化分配给特定提供方的最小资源量。

为使该方法适用于在线推荐场景，我们提出了一种名为 P-MMF 的在线重新排名算法。该算法将提供方公平在线推荐视为一个正则化在线优化问题，重点关注提供者最大最小公平性。然而，该优化问题包含许多整数变量，因而极难求解。随后，我们推导出其对偶问题，并提出一种在对偶空间中优化该问题的高效算法。在在线设置下，我们开发了一种动量梯度下降方法，以实现有效且高效的在线推荐。理论分析表明，P-MMF 的遗憾值有良好的上界。此外，由于 P-MMF 对物品数量不敏感，因此其计算效率也很高。

本文的主要贡献总结如下：

(1) 我们分析了在多利益相关者推荐系统中确保提供方公平的重要性，并提出在提供方公平推荐中使用最大最小公平性指标。

(2) 我们将提供方公平推荐建模为一个受最大最小公平性正则化的资源分配问题，并提出了一种名为 P-MMF 的重排序模型，以平衡提供方公平性与用户偏好。我们的理论分析表明，P-MMF 的遗憾值可以得到界定。

(3) 在小型数据集上的仿真表明，与比例公平相比，MMF 在为提供方实现公平推荐方面更具优势。在四个公开可用数据集上的广泛实验表明，P-MMF 的表现优于最先进基线方法，包括基于比例公平和基于 MMF 的方法。

2 相关工作

公平性已成为多利益相关者推荐系统中的热门研究主题 [1, 2]。研究人员提出了若干面向用户和面向物品/提供方的公平性重排序模型。对于面向用户的公平性，Abdollahpouri 等 [3] 提出了一种以用户为中心的评估方法，用于衡量用户对热门物品的不同兴趣程度。Li 等 [29], 、Sacharidis [40], 以及 Serbos 等 [41] 从群体公平性的角度研究了以用户为中心的公平性。对于面向提供方的公平性，Gómez 等 [21] 提出了一种基于规则的算法，以确保物品的曝光量得到均等分配。Ge 等 [20] 旨在应对物品流行度随时间变化时的动态公平性。然而，这些面向提供方的方法没有考虑用户视角。

近年来，也有研究同时考虑用户偏好与提供方公平性之间的权衡。Chakraborty 等[11] 声称，在经济平台中，所有提供方获得的曝光量都应与其相关性成正比。Rahmani 等 [38] 研究了兴趣点推荐中用户公平性与生产者公平性之间的权衡。TFROM [47] 和 CPFair [32] 分别将这种权衡建模为背包问题和松弛的线性规划问题。然而，它们在在线场景中采用了基于贪心的算法，这只能改善提供方的比例公平性。一些研究注意到，还应保障处境较差的提供方的效用。例如，FairRec [35] 及其扩展版本 FairRec+ [8] 提出了硬约束，以确保每个提供方都能获得最低曝光量。Welf [16] 提出了一种 Frank-Wolfe 算法，以最大化处境较差物品的福利函数。然而，这些方法都是为离线场景设计的，并且计算成本高昂，从而阻碍了它们在实际在线推荐系统 [50, 51]中的应用。

在本文中，我们将重新排名任务表述为资源分配问题 [10], ，该问题在通信和交通运输领域至关重要。在在线资源分配中，大多数研究 [14, 28]专注于设计时间可分奖励函数，即各个时期奖励的总和。Li 等人 [27] 针对线性奖励提出了一种基于局部的次梯度算法。Cheung 等人[13] 设计了一种基于对偶的在线算法，通过请求分布进行学习。Balseiro 等人 [5] 提出了一种镜像下降方法，用于解决公平性正则化在线分配问题。

3 提供方公平重排序

在本节中，我们首先定义多利益相关者推荐系统中的符号。然后给出比例公平性和最大最小公平性的形式化定义。

在多利益相关者推荐系统中，存在多个参与方，包括用户、物品提供方和其他利益相关者。设 U、I 和 P 分别表示用户、物品和提供方的集合。每个物品 $i \in \mathrm{I}$ 都与唯一的提供方 $p \in \mathrm{P}$ 相关联。与特定提供方 p 相关联的物品集合记为 I_p。当特定用户 $u \in \mathrm{U}$ 访问推荐系统时，系统会向用户提供一个包含 K 个物品的列表，记为 $L_K(u) \in \mathrm{I}^K$。对于每个用户—物品对 (u, i)，推荐模型会估计偏好分数 $s_{u,i} \in \mathbb{R}$。这些物品会根据其偏好分数进行排名。在本文中，我们将向 u 曝光物品列表 $L_K(u)$ 所产生的用户侧效用定义为该列表中偏好分数的总和，记为 $f(L_K(u)) = \sum_{i \in L_K(u)} s_{u,i}$。遵循文献惯例 [32, 35]，我们将提供方的公平性向量定义为 $\mathbf{e} \in \mathbb{R}^{|\mathrm{P}|}$；对于特定提供方 p，$\mathbf{e}_p \in \mathbb{R}^+$ 表示提供方 p 的曝光物品数量。提供方公平重排序的目标是计算一个新的公平列表 $L_K^F(u) \in \mathrm{I}^K$ 在良好平衡用户效用 $f\left(L_K^F(u)\right)$ 与基于 $\mathbf{e}$ 定义的提供方公平性指标方面。

在现实应用中，用户会依次到达推荐系统。假设在时刻 t 用户 u_t 到达。推荐系统需要在从 $t = 0$ 到 T 的整个时间范围内考虑提供方长期曝光量。我们的任务可以表述为一个具有时间可分公平性的资源分配问题 [5]。具体而言，推荐系统的最优效用可定义为时间可分效用函数，即 0 到 T 各时段内的累积奖励 [5]。在这种情况下，$\mathbf{e}_p$ 可视为提供方 p 在 0 到 T 期间累计获得的曝光物品总数。形式化地，在权衡用户偏好与提供方公平性时，我们有如下数学规划：

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

其中，$\boldsymbol{\gamma} \in \mathbb{R}^{|\mathrm{P}|}$ 表示不同提供方的权重，例如加权PF 或 MMF [7],，而 $r(\mathbf{e}) \in \mathbb{R}$ 是用作公平性正则化项的提供方公平性指标。请注意，约束 $\mathbf{e} \le \boldsymbol{\gamma}$ 也可以视为分配给该提供方的最大资源量

提供方。根据比例公平和 MMF 的定义，[7]，$r(\mathbf{e})$ 也有两种不同的形式：

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

最大-最小公平（MMF）：$r(\mathbf{e}) = \text{最小}_{p \in \mathcal{P}}\left[\mathbf{e}_p/\boldsymbol{\gamma}_p\right]$。

使用比例公平性作为正则化项，可以减小 $\mathbf{e}$ 与 $\boldsymbol{\gamma}$ 之间的差异；而使用最大最小公平性，则可以提高曝光量最低的提供方的相对曝光量（相对于权重 $\boldsymbol{\gamma}$）。

4 我们的方法：P-MMF

在本节中，我们首先将推荐中的提供方公平性表述为一个资源分配问题。然后，我们提出一种名为 P-MMF 的在线推荐算法。

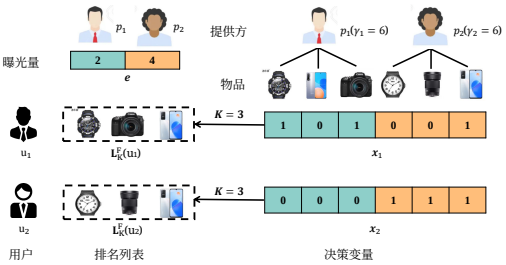


图 2：推荐中时间可分公平性的一个简单示例

4.1 具有提供方公平性的资源分配

我们将提供方的公平推荐问题表述为一个资源分配过程 [10]。在这一过程中，资源可视为分配给提供方的有限排序槽位。分配成本根据用户的偏好和最大-最小公平性进行定义。

形式化地，根据公式（1），公平推荐问题可以写成一个线性规划：

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

其中，$\mathbf{x}_t \in \{0,1\}^{|\mathcal{I}|}$ 是用户 u_t 的决策向量，$g(\cdot)$ 是用户侧效用函数。具体而言，对于每个物品 i，若将其加入公平排序列表 $L_K^F(u_t)$，则 $\mathbf{x}_{ti} = 1$；否则，$\mathbf{x}_{ti} = 0$。请注意，$g(\cdot)$ 与 $f(\cdot)$ 等价，即二者会产生相同结果，而 $g(\cdot)$ 以二元决策向量作为输入。公式（2）中的第一个约束确保推荐列表的大小为 K。公式（2）中的第二个约束表明，每个提供方 p 的曝光量是其对应物品在所有时期内累计曝光量的总和。一般而言，在时效性较弱的场景下，我们认为时间可分公平性更受青睐。例如，推荐使用寿命较长的物品（如游戏、服装等）。

图 2 展示了线性规划问题 (2) 的一个简单示例。在此示例中，我们设定 $T = 2$。假设有两位用户 u_1 和 u_2 依次到达系统。在每个时间步，系统都会推荐一个包含 $K = 3$ 个物品的列表。因此，系统总共有 $2 \times 3 = 6$ 个可用于曝光的名额。假设系统有两个提供方 p_1 和 p_2，每个提供方拥有三个物品。令 p_1 和 p_2 的权重相同（即 $\boldsymbol{\gamma} = [\gamma_1 = 6, \gamma_2 = 6]$）。该模型求解问题 (2)，得到两个二进制向量：$\mathbf{x}_1$ 对应 u_1，以及 $\mathbf{x}_2$ 对应 u_2。最后，我们统计整个时间 $T = 2$ 内的曝光量：$\mathbf{e} = [e_{p_1} = 2, e_{p_2} = 4]$，这表示通过推荐排名列表 $L_K^F(u_1)$（基于 $\mathbf{x}_1$ 生成）和 $L_K^F(u_2)$（基于 $\mathbf{x}_2$ 生成）

对于 u_1 和 u_2，提供方 p_1 和 2 分别获得 2 次和 4 次曝光。

尽管我们已经给出了该问题的线性规划解公式（2），但它只能以离线方式求解小规模问题。在在线推荐系统中，对于每个用户 u_t 的访问，模型需要立即从大规模物品语料库中生成公平排序列表 $L_K^F(u_t)$。这意味着我们无法获知 t 之后的信息。接下来，我们将讨论如何在在线推荐问题中使用 MMF。

4.2 面向在线应用的 P-MMF

本节首先对在线场景中的 MMF 进行形式化。接下来，我们考虑原始问题的对偶问题。关注对偶问题具有几个优势：对偶问题的变量显著更少，并且这些变量不再像原始问题中的变量那样需要是整数。最后，我们提出了一种动量梯度下降算法，以实现高效的在线学习。

4.2.1 在线场景中的最大最小公平性。在时间步 t，推荐系统收到来自用户 u_t 的请求。在线算法 h 基于当前用户 u_t 和此前的历史记录 $\mathrm{H}_{t-1}=\{u_s, \mathbf{x}_s\}_{s=1}^{t-1}$，生成实时决策向量 $\mathbf{x}_t \in \{0,1\}^{|\mathcal{I}|}$：$\mathbf{x}_t = h(u_t \mid \mathrm{H}_{t-1})$。我们将最大最小公平性的在线奖励定义为所有时间步奖励之和：

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

其中，$\min\left[\mathbf{e}/\boldsymbol{\gamma}\right]$ 对应于最大最小公平性正则化项。

我们的目标是设计一种算法 h，使其达到较低的遗憾值。记 W_{OPT} 为最优值，我们将算法的遗憾值定义为离线问题的最优性能 W_{OPT} 与在线算法 W 在用户分布 U 上的性能之期望差：

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

4.2.2 对偶问题。根据公式 (1) 中的原问题，我们知道，对于每个时间 t，整数决策变量 $\mathbf{x}_t$ 的规模为 $|\mathcal{I}|$，因此难以求解。然而，考虑其对偶问题后，我们可以显著降低其计算成本。

定理1（对偶问题）。公式的对偶问题可写为：

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

其中，$\mathbf{A}\in\mathbb{R}^{|\mathcal{I}|\times|\mathcal{P}|}$ 是物品-提供商邻接矩阵，$A_{ip} = 1$ 表示物品 $i \in \mathcal{I}_p$，否则为 0。令 $\mathcal{X}=\{\mathbf{x}_t | \mathbf{x}_t \in\{0,1\}^{|\mathcal{I}|} \wedge \sum_{i\in\mathcal{I}}\mathbf{x}_{ti} = K\}$，$g^*(\cdot)$ $r^*(\cdot)$ 为共轭函数：

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

且 $\mathcal{D}=\{\boldsymbol{\mu} | r^*(-\boldsymbol{\mu}) < \infty\}$ 是对偶变量 $\boldsymbol{\mu}$ 的可行域。

定理1的证明见附录 A.1。根据定理1，我们可以得到一个新的非整数决策变量 $\boldsymbol{\mu}\in\mathbb{R}^{|\mathcal{P}|}$。在实践中，提供方规模为 $|\mathcal{P}| \ll |\mathcal{I}|$。此外，由于 A 的稀疏性，计算 $\mathbf{A}\boldsymbol{\mu}$ 非常高效，其目的是将变量 $\boldsymbol{\mu}$ 从提供方空间投影到物品空间。

在第 4.2.3 节讨论的在线算法中，我们可以在常数时间内得到共轭函数 $g^*(\cdot)$ 的闭式。对于可行域 D 和共轭函数 $r^*(\cdot)$，我们有定理2和引理 1。

定理2（对偶可行域）。在 *MMF* 中，对偶问题的可行域

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

其中，$\mathcal{P}_s$ 是 $\mathcal{P}$ 的幂集，即 $\mathcal{P}$ 的所有子集构成的集合。

定理2的证明见附录 A.1，这意味着如下引理 1：

引理 1。共轭函数 $r^*(\cdot)$ 具有如下闭式：$\max_{\boldsymbol{\mu}\le\boldsymbol{\gamma}} r^*(-\boldsymbol{\mu}) = \boldsymbol{\gamma}^T\boldsymbol{\mu}/\lambda + 1$，且最优对偶变量为：$\arg\max_{\boldsymbol{\mu}\le\boldsymbol{\gamma}} r^*(-\boldsymbol{\mu}) = \boldsymbol{\gamma}/\lambda$.

4.2.3 P-MMF 算法。算法 1 展示了 P-MMF 算法。遵循 Balseiro 等人 [5], 的方法，P-MMF 为每个时间 t 保留一个对偶变量 $\boldsymbol{\mu}_t$、剩余资源 $\boldsymbol{\beta}_t$ 和梯度 $\mathbf{g}_t$。

每当用户到达时，算法会根据剩余资源和对偶变量 $\boldsymbol{\mu}_t$（第 7 行）计算推荐变量 $\mathbf{x}_t$。最终对偶变量被估计为每个时间 t 的平均对偶变量：$\boldsymbol{\mu} = \sum_{t=1}^{T}\boldsymbol{\mu}_t/T$。直观而言，对于 $\boldsymbol{\mu}_t$，当对偶变量的值较高时，算法自然会推荐较少与相应提供方相关的物品。剩余资源 $\boldsymbol{\beta}_t$ 确保算法只从拥有剩余资源的提供方中推荐物品。请注意，在第 7 行中，该公式关于 $\mathbf{x}_t$ 是线性的。因此，可以通过 top-K 排序算法在常数时间内高效计算 $\mathbf{x}_t$。

在线学习过程如下。首先，根据引理 1，我们得到最大-最小正则化器 $r^*(-\boldsymbol{\mu}_t)$ 的共轭函数的闭式。然后，我们可以得到对偶函数 $g^*(\mathbf{A}\boldsymbol{\mu}) + \lambda r^*(-\boldsymbol{\mu})$ 的次梯度：

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

我们还将上一次的动量 [37] 加入更新后的梯度 $\mathbf{g}_t$。最后，我们利用 $\mathbf{g}_t$ 通过执行第 14 行的在线下降来更新对偶变量，其中我们使用了加权 ℓ_2-norm：$\|\boldsymbol{\mu}\|_{\boldsymbol{\gamma}^2}^2 = \sum_{j=1}^{|\mathcal{P}|}\boldsymbol{\gamma}_j^2\boldsymbol{\mu}_j^2$。因此，对偶变量将朝着曝光量较少的提供方所在的方向移动，而原变量 $\mathbf{x}_t$ 将移动到更优的解。

请注意，由于 D 是坐标对称的，因此第 14 行的投影步骤可以使用凸优化求解器 [5] 高效求解。

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

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

其中 $\widetilde{\boldsymbol{\mu}}$ 满足：$\boldsymbol{\gamma}_1\widetilde{\boldsymbol{\mu}}_1 \le \boldsymbol{\gamma}_2\widetilde{\boldsymbol{\mu}}_2 \le \cdots \le \boldsymbol{\gamma}_{|\mathcal{P}|}\widetilde{\boldsymbol{\mu}}_{|\mathcal{P}|}$.

我们在定理3中给出了 Regret(h) 的遗憾界，但由于篇幅限制，将证明留至附录A.3。直观而言，更大的排名大小 K 和时间 T 会导致 P-MMF 中更大的偏差。

定理3（遗憾界）。假设函数 $\|\cdot\|_{\boldsymbol{\gamma}^2}^2$ 是 σ-强凸的，并且存在常数 $G\in\mathbb{R}^+$ 和 $H>0$ 使得

算法 1： P-MMF 的在线学习

输入： 用户到达顺序 $\{u_i\}_{i=1}^{N}$、时间分段大小 T、排名大小 K、用户-物品偏好分数 $s_{u,\ i},\ \forall u,\ i$、物品-提供商邻接矩阵 $\mathbf{A}$、最大资源量 $\boldsymbol{\gamma}$ 以及权衡系数 λ。

输出： 决策变量 $\{\mathbf{x}_i,\ i=1,\ 2,\ \ldots\ldots,\ N\}$

对于 $n=1,\ \cdots\ N/T$ **执行**
  初始化对偶解 $\boldsymbol{\mu}_1=0$、剩余资源 $\boldsymbol{\beta}_1=\boldsymbol{\gamma}$、以及动量梯度 $\mathbf{g}_0=0$。
  对于 $t=1\cdots,T$ **do**
    用户 u_{nT+t} 到达
    $\mathbf{m}_p=\begin{cases}0, & \boldsymbol{\beta}_{tp}>0\\ \infty, & \text{否则}\end{cases}$
    // 生成推荐：
    $\mathbf{x}_t=$ 参数最大值$_{\mathbf{x}_t\in\mathbf{X}}\ \left[g(\mathbf{x}_t)/T-(\mathbf{A}(\boldsymbol{\mu}_t+\mathbf{m}))^{\top}\mathbf{x}_t\right]$
    $\boldsymbol{\beta}_{t+1}=\boldsymbol{\beta}_t-\mathbf{A}^{\top}\mathbf{x}_t$
    $\mathbf{e}_t=$ 最大值$_{\mathbf{e}_t\le\boldsymbol{\beta}_t}\ r^*(-\boldsymbol{\mu}_t)$
    $\widetilde{\mathbf{g}}_t=-\mathbf{A}^{\top}\mathbf{x}_t+\mathbf{e}_t$
    $\mathbf{g}_t=\alpha\widetilde{\mathbf{g}}_t+(1-\alpha)\ \mathbf{g}_{t-1}$
    $\boldsymbol{\mu}_{t+1}=\arg\min_{\boldsymbol{\mu}\in\mathcal{D}}\left[\mathbf{g}_t,\boldsymbol{\mu}\ +\eta\|\boldsymbol{\mu}-\boldsymbol{\mu}_t\|^2_{\boldsymbol{\gamma}^2}\right]$
  结束循环
结束循环

$\|\widetilde{g}_t\|<G$、$\|\boldsymbol{\mu}_t-\boldsymbol{\mu}_0\|^2_{\boldsymbol{\gamma}^2}\le H$……则遗憾值可界定如下：

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

其中，$\bar{r}$ 是 *MMF* 正则化项的上界，在实践中，$\bar{r}\le 1$。

将学习率设为 $\eta=O(T^{-1/2})$，我们可以获得阶数为 $O(T^{1/2})$ 的次线性遗憾值上界。

5 实验

我们进行了实验，以展示所提出的 P-MMF 在提供商公平推荐方面的有效性。源代码和实验已分享到 GitHub [2]。

5.1 实验设置

5.1.1 数据集。实验采用了四个大规模、公开可用的推荐数据集，包括：

Yelp[3]：一个大规模的商家推荐数据集。我们仅使用了点击数据，该数据模拟为评分为 4—5 星的样本。物品所在的城市被视为提供方。该数据集包含 154543 个样本，其中有 17034 个用户、11821 个物品和 23 个提供方。

Amazon-Beauty/Amazon-Baby：亚马逊产品数据集[4]中的两个子集（Beauty 和数字音乐领域）。我们仅使用了点击数据，即模拟为 4-5 星评分样本的数据。此外，品牌被视为提供方。它们包含 49217/59836 个样本，其中包含 9625/11680 个用户、2756/2687 个物品和 1024/112 个提供方。

Steam[5] [24]：我们在实验中使用了游戏时长超过 10 小时的数据。游戏的发行商被视为提供方。该数据集包含 29530 个样本，其中有 5902 个用户、591 个物品和 81 个提供方。

作为预处理步骤，我们从整个数据集中移除了与少于 5 个物品/用户发生交互的用户、物品和提供方，以避免极端稀疏情况。我们还移除了关联物品少于 5 个的提供方。

5.1.2 评估。我们按照时间对所有交互进行排序，并使用前 80% 的交互作为训练数据来训练基础模型（即 BPR [39]）。剩余 20% 的交互作为测试数据用于评估。基于训练好的基础模型，我们可以获得每个用户-物品对 (u,i) 的偏好分数 $s_{u,i}$。测试数据中的按时间顺序排列的交互被划分为交互序列，其中时间范围长度设置为 T。我们分别计算每个序列的评估指标，并将平均结果报告为最终性能。

对于评估指标，我们从三个方面评估模型的性能：用户侧偏好、提供方公平性，以及二者之间的权衡。对于用户侧偏好，遵循 [47]，中的做法，我们采用 NDCG@K，其定义为原始排序列表 $\mathbf{L}_K(u_t)$ 中基于位置的用户-物品分数之和 $\mathbf{L}_K^F(u_t)$ 与重排序列表中的相应分数之和之间的比值：

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

其中，rank_i 和 rank_i^F 分别表示项目 i 在 $\mathbf{L}_K(u_t)$ 和 $\mathbf{L}_K^F(u_t)$ 中的排名位置。

对于提供方公平性，我们直接使用第3节中MMF的定义作为指标：

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

其中，$\mathbb{I}(\cdot)$ 为指示函数。

对于权衡性能，我们使用与公式（3）进行权衡的在线目标：

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

其中，$\lambda\ge 0$ 为权衡系数。

5.1.3 基线方法。我们选择了以下具有代表性的提供方公平重排序模型作为基线方法：**FairRec** [35]和**FairRec+** [8] 旨在保证提供方曝光量至少达到最大最小份额（MMS）。**CPFair** [32] 将权衡问题形式化为背包问题，并提出了一种贪心解法。

我们还选择了以下 MMF模型：**Welf** [16] 使用 Frank-Wolfe 算法来最大化最不受益物品的福利函数。然而，该方法是在离线设置下开发的；**RAOP** [5]

[2]https://github.com/XuChen0427/P-MMF。关于 MindSpore [12] 版本，请参阅 https://gitee.com/mindspore/models/tree/master/research/recommend/pmmf [3] https://www.yelp.com/dataset [4]http://jmcauley.ucsd.edu/data/amazon/

[5]http://cseweb.ucsd.edu/~wckang/Steam_游戏.json.gz

是一种最先进的在线资源分配方法。我们将物品视为资源、用户视为需求方，将其应用于推荐。

我们还将所提出的 P-MMF 与两种启发式 MMF 基线方法进行了比较：**K-邻居**：在每个时间步 t，仅推荐与累积曝光量最少的前K 个提供方相关联的物品；**最小正则化项**：在每个时间步 t，我们添加一个用于衡量目标提供商与最差提供商之间曝光差距的正则化项。附录 A.4 展示了算法细节。

5.1.4 实现细节。 对于所有模型中的超参数，学习率在 $[1e-2, 1e-4]/T^{1/2}$范围内进行调优，动量系数 α在 $[0.2, 0.6]$ 范围内进行调优。对于最大资源量（即权重） $\boldsymbol{\gamma}$，按照 [47], 中的做法，我们根据提供方提供的物品数量设置 $\boldsymbol{\gamma}$：$\boldsymbol{\gamma}_p = KT\eta|\mathbb{I}_p|/|\mathbb{I}|$， 其中 η 是控制资源丰富度的因子。在所有实验中，我们设置 $\eta = 1 + 1/|\mathbb{P}|$。我们分别基于 cvxpy [15] 及其 PyTorch 版本 [4], 实现了 P-MMF 的 CPU 和 GPU版本。

5.2 在小数据集上进行模拟

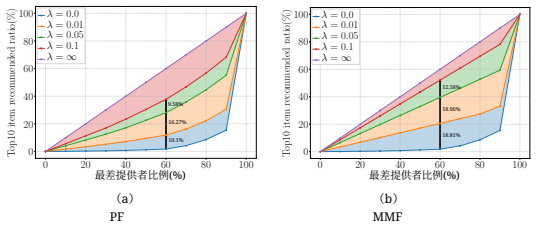
请注意，问题（2）可以使用MMF或比例公平作为其正则化项，尽管直接在大规模数据集上求解该问题较为困难。为验证该公式的正确性，并研究基于MMF和基于PF的正则化项对提供方公平推荐的不同影响，我们首先使用Yelp数据的5%进行了数值模拟。该数据包含844名用户、813件物品、10个提供方，时间范围的长度为 $T = 256$ 。

具体而言，我们使用基于PF或基于MMF的正则化项求解问题（2）（即将其提供方权重设置为均匀分布 $\boldsymbol{\gamma}_p = KT, \forall p \in \mathbb{P}$），并使用求解器 CVXPY [15]。在每个时间点 t，接收到推荐决策变量 $\mathbf{x}_t$ 后，我们计算了该数据集上的总体提供方曝光量 $\mathbf{e}$。随后绘制提供方曝光量的洛伦兹曲线 [19]，如图 3 所示。洛伦兹曲线常用于表示曝光分布。此处，它表示底部 x% 的提供方所占总体曝光量的百分比。换言之，对于底部 x%的提供方，它们获得的曝光量占总曝光量的百分比（y%）。请注意，右上方的对角线被称为绝对公平线。

从图 3 所示的曲线中，我们可以观察到，当 $\lambda \to \infty$ （即仅考虑提供方的公平性），PF 和 MMF 都能达到预期比例 $\boldsymbol{\gamma}$（即此处的曝光量相同）。然而，在考虑用户偏好后，基于MMF的正则化器倾向于优先考虑最差提供者的曝光量，而 PF 则不会。例如，当权衡系数 λ在 $[0.01, 0.05, 0.1]$之间变化时，MMF 会使曝光量最低的提供方中 60% 的曝光量增加 [18.81% 18.96% 12.58%]。如果使用基于PF的正则化器，则曝光量增加比例变为 [10.1% 16.27% 9.58%]。结果验证了，将推荐建模为由 MMF 正则化的资源分配问题，相较于 PF，能够带来更好的提供商公平推荐。

5.3 完整数据集上的实验结果

在本节中，我们使用第 4.2 节开发的在线算法进行了实验。在所有实验中，选择 BPR [39] 作为用于生成偏好的基础排序模型】【。



(a) PF (b) MMF

图 3：分别采用PF正则化项（a）和MMF正则化器（b）时提供方曝光量的洛伦兹曲线。实验基于Yelp数据集的5%进行。

分数。我们将时间范围的长度设为 $T = 256$。

表1报告了P-MMF和基线方法在所有数据集上根据指标 W_1@K得出的实验结果。带下划线的数字表示表现最佳的基线方法。为进行公平比较，所有基线方法均经过调参，并使用 W_1@K 作为评估指标。需要注意的是，在其他 λ 值下也观察到了类似的实验现象。

从报告的结果中，我们发现P-MMF在 W_1@K （$K = 5$, 10，20）方面优于所有基于PF的基线方法，验证了P-MMF能够为条件较差的提供方提供支持。我们还观察到，P-MMF优于所有基于MMF的基线方法，这表明P-MMF在保持较高用户偏好的同时，能够有效提升提供方公平性。

图 4 展示了四个具有不同排名大小 K的数据集上 NDCG@K 和 MMF@K 的帕累托前沿 [30]。通过调整模型的不同参数，并选择表现最佳的（NDCG@K，MMF@K）点，绘制出了帕累托前沿。在实验中，我们选取了 CPFair、最小正则化器、Welf 和 ROAP 作为基线方法，它们在所有基线方法中表现相对较好。

从帕累托前沿可以看出，所提出的 P-MMF 帕累托支配了基线方法（即 P-MMF 曲线位于右上角）。帕累托支配意味着，在相同的 NDCG@K 水平下，P-MMF 能够取得更好的 MMF@K；在相同的 MMF@K 水平下，P-MMF 能够取得更好的 NDCG@K。结果表明，P-MMF 在不过多牺牲用户效用的情况下，显著提升了条件较差的提供方的效用。

5.4 实验分析

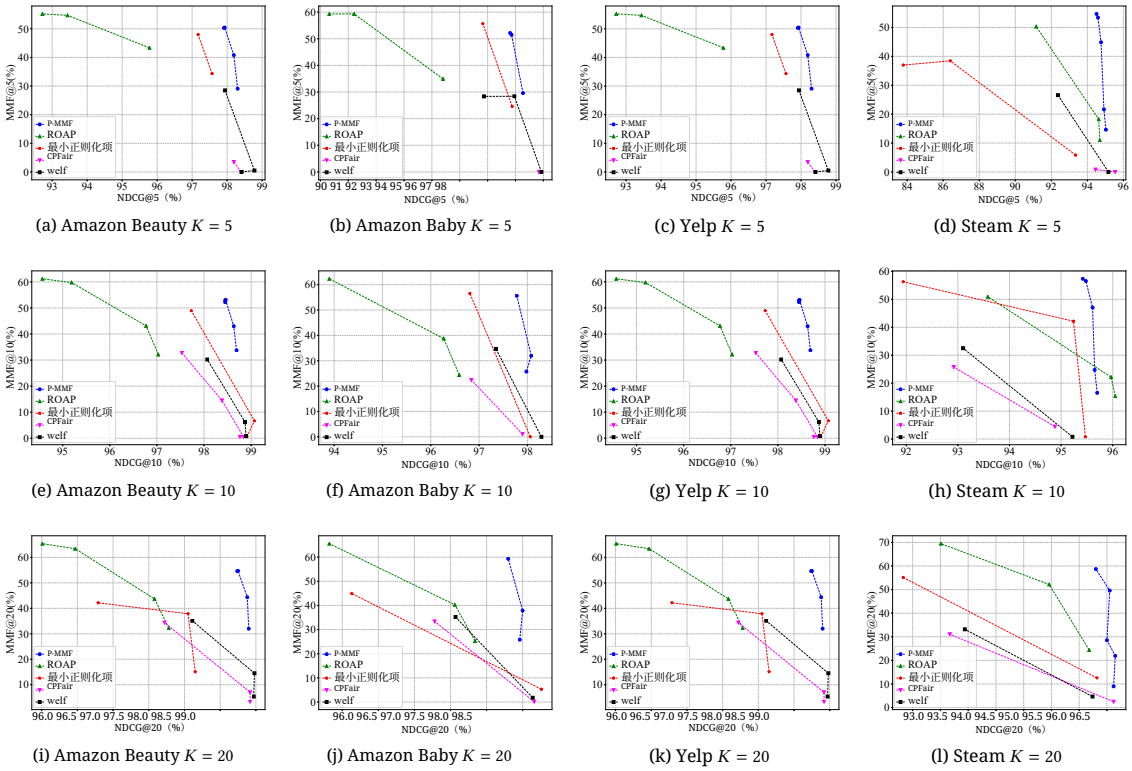
我们还在 Yelp 上进行了实验，以分析 P-MMF。

5.4.1 不同基础模型上的消融研究。 P-MMF 和其他基线方法都是重排序模型，用于对基础推荐模型输出的结果进行重新排名。为验证 P-MMF 在不同基础排序模型下的有效性，我们选择了另外三个基础模型来生成用户-项目偏好分数 $s_{u,\ i}$，包括 **DMF** [48]，该模型使用深度神经网络优化矩阵分解；**LINE** [42] 是一种基于图嵌入的矩阵分解模型；以及**LightGCN** [22]，该模型构建用户-项目交互图，并采用图卷积网络 [49] 进行推荐。所有实验也都在完整的 Yelp数据集上进行，并使用 $T = 256$。

表2展示了P-MMF搭配不同基础模型时的实验结果。我们观察到，P-MMF始终优于

表 1：我们的方法与基线方法在 Yelp、Beauty、Baby 和 Steam 上的性能比较。我们选择了权衡系数 $\lambda = 1$，在三种不同的排名规模 K 下考察我们方法的有效性。

		基于比例公平的基线方法			基于MMF的基线方法				我们的方法	
数据集	指标	FairRec	FairRec+	CPFair	k近邻	Welf	最小正则化器	ROAP	**P-MMF（我们的）**	改进幅度
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 $K = 20$

图 4：四个不同数据集在不同 top-K 排名下的帕累托前沿。

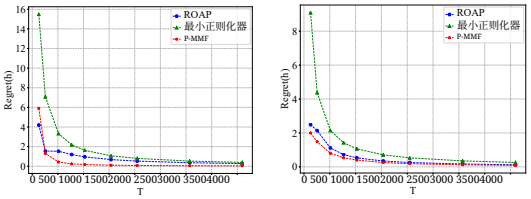
基于PF和基于MMF的基线方法。实验结果还验证了，在对不同基础模型输出的结果进行重新排名时，P-MMF比基线方法更为有效。

5.4.2 遗憾界实验。 为进一步展示 P-MMF 的有效性，我们直接计算遗憾值 Regret(h) $= W_{OPT} - W$，其中 W_{OPT}是公式（1）中定义的预言机性能，W是公式（3）中定义的在线性能。由于在大规模

数据集方面，我们在第5.2节创建的Yelp数据的5%上进行了实验。我们将P-MMF与当前最先进的提供商公平在线MMF基线方法进行了比较：最小正则化器和ROAP。实验采用的排名大小为 $K = 10$。请注意，在实验中，我们固定了用户到达数量 N，遗憾值通过对 N/T个样本求和进行计算，其中 T为时间范围的长度。根据定理3，求和遗憾值Regret(h)与 $O(T^{1/2}N/T) = O(N/T^{1/2})$具有可比性。

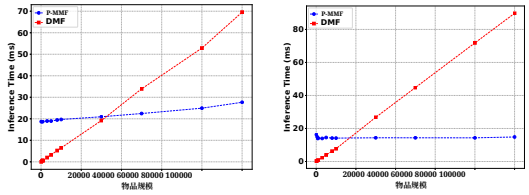
表2：在不同基础推荐模型下，我们的方法与基线方法在Yelp上的性能比较。我们在三种不同的排名规模 K 上选择权衡系数 $\lambda = 1$，以研究我们方法的有效性。

		基于PF的基线方法			基于MMF的基线方法				我们的方法	
基础模型	指标	FairRec	FairRec+	CPFair	k近邻	Welf	最小正则化器	ROAP	**P-MMF（我们的方法）**	改进
DMF [48]	W_1@5	3.1982	4.087	5.070	4.114	4.981	5.198	5.190	**5.272**	1.4%
	W_1@10	6.364	7.829	9.929	7.957	9.867	9.968	9.878	**10.003**	0.3%
	W_1@20	12.643	15.148	19.372	15.208	19.292	19.410	19.324	**19.448**	0.2%
LINE [42]	W_1@5	2.323	4.074	5.237	4.269	5.153	5.239	5.159	**5.407**	3.2%
	W_1@10	4.614	7.866	10.087	8.099	10.064	10.124	10.042	**10.280**	1.5%
	W_1@20	9.164	15.041	19.814	15.829	19.826	19.450	19.723	**19.973**	0.7%
LightGCN [22]	W_1@5	2.729	3.714	4.531	3.963	4.633	4.782	4.718	**4.837**	1.2%
	W_1@10	5.468	7.2368	8.982	7.584	8.975	9.016	9.144	**9.283**	1.5%
	W_1@20	10.927	13.995	17.799	14.856	17.774	17.812	17.871	**18.035**	0.9%



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

图5：遗憾值与时间分段长度的关系 T。



（a）CPU版本 （b）GPU版本

图6：CPU和GPU实现下，关于物品总数的每次用户访问在线推理时间。

图5展示了求和遗憾值Regret(h)相对于 T的曲线。图（a）和图（b）分别展示了权衡系数 λ 设置为1和0.1时的曲线。从图5可以看出，P-MMF的遗憾界低于其他在线模型，尤其是在 λ 较大时。此外，我们可以看到，P-MMF的遗憾界随着 T的增加而降低，这验证了理论分析结果。对于其他 λ 值，也观察到了类似的结果。

5.4.3 在线推理时间。 我们进行了实验，以研究 P-MMF 和最具实用性的推荐系统模型 DMF [48] 的在线推理时间。在实际推荐系统中，提供方的数量相对较少且较为稳定。然而，用户和物品的数量通常非常庞大，并且增长迅速。因此，在保持用户和提供方数量不变的情况下，我们在不同物品数量下测试了 P-MMF 和 DMF 的 CPU和GPU实现的推理时间。GPU实现基于 PyTorch [34]。

图6展示了每次用户访问的推理时间（毫秒）与物品规模之间的关系曲线。可以看到，P-MMF 的 CPU和GPU版本在线推理分别只需约 20–40 毫秒和 17–18 毫秒。此外，我们可以看到，随着物品数量的增加，P-MMF 的推理时间并未显著增加

物品。例如，将物品规模从 0 增加到 200000，P-MMF CPU 版本的在线推理仅需略多一点时间（19ms）。GPU版本的推理时间几乎保持不变。相比之下，DMF 的推理时间迅速增加：CPU 和 GPU版本都需要更多时间（约 65ms）。通过定理1中的对偶问题分析，可以轻松解释这一现象。我们看到参数

P-MMF 的规模是提供方规模，远小于物品规模 $|\mathcal{P}| \ll |\mathcal{I}|$。因此，在线推理时间对物品数量并不敏感。我们得出结论：由于在线计算成本低且稳定，即使物品数量快速增长，P-MMF 也能高效适应实际的在线推荐场景。

6 结论

本文提出了一种名为P-MMF的新型重排序模型，旨在平衡提供方公平性与用户偏好。首先，我们将该任务形式化为一个由最大最小公平性指标正则化的资源分配问题。随后，为适应推荐中的在线场景，我们提出了一种动量梯度下降方法，在对偶空间中对资源分配进行在线学习。理论分析表明，P-MMF的遗憾值是有界的。在四个大规模数据集上的广泛实验结果表明，P-MMF优于基线方法，并且帕累托支配最先进的提供方公平基线。

致谢

本工作得到了中国国家重点研发计划（2019年YFE0198200）、国家自然科学基金（61872338、62006234、61832017）、智能社会治理交叉学科平台、“双一流”建设重大创新与规划交叉学科平台，以及中国人民大学2023年优秀创新人才培养资助项目的资助。

参考文献

[1] Himan Abdollahpouri、Gediminas Adomavicius、Robin Burke、Ido Guy、Dietmar Jannach、Toshihiro Kamishima、Jan Krasnodebski 和 Luiz Pizzato。2020年。多利益相关者推荐：综述与研究方向。用户建模与用户适应性交互 30，1（2020年），127–158。[2] Himan Abdollahpouri 和 Robin Burke。2019年。多利益相关者推荐及其与多方公平性的联系。arXiv预印本 arXiv:1907.13158（2019年）。[3] Himan Abdollahpouri、Masoud Mansoury、Robin Burke 和 Bamshad Mobasher。2019年。推荐中流行度偏差的不公平性。arXiv预印本 arXiv:1907.13286（2019年）。[4] Akshay Agrawal、Brandon Amos、Shane Barratt、Stephen Boyd、Steven Diamond 和 J Zico Kolter。2019年。可微凸优化层。神经信息处理系统进展 32（2019年）。[5] Santiago Balseiro、Haihao Lu 和 Vahab Mirrokni。2021年。正则化在线分配问题：公平性及其他方面。载于国际机器学习会议。PMLR，630–639。[6] Santiago R. Balseiro、Haihao Lu、Vahab Mirrokni 和 Balasubramanian Sivan。2022年。从在线优化到PID控制器：带动量的镜像下降。https://doi.org/ 10 48550 2202 06152 。/ARXIV。 。[7] Dimitris Bertsimas、Vivek F Farias 和 Nikolaos Trichakis。2011年。公平性的代价。Operations Research 59，1（2011年），17–31。[8] Arpita Biswas、Gourab K Patro、Niloy Ganguly、Krishna P Gummadi 和 Abhijnan Chakraborty。2021年。迈向双边平台中的公平推荐。ACM网络学报（TWEB）16，2（2021年），1–34。[9] Thomas Bonald、Laurent Massoulié、Alexandre Proutiere 和 Jorma Virtamo。2006年。最大最小公平性、比例公平性和平衡公平性的排队分析。Queueing Systems 53，1（2006年），65–84。[10] Joseph L Bower。1972年。管理资源分配过程：企业规划与投资研究。Irwin Ho mewood。[11] Abhijnan Chakraborty、Aniko Hannak、Asia J Biega 和 Krishna Gummadi。2017年。共享经济平台中的公平分享。载于推荐系统中的公平性、问责制与透明度——负责任推荐研讨会。[12] Lei Chen。2021年。使用MindSpore进行深度学习与实践。Springer Nature。[13] Wang Chi Cheung、Guodong Lyu、Chung-Piaw Teo 和 Hai Wang。2020年。利用离线模拟进行在线规划。可见于SSRN 3709882（2020年）。[14] Nikhil R Devanur、Kamal Jain、Balasubramanian Sivan 和 Christopher A Wilkens。2011年。资源分配问题的近最优在线算法与快速近似算法。载于第12届ACM电子商务会议论文集。29–38。[15] Steven Diamond 和 Stephen Boyd。2016年。CVXPY：一种用于凸优化的嵌入Python的建模语言。《机器学习研究杂志》17，1（2016年），2909–2913。[16] Virginie Do、Sam Corbett-Davies、Jamal Atif 和 Nicolas Usunier。2021年。通过洛伦兹支配实现排名中的双边公平性。神经信息处理系统进展 34（2021年），8596–8608。[17] VR Dondeti 和 H Emmons。1996年。具有多重分配的最大-最小匹配问题。《优化理论与应用杂志》91，2（1996年），491–511。[18] KJ Erickson。2018年。用他们自己的话说：卖家为何对亚马逊感到厌倦。
h
t
t
p
s://
m
e
dium.com/public-market/in-their-own-words-why-sellers-are-fed-up-with-amazon-e 97 44 7 18 da f f Joseph L Gastwirth。 。洛伦兹曲线的一般定义。[19] 1971《Econometrica：计量经济学会杂志》（1971年），1037–1039。[20] Yingqiang Ge、Shuchang Liu、Ruoyuan Gao、Yikun Xian、Yunqi Li、Xiangyu Zhao、Changhua Pei、Fei Sun、Junfeng Ge、Wenwu Ou 等。2021年。迈向推荐中的长期公平性。载于第14届ACM网络搜索与数据挖掘国际会议论文集。445–453。[21] Elizabeth Gómez、Ludovico Boratto 和 Maria Salamó。2022年。协同推荐系统中跨洲的提供方公平性。《信息处理与管理》59，1（2022年），102719。[22] Xiangnan He、Kuan Deng、Xiang Wang、Yan Li、Yongdong Zhang 和 Meng Wang。2020年。LightGCN：简化并增强用于推荐的图卷积网络。载于第43届国际ACM SIGIR信息检索研究与开发会议论文集。639–648。[23] Ehud Kalai 和 Meir Smorodinsky。1975 年。纳什谈判问题的其他解。《Econometrica：计量经济学会杂志》（1975年），513–518。[24] Wang-Cheng Kang 和 Julian McAuley。2018年。自注意力序列推荐。载于2018年IEEE国际数据挖掘会议（ICDM）。IEEE，197–206。[25] Julian Lamont。2017年。分配正义。Routledge。[26] Johannes Langemeyer 和 James JT Connolly。2020年。将正义理念融入城市生态系统服务研究与实践。《Environmental Science & Policy》109（2020年），1–14。[27] Xiaocheng Li、Chunlin Sun 和 Yinyu Ye。2020年。用于二进制整数规划和在线线性规划的简单快速算法。《神经信息处理系统进展》33（2020年），9412–9421。[28] Xiaocheng Li 和 Yinyu Ye。2021年。在线线性规划：对偶收敛、新算法与遗憾界。《Operations Research》（2021年）。[29] Yunqi Li、Hanxiong Chen、Zuohui Fu、Yingqiang Ge 和 Yongfeng Zhang。2021年。推荐中的面向用户的公平性。载于《2021年万维网大会论文集》。624–632。[30] Alexander V Lotov 和 Kaisa Miettinen。2008年。可视化帕累托前沿。载于《多目标优化》。Springer，213–243。[31] David Matsumoto 和 Linda Juang。2016年。文化与心理学。Cengage Learning。[32] Mohammadmehdi Naghiaei、Hossein A Rahmani 和 Yashar Deldjoo。2022年。CPFair：面向推荐系统的个性化消费者与生产者公平性重新排名。arXiv预印本 arXiv:2204.08085（2022年）。[33] John F Nash Jr。1950年。谈判问题。《Econometrica：计量经济学会期刊》（1950年），155–162。[34] Adam Paszke、Sam Gross、Soumith Chintala、Gregory Chanan、Edward Yang、Zachary DeVito、Zeming Lin、Alban Desmaison、Luca Antiga 和 Adam Lerer。2017年。PyTorch中的自动微分。（2017年）。[35] Gourab K Patro、Arpita Biswas、Niloy Ganguly、Krishna P Gummadi 和 Abhijnan Chakraborty。2020年。FairRec：双边平台中个性化推荐的双边公平性。载于《2020年万维网大会论文集》。1194–1204。[36] Larry L Peterson 和 Bruce S Davie。2007年。计算机网络：一种系统方法。Elsevier。[37] Ning Qian。1999年。梯度下降学习算法中的动量项。《神经网络》12，1（1999年），145–151。[38] Hossein A Rahmani、Yashar Deldjoo、Ali Tourani 和 Mohammadmehdi Naghiaei。2022年。兴趣点推荐中活跃用户的不公平性与流行度偏差。arXiv预印本 arXiv:2202.13307（2022年）。[39] Steffen Rendle、Christoph Freudenthaler、Zeno Gantner 和 Lars Schmidt-Thieme。2012年。BPR：基于隐式反馈的贝叶斯个性化排序。arXiv预印本 arXiv:1205.2618（2012年）。[40] Dimitris Sacharidis。2019年。具有公平性的Top-N群组推荐。载于《第34届ACM/SIGAPP应用计算研讨会论文集》。1663–1670。[41] Dimitris Serbos、Shuyao Qi、Nikos Mamoulis、Evaggelia Pitoura 和 Panayiotis Tsaparas。2017年。套餐到群组推荐中的公平性。载于《第26届国际万维网大会论文集》。371–379。[42] Jian Tang、Meng Qu、Mingzhe Wang、Ming Zhang、Jun Yan 和 Qiaozhu Mei。2015年。LINE：大规模信息网络嵌入。载于《第24届国际万维网大会论文集》。1067–1077。[43] Tom R Tyler 和 E Allan Lind。2002年。程序正义。载于《法律正义研究手册》。Springer，65–92。[44] Tom R Tyler 和 Heather J Smith。1995年。社会正义与社会运动。（1995年）。[45] Lequn Wang 和 Thorsten Joachims。2021年。双边市场中排名的用户公平性、项目公平性与多样性。载于《2021年ACM SIGIR信息检索理论国际会议论文集》。23–41。[46] David Williams。1991年。鞅概率论。剑桥大学出版社。[47] Yao Wu、Jian Cao、Guandong Xu 和 Yudong Tan。2021年。TFROM：面向客户和提供方的双边公平性感知推荐模型。载于《第44届国际ACM SIGIR信息检索研究与发展会议论文集》。1013–1022。[48] Hong-Jian Xue、Xinyu Dai、Jianbing Zhang、Shujian Huang 和 Jiajun Chen。2017年。推荐系统的深度矩阵分解模型。载于《国际人工智能联合会议》，第17卷。澳大利亚墨尔本，3203–3209。[49] Si Zhang、Hanghang Tong、Jiejun Xu 和 Ross Maciejewski。2019年。图卷积网络：综述。《Computational Social Networks》6，1（2019年），1–23。[50] Xiao Zhang、Sunhao Dai、Jun Xu、Zhenhua Dong、Quanyu Dai 和 Ji-Rong Wen。2022年。通过奖励修正抵消音乐流媒体推荐中的用户注意力偏差。载于《第28届ACM SIGKDD知识发现与数据挖掘会议论文集》。2504–2514。[51] Xiao Zhang、Haonan Jia、Hanjing Su、Wenhan Wang、Jun Xu 和 Ji-Rong Wen。2021年。针对具有延迟反馈的流媒体推荐进行反事实奖励修正。载于《第44届国际ACM SIGIR信息检索研究与发展会议论文集》。41–50。

A 附录

A.1 定理1的证明

证明。对于最大最小公平性，正则化项为 $r(\mathbf{e}) = \min_{p\in\mathcal{P}}\left(\mathbf{e}p/\boldsymbol{\gamma}_p\right)$，我们可以很容易地证明，曝光向量$\mathbf{e}$可以表示为决策变量$\mathbf{x}_t$与物品-提供商邻接矩阵$\mathbf{A}$的点积：$\mathbf{e} = \sum_{t=1}^{T}\left(\mathbf{A}^\top\mathbf{x}_t\right)$. 然后，我们将$\mathbf{e}$ 视为辅助变量，式（1）可以写为：

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

其中 $\mathrm{X} = \{\mathbf{x}_t|\mathbf{x}_t \in 0,1^n \wedge \sum_{i\in\mathcal{I}}\mathbf{x}_{ti} = K\}$。然后，我们使用拉格朗日乘子$\boldsymbol{\mu}\in\mathbb{R}^{|\mathcal{P}|}$将约束移至目标函数中：

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

其中 $\mathrm{D} = \{\boldsymbol{\mu}|r^*(-\boldsymbol{\mu}) < \infty\}$ 是对偶变量 $\boldsymbol{\mu}$ 的可行域。根据 Balseiro 等人 [5], 的引理 1，我们有 D 是凸集，且正正交象限位于 D 的衰退锥内。

□

A.2 定理2的证明

证明。令变量 $\mathbf{z}p = (\mathbf{e}p/\boldsymbol{\gamma}_p - 1)$，则有：

$$r^*(\boldsymbol{\mu}) = \max_{\mathbf{e}\le\boldsymbol{\gamma}}\left[\min\left(\mathbf{e}_p/\boldsymbol{\gamma}_p\right) + \boldsymbol{\mu}^\top\mathbf{e}/\lambda\right]$$
$$= \boldsymbol{\mu}^\top\boldsymbol{\gamma}/\lambda + 1 + \max_{\mathbf{z}_p\le 0}\left[\min(\mathbf{z}_p) + 1/\lambda\sum_{p\in\mathcal{P}}\boldsymbol{\mu}_p\boldsymbol{\gamma}_p\mathbf{z}_p\right]$$

令 $s(\mathbf{z}) = \min p\mathbf{z}p$，且 $\mathrm{v} = (\boldsymbol{\mu}\odot\boldsymbol{\gamma})/\lambda$，其中$\odot$为Hadamard积。然后我们定义 $s^*(\mathbf{v}) = \max_{\mathbf{z}\le 0}\left(s(\mathbf{z}) + \mathbf{z}^T\mathbf{v}\right)$。我们首先证明，如果 $\sum_{p\in\mathcal{S}}\mathrm{v}$ $_p \ge -1, \forall\mathrm{S} \in \mathrm{P}s$，那么 $s^*(\mathbf{v}) = 0$，且 $\mathbf{z} = 0$是最优解；否则 $s^*(\mathbf{v}) = \infty$。

我们可以等价地将 $\mathrm{D} = \{\mathrm{v}|\sum_{p\in\mathcal{S}}\mathrm{v}_p \ge -1, \forall\mathrm{S} \in \mathrm{P}s\}$ 写为。我们首先证明，对于$\mathbf{v} \notin \mathrm{D}$，有 $s^*(\mathbf{v}) = \infty$。假设存在一个子集 $\mathrm{S} \in \mathrm{P}s$，使得 $\sum_{p\in\mathcal{S}}\mathrm{v}_p < -1$。对于任意 $b > 1$，我们可以得到一个可行解：

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

然后，由于该解是可行的，并且 $s(\mathbf{z}) = -b$，我们得到

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

令 $b \to \infty$，我们有 $s^*(\mathbf{v}) \to \infty$.

接着我们证明，对于$\mathbf{v} \in \mathrm{D}$，有 $s^*(\boldsymbol{\mu}) = 0$。注意，$\mathbf{z} = 0$ 是可行的。因此，我们有

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

然后我们有 $\mathbf{z} \le 0$，不失一般性地，向量 $\mathbf{z}$ 按递增顺序排列，即 $\mathbf{z}_1 \le \mathbf{z}_2, \cdots \le \mathbf{z}|\mathcal{P}|$。目标值为

$$s^*(\mathbf{v}) = \mathbf{z}_1 + \sum_{j\in|\mathcal{P}|}\mathbf{z}_j\mathbf{v}_j$$
$$= \sum_{j=1}^{|\mathcal{P}|}\left(\mathbf{z}_j - \mathbf{z}_{j+1}\right)\left(1 + \sum_{i=1}^{j}\mathbf{v}_j\right) \le 0.$$

因此，对于$\mathbf{v} \in \mathrm{D}$，我们可以有 $s^*(\boldsymbol{\mu}) = 0$，并且我们有

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

□

A.3 定理3的证明

证明。首先，在实践中，我们将用户-项目偏好分数 $s_{u,i}$归一化为 $[0, 1]$。因此，$\sum_{t=1}^{T} g(\mathbf{x}_t)/T \le K$。在最大-最小正则化器 $r(\mathbf{e})$ 中。将其上界简写为 $\bar{r}$。在实践中，$\bar{r} \le 1$ 我们有

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

我们将算法 1 的停止时间 τ 定义为提供方获得最大曝光量的首次时刻，即

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

注意，τ 是一个随机变量 .

类似地，遵循 Balseiro 等人的证明思路。 [5], 首先，我们分析目标函数的原始性能。其次，我们通过动量梯度下降对互补松弛项进行界定。最后，我们将其结合起来，得到最终的遗憾界。

原始性能证明：考虑时间 $t < \tau$，推荐器的操作不会违反资源约束。因此，我们有：

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

并且我们有 $\mathbf{e}_t = \arg\max_{\mathbf{e}\le\boldsymbol{\gamma}}\{r(\mathbf{e}) + \boldsymbol{\mu}^\top\mathbf{e}/\lambda\}$

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

我们对当前时间步 t 的原始函数取期望：

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

考虑过程 $Z_t = \sum_{j=1}^{T}\boldsymbol{\mu}_j^t(-\mathbf{A}^\top\mathbf{x}_t + \mathbf{e}_t) - \mathbf{E}\left[\boldsymbol{\mu}_t^T(-\mathbf{A}^\top\mathbf{x}_t + \mathbf{e}_t)\right]$是一个鞅过程。鞅过程中的可选停止定理 [46] 意味着 $\mathbb{E}[Z_\tau] = 0$。考虑变量 $w_t(\boldsymbol{\mu}_t) = \boldsymbol{\mu}_t^T(-\mathbf{A}^\top\mathbf{x}_t + \mathbf{e}_t)$，我们有

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

算法2： 最小正则化项的在线学习

输入： 用户到达集合 $\{u_i\}_{i=1}^N$、时间分段大小 T、排名大小 K、用户-物品偏好分数 $s_{u,\ i},\ \forall u,\ i$、物品-提供商邻接矩阵 $\mathbf{A}$、最大资源量 $\boldsymbol{\gamma}$ 以及权衡系数 λ。

输出： 决策变量 $\{\mathbf{x}_i,\ i=1,\ 2,\ \cdots\ N\}$

对于 $n=1,\ \cdots\ N/T$ **执行**

初始对偶解 $\boldsymbol{\mu}_1=\mathbf{0}$。剩余资源 $\boldsymbol{\beta}_1=\boldsymbol{\gamma}$ 和动量梯度 $\mathbf{g}_0=\mathbf{0}$。

对于 $t=1\cdots,T$ **do**

接收 u_{nT+t}

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

// 做出推荐：

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

// 更新剩余资源：

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

结束循环

结束循环

此外，在 MMF 中，对偶函数 W_{Dual} 在定理1中已证明是凸函数，我们有

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{\tau}g(\mathbf{x}_t)/T+\lambda r(\mathbf{e}_t)\right]&=\mathbb{E}\left[\sum_{t=1}^{\tau}W_{Dual}(\boldsymbol{\mu}_t)\right]-\mathbb{E}\left[\sum_{t=1}^{\tau}w_t(\boldsymbol{\mu}_t)\right]\\&\le\mathbb{E}\left[\tau W_{Dual}(\widetilde{\boldsymbol{\mu}_\tau})\right]-\mathbb{E}\left[\sum_{t=1}^{\tau}w_t(\boldsymbol{\mu}_t)\right],\end{aligned}\tag{8}$$

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

互补松弛证明 接下来我们旨在证明互补松弛 $\sum_{t=1}^{T}w_t(\boldsymbol{\mu}_t)-w_t(\boldsymbol{\mu})$ 是有界的。假设存在 $G,\ s.t.$ 梯度范数有界 $\|\tilde{\mathbf{g}}_t\|\le G$。则有：

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

其中投影函数 $\|\boldsymbol{\mu}-\boldsymbol{\mu}_t\|_{\boldsymbol{\gamma}}^2$ 是 σ–强凸的。

接下来我们证明公式(9)中的不等式。根据 [6], 中的定理1，我们有

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

and

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

假设存在 $H>0$，使得 $\|\boldsymbol{\mu}_t-\boldsymbol{\mu}_0\|_{\boldsymbol{\gamma}^2}^2\le H$。根据柯西–施瓦茨不等式，结果得证。令 $M=\frac{H}{\eta}+\frac{G^2}{1-\alpha()\sigma}\eta(T-1)+\frac{G^2}{2\ 1-\alpha}()^2\sigma\eta$。现在我们选择一个合适的 $\boldsymbol{\mu}$，使得互补堆叠性可以得到进一步界定。

对于 $\boldsymbol{\mu}=\hat{}\boldsymbol{\mu}+\theta$，其中 $\theta\in\mathbb{R}^{|P|}$ 为非负数，稍后确定，且 $\hat{}\boldsymbol{\mu}=\arg\max_{\boldsymbol{\mu}}-\boldsymbol{\mu}^\top(\sum_{i=1}^{T}\mathbf{A}^\top\mathbf{x}_t)/\lambda$。根据约束 $\mathbf{e}=\sum_{i=1}^{T}\mathbf{A}^\top\mathbf{x}_t$，可得

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

注意，在定理2的证明中，可行域 D 是衰退锥，因此，$\boldsymbol{\mu}\in\mathrm{D}$。

因此，我们有

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

将它们合并： 在最大-最小公平条件下，我们得到

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

因此，结合式（10、12、13），遗憾值 Regret(h) 可界定为：

$$\begin{aligned}\mathrm{Regret}(h)&=\mathbb{E}\left[W_{OPT}-W\right]\\&\le\mathbb{E}\left[W_{OPT}-\sum_{t=1}^{\tau}\left(g(\mathbf{x}_t)/T-\lambda r(\mathbf{A}^\top\mathbf{x}_t/\boldsymbol{\gamma})\right)\right]\\&\le\mathbb{E}\left[W_{OPT}-\tau W_{Dual}(\widetilde{\boldsymbol{\mu}_t})+\sum_{t=1}^{\tau}w_t(\boldsymbol{\mu}_t)+\sum_{t=1}^{T}(\mathbf{e}_t-\mathbf{A}^\top\mathbf{x}_t)\right]\\&\le\mathbb{E}\left[(T-\tau)(K+\lambda\bar{r})+\sum_{t=1}^{T}w_t(\hat{\boldsymbol{\mu}})+\sum_{t=1}^{\tau}w_t(\theta)\right]+M\\&\le(T-\tau)(K+\lambda\bar{r}+\lambda K)+\sum_{t=1}^{\tau}w_t(\theta)+M.\end{aligned}\tag{12}$$

令 $C=K+\lambda\bar{r}+\lambda K$，然后设置 $\theta=C\min^p\boldsymbol{\gamma}_p\mathbf{u}^p$，其中 $\mathbf{u}^p$ 是第 p 个单位向量。我们有

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

那么，当我们设置 $\eta=O(T^{-1/2})$ 时，Regret(h) $\le M+C/($ $\min^p\boldsymbol{\gamma}_p)$，Regret(h) 与 $O(T^{1/2})$ 相当。由此可得结果。

□

A.4 Min-reguarlizer 的算法

在本节中，我们提出了一种用于提供商最大-最小公平在线应用的启发式方法，称为 Min-reguarlizer。该方法包含一个正则化项，用于衡量目标提供方与最差提供商之间的曝光差距。具体算法如算法2所示。相关符号与算法 1 中的 P-MMF 相同。