机器学习研究杂志 1（2021年）1-48 投稿于 4/00；发表于 10/00

arXiv:2302.01971v2 [cs.GT] 2 May 2023

竞争中的内容创作者会让 Top-K 推荐变得多糟?

Fan Yao[1] FY4BC@VIRGINIA.EDU
Chuanhao Li[1] CL5EV@VIRGINIA.EDU
Denis Nekipelov[2] DN4W@VIRGINIA.EDU
Hongning Wang[1] HW5X@VIRGINIA.EDU
Haifeng Xu[3] HAIFENGXU@UCHICAGO.EDU

[1]美国弗吉尼亚大学计算机科学系[2]美国弗吉尼亚大学经济学系[3]
美国芝加哥大学计算机科学系

摘要

内容创作者在推荐平台上竞争曝光度，这种策略行为会导致内容分布动态变化。然而，创作者竞争如何影响用户福利，以及相关性驱动的推荐如何影响长期动态，目前仍鲜为人知。

本研究为这些问题提供了理论见解。我们在以下假设下对创作者竞争进行建模：1）平台采用无害的 top-K 推荐策略；2）用户决策遵循随机效用模型；3）内容创作者为争取用户参与度而竞争，且由于事后并不知道自己的效用函数，因此采用任意的无悔学习算法来更新策略。我们从无政府价格的角度研究用户福利保障，并表明，创作者竞争造成的用户福利损失比例始终有一个较小的上界，该上界取决于 K 以及用户决策中的随机性；我们还证明了这一界是紧的。我们的结果揭示了短视推荐方法的内在优点：只要用户决策具有随机性，且平台为用户提供足够多的备选项，相关性驱动的匹配在长期内就能取得合理效果。

1. 引言

“（生产者）受到看不见的手的引导，几乎以相同的方式分配生活必需品...
...因此，他们无意于此，也毫不知情，却推动了社会利益。”

—— 亚当·斯密，《道德情操论》，1759年。

Instagram 和 YouTube 等在线推荐平台已在我们的日常生活中普及（Bobadilla 等人，2013年）。这些平台的核心是推荐系统（RS），旨在根据预测的相关性，为每位用户匹配最相关的内容。这种做法通常称为 Top-RS，被认为能够提升用户

满意度，数十年来一直是学术界和工业界奉行的经验法则（Konstan 等人，1997年；Koren 等人，2009年；He 等人，2017年）。

直到最近，学界才意识到，由于内容创作者可能采取策略行为，用户效用无法通过单方面行动实现最大化（Qian 和 Jain，2022年）。由于内容创作者的效用与其内容的曝光度直接相关，他们有动机灵活调整策略，以最大化自身效用。这会引发竞争，而竞争可能损害社会福利（定义为用户满意度／参与度总和）（Fleder 和 Hosanagar，2009年）。例如，设想一个用户群体中有一大群体育迷和一小群旅行爱好者。当可供推荐的内容涵盖这两个主题时，社会福利达到最大化。然而，一种可能出现的竞争均衡是：当创作小众内容带来的收益不足以弥补放弃多数用户曝光度所造成的效用损失时，所有创作者都会发布同质化的体育内容。因此，亟需了解从长远来看，在由 top-推荐系统驱动的环境中，策略性内容创作者可能导致的社会福利损失究竟有多严重。

在本文中，我们提出了竞争性内容创作博弈，以模拟创作者竞争对 Top-K RS 中用户参与度的影响。我们通过无政府价格（PoA）的视角衡量社会福利保证（Koutsoupias 和 Papadimitriou，1999年）。无政府价格通过博弈均衡中最差福利值与最优结果福利值之比，量化自私行为造成的低效。此前有一些研究在不同的竞争模型下探讨了这一问题，但得出的结论都较为悲观。例如，Ben-Porat 和 Tennenholtz （2018年）指出，在由 Shapley 中介机制实现的 RS 中，社会福利的无政府代价没有上界。Ben- Porat 等人 （2019年）研究了一维空间中的竞争模型，并证明在 Top-1 匹配原则下，无政府代价可能高达常数 2。这些负面结果要么基于确定性用户选择模型，要么假设创作者争夺内容曝光份额。我们推翻了这些悲观结论，证明当（1） $K > 1$，（2）用户选择受到轻微随机噪声影响，且（3）创作者的竞争目标是用户参与度而非内容曝光时，Top-K RS 所导致的无政府代价至多为 $1 + O(\ ^{1}/_{\text{对数}\,K})$。我们还通过分析一个下界实例证明了该界是紧的。因此，在这些假设下，随着 K 增大，RS 将趋近于最优效率（即无政府代价比趋近于 1），但这一收敛速度相对较慢，为 $1/$对数 K。值得注意的是，我们的无政府代价上界在动态场景中同样成立：创作者会以在线方式逐步学习并改进策略。基于大量合成数据和真实世界数据的模拟也支持了这些理论发现。总而言之，我们的结果有力地表明，在激励机制设置得当时，内容创作竞争是有效率的。这印证了亚当·斯密在本节开头所引述的著名观点：市场的“看不见的手”能够推动商品的生产实现社会效率。

我们的结果基于三个关键假设，它们都源自推荐领域的研究与实践。首先，在平台方面，我们假设排名前列的K 推荐系统基于相关性函数；如果用户消费了推荐内容，该函数能最好地预测用户满意度。为简化研究设定，我们假设推荐系统已知真实的相关性函数，因为已有大量研究致力于这一问题（Bobadilla 等人，2013年；Konstan 等人，1997年；Koren 等人，2009年；He 等人，2017年），而本研究的目标并非改进对它的估计。其次，在用户方面，我们采用成熟的随机

效用（*RU*）模型 (Manski，1977) 来描述用户面对一系列推荐时的选择分布结构及由此产生的效用。RU 模型已得到广泛采用，并成功应用于营销研究中的消费者选择建模（Baltas 和 Doyle，2001年）。第三，在创作者方面，我们假设创作者将其内容与用户匹配后获得的效用与用户效用成正比，因为平台通常会按照用户满意度或参与度的比例，与内容创作者分成（Meta，2022年；Savy，2019年；YouTube，2023；TikTok，2022年）。在转向创作者无法通过预言机获取其效用函数的动态设定时，我们允许创作者采用任意无悔学习算法，涵盖多种理性学习行为。

我们结果的另一项优势在于，对社会福利的刻画不要求竞争具备任何稳定性（例如纳什均衡的存在性或唯一性），也就是说，无论一系列内容创作策略是否收敛，该结果都适用于任何合理的内容创作策略序列所诱导的内容分布。虽然先前的研究（Ben-Porat 和 Tennenholtz，2018年，2017年；Ben-Porat 等人，2019年）致力于在类似的博弈论设定中证明纯纳什均衡（PNE）的唯一性，但我们认为，这种稳定性要求对系统设计并非至关重要，原因有二。首先，正如先前研究所示（Ben-Porat 和 Tennenholtz，2018年，2017年；Ben-Porat 等人，2019年），对于 top-K 推荐系统而言，通常无法保证唯一的纯纳什均衡。然而，我们的主要结果表明，无论纯纳什均衡是否存在，top-K推荐系统都能确保社会福利，因此无需在稳定性与推荐算法的复杂性之间做出权衡。其次，即使纯纳什均衡确实存在，创作者在实践中如何实现这种稳定状态也并不明确。例如，Ben-Porat 和 Tennenholtz (2018年，2017年)证明，采用最佳响应动态的创作者（Monderer 和 Shapley，1996）会收敛到其博弈论设定中的唯一纯纳什均衡，但这要求创作者能够通过预言机访问其效用函数；而这并不现实，因为在实践中，创作者只能评估自己已采取动作的效用（即多臂老虎机反馈）。因此，我们不讨论稳定性，而是侧重于刻画创作者演化策略下的平均社会福利。

2. 相关工作

关于内容创作者在推荐系统（RS）中介下策略行为的理论研究，可追溯至Ben-Porat 和 Tennenholtz（2017年，2018年）的开创性著作。他们扩展了搜索与排序系统中的博弈设定（Ben Basat 等人，2015；Raifer 等人，2017年），并提出了一种基于 Shapley值的推荐系统，该系统能够导向唯一的纯纳什均衡（PNE），并满足若干与公平相关的要求。然而，他们证明，在所提出的 Shapley 中介机制下，社会福利可能会变得任意糟糕。

另一类研究是在霍特林空间竞争框架下，探讨策略性内容创作者参与的推荐系统（1929年）。霍特林（1929年）最早提出的霍特林模型研究了两家餐馆如何确定各自的位置，以吸引均匀分布在区间 $[0,1]$ 上的顾客。由此产生的博弈的纳什均衡是两家餐馆都选址于中心位置，这被称为“最小差异原则”。近期，Shen 和 Wang（2016年）提出了霍特林竞争的一种变体，其中每个参与者都有自己的吸引区域，并证明了

在最坏情况下，纳什均衡下的社会福利仅为可能达到的最大社会福利的一半。我们在附录 F 中证明，他们的博弈设定是我们提出的竞争性内容创作博弈的特例，因此我们的主要结果可以直接推出他们的界限。与本文更为相关的是Ben-Porat 等人（2019年）的研究，他们将推荐系统引入竞争，将其作为引导用户前往设施的中介。他们研究了不同干预程度的中介，并提出一种有限干预中介，在社会福利和干预成本之间实现了良好权衡。有趣的是，他们在无干预中介下的博弈设定也恰好是我们的一个特例。我们还指出，（Shen 和 Wang，2016年）和（Ben-Porat 等人，2019年）的设定和理论讨论都仅限于一维情形下的纯策略，且用户效用函数由距离决定；而我们的模型和结果适用于任意维度以及一般形式的用户效用函数。

近期有两项研究（Hron 等人，2022年；Jagadeesan 等人，2022年）探讨了供给侧竞争，其中创作者的策略空间维度很高。这些模型假设创作者直接争夺用户曝光，而不考虑推荐系统的作用。他们重点刻画纳什均衡，并识别创作者策略专业化可能出现的条件。相比之下，我们研究了顶级 K 推荐系统影响下的社会福利，不局限于纳什均衡的存在性，而且我们的结果适用于一般的用户效用函数。

我们的用户决策模型源自计量经济学中的 RU 模型（Baltas 和 Doyle，2001年），该模型解释个体如何在一组离散的备选项中做出选择。在 RU 模型中，决策者从备选项 j 中获得的效用被分解为 $U_j = V_j + \epsilon_j$，其中 V_j 是已知的参数化部分，而 ϵ_j 是未知的随机部分。随后，观测到的选择就是效用最大的备选项。研究表明，如果未观测到的随机效用服从极值分布（即 Gumbel 分布），那么选择概率由 Logit 公式给出，即 $P_j \propto \exp(V_j)$（Luce 等人，1965）。在我们的研究中，我们运用 RU 模型解释典型用户如何在推荐列表中分配注意力。

最后，为了分析竞争性内容创作博弈的均衡效率，我们采用了无政府代价（PoA）的标准框架。这一框架源自Koutsoupias 和 Papadimitriou （1999年）的开创性研究，此后催生了大量文献，致力于理解众多策略博弈的效率。我们无法在此充分评述这一丰富的文献；这里只提及少数与我们的研究密切相关的工作。由于在我们的智能体效用不连续的问题中，纳什均衡（NE）并不一定存在（Hron 等人，2022年），因此我们必须考虑一种弱于 NE 的解概念，并由此证明一个更强的无政府代价界。具体而言，我们考虑粗相关均衡（CCE）。Blum 等人Blum 等人 （2008年）首次研究了 CCE 的无政府代价，他们考察了由无悔学习者参与的动态设置的效率，并提出了完全无政府代价这一新概念；事实证明，它等价于 CCE 的无政府代价界。这正是我们想要探讨的问题，但我们提出的竞争性内容创作博弈在结构上与他们研究的博弈显著不同，例如图上的 Hotelling 博弈以及 Vetta 的有效效用博弈Vetta （2022年）。因此，他们的技术并不能直接应用于我们的问题。我们转而采用 RoughgardenRoughgarden （2015）近期提出的光滑性论证框架。众所周知，该框架能够给出适用于 CCE 的强无政府代价界。然而，这些界

由这一强大框架得到的界通常并不紧；迄今为止，已知它仅能为线性成本拥塞博弈（Christodoulouand Koutsoupias，2005年）、第二价格拍卖（Christodoulou 等人，2008年）和有效效用博弈（Vetta，2022年）给出紧的无政府代价界。有趣的是，我们证明，光滑性论证也能为我们的竞争性内容创作博弈给出紧的无政府代价界，从而为这份重要的博弈清单增添了一个新成员。

3. 内容创作者竞争模型

本节将对竞争性内容创作博弈进行形式化定义。该博弈 $\mathcal{G}$ 由元组 $(\{\mathcal{S}_i\}_{i=1}^n,\ \mathcal{X},\ \sigma,\ \beta,\ K)$ 定义，其组成部分如下：

1. 一个用户有限集 $\mathcal{X} = \{\boldsymbol{x}_j \in \mathbb{R}^d\}_{j=1}^m$，以及一个玩家集（即内容创作者[1]），记为 $[n] = \{1, \cdots, n\}$。每个玩家 i 都可以从行动集[2] $\mathcal{S}_i \subset \mathbb{R}^d$中采取一个动作 $\boldsymbol{s}_i$，在博弈论文献中通常称为纯策略。 $\boldsymbol{s}_i$ 可以理解为创作者 i 能够创作的内容类型的嵌入。令 $\mathcal{S} = \prod_{i=1}^n \mathcal{S}_i$ 表示联合策略空间。按照惯例，对于任意 $\boldsymbol{s} =(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$，我们用 $\boldsymbol{s}_{-i}$ 表示不包括 $\boldsymbol{s}_i$的联合策略 $\boldsymbol{s}$。此外，我们用 $\boldsymbol{\alpha}_i \in \Delta(\mathcal{S}_i)$表示玩家 i的混合策略，它是一个支撑集为 $\mathcal{S}_i$的概率测度。类似地， $\boldsymbol{\alpha} \in \Delta(\mathcal{S})$用于表示所有玩家的（可能相关的）联合策略组合分布。

2. 相关性函数 $\sigma(\boldsymbol{s}, \boldsymbol{x}) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$ 用于衡量用户 $\boldsymbol{x} \in \mathcal{X}$ 与内容 $\boldsymbol{s}$之间的相关性。不失一般性，我们将 σ 归一化为 $[0, 1]$, ，其中 1 表示完全匹配。我们专注于对创作者的策略行为进行建模，因此不考虑对 σ的估计。

3. 推荐策略：给定所有玩家的联合策略 $\boldsymbol{s} =(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$，对于每位用户 $\boldsymbol{x}_j$，推荐系统首先计算所有可用内容的相关性得分 $\{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j)\}_{i=1}^n$ ，然后生成 $\mathcal{T}_j(\boldsymbol{s}; K)$，即包含面向用户 j的前 K 条推荐的 $\boldsymbol{s}$ 子集。形式化地，

$$\mathcal{T}_j(\boldsymbol{s}; K) = \{\boldsymbol{s}_{l_i} | i = 1, \cdots, K\}, \tag{1}$$

其中，$(l_i)_{i=1}^n$ 是 $[n]$ 的一个排列，满足 $\sigma(\boldsymbol{s}_{l_1}, \boldsymbol{x}_j) \geq \sigma(\boldsymbol{s}_{l_2}, \boldsymbol{x}_j) \geq \cdots \geq \sigma(\boldsymbol{s}_{l_n}, \boldsymbol{x}_j)$。[3]

4. 用户效用与选择模型：我们采用广泛使用的随机效用（RU）模型来描述用户的效用和对推荐内容的选择。形式上，RU 模型假设用户 $\boldsymbol{x}_j$ 消费内容 $\boldsymbol{s}_i$ 时的效用为 $\sigma(\boldsymbol{s}_i,\ \boldsymbol{x}_j) + \varepsilon_i$，其中$\varepsilon_i$ 是一个噪声项，包含无法由推荐系统对 $\sigma(\boldsymbol{s}_i,\ \boldsymbol{x}_j)$ 的预测所捕捉的任何额外不确定性（例如，用户当时的心情）。RU 模型假设 $\{\varepsilon_i\}$ 是独立同分布的随机变量，通常假设它们服从*Gumbel* 分布

1. 在不存在歧义时，我们交替使用这两个术语。2. 不假设动作集是有限的，因此动作集可以是连续的。3. 当 $(l_i)_{i=1}^n$ 不唯一时， $\mathcal{T}_j(\boldsymbol{s}; K)$ 可以等概率地取任意此类排列的前 K 项截断结果。

其累积分布函数为 Gumbel(μ, β) $= e^{-e^{-\frac{x-\mu}{\beta}}}$。[4] 我们进一步假设$\varepsilon_i$ 均值为零，因此可得 $\mu = -\beta\gamma$ 其中 $\gamma \approx 0.577$ 是欧拉–马歇罗尼常数。Gumbel($-\beta\gamma$, β) 的方差为 $\frac{\pi\beta}{\sqrt{6}}$，参数 β 衡量噪声水平。

在收到推荐列表 $\mathcal{T}_j(\boldsymbol{s}; K)$ 后，用户 j 选择使其效用最大化的 $i_j^* \in \mathcal{T}_j(\boldsymbol{s}; K)$:

$$i_j^* = \arg \max_{\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s};K)} \{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\}. \tag{2}$$

注意， i_j^* 是随机的，其随机性继承自 $\{\varepsilon_i\}$。因此，用户 j消费所选内容可获得如下期望效用π_j

$$\pi_j(\boldsymbol{s}) \triangleq \mathbb{E}_{\{\varepsilon_i\}} \left[\max_{\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s};K)} \{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\} \right]. \tag{3}$$

5. 玩家效用：遵循惯例，我们假设每个参与者 i的期望效用等于来自其 i 服务的用户的效用之和，即

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \Pr[\boldsymbol{x}_j \to \boldsymbol{s}_i], \tag{4}$$

其中"$\boldsymbol{x}_j \to \boldsymbol{s}_i$"表示事件 i_j^* 在（2）中定义为 i。简洁地说，对于任意 $[\boldsymbol{x}_j \to \boldsymbol{s}_i] \propto e^{\beta\sigma(\boldsymbol{x}_j, \boldsymbol{s}_i)}$，$i \in \mathcal{T}_j$（$\boldsymbol{s}$； K）（McFadden，1974），且当 $[\boldsymbol{x}_j \to \boldsymbol{s}_i] = 0$ 时，$i \notin \mathcal{T}_j$（$\boldsymbol{s}$； K）。

6. 社会福利：社会福利函数定义为所有用户效用的总和：

$$W(\boldsymbol{s}) = \sum_{j=1}^{m} \pi_j(\boldsymbol{s}). \tag{5}$$

注意，在玩家效用函数（4）下，我们有 $W(\boldsymbol{s}) = \sum_{i=1}^{n} u_i(\boldsymbol{s})$。也就是说，社会福利也是玩家的总效用。

我们指出，式 (4) 中定义的参与者 i 的效用不仅取决于与 i 匹配的用户比例，还取决于由项 $\mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i]$ 所体现的用户参与度。这与霍特林竞争（Hotelling，1929年）及其近期在推荐系统中的应用（Shen 和 Wang，2016年；Ben-Porat 等人，2019年；Hron 等人，2022年；Jagadeesan 等人，2022年）的设定有本质区别；在这些设定中，玩家的效用被设定为总用户曝光量，即用户访问的总次数或比例（不论用户对推荐内容是否满意）。这两项指标在当前工业界实践中都被广泛用于奖励创作者（Meta，2022年；Savy，2019年）。在本文中，我们

4. 使用 Gumbel 噪声模型有许多自然的理由。从经验上看，这种噪声模型与高斯分布几乎无法区分，但其尾部稍厚，因此可以容纳更多异常的用户行为。带有 Gumbel 噪声的 RU 模型也称为多项式Logit模型（McFadden，1974）。它与离散选择模型（McFadden，1984）密切相关；量子响应均衡用于刻画有限理性行为（McKelvey andPalfrey，1995），熵正则项则用于优化随机化策略（Ling et al.，2018年）。

主要将用户参与度（即此前研究较少的情形）视为创作者的效用，并将在第 4.2 节将其与用户曝光指标进行比较，以突出二者影响的差异。

我们的研究问题与均衡概念。我们尤其关注创作者以适应性方式学习并更新策略时的平均社会福利。具体而言，我们考虑由 n 位创作者参与、持续一段时间 T 的重复性竞争性内容创作博弈。在每个时刻 t，每位创作者都会选择一个动作，观察该轮所有创作者策略所产生的效用，并利用反馈调整后续动作。自然地，创作者希望在整个交互过程中最大化累积效用。然而，在现实世界的在线推荐平台上，创作者只能评估自己所选动作的效用，必须通过反复试错，在信息有限的情况下（即多臂老虎机反馈）逐步学习最优策略。在这种环境下，用于刻画"合理"学习行为的一个自然概念是无悔。参与者 i 的（外部）遗憾 $R_i(T)$ 定义为事后最优效用与实际累积效用之差，即

$$R_i(T) = \max_{\boldsymbol{s}_i'} \sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}_{-i}\sim\boldsymbol{\alpha}_{-i}^t}[u_i(\boldsymbol{s}_i', \boldsymbol{s}_{-i})] - \sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[u_i(\boldsymbol{s})] \tag{6}$$

其中，$\boldsymbol{\alpha}^t = \prod_{i=1}^n \boldsymbol{\alpha}_i^t$ 表示时刻 t的联合策略分布。参与者 i的学习是无悔的，如果 $R_i(T) = o(T)$，或者等价地，平均遗憾 $R_i(T)/T \to 0$ 在 T 趋于无穷时趋于零。请注意，这类无悔算法确实存在，因为任何无悔的对抗式在线学习算法（例如，多臂老虎机文献中的 Exp3（Auer 等人，2002））都能在这种多智能体学习设置中保证无悔。

为了刻画无遗憾学习玩家的结果，我们聚焦于一种称为粗相关均衡（CCE）的均衡概念，因为众所周知，在重复博弈中，任何无遗憾重复博弈序列的经验动作分布都会收敛到其粗相关均衡集合（Blum 等人，2008年）。粗相关均衡的正式定义如下：

定义 1 *粗相关均衡（CCE）是联合策略组合空间上的一个分布 $\boldsymbol{\alpha}$，满足对于每个玩家 i 和每个动作 $\boldsymbol{s}_i' \in \mathcal{S}_i$,*

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i', \boldsymbol{s}_{-i})]. \tag{7}$$

得益于无遗憾动态与粗相关均衡之间的良好联系，我们先在第4.1 节确立粗相关均衡的福利保证，然后在第 4.2节将其扩展到由重复博弈所带来的累积福利。

我们还注意到，粗相关均衡（CCE）的概念之所以特别有用，还有两个原因。首先，在任何有限博弈中（因此也包括我们的博弈中），粗相关均衡始终存在，因此无需像以往研究那样处理纳什均衡（NE）是否存在的问题；纳什均衡或许是最著名的解概念（Hron 等人，2022年）。事实上，当行动集是连续的时，我们的博弈中无法保证纳什均衡（无论是纯策略还是混合策略）存在，因为式（4）中定义的玩家效用函数并不连续。这是该问题固有的挑战，因为对 $\sigma(\boldsymbol{s},\ \boldsymbol{x})$ 的任何改变都可能导致不同的 Top-K 推荐列表 $\mathcal{T}_j(\boldsymbol{s};\ K)$，进而使玩家效用发生巨大变化。Hron 等人（2022年）也观察到了类似挑战以及混合纳什均衡不存在的情况，不过他们的

效用模型和研究问题与我们的不同。其次，即使在纳什均衡存在的情况下，考虑到人们对纳什均衡的种种质疑（包括其计算方面的问题（Daskalakis 等人，2009年）），假设玩家最终达到某种粗相关均衡，比假设他们最终达到纳什均衡更符合现实。

4. 无政府代价分析

我们分析任意 top-K RS 在任意可能的粗相关均衡下的社会福利；更具体地说，自利内容创作者之间的竞争可能导致福利变得多糟 ——与 理想化的 非策略情形相比，在该情形下，平台可以“指示”所有创作者选择内容，从而全局优化福利函数（5）。这可以通过著名的无政府价格（PoA）概念来刻画（Koutsoupias 和 Papadimitriou，1999年）。顾名思义，PoA 刻画了玩家的自私行为所导致的福利低效。本节的主要结果是对竞争性内容创作博弈的无政府价格进行全面刻画。

定义 2（粗相关均衡下的无政府价格）将博弈 $\mathcal{G}$ 的无政府价格定义为

$$PoA(\mathcal{G}) = \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]}, \tag{8}$$

其中，CCE（$\mathcal{G}$）是 $\mathcal{G}$ 的粗相关均衡集合。

根据定义，PoA($\mathcal{G}$) ≥ 1 始终成立，且数值越大表示福利越差。我们选择粗相关均衡这一解概念，带来了最强的福利保证：任何在粗相关均衡下成立的 PoA 上界，也必然平凡地适用于相关均衡（CE）、纯纳什均衡或混合纳什均衡（如果它们存在）下的 PoA，因为这些均属于粗相关均衡。除非另有特别说明，本文中的任何 PoA 均指粗相关均衡下的 PoA。

4.1 匹配的无政府代价上下界

我们的主要理论发现是无政府代价的上下界，二者相匹配，从而证明了我们分析的紧确性。我们首先给出如下上界。

定理1 任意满足参数$\beta \geq 0$ 和 $K \geq 1$ 的竞争性内容创作博弈实例 $\mathcal{G}$，其无政府代价满足

$$PoA(\mathcal{G}) < 1 + \frac{1}{c(\beta, K)}, \tag{9}$$

其中，c（β、K）定义为

$$c(\beta, K) = \frac{(b+1)\log(b+K)}{(b+K)(\log(b+K) - \log K)}, b = e^{\frac{1}{\beta}} - 1. \tag{10}$$

定理 1 的证明较为复杂，因此详细论证放在附录 A中。证明的主要难点在于分析福利和

玩家效用函数的各种光滑性，尤其是排除任意一名玩家 i参与后福利函数如何变化。在第4.3节中，我们着重介绍福利函数的一些显著性质，包括其子模性。这些性质是在证明定理1 的过程中推导出来的，同时也有助于独立理解竞争性内容创作博弈。

.....的形式 $c(\beta,\ K)$ 对读者而言可能不够直观，难以理解推导出的无政府代价上界。因此，我们给出以下观察，揭示了 $c(\beta,\ K)$ 的各种性质，有助于理解 (9)：

1. 对任意 $\beta > 0$ 和 $K \geq 1$，我们有 $c(\beta,\ K) \geq 1$，因此 PoA($\mathcal{G}$) < 2 总是成立。

2. 当且仅当 $K = 1$ 或 $\beta \to 0$ 时，$c(\beta,\ K) = 1$。

3. 固定任意 $\beta > 0$，$c(\beta,\ K)$ 随 K 单调递增；类似地，固定任意 $K \geq 1$，$c(\beta,\ K)$ 随 β 单调递增。

4. 对于 足够大的 β 和 K，$c(\beta,\ K) \approx (1+\beta)$ 渐近地为 log K，因此 e

$$PoA(\mathcal{G}) < 1 + \frac{1}{(1+\beta)\log K}. \tag{11}$$

基于这些观察，定理1 有多个有趣且直接的推论。首先，在任何情况下，任一粗相关均衡下的福利损失最多为一半，因为无政府代价始终不超过 2。上述第二和第三个事实表明，这种最坏情况的无政府代价确实会出现，而且仅当用户以“极端”的方式做出选择时才会出现：要么推荐系统通过设置 $K = 1$ 来决定用户的选择，要么用户选择中的随机性极低（即 $\beta \to 0$）。请注意，在后一种情况下，由于决策随机性很低，用户只会消费最相关的内容（即排名最高的内容）。

其次，随着 K 增大（即推荐的项目增多），或 β 增大（即用户选择的随机性增强），福利保障都会改善。后一种情况下福利的改善很容易理解，因为当用户获得多个项目时，就可以选择 ε_i 较大的内容（即推荐系统无法预测的奖励部分），从而获得效用。这些发现共同揭示了一项有趣的实践启示：当推荐系统无法完美预测用户效用（即 $\beta > 0$）时，提供更多项目有助于提升社会福利。这为 top-K 推荐以及推荐多样性的必要性提供了依据（Hurley和Zhang，2011）。

我们的第二个主要结果表明，这个无政府代价上界是紧的，误差仅为可忽略的常数。

1**定理 2** 对于任意 $0 \leq \beta \leq 1$、$n > 2$，以及任意 $1 \leq K \leq$最小值$\{n=1, e^{5\beta}\}$，都存在一个竞争性内容创作博弈实例 $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$，使得

$$PoA(\mathcal{G}) > \frac{n-1}{n} + \frac{1}{1+5\beta\log K}. \tag{12}$$

该定理还意味着，我们为定理 1所采用的、基于 光滑性 证明的方法由 Roughgarden (2015) 提出，能够为我们提出的博弈给出紧的无政府代价界。光滑性论证本身是否紧，也是一个引人入胜的

研究问题。目前已知，只有三类博弈能够通过光滑性论证得到紧的无政府代价界：具有仿射成本的拥塞博弈（Christodoulouand Koutsoupias，2005年）、第二价格拍卖（Christodoulou 等人，2008年）以及有效效用博弈（Vetta，2002），它们都是基础性的博弈类别。定理 2表明，我们的竞争性内容创作博弈也属于这一范畴。定理 2 的证明通过明确构造一个博弈实例，证明其确实能够达到所述的无政府代价下界（见附录 D）。

4.2 无政府代价界的启示

我们已经讨论了定理 1的一些直接推论。现在，我们将提出一些新结果，它们要么由定理 1和 2推导而来，要么可与之进行比较。这些结果将从我们的主要理论结果中揭示更多见解。

对 学习型内容创作者的福利启示。定理 1和2 中给出的无政府代价界基于这样一个假设：创作者了解博弈参数，并采用该博弈的某种粗相关均衡。尽管粗相关均衡是一种合理的均衡概念，但一种潜在的批评是，寻找粗相关均衡需要假设每位创作者都掌握系统参数（例如所有其他创作者的策略以及σ 函数），而这可能并不现实。幸运的是，在现实场景中，即使创作者利用无悔算法，通过多臂老虎机反馈来参与重复竞争性内容创作博弈，我们仍可利用以下事实，得到一个略差的无政府代价上界：无悔玩家的平均策略历史会收敛到粗相关均衡。如下文推论 1所示。

推论 1 *[定理的动态版本 1]* 假设重复竞争内容创作博弈中的每个参与者 $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ 均独立执行某种无悔学习算法，其最坏遗憾为 $R(T) = \max_i R_i(T)$，如 (6)中所定义。那么我们有

$$\frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]} < 1 + \left(1 + \frac{n}{\beta \log K} \cdot \frac{R(T)}{T}\right) \cdot \frac{1}{c(\beta, K)}, \tag{13}$$

其中 $\boldsymbol{\alpha}^t$ 表示第 t 步的联合策略分布，且 $c(\beta, K)$ 是 (10)中定义的常数。

换言之，所有轮次的平均福利 $\frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]$ 接近最大可能社会福利 $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$，二者相差至多一个常数因子。式 (13) 左侧的量也称为“完全无政府代价”（Blum 等人（2008））。当我们希望刻画重复博弈所得结果的福利、而该结果不一定属于任何均衡概念时，它可作为无政府代价的替代指标。推论 1 的证明见附录 B。由于对于任何无悔算法都有 $R(T)/T \to 0$ as $T \to \infty$（大多数无悔算法满足 $R(T) = O(\sqrt{T})$），对于任意固定常数 (n, β, K)，式 (13) 右侧仍严格小于 2。

定理 1 依赖于平台的两个关键特征：1. 式（4）中的参与者效用定义为考虑用户效用的总参与度 $\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_j$，而非仅仅是“用户曝光”（即匹配总数的期望值）；2. 平台采用 top-K 推荐策略。接下来，我们将围绕这两个关键特征，阐释定理1 揭示的见解。

奖励用户参与度而非单纯奖励曝光度的重要性。 我们的竞争性内容创作博弈之所以能获得良好的无政府代价保证，一个关键原因是每位参与者i的效用被定义为用户参与度 $\sum_j \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \Pr(\boldsymbol{x}_j \to \boldsymbol{s}_i)$，如式（4所示，而非采用以下用户曝光度指标：

$$\text{User exposure for player } i: \quad \sum_{j=1}^{m} \Pr(\boldsymbol{x}_j \to \boldsymbol{s}_i). \tag{14}$$

我们的下一个结果表明，激励创作者最大化用户曝光，可能会导致福利显著恶化。

命题1 令 $\mathcal{G}$ 表示竞争性内容创作博弈 $\mathcal{G} = (\{\mathcal{S}_i\}_{i=1}^n, \{\boldsymbol{x}_j\}_{j=1}^m, \sigma, \beta, K)$ 的一个变体，其中将(4) 中的玩家效用函数替换为上述(14)中的用户曝光。那么，对于任意 $K \geq 1$，$0 \leq \beta \leq$最小值$\{0.14, \frac{1}{5 \text{ 对数 } K}\}$，均存在 $\mathcal{G}$和 $\tilde{\mathcal{G}}$，使得

$$PoA(\tilde{\mathcal{G}}) > 2 > PoA(\mathcal{G}). \tag{15}$$

此外，当 $K = 1$ 或 β 趋近于0时，无政府代价$(\mathcal{G})$ 可以任意大。

与保证无政府代价$(\mathcal{G}) < 2$的定理 1 形成鲜明对比的是，命题 1表明，当激励内容创作者竞相争取其内容的预期曝光时，用户福利会下降。然而，我们发现实践中这两种指标都会被使用：例如，用户参与度更常被用作成熟创作者的奖励指标，而用户曝光则更多用于新创作者 Meta (2022); Savy (2019年). 如果系统旨在提升推荐内容的整体福利，我们的结果为依据用户参与度奖励创作者提供了理论支持。

为了证明命题 1，我们构造了一个博弈实例，使得纳什均衡下的用户福利可以任意接近于零。我们的构造还揭示了用户福利可能极差的情形所蕴含的有趣见解。因此，我们在此简要说明这一构造，并将正式论证留在附录E中。我们构造的博弈包含两组用户：一组是兴趣分散的用户，他们对任何内容都不挑剔，但也从不会特别满意（即对所有内容的相关性得分都很低）；另一组是兴趣集中的用户，他们寻求特定类型的高质量内容（对这类内容的相关性得分很高）；但只有一小群专业创作者能够产出这类高质量内容。然而，如果激励玩家争夺曝光度，即使是这小群创作者中的创作者，也往往会转而产出吸引兴趣分散用户的低质量内容，而不是能让兴趣集中用户受益的高质量内容。在最糟糕的情况下，这可能导致平台福利任意恶化。

top-K 推荐策略的福利效率。 人们或许会好奇，top-K 推荐是否确实是保障平台福利的良策，也就是说，其他推荐策略（例如基于 Plackett-Luce模型的概率策略（Plackett, 1975年; Luce, 1959年））是否可能带来更好的均衡结果？作为定理 1的推论，我们接下来的分析表明，在某种程度上，答案是 否，因为 任何 推荐策略都不可能比 top-K 规则好出超过理论最优值一小部分的幅度。我们认为，这一发现也为实践中广泛采用 top-K 原则提供了理论依据。

推论 2 考虑任意一种推荐策略，其提供的推荐数量至多为 K，该策略会导出一个不同的竞争性内容创作博弈 $\mathcal{G}'$。令 $CCE(\mathcal{G}')$ 表示相应的 $\mathcal{G}'$ 的粗相关均衡集合，且 $W(\mathcal{G}') = \min_{\boldsymbol{\alpha} \in CCE(\mathcal{G}')} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$ 表示其最坏情况下的粗相关均衡社会福利。那么我们有

$$W(\mathcal{G}') \leq W(\mathcal{G}) + W_K^* / \Big(1 + \frac{K \log(K+b)}{K+b}\Big), \tag{16}$$

其中，W_K^* 是通过具有 K 个槽位的任意集中式推荐策略所能实现的最佳社会福利。

如（16）所示，福利损失所占的比例近似为 $O(\frac{1}{\log K})$，当 $\frac{K}{K+b} \sim O(1)$ 较大时。~~K~~ 证明基于定理1 即可直接得到，详见附录C。

4.3 定理1的证明要点

我们的第一步是利用 Gumbel 分布的性质，为博弈基本要素推导出简洁的刻画。用户效用 π_j 和福利 W 是随机效用模型的推论（Baltas 和 Doyle，2001年），然而，创作者效用 u_i 的闭式表达式是我们推导出的一个新性质。相关结果总结于以下引理1中，详细证明见附录A.1。

引理1 给定 $\{\varepsilon_i\}$ 独立同分布地取自均值为零的 *Gumbel* 分布（$-\beta\gamma$，β），定义于（3）、（4）和（5）中的效用函数和福利函数具有以下闭式形式

$$\pi_j(\boldsymbol{s}) = \beta \log \Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right) \Big], \tag{17}$$

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \pi_j(\boldsymbol{s}) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s};K)] \exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{18}$$

$$W(\boldsymbol{s}) = \beta \sum_{j=1}^{m} \log \Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right) \Big]. \tag{19}$$

定理1的主要证明基于 Roughgarden（2015）开创性工作中提出的光滑性论证框架。对于任意策略组合 $\boldsymbol{s}$, $W(\boldsymbol{s}) = \sum_i u_i(\boldsymbol{s}_i, \boldsymbol{s}_{-i})$ 是其总社会福利函数。若一个博弈满足 (λ, μ)-光滑性，则对于任意 $(\boldsymbol{s}, \boldsymbol{s}') \in \mathcal{S}$，$\lambda W(\boldsymbol{s}') - \mu W(\boldsymbol{s}) \leq \sum_i u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})$。Roughgarden（2015）指出，任何 (λ, μ)-光滑博弈的无政府代价都可以用 $\frac{1+\mu}{\lambda}$ 作为上界。代入 $W(\boldsymbol{s})$ 的表达式后，(λ, μ)-光滑性条件可改写为

$$\sum_i [\lambda u_i(\boldsymbol{s}'_i, \boldsymbol{s}'_{-i}) - u_i(\boldsymbol{s}'_i, \boldsymbol{s}_{-i})] \leq \sum_i \mu u_i(\boldsymbol{s}).$$

直观而言，光滑性参数界定了其他玩家的动作（即 或 ）对任意玩家 的效用所施加的外部性有多大。而且，这个界限越紧，

博弈就越光滑，无政府代价也越小。为了获得一些直观认识并进行合理性检查，不妨考虑一种极端情形：每个玩家的效用完全不受其他玩家动作的影响（即无外部性情形）。此时有 $\lambda = 1$ 且 $\mu = 0$，这意味着无政府代价=1。也就是说，如果任何玩家的效用都不受其他玩家影响，那么以自身利益为出发点、使效用最大化的行为也会使社会福利最大化，这是一个显而易见的结论。当然，我们不能指望这种优良性质普遍成立，但幸运的是，许多著名的博弈已被证明是光滑的。例如，Christodoulou 等人（2008年）证明，第二价格拍卖是 $(1, 1)$-光滑的；Roughgarden（2015）证明，拥塞博弈是 $(\frac{5}{3}, \frac{1}{3})$-光滑的；Roughgarden 等人（2017年）证明，全支付拍卖是 $(1/2, 1)$-光滑的。

因此，证明定理1 的关键挑战在于，为我们的竞争性内容创作博弈确定尽可能精确的（λ，μ）参数。这归结为 top-K 推荐系统中的一个基本问题，即：其他竞争性内容创作者的存在会在多大程度上影响创作者的效用？为回答这一问题，我们发现了福利函数和创作者效用函数的若干有趣性质，具体如下。除了证明定理1中的主要结果外，我们认为这些性质也有助于理解推荐系统。

我们的第二个引理 2 证明了 $W(\boldsymbol{s})$ 的子模性。也就是说，随着创作者总数增加，新增一名参与者所带来的福利边际增益会降低。引理 3 进一步将这一福利边际增量与新增参与者自身的效用联系起来。它表明，引入一名采用策略 $\boldsymbol{s}_i$ 的新参与者 i 后，福利的增量至多为 i 在 $\boldsymbol{s}_i$ 下的效用乘以收缩因子 $c^{-1}(\beta, K) \in (0, 1]$。这两个引理共同使我们得以证明，竞争性内容创作博弈是（$c^{-1}(\beta, K)$ $c^{-1}(\beta, K)$)-平滑的，从而得到定理1。详细证明见附录 A.2。

引理 2 *[福利的子模性]对于任意 $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$，令 $S = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$。则式 (5) 中定义的社会福利函数是一个集合函数，且具有子模性；也就是说，对于任意 $S, \boldsymbol{s}_x, \boldsymbol{s}_y$，均有*

$$W(S \cup \{\boldsymbol{s}_x\}) - W(S) \geq W(S \cup \{\boldsymbol{s}_x, \boldsymbol{s}_y\}) - W(S \cup \{\boldsymbol{s}_y\}).$$

引理 3 *[福利的光滑性]对于任意 $\boldsymbol{s} = (\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n) \in \mathcal{S}$，$i \in [n]$ 以及 $c(\beta, K)$（定义见式 (10)）中，参与者 i的效用函数 $u_i(\boldsymbol{s})$（定义见式 (4)）满足*

$$W(\boldsymbol{s}) - W(\boldsymbol{s}_{-i}) \leq c^{-1}(\beta, K) \cdot u_i(\boldsymbol{s}_i; \boldsymbol{s}_{-i}).$$

5. 实验

为验证我们的理论发现，并通过实证测量创作者竞争所带来的社会福利，我们针对博弈实例进行了模拟。
$\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ 由两个合成数据集和 MovieLens-1m 数据集构建而成（Harper and Konstan, 2015）。在介绍结果之前，我们先详细概述模拟环境，包括所用数据集的特征以及用于评估的指标。

5.1 数据集说明

• **合成数据集-1** 数据集 1 模拟了内容创作者在不均衡的用户兴趣分布下相互竞争的情形。我们构建了 n 个用户簇，其中最大的簇包含半数用户，并令创作者行动集中的每种策略所生成的内容只吸引特定的用户群体。

具体而言，用户群体由互不相交的簇 $\mathcal{X} = \cup_{i=1}^{n}\mathcal{X}_i$ 组成，满足$|\mathcal{X}_1| = \frac{m}{2}$；较小簇的规模 $|\mathcal{X}_l|$ 是在满足$\sum_{l=2}^{n}|\mathcal{X}_l| = \frac{m}{2}$的条件下均匀随机采样得到的。玩家共享相同的行动集 $\mathcal{S}_i = \{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$，且 σ 函数满足：对于任意 $i \in [n]$,

$$\sigma(\boldsymbol{s}_i, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in \mathcal{X}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

数据集 1取决于划分 $\cup_{i=1}^{n}\mathcal{X}_i$的随机性。

• **合成数据集-2**数据集 2模拟了这样一种情况：内容创作者可以通过生成平庸内容来"追逐潮流"，也可以针对特定用户兴趣群体提供高质量内容。与数据集 1的构建方式类似，我们令用户群体由 n 个簇组成，并允许每位玩家采取针对任意特定用户群体的动作。此外，我们还允许每位玩家采取一种"稳妥"动作 $\boldsymbol{s}_0$，生成一些热门内容，在一定程度上满足所有用户的需求 δ。

具体而言，用户群体也由互不相交的簇 $\mathcal{X} = \cup_{i=1}^{n}\mathcal{X}_i$ 给出，所有簇 $|\mathcal{X}_l|$ 的大小均从满足 $\sum_{l=i}^{n}|\mathcal{X}_l| = m$ 的分布中均匀随机采样。玩家共享相同的行动集 $\mathcal{S}_i = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \cdots, \boldsymbol{s}_n\}$，且 σ 函数满足：对于任意 $i \in [n]$,

$$\sigma(\boldsymbol{s}_i, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \in \mathcal{X}_i, i \geq 1 \\ \delta, & \text{if } i = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

数据集- 2 取决于划分 $\cup_{i=1}^{n}\mathcal{X}_i$ 的随机性以及一个参数 $\delta \in [0, 1]$。

• 从 MovieLens-1m 生成的数据集 我们使用深度矩阵分解（Fan 和 Cheng，2018）训练用户嵌入和电影嵌入，以拟合 1 至 5 分的电影评分。用户总数 $m = 6040$、电影数量 $k = 3883$ 以及嵌入维度均设为 $d = 32$。为验证训练所得表示的质量，我们首先进行了 5 折交叉验证，并在测试集上获得了平均 RMSE $= 0.883$，随后使用完整数据集训练用户/物品嵌入。所得的用户嵌入 $\mathcal{X} = \{\boldsymbol{x}_j\}_{j\in[m]}$ 被用作用户群体。为构建每个参与者-i 的行动集 $\mathcal{S}_i$，我们从训练得到的电影嵌入集合 $\mathcal{M}$（$|\mathcal{M}| = 3883$）中独立地随机采样 500 个向量。为将相关性得分归一化为 $[0, 1],$，当电影 $\boldsymbol{s}$ 对用户 $\boldsymbol{x}$ 的预测评分至少为 4 时，即 $\sigma(\boldsymbol{s}, \boldsymbol{x}) = \mathbb{I}[\langle \boldsymbol{s}, \boldsymbol{x}\rangle \geq 4]$，我们令 $\sigma(\boldsymbol{s}, \boldsymbol{x}) = 1$。

5.2 评估指标

我们在实验中同时使用无政府代价和 PotA。评估无政府代价需要求解两个优化问题；由于 $W(\cdot)$ 的非凹性以及 CCE($\mathcal{G}$) 的结构不确定，这两个问题通常都难以求解。因此，当无法精确计算时，我们使用模拟退火来逼近 $\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})$。为了计算

表 1：$\beta = 0.1$下的无政府代价。结果反映了从 10 个独立抽样的博弈实例中得到的最坏情况。

K \ n	$*$	2	3	4	5
1	2.00	1.33	1.54	1.66	1.72
2	1.93	1.28	1.46	1.56	1.60
3	1.89		1.42	1.47	1.51
4	1.86			1.43	1.42
5	1.84				1.42

$*$ 表示理论上界。

表 2：$\beta = 0.1$ 下的 PotA。结果反映了从 10 个独立采样的博弈实例中得到的最坏情况。

K \ n	$*$	5	10	15	20	40
1	2.00	1.59	1.59	1.60	1.50	1.38
3	1.89	1.37	1.39	1.42	1.41	1.32
5	1.84	1.35	1.34	1.33	1.36	1.31
7	1.80		1.30	1.31	1.30	1.29

$*$ 表示理论上界。

$\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$, 对于较小的 n 和中等规模的行动集 k，我们通过求解一个包含k^n 个变量和 kn 个约束的线性规划，计算其精确解（Papadimitriou 和 Roughgarden，2005）。为了处理规模更大的问题，我们让每个参与者在 $T = 5000$ 轮中运行 Exp3（Auer 等人，2002），并计算完全无政府代价 $\text{PotA}(\mathcal{G}) = \max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s}) \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}^t}[W(\boldsymbol{s})]$。更多细节见附录 G。

5.3 结果

模拟得到的实证无政府代价。我们首先展示数据集 1 在不同博弈参数（n、K、β）下的实证福利。我们固定 $\beta \in \{0.1, 0.5\}$，并报告 n 和 K变化时的无政府代价和完全无政府代价。结果见表 1、2、3 和 4。我们观察到，对于任意固定的 n，无政府代价和完全无政府代价均随 K 和 β增加而下降，这与定理1所揭示的结果一致。此外，在（β、K）固定时，随着 n 增加，无政府代价趋近其理论上界。然而，当 n 小于 15 时，完全无政府代价也呈现这一趋势；但随着 n进一步增加，它开始下降。这种差异可能是因为当 n较大时（即表 2 中），近似的最优福利不够准确，因此完全无政府代价往往被低估。

用户参与度/曝光度指标之间的比较。接下来，我们考察采用两种不同激励指标（即用户参与度与用户曝光）所带来的影响。然而，数据集 1 不再是适合此目的的理想基准，因为在简单的二值 $\sigma(\cdot,\cdot)$ 下推导出的效用函数，在这两种指标下几乎没有差别。为此，我们使用数据集 2，它具有更复杂的 $\sigma(\cdot,\cdot)$

表3： $\beta = 0.5$ 下的无政府代价。结果反映了从 10 个独立采样的博弈实例中得到的最坏情况。

K \ n	$*$	2	3	4	5
1	2.00	1.33	1.54	1.66	1.72
2	1.77	1.11	1.24	1.32	1.34
3	1.65		1.08	1.13	1.18
4	1.57			1.05	1.08
5	1.52				1.02

$*$ 表示理论上界。

表4： $\beta = 0.5$下的 PotA。结果反映了从 10 个独立抽样的博弈实例中得到的最坏情况。

K \ n	$*$	5	10	15	20	40
1	2.00	1.59	1.59	1.60	1.50	1.38
3	1.65	1.13	1.20	1.21	1.22	1.20
5	1.52	1.03	1.10	1.12	1.14	1.14
7	1.45		1.05	1.08	1.09	1.11

$*$ 表示理论上界。

用于模拟创作者可能专注于追逐潮流，而不关注内容质量这一情形的函数。

我们固定 $(\beta, K) = (0.1, 2)$，并报告不同 n和 δ下的 PotA。图1所示结果表明，使用用户参与度指标具有优势：在 n 和 δ取不同值时，它始终能带来更小的 PotA。当 n 大于 10 时，采用用户曝光度指标的 PotA 可能超过 2，正如命题 1[5]所示。当 δ 变小时，两种指标之间的性能差距更加明显。这可以理解为：当创作者能够以较低的努力产出热门内容时，仅凭曝光度来奖励创作者可能会严重损害整体用户福利。图2展示了不同 (β, K) 下 PotA 的变化趋势，进一步表明采用用户参与度指标更具优势。值得指出的是，图2表明，随着 β 或K 增加，参与度指标相较于曝光度指标的优势会减弱。这意味着，采用曝光度指标时的 PotA 也会随 K 和 β增加而降低，不过这一结论尚未得到证明。尽管如此，这仍是一个值得进一步研究的问题。

不同理性程度下的社会福利。 在本实验中，我们旨在研究当玩家采用理性程度不同的在线学习算法时，竞争结果会如何变化。为了更好地模拟实践中的情况，我们采用了由 MovieLens-1m（Harper and Konstan，2015）生成的数据集。在模拟中，我们设定每个玩家以不同的探索率运行 Exp3 的情形

5. 同样，由于计算最优 W 时存在近似误差，当 n 增大时，PotA 可能会被低估。

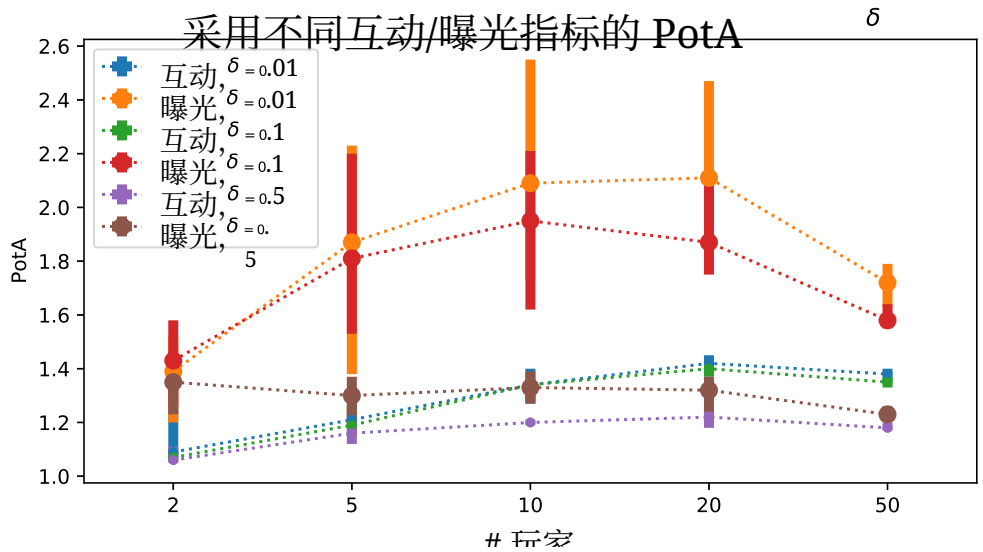

图1：采用曝光/参与度指标时的 PotA，其中 $(\beta,\ K) = (0.1,\ 2)$。δ 是创作者采取“安全”动作时获得的相关性得分。误差条表示 10 次独立试验中的最大值和最小值，圆点表示平均值。

ϵ （即以 ϵ 的概率，每个玩家会在每一轮采取随机动作）。我们使用这一模拟设置来考察一种实际情况，即创作者试图优化其累积遗憾，但具有有限理性：由于已知当 $\epsilon \sim O(\sqrt[k]{\log_T{}^k})$ 时，Exp3 能够取得次线性遗憾（Auer 等人，2002），因此，创作者将 ϵ 设置得过大或过小都不够理性，因为这会导致更大的遗憾 $R(T)$。我们固定 $(\beta,\ K,\ T) = (0.1, 5, 1000)$，并报告在 T 轮中，即$\bar{W} = \frac{1}{mT}\sum_{t=1}^{T} W(\boldsymbol{s}^{(t)})$ 的平均社会福利，结果如图3所示。在 $(\beta,\ K) = \{(0.5, 1)、(0.5, 5)\}$ 下获得的平行结果如图4所示。图3和图4中的曲线均表明，对于$K = 5$，在不同的 n 值下，使福利最大化的最优探索率约为 $\epsilon = 0.1$。当 ϵ 设置得过小或过大时，平均福利都会下降，从而证实了我们在推论 1 中的主张：随着每个玩家学习算法的累积遗憾增加，福利保证会恶化。此外，我们观察到，随着参与创作者数量增加，平均福利也会提高；鉴于平台上的内容池更大，用户获得满意推荐的机会更高，这是符合预期的。进一步地，当玩家数量足够多（$n = 100$）时，即使玩家采用近乎随机化的策略（$\epsilon = 0.9$），福利仍然相当不错。

有趣的是，$K = 1$ 的最优探索率约为 $\epsilon = 0.3$，高于左侧面板图 ϵ中 $K = 5$ 所示的最优4。造成这一差异的一个可能解释是，当推荐系统在推荐列表中提供更多备选项时，用户决策中的随机性可以更好地弥补创作者在探索更多样化策略方面的探索，从而提升社会福利。

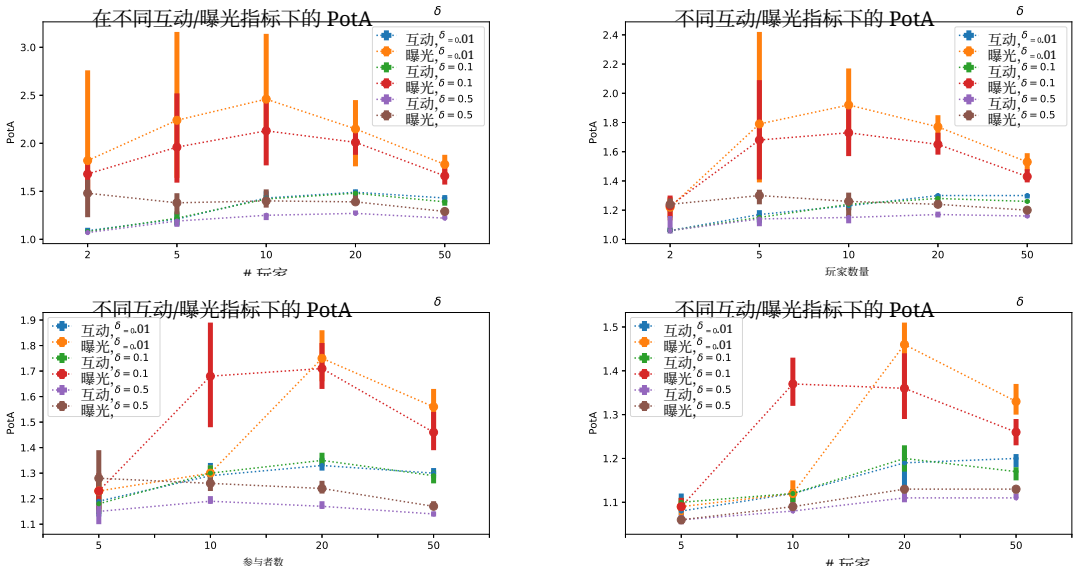

图2：不同（β，K）条件下，用户曝光和用户参与度指标对应的 PotA。δ是创作者采取“安全”动作片时获得的相关性得分。左上：（0.1，1），右上：（0.3，2），左下：（0.1，5），右下：（0.3，5）。误差条表示 10 次独立试验中的最大值和最小值，圆点表示均值。

表 5：MovieLens 数据集在 (β，K) =(0.1，5) 下的 PotA。结果为不同探索因子 ϵ 条件下 10 次独立模拟的平均值。

n \ ϵ	.003	.01	.03	.1	.3	.9
5	1.51	1.32	1.18	1.10	1.12	1.51
10	1.15	1.14	1.08	1.03	1.05	1.25
30	1.06	1.05	1.04	1.02	1.03	1.07
100	1.02	1.02	1.02	1.02	1.02	1.03

MovieLens 数据集上的 PotA 在图 3 和4中采用相同的设置，即计算不同玩家数量和不同探索率下的平均社会福利；我们还在表 5 和 6中给出了相应的 PotA。如结果所示，当创作者能够使用使遗憾最小化的最优探索率时，社会福利高得出乎意料。即便创作者只有五位，福利损失比例也仅约为 10%（PotA=1.10）。随着创作者数量增加，PotA 会迅速趋近理想值 1.0。尽管由于确定最优福利时的近似精度，PotA 可能被低估，但这些结果仍传递出非常乐观的信息。

游玩序列产生的类型分布 除了评估福利指标外，我们还研究了不同探索率下创作者游玩序列所诱导的内容分布，如图5所示。内容分布

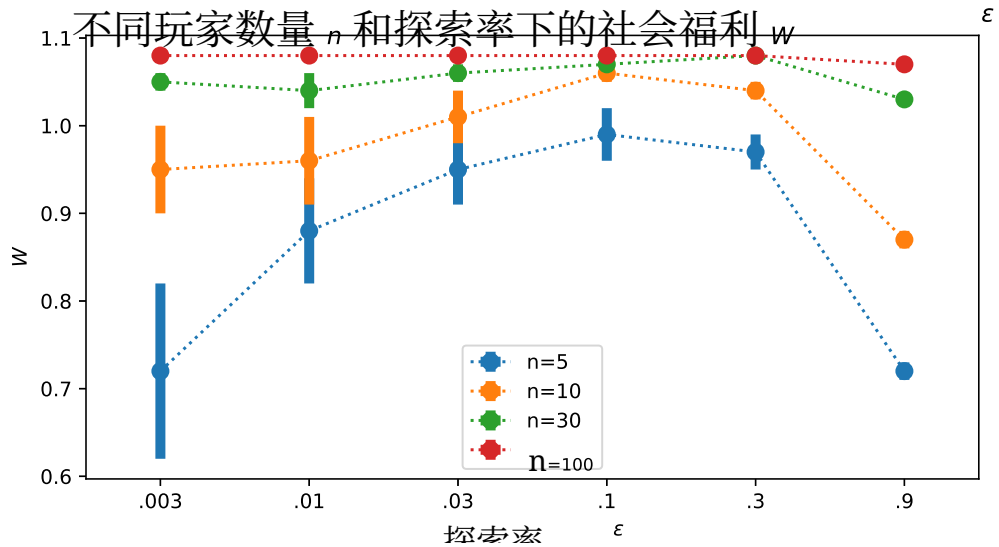

图3：不同探索率ϵ 和玩家数量 n下， $T = 1000$ 轮的平均福利 W 。结果是在（β, K）=(0.1, 5)条件下进行10次独立运行后取平均值。

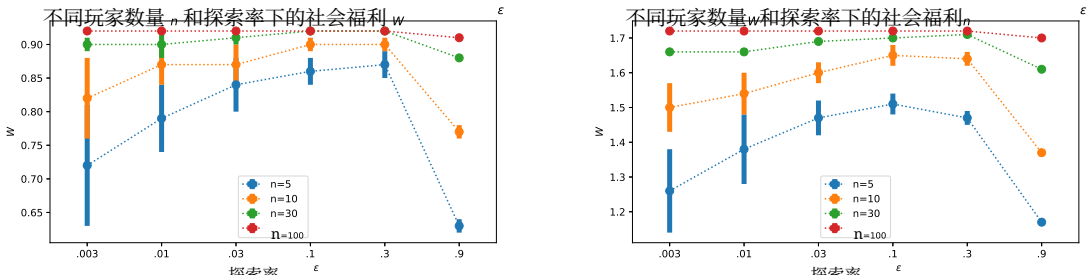

图4：不同探索率ϵ 和玩家数量 n下， $T = 1000$ 轮的平均福利 W 。结果是在（β, K）$\in \{(0.5, 1), (0.5, 5)\}$条件下进行10次独立模拟后取平均值。

通过统计整个仿真轨迹中类型的频率生成（在 MovieLens-1m 数据集中，每部电影都关联了多个类型标签）。具体而言，对于每条具有固定时间范围 $T = 1000$ 的仿真路径，我们收集所有创作者创作的完整内容集合，并总结其相对于类型的频率。我们还计算了不同 n 下最大化社会福利的最优联合策略所对应的类型分布，作为参考。

图 5 左侧面板展示了创作者人数不同时，全局最优内容创作策略的类型分布 $n_\circ$ 随着 n 增加，最优类型分布变得更加“平缓”，这符合预期：希望最大化效用的新创作者应当专注于细分兴趣群体，以实现多样化

表 6：MovieLens 数据集在 $(\beta,\ K)\ =(0.5,\ 5)$ 条件下的 PotA。结果取自探索因子 ϵ 不同时进行的 10 次独立模拟的平均值。

n \ ϵ	.003	.01	.03	.1	.3	.9
5	1.37	1.25	1.18	1.15	1.18	1.48
10	1.15	1.12	1.08	1.05	1.06	1.26
30	1.04	1.04	1.02	1.02	1.01	1.07
100	1.01	1.01	1.01	1.01	1.01	1.02

整体内容分布。右侧三个面板展示的模拟实证类型分布，相较于最优分布呈现出更强的两极化。例如，在所有情况下，创作者创作的剧情类内容（最受欢迎的类型）都多于理想频率。随着创作者侧探索因子的增加，这种两极化效应会有所缓解。有趣的是，如果我们考察不同 n 下固定 ϵ时的实证频率，可以发现，随着 n 增大，内容分布也变得日益多样化，这有利于社会福利。这也解释了表5 和6 中的发现：随着创作者数量增加，在各种 ϵ 下，PotA 都持续减小。这些观察结果揭示了创作者采用遗憾最小化策略所带来的积极影响：即使受到有限理性的制约，由此产生的社会福利仍相当可观。

6. 结论与未来工作

我们提出了竞争性内容创作博弈，这是一种用于分析创作者在在线内容推荐平台上策略行为的博弈论框架。我们的主要贡献是全面刻画创作者之间竞争所带来的社会福利，由此表明，当平台将用户参与度作为激励指标，并为用户提供足够多的选择时，传统的 top-K 推荐原则是有效的。这一结论呼应了亚当·斯密提出的著名“看不见的手”论述。

我们的正向结果依赖于两个关键假设：1. 平台上的用户在做出决策时采用带有 Gumbel 噪声的 RU 模型；2. 平台向创作者效用提供的激励必须与社会福利指标中使用的用户效用（即用户参与度）一致。未来一个有趣且重要的研究方向，或许是放宽其中任一假设。因此，我们提出两个开放性问题：

开放性问题 1： 我们的主要结果在不同的用户选择模型下是否仍然成立？例如，如果允许用户做出多个选择，或者用户采用 RU 模型但使用不同的决策噪声（如高斯分布），会怎样？尽管我们有一些初步的实证证据表明，无政府代价上界仍严格小于 2，但严格的理论分析可能颇具挑战，因为玩家效用函数未必具有闭式表达式，而且我们可能需要新的技术来推导新的光滑性常数。

开放性问题 2： 当参与者激励所采用的用户效用指标与社会福利所采用的指标不一致时，我们对社会福利能得出什么结论？尽管

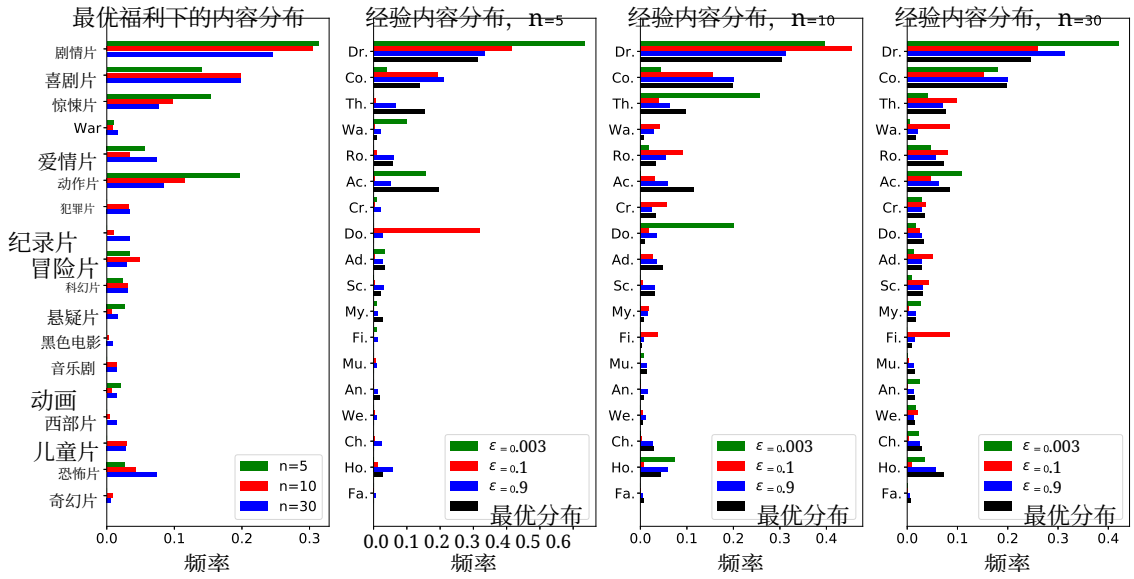

图5：在不同 n下，创作者采用自适应内容创作策略所产生的类型分布。最左侧面板展示了在 $n = 5$，10，30条件下，使社会福利最大化的联合内容创作策略所对应的最优类型分布。右侧其余3个面板从左至右展示了在$n = 5$，10，30条件下通过模拟获得的类型分布。右侧3个面板中的绿色、红色和蓝色分别表示探索率 $\epsilon = 0.003$，0。1，0。9取不同值时的模拟结果，黑色柱表示最优分布。所有频率值均为5次独立运行的平均值。

平台设计这样一致的激励是合理的，但实际原因有时可能会使其难以实现。尽管我们的命题 1 证明了，如果参与者激励与用户参与度不对应，福利可能会变得任意糟糕，但考察介于两者之间的情形仍然很有意义：如果参与者激励与用户参与度相关，却并非完全一致，会发生什么？对这一扩展情形的任何补充理解，都能对现实世界系统产生更深远的影响。

7. 致谢

本研究部分由美国国家科学基金会通过项目编号 IIS- 2007492、IIS-2128019 和 IIS-1838615 资助。Haifeng Xu 部分获得美国陆军研究办公室项目 W911NF-23-1-0030 的资助。

参考文献

Peter Auer、Nicolo Cesa-Bianchi、Yoav Freund 和 Robert E Schapire。《非随机多臂赌博机问题》。《*SIAM* 计算杂志》，32(1):48–77，2002。

Matej Balog、Nilesh Tripuraneni、Zoubin Ghahramani 和 Adrian Weller。Gumbel 技巧失落的亲缘关系。载于国际机器学习会议，第 371–379 页。PMLR，2017年。George Baltas 和 Peter Doyle。营销研究中的随机效用模型：综述。《商业研究杂志》，51(2)：115–125，2001年。Ran Ben Basat、Moshe Tennenholtz 和 Oren Kurland。概率排序原则在对抗性检索场景中并非最优。载于*2015*年国际信息检索理论会议论文集，第 51–60 页，2015。Omer Ben-Porat 和 Moshe Tennenholtz。Shapley 设施选址博弈。载于国际网络与互联网经济学会议，第 58–73 页。Springer，2017年。Omer Ben-Porat 和 Moshe Tennenholtz。面向策略性内容提供商的推荐系统博弈论方法。神经信息处理系统进展，31，2018年。Omer Ben-Porat、Gregory Goren、Itay Rosenberg 和 Moshe Tennenholtz。从推荐系统到设施选址博弈。载于*AAAI* 人工智能会议论文集，第 33 卷，第 1772–1779 页，2019年。Avrim Blum、MohammadTaghi Hajiaghayi、Katrina Ligett 和 Aaron Roth。遗憾最小化与完全无政府代价。载于第四十届 *ACM* 计算理论研讨会论文集，第 373–382 页，2008年。Jes´us Bobadilla、Fernando Ortega、Antonio Hernando 和 Abraham Guti´errez。推荐系统综述。《知识型系统》，46：109–132，2013年。George Christodoulou 和 Elias Koutsoupias。有限拥塞博弈的无政府价格。载于第三十七届 *ACM* 计算理论研讨会论文集，第 67–73 页，2005年。George Christodoulou、Annam´aria Kov´acs 和 Michael Schapira。贝叶斯组合拍卖。载于国际自动机、语言与编程研讨会，第 820–832 页。Springer，2008年。Constantinos Daskalakis、Paul W Goldberg 和 Christos H Papadimitriou。计算纳什均衡的复杂性。《*SIAM* 计算杂志》，39(1)：195–259，2009年。Jicong Fan 和 Jieyu Cheng。通过深度矩阵分解实现矩阵补全。《神经网络》，98：34–41，2018年。Daniel Fleder 和 Kartik Hosanagar。大片文化的下一次兴衰：推荐系统对销售多样性的影响。《管理科学》，55(5)：697–712，2009年。F Maxwell Harper 和 Joseph A Konstan。MovieLens 数据集：历史与背景。《*ACM* 交互式智能系统汇刊》（*TIIS*），5(4)：1–19，2015。

Xiangnan He、Lizi Liao、Hanwang Zhang、Liqiang Nie、Xia Hu 和 Tat-Seng Chua。神经协同过滤。发表于第*26*届国际万维网大会，第 173–182 页，2017年。Harold Hotelling。（1929）：竞争中的稳定性。《经济学杂志》，39(4):57，1929年。Jiri Hron、Karl Krauth、Michael I Jordan、Niki Kilbertus 和 Sarah Dean。对算法策展平台上的内容创作者激励机制进行建模。*arXiv* 预印本 *arXiv:2206.13102*，2022年。Neil Hurley 和 Mi Zhang。Top-N 推荐中的新颖性与多样性——分析与评估。《*ACM*互联网技术汇刊》（*TOIT*），10(4):1–30，2011年。Meena Jagadeesan、Nikhil Garg 和 Jacob Steinhardt。推荐系统中的供给侧均衡。*arXiv* 预印本 *arXiv:2206.13489*，2022年。Joseph A Konstan、Bradley N Miller、David Maltz、Jonathan L Herlocker、Lee R Gordon 和 John Riedl。GroupLens：将协同过滤应用于 Usenet 新闻。《*ACM*通讯》，40(3):77–87，1997年。Yehuda Koren、Robert Bell 和 Chris Volinsky。推荐系统的矩阵分解技术。《计算机》，42(8):30–37，2009年。Elias Koutsoupias 和 Christos Papadimitriou。最坏情形均衡。发表于计算机科学理论方面年度研讨会，第 404–413 页。Springer，1999年。Chun Kai Ling、Fei Fang 和 J Zico Kolter。我们正在进行什么博弈？标准形博弈和扩展式博弈中的端到端学习。*arXiv* 预印本 *arXiv:1805.02777*，2018年。Robert D Luce。个体选择行为：理论分析，纽约，纽约州：John Willey and Sons，1959年。Robert Duncan Luce、Patrick Suppes 等。偏好、效用与主观概率。1965年。Charles F Manski。随机效用模型的结构。《理论与决策》，8(3):229，1977年。Daniel McFadden。城市出行需求的测量。《公共经济学杂志》，3(4):303–328，1974年。Daniel L McFadden。定性响应模型的计量经济学分析。《计量经济学手册》，2:1395–1457，1984年。Richard D McKelvey 和 Thomas R Palfrey。标准形博弈的量化响应均衡。《博弈与经济行为》，10(1):6–38，1995年。Meta。Meta 正在为创作者试行新的变现选项，2022年。
https://www.digitalinformationworld.com/2022/03/meta-is-experimenting-with-new.html。

Dov Monderer 和 Lloyd S Shapley。势博弈。《博弈与经济行为》，14 (1):124–143，1996年。Christos H Papadimitriou 和 Tim Roughgarden。计算多人博弈的均衡。发表于 *SODA*，第 5 卷，第 82–91 页。Citeseer，2005年。Robin L Plackett。排列分析。《皇家统计学会期刊：*C*辑（应用统计）》，24(2):193–202，1975年。Kun Qian 和 Sanjay Jain。数字内容创作：推荐系统影响的分析。可见 *SSRN 4311562*，2022年。Nimrod Raifer、Fiana Raiber、Moshe Tennenholtz 和 Oren Kurland。信息检索与博弈论的交汇：文档作者之间的排名竞争。发表于国际 *ACM SIGIR* 信息检索研究与开发大会，第 465–474 页，2017年。Tim Roughgarden。无政府价格的内在稳健性。《*ACM*学报》（*JACM*），62(5):1–42，2015。Tim Roughgarden、Vasilis Syrgkanis 和 Eva Tardos。拍卖中的无政府价格。《人工智能研究期刊》，59:59–101，2017年。Savy。新的 YouTube 算法会影响你的内容吗？，2019年。https://savyagency.com/new-youtube-algorithm/。Weiran Shen 和 Zihe Wang。具有有限吸引力的霍特林-唐斯模型。*arXiv* 预印本 *arXiv:1611.05959*，2016年。TikTok。什么是 TikTok 创作者基金？以下是如何加入 +并开始赚钱，2022年。https://www.backstage.com/magazine/article/tiktok-creator-fund-explained-how-to-join-75090/。Adrian Vetta。竞争性社会中的纳什均衡及其在设施选址、交通路由和拍卖中的应用。发表于第*43*届 *IEEE* 计算机科学基础年会，*2002*年。会议论文集。，第 416–425 页。IEEE，2002年。YouTube。YouTube 合作伙伴计划概览与资格要求，2023年。https://support.google.com/youtube/answer/72851。

附录A. 定理1的证明

首先，我们推导引理1中给出的竞争性内容创作博弈的效用函数和福利函数的闭式表达式。

A.1 引理1的证明：效用函数和福利函数的闭式表达式

我们先介绍 Gumbel 分布的一些已知且有用的性质。

引理 4 *[例如，（Balog等人，2017年）]设$(v_1, \cdots, v_n) \in \mathbb{R}^n$为任意实值向量，且$\varepsilon_1$、$\cdots$、$\varepsilon_n$为来自 Gumbel 分布的独立样本$(\mu, \beta)$。那么*

$$\arg\max_i (v_i + \varepsilon_i) \sim Categorical\Big(\frac{\exp(\beta^{-1} v_i)}{\sum_{j=1}^n \exp(\beta^{-1} v_j)}\Big), \tag{22}$$

and

$$\max_i (v_i + \varepsilon_i) \sim Gumbel\Big(\mu + \beta \log\Big(\sum_{j=1}^n \exp(\beta^{-1} v_j)\Big), \beta\Big). \tag{23}$$

用户效用和福利的推导。这些推导可由引理4直接得出。由于我们假设 ε_i ~服从 Gumbel 分布（$-\beta\gamma$, β），利用引理4 中的性质，我们得出 $\boldsymbol{x}_j$在 K 备选项 $\{\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K\} = \mathcal{T}_j(\boldsymbol{s};\ K)$上的选择分布遵循 softmax规则

$$\Pr[\boldsymbol{x}_j \to \boldsymbol{s}_i] = \frac{\exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{24}$$

作出选择后的期望用户效用具有以下形式

$$\pi_j(\boldsymbol{x}_j) = \mathbb{E}\left[\max_{i\in[K]}\{\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i\}\right] = \beta \log\left[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))\right]. \tag{25}$$

对所有用户取期望，我们得到以下福利函数

$$W(\boldsymbol{s}) = \sum_{j=1}^m \mathbb{E}\left[\max_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)}\{\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j) + \varepsilon_i\}\right] = \beta \sum_{j=1}^m \log\left[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\right]. \tag{26}$$

令 $\tilde{W}(\boldsymbol{s}) = \beta W(\boldsymbol{s})$ $\tilde{\sigma}(\boldsymbol{s},\ \boldsymbol{x}) = \beta^{-1}\sigma(\boldsymbol{s},\ \boldsymbol{x})$，则有

$$\tilde{W}(\boldsymbol{s}) = \sum_{j=1}^m \log[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\tilde{\sigma}(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)].$$

因此，在对常数 β 进行重新缩放后，不失一般性，可考虑如下形式的评分函数$\sigma \in [0, \frac{1}{\beta}]$、用户效用函数和社会福利函数：

$$\pi_j(\boldsymbol{s}) = \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big], \tag{27}$$

$$W(\boldsymbol{s}) = \sum_{j=1}^{m} \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big]. \tag{28}$$

创作者效用的推导。这最终得出了一个新结果，需要运用非平凡的论证。玩家效用为

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \Pr[\boldsymbol{x}_j \to \boldsymbol{s}_i] \tag{29}$$

$$= \sum_{j=1}^{m} \mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] \cdot \frac{\exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}, \tag{30}$$

根据（30）中的定义，我们需要证明，对于从 Gumbel($-\beta\gamma$，β) 中抽取的独立同分布随机变量 $\{\varepsilon_i\}_{i=1}^{K}$

$$\mathbb{E}[\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j) + \varepsilon_i | \boldsymbol{x}_j \to \boldsymbol{s}_i] = \mathbb{E}[\max_{k\in[K]} \{\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j) + \varepsilon_i\}] = \log\Big[\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp\left(\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j)\right)\Big], \tag{31}$$

也就是说，对于任意的 $(v_1, \cdots, v_K) \in \mathbb{R}^K$ 以及从 Gumbel(0，1) 中抽取的独立同分布随机变量 $\{\varepsilon_i\}_{i=1}^{K}$，

$$\mathbb{E}[v_i + \varepsilon_i | i = \arg\max_{k\in[K]} (v_k + \varepsilon_k)] = \gamma + \log\Big(\sum_{k=1}^{K} \exp(v_k)\Big). \tag{32}$$

令 $Y_i = \max_{k\in[K],\ k\neq i}(v_k + \varepsilon_k) \sim$服从 Gumbel($\log(\sum_{k\neq i} \exp(v_k))$，1)，且 $X_i = v_i + \varepsilon_i \sim$服从 Gumbel($v_i$，1)。则 X_i 的概率密度函数为

$$f_i(x) = \exp(-((x - v_i) + e^{-(x-v_i)})), \tag{33}$$

且 Y 的累积分布函数为

$$F_i(y) = \exp(-e^{-(y-\log(\sum_{k\neq i} \exp(v_k)))}). \tag{34}$$

T因此，我们可以显式计算 X_i 的条件期望，如下所示 :

$$
\begin{aligned}
&\mathbb{E}[v_i+\varepsilon_i | i=\arg\max_{k\in[K]}(v_k+\varepsilon_k)] \\
=&\mathbb{E}[v_i+\varepsilon_i | v_i+\varepsilon_i \geq \max_{k\in[K],k\neq i}(v_k+\varepsilon_k)] \\
=&\mathbb{E}[X|X\geq Y, X\sim \text{Gumbel}(v_i,1), Y\sim \text{Gumbel}(\log(\sum_{k\neq i}\exp(v_k)),1)] \qquad (35)\\
=&\frac{\int_{\mathbb{R}} x f_i(x)F_i(x)dx}{\int_{\mathbb{R}} f_i(x)F_i(x)dx} \\
=&\frac{\int_{\mathbb{R}} x\exp(-((x-v_i)+e^{-(x-v_i)}))\exp(-e^{-(x-\log(\sum_{k\neq i}\exp(v_k)))})dx}{\int_{\mathbb{R}} \exp(-((x-v_i)+e^{-(x-v_i)}))\exp(-e^{-(x-\log(\sum_{k\neq i}\exp(v_k)))})dx} \\
=&\frac{\int_{\mathbb{R}_{\geq 0}} -\ln t\cdot\exp(-t\sum_{k=1}^K\exp(v_k))dt}{\int_{\mathbb{R}_{\geq 0}}\exp(-t\sum_{k=1}^K\exp(v_k))dt} \qquad (36)\\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)+\frac{\int_{\mathbb{R}_{\geq 0}} -\ln s\cdot\exp(-s)ds}{\int_{\mathbb{R}_{\geq 0}}\exp(-s)ds} \qquad (37)\\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)-\frac{d}{d\alpha}\int_{\mathbb{R}_{\geq 0}} s^\alpha e^{-s}ds \\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)-\frac{d}{d\alpha}\Gamma(\alpha+1)\Big|_{\alpha=0} \\
=&\ln\Big(\sum_{k=1}^K\exp(v_k))\Big)+\gamma. \qquad (38)
\end{aligned}
$$

其中（35）由引理4、（36）和（37）通过变量替换 $t=e^{-x}$以及 $s=t\sum_{k=1}^K\exp(v_k)$ 得到，而（38）则来自欧拉–马歇罗尼常数的定义。因此，我们证明了（32），且玩家效用函数具有以下形式

$$
u_i(\boldsymbol{s})=\sum_{j=1}^m\Big(\log\Big[\sum_{\boldsymbol{s}_k\in\mathcal{T}_j(\boldsymbol{s};K)}\exp\left(\sigma(\boldsymbol{s}_k,\boldsymbol{x}_j)\right)\Big]\Big)\frac{\mathbb{I}[\boldsymbol{s}_i\in\mathcal{T}_j(\boldsymbol{s};K)]\exp(\sigma(\boldsymbol{s}_i,\boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k\in\mathcal{T}_j(\boldsymbol{s};K)}\exp(\sigma(\boldsymbol{s}_k,\boldsymbol{x}_j))}. \qquad (39)
$$

A.2 引理2 和3的证明：效用函数和福利函数的性质

我们考虑式 (27)、(39) 和 (28) 中给出的效用函数和福利函数，其中常数 β经过重新缩放，并采用新的假设： $\sigma(\boldsymbol{s},\boldsymbol{x})\in[0,\frac{1}{\beta}],\forall\boldsymbol{s}\in\cup_{i=1}^n\mathcal{S}_i,\boldsymbol{x}\in\mathcal{X}$。为简化后续分析，我们先说明一些有用的记号和约定。对于任意联合策略组合 $\boldsymbol{s}=(\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n)$，我们用大写字母 S表示其集合表示，即 $S=\{\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n\}$。这样，我们可以将式 (1)、(3)、(4)、(5) 中定义的 $\mathcal{T}_j(\boldsymbol{s};K)\,\pi_j(\boldsymbol{s})\,u_i(\boldsymbol{s})\,W(\boldsymbol{s})$视为集合函数 $\mathcal{T}_j(S;K)\,\pi_j(S)\,u_i(S)\,W(S)$。从现在起，我们将根据上下文交替使用集合记号 S 和向量记号 $\boldsymbol{s}$ 。类似地，我们用 S_{-i} 表示集合 $\{\boldsymbol{s}_1,\cdots,\boldsymbol{s}_n\}$ 去掉元素 $\boldsymbol{s}_i$后得到的集合。

此外，我们扩展 $\mathcal{T}_j(S;K)$ 的定义，允许 $|S| = K-1$，具体如下：当 $|S| = K-1$时，令 $\mathcal{T}_j(S;K) = S \cup \{\bar{\boldsymbol{s}}\}$，其中 $\bar{\boldsymbol{s}}$ 是一个默认外部选项，满足对所有 $\boldsymbol{x} \in \mathcal{X}$， $\sigma(\bar{\boldsymbol{s}}, \boldsymbol{x}) = 0$。这一扩展涵盖了系统没有足够的活跃内容创作者分配给用户的情形。当这种情况发生时，系统会将默认选项 $\bar{\boldsymbol{s}}$ 放入 top-K 列表中，但不保证其效用。我们指出，引入这一扩展定义仅仅是为了便于表述，并不影响主要结果的含义。

在证明引理2 和3之前，我们先给出命题2 和引理5中的两个中间结果。命题2 揭示了社会福利的一个相当基本的性质 W，该性质有助于定理1的证明；引理5 则有助于引理3的证明。

命题2 在任意一个有 $n-$ 个参与者的竞争性内容创作博弈 $\mathcal{G}$ 中，固定一个联合策略 $S = \{\boldsymbol{s}_1, ..., \boldsymbol{s}_n\}$。若向博弈中增加一名编号为 $n+1$、纯策略为 $\boldsymbol{s}_{n+1}$ 的参与者，并令 $S' = \{\boldsymbol{s}_1, ..., \boldsymbol{s}_n, \boldsymbol{s}_{n+1}\}$，则社会福利 W 将严格增加，即：

$$W(S') > W(S). \tag{40}$$

证明 根据定义，

$$W(S') = \sum_{j=1}^{m} \left(\log \left[\sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \right] \right), \tag{41}$$

$$W(S) = \sum_{j=1}^{m} \left(\log \left[\sum_{\boldsymbol{s} \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \right] \right), \tag{42}$$

显然，对于任意固定的用户 j，来自 S' 的前 K 个选项的指数得分之和优于来自 $\mathcal{S}$ 的得分之和，即：

$$\sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right) \geq \sum_{\boldsymbol{s} \in \mathcal{T}_j(S';K)} \exp\left(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)\right).$$

因此，式（40）由对数函数的单调性立即成立。 ■

命题2 揭示了现实世界中内容提供商竞争的一项重要而自然的性质：市场中的竞争者越多，用户面临的选择就越多，因此他们的福利总会增加。

引理5如下函数

$$f(x,y) = \frac{(x+1)\log(x+y)}{(x+y)(\log(x+y) - \log y)}, (x,y) \in \mathbb{R}_+ \times \mathbb{N}_+, \tag{43}$$

对于任意 $x \in \mathbb{R}_+$，它关于 y 单调递增；对于任意整数 $y \in \mathbb{N}_+$，它关于 x 单调递减。

证明 我们首先通过直接计算偏导数来证明 $f(\cdot,\ y)$ 的单调性。注意到对于任意 $t \geq 0,\ t \geq \log(1+t)$ 均成立，因此我们有

$$\frac{1}{x+1}\frac{\partial f(x,y)}{\partial y} = \frac{\log(1+\frac{x}{y}) + \log(x+y)[\frac{x}{y} - \log(1+\frac{x}{y})]}{[(x+y)(\log(x+y) - \log y)]^2} > 0, \tag{44}$$

这意味着 $f(x,\ y)$ 随 y 递增。现在只需证明固定 $y = K$ 时关于 x 的单调性，这稍微复杂一些。 $f(x,\ K)$ 关于 x 的导数现在写作

$$\begin{aligned} f'(x,K) &= \frac{(K-1)\log(x+K)\log(1+\frac{x}{K}) - (x+1)\log K}{[(x+K)(\log(x+K) - \log K)]^2} \\ &\triangleq \frac{-g(x,K)}{[(x+K)(\log(x+K) - \log K)]^2}, \end{aligned}$$

and

$$\begin{aligned} g'(x,K) &= \frac{1}{x+K}\Big[(2K+x-1)\log K - 2(K-1)\log(x+K)\Big] && (45) \\ &= \frac{2(K-1)}{x+K}\Big[\frac{2K+x-1}{2(K-1)}\log K - \log(x+K)\Big] \\ &= \frac{2(K-1)}{x+K}\Big[\frac{x+1}{2(K-1)}\log K - \log(1+\frac{x}{K})\Big] && (46) \\ &\geq \frac{2(K-1)}{x+K}\Big[\frac{x+1}{2(K-1)}\log K - \frac{x}{K}\Big] \\ &\geq \frac{x}{x+K}\Big[\log K - \frac{2(K-1)}{K}\Big]. && (47) \end{aligned}$$

我们声称 $g'(x,\ K) \geq 0,\ \forall K \in \mathbb{N}^+$，这是因为

1. 若 $K = 1$，由（45）可得 $g'(x,\ K) = 0$。
2. 对于 $K \geq 5$，可验证 $\log K - \frac{2(K-1)}{K} > 0$。由（47）可得 $g'(x,\ K) > 0$。
3. 对于 $K \in \{2,3,4\}$，可验证 $2^{\frac{x+1}{K-1}}\log K - \log(1+\frac{x}{K}) > 0$ 对任意 $x \geq 0$ 成立。

因此，由（46）可得 $g'(x,\ K) > 0$。

现在，由于 $g'(x,\ K) \geq 0$，我们得出 $g(x,\ K) \geq g(0,\ K) = \log K \geq 0$，这意味着 $f'(x,\ K) \leq 0,\ \forall K \in \mathbb{N}^+$。因此，$f(x,\ K)$ 随 x 递减。 ■

现在我们可以证明引理 2 和3了。

引理 [的证明 2] 根据定义，我们只需证明对于任意 $j \in [m]$，$\pi_j(S)$ 具有子模性，即

$$\pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_x\}; K)) - \pi_j(\mathcal{T}_j(S; K)) \geq \pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_x, \boldsymbol{s}_y\}; K)) - \pi_j(\mathcal{T}_j(S \cup \{\boldsymbol{s}_y\}; K)). \tag{48}$$

不失一般性，我们假设 $\sigma(\boldsymbol{s}_x,\ \boldsymbol{x}_j) \geq \sigma(\boldsymbol{s}_y,\ \boldsymbol{x}_j)$，并令

$$\{v_1, \cdots, v_K\} = \{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S; K)\},$$

其中 $v_1 \leq \cdots \leq v_K$。接下来，根据 $v_x = \exp(\sigma(\boldsymbol{s}_x, \boldsymbol{x}_j))\ v_y = \exp(\sigma(\boldsymbol{s}_y, \boldsymbol{x}_j))$ 和 K 的取值，可分为三种情况：

1. $v_x \le v_1$: (48) 成立，因为其左边和右边均等于 0。
2. $v_x > v_1$，$K = 1$：式（48）的左侧等于 $\log \frac{v_x}{v_1} > 0$，式（48）的右侧等于 0。
3. $v_x > v_1$，$K \ge 2$：式（48）的左侧等于 $\log \frac{v_x+v_2+a}{v_1+v_2+a}$，式（48）的右侧等于 $\log \frac{v_x+v_y+a}{v_y+v_2+a}$，其中 $a = \sum_{k=3}^{K} v_k$ 当 $K \ge 3$ 时，且 $a = 0$ 当 $K = 2$时。我们可以验证

$$\begin{aligned}&(v_x + v_2 + a)(v_y + v_2 + a) - (v_1 + v_2 + a)(v_x + v_y + a)\\ =&(v_2 - v_1)(a + v_1 + v_2) + (v_x - v_1)(v_y - v_1) \ge 0.\end{aligned}$$

因此，式（48）成立，且将式（48）对所有 $j \in [m]$求和即可得到引理 2 。 ■

证明 [**引理 3**]根据定义，

$$u_i(\boldsymbol{s}_i; \boldsymbol{s}_{-i}) = \sum_{j=1}^{m} \left(\log \Big[\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}', \boldsymbol{x}_j)\right) \Big] \right) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(S;K)] \exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp(\sigma(\boldsymbol{s}', \boldsymbol{x}_j))}, \tag{49}$$

and

$$W(S) = \sum_{j=1}^{m} \pi_j(\mathcal{T}_j(S;K)). \tag{50}$$

只需证明对于任意用户 j,

$$\begin{aligned}&\left(\log \Big[\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp\left(\sigma(\boldsymbol{s}', \boldsymbol{x}_j)\right) \Big] \right) \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(S;K)] \exp(\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}' \in \mathcal{T}_j(S;K)} \exp(\sigma(\boldsymbol{s}', \boldsymbol{x}_j))}\\ &\ge c(\beta, K) \cdot \left[\pi_j(\mathcal{T}_j(S;K)) - \pi_j(\mathcal{T}_j(S_{-i};K)) \right].\end{aligned} \tag{51}$$

注意，当 $\boldsymbol{s}_i \notin \mathcal{T}_j(S;\ K)$ 时，(51) 是显然成立的，因为其左端=右端=0。现在我们假设$\boldsymbol{s}_i \in \mathcal{T}_j(S;\ K)$，因此 $\mathcal{T}_j(S_{-i};\ K)$ 与 $\mathcal{T}_j(;\ K)$ 仅在一个元素上不同。不失一般性，令

$$\{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S_{-i};K)\} = \{v_1', v_2, \cdots, v_K\},$$

and

$$\{\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j)) | \boldsymbol{s} \in \mathcal{T}_j(S;K)\} = \{v_1, v_2, \cdots, v_K\}, v_1 \ge v_1'.$$

由于我们对 $\mathcal{T}_j(S;\ K)$ 进行了扩展定义，$\mathcal{T}_j(S_{-i},\ K)$ 在 $K = n$ 时定义良好，在这种情况下，我们有 $v_1' = \exp(\sigma(\bar{\boldsymbol{s}},\ \boldsymbol{x}_j)) = 1$。现在令 $z = v_2 + \cdots + v_K$，（51）等价于

$$\frac{v_1}{v_1 + z} \log(v_1 + z) \ge c(\beta, K) \cdot \log \left[\frac{v_1 + z}{v_1' + z} \right]. \tag{52}$$

由于 $v_1' = \exp(\sigma(\cdot, \boldsymbol{x}_j)) \ge 1$，使 (52) 成立的一个充分条件是

$$\frac{v_1}{v_1 + z} \cdot \frac{\log(v_1 + z)}{\log(v_1 + z) - \log(1 + z)} \ge c(\beta, K). \tag{53}$$

注意， $x = v_1 - 1 \in [0, e^{1/\beta} - 1]$， $y = z + 1 \in [K,(K-1)e^{1/\beta} + 1]$，左侧变为关于 (x, y) 的函数，其形式如下 (53)

$$f(x,y) = \frac{(x+1)\log(x+y)}{(x+y)(\log(x+y) - \log y)}. \tag{54}$$

根据引理5可知， $f(x,y)$ 关于 y 对任意 $x > 0$ 均单调递增，并且关于 x 任意整数 $K \geq 1$ 均单调递减。因此，有

$$f(x,y) \geq f(x,K) \geq f(e^{1/\beta} - 1, K) = c(\beta, K). \tag{55}$$

因此，(51)成立，证明完毕。 ■

A.3 定理1的证明

借助命题2、引理3 和2，现在我们可以证明定理1中的结论。我们将证明，任意竞争性内容创作博弈实例$\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$都是参数为$(\lambda, \mu) = (c(\beta,K)\ c(\beta,K))$的光滑博弈，因此其无政府代价可由 $\frac{1+\mu}{\lambda} = 1 + \frac{1}{c_{(\beta,K)}}$ 上界。</output>

证明 设 $\boldsymbol{s} = (\boldsymbol{s}_1, ..., \boldsymbol{s}_n)$ 和 $\boldsymbol{s}^* = (\boldsymbol{s}_1^*, ..., \boldsymbol{s}_n^*)$ 是两个不同的策略组合。首先，根据引理2中揭示的函数 W的次模性质，对每个 $i \in [n]$ ，都有

$$W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i}) \geq W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}_i^*, \boldsymbol{s}]) - W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}]). \tag{56}$$

对所有参与者 i 求和，我们得到

$$\begin{aligned}\sum_{i=1}^n (W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})) &\geq \sum_{i=1}^n (W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}_i^*, \boldsymbol{s}]) - W([\boldsymbol{s}_1^*, \cdots, \boldsymbol{s}_{i-1}^*, \boldsymbol{s}])) \\ &= W([\boldsymbol{s}^*, \boldsymbol{s}]) - W(\boldsymbol{s}) \\ &> W(\boldsymbol{s}^*) - W(\boldsymbol{s}),\end{aligned} \tag{57}$$

其中，最后一个不等式由命题2得出。另一方面，根据引理3 ，还可得

$$u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) \geq c(\beta, K) \cdot \big[W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})\big], \tag{58}$$

因此

$$\sum_{i=1}^n u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) \geq c(\beta, K) \cdot \sum_{i=1}^n \big[W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}]) - W(\boldsymbol{s}_{-i})\big] \tag{59}$$

$$> c(\beta, K)[W(\boldsymbol{s}^*) - W(\boldsymbol{s})]. \tag{60}$$

其中，不等式（59）由（58）得出，而不等式（60）由（57）得出。

由于（60）对任意 $\boldsymbol{s} \in \mathcal{S}$ 均成立，因此对任意 $\boldsymbol{\alpha} \in CCE$（$\mathcal{G}$），我们可以对 $\boldsymbol{s} \sim \boldsymbol{\alpha}$取期望，得到

$$\sum_{i=1}^n \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i})] > c(\beta, K)[W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]]. \tag{61}$$

因此，

$$\begin{aligned}\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] &= \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[\sum_{i=1}^{n} u_i(\boldsymbol{s})] \\ &\geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[\sum_{i=1}^{n} u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i})] \qquad (62) \\ &\geq c(\beta, K) \cdot \sum_{i=1}^{n} \left[\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W([\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}])] - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s}_{-i})]\right] \\ &> c(\beta, K)[W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]]. \qquad (63)\end{aligned}$$

其中，不等式（62）根据粗相关均衡的定义成立，而不等式（63）由（60）可得。整理各项后，我们得到

$$PoA(\mathcal{G}) = \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]} < 1 + \frac{1}{c(\beta, K)}. \qquad (64)$$

■

A.4 $c(\beta, K)$ 性质的证明

证明 $c(\beta, K)$ 函数具有以下形式：

$$c(\beta, K) = \frac{(b+1)\log(b+K)}{(b+K)(\log(b+K) - \log K)}, b = e^{\frac{1}{\beta}} - 1. \qquad (65)$$

我们逐一证明以下事实。

1. 任取 $\beta > 0$，$c(\beta, K)$ 关于 K 单调递增；类似地，任取 $K \geq 1$，$c(\beta, K)$ 关于 β 单调递增。注意到 $e^{\beta} - 1$ 关于 β 递减，根据引理 5 结论得证。

2. $c(\beta, K) = 1$ 当且仅当 $K = 1$ 或 $\beta \to 0$。当 $K = 1$ 时，$c(\beta, K) = 1$ 直接成立。当 $\beta \to 0$ 时，$b \to +\infty$ 和 $c(\beta, K) \to 1$。“仅当”方向由 $c(\beta, K)$ 的单调性性质得出。

3. 对于任意 $\beta > 0$ 和 $K \geq 1$，我们有 $c(\beta, K) \geq 1$，因此 $PoA(\mathcal{G}) < 2$ 始终成立。根据 c 的单调性，$c(\beta, K) \geq c(\beta, 1) = 1$。因此，$PoA(\mathcal{G}) < 1 + c^{-1}(\beta, K) \leq 2$。

4. 对于充分大的 β 和 K，$c(\beta, K) \approx (1+\beta)$ 渐近地为 $\log K$，因此

$$PoA(\mathcal{G}) < 1 + \frac{1}{(1+\beta)\log K}. \qquad (66)$$

当 β 足够大时，$b = e^{\frac{1}{\beta}} - 1 \approx \frac{1}{\beta} \to 0$。因此，

$$
\begin{aligned}
c(\beta, K) &= \frac{(b+1)\log(b+K)}{(b+K)\log(1+\frac{b}{K})} \\
&\approx \frac{(b+1)K\log(b+K)}{(b+K)b} && \text{since } \log(1+x)\approx x \text{ as } x\to 0 \\
&\approx \frac{(b+1)\log K}{b} && \text{since } K >> b \\
&\approx (1+\beta)\log K. && \text{since } b\approx \tfrac{1}{\beta}
\end{aligned}
$$

■

附录 B. 推论 1 的证明

证明 令 $\epsilon(T) = R(T)/T$，并令 $\boldsymbol{s}^* = (\boldsymbol{s}_1^*, ..., \boldsymbol{s}_n^*)$ 为 $W(\boldsymbol{s})$ 的全局最大值点。根据定义，

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i})] - \epsilon(T). \tag{67}$$

对所有参与者 $i \in [n]$ 求和可得

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] = \sum_{i=1}^{n}\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s})] \geq \sum_{i=1}^{n}\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i})] - n\epsilon(T). \tag{68}$$

另一方面，根据定理 1 的证明中的（60），我们有

$$\sum_{i=1}^{n} u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i}) > c(\beta, K)[W(\boldsymbol{s}^*) - W(\boldsymbol{s})], \forall \boldsymbol{s} \in S. \tag{69}$$

对分布 $\boldsymbol{\alpha}$ 取期望可得 $\boldsymbol{s}$

$$\sum_{i=1}^{n}\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[u_i(\boldsymbol{s}_i^*; \boldsymbol{s}_{-i})] > c(\beta, K)\big(W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]\big). \tag{70}$$

（68）和（70）共同推出

$$\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})] + n\epsilon(T) > c(\beta, K)\big(W(\boldsymbol{s}^*) - \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]\big). \tag{71}$$

注意，对于任意 $\boldsymbol{s} \in \mathcal{S}$，我们有 $W(\boldsymbol{s}) = \sum_{j=1}^{m}\left(\log\left[\sum_{\boldsymbol{s}\in\mathcal{T}_j(\boldsymbol{s};K)}\exp(\sigma(\boldsymbol{s}, \boldsymbol{x}_j))\right]\right) \geq \beta\log K$，因此，$n\epsilon(T) \leq n\epsilon(T)\beta\log K\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$。将其代入（71），即可得到（13）。

■

附录 C. 推论的证明2

证明 注意，固定任意玩家策略组合 $\boldsymbol{s}$，top-K 匹配机制使社会福利 W最大化。因此， $W(\mathcal{G}')\leq W_K^*$。另一方面，根据定理 1中的无政府代价界，可得

$$\begin{aligned}\frac{W_K^*}{W(\mathcal{G})} &< 1+\frac{(b+K)(\log(1+b/K))}{(b+1)\log(b+K)}\\ &< 1+\frac{(b+K)(b/K)}{(b+1)\log(b+K)}\\ &= 1+\frac{b+K}{K\log(b+K)}.\end{aligned}$$

重新整理该项可得 $W(\mathcal{G}')-W(\mathcal{G})\leq W_K^*-W(\mathcal{G})\leq W_K^*/(1+{}^{K}\log(b+K)b+K)$。 ■

附录 D. 定理2的证明

证明 令 $b=\exp(1/\beta)-1$。考虑一个有 n 个参与者的博弈，其中每个参与者-i 具有相同的行动集 $\mathcal{S}_i=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$。设用户群体 $\mathcal{X}$为一个大小为 $m=n+(n-1)a$ 的集合，其中 n 个用户具有策略组合 $\boldsymbol{x}_1$，而 a 个用户具有策略组合 $\boldsymbol{x}_i$，对于 $i=2,\cdots,n$。此处 $a=\beta\log K+1$ 是一个常数，其取值稍后将会明确。设评分函数 σ 为如下定义的示性函数：

$$\sigma(\boldsymbol{s},\boldsymbol{x})=\begin{cases}1, & \text{if } \boldsymbol{s}=\boldsymbol{x},\\ 0, & \text{otherwise.}\end{cases}\tag{72}$$

首先，我们给出最优福利 $\max_{\boldsymbol{s}\in\mathcal{S}}W(\boldsymbol{s})$ 的下界。考虑联合策略组合 $\boldsymbol{s}^*=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$，在该组合下，每个用户都获得一个得分为 1 的参与者，以及一个得分为 0 的参与者。因此，每个用户-j 的效用 $\pi_j(\boldsymbol{s})=\beta\log(b+K)$，而社会福利 $W(\boldsymbol{s})=m\beta\log(b+K)$。因此，最优社会福利

$$\max_{\boldsymbol{s}\in\mathcal{S}}W(\boldsymbol{s})\geq W(\boldsymbol{s}^*)=m\beta\log(b+K).\tag{73}$$

接下来我们证明， $\boldsymbol{s}=(\boldsymbol{x}_1,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_1)$ 是 $\mathcal{G}$ 的纯纳什均衡，因此 $\boldsymbol{s}\in CCE(\mathcal{G})$。给定玩家的联合策略 $\boldsymbol{s}$， n用户将被分配给得分为 1 的 K 名玩家，而 $(n-1)a$用户将被分配给得分为 0 的 K 名玩家。因此，任意参与者-i 的效用为

$$\begin{aligned}u_i(\boldsymbol{s}) &= \Big[n\cdot(\beta\log K+1)+a(n-1)\cdot\beta\log K\Big]/n\\ &= \beta\log K+1+\frac{a(n-1)\beta\log K}{n}.\end{aligned}\tag{74}$$

如果参与者-i 从策略 $\boldsymbol{s}_1$ 切换到 $\boldsymbol{s}_j$， n 用户仍会获得得分为 1 的 K 位参与者， $(n-2)$ a 用户会获得得分为 0 的参与者，而 a用户会获得得分为 $(1,0,\cdots,0)$ 的 K位参与者。因此，参与者-i偏离后的效用为

$$u_i(\boldsymbol{s}_j, \boldsymbol{s}_{-i}) = n \cdot 0 + a(n-2) \cdot \beta \log K \cdot \frac{1}{n} + a \cdot \beta \log(b+K) \cdot \frac{e^{\frac{1}{\beta}}}{e^{\frac{1}{\beta}} + K - 1}$$
$$= \frac{a(n-2)\beta \log K}{n} + \frac{b+1}{b+K} \cdot a\beta \log(b+K).$$

我们可以验证，对于任意 $2 \le j \le n$，若取

$$a = \beta \log K + 1 \le \frac{\beta \log K + 1}{\beta\left(\frac{b+1}{b+K}\log(b+K) - \frac{1}{n}\log K\right)}. \tag{75}$$

这是因为：1. 当 $\beta \in [0,1]$ 时，(75) 中的不等式始终成立，因为 $\frac{b+1}{b+K}\log(b+K) < \log(b+1) = \frac{1}{\beta}$（这是由 $\log x/x$ 的单调性所致）；2. 当 $a = \frac{\beta \log \frac{b+1}{b+K} \log(b+K) - \frac{1}{n}\log K}{\beta \log \frac{b+1}{b+K} \log(b+K) - \frac{1}{n}\log K}$ 时，$u_i(\boldsymbol{s}) = u_i(\boldsymbol{s}_j, \boldsymbol{s}_{-i})$。因此，$\boldsymbol{s}$ 是 $\mathcal{G}$ 的纳什均衡。结合 (73)、(74) 和 (75)，我们得到

$$\begin{aligned} PoA(\mathcal{G}) &= \frac{\max_{\boldsymbol{s}\in\mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})} \mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]} \ge \frac{W(\boldsymbol{s}^*)}{W(\boldsymbol{s})} \\ &= \frac{m\beta \log(b+K)}{n u_i(\boldsymbol{s})} \ge \frac{m}{n u_i(\boldsymbol{s})} \qquad (76) \\ &= \frac{n + (n-1)a}{n(\beta \log K + 1) + a(n-1)\beta \log K} \\ &= \frac{n-1}{n} + \frac{1 - t^2 a(a-1)}{a + ta(a-1)} \quad (t = \frac{n-1}{n}) \\ &> \frac{n-1}{n} + \frac{1}{5a-4} \qquad (77) \\ &= \frac{n-1}{n} + \frac{1}{1 + 5\beta \log K}. \end{aligned}$$

其中，不等式 (76) 成立，是因为 $\beta\log(b+K) \ge \beta\log(b+1) = 1$；而 (77) 成立，是因为当 $a = 1 + \beta\log K \in [1, 1.2]$ and $t = \frac{n-1}{n} \in [0.5, 1)$ 时，很容易验证 $\frac{1-t^2a(a-1)}{a+ta(a-1)} \ge \frac{1}{5a-4}$ 这等价于 $4t^2a + 4 > 5t^2a^2 + ta$，而这是真的，因为 $t^2a(5a-4) + ta < a(5a-4) + a < 4$。 ■

附录 E. 命题的证明1

证明 考虑一个由两个用户 $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ 和 n 个玩家组成的群体，其中一名玩家有两种纯策略，而另一名 $n-1$ 名玩家只能使用一种策略，即 $\mathcal{S}_i = \{\boldsymbol{s}_0\}$，$i = 2, \ldots\ldots, n$ 和 $\mathcal{S}_1 = \{\boldsymbol{s}_1, \boldsymbol{s}_2\}$。令评分函数 σ 为

$$\sigma(\boldsymbol{s}, \boldsymbol{x}) = \begin{cases} 1, & \text{if } \boldsymbol{s} = \boldsymbol{s}_1, \boldsymbol{x} = \boldsymbol{x}_1, \\ \delta, & \text{if } \boldsymbol{s} = \boldsymbol{s}_2, \\ 0, & \text{otherwise.} \end{cases} \tag{78}$$

我们将证明，对于任意给定的 $K \geq 1$，$0 \leq \beta \leq \min\{0.14, \frac{1}{5\log K}\}$，都存在 $\delta \in (0, 1)$，使得博弈 $\mathcal{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ 的无政府代价始终严格大于 2。

根据引理 1 的证明，$\mathcal{G}$ 具有相同的形式，如式（17）、（19），而 $\mathcal{G}$ 的参与者效用函数具有如下形式

$$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \frac{\mathbb{I}[\boldsymbol{s}_i \in \mathcal{T}_j(\boldsymbol{s}; K)] \exp(\beta^{-1}\sigma(\boldsymbol{s}_i, \boldsymbol{x}_j))}{\sum_{\boldsymbol{s}_k \in \mathcal{T}_j(\boldsymbol{s};K)} \exp(\beta^{-1}\sigma(\boldsymbol{s}_k, \boldsymbol{x}_j))}. \tag{79}$$

令 $b = \exp(1/\beta) - 1$，并任取 $\delta \in [\delta_0, 1)$，使得 $\exp(\delta_0/\beta) + K - 1 = \frac{1}{K} + \frac{2}{b+K}$。这样的 $\delta_0 \in (0, 1)$ 必然存在，因为函数 $f(\delta) = \exp(\delta/\beta) + K - 1$ 关于 $[0, 1]$ 单调递增，且值域为 $[K, b+K] \supset \frac{1}{K} + \frac{2}{b+K}$。作出这样的选择 δ 后，我们可以验证

$$\begin{aligned}
2u_1(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0) &= \frac{2\exp(\delta/\beta)}{\exp(\delta/\beta) + K - 1} \\
&\geq \frac{2\exp(\delta_0/\beta)}{\exp(\delta_0/\beta) + K - 1} \\
&= \frac{\exp(1/\beta)}{\exp(1/\beta) + K - 1} + \frac{1}{K} \\
&\geq \frac{\exp(1/\beta)}{\exp(1/\beta) + K - 1} + \frac{1}{n} = 2u_1(\boldsymbol{s}_1, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0),
\end{aligned}$$

这表明 $(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)$ 是 $\mathcal{G}$ 的纯纳什均衡。因此，取 $\delta = \delta_0$，我们有

$$\begin{aligned}
PoA(\tilde{\mathcal{G}}) &= \frac{\max_{\boldsymbol{s} \in \mathcal{S}} W(\boldsymbol{s})}{\min_{\boldsymbol{\alpha} \in CCE(\tilde{\mathcal{G}})} \mathbb{E}_{\boldsymbol{s} \sim \boldsymbol{\alpha}}[W(\boldsymbol{s})]} \geq \frac{W(\boldsymbol{s}_1, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)}{W(\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0)} \\
&= \frac{\log(\exp(1/\beta) + K - 1) + \log K}{2\log(\exp(\delta_0/\beta) + K - 1)} && (80) \\
&= \frac{\log(b+K) + \log K}{2\log[2K(b+K)] - 2\log(b+2K)} && \text{by the choice of } \delta_0 \quad (81) \\
&> 2, && (82)
\end{aligned}$$

因为（82）成立，是因为（82）等价于

$$(b+2K)^4 > 16K^3(b+K)^3. \tag{83}$$

我们通过验证以下情况来证明（83）的正确性：

1. 当 $K \in \{2, 3\}$时，（83）对所有 $\beta \in [0, 0.14]$、$b = \exp(1/\beta) - 1$均成立。
2. 当 $K \geq 4$时，根据$\beta \leq \frac{1}{5\log K}$可知 $b + K = \exp(1/\beta) + K - 1 > K^5$，因此$\frac{(b+2K)^4}{(b+K)^3} \geq b > K^5 \geq 16K^3$。因此，（83）成立。

最后，我们证明，当 $K = 1$ 或 $\beta \to 0$ 时，$PoA(\mathcal{G})$ 可以任意大。

1. 当 $\beta \to 0$ 时，我们有 $b \to \infty$。由（81）可知，对于任意固定的 K，

$$\lim_{\beta\to 0} PoA(\tilde{\mathcal{G}}) = \lim_{b\to+\infty} \left\{ \frac{\log(b+K)+\log K}{2\log[2K(b+K)]-2\log(b+2K)} \right\} = \lim_{b\to+\infty} \log(b+K) \to +\infty.$$

2. 当 $K = 1$ 时，用户的选择是确定性的，因此任意 $\delta \in (0, 1)$ 都会使（$\boldsymbol{s}_2, \boldsymbol{s}_0, \cdots, \boldsymbol{s}_0$）成为 $\mathcal{G}$ 的纯纳什均衡。令 $\delta \to 0$，由（80）可知，对于任意固定的 β，

$$\lim_{\delta\to 0} PoA(\tilde{\mathcal{G}}) = \lim_{\delta\to 0} \left\{ \frac{\log(\exp(1/\beta)+K-1)+\log K}{2\log(\exp(\delta/\beta)+K-1)} \right\} = \lim_{\delta\to 0} \frac{1}{2\delta} \to +\infty.$$

■

附录 F. 与现有模型的联系

作为对相关工作的扩展讨论，我们展示了我们的竞争性内容创作博弈与以下三种先前提出的内容创作者竞争模型之间的联系。以下所有模型均未考虑推荐系统的存在，并将每位用户与所有内容创作者（玩家）匹配，这对应于我们设定中的情形 $K = n$。有趣的是，我们发现，这些模型各自都是我们竞争性内容创作博弈的一个特例。

无干预中介下的设施选址博弈（Ben-Porat 等人，2019年）

考虑以下竞争性内容创作博弈实例：

1. 用户群体 $\mathcal{X} \subseteq [0,1]$ 是一个大小为 m 的有限集，
2. 每位参与者 $i \in [n]$ 共享同一个行动集 $\mathcal{S}_i = [0,1]$，
3. 评分函数定义为 $\sigma(s, x) = |s - x|$，
4. $(\beta, K) = (0, n)$，
5. 效用函数由用户曝光度指标诱导，即 $u_i(\boldsymbol{s}) = \sum_{x\in\mathcal{X}} \Pr(x \to s_i)$。

若令 $m \to \infty$使得 $\mathcal{X}$ 成为单位区间 $[0,1]$，上的、密度函数为 g $[0,1]$，则上述定义的博弈实例 $\tilde{G}(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ [6] 等价于 Ben-Porat 等人(2019年提出的无干预中介下的设施选址博弈。

支持效用函数下具有有限吸引力的霍特林-唐斯模型（Shen 和 Wang，2016年）

考虑以下竞争性内容创作博弈实例：

6. 请注意，我们用 $\tilde{G}$ 指代 G 的一个变体，该变体在玩家效用函数中使用用户曝光度指标，而非用户参与度指标。

1. 用户群体 $\mathcal{X} = \{x_1, \cdots, x_m\} \subseteq [0,1]$ 是一个大小为 m 的有限集,

2. 每个参与者 $i \in [n]$ 共享相同的行动集 $\mathcal{S}_i = [0,1] \times [0,1]$。参与者-$i$采取的每个动作 $\boldsymbol{s}_i = (s_i, w_i)$都对应一个吸引区域 $R_i = [s_i - \frac{w_i}{2}, s_i + \frac{w_i}{2}] \cap [0,1]$。

3. 对于每个 $i \in [n]$，评分函数定义为 $\sigma(\boldsymbol{s}_i, x) = \mathbb{I}[x \in R_i]$,

4. $(\beta, K) = (0, n)$,

5. 效用函数即用户参与度指标，也就是说，$u_i(\boldsymbol{s}) = \sum_{j=1}^{m} \pi_j(\boldsymbol{s}) \Pr(x_j \to s_i)$.

事实上，给定 $\beta = 0$ 以及上述 σ的定义，可以看出，在曝光度指标和参与度指标下，效用函数完全相同，因为 $\pi_j(\boldsymbol{s}) \in \{0,1\}$当且仅当 $\pi_j(\boldsymbol{s}) = 1$当且仅当 $\Pr(x_j \to s_i) > 0$。我们可以验证，上文定义的博弈实例$G(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$等价于 Shen 和 Wang(2016年)提出的基于支持效用的有限吸引力 Hotelling-Downs 模型。

曝光博弈（Hron 等人，2022年）

考虑以下竞争性内容创作博弈实例:

1. 用户群体 $\mathcal{X} \subseteq \mathbb{R}^d$ 是一个大小为 m的有限集,

2. 每个参与者 $i \in [n]$ 都关联一个位于 $\mathbb{R}^d$中的单位球面上的行动集 $\mathcal{S}_i$，即 $\mathcal{S}_i \in \mathbb{S}^{d-1}$,

3. 评分函数由内积给出，即 $\sigma(\boldsymbol{s}, \boldsymbol{x}) = \langle \boldsymbol{s}, \boldsymbol{x} \rangle$,

4. $(\beta, K) = (\tau, n)$,

5. 效用函数由用户曝光度指标诱导，即 $u_i(\boldsymbol{s}) = \sum_{x \in \mathcal{X}} \Pr(x \to s_i)$.

请注意，在曝光博弈中，参数 β 不再表示用户决策噪声，而是变为温度参数 τ，用于控制曝光概率在~各项目之间的分布。上述定义的博弈实例 $G(\{\mathcal{S}_i\}_{i=1}^n, \mathcal{X}, \sigma, \beta, K)$ 等价于 Hron 等人提出的曝光博弈（Hron 等人(2022年）。

附录 G. 详细实验设置

G.1 全局最优社会福利的计算

当精确解在计算上不可行时，我们采用两种启发式方法来计算 $\max_{\boldsymbol{s} \in \mathcal{S}} W(\boldsymbol{s})$：算法1 所示的模拟退火（SA），以及算法2所示的最佳响应搜索（BRS）。对于 SA，我们设置 $T = 5000$，并采用温度调度$\tau_t = 0.1/\sqrt{t}$；对于 BRS，我们设置 $T = \max\{30, 2n\}$，并从 5 次独立运行中选取最佳输出。当玩家数量 n 和行动集大小 k 均小于 6 时，两种方法的表现与暴力搜索相当，因此在这些情况下均能得到精确的全局最优解。对于规模更大的问题，我们运行这两种方法，并选取最佳输出，以近似计算 $\max_{\boldsymbol{s} \in \mathcal{S}} W(\boldsymbol{s})$。

算法 1 用于计算最优福利的模拟退火

输入：时间范围 T，联合行动空间 $\mathcal{S}=\prod_{i=1}^{n}\mathcal{S}_i$，福利函数 $W(\boldsymbol{s})$，温度调度表 $\{\tau_t\}_{t=1}^{T}$。初始化：随机选择一个联合动作 $\boldsymbol{s}^{(0)}=(\boldsymbol{s}_1^{(0)},\cdots,\boldsymbol{s}_n^{(0)})\in\mathcal{S}$。对于 $t=0$ 到 T，执行：随机选择一名参与者 $i\in[n]$，并在 $\boldsymbol{s}^{(t)}$ 中随机扰动其动作 $\boldsymbol{s}_i^{(t)}$，得到 $\boldsymbol{s}'^{(t)}$。计算 $W(\boldsymbol{s}'^{(t)})$ $W(\boldsymbol{s}^{(t)})$。如果 $W(\boldsymbol{s}'^{(t)})>W(\boldsymbol{s}^{(t)})$，则设置 $\boldsymbol{s}^{(t+1)}=\boldsymbol{s}'^{(t)}$。否则，以 $e^{(W(\boldsymbol{s}'^{(t)})-W(\boldsymbol{s}^{(t)}))/\tau_t}$的概率，设置 $\boldsymbol{s}^{(t+1)}=\boldsymbol{s}'^{(t)}$；否则，$\boldsymbol{s}^{(t+1)}=\boldsymbol{s}^{(t)}$。输出：$\max_{t\in[T]}W(\boldsymbol{s}^{(t)})$。

算法2 用于计算最优福利的最佳响应搜索

输入： 时间范围 T、联合行动空间 $\mathcal{S}=\prod_{i=1}^{n}\mathcal{S}_i$、福利函数 $W(\boldsymbol{s})$。**初始化：** 随机选择一个联合动作 $\boldsymbol{s}^{(0)}=(\boldsymbol{s}_1^{(0)},\cdots,\boldsymbol{s}_n^{(0)})\in\mathcal{S}$。**对于** $t=0$ **到** T，**执行**随机选择一个参与者 $i\in[n]$，并搜索使 W 最大化的最佳响应，即设置 $\boldsymbol{s}^{(t+1)}=(\boldsymbol{s}_i^{(t+1)},\boldsymbol{s}_{-i}^{(t)})$。**输出：** $W(\boldsymbol{s}^{(T)})$。

$$\boldsymbol{s}_i^{(t+1)}=\arg\max_{\boldsymbol{s}_i\in\mathcal{S}_i}W(\boldsymbol{s}_i,\boldsymbol{s}_{-i}^{(t)}).$$

G.2 粗相关均衡下最坏情况福利的计算

我们可以将粗相关均衡的定义表示为一组线性约束。令 $\boldsymbol{\alpha}\in$粗相关均衡$(\mathcal{G})$ 表示联合行动空间 $\mathcal{S}=\prod_{i=1}^{n}\mathcal{S}_i$上的一个概率分布。那么，$\boldsymbol{\alpha}$ 满足

$$\forall i,\forall\boldsymbol{s}_i',\sum_{\boldsymbol{s}\in\mathcal{S}}\boldsymbol{\alpha}(\boldsymbol{s})u_i(\boldsymbol{s})\geq\sum_{\boldsymbol{s}\in\mathcal{S}}\boldsymbol{\alpha}(\boldsymbol{s})u_i(\boldsymbol{s}_i',\boldsymbol{s}_{-i}).\tag{84}$$

并且 $\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})}\mathbb{E}_{\boldsymbol{s}\sim\boldsymbol{\alpha}}[W(\boldsymbol{s})]$ 正在求解

$$\min_{\boldsymbol{\alpha}\in CCE(\mathcal{G})}\boldsymbol{\alpha}(\boldsymbol{s})W(\boldsymbol{s})\tag{85}$$

在线性约束（84）下。假设每个 $\mathcal{S}_i$ 具有相同的大小 k，那么我们将得到一个包含 k^n 个变量和 kn 个约束的线性规划。

G.3 T 无悔动态模拟和 PotA 计算的详细信息

为了计算 PotA，我们需要模拟每位玩家的游玩序列。我们让所有玩家同时运行以下 Exp3算法 3 来更新各自的策略

在固定时间范围内 $T = 5000$。根据（Auer 等人，2002），算法 3 在 $\eta = \epsilon \sim O(\sqrt{k \log T / T})$ 的条件下具有次线性遗憾。然而，假设现实中的内容创作者足够老练，能够确定博弈参数以及最优学习率/探索率，并不现实。因此，除非另有说明，我们在实验中始终使用固定值（η，ϵ）$= (0.1, 0.1)$。

算法 3 Exp3：模拟参与者-i的演化策略

输入： 时间范围 T，动作数量 k，探索参数 ϵ，学习率 η。**初始化：** 得分向量 $\boldsymbol{y}_0 = (y_1(0) \cdots, y_k(0)) = (0, \cdots, 0)$。**对于** $t = 0$ **到** T **执行**

根据累计得分计算混合策略：

$$p_j(t) = (1-\epsilon)\frac{\exp(y_j(t))}{\sum_{l\in[k]}\exp(y_l(t))} + \frac{\epsilon}{k}, \quad \forall j \in [k].$$

根据分布 $\boldsymbol{p}_t = (p_1(t)\ p_2(t)\ \ldots p_k(t))$ 随机抽取动作片 $\boldsymbol{s}_{i,\,t} \in [k]$。执行动作片 $\boldsymbol{s}_{i,\,t}$ 并观察效用 $u_i(\boldsymbol{s}_{i,\,t}, \boldsymbol{s}_{-i,\,t})$。更新得分

$$y_{\boldsymbol{s}_{i,t}}(t+1) = y_{\boldsymbol{s}_{i,t}}(t) + \eta \cdot \frac{u_i(\boldsymbol{s}_{i,t}, \boldsymbol{s}_{-i,t})}{p_{\boldsymbol{s}_{i,t}}(t)}.$$