arXiv:2301.03862v1 [cs.DS] 10 Jan 2023

多群体比例公平匹配

SayanBandyapadhyay * Fedor V.Fomin†. Tanmay Inamdar‡ KirillSimonov§

摘要

由于应对偏见和歧视的需求日益增长，公平算法的研究已成为机器学习和人工智能领域的主流。沿着这一方向，研究人员考察了传统优化问题的公平版本，包括聚类、回归、排名和投票。然而，大多数研究都致力于设计启发式算法，而这类算法往往无法对解的质量提供任何保证。在本文中，我们研究具有比例公平概念的匹配问题。比例公平是最流行的群体公平概念之一，它要求每个组在最终选择中的代表比例与其规模相称。具有比例公平性的匹配，或更常见的称谓——比例公平匹配，由 [Chierichetti 等人，AISTATS，2019], 提出；他们研究的问题仅涉及两个组。然而，在许多实际应用中，组的数量虽然通常是一个较小的常数，却多于两个。在本文中，我们迈出第一步，开始理解涉及两个以上组的比例公平匹配的计算复杂性。我们设计了精确算法和近似算法，在匹配质量和时间复杂性方面均能提供合理保证。我们还为这些算法给出了相应的困难性界。

1 引言

机器学习（ML）算法如今无处不在，在决策过程中持续发挥着关键作用，而决策对人类生活有着不可估量的影响。这些基于过往实例训练的算法功能极其强大，且大多数时候都能给出正确的解，不会出错。然而，近年来，这些算法因对代表性不足群体存在偏见而受到批评 [4, 21, 27, 10]。因此，研究人员致力于了解偏见是如何在机器学习流程中引入的，以及是否有可能消除这些偏见。这项研究催生了一个名为机器学习公平性 的新兴子领域。迄今为止，这一新兴子领域的研究大致可分为两类。第一类研究探讨不同的公平性概念及其相互作用 [11, 22, 25, 19, 39, 18]。这些研究主要表明，不存在能够涵盖所有情形的普适公平性定义，而且不同的公平性概念无法同时得到满足。第二类研究则探讨经典问题的公平版本，并融入第一类研究提出的恰当公平性概念。值得注意的是，所考察的问题包括聚类 [15, 32, 7, 5], 回归 [34, 9, 1], 排名[13], 投票 [12]和匹配 [16]。

本文研究比例公平匹配问题。匹配自然地出现在机器学习的多种应用中，例如：将产品分配给客户 [44]；将学生分配给学校[40]；将审稿人分配给稿件 [14]；以及将工人分配给企业 [2]。已有大量研究探讨匹配的公平版本 [16, 31, 28, 35, 38, 26]。在这些不同的匹配概念中，我们的工作与（α， β）平衡匹配最为相关 [16]。（α， β）平衡匹配的提出

*波特兰州立大学，美国
†卑尔根大学，挪威
‡卑尔根大学，挪威§维也纳工业大学¨，
奥地利

由 [16] 将 公平性 与最大基数匹配结合而成。比例公平性基于差别影响这一概念 [25], ，其在匹配情境下的定义如下。若一个匹配是（α， β）平衡的，或称比例公平的，则来自每个群体的边数与匹配规模之比至少为 α 且至多为 β。

作为比例公平匹配的一个启发性示例，我们来看电子商务中的产品推荐问题。随着数字营销和广告的发展，如今企业希望采用更精细的方法，以便触达目标客户群体。这些群体可能代表基于性别、年龄组、地理位置等因素划分的某些人口统计分类。因此，群体数量通常是一个较小的常数。尤其是在当今的情境下，人们希望找到涉及所有目标群体中的客户，并对所有这些群体产生平衡影响的赋值。这个分配问题可以建模为客户与产品之间的比例公平匹配。在现实情况下，可能需要为一个客户分配多个产品，也需要将一个产品分配给多个客户。这可以通过迭代地计算多个匹配来实现，同时从输入图中移除已匹配的边。

在一项开创性工作中， [16] 在群体数量为 2 时，得到了一个多项式时间的 3/2-近似算法。然而，在许多现实场景中，例如上述示例，很自然地可以假设目标群体数量超过 2。不幸的是，[16]的算法充分利用了群体数量 $\ell = 2$这一事实。当群体数量超过两个时，如何调整或扩展他们的算法尚不明确。在我们的工作之前，对于 $\ell > 2$ 组，唯一已知的算法是一个 $n^{O(\ell)}$时间的随机精确算法 [16]。该算法的运行时间对群体数量呈现“糟糕的”指数级依赖，也就是说，其运行时间并非关于 n的固定多项式。因此，如果 ℓ 增大，该算法很快就会变得不切实际。我们对比例公平匹配的研究受以下问题驱动。当群体数量 ℓ 超过两个时，是否存在具有性能保证的高效比例公平匹配算法?

1.1 我们的结果与贡献

在这项工作中，我们针对具有任意数量组的 Proportionally Fair Matching 问题取得了若干结果。

- 首先，我们证明，对于任意数量的组，该问题都极难求解：除非指数时间假设（ETH） [33] 不成立，否则即使在路径图上，也无法在 $2^{o(\ell)}n^{O(1)}$ 时间内获得任何近似算法。

- 为补充我们的困难性结果，我们设计了一种 $1/4\ell$-近似算法，其运行时间为 $2^{O(\ell)}n^{O(1)}$ 。当 $|$最优解$|$ 大于 $4\ell^2$ 时，我们的算法对下界（α）和上界（β）的违反程度至多为一个乘法因子（$1 + 4\ell/|$最优解$|$），其中最优解是任意一个最优解。因此，违反因子至多为 $1 + 1/\ell$，并且随着 $|$最优解$|$趋于渐近值而趋近于 1。

- 我们还研究了该问题的一个受限情形，称为 β-受限情形： [16],此时我们只有上界，也就是说，某些组可能没有任何边。在这种情况下，我们可以将近似因子提升至 $1/2\ell$，并将运行时间缩短至多项式时间。

- 最后，我们证明，问题的参数化版本——即寻找大小为 k 的比例公平匹配——可以在 $2^{O(k)}n^{O(1)}$ 时间内精确求解。因此，以 k 为参数，该问题是固定参数可解的。

我们的所有算法都基于简单的方案。我们的近似算法采用迭代剥离方案，在每次迭代中提取一个彩虹匹配，其中最多包含一条边

来自每个组。精确算法基于对著名的颜色编码方案的一种非平凡应用 [3]。这些算法分别见第3、4和5节。困难性证明见第6节。

1.2 相关工作

近年来，研究人员提出并研究了多种不同的公平性概念，例如差别影响 [25]，统计均等 [47, 36]，个体公平 [22] 和群体公平 [23]。Kleinberg 等人 [39] 提出了三种公平性概念，并证明从理论上讲，不可能同时满足这三种概念。类似的研究还可参见 [18, 17]。

关于多个受保护群体的比例公平性已得到广泛研究，其依据是差别影响 [25]。Bei 等人 [6] 研究了比例候选人选择问题，其目标是从给定集合中选出具有不同属性的候选人子集，同时满足特定的比例公平约束。Goel 等人 [29] 研究了非歧视且符合比例公平的分类器学习问题，并提出了对数加权和技术。比例公平性也在联邦学习 [49]背景下得到研究。此外，比例公平性还在许多优化问题中得到研究，包括投票 [24]，调度 [37, 41]，肾脏交换 [45],和旅行商问题 [43]。

文献中已研究了若干不同的公平匹配问题。 [31] 研究了公平b-匹配，其中每个顶点的匹配偏好以等级表示，目标是尽可能避免将顶点分配给排名较高的偏好。 [28]，研究了设计即公平匹配：该方法不计算单个匹配，而是计算所有可行匹配上的概率分布，从而保证个体公平。另请参见 [35, 38]。

除了公平匹配的相关研究外，许多受约束版本也得到了研究 [46, 8]。[46] 研究了有界颜色匹配（BCM）问题，其中边带有颜色，且每个颜色类中只能选取给定数量的边。BCM 是 3-集合打包的特例，因此存在一个 3/4-近似算法 [46]。我们注意到， ProportionallyFair Matching 的 β限制情形是 BCM 的特例，因此在这种情形下可得到一个 3/4-近似算法，而上界可能被违反 3/4 倍。应将这一违反因子与我们的违反因子进行比较，后者渐近趋于 1。

2 预备知识

对于整数 $\ell \geq 1$，令 $[\ell] := \{1, 2, \ldots, \ell\}$。考虑任意无向 n 顶点图 $G = (V, E)$，其中 E 中的边由 $C = \{1, \ldots, \ell\}$ 中的颜色着色。函数 $\chi\colon E \to C$ 描述颜色分配。对于每种颜色 $c \in C$，令 E_c 为由颜色 c 着色的边集，即 $E_c = \chi^{-1}(c)$。若子集 $E' \subseteq E$ 中任意两条边都不共享顶点，则称其为 G 中的匹配 。

定义 1. (α, β)-平衡匹配。给定 $0 \leq \alpha \leq \beta \leq 1$，若匹配 $M \subseteq E$ 满足对每种颜色 $c \in C$，都有 $\alpha \leq \frac{|M \cap E_c|}{|M|} \leq \beta.$，则称其为$(\alpha, \beta)$-平衡匹配。 ———

因此，如果一个匹配中每种颜色的边至少占 α 、至多占 β，那么该匹配就是 (α, β)-平衡的。在 Proportionally Fair Matching 问题中，目标是找到一个规模最大的 (α, β)-平衡匹配。在该问题受限的 β-受限情形下， $\alpha = 0$，也就是说，我们只有上界。

对于 $\gamma \leq 1$ 和 $\Delta \geq 1$， Proportionally Fair Matching的 (γ, Δ)-近似算法会计算出一个规模至少为 $\gamma \cdot$|最优解|的匹配，其中每种颜色至少占边的 α/Δ，至多占边的 $\beta \cdot \Delta$。最优解是一个最优的 (α, β)-平衡匹配。

如果一个匹配中的所有边颜色各不相同，则称其为彩虹匹配。我们将用到 Gupta 等人 [30] 得出的以下结果。

定理 1. 对于某个整数 $k > 0$，假设在 G 中存在一个大小为 k 的彩虹匹配。存在一个 $2^k \cdot n^{O(1)}$ 时间算法，可以计算出一个大小为 k 的彩虹匹配。

3 A $(\frac{1}{4\ell}, 1 + \frac{4\ell}{|OPT|})$-近似比例公平匹配

本节中，我们为 Proportionally Fair Matching 设计一种近似算法。令 OPT 为一个最优的 (α, β)-平衡匹配，$\text{OPT}_c = \text{OPT} \cap E_c$。我们设计了两种算法：一种用于 $\alpha > 0$ 的情况，另一种用于互补的 β-受限情况。在本节中，我们稍微滥用一下符号，用 OPT（对于某种颜色 $c \in C$，则用 OPT_c）指代 $|\text{OPT}|$（分别指代 $|\text{OPT}_c|$）。其具体含义应可从上下文中明确；不过，如果可能产生歧义，我们会加以说明。

首先，我们考虑 $\alpha > 0$ 的情况。显然，我们有以下观察。

观察1。 对于任意颜色 $c \in C$，最优解 至少包含一条颜色为 c 的边，因此， G包含一个大小为 ℓ 的彩虹匹配。

我们的算法在轮次内运行。下面，我们定义一轮。每轮的输入是 G 的一个子图 $G' = (V', E')$。

轮次。 初始时 $M = \emptyset$。对于每种颜色 $1 \le c \le \ell$，迭代执行以下操作。如果 G' 中没有颜色为 c 的边，则转到下一种颜色；如果 $c = \ell$，则终止并返回 (G', M)。否则，从 G' 中任选一条颜色为 c 的边 e，并将 e 加入已计算出的匹配 M。从 G' 中移除所有与 e 共享顶点的边（包括 e）。对当前图（或更新后的图）G'，针对下一种颜色重复此过程；如果 $c = \ell$，则终止并返回 (G', M)。

因此，在每轮中，我们都尝试以贪心方式选取一个彩虹匹配。接下来，我们介绍我们的算法。算法中最具挑战性的部分，是确保最终计算出的匹配在一个较小的因子范围内满足 (α, β)-平衡；也就是说，我们需要确保每种颜色的下界和上界都在一个较小的因子范围内。请注意，仅采用上述贪心方式选取边，甚至无法保证每种颜色至少选中一条边。我们在算法开始时使用 [30]的算法来克服这一障碍。不过，我们算法的其余部分极为简单。

算法。我们假设已知最优解 OPT 的大小。稍后我们会说明如何去除这一假设。在 G 上应用定理 1 中的算法，计算一个大小为 ℓ 的彩虹匹配 M'。如果 $\text{OPT} \le 4\ell^2$，则返回 $M := M'$ 作为解并终止。否则，从 G 中移除 M' 的所有边及与它们相邻的边，得到图 G_0。将 M 初始化为 M'。使用轮次过程，逐轮贪心地选取匹配边并将它们加入 M，直到总共选取恰好 $\lceil \text{OPT}/(4\ell) \rceil$ 条边。具体而言，图 G_0 是第 1 轮的输入， G_1 是第 1 轮的输出图。 G_1 是第 2 轮的输入， G_2 是第 2 轮的输出图，依此类推。注意，如果在该轮完成之前， M 的大小已达到 $\lceil \text{OPT}/(4\ell) \rceil$，最后一轮可能无法完整执行。

请注意，上述算法对 α 和 β 不敏感，也就是说，它从未使用这些值。尽管如此，我们证明，计算得到的匹配在忽略一个较小因子的情况下是 (α, β)-平衡的。现在我们来分析该算法。

3.1 分析

令 $M_c = M \cap E_c$。此外，令 c^* 为颜色 $c \in C$，使得 $|\text{最优解}_c|$ 在 $c = c^*$处取最小值。

观察 2. $\alpha \le 1/\ell \le \beta$.

证明。 设$\hat{c}$ 是一个颜色 $c \in C$，使得 $|\text{最优解}_c|$ 在 $c = \hat{c}$处取最小值。根据定义，$\text{最优解} \ge \ell \cdot \text{最优解}_{c^*}$，或者$\text{最优解}_{c^*}/\text{最优解} \le 1/\ell$。因此， $\alpha \le \text{最优解}_{c^*}/\text{最优解} \le 1/\ell$。类似地，$\text{最优解} \le \ell \cdot \text{最优解}_{\hat{c}}$，或者最优解$_{\hat{c}}$/最优解 $\ge 1/\ell$。因此， $\beta \ge \text{最优解}_{\hat{c}}/\text{最优解} \ge 1/\ell$。 □

首先考虑最优解 $\leq 4\ell^2$ 的情况。在这种情况下，返回的匹配 M 是一个大小恰好为 ℓ的彩虹匹配。根据观察 1，这样的匹配确实存在。因此，我们立即得到一个 4ℓ-近似。由于 $|M_c|/|M| = 1/\ell$ 在这种情况下，根据观察2， $\alpha \leq |M_c|/|M| \leq \beta$。因此，我们得到了所需的结果。在证明的其余部分，我们分析最优解 $> 4\ell^2$ 的情况。我们从以下引理开始。

引理 1. 该算法成功计算出一个大小恰为 $\lceil$最优解$/(4\ell)\rceil$ 的匹配。此外，对于每种满足 最优解$_c > 4\ell$ 的颜色 c，以及每个轮次 $i \in [1, \lceil$最优解$_c/(4\ell)\rceil - 1]$， G_{i-1} 都包含一条颜色为 c 的边。

证明。注意，根据观察1，定理 1 中的算法成功计算出一个大小为 ℓ 的彩虹匹配 M'。现在考虑任意满足 $\text{OPT}_c \leq 4\ell$ 的颜色 c。对于这样的颜色， M 已经包含至少 $1 \geq \lceil \text{OPT}_c/(4\ell)\rceil$ 条边。现在考虑任意其他满足 $|\text{OPT}_c| > 4\ell$ 的颜色 c。考虑在开始时计算出的彩虹匹配 M'。由于 $|M'| = \ell$， M' 中的边至多与来自 OPT 的 2ℓ 条边相邻，因为它是一个匹配。特别地， M' 中的边至多与集合 OPT_c 中的 2ℓ 条边相邻。因此， G_0 包含集合 OPT_c 中至少 $\text{OPT}_c - 2\ell$ 条边。现在考虑第 $i \geq 1$ 轮的执行。本轮至多选取 ℓ 条边。因此，这些边至多与 OPT_c 中的 2ℓ 条边相邻。由此可知，与 G_{i-1} 相比，集合 OPT_c 中包含在 G_i 里的边至多少 2ℓ 条。由于 G_0 中包含集合 OPT_c 中颜色为 c 且满足 $\text{OPT}_c > 4\ell$ 的至少 $\text{OPT}_c - 2\ell$ 条边，因此在前 $\lceil(\text{OPT}_c - 2\ell)/(2\ell)\rceil = \lceil \text{OPT}_c/(2\ell)\rceil - 1$ 轮中，算法能够选取一条颜色为 c 的边。因此，对于任意满足 $\text{OPT}_c > 4\ell$ 的此类颜色 c，算法至少可以安全地选取总共 $\lceil \text{OPT}_c/(2\ell)\rceil \geq \lceil \text{OPT}_c/(4\ell)\rceil$ 条边。现在，由于 $\text{OPT} = \sum_c \text{OPT}_c$， $\sum_{c\in C}\lceil \text{OPT}_c/(4\ell)\rceil \geq \lceil \text{OPT}/(4\ell)\rceil$。由此可知，该算法至少可以选取 $\lceil \text{OPT}/(4\ell)\rceil$ 条边。由于当 M 的大小达到 $\lceil \text{OPT}/(4\ell)\rceil$ 时，我们便停止算法，因此引理得证。 □

注意，所声称的近似因子可直接由上述引理得出。接下来，我们将证明， M 是（α， β）-平衡的，误差因子较小，并且随着最优解的规模增大而渐近趋于 1。

引理 2. 对于每种颜色 $c \in C$, $|M_c| \geq |$最优解$_{c^*}|/(4\ell)$。

证明。如果 $\text{OPT}_{c^*} \leq 4\ell$, $|M_c| \geq 1 \geq \text{OPT}_{c^*}/(4\ell)$。因此，假设 $\text{OPT}_{c^*} > 4\ell$。现在，假设对于某个 c, 有 $|M_c| <\text{OPT}_{c^*}/(4\ell)$。根据引理 1，在前 $\lceil \text{OPT}_c/(4\ell)\rceil - 1 \geq \lceil \text{OPT}_{c^*}/(4\ell)\rceil - 1$轮中的每一轮， G_{i-1} 都包含一条颜色为 c 的边。因此，算法在某一轮 $i \leq (\text{OPT}_{c^*}/(4\ell)) - 1$ 中被迫终止。于是，从每种颜色 $c' \neq c$ 中选出的边数至多为 $\text{OPT}_{c^*}/(4\ell)$。因此，

$$\begin{aligned} |M| &= \sum_{c'\neq c} |M_{c'}| + |M_c| \\ &< (\ell - 1)\cdot(\text{OPT}_{c^*}/(4\ell)) + (\text{OPT}_{c^*}/(4\ell)) \\ &\leq \lceil \text{OPT}/(4\ell)\rceil. \end{aligned}$$

这与引理1矛盾，因为引理指出，我们恰好选择了 $\lceil$最优解$/(4\ell)\rceil$ 条边。 □

推论 1. 对于每种颜色 $c \in C$, $(|M_c|/|M|) \geq$ $\alpha/(1+4\ell/$最优解$)$。

证明。 根据引理 2， $|M_c| \geq$最优解$_{c^*}/(4\ell)$。

$$\begin{aligned} \frac{|M_c|}{|M|} &\geq \frac{(\text{OPT}_{c^*}/(4\ell))}{\lceil \text{OPT}/(4\ell)\rceil} \geq \frac{(\text{OPT}_{c^*}/(4\ell))}{(\text{OPT}/(4\ell)) + 1} \\ &= \frac{(\text{OPT}_{c^*})/(\text{OPT})}{(1 + 4\ell/\text{OPT})} \geq \frac{\alpha}{(1 + 4\ell/\text{OPT})}. \end{aligned}$$

最后一个不等式成立，是因为最优解满足所有颜色的下界。 □

现在我们来证明上界。令 α^* =最优解$_{c^*}$/最优解。

引理 3. 对于每种颜色 $c \in C$, $|M_c| \leq \frac{\beta}{\alpha^*} \cdot$ (最优解$_{c^*}/(4\ell)) + 1$。

证明。 假设对于某个 $c \in C$, $|M_c| > \frac{\beta}{\alpha^*} \cdot$ (最优解$_{c^*}/(4\ell))$ $+ 1$。那么轮次数严格大于 $\frac{\beta}{\alpha^*} \cdot$(最优解$_{c^*}/(4\ell)$)。现在，对于任意 c'，最优解$_{c'} \geq \alpha^* \cdot$最优解且最优解$_{c'} \leq \beta \cdot$最优解。因此，根据 α^*的定义，$\frac{\beta}{\alpha^*} \cdot$最优解$_{c^*} \geq$ 最优解$_{c'}$。由此可知，对于每个 c'，轮次数严格大于最优解$_{c'}/(4\ell)$。因此，对于每个 $c' \in C$，已选择的边数超过 (最优解$_{c'}/(4\ell)) + 1$。所以，已选择边的总数严格大于

$$\sum_{c' \in C} ((\mathrm{OPT}_{c'}/(4\ell)) + 1) \geq \lceil \mathrm{OPT}/(4\ell) \rceil.$$

这与引理1矛盾。该引理指出，我们恰好选择 $\lceil$最优解$/(4\ell)\rceil$ 条边。 □

推论 2。 对于每种颜色 $c \in C$, $(|M_c|/|M|) \leq \beta \cdot (1 + \frac{4\ell}{\text{最优解}})$。

证明。 根据引理 3,

$$\begin{aligned}
\frac{|M_c|}{|M|} &\leq \frac{(\beta/\alpha^*) \cdot (\mathrm{OPT}_{c^*}/(4\ell)) + 1}{\lceil \mathrm{OPT}/(4\ell) \rceil} \\
&\leq \frac{(\beta/\alpha^*) \cdot (\mathrm{OPT}_{c^*}/(4\ell)) + (\beta/\alpha^*)}{\mathrm{OPT}/(4\ell)} \\
&= \frac{\beta}{\alpha^*} \cdot \frac{\mathrm{OPT}_{c^*}}{\mathrm{OPT}} \cdot \left(1 + \frac{4\ell}{\mathrm{OPT}}\right) \\
&= \frac{\beta}{\alpha^*} \cdot \alpha^* \left(1 + \frac{4\ell}{\mathrm{OPT}}\right) = \beta \cdot \left(1 + \frac{4\ell}{\mathrm{OPT}}\right).
\end{aligned}$$

第二个不等式成立，因为 $\alpha^* \leq \beta$ 或 $\beta/\alpha^* \geq 1$。 □

现在我们去掉已知最优解大小这一假设。注意，$\ell \leq$ 最优解 $\leq n$。我们枚举 ℓ 和 n之间的所有取值，并对每个这样的取值 T 运行我们的算法。对于算法返回的每个匹配 M，我们检查 M 是否为 ($\alpha_{(1+4\ell/T)}\ \beta \cdot (1 + \frac{4\ell}{T})$)- 平衡。如果是，则保留该解；否则，舍弃该解。最后，我们从未被舍弃的解中选取规模最大的一个。根据上述分析，对于$T =$ 最优解，返回的匹配满足所需的下界和上界，且大小恰为$\lceil$最优解$/(4\ell)\rceil$。最后，我们算法的运行时间主要由 $2^\ell n^{O(1)}$ 计算彩虹匹配算法所需的时间决定，如定理1所述。

定理 2。 对于 Propor-tionally Fair Matching 且满足 $\alpha > 0$ 的情形，存在一个 $2^\ell \cdot n^{O(1)}$ 时间的$(1/4\ell, 1 + 4\ell/$最优解$)$-近似算法。

4 β受限情形下的多项式时间近似

在 β-受限情形下，我们再次使用**轮次**过程。不过，算法略有不同。最重要的是，我们一开始不会应用定理 1中的算法。因此，我们的算法在多项式时间内运行。

算法。 假设我们知道最优解 OPT 的大小。如果 OPT $\leq 2\ell$，我们任选一条边并将其作为解返回。否则，我们只需使用**轮次** 过程，分轮次贪心地选取匹配边，但需注意以下两点。对于某种颜色，如果已选边数至少达到 $\beta \cdot$OPT$/(2\ell)$，则不再选择该颜色的边。如果已选边数至少达到 (OPT$/2\ell)$ $- 1$，则终止。

现在我们分析该算法。首先注意，如果 OPT $\leq 2\ell$，返回的匹配只有一条边。显然满足上界，而且我们也得到了一个 2ℓ-近似。因此，以下我们假设 OPT $> 2\ell$。在证明正确性并分析近似因子之前，我们先证明每种颜色的上界。再次令 M 为计算得到的匹配，并令$M_c = M \cap E_c$。稍后我们将证明以下引理。

引理 4. 该算法始终返回一个大小至少为 (最优解$/2\ell) - 1$ 的匹配。

假设这一点成立，我们有如下观察。

观察 3. 对于每种颜色 $c \in C$, $|M_c|/|M| \leq \beta \cdot (1 + {}^{2\ell}/_{\text{最优解}|})$。

证明。 根据引理 4以及算法中对每种颜色设置的阈值，

$$\frac{|M_c|}{|M|} \leq \frac{\beta \cdot \mathrm{OPT}/(2\ell)}{(\mathrm{OPT}/2\ell) - 1} \leq \beta \cdot (1 + \frac{2\ell}{\mathrm{OPT}})$$

最后一个不等式成立，因为最优解 $> 2\ell$。 □

接下来，我们证明引理 4。

证明。 令 C_1 为颜色的子集，使得对于每个 $c \in C_1$，算法至少选取$\beta \cdot$OPT$/(2\ell)$条边。注意，算法可能以两种方式终止：(i) 已经选取至少 (OPT$/2\ell) - 1$ 条边；(ii) 所有边均已耗尽。注意，若发生(i)，则结论成立。我们证明(ii)不可能在(i)未发生的情况下出现。假设(ii)发生，但(i)未发生。令OPT′ 为OPT中包含颜色属于 $C' = C \setminus C_1$的边的子集。回顾一下， G_{i-1}是第 i轮的输入图， G_i 是 $i \geq 1$的输出图。在第 i轮选取的边数至多为 ℓ。因此，这些边至多与 G_{i-1}中的 2ℓ 条边相邻。特别地，与 G_{i-1}相比， G_i 中包含的OPT′ 边至多少 2ℓ 条。因此，算法至少可以选取 $\lfloor$条颜色属于 C'的OPT′$/2\ell\rfloor$ 边。由于对于 C'中的每种颜色，选取的边数少于$\beta \cdot$OPT$/(2\ell)$条，因此算法确实选取了这些颜色中至少 $\lfloor$OPT′$/2\ell\rfloor$ 条边。算法选取的边总数为，

$$\begin{aligned}\sum_{c\in C_1} |M_c| + \sum_{c\in C'} |M_c| &\geq \sum_{c\in C_1} \beta \cdot \mathrm{OPT}/(2\ell) + \lfloor \mathrm{OPT}'/2\ell \rfloor \\ &\geq \sum_{c\in C_1} \mathrm{OPT}_c/2\ell + \lfloor \mathrm{OPT}'/2\ell \rfloor \\ &\geq (\mathrm{OPT}/2\ell) - 1\end{aligned}$$

但是，这与我们的假设矛盾，因此引理得证。 □

定理 3. 在 Proportionally Fair Matching *in the*β-受限的情况下，存在一种多项式时间算法，可返回一个大小至少为 (最优解$/2\ell) - 1$ 的匹配，其中每种颜色出现的边数至多占 $(1 + 2\ell/$最优解$)$ 的 $\beta \cdot$ 分数。

5 一种求比例公平匹配的精确算法

定理 4. 存在一种 $2^{O(k)}n^{O(1)}$-时间算法，对于一个 Proportionally Fair Matching 实例，要么找到大小为 k 的解，要么判定这样的解不存在。

证明。我们采用著名的颜色编码技术，提出两种不同的算法：一种适用于 $\alpha = 0$ （β-受限情形），另一种适用于 $\alpha > 0$。

β-受限情形。我们的目标是将问题转化为寻找大小为 k 的彩虹匹配，然后通过定理 1 求解。图 G 保持不变，但着色方式有所不同。具体来说，对于每种原始颜色 $c \in C$，我们从一个包含 k' 种新颜色的集合中，均匀且独立地随机选取颜色，为 E_c 中的边着色，其中 $k' = \lfloor \beta k \rfloor$。因此，新实例 I' 使用 $\ell \cdot k'$ 种颜色着色。我们使用定理 1 的算法，在着色后的图 I' 中寻找大小为 k 的彩虹匹配。显然，如果找到大小为 k 的彩虹匹配 M，那么同一匹配 M 在原始着色中就是大小为 k 的 β-受限匹配。这是因为根据构造，对于任意原始颜色 $c \in C$，边集 E_c 中有 k' 种新颜色，因此 $|M \cap E_c|$ 中的边不超过 k' 条。

反过来，我们证明：如果相对于原始着色，存在一个大小为 k 的 β-受限匹配 M，那么以较高概率，M 在新着色中是一个大小为 k 的彩虹匹配。假设原始颜色 c_1、.....、 c_t 对某个 $1 \le t \le \ell$ 与 M 有非空交集，并对每个 $j \in [t]$ 记 $k_j = |M \cap E_{c_j}|$。注意到 $\sum_{j=1}^{t} k_j = k$，且对每个 $j \in [t]$， $1 \le k_j \le k'$。

断言 1. 存在某个 $\delta > 0$，使得对每个 $j \in [t]$:

$$\mathsf{Pr}\left[M \cap \left(\bigcup_{i=1}^{j} E_{c_i} \right) \text{ is a rainbow matching in } I' \right] \ge \exp\left(-\delta \sum_{i=1}^{j} k_i \right),$$

证明. 我们对 j 进行归纳法来证明该断言。对于基础情形，显然 (1) 对 $j = 0$ 成立。现在，固定 $j \in [t]$，并假设该陈述对每个 $j' < j$ 都成立，我们证明 (1) 对 j 也成立。考虑 k_j 的边，它们在 $M \cap E_{c_j}$ 中被独立且均匀地染成 $k' \ge k_j$ 种颜色。通过计算 $M \cap E_{c_j}$ 的可能着色数，可得

$$\begin{aligned} &\mathsf{Pr}\left[M \cap E_{c_j} \text{ is a rainbow matching } \right] \\ &\quad \ge \frac{(k')!/(k'-k_j)!}{(k')^{k_j}} \ge \frac{k_j!}{k_j^{k_j}} \ge 2^{-\delta k_j}, \end{aligned}$$

其中，最后一个界由斯特林公式得出。现在，由于用于 E_{c_j} 的颜色不会出现在其他任何地方，利用归纳假设可得

$$\begin{aligned} &\mathsf{Pr}\left[M \cap \left(\bigcup_{i=1}^{j} E_{c_i} \right) \text{ is a rainbow matching } \right] \\ =&\mathsf{Pr}\left[M \cap \left(\bigcup_{i=1}^{j-1} E_{c_i} \right) \text{ is a rainbow matching } \right] \\ &\cdot \mathsf{Pr}\left[M \cap E_{c_j} \text{ is a rainbow matching } \right] \\ \ge& 2^{-\delta \sum_{i=1}^{j-1} k_i} \cdot 2^{-\delta k_j} = 2^{-\delta \sum_{i=1}^{j} k_i}. \end{aligned}$$

□

应用（1）并结合 $j = t$，我们得到 M 是一个彩虹匹配，其概率至少为$2^{-\delta k}$。将上述归约独立重复 $2^{O(k)}$ 次，即可使算法以常数概率成功。

这种情况 $\alpha > 0$。我们注意到，在这种情况下，如果一个匹配是公平的，那么它必然包含每个组中至少一条边。因此，如果组的数量 ℓ 大于 k，我们便可立即得出结论：不存在大小为 k 的公平匹配。否则，我们猜测所需的 k 条边如何划分到 ℓ 个组 $C = \{c_1, \ldots, c_\ell\}$中。也就是说，我们猜测满足 $\sum_{j=1}^{\ell} k_j = k$的数值 k_j（其中 $j \in [\ell]$），并对每个 $j \in [\ell]$猜测 $\alpha k \le k_j \le \beta k$。从现在起，该算法与 β限制情形下的算法非常相似。对于每个组 c_j，我们对 E_{c_j}中的边进行着色

从 k_j 种颜色中均匀且独立地随机选取颜色，其中每个 E_{c_j}所用的颜色互不重叠。现在，我们使用定理 1 的算法寻找大小为 k的彩虹匹配。如果存在大小为 k的彩虹匹配 M，那么同一个匹配就是原始实例中大小为 k 的公平匹配，因为在每个 E_{c_j} 中恰好选取了 k_j 条边，其数量至少为 αk 且至多为βk。反过来，如果原始实例中存在大小为 k 的公平匹配 M，根据 (1)，匹配 M 在新实例中成为彩虹匹配的概率至少为 $2^{-\delta k}$。同样，通过独立重复着色子过程 $2^{O(k)}$ 次，我们可以获得常数概率的成功率。由于将 k 条边划分到至多 $\ell \leq k$ 个组中共有 $2^{O(k)}$ 种方案，整个算法的运行时间为 $2^{O(k)}n^{O(1)}$。

最后，我们指出，在这两种情况下，着色部分都可以通过使用完美哈希族，以标准方式去随机化 [42],从而得到一个完全确定性的算法。 □

6 Har比例公平匹配的近似难度 g

本节中，我们在指数时间假设 (ETH) 下证明了 Proportionally Fair Matching 的不可近似性结果 [33]。ETH 指出，求解任何包含 n 个变量的一般 3-SAT 实例都需要 $2^{\Omega(n)}$ 时间。为此，我们需要以下受限版本的 3-SAT。

3SAT-3
输入：变量为 x_1、.....、x_n 的子句集合 $T = \{C_1, \ldots, C_m\}$，每个子句都是 3 个或 2 个文字的析取，其中文字是一个变量 x_i 或其否定 $\bar{x}_i$。此外，每个变量出现 3 次。

问题：是否存在一个真值赋值，能够同时满足所有子句?
已知 3SAT-3 是 NP 困难的 [48]。我们需要以下在 [20] 中证明的关于 3SAT-3 的更强下界。

命题 1 ([20])。 *在指数时间假设下，3SAT-3 无法在 $2^{o(n)}$ 时间内求解。*

我们将 3SAT-3 归约到 Proportionally Fair Matching，这排除了在 $2^{o(\ell)}n^{O(1)}$ 时间内对后一个问题进行任何近似的可能性。我们的归约如下。对于每个子句 C_i，我们分配一种颜色 i。此外，我们还使用 $n-1$ 种额外颜色 $m+1$、.....、$m+n-1$。因此，颜色集合$C = \{1, \ldots, m+n-1\}$。对于每个变量 x_i，我们构造一个构件，它是一条 3-路径（一条有 3 条边的路径）。注意，x_i 可以在其正形式或否定形式中出现两次，因为它总共出现 3 次。令 C_{i_1}、C_{i_2} 和 C_{i_3} 为 x_i 出现的子句。另设它以一种形式出现在 C_{i_1} 和 C_{i_3} 中，以另一种形式出现在 C_{i_2} 中。我们为 x_i 构造一条 3-路径 P_i，其中对于 $1 \leq j \leq 3$，第 j 条边的颜色为 i_j。此外，我们为$1 \leq i \leq n-1$ 构造 $n-1$ 3 条路径 $Q_{i,\ i+1}$。$Q_{i,\ i+1}$ 的所有边都是颜色 $m+i$。最后，我们按以下方式将所有路径拼接起来，得到一条单独的路径。对于每个 $1 \leq i \leq n-1$，我们将 $Q_{i,\ i+1}$ 拼接在 P_i 和P_{i+1} 之间：将 P_i 的最后一个顶点与Q_i 的第一个顶点重合，并将 $Q_{i,\ i+1}$ 的最后一个顶点与 P_{i+1} 的第一个顶点重合。由此得到一条路径 P，恰有 3$(n+n-1)$ 条 $= 6n-3$ 边。最后，我们令 $\alpha = \beta = 1/(m+n-1)$。

引理 5。 *当且仅当存在大小至少为 $m+n-1$ 的 (α, β)-平衡匹配时，3SAT-3 中的子句存在满足赋值。*

证明。 假设所有子句都存在一个满足赋值。对于每个子句 C_j，考虑一个变量，例如 x_i，它满足 C_j。那么在 P_i 上存在一条颜色为 j 的边。将这条边加入集合 M。因此，这一步之后，M 对于 $1 \leq j \leq m$ 恰好包含一条颜色为 j 的边。另请注意，对于每条路径 P_i，如果选择了中间边，那么就不能再选择来自 P_i 的其他边。这是因为变量 x_i 可以满足它以正形式出现的子句，或者满足它以否定形式出现的子句，但不能同时满足这两类子句。因此，M 是一个匹配。最后，对于每条路径 $Q_{i,\ i+1}$，我们将其中间边加入 M。请注意，M 仍然是一个匹配。

此外， M 对于 $1 \le j \le m+n-1$ 恰好包含一条颜色为 j 的边。由于 $\alpha = \beta = 1/(m+n-1)$，$M$ 是一个（α， β）-平衡匹配。

现在假设存在一个大小至少为 $m+n-1$ 的（α, β）-平衡匹配 M。首先，我们证明 $|M| = m+n-1$。注意，如果 $|M| > m+n-1$，那么唯一的可能性是 $|M| = 2(m+n-1)$，因为 $\alpha = \beta$，且对于 $m+1 \le j \le m+n-1$，任何匹配中至多能选取 2 条颜色为 j 的边。假设 $|M| = 2(m+n-1)$。那么，对于每个 Q_i，$i+1$， M 都包含第一条边和第三条边。这意味着，对于每个 P_t， $1 \le t \le n$，我们至多能选取一条边。因此， M 中的边数至多为 $2(n-1)+n$。由此可得 $2m+2n-2 \le 2n-2+n$ 或 $n \ge 2m$。现在，在 3SAT-3 中，文字总数为 $3n$，至多为 $3m$，因为每个变量出现 3 次，且每个子句至多包含 3 个文字。这意味着 $n \le m$，从而得到矛盾。因此，$|M| = m+n-1$。现在，考虑任意一个 P_i。在第一种情况下， P_i 中的第一条边和第三条边分别对应文字 x_i，因此，中间边对应文字 $\bar{x}_i$。如果中间边属于 M，则给 x_i 赋值 0；否则，给 x_i 赋值 1。在另一种情况下，如果中间边属于 M，则给 x_i 赋值 1；否则，给 x_i 赋值 0。我们声称，所构造的赋值满足所有子句。考虑任意一个子句 C_j。设 $e \in P_i$ 是 $1 \le j \le m$ 对应的、 M 中颜色为 j 的边。注意， e 可以是 P_i 中间的边，也可以不是。无论哪种情况，如果 e 对应于 $\bar{x}_i$，我们就给 x_i 赋值 0；如果 e 对应于 x_i，我们就给 x_i 赋值 1。因此，两种情况下，C_j 都得到满足。引理得证。

□

注意，对于一个 3SAT-3 实例，文字总数为 $3n$。由于每个子句至少包含 2 个文字，$m \le 3n/2$。现在，对于上述证明中构造的实例，颜色数 $\ell = m+n-1 \le 3n/2+n-1 = 5n/2-1$。因此，上述引理和命题 1 表明，无法在时间 $2^{o(\ell)}n^{O(1)}$ 内判定是否存在一个大小给定的 (α, β)-平衡匹配。利用这一点，我们还证明了甚至不存在 $2^{o(\ell)}n^{O(1)}$ 时间的近似算法。假设存在一个 $2^{o(\ell)}n^{O(1)}$ 时间的 γ-近似算法，其中 $\gamma < 1$。对于我们构造的路径实例，应用该算法来获得一个匹配。注意，γ-近似解 M 必须包含每种颜色至少一条边，因为 $\alpha = \beta$。根据上述引理中的证明， $|M|$ 恰好是 $m+n-1$。因此，利用该算法，我们可以在 $2^{o(\ell)}n^{O(1)}$时间内判定是否存在一个大小为 $m+n-1$的 (α, β)-平衡匹配。但这会导致矛盾，从而得到以下定理。

定理 5. 对于任意 $\gamma > 1$，在 *ETH* 下，即使在路径上，也不存在用于 Proportionally Fair Matching 的 $2^{o(\ell)}n^{O(1)}$ 时间 γ-近似算法。

7 结论

本文研究了图中匹配的比例公平这一概念，Chierichetti 等人 [15]也对此进行了研究。我们为比例公平匹配问题提出了近似算法和精确算法，并通过证明困难性结果补充了这些成果。若能获得一个 $o(\ell)$-近似算法，或一个真正的 $O(\ell)$-近似算法来解决 Pro-portionally Fair Matching，从而改进我们的结果，将会很有意义。正如我们的困难性结果所表明的，此类算法的运行时间存在 $2^{\Omega(\ell)}n^{O(1)}$ 下界。

致谢。 这项工作的绝大部分是在四位作者均供职于挪威卑尔根大学期间完成的。促成本研究成果的研究获得了挪威研究委员会通过 MULTIVAL 项目，以及欧洲研究委员会（ERC）通过 LOPPRE 项目（项目编号 819416）提供的资助。

参考文献

[1] A. Agarwal, M. Dudik, and Z. S. Wu, 公平回归：定量定义与基于归约的算法，载于国际机器学习大会，机器学习研究论文集，2019年，第 120–129 页。 1[2] S. Ahmadi, F. Ahmed, J. P. Dickerson, M. Fuge, and S. Khuller, 多属性多样化匹配算法，载于第二十九届国际人工智能联合会议论文集，IJCAI 2020，C. Bessiere 编，ijcai.org，2020年，第 3–9 页。 1[3] N. Alon, R. Yuster, and U. Zwick, 彩色编码，《美国计算机协会学报》（JACM），42（1995年），第 844–856 页。 3[4] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, 机器偏见：全国各地都在使用一种预测未来罪犯的软件，而该软件对黑人存在偏见，ProPublica，（2016年5月23日）。 1[5] S. Bandyapadhyay, F. V. Fomin, and K. Simonov, 关于度量空间和欧几里得空间中公平聚类的核心集及其应用，CoRR，abs/2007.10137（2020年）。 1[6] X. Bei, S. Liu, C. K. Poon, and H. Wang, 满足比例公平性约束的候选人遴选，载于第十九届自主代理与多智能体系统国际会议论文集，AAMAS '20，新西兰奥克兰，2020年5月9日至13日，A. E. F. Seghrouchni、G. Sukthankar、B. An 和 N. Yorke-Smith 编，国际自主代理与多智能体系统基金会，2020年，第 150–158 页。 3[7] S. Bera, D. Chakrabarty, N. Flores, and M. Negahbani, 聚类公平算法，载于《神经信息处理系统进展》，2019年，第 4954–4965 页。 1[8] A. Berger, V. Bonifaci, F. Grandoni, and G. Schäfer, 通过汽油谜题实现预算约束匹配和预算约束拟阵交，Math. Program., 128（2011年），第 355–372 页。 3[9] R. Berk, H. Heidari, S. Jabbari, M. Joseph, M. J. Kearns, J. Morgenstern,S. Neel, and A. Roth, 公平回归的凸框架，CoRR，abs/1706.02409（2017年）。 1[10] J. Buolamwini and T. Gebru, 性别阴影：商业性别分类中的交叉性准确率差异，载于公平、问责与透明度会议，FAT 2018，2018年2月23日至24日，美国纽约州纽约市，S. A. Friedler 和 C. Wilson 编，第 81 卷《机器学习研究论文集》，机器学习研究论文集，2018年，第 77–91 页。 1[11] T. Calders and S. Verwer, 三种实现无歧视分类的朴素贝叶斯方法，《数据挖掘与知识发现》，21（2010年），第 277–292 页。 1[12] L. E. Celis, L. Huang, and N. K. Vishnoi, 满足公平性约束的多席位投票，载于第二十七届国际人工智能联合会议论文集，IJCAI 2018，2018年7月13日至19日，瑞典斯德哥尔摩，J. Lang 编，ijcai.org，2018年，第 144–151 页。 1[13] L. E. Celis, D. Straszak, and N. K. Vishnoi, 满足公平性约束的排名，载于第45届自动机、语言与编程国际研讨会，ICALP 2018，2018年7月9日至13日，捷克共和国布拉格，I. Chatzigiannakis、C. Kaklamanis、D. Marx 和 D. Sannella 编，LIPIcs 第 107 卷，Schloss Dagstuhl - Leibniz-Zentrum f¨ur Informatik，2018年，第 28:1–28:15 页。 1

[14] L. Charlin and R. Zemel, 《多伦多论文匹配系统：一个自动化论文审稿人分配系统》，（2013）。1[15] F. Chierichetti, R. Kumar, S. Lattanzi, and S. Vassilvitskii, 基于公平小簇实现公平聚类，载于《神经信息处理系统进展》，2017年，第5029–5037页。1, 10
[16] F. Chierichetti, R. Kumar, S. Lattanzi, and S. Vassilvitskii, 拟阵、匹配与公平性，载于《第22届国际人工智能与统计学会议（AISTATS 2019）》，2019年4月16日至18日，日本冲绳那霸，K. Chaudhuri和M. Sugiyama编，第89卷《机器学习研究论文集》，PMLR，2019年，第2212–2220页。1, 2[17] A. Chouldechova，具有差别影响的公平预测：再犯预测工具中偏见的研究，《大数据》，5（2017年），第153–163页。3[18] S. Corbett-Davies, E. Pierson, A. Feller, S. Goel, and A. Huq，算法决策与公平成本，载于《第23届ACM SIGKDD国际知识发现与数据挖掘会议论文集》，加拿大新斯科舍省哈利法克斯，2017年8月13日至17日，ACM，2017年，第797–806页。1, 3[19] C. S. Crowson, E. J. Atkinson, and T. M. Therneau，预后风险评分的校准评估，《医学研究中的统计方法》，25（2016年），第1692–1706页。1[20] M. Cygan, D. Marx, M. Pilipczuk, and M. Pilipczuk，在有界树宽图中命中禁用子图，《信息与计算》，256（2017年），第62–82页。9[21] A. Datta, M. C. Tschantz, and A. Datta，针对广告隐私设置的自动化实验：不透明性、选择与歧视的故事，《隐私增强技术论文集》，2015（2015年），第92–112页。1[22] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel，基于认知的公平性，载于《第3届理论计算机科学创新会议论文集》，2012年，第214–226页。1, 3
[23] C. Dwork and C. Ilvento，组合下的群体公平，载于《2018年公平性、问责制与透明度会议论文集》（FAT* 2018），2018年。3[24] S. Ebadian, A. Kahng, D. Peters, and N. Shah，投票中的优化失真与比例公平，载于《第23届ACM经济学与计算会议论文集》，2022年，第563–600页。3[25] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubra-manian，差别影响的认证与消除，载于《第21届ACM SIGKDD国际知识发现与数据挖掘会议论文集》，2015年，第259–268页。1, 2, 3[26] R. Freeman, E. Micha, and N. Shah，双边匹配与公平分配的结合，（2020年）。1[27] H. N. Garb，临床判断中的种族偏见、社会阶层偏见与性别偏见，《临床心理学：科学与实践》，4（1997年），第99–120页。1[28] D. García-Soriano and F. Bonchi，公平设计匹配，《数据挖掘与知识发现》，34（2020年），第1291–1335页。1, 3[29] N. Goel, M. Yaghini, and B. Faltings，通过凸公平性准则实现非歧视性机器学习，载于《2018年AAAI/ACM人工智能、伦理与社会会议论文集》（AIES 2018），美国路易斯安那州新奥尔良，2018年2月2日至3日，J. Furman、G. E. Marchant、H. Price和F. Rossi编，ACM，2018年，第116页。3

[30] S. Gupta, S. Roy, S. Saurabh, and M. Zehavi，彩虹匹配的参数化算法与核，《Algorithmica》，81（2019年），第1684–1698页。3, 4[31] C. Huang, T. Kavitha, K. Mehlhorn, and D. Michail，公平匹配及相关问题，《Algorithmica》，74（2016年），第1184–1203页。1, 3[32] L. Huang, S. Jiang, and N. Vishnoi，满足公平性约束的聚类核心集，载于《神经信息处理系统进展》，2019年，第7589–7600页。1
[33] R. Impagliazzo and R. Paturi，关于k-SAT的复杂性，《计算机与系统科学杂志》，62（2001年），第367–375页。2, 9[34] M. Joseph, M. J. Kearns, J. Morgenstern, and A. Roth，学习中的公平性：经典老虎机与上下文老虎机，载于《神经信息处理系统第29届会议：2016年神经信息处理系统年度会议》，2016年12月5日至10日，西班牙巴塞罗那，D. D. Lee、M. Sugiyama、U. von Luxburg、I. Guyon和R. Garnett编，2016年，第325–333页。1
[35] Y. Kamada and F. Kojima，约束下的公平匹配：理论与应用，（2020年）。1, 3
[36] T. Kamishima, S. Akaho, and J. Sakuma，通过正则化方法实现公平性感知学习，载于《数据挖掘研讨会（ICDMW），IEEE第11届国际会议》，加拿大不列颠哥伦比亚省温哥华，2011年12月11日，M. Spiliopoulou、H. Wang、D. J. Cook、J. Pei、W. Wang、O. R. Za¨ıane和X. Wu编，IEEE计算机学会，2011年，第643–650页。3[37] D. Kesavan, E. Periyathambi, and A. Chokkalingam，基于多目标梯度非洲水牛优化算法的比例公平调度策略：实现有效资源分配与干扰最小化，《国际通信系统杂志》，35（2022年），第e5003页。3[38] B. Klaus and F. Klijn，程序公平且稳定的匹配，《经济理论》，27（2006年），第431–447页。1, 3[39] J. Kleinberg, S. Mullainathan, and M. Raghavan，风险评分公平确定中的内在权衡，载于《第八届理论计算机科学创新会议（ITCS 2017）》，Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik，2017年。1, 3[40] R. Kurata, N. Hamada, A. Iwasaki, and M. Yokoo，具有软边界与重叠类型的受控学校选择，《人工智能研究杂志》，58（2017年），第153–184页。1[41] Y. Lu，基于比例公平优先级的自动化集装箱码头调度优化，《工程数学问题》，2022（2022年）。3[42] M. Naor, L. J. Schulman, and A. Srinivasan，分割器与近乎最优的去随机化，载于《第36届计算机科学基础年度研讨会论文集》（FOCS 1995），IEEE，1995年，第182–191页。9[43] M. H. Nguyen, M. Baiou, V. H. Nguyen, and T. Q. T. Vo，平衡TSP的纳什公平解，载于《第10届国际网络优化会议》（INOC），2022年。3[44] P. Ristoski, P. Petrovski, P. Mika, and H. Paulheim，用于产品匹配与分类的机器学习方法，《语义网》，9（2018年），第707–728页。1

[45] W. St-Arnaud, M. Carvalho, and G. Farnadi, 肾脏交换项目中公平方案的适应、比较与实际实施，arXiv预印本 arXiv:2207.00241，（2022年）。3[46] G. Stamoulis，通过凸分解求解有界彩色匹配的近似算法，载于《计算机科学的数学基础 2014》——第39届国际研讨会 MFCS 2014，匈牙利布达佩斯，2014年8月25日至29日。论文集，第二部分，E. Csuhaj-´Varj´u、M. Dietzfelbinger 和 Z. Esik 编，第8635卷《计算机科学讲义》，Springer，2014年，第625–636页。3[47] B. L. Thanh, S. Ruggieri, and F. Turini，将k近邻用于情境测试，以发现和预防歧视，载于第17届 ACM SIGKDD 国际知识发现与数据挖掘会议论文集，美国加利福尼亚州圣迭戈，2011年8月21日至24日，C. Apt´e、J. Ghosh 和 P. Smyth 编，ACM，2011年，第502–510页。3[48] M. Yannakakis，节点删除和边删除 NP完全问题，载于第10届 ACM年度计算理论研讨会论文集，1978年5月1日至3日，美国加利福尼亚州圣迭戈，R. J. Lipton、W. A. Burkhard、W. J. Savitch、E. P. Friedman 和 A. V. Aho 编，ACM，1978年，第253–264页。9[49] G. Zhang, S. Malekmohammadi, X. Chen, and Y. Yu，平等不等于公平：联邦学习中的比例公平性，arXiv预印本 arXiv:2202.01666，（2022年）。3