arXiv:2305.04453v2 [cs.DS] 9 May 2023

具有可控处理时间的在线任务分配

Ruoyu Wu* Wei Bao† Liming Ge†

摘要

我们研究一种新的在线分配问题，称为 *Online Task Assignment with ControllableProcessing Time*。在一个二分图中，在已知对抗性分布（KAD）假设下，应将一组在线顶点（任务）分配给一组离线顶点（机器）。我们首次研究了这一场景下的可控处理时间：每个任务有多个处理级别，级别越高，效用越大，但处理延迟也越长。机器可以拒绝分配，但需付出拒绝惩罚，该惩罚从预先设定的拒绝预算中扣除。不同的处理级别会产生不同的惩罚。我们提出在线机器与级别分配（OMLA）算法，同时为每个在线任务分配一台离线机器和一个处理级别。我们证明，如果每台机器的拒绝预算无限，OMLA 可达到 $1/2$ 竞争比；如果每台机器的初始拒绝预算最多为 Δ，则可达到 $\Delta/(3\Delta-1)$ 竞争比。有趣的是，在可控处理时间的不同设置下，竞争比保持不变，因此我们可以得出结论：OMLA 对可控处理时间“不敏感”。

1 引言

本文研究带有可控处理时间的在线任务分配问题。在该问题中，我们有一组在线顶点（任务）和一组离线顶点（机器）。在线任务依次到达，每个任务都可以由一台机器处理，但每个任务只能由一部分机器处理 [7, 31]。我们关注可控处理时间：每个任务有多个处理时间级别 [37, 30]。如果任务以更高级别处理，则可获得更高的奖励，但需要等待更长时间。每台机器一次只能处理一个任务，并且只有在前一个任务完成后才能处理另一个任务。机器并不总是接受任务分配；每次拒绝时，都会消耗一定的拒绝预算。被拒绝的分配的处理级别越高，消耗的预算就越多。当机器的拒绝预算耗尽时，它会立即且永久地从系统中移除。本文考虑在已知对抗性分布(KAD) [7, 34] 下到达的在线任务。每个任务在每个时刻的到达概率都是已知的。目标是在不了解未来任务到达情况的条件下（在线设定），最大化期望奖励。

尽管已有许多研究对在线分配问题进行了研究，但本文首次考虑了可控处理时间。这一研究动机源于各个领域的现实场景，例如：

1. 边缘计算中的任务卸载 [24]。借助边缘计算，终端用户设备可以将计算密集型任务卸载到边缘计算机，尤其是处理机器学习（ML）模型的边缘计算机。每个用户（任务）附近只有一组边缘计算机（机器），因此该用户的任务只能由这些计算机处理。当我们为任务分配一台边缘计算机时，还可以通过采用不同的机器学习模型进一步控制处理时间。更好的处理质量需要处理时间更长的模型，而轻量级模型（例如剪枝与稀疏化模型）完成处理的速度更快，但准确率较低。

2. 拼车及收费 [4]。拼车系统将乘客请求（任务）分配给可用的司机（机器）。每个请求都有一个起点，且只能由该区域附近的司机处理。For

*计算机科学学院，悉尼大学，ruwu6940@uni.sydney.edu.au
†计算机科学学院，悉尼大学，{wei.bao, liming.ge}@sydney.edu.au

对于每次行程，我们可以选择走收费公路（成本较高）以缩短处理时间，也可以避开收费公路以降低成本。

3. 翻译服务 [33]。客户向语言服务机构（机器）下达订单（任务），机构可以提供不同级别的服务（例如，仅翻译、翻译并校对等）。每个翻译请求都有一个目标语言，且只能由持有该语言认证的译员处理。机构可以为翻译任务分配更多时间，从而提高处理质量；也可以为任务分配更短的处理时间，以便处理更多客户请求。

在许多情况下，由于只有在任务到达时才能获知其到达情况，因此需要使用在线算法。我们的目标是设计一种在线算法，使其相对于离线最优性能的最坏情况比率（竞争比）最大化。简而言之，本文研究的在线分配问题具有以下特点：

A. 已知对抗性分布（KAD）： 预先已知每个任务在每个时刻到达的概率。**B. 可复用机器：** 机器完成任务后会返回系统；处理延迟取自已知分布。**C. 可控处理时间：** 每个任务都可以按不同的处理级别进行处理；较高的处理级别会带来较高的奖励，但预期延迟也会更长。**D. 有预算限制的机器拒绝：** 机器被分配任务后，可以接受惩罚并拒绝该分配；拒绝处理级别较高的任务将受到更高的惩罚。预算耗尽后，机器将被永久移出系统。

本文的主要贡献是针对上述问题设计一种在线机器与级别分配（OMLA）算法，尤其适用于具有多个处理级别的情形。我们证明，当每台机器的拒绝预算无限时，该算法可达到 $1/2$ 的竞争比；当每台机器的拒绝预算有限时，可达到 $\Delta/(3\Delta-1)$ 的竞争比，其中 Δ是初始时机器的最大拒绝预算。结论表明，无论拒绝预算是否有限，竞争比都不依赖于处理级别，这说明 OMLA 对可控处理时间不敏感。

可控处理时间使本文研究的问题更加贴近现实，但也给在线算法设计和竞争分析带来了更多挑战。可控处理时间拓展了搜索空间，因为不同的处理级别会带来不同的奖励、延迟和拒绝预算。由于目标与约束相互耦合，算法需要在这些维度之间进行权衡。为应对这一挑战，在设计在线算法时，我们首先将选择机器和处理级别的联合概率作为决策变量，构建一个离线线性规划（LP）。随后，利用该线性规划的最优解，计算每台机器在每个处理级别下的激活值和基准值。这两个值将决定我们在线做出决策时对机器和处理级别的选择。为界定竞争比，我们引入一个参考系统，将具有L个处理级别的每个任务重构为 L 个只有一个处理级别的任务。然后，将参考系统的性能作为中间值，用于界定竞争比。数学推导表明，多个处理级别不会使竞争比变差，因为参考系统能够对不同的处理级别进行统一界定，因此竞争比对可控处理时间不敏感。

2 相关工作

与本文相关的一类研究是在线二分图匹配，其中系统需要将离线机器分配给在线任务，以最大化效用 [25]。一类子领域的研究关注对抗性到达顺序 [15], , 另一类则假设已知对抗性分布（KAD） [21, 1]或已知同分布独立分布（KIID） [30], , 即任务到达遵循已知分布。受现实场景启发， [8] 和 [7] 研究了机器可重复使用的情形， [13, 27, 4]研究了机器可以拒绝任务分配的情形。 [31]同时研究了机器可重复使用和拒绝分配的情形。

该领域研究的其他主题包括任务分配的公平性 [28, 23], 、多单元需求（任务可能需要多台机器处理） [9, 12], 以及多容量代理（机器可以处理多个任务） [3, 20]。然而，现有研究尚未考虑在线二分图匹配中的可控处理时间。 [31]的大部分内容可视为本文未考虑可控处理时间时的一个特例。当每台离线机器可以无限次拒绝任务时，该研究给出了一种 1/2-竞争算法；当每台机器最多可以拒绝 Δ次时，则给出了一种 $\Delta/(3\Delta - 1)$-竞争算法。有趣的是，我们提出的算法也能达到相同的竞争比，但其设计和分析要复杂得多。由此，本文得出的一个关键结论是，竞争比对处理级别“并不敏感”。请注意，另一项研究 [11]探讨了可控奖励和不同的到达概率；在该研究中，分配会影响奖励和到达概率，这与可控处理时间在本质上有所不同。除了机器分配，在线二分图匹配还被用于解决许多现实问题，例如拼车 [21, 7, 28]、众包 [9, 19, 11] 和 AdWords [26]。然而，现有研究仍未考虑可控处理时间。

与本文相关的另一类研究是可控处理时间。可控处理时间是在调度问题的背景下进行研究的。我们可以缩短作业的处理时间，但代价是降低处理奖励 [30, 32]。 [14], [5] 和 [10] 研究单机调度。 [2] 和 [29] 研究多台并行机器的调度。 [6] 采用二分图匹配来分析多台并行机器的调度。上述研究关注离线调度问题。 [22] 和 [36]研究可控处理时间下的在线调度。 [22] 关注单机调度， [36] 则关注流水车间调度。随机批量决策问题的研究也分析了可控处理时间。我们可以通过增加额外成本来压缩生产时间，从而获得更好的规划效果 [35, 16, 17]。这些研究以离线方式优化性能。据我们所知，尚无现有研究探讨二分图匹配问题中的可控处理时间。

3 模型

本节对我们的问题进行形式化描述。我们有一个二分图 $G = (U, V; E)$，其中 U 是机器集合， V 是可重复执行的任务集合。[1] E 是边的集合，用于表示任务 $v \in V$是否可以由机器 $u \in U$处理。对于每个任务，我们有 L 个处理级别 $\mathcal{L} = \{1, ..., L\}$，表示其质量层级。如果任务 v 由机器 u $((u, v) \in E)$以处理级别 l处理，则会产生奖励 $r_{u,v,l}$。不失一般性（WLOG），当 $l < l'$ 时，我们有 $r_{u,v,l} < r_{u,v,l'}$ （较高的处理级别会带来较高的奖励）。系统的运行时间范围有限，为 $T \in \mathbb{N}^+$。每个处理级别 $l \in \mathcal{L}$ 都会产生随机处理延迟 d_l，表示以级别 l处理任务所需的占用时间。换言之，如果机器开始以级别 l处理任务，则直到时间 $t + d_l$之前都无法处理其他任务。 d_l 服从已知分布 D_l。当 $l < l'$ 时，我们有 $\mathbb{E}[d_l] < \mathbb{E}[d_{l'}]$ （较高的处理级别需要更长的处理时间）。为方便起见，对于所有任务 v （$E_v = \{(u, v) | (u, v) \in E\}$），我们将与任务 v 相连的边集合记为 E_v；类似地， E_u 是与机器 u相连的边集合。

在每个时间 t，任务 v 可能以概率 $p_{v,t}$到达。以概率 $1 - \sum_{v \in V} p_{v,t}$，在 t时没有任务到达。概率分布集合 $\{\{p_{v,t}\}_{v \in V}\}_{t \in [T]}$ 提前为我们所知（随时间变化，但时间之间相互独立）。当任务 v 到达时，我们立即且不可撤销地选择将其分配给一个与 v 相邻且可用的机器，或丢弃 v。此外，当我们将任务 v 分配给机器 u时，还会指定一个处理级别 l。机器 u 收到任务 v的分配后，有两种可能的操作：以概率 q_e， u 接受该分配；以概率 $1 - q_e$， u 拒绝该分配（$e = (u, v)$）。假设该分配指定了处理级别 l，这两种操作会产生不同的结果。如果u 接受该分配，它会立即获得奖励 $r_{u,v,l}$，并在一段随机时长d_l内变得不可用。机器 u 拥有有限的拒绝预算（初始化为 Δ_u）。如果 u 拒绝该分配，则会产生拒绝惩罚 θ_l。我们假设 Δ_u 和 θ_l 是整数。该惩罚从剩余的

[1]请注意， $v \in V$ 实际上表示一类会在不同时间重复到达的任务。为方便表述并遵循惯例，本文也将 v 称为“任务”。

拒绝预算为 u，记作 δ_u。为方便起见，我们记 $\theta = \max_{l\in\mathcal{L}} \theta_l$。当一台机器 u 耗尽其剩余预算（$\delta_u \le 0$）时，它会立即且永久地从系统中移除。如果一台机器u 被移出系统，则不会再收到任务分配。

每台机器都有一个正的初始预算（$\Delta_u > 0$）。请注意，我们允许使用 $\Delta_u = \infty$ 表示无限制拒绝。我们也允许使用 $\theta_l = 1$ 表示同质拒绝惩罚（以限制拒绝次数）。

3.1 解法概述

我们的目标是最大化奖励总和。我们关注在线设定：只有在任务到达时，我们才知道该任务的到达情况。我们预先知道任务到达的分布以及占用时间的分布（KAD）。

我们首先构建线性规划（LP）Off，以获得一个最优解$\mathbf{x}^*$ 以及离线最优值的一个上界，该上界称为 LP(Off)。随后，我们利用最优解$\mathbf{x}^*$ 构建在线算法。同时，离线最优值的上界 LP(Off) 将作为基准，用于评估在线算法，从而衡量在线算法的表现与离线最优值之间的竞争比。

3.2 离线最优值与竞争比

我们将原问题的离线优化版本作为基准，并定义竞争比。在离线设置中，完整的任务序列 I 已知。然而，只有当分配发生时，我们才知道机器会接受还是拒绝该分配。我们只知道接受的概率 q_{e}。给定完整的任务序列 I，如果某个离线算法使期望奖励最大化，那么它就是 I 的离线最优算法。对于 I，这个最大化后的期望奖励记为 OPT(I)。每个序列 I 上的期望 OPT(I) 为 $\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$, ，称为离线最优值。

如果在线算法 ALG 获得的期望奖励至少为离线最优值的 α 倍，则称 ALG 是 α-竞争的。也就是说，对于任意 $\mathcal{I}$，都有 $\mathbb{E}_{I\sim\mathcal{I}}[\text{ALG}(I)] \ge \alpha\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$ 。

3.3 线性规划

离线最优值 $\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$ 并不容易量化。下面我们构建一个离线线性规划，以得到离线最优值的上界。

$$
\begin{aligned}
&\max_{\substack{x_{e,l,t},\\ \forall e\in E,\\ l\in\mathcal{L},\\ t\in[T]}} \sum_{t\in[T]}\sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} x_{e,l,t} \\
&\text{s.t.} \sum_{t'<t}\sum_{e\in E_u} q_e \sum_{l\in\mathcal{L}} x_{e,l,t'} \Pr\{d_l \ge t-t'+1\} + \sum_{e\in E_u} q_e \sum_{l\in\mathcal{L}} x_{e,l,t} \le 1, \quad (\forall u\in U, t\in[T]), && (1)\\
&\quad \sum_{t\in[T]}\sum_{e\in E_u}\sum_{l\in\mathcal{L}} x_{e,l,t}\left[\theta q_e \Pr\{d_l > T-t\} + (1-q_e)\theta_l\right] \le \Delta_u + \theta - 1, (\forall u\in U), && (2)\\
&\quad 0 \le \sum_{e\in E_v}\sum_{l\in\mathcal{L}} x_{e,l,t} \le p_{v,t}, \quad (\forall v\in V, t\in[T]), && (3)\\
&\quad 0 \le \sum_{l\in\mathcal{L}} x_{e,l,t} \le p_{v,t}, (\forall v\in V, e\in E_v, t\in[T]), && (4)\\
&\quad 0 \le \sum_{e\in E_u}\sum_{l\in\mathcal{L}} x_{e,l,t} \le 1, \quad (\forall u\in U, t\in[T]), && (5)
\end{aligned}
$$

该线性规划称为 Off。Off 的最优解为 $\mathbf{x}^* \overset{\text{def}}{=} \{x^*_{e,l,t}\}$，Off 的目标函数最优值称为 LP(Off)。LP(Off) 是 $\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$ 的上界（如下文引理所示）。此外，$\mathbf{x}^*$ 将用于在线算法中。

算法1 在线机器与级别分配（OMLA）算法

输入： U， V， E， $\{Q^{\delta}_{e,l,t}\}$, $\{R^{\delta}_{u,t}\}$, $\mathbf{x}^*$

1: **对于**所有 $t \leftarrow 1$ 到 T **执行**
2: **如果**没有任务到达**则**
3: skip
4: **否则(**v 到达**)**
5: 以概率 x^* 选择配对 $(u,\ l)$ $_{e,l,t}/p_{v,t}$
6: **如果** u 未被占用， u 具有正的剩余预算 $\delta_u > 0$ 且$Q^{\delta}_{e,\ l,\ t} \geq R^{\delta}_{u,\ t+1}$ **则**
7: 我们将 $(u,\ l)$ 分配给 v
8: **如果** u接受**则**
9: 从 d_l 中抽取 D_l， u 被占用 d_l
10: **else**
11: $\delta_u \leftarrow \delta_u - \theta_l$
12: **结束 if**
13: **结束 if**
14: **结束 if**
15: **结束 for**

在下面的引理中，我们证明 LP(Off) 是离线最优值 $\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$ 的有效上界。

引理 1. *(LP(Off)*上界*) LP(Off)* $\geq \mathbb{E}_{I\sim\mathcal{I}}[OPT(I)]$

证明。 所有引理和定理的证明见附录。 □

3.4 在线机器与级别分配（OMLA）算法

本节中，我们利用 Off 的最优解来构造 OMLA 算法。

设计概述。 在线算法首先确定当任务 v 在时间 t 到达时，选择机器-层级对（u， l）的概率。然后，通过比较不同决策所带来的（机器 u 的）不同期望奖励，来决定是否将机器 u 以处理级别 l 分配给任务 v 。我们在算法1中给出在线算法（OMLA），并逐行进行讨论。

OMLA。任务 v到达时，我们首先以概率$x^*_{e,l,t}/p_{v,t}$ 选择一台机器 u 和一个处理级别 l （算法1第5行）。假设 u 在 t时的剩余拒绝预算为 δ 。如果 u 的拒绝预算（$\delta \leq 0$）已用尽，或正被先前任务占用，我们就不将 u 分配给 v ，也不将任何其他机器分配给 v （算法1第6行）。接着，我们将 u 在 t 的基准值（R值）定义为：在不知道 t到达情况时， u 在 t 及之后的期望总奖励，记作$R^{\delta}_{u,t}$；我们将 u 在 t 的激活值（Q值）定义为：如果将 u 以处理级别 l分配给 v ，则 u 在 t 及之后的期望总奖励，记作$Q^{\delta}_{e,l,t}$。稍后将详细说明基准值和激活值的推导过程。我们将比较基准值和激活值，以作出决策。我们比较 t 的激活值与 $t+1$的基准值，以判断在 t 采取激活动作（进行分配）是否有利。如果 u 在 $t+1$ 的基准值更大，我们就不分配 u ，并丢弃 v; 反之，如果 t处的激活值更大，我们就以 l 将 u 分配给 v （算法1第7行）。当 u接受这项分配时，它将在随机时长 d_l 内处于占用状态（算法1第9行）。当 u 拒绝这项分配时，将从 u的拒绝预算 δ 中扣除拒绝惩罚 θ_l （算法1第11行）。

激活值和基线值的计算。 由于我们关注 KAD 模型，因此可以预先计算每个激活值和基准值（在执行算法1之前）。计算过程如算法2所示。当 u 的剩余预算（$\delta > 0$）为正时，激活值$Q^{\delta}_{e,\ l,\ t}$ 由两部分组成： 1◯ 以概率 q_e $(e = (u,\ v))$， u 接受分配并立即获得奖励$r_{u,\ v,\ l}$ (6)。经过随机占用时间 d_l后， u 完成该任务，其基准值变为$R^{\delta}_{u,\ t+d_l}$（算法2第25–27行）； 2◯ 以概率 $1-q_e$， u拒绝该分配

算法2 激活与基准值的计算 es

输入： $U, V, E, \mathcal{L}, T, \{\theta_l\}, \{q_e\}, \{r_{u,v,l}\}, \{D_l\}, \{\Delta_u\}$1: 求解 LP(Off)，得到 $\mathbf{x}^*$2: $\Delta \leftarrow$ 最大值 Δ_u3: 对于所有满足（δ, u, v, l）且（u, v）$\in E, \delta > 0$ 并且 $l \in \mathcal{L}$ 的情况，执行 4: $Q^{\delta}_{e,l,T} \leftarrow q_e r_{u,v,l}$5: 结束循环 6: 对于所有满足 $\delta > 0$ 且 $u \in U$ 的（δ, u），执行 7: $a \leftarrow 0$8: 对于所有满足（u, v）$\in E_u$ 的 v，执行 9: $b \leftarrow 0$10: 对于所有 $l \in \mathcal{L}$，执行 11: $b \leftarrow b + x^*_{e,l,T} q_e r_{u,v,l}$12: 结束循环 13: $a \leftarrow a + b$14: 结束循环 15: $R^{\delta}_{u,T} \leftarrow a$16: 结束循环 17: 对于 $t \leftarrow T-1$ 到 1，执行 18: 对于所有满足 $\delta > 0$ 且 $u \in U$ 的（δ, u），执行 19: 对于所有满足（u, v）$\in E_u$ 且 $l \in \mathcal{L}$ 的（v, l），执行 20: $a \leftarrow 0$21: 如果 $\delta \geq \theta_l$，则 22: $b \leftarrow R^{\delta-\theta_l}_{u,t+1}$23: 否则 $b \leftarrow 0$24: 结束条件判断 25: 对于 $d \leftarrow 1$ 到 $T-t+1$，执行 26: $a \leftarrow a + R^{\delta}_{u,t+d} \Pr\{d_l = d\}$27: 结束循环 28: $Q^{\delta}_{e,l,t} \leftarrow q_e (r_{u,v,l} + a) + (1 - q_e) b$29: 结束循环 30: 根据（7）计算 $R^{\delta}_{u,t}$ 31: 结束循环 32: 结束循环 输出： $\{Q^{\delta}_{e,l,t}\}, \{R^{\delta}_{u,t}\}, \mathbf{x}^*$

并从其剩余预算 δ中扣除拒绝惩罚 θ_l，其基准值变为 $R^{\delta-\theta_l}_{u,\ t+1}$ 在 $t+1$处（算法2第21–23行）。结合上述两部分，我们可以计算激活值 $Q^{\delta}_{e,\ l,\ t}$如下

$$Q^{\delta}_{e,l,t} = q_e (r_{u,v,l} + \sum_{d' \in [T-t]} \Pr\{d_l = d'\} R^{\delta}_{u,t+d'}) + (1 - q_e) R^{\delta-\theta_l}_{u,t+1}, \quad (\delta > 0, t \in [T]). \tag{6}$$

公式 (6) 在算法2的第 28 行计算。当 u 用尽拒绝预算（$\delta \leq 0$）时，会将其从市场中移除。我们将 $Q^{\delta}_{e,l,\ t} = 0$ 设为边界值，条件是 $\delta \leq 0$ 或 $t > T$。由于我们选择 t 处的激活值与 $t+1$ 处的基准值中较高的一个，因此所选 u 的期望奖励为 $\max\{Q^{\delta}_{e,l,t}, R^{\delta}_{u,t+1}\}$。由于 v 在 t 到达且选择了（u, l）的概率为 $x^*_{e,l,t}$，我们可以通过以下方式计算每个基准值 $R^{\delta}_{u,t}$（算法2的第 30 行）】【。

$$R^{\delta}_{u,t} = \sum_{e \in E_u} \sum_{l \in \mathcal{L}} x^*_{e,l,t} \max\{Q^{\delta}_{e,l,t}, R^{\delta}_{u,t+1}\} + (1 - \sum_{e \in E_u} \sum_{l \in \mathcal{L}} x^*_{e,l,t}) R^{\delta}_{u,t+1}, (\delta > 0, t \in [T]). \tag{7}$$

我们将 $R^{\delta}_{u,\ t}=0$ 将 $\delta \le 0$ 或 $t>T$ 设为边界值。

为了执行我们的在线算法，我们需要预先计算每个 $Q^{\delta}_{e,l,t}$ 和 $R^{\delta}_{u,t}$ 。这可以通过 T 处的初始条件完成（算法2第 3–16 行）

$$\begin{cases} Q^{\delta}_{e,l,T}=q_e r_{u,v,l}, & \forall e\in E, l\in\mathcal{L}, \delta>0, \\ R^{\delta}_{u,T}=\sum\limits_{e\in E_u}\sum\limits_{l\in\mathcal{L}} x^{*}_{e,l,T} q_e r_{u,v,l}, & \forall u\in U, \delta>0. \end{cases} \tag{8}$$

我们在算法2中给出了计算每个 $Q^{\delta}_{e,l,t}$ 和 $R^{\delta}_{u,t}$ 的算法。

4 竞争比分析

本节分析我们在线算法的竞争比。图 1 展示了分析思路。我们已经推导出引理 1，其中给出了离线最优值的上界 LP(Off)。接下来，我们首先介绍参考系统，它给出了原始系统的下界（引理2）。在该参考系统中，每个具有 L 个层级的任务都被重构为 L 个仅有一个层级的任务。引理2之后，我们分别从两个分支分析竞争比： 1○ 每台机器 u都有无限初始拒绝预算 $\Delta_u=\infty$ （无限拒绝情形）； 2○ 每台机器 u 都有有限拒绝预算 $\Delta_u<\infty$。对于无限拒绝情形，我们首先求出参考系统的下界（引理3）。然后利用引理 1–3，为无限拒绝情形构造辅助不等式（引理 4）。接着根据引理 4，我们证明 OMLA 在无限拒绝情形下是 $1/2$-竞争的（定理 1）。对于有限拒绝情形，我们首先求出参考系统的性能归纳不等式（引理 5）。利用该不等式，我们求出参考系统的下界（引理 6）。然后利用引理 1、2 和 6，为有限拒绝情形构造辅助不等式（引理 7）。根据引理 7，我们证明 OMLA 在有限拒绝情形下是 $\Delta/(3\Delta-1)$-竞争的，其中$\Delta=\max_{u\in U}\Delta_u$ （定理 2）。

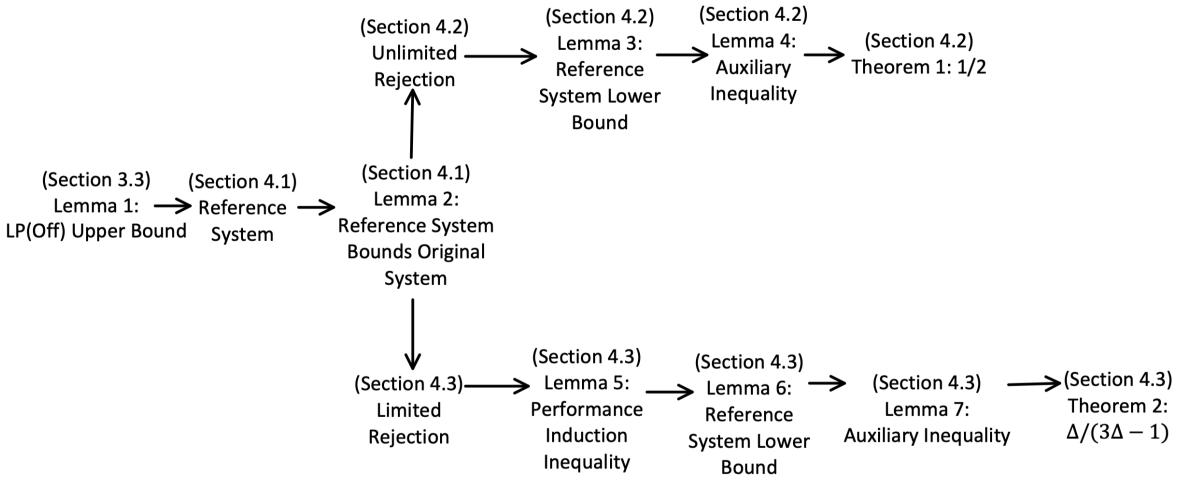


图 1：分析示意图。

4.1 参考系统

直接分析原始系统的性能并不容易，因此我们需要找到一个中间值来对其进行界定。原问题的一个关键挑战来自不同的处理级别。在可控处理时间下， L 处理级别构成了一个额外维度。参考系统旨在构造另一个不含这一额外维度的二部匹配系统，以

为原始三部匹配系统提供一个下界。我们还需要证明：1）参考系统是一个有效系统；2）该参考系统的期望奖励是, 的一个有效下界。

对于每台机器 u，我们按如下方式构造其参考系统 u。该参考系统包含一个二分图 $G'_u = (U'_u, V'_u; E'_u,)$，其中 U'_u 仅包含一台机器 u， V'_u 是一组不可重复任务。在每个 t，可能到达 L 个任务中的一个，记作 $v'_{u,l,t}$，其概率为 $p'_{u,l,t}$。 V'_u 中的每个 $L \times T$ 任务各不相同。$v'_{u,l,t}$ 只能以处理级别 l进行处理。每个 $v'_{u,l,t}$ 都与 u 相连。如果任务 $v'_{u,l,t}$ 在 t 到达，且 u 可用，我们必须选择 u，并决定是否将 u 分配给 $v'_{u,l,t}$。假设在 t时，u的剩余预算为 δ。类似于基准值和激活值（第3.4节），我们将参考基准值（以及参考激活值）定义为 $\tilde{R}^{\delta}_{u,t}$（以及 $\tilde{Q}^{\delta}_{u,l,t}$）。参考基准值和参考激活值的计算方式与基准值和激活值类似，稍后将给出。我们比较 t 时 u 的参考激活值与 $t+1$时 u 的参考基准值。较高者决定我们的选择：如果前者较大，我们就将 u 分配给 $v'_{u,l,t}$；否则，丢弃 $v'_{u,l,t}$。 u 接受此分配的概率为 $q'_{u,l,t}$。处理 $v'_{u,l,t}$ 的奖励为 $r'_{u,l,t}$。处理时间 d'_l 取自已知分布 D_l。拒绝惩罚为 θ_l，与原始系统相同。 u 的初始预算为 Δ_u，与原始系统相同。我们在参考系统中按如下方式设置参数 $p'_{u,l,t}$, $q'_{u,l,t}$以及 $r'_{u,l,t}$ ：$v'_{u,l,t}$ 在 t到达的概率：$p'_{u,l,t} = \sum_{e \in E_u} x^*_{e,l,t}$； u 接受任务 $v'_{u,l,t}$的分配的概率：$q'_{u,l,t} = (\sum_{e \in E_u} q_e x^*_{e,l,t})/p'_{u,l,t}$，若 $p'_{u,l,t} > 0$；否则为 $q'_{u,l,t} = 0$；以级别 l处理 $v'_{u,l,t}$ 的奖励：$r'_{u,l,t} = (\sum_{e \in E_u} q_e r_{u,v,l} x^*_{e,l,t})/(p'_{u,l,t} q'_{u,l,t})$，若 $p'_{u,l,t} q'_{u,l,t} > 0$；否则为 $r'_{u,l,t} = 0$。处理级别 l 的占用时间分布为 $\mathbf{Pr}\{d'_l = d\} = \mathbf{Pr}\{d_l = d\}$.

利用上述参数，我们可以通过以下方式计算 $\tilde{Q}^{\delta}_{u,l,t}$ 和 $\tilde{R}^{\delta}_{u,t}$

$$\tilde{R}^{\delta}_{u,t} = \sum_{l \in \mathcal{L}} p'_{u,l,t} \max\{\tilde{Q}^{\delta}_{u,l,t}, \tilde{R}^{\delta}_{u,t+1}\} + \Big(1 - \sum_{l \in \mathcal{L}} p'_{u,l,t}\Big) \tilde{R}^{\delta}_{u,t+1}, \quad (\delta > 0, t \in [T]), \tag{9}$$

$$\tilde{Q}^{\delta}_{u,l,t} = q'_{u,l,t} \Big(r'_{u,l,t} + \sum_{d' \in [T-t]} \Pr\{d_l = d'\} \tilde{R}^{\delta}_{u,t+d'} \Big) + (1 - q'_{u,l,t}) \tilde{R}^{\delta-\theta_l}_{u,t+1}, \quad (\delta > 0, t \in [T]). \tag{10}$$

若 $\delta \le 0$ 或 $t > T$，我们令 $\tilde{R}^{\delta}_{u,t} = 0$ 和 $\tilde{Q}^{\delta}_{u,l,t} = 0$。我们注意到，无需计算 $\tilde{Q}^{\delta}_{u,l,t}$ 和 $\tilde{R}^{\delta}_{u,t}$ 的具体值（不会引入计算复杂度）。我们只需在分析过程中使用上述值。我们注意到，对于每个 u，参考系统都是有效系统，因为每个 $p'_{u,l,t}$ 和 $q'_{u,l,t}$ 都是有效的概率值。由 (5) 可得 $p'_{u,l,t} \le 1$ 和 $\sum_l p'_{u,l,t} \le 1$。由 (1) 可得 $q'_{u,l,t} \le 1$。因此，$p'_{u,l,t}$ 和 $q'_{u,l,t}$ 都是有效的概率值，并且每台机器 u 的参考系统都是有效系统。

在引理2中，我们证明，对于每个 u、 t 和 δ，参考系统 $\tilde{R}^{\delta}_{u,t}$ 的性能是原始系统 u 中的 $R^{\delta}_{u,t}$性能的下界。

引理2。（参考系统为原始系统提供下界）$R^{\delta}_{u,t} \ge \tilde{R}^{\delta}_{u,t}$、 $\forall \delta$ 和 t。

根据引理2，我们可以将 $\tilde{R}^{\delta}_{u,t}$ 的下界作为 $R^{\delta}_{u,t}$的有效下界。基于该参考系统，我们先分析每台机器的初始拒绝预算为无限时的竞争比，再分析每台机器的初始拒绝预算不超过 Δ 时的竞争比，其中 $\Delta = \max_{u \in U} \Delta_u$。

4.2 无限拒绝情形

本节分析无限拒绝情形下的竞争比。在无限拒绝情形下，每个 u 都有无限的初始拒绝预算 $\Delta_u = \infty$。一种直接的方法是在运行算法1和算法2时，让 Δ_u 足够大。更高效的方法是将$R^{\delta}_{u,t}$, $Q^{\delta}_{e,l,t}$, $\tilde{R}^{\delta}_{u,t}$和$\tilde{Q}^{\delta}_{u,l,t}$ （$\forall \delta$）分别替换为$R_{u,t}$, $Q_{e,l,t}$, $\tilde{R}_{u,t}$以及$\tilde{Q}_{u,l,t}$，因为它们在不同的 δ下并无差别。略作修改后的算法1和算法2见附录。

本节的主要结果是，算法1在无限拒绝情形下是 $1/2$-竞争的（定理 1）。为得到这一结果，我们首先给出参考系统性能的下界（引理3），然后构造一个辅助不等式（引理 4），以证明竞争比的下界可以通过 $\sum_u R_{u,1}$ 的下界与离线最优值LP(Off)之比得到。最后，根据引理 4，我们得到定理 1中的结果。

我们首先在引理3中证明，我们对$R_{u,1}$ 有一个下界。$R_{u,1}$ 是 u在参考系统中从 $t=1$ 时刻起（含该时刻）获得的期望总奖励（即 u的整体期望总奖励）。

引理3。（参考系统下界）对于每个 u，我们有

$$\tilde{R}_{u,1} \geq \frac{1}{2} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{11}$$

为证明引理3，关键步骤是建立一个对偶线性规划，以推导出该下界。我们还利用了“最大值之和不小于和的最大值”这一性质。接下来，我们通过引理1–3证明无限拒绝情形下的辅助不等式。

引理 4。（辅助不等式）对于原始系统，在无限拒绝情形下，我们有

$$\frac{\mathbb{E}_{I\sim\mathcal{I}}[ALG(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[OPT(I)]} \geq \frac{\frac{1}{2}\sum_{u\in U}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{LP(\mathit{Off})}. \tag{12}$$

接下来我们引入定理 1。我们证明，对于无限拒绝情形，算法1是 $1/2$-竞争的。

定理 1.（无限制拒绝下的竞争比）对于拒绝预算无限制的问题，*OMLA* 是 $1/2$-竞争的。

4.3 有限拒绝情形

本节分析有限拒绝情形下的竞争比。每个 u 都有一个有限的初始拒绝预算 $\Delta_u < \infty$。本小节的主要结果是，算法1在有限拒绝情形下是 $\Delta/(3\Delta-1)$-竞争的，其中 $\Delta = \max_{u\in U} \Delta_u$（定理 2）。为得到这一结果，我们首先给出性能归纳不等式（引理 5），并用它推导参考系统性能的下界（引理 6）。随后，我们构造一个辅助不等式（引理 7），以证明竞争比的下界可由 $\sum_u \tilde{R}^{\Delta_u}_{u,1}$ 的下界与离线最优值 LP(Off) 之比得到。最后，根据引理 7，我们得出定理 2 中的结果。

我们首先证明关于 $\tilde{R}^{\delta-\theta_l}_{u,t}$和 $\tilde{R}^{\delta}_{u,t}$ 的一个不等式，用于获得参考系统性能的下界。请注意，由于现在需要考虑剩余预算 δ，这个不等式是针对有限拒绝情形新提出的。

引理 5.（性能归纳不等式）对于所有 δ、t 和 l，我们有

$$\tilde{R}^{\delta-\theta_l}_{u,t} \geq \frac{\delta-\theta_l}{\delta} \tilde{R}^{\delta}_{u,t}. \tag{13}$$

接下来，我们证明 $\tilde{R}^{\Delta_u}_{u,1}$ 有一个下界。

引理 6。（参考系统下界）对于每个 u，我们有

$$\tilde{R}^{\Delta_u}_{u,1} \geq \frac{\Delta_u}{3\Delta_u - 1} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{14}$$

关键步骤是建立一个对偶线性规划来推导界。根据引理 6，我们可以从（6）中消除不同层级的不同拒绝惩罚值（非齐次 θ_l）的影响。在证明中，只需利用对偶线性规划推导一个界，而对偶线性规划可以消除（2）中非齐次 θ_l 的影响。

接下来, 根据引理 1、2 和 6，我们证明有限拒绝情形下的辅助不等式 .

引理 7。*（ 辅助不等式）对于原始系统，在有限拒绝情形下，我们有*

$$\frac{\mathbb{E}_{I\sim\mathcal{I}}[ALG(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[OPT(I)]} \geq \frac{\sum_{u\in U}\frac{\Delta_u}{3\Delta_u-1}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{LP(Off)}. \tag{15}$$

然后我们介绍定理 2。我们证明，在有限拒绝情形下，OMLA 是 $\Delta/(3\Delta-1)$-竞争的，其中 $\Delta=\max_{u\in U}\Delta_u$。

定理 2。 *在有限拒绝情形下，OMLA 是一个 $\Delta/(3\Delta-1)$-竞争算法，其中 $\Delta=\max_{u\in U}\Delta_u$。*

5 评估

5.1 基准算法

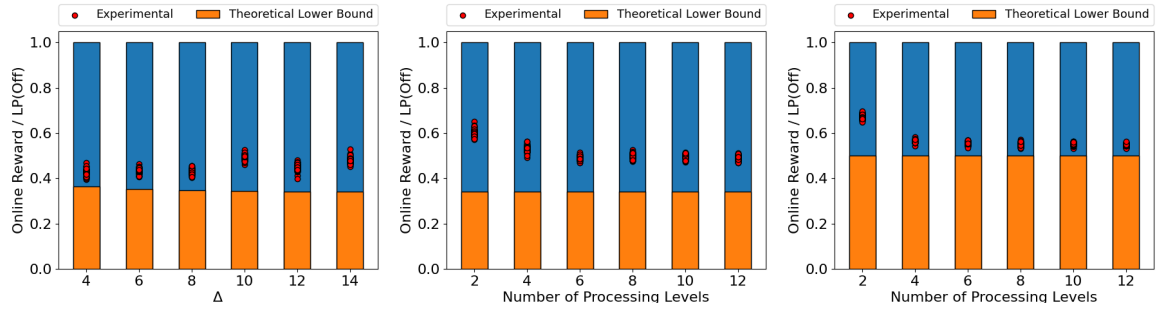
本节将 OMLA 与五种基准算法进行比较：1）随机算法（R）：在线任务到达时，系统随机选择一个机器-层级对。如果所选机器可用，则分配该机器-层级对；否则，丢弃该任务。2）效用贪心算法（UG）：任务 v 到达时，系统根据 $r_{u,v,l}$ 对所有机器-层级对进行排序，并选择其中最高且可用的对。3）效率贪心算法（EG）：预先计算处理级别 l 的占用时间期望值 $\mathbb{E}[d_l]$；任务 v 到达时，系统根据 $r_{u,v,l}/\mathbb{E}[d_l]$ 对所有机器-层级对进行排序，并选择其中最高且可用的对。4）效用贪心算法 + (UG+)：任务 v 到达时，我们根据 [31], 选择机器 u，然后选择具有最高 $r_{u,v,l}$的级别 l。5）效率贪心算法 + (EG+)：任务 v 到达时，我们根据 [31], 选择机器 u，然后选择具有最高 $r_{u,v,l}/\mathbb{E}[d_l]$的级别 l。请注意，[31] 未考虑处理级别。我们采用 [31] 中的方法选择机器，然后在 UG+ 和 EG+中使用贪心方法选择级别。

5.2 合成数据集

我们在实验中生成合成数据集。（ [31] 中也采用了这种方法。）我们设置$|U|=10$、$|V|=25$和 $T=100$。对于每个 u和 v，在 E中存在一条边（u，v），其概率为 0。1。对于每个 $e\in E$，我们设置 $q_e\sim U$(0。5，1) 和 $r_{u,v,l}\sim U$ $(a\cdot l^{0.2},\ a\cdot l^{0.4})$，其中 $a\sim U$(0。5，1)。对于每个 $l\in\mathcal{L}$，我们将分布 D_l设为二项分布 $B(T,\ l^{1.2}/20)$。对于图 2 中的 (a)、(b) 以及图 3 中的 (a)、(b) 这些设置，Δ_u 从 [Δ]中均匀抽取。我们将级别 l 的拒绝惩罚设为 $\theta_l=l+2$。

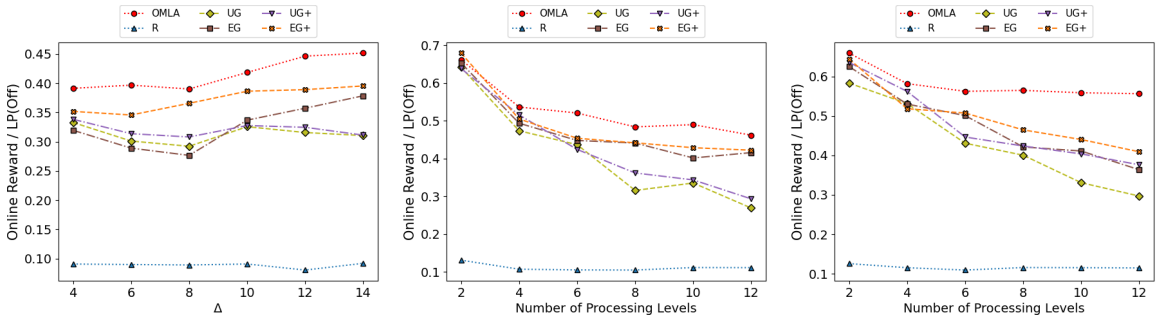
5.3 竞争比验证

在图 2 中，我们研究了 OMLA 与 LP(Off) 在线性能之间的比值。对于每个子图，我们随机生成一组 $\{p_{v,t}\}$、E 和 $\{q_e\}$。对于每一对 Δ 和 L，我们生成一组 $\{\Delta_u\}$和 $\{r_{u,v,l}\}$，然后进行 50 轮实验。在每轮实验中，我们从 $\{p_{v,t}\}$中随机生成 50 个任务序列，并计算 OMLA 的平均总奖励与 LP(Off) 之间的比值。图 2 中的橙色柱表示每一对 Δ 和 L对应的 OMLA 竞争比。图 2 中的红点表示平均在线总奖励与 LP(Off) 之间的比值。图 2 表明，OMLA 的平均性能比竞争比的理论下界确实更高。图 2 中的结果验证了我们关于竞争比的结论。



(a) L=12 (b) Δ=20 (c) Δ=∞

图 2：不同 L 和 Δ 的在线奖励 / LP(Off)。



(a) L=12 (b) Δ=20 (c) Δ=∞

图 3：不同基准算法在不同 L and Δ 下的在线奖励 / LP(Off)。

5.4 OMLA 与基准算法的比较

在图 3 中，我们将 OMLA 的性能与基准算法进行比较。对于每个子图，我们随机生成一组 $\{p_{v,t}\}$、E 和 $\{q_e\}$ 。对于每一对 Δ 和 L，我们生成一组 $\{\Delta_u\}$ 和 $\{r_{u,v,l}\}$。然后，对于每种算法，我们从 $\{p_{v,t}\}$中随机生成 250 个任务序列，并计算平均总奖励与 LP(Off) 之间的比值。对于每一对 Δ 和 L，OMLA 的性能都优于所有基准算法。在图 3b 和图 3c 中，当处理级别数量为 2 时，OMLA 的性能略高于 UG+ 和 EG+的性能；但当处理级别更多时，性能增益会进一步扩大。这表明，处理级别越多，OMLA 的优势越明显，因为它是为机器和级别的联合分配而设计的。结果表明，在可控处理时间下，OMLA 具有最佳性能。OMLA 既提供了理论性能保证（竞争比），又在合成数据集上取得了最佳平均性能。

6 结论

本文研究具有可控处理时间的在线二分图匹配问题，该问题受到现实场景的启发。我们设计了 OMLA，这是一种在线算法，可同时为每个在线任务分配一台离线机器和一个处理级别。我们证明，若每台机器的拒绝预算无限，OMLA 可实现 $1/2$ 竞争比；若每台机器的初始预算至多为 Δ，则可实现 $\Delta/(3\Delta-1)$ 竞争比。此外，我们在合成数据集上进行了实验，结果表明，在各种环境下，OMLA 的表现均优于基准算法。

致谢

参考文献

[1] S. Alaei、M. Hajiaghayi 和 V. Liaghat。具有应用于广告分配的在线先知不等式匹配。第 18–35 页，1993 年。[2] B. Alidaee 和 A. Ahmadian。涉及可控作业加工时间的双并行机器排序问题。《欧洲运筹学杂志》，70(3):335–341，1993 年。ISSN 0377-2217。[3] J. Alonso-Mora、S. Samaranayake、A. Wallar、E. Frazzoli 和 D. Rus。通过动态行程—车辆分配提供按需高容量拼车服务。《美国国家科学院院刊》，114(3): 462–467，2017 年。ISSN 0027-8424。[4] I. G. A. Andani、L. L. P. Puello 和 K. Geurs。在印度尼西亚雅加达—万隆地区，建模分析出行时间及收费公路使用成本变化对不同人口群体的居住地、路线和出行方式选择的影响。《交通研究*A*辑：政策与实践》，145: 81–102，2021 年。ISSN 0965-8564。[5] Z.-L. Chen、Q. Lu 和 G. Tang。具有离散可控加工时间的单机调度。《运筹学快报》，21(2):69–76，1997 年。ISSN 0167-6377。[6] T. C. E. Cheng、Z. L. Chen 和 C.-L. Li。具有可控处理时间的并行机调度。《*IIE*汇刊》，28(2):177–180，1996 年。ISSN 0740-817X。[7] J. P. Dickerson、K. A. Sankararaman、A. Srinivasan 和 P. Xu。拼车平台中的分配问题：利用离线可重复使用资源进行在线匹配。《*ACM*经济学与计算汇刊》（*TEAC*），9(3):1–17，2021 年。ISSN 2167-8375。[8] Z. Dong、S. Das、P. Fowler 和 C.-J. Ho。稀缺可重复使用资源的高效非短视在线分配。2021 年。[9] V. Goyal、G. Iyengar 和 R. Udwani。可重复使用资源在线分配的渐近最优竞争比。*arXiv* 预印本 *arXiv:2002.02430*，2020 年。[10] Y. He、Q. Wei 和 T. C. E. Cheng。延迟作业数量与压缩成本之间存在权衡的单机调度。《调度杂志》，10(4):303–310，2007 年。ISSN 1099-1425。[11] Y. Hikima、Y. Akagi、N. Marumo 和 H. Kim。具有可控奖励和到达概率的在线匹配。载于 L. D. Raedt 编，第三十一届国际人工智能联合会议（*IJCAI-22*）论文集，第 1825–1833 页。国际人工智能联合会议组织，2022 年 7 月。doi: 10.24963/ijcai.2022/254。URL https://doi.org/10.24963/ijcai.2022/ 254。主会场。[12] H. Hosseini、Z. Huang、A. Igarashi 和 N. Shah。在线匹配中的类别公平性。*arXiv* 预印本 *arXiv:2203.03751*，2022 年。[13] P. Jaillet 和 X. Lu。在线随机匹配：具有更优界的新算法。《运筹学数学》，39(3):624–646，2014 年。ISSN 0364-765X。[14] A. Janiak 和 M. Y. Kovalyov。受截止期限和资源相关加工时间约束的单机调度。《欧洲运筹学杂志》，94(2):284–291，1996 年。ISSN 0377-2217。[15] R. M. Karp、U. V. Vazirani 和 V. V. Vazirani。在线二分图匹配的最优算法。第 352–358 页，1990 年。

[16] E. Koca、H. Yaman 和 M. S. Akt¨urk。具有可控处理时间的随机批量计划问题。*Omega*，53：1–10，2015。ISSN 0305-0483。[17] E. Koca、H. Yaman 和 M. S. Akt¨urk。考虑紧张效应的随机批量计划问题。*Computers & Operations Research*，94：23–37，2018。ISSN 0305-0548。[18] U. Krengel 和 L. Sucheston。半鞅与有限值。*Bulletin of the American MathematicalSociety*，83(4)：745–747，1977。[19] H. Liu、Q. Zhao、Y. Ma 和 F. Dai。用于点监督人群计数的二分图匹配。发表于*IJCAI*，第 860–866 页，2021。[20] M. Lowalekar、P. Varakantham 和 P. Jaillet。Zac：一种用于实现高效实时拼车的区域路径构建方法。第 29 卷，第 528–538 页，2019。ISBN 2334-0843。[21] M. Lowalekar、P. Varakantham 和 P. Jaillet。在线多容量拼车的竞争比。*arXiv*预印本 *arXiv:2009.07925*，2020。[22] C. Lu、L. Gao、X. Li 和 S. Xiao。用于真实焊接行业动态调度的混合多目标灰狼优化器。人工智能工程应用，57：61–79，2017。ISSN 0952-1976。[23] W. Ma、P. Xu 和 Y. Xu。在线二分图匹配中的群体公平性最大化。*arXiv*预印本 *arXiv:2011.13908*，2020。[24] B. Mahesh。机器学习算法综述。国际科学与研究期刊（*IJSR*）。*[互联网]*, 9：381–386，2020。[25] A. Mehta。在线匹配与广告分配。理论计算机科学基础与趋势®，8(4)：265–368，2013。ISSN 1551-305X。[26] A. Mehta、A. Saberi、U. Vazirani 和 V. Vazirani。Adwords 与广义在线匹配。*ACM*学报（*JACM*），54(5)：22–es，2007。ISSN 0004-5411。[27] A. Mehta、B. Waggoner 和 M. Zadimoghaddam。概率不等的在线随机匹配。第 1388–1404 页。SIAM，2015。[28] V. Nanda、P. Xu、K. A. Sankararaman、J. Dickerson 和 A. Srinivasan。高峰时段拼车平台中利润与公平性之间的权衡。第 34 卷，第 2210–2217 页，2020。ISBN 2374-3468。[29] D. Shabtay 和 M. Kaspi。具有凸资源消耗函数的并行机器调度。《欧洲运筹学杂志》，173(1)：92–107，2006。ISSN 0377-2217。[30] D. Shabtay 和 G. Steiner。可控处理时间调度综述。离散应用数学，155(13)：1643–1666，2007。ISSN 0166-218X。[31] H. Sumita、S. Ito、K. Takemura、D. Hatano、T. Fukunaga、N. Kakimura 和 K.-i. Kawarabayashi。可重复使用资源的在线任务分配问题。发表于*AAAI*人工智能会议论文集，第 36 卷，第 5199–5207 页，2022。[32] A. Tafreshian、N. Masoud 和 Y. Yin。服务科学前沿：点对点共享出行的乘车匹配：综述与未来方向。服务科学，12(2-3)：44–60，2020。ISSN 2164-3962。[33] TAIA。轻松选择翻译交付日期。https://taia.io/features/translation-delivery-dates/，2022。

[34] Y. Tong、Z. Zhou、Y. Zeng、L. Chen 和 C. Shahabi。空间众包：综述。*VLDB Journal*，29(1)：217–250，2020。ISSN 0949-877X。[35] H. Tunc。带有可控处理时间的随机批量规模确定问题的混合整数规划公式。*Computers & Operations Research*，132：105302，2021。ISSN 0305-0548。[36] D.-J. Wang、F. Liu 和 Y. Jin。由定向搜索引导的动态调度多目标进化算法。*Computers & Operations Research*，79：279–290，2017。ISSN 0305-0548。[37] Z. Wang、W. Bao、D. Yuan、L. Ge、N. H. Tran 和 A. Y. Zomaya。参见：服务中断期间移动 DNN 推理的提前退出调度。第 279–288 页，2019 年。

困难性分析

定理 3. 对于任意 $\zeta > 0$，不存在任何在线算法能够使该问题达到 $(\frac{1}{2} + \zeta)$-竞争比，即使对所有 $e \in E$ 和 $L = 1$，$q_e = 1$。

证明。考虑如下系统 [18]：$U = \{u\}$、$V = \{x, y, z\}$、$T = 2$、$p_{x,\,1} = 1$、$p_{x,\,2} = 0$、$p_{y,\,1} = 0$、$p_{y,\,2} = \epsilon$、$p_{z,\,1} = 0$、$p_{z,\,2} = 1 - \epsilon$。$r_{u,\,x} = 1$、$r_{u,\,y} = 1/\epsilon$、$r_{u,\,z} = 0$。占用时间等于 $T(\Pr\{d = T\} = 1)$。对于所有 $e \in E$，接受概率为 $q_e = 1$。

在离线设置中，有两种可能的任务序列。以 ϵ 的概率，完整任务序列为$I_1 = (x,\ y)$。以 $1 - \epsilon$ 的概率，完整任务序列为 $I_2 = (x,\ z)$。对于 I_1，离线最优算法在 $t = 1$ 丢弃 x，并在 $t = 2$ 将 u 分配给 y。对于 I_2，离线最优算法在 $t = 1$ 将 u 分配给 x，并在 $t = 2$ 丢弃 z。于是有 $\mathbb{E}_{I \sim \mathcal{I}}[\mathrm{OPT}(I)] = 2 - \epsilon$。

对于任何在线算法，如果它在 $t = 1$ 时将 u 分配给 x，则获得的总奖励为 1。如果它没有在 $t = 1$ 时将 u 分配给 x，那么它以概率 ϵ 获得（至多）$1/\epsilon$ 的总奖励，或者以概率 $1 - \epsilon$ 获得 0 的总奖励。因此，任何在线算法的期望总奖励至多为 1。由此可知，竞争比至多为 $1/(2 - \epsilon)$。

□

B 真实世界数据集

除了第 5.4 节中的评估外，我们还使用纽约市黄色出租车数据集[2] 将 OMLA 与五种基准算法进行比较（该数据集也曾被 [31] 和 [28] 使用）。该数据集包含 2013 年纽约出租车行程的记录。每条行程记录包括出租车 ID、司机 ID、起点和终点位置、行程开始和结束的时间及日期、行程距离、行程时长、车费金额和通行费金额。

我们从该数据集中提取了 1 月份 23 个工作日晚上 7 点至 8 点之间的数据，并剔除了这些时段内行程记录少于六条的出租车。我们关注经度从 $-73°$ 到 $-75°$、纬度从 40.4° 到 41° 的区域，经纬度的步长均为 0.02°（每个 $0.02° \times 0.02°$ 的正方形称为一个“网格”）。在该数据集子集中，共有 2308 条行程记录和 326 辆出租车。如果一条行程记录的起点位于某个网格内，且终点也位于某个网格内，则将其归为同一种行程类型。

[2] http://www.andresmh.com/nyctaxitrips/

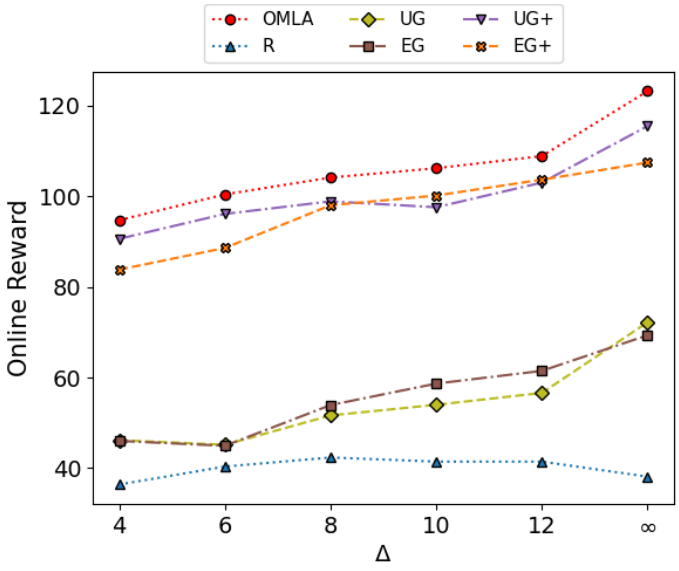


图4：不同基准算法在不同的 Δ、$L = 3$ 下的在线奖励。

将行程记录分组后，共得到 316 种行程类型。对于每种行程类型 v 以及每辆出租车 u，如果该行程类型的起点位置与出租车所在位置位于同一网格内，则存在一条边 (u, v)。我们将这一小时划分为 60 个时间段，每个时间段代表一分钟（$T = 60$）。概率 $p_{v,t}$设定为与 1 月份 23 个工作日中任意一天晚上 7 点至 8 点期间，类型为 v 的行程在时间 t 到达的频率成正比，即行程类型为 v 的行程在 t 到达。

在一月份记录的所有出行中，收费金额为 0、2.2 美元和 4.8 美元的出行占 99.9%，因此我们选择这三个收费层级作为三个处理级别。对于可以经过收费公路的出行，我们使用 Google 地图规划具有上述三个收费层级的替代路线（工作日晚上 7–8 点之间）。利用不同收费层级下的出行时长，我们计算每个层级的出行时长分布。奖励$r_{u,v,l}$ 设为出行类型 v 的平均车费减去级别 l收取的收费。某种出行类型的平均车费设为该类型所有出行的平均车费。我们设定 $L = 3$，其中 $l = 1$表示选择费用为 4.8 美元的收费公路，$l = 2$ 表示选择费用为 2.2 美元的收费公路，$l = 3$ 表示不选择任何收费公路。对于有限拒绝情形，初始拒绝预算 Δ_u 从 $[\Delta]$中均匀采样，拒绝惩罚 θ_l 设为 $\theta_1 = 3$、$\theta_2 = 5$ 和 $\theta_3 = 10$。

对于每个 Δ，我们进行 35 轮实验。在每轮实验中，我们首先按如下方式构建二分图 $G = (U, V; E)$。我们从上述 326 辆出租车中随机抽取 10 辆，构成 U，然后将与 U 中至少一辆出租车相连的行程类型构成 V（请注意，[31] 也会抽取出租车的一个子集来缩小二分图的规模；我们通过对 35 个样本的在线奖励取平均，改进了这种方法）。对于每条边 $e = (u, v)$，我们将接受概率设为 $q_e \sim U(0.5, 1)$。我们根据上述参数随机生成 40 个任务序列，并将 OMLA（或基准算法）在这 40 个序列上的平均在线奖励，记录为该样本中 OMLA（或基准算法）的平均在线奖励。然后，将 Δ 中 OMLA（或基准算法）的平均在线奖励计算为 35 个样本的平均奖励。

利用上述二分图和参数，我们在图4中绘制了 OMLA 和五种基准算法的平均在线奖励。我们注意到，随着 Δ 增加，除随机算法外，OMLA 和所有基准算法的性能均有所提升。这是因为较大的初始拒绝预算允许出租车拒绝更多分配（平均而言），从而在系统中停留更长时间（平均而言），因此获得更高的平均奖励。如图4所示，对于 Δ 的所有取值，OMLA 的表现均优于所有基准算法。这些结果表明，OMLA 在真实世界数据集上具有最佳的平均性能。

C 算法1和算法2的复杂度分析

在第3.4节中，我们在算法1中提出了 OMLA。在每个 t，每当一个任务 v 到达时，算法1为该任务做出决策的复杂度为 $\mathcal{O}(1)$。

算法3 无限拒绝情形下的在线机器与级别分配（OMLA）算法

输入： U，V，E，$\{Q_{e,l,t}\}$，$\{R_{u,t}\}$，$\mathbf{x}^*$

对于所有 $t \leftarrow 1$ 到 T **执行**
  如果没有任务到达，**则**
    skip
  否则（v 到达）
    以概率 x^* 选择一对（u，l） $_{e,l,t}/p_{v,\ t}$
    如果 u 未被占用且 $Q_{e,\ l,\ t} \geq R_{u,\ t+1}$，**则**
      我们将（u，l）分配给 v
      如果 u接受，**则**
        从 D_l 中抽取 d_l，u 被占用 d_l】【。
      结束 if
    结束如果
  结束如果
结束循环

为了执行算法1，我们需要预先（以离线方式）运行算法2。算法2的复杂度分析如下。在第3–5行中，我们最多需要 $\Delta \cdot |U| \cdot |V| \cdot L$ 步来计算 $t = T$ 处的每个激活值。在第6–16行中，我们最多需要 $\Delta \cdot |U| \cdot [(L + 2) \cdot |V| + 2]$ 步来计算 $t = T$ 处的每个基准值。在第19–29行中，对于给定的 (δ, u, t) 集合，我们最多需要 $|V| \cdot L \cdot (T + 3)$ 步来计算每个激活值。在第30行中，对于给定的 (δ, u, t) 集合，我们最多需要 $|V| \cdot L$ 步来计算每个基准值。在第17–32行中，我们最多需要 $(T - 1) \cdot \Delta \cdot |U| \cdot |V| \cdot L \cdot (T + 4)$ 步来计算所有激活值和基准值。因此，算法2的总体复杂度为 $\mathcal{O}(T^2 \cdot \Delta \cdot |U| \cdot |V| \cdot L)$（对于需要做出 T 个决策的系统）。

D 无限拒绝情形下算法1和算法2的修改版本

本小节给出算法3（分别为算法4）中适用于无限拒绝情形的修改版算法1（分别为算法2）。对于

算法 ithm 4 Calculation of Activati值和 d Base无限拒绝情形下的基线值 e

输入： $U, V, E, \mathcal{L}, T, \{q_e\}, \{r_{u,v,l}\}, \{D_l\}$1：求解 LP(Off) 以获得 $\mathbf{x}^*$2： $\Delta \leftarrow$ 最大化 Δ_u3：对于所有满足 (u, v) 的 (u, v, l)， $\in E$ 并执行 $l \in \mathcal{L}$ 4： $Q_{e,l,T} \leftarrow q_e r_{u,v,l}$5：结束循环 6：对于所有满足 $u \in U$ 的 u ，执行 7： $a \leftarrow 0$8：对于所有满足 (u, v) 的 v ， $\in E_u$ 执行 9： $b \leftarrow 0$10：对于所有 $l \in \mathcal{L}$ ，执行 11： $b \leftarrow b + x^*_{e,l,T} q_e r_{u,v,l}$12：结束循环 13： $a \leftarrow a + b$14：结束循环 15： $R_{u,T} \leftarrow a$ 16：结束循环 17：对于 $t \leftarrow T - 1$ 到 1，执行 18：对于满足 $u \in U$ 的 u ，执行 19：对于满足 (u, v) 的 (v, l)， $\in E_u$ 并执行 $l \in \mathcal{L}$ 20： $a \leftarrow 0$ 21： $b \leftarrow R_{u,t+1}$22：对于 $d \leftarrow 1$ 到 $T - t + 1$ ，执行 23： $a \leftarrow a + R_{u,t+d} \Pr\{d_l = d\}$24：结束循环 25： $Q_{e,l,t} \leftarrow q_e(r_{u,v,l} + a) + (1 - q_e)b$26：结束循环 27：根据（ 17 ）计算 $R_{u,t}$ 28：结束循环 29：结束循环 输出： $\{Q_{e,l,t}\}, \{R_{u,t}\}, \mathbf{x}^*$

无限拒绝情形是

$$Q_{e,l,t} = q_e(r_{u,v,l} + \sum_{d' \in [T-t]} \Pr\{d_l = d'\} R_{u,t+d'}) + (1-q_e) R_{u,t+1}, \quad (t \in [T]), \tag{16}$$

$$R_{u,t} = \sum_{e \in E_u} \sum_{l \in \mathcal{L}} x^*_{e,l,t} \max\{Q_{e,l,t}, R_{u,t+1}\} + (1 - \sum_{e \in E_u} \sum_{l \in \mathcal{L}} x^*_{e,l,t}) R_{u,t+1}, (t \in [T]). \tag{17}$$

E 引理与定理的证明

E.1 引理 1 的证明

为了实现 I 的 OPT(I)，离线最优算法会确定一个概率，即 u 在 t 时以 l 被分配给 v，记为 $y_{I,e,l,t}$。最大化后的期望奖励 OPT(I) 等于 $\sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} y_{I,e,l,t}$。在每个序列 I 上，$y_{I,e,l,t}$ 的均值为 $\mathbb{E}_{I\sim\mathcal{I}}[y_{I,e,l,t}]$，称为$y^*_{e,l,t}$。离线最优值$\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)]$ 等于 $\sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} y^*_{e,l,t}$。为方便起见，我们对每个任务序列 I 定义 $\mathbf{y}_I \text{ def}= \{y_{I,e,l,t}\}$，并令 $\mathbf{y}^* \stackrel{\text{def}}{=} \{y^*_{e,l,t}\}$。

为证明引理 1，我们证明每个$\mathbf{y}_I$ 都是 LP(Off) 的可行解，因此$\mathbf{y}^*$ 也是 LP(Off) 的可行解。也就是说，它们满足约束 (1) 至 (5)。

给定任务序列 I 和 $\mathbf{y}_I$：　1○ 对于每个 u 和 t，如果 u 已被先前任务占用，我们就不会在 t将新任务分配给 u；如果 u 在 t未被占用，我们最多可以向 u分配一个任务，因此 $\mathbf{y}_I$ 满足 (1)；2○ 对于每个 u，如果它在 T之后被占用，那么它接受该任务时至少有 1 的剩余预算；否则，u 承担的拒绝惩罚总额至多为 $\Delta_u + \theta - 1$（承担拒绝惩罚 θ，且剩余预算为 1），因此 $\mathbf{y}_I$ 满足 (2)；3○ 对于每个 v 和 t，如果 v 在 t到达，我们最多为 v分配一台机器；如果 v 没有在 t到达，我们就不会向 v分配任何机器，因此 $\mathbf{y}_I$ 满足 (3)；　4○ 对于每个 $e=(u,v)$ 和 t，如果 v 在 t 到达且我们将 u 分配给 v，那么我们将 u 分配给 v 时至多采用一个处理级别，因此 $\mathbf{y}_I$ 满足 (4)；5○ 对于每个 u 和 t，我们在 t至多分配一次 u，因此 $\mathbf{y}_I$ 满足 (5)。因此，$\mathbf{y}_I$ 是 Off 的可行解。

每个 $y^*_{e,l,t}$ 都是 $\mathcal{I}$ 上的期望 $y_{I,e,l,t}$。由于 $\mathbf{y}_I$ 满足约束 (1) 至 (5)，$\mathbf{y}^*$ 也满足约束 (1) 至 (5)。以$\mathbf{y}^*$ 作为 Off 的解，Off 的目标函数结果为 $\sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} y^*_{e,l,t}$。于是，以下等式成立：

$$\begin{aligned}
\mathbb{E}_{I\sim\mathcal{I}}[\text{OPT}(I)] &= \sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} y^*_{e,l,t} \\
&\le \sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} x^*_{e,l,t} \\
&= \text{LP}(\texttt{Off}),
\end{aligned}$$

其中，$\{x^*_{e,l,t}\}$ 是 Off 的最优解。由此证明了引理 1。

E.2 引理2的证明

由于我们设置了 $R^\delta_{u,t} = 0$ 和 $\tilde{R}^\delta_{u,t} = 0$，若 $\delta \le 0$, $R^\delta_{u,t} \ge \tilde{R}^\delta_{u,t}$ 对所有 $t \in [T]$（当 $\delta \le 0$ 时）均成立，则引理2的其余部分可通过对 t 进行反向归纳法证明。我们从 T 处的初始条件开始。对于所有 $\delta > 0$，有

$$\begin{cases}
\tilde{Q}^\delta_{u,l,T} = q'_{u,l,T} r'_{u,l,T} = (\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,T}) / p'_{u,l,T}, \\
\tilde{R}^\delta_{u,T} = \sum_{l\in\mathcal{L}} p'_{u,l,T} \tilde{Q}^\delta_{u,l,T} = \sum_{e\in E_u} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} x^*_{e,l,T}.
\end{cases} \tag{18}$$

由 (8) 和 (24)，我们有 $R^\delta_{u,\,T} = \tilde{R}^\delta_{u,\,T}$，对所有 δ。

对于 $t \in [T]$，我们假设对于每个 δ 和 $t' \in [t+1, T]$，都有 $R^{\delta}_{u,t'} \geq \tilde{R}^{\delta}_{u,t'}$。根据$R^{\delta}_{u,t}$ (7) 的计算，我们有

$$\begin{aligned}
R^{\delta}_{u,t} =& \sum_{e\in E_u}\sum_{l\in\mathcal{L}} x^*_{e,l,t} \max\{Q^{\delta}_{e,l,t}, R^{\delta}_{u,t+1}\} + (1 - \sum_{e\in E_u}\sum_{l\in\mathcal{L}} x^*_{e,l,t}) R^{\delta}_{u,t+1} \\
\geq& \sum_{l\in\mathcal{L}} \max\{ \sum_{e\in E_u} x^*_{e,l,t} Q^{\delta}_{e,l,t}, p'_{u,l,t} R^{\delta}_{u,t+1}\} + (1 - \sum_{l\in\mathcal{L}} p'_{u,l,t}) R^{\delta}_{u,t+1} \\
\geq& \sum_{l\in\mathcal{L}} \max\{ \sum_{e\in E_u} x^*_{e,l,t} Q^{\delta}_{e,l,t}, p'_{u,l,t} \tilde{R}^{\delta}_{u,t+1}\} + (1 - \sum_{l\in\mathcal{L}} p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1}.
\end{aligned} \tag{19}$$

(19)中的最后一个不等式来自归纳假设，其中 $R^{\delta}_{u,t'} \geq \tilde{R}^{\delta}_{u,t'}$ 对所有 δ 和 $t' \in [t+1, T]$均成立。对于每个 $l \in \mathcal{L}$，我们有

$$\begin{aligned}
\sum_{e\in E_u} x^*_{e,l,t} Q^{\delta}_{e,l,t} =& \sum_{e\in E_u} x^*_{e,l,t} \Big[q_e \big(\sum_{d'=1}^{T-t} \Pr\{d_l = d'\} R^{\delta}_{u,t+d'} + r_{u,v,l} \big) + (1 - q_e) R^{\delta-\theta_l}_{u,t+1} \Big] \\
=& \sum_{e\in E_u} x^*_{e,l,t} q_e \sum_{d'=1}^{T-t} \Pr\{d_l = d'\} R^{\delta}_{u,t+d'} + \sum_{e\in E_u} x^*_{e,l,t} q_e r_{u,v,l} \\
&+ \sum_{e\in E_u} x^*_{e,l,t} R^{\delta-\theta_l}_{u,t+1} - \sum_{e\in E_u} x^*_{e,l,t} q_e R^{\delta-\theta_l}_{u,t+1} \\
=& p'_{u,l,t} q'_{u,l,t} \sum_{d'=1}^{T-t} \Pr\{d_l = d'\} R^{\delta}_{u,t+d'} + r'_{u,l,t} p'_{u,l,t} q'_{u,l,t} + p'_{u,l,t} R^{\delta-\theta_l}_{u,t+1} - p'_{u,l,t} q'_{u,l,t} R^{\delta-\theta_l}_{u,t+1} \\
=& p'_{u,l,t} q'_{u,l,t} \sum_{d'=1}^{T-t} \Pr\{d_l = d'\} R^{\delta}_{u,t+d'} + r'_{u,l,t} p'_{u,l,t} q'_{u,l,t} + p'_{u,l,t} (1 - q'_{u,l,t}) R^{\delta-\theta_l}_{u,t+1} \\
\geq& p'_{u,l,t} \Big[q'_{u,l,t} \sum_{d'=1}^{T-t} \Pr\{d_l = d'\} \tilde{R}^{\delta}_{u,t+d'} + r'_{u,l,t} q'_{u,l,t} + (1 - q'_{u,l,t}) \tilde{R}^{\delta-\theta_l}_{u,t+1} \Big] \\
=& p'_{u,l,t} \tilde{Q}^{\delta}_{u,l,t}.
\end{aligned} \tag{20}$$

由(19)和(20)，我们有

$$\begin{aligned}
R^{\delta}_{u,t} \geq& \sum_{l\in\mathcal{L}} \max\{p'_{u,l,t} \tilde{Q}^{\delta}_{u,l,t}, p'_{u,l,t} \tilde{R}^{\delta}_{u,t+1}\} + (1 - \sum_{l\in\mathcal{L}} p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1} \\
=& \sum_{l\in\mathcal{L}} p'_{u,l,t} \max\{\tilde{Q}^{\delta}_{u,l,t}, \tilde{R}^{\delta}_{u,t+1}\} + (1 - \sum_{l\in\mathcal{L}} p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1} \\
=& \tilde{R}^{\delta}_{u,t}.
\end{aligned} \tag{21}$$

(21)中的最后一个等式来自(9)。由(21)、归纳假设以及在 T,$R^{\delta}_{u,t} \geq \tilde{R}^{\delta}_{u,t}$ 处的初始条件可知，对所有 $\delta > 0$ 和 $t \in [T]$均成立。

因此，引理2得证。

E.3 引理3的证明

~为证明引理3，我们首先推导一个不等式$R_{u,t}$ (25)。我们构造一个线性规划 (28)，以最小化 $R_{u,1}$，~并满足 (25) 中的条件，从而可以通过求得~(28) 的对偶线性规划 (29) 的一个可行解，得到 $R_{u,1}$ 的下界。有了对偶线性规划的这个可行解，我们便可以得到 $R_{u,1}$的下界。下面我们逐步详细说明这一证明。

˜1) 首先，我们推导出 $R_{u,\ t}$ 的不等式 (25)。与 (9) 和 (10) 类似，无限拒绝情形下的参考基准值和参考激活值可计算如下：

$$\tilde{R}_{u,t} = \sum_{l\in\mathcal{L}} p'_{u,l,t} \max\{\tilde{Q}_{u,l,t}, \tilde{R}_{u,t+1}\} + (1 - \sum_{l\in\mathcal{L}} p'_{u,l,t})\tilde{R}_{u,t+1}, \tag{22}$$

and

$$\tilde{Q}_{u,l,t} = q'_{u,l,t}(r'_{u,l,t} + \sum_{d'\in[T-t]} \Pr\{d_l = d'\}\tilde{R}_{u,t+d'}) + (1 - q'_{u,l,t})\tilde{R}_{u,t+1}, \tag{23}$$

具有初始条件

$$\begin{cases} \tilde{Q}_{u,l,T} = q'_{u,l,T} r'_{u,l,T} = (\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,T})/p'_{u,l,T}, \\ \tilde{R}_{u,T} = \sum_{l\in\mathcal{L}} p'_{u,l,T} \tilde{Q}_{u,l,T} = \sum_{e\in E_u} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} x^*_{e,l,T}. \end{cases} \tag{24}$$

公式（22）意味着

$$\tilde{R}_{u,t} \geq \max\{\tilde{R}_{u,t+1} + \sum_{l\in\mathcal{L}} p'_{u,l,t}(\tilde{Q}_{u,l,t} - \tilde{R}_{u,t+1}), \tilde{R}_{u,t+1}\}. \tag{25}$$

为方便起见，对于每个 $t \in [T]$，我们定义 $B_{u,l,t} \overset{\text{def}}{=} p'_{u,l,t} q'_{u,l,t}$ 以及 $A_{u,l,t} \overset{\text{def}}{=} p'_{u,l,t} q'_{u,l,t} \Pr\{d_l = 1\} + p'_{u,l,t}(1 - q'_{u,l,t}) + (1/L - p'_{u,l,t})$，从而可以将（25）中的第一项写为

$$\tilde{R}_{u,t+1} + \sum_{l\in\mathcal{L}} p'_{u,l,t}(\tilde{Q}_{u,l,t} - \tilde{R}_{u,t+1}) = \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t} + \sum_{l\in\mathcal{L}} B_{u,l,t} \sum_{d'=2}^{T-t} \Pr\{d_l = d'\}\tilde{R}_{u,t+d'} + \sum_{l\in\mathcal{L}} A_{u,l,t}\tilde{R}_{u,t+1}. \tag{26}$$

此外，$A_{u,\ l,\ t}$ 也可以用另一种方式表示。

$$\begin{aligned} A_{u,l,t} &= p'_{u,l,t} q'_{u,l,t} \Pr\{d_l = 1\} + \frac{1}{L} - p'_{u,l,t} q'_{u,l,t} \\ &= \frac{1}{L} - B_{u,l,t} \Pr\{d_l \geq 2\}. \end{aligned} \tag{27}$$

˜2）在第二步中，我们构建一个线性规划 (28)，以最小化 $R_{u,1}$ ，并满足下文所示的 (25) 中的条件。

$$\begin{aligned} \overset{\text{˜最小化}}{R_{u,1}} \quad & \\ \text{s.t} \quad & \tilde{R}_{u,t} \geq \sum_{l\in\mathcal{L}} B_{u,l,t} \sum_{d'=2}^{T-t} \Pr\{d_l = d'\}\tilde{R}_{u,t+d'} + \sum_{l\in\mathcal{L}} A_{u,l,t}\tilde{R}_{u,t+1} + \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t}, \quad (t \in [T-1]), \\ & \tilde{R}_{u,t} \geq \tilde{R}_{u,t+1}, \quad (t \in [T-1]), \\ & \tilde{R}_{u,T} \geq \sum_{l\in\mathcal{L}} B_{u,l,T} r'_{u,l,T}, \\ & \tilde{R}_{u,t} \geq 0, \quad (t \in [T]). \end{aligned} \tag{28}$$

注意，我们无需计算 (28) 的具体解。我们只需使用上述线性规划进行分析。(28) 中的最小化问题是一个线性规划问题，其对偶线性规划为

$$\max \quad \sum_{t\in[T]} \alpha_{u,t} \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t}$$

满足以下条件：$\alpha_{u,\ 1}+\beta_{u,\ 1}\le 1$,

$$\begin{aligned}
&\alpha_{u,2}+\beta_{u,2}\le \alpha_{u,1}\sum_{l\in\mathcal{L}}A_{u,l,1}+\beta_{u,1},\\
&\alpha_{u,t}+\beta_{u,t}\le \sum_{t'\in[t-2]}\alpha_{u,t'}\sum_{l\in\mathcal{L}}B_{u,l,t'}\Pr\{d_l=t-t'\}+\alpha_{u,t-1}\sum_{l\in\mathcal{L}}A_{u,l,t-1}+\beta_{u,t-1},\quad (t\in[3,T-1]),\\
&\alpha_{u,T}\le \sum_{t'\in[T-2]}\alpha_{u,t'}\sum_{l\in\mathcal{L}}B_{u,l,t'}\Pr\{d_l=T-t'\}+\beta_{u,T-1}+\alpha_{u,T-1}\sum_{l\in\mathcal{L}}A_{u,l,T-1},\\
&\alpha_{u,t}\ge 0,\quad (t\in[T]),\\
&\beta_{u,t}\ge 0,\quad (t\in[T-1]).
\end{aligned}\tag{29}$$

3）在第三步　　，为了得到式（28）的下界，我们构造其对偶线性规划（29）的一个可行解。
我们设

$$\begin{cases}
\alpha_{u,t}=\gamma, & t\in[T],\\
\beta_{u,1}=1-\gamma,\\
\beta_{u,2}=\beta_{u,1}-\gamma+\gamma\sum_{l\in\mathcal{L}}A_{u,l,1},\\
\beta_{u,t}=\gamma\sum_{l\in\mathcal{L}}\sum_{t'\in[t-2]}B_{u,l,t'}\Pr\{d_l=t-t'\}\\
\qquad +\beta_{u,t-1}-\gamma+\gamma\sum_{l\in\mathcal{L}}A_{u,l,t-1}, & t\in[3,T-1],
\end{cases}\tag{30}$$

并证明，对于 $t\ge 2$, $\beta_{u,\ t}$，可以通过以下方式计算

$$\beta_{u,t}=1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t}B_{u,l,t'}\Pr\{d_l\ge t-t'+1\}.\tag{31}$$

首先，当 $t=2$ 时，我们有

$$\begin{aligned}
\beta_{u,2}=&1-\gamma-\gamma+\gamma\sum_{l\in\mathcal{L}}A_{u,l,1}\\
=&1-\gamma-\gamma+(1-\sum_{l\in\mathcal{L}}B_{u,l,1}\Pr\{d_l\ge 2\})\gamma\\
=&1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<2}B_{u,l,t'}\Pr\{d_l\ge 2-t'+1\}.
\end{aligned}\tag{32}$$

第二个等式（32）由式（27）得出。这表明，当 $t=2$ 时，式（31）成立。
接下来，我们通过归纳法证明（31）的其余部分。对于 $t\ge 3$，我们假设

$$\beta_{u,t-1}=1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t-1}B_{u,l,t'}\Pr\{d_l\ge t-t'\}.\tag{33}$$

那么根据（30），对于 $\beta_{u,\ t}$ 我们有

$$\beta_{u,t}=1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t-1}B_{u,l,t'}\Pr\{d_l\ge t-t'\}+\gamma\sum_{l\in\mathcal{L}}\sum_{t'\in[t-2]}B_{u,l,t'}\Pr\{d_l=t-t'\}-\Big(1-\sum_{l\in\mathcal{L}}A_{u,l,t-1}\Big)\gamma.\tag{34}$$

由（27）和（34）可得。

$$\begin{aligned}
\beta_{u,t}=&1-\gamma-\gamma\sum_{l\in\mathcal{L}}B_{u,l,t-1}\Pr\{d_l\ge 2\}-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t-1}B_{u,l,t'}\Pr\{d_l\ge t-t'\}\\
&+\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t-1}B_{u,l,t'}\Pr\{d_l=t-t'\}
\end{aligned}$$

$$
\begin{aligned}
=&1-\gamma-\gamma\sum_{l\in\mathcal{L}}B_{u,l,t-1}\Pr\{d_l\geq 2\}-\gamma\sum_{t'<t-1}\sum_{l\in\mathcal{L}}B_{u,l,t'}\Pr\{d_l\geq t-t'+1\}\\
=&1-\gamma-\gamma\sum_{t'<t}\sum_{l\in\mathcal{L}}B_{u,l,t'}\Pr\{d_l\geq t-t'+1\}. \qquad (35)
\end{aligned}
$$

然后根据 (35)、归纳假设以及 $\beta_{u,\ 2}$ 处的初始条件，当 $t\geq 2$ 时，(31) 成立。

4) 在第四步中，我们证明由 (30) 定义的 $\gamma\in[0,1/2]$, $\alpha_{u,t}$ 和 $\beta_{u,t}$ 是 (29) 的一个可行解。

首先，根据 $\gamma\in[0,1/2]$ 和 (30)，显然 $\alpha_{u,t}$ 和 $\beta_{u,t}$ 满足对偶线性规划 (29) 中的第一、第二、第三和第五个约束。

接下来，我们证明对于所有 t，$\beta_{u,t}\geq 0$ （(29) 中的最后一个约束）成立。对于 $t\in[2,T-1]$，根据 (31) 和$B_{u,l,t}=\sum_{e\in E_u}q_e x^*_{e,l,t}$，我们有

$$
\beta_{u,t}=1-\gamma-\gamma\sum_{t'<t}\sum_{l\in\mathcal{L}}\Pr\{d_l\geq t-t'+1\}\sum_{e\in E_u}q_e x^*_{e,l,t'}. \qquad (36)
$$

由 (1) 和 (36)，我们有 $\beta_{u,t}\geq 1-2\gamma$。因此，根据 $\gamma\in[0,1/2]$，由 (30) 定义的 $\beta_{u,t}\geq 1-2\gamma\geq 0$ 和 $\beta_{u,t}$满足 (29) 中的最后一个约束。

接下来，我们证明（29）中的第四个约束得以满足。对于 $\alpha_{u,t}=\gamma$，（29）中的第四个约束等价于 (29)

$$
\gamma\ \leq\gamma\sum_{t'\in[T-2]}\sum_{l\in\mathcal{L}}B_{u,l,t'}\Pr\{d_l=T-t'\}+\beta_{u,T-1}+\gamma\sum_{l\in\mathcal{L}}A_{u,l,T-1}. \qquad (37)
$$

根据 (27) 和 (31)，我们可以将 (37) 的右端项重新整理为

$$
\begin{aligned}
\text{RHS}=&1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<T-1}B_{u,l,t'}\Pr\{d_l\geq T-t'+1\}+\gamma-\gamma\sum_{l\in\mathcal{L}}B_{u,l,T-1}\Pr\{d_l\geq T-(T-1)+1\}\\
=&1-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<T}B_{u,l,t'}\Pr\{d_l\geq T-t'+1\}. \qquad (38)
\end{aligned}
$$

根据（38），（37）等价于

$$
1\geq\gamma(1+\sum_{l\in\mathcal{L}}\sum_{t'<T}B_{u,l,t'}\Pr\{d_l\geq T-t'+1\}). \qquad (39)
$$

当 $B_{u,l,t}=\sum_{e\in E_u}q_e x^*_{e,l,t}$ 时，满足（37），如果

$$
1\geq\gamma(1+\sum_{l\in\mathcal{L}}\sum_{t'<T}\sum_{e\in E_u}q_e x^*_{e,l,t'}\Pr\{d_l\geq T-t'+1\}). \qquad (40)
$$

由 (1) 可得 (40) 右端项满足以下不等式。

$$
\text{RHS of }(40)\leq 2\gamma, \qquad (41)
$$

因此，当 $1\geq 2\gamma$ 时，(37) 得到满足。当 $\gamma\in[0,1/2]$, $1\geq 2\gamma$ 有效时，$\alpha_{u,t}$ 和 $\beta_{u,t}$ 在 (30) 中的定义满足 (29) 中的第四个约束。

我们已经证明，当 $\gamma\in[0,1/2]$、$\alpha_{u,\ t}$ 和 $\beta_{u,\ t}$ 在（30）中定义的解是可行解，因为它满足（29）中的所有约束。 (29),

˜5）在第五步中，我们为$R_{u,\ 1}$ 提供了一个有效的下界，其中$\alpha_{u,\ t}$ 和$\beta_{u,\ t}$ 定义于（30）中。在$\gamma=1/2$ 的条件下，$\alpha_{u,\ t}$ 和$\beta_{u,\ t}$ 定义于（30）中，是（29）的一个可行解，而（29）的目标函数结果为

$$
\frac{1}{2}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}B_{u,l,t}r'_{u,l,t}=\frac{1}{2}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u}q_e r_{u,v,l}x^*_{e,l,t}
$$

$$\leq \max \sum_{t\in[T]} \alpha_{u,t} \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t}. \tag{42}$$

根据线性规划对偶定理，我们有

$$\min \tilde{R}_{u,1} = \max \sum_{t\in[T]} \alpha_{u,t} \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t}, \tag{43}$$

and

$$\min \tilde{R}_{u,1} \geq \frac{1}{2} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{44}$$

因此，引理3得证。

E.4 引理 4 的证明

由引理 1 可得

$$\mathrm{LP}(\texttt{Off}) = \sum_{t\in[T]} \sum_{e\in E} q_e \sum_{l\in\mathcal{L}} r_{u,v,l} x^*_{e,l,t} \geq \mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]. \tag{45}$$

由引理 2 和引理 3 可得

$$R_{u,1} \geq \tilde{R}_{u,1} \geq \frac{1}{2} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{46}$$

根据激活值的定义，原始系统中 OMLA 的期望总奖励$\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]$ 等于 $\sum_{u\in U} R_{u,\ 1}$ （对于无限拒绝情形）。然后由（45）和（46）可得

$$\begin{aligned}
\frac{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]} \geq & \frac{\sum_{u\in U} R_{u,1}}{\mathrm{LP}(\texttt{Off})} \\
\geq & \frac{\frac{1}{2}\sum_{u\in U} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\mathrm{LP}(\texttt{Off})}.
\end{aligned} \tag{47}$$

引理 4 得证。

E.5 定理 1 的证明

由引理 4 可得

$$\begin{aligned}
\frac{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]} \geq & \frac{\frac{1}{2}\sum_{u\in U} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\mathrm{LP}(\texttt{Off})} \\
= & \frac{\frac{1}{2}\sum_{u\in U} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\sum_{u\in U} \sum_{t\in[T]} \sum_{l\in\mathcal{L}} \sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}} \\
= & \frac{1}{2}.
\end{aligned} \tag{48}$$

定理 1 得证。

E.6 引理 5 的证明

当 $\delta \le 0$ 时，我们有 $\tilde{R}_{u,t}^{\delta-\theta_l} = 0$ 以及 $\tilde{R}_{u,t}^{\delta} = 0$，因此 (13) 对所有 $t \in [T]$ 和 $l \in \mathcal{L}$均成立。当 $\delta > 0$ 且$\delta \le \theta_l$时，我们有 $\tilde{R}_{u,t}^{\delta-\theta_l} = 0$以及 $\frac{\delta-\theta_l}{\delta}\tilde{R}_{u,t}^{\delta} \le 0$，因此 (13) 对所有 $t \in [T]$均成立。

接下来，我们通过归纳法证明 $\delta > \theta_l$ 的引理 5。初始条件为 $t = T$。我们有 $\tilde{R}_{u,T}^{\delta} = \tilde{R}_{u,T}^{\delta'}$，$T = \sum_{l\in\mathcal{L}} p'_{u,l,T} q'_{u,l,T} r'_{u,l,T}$对所有 δ，$\delta' > 0$均成立，因此 (13) 对所有 $t = T$, $\delta > \theta_l$均成立。

对于 $t < T$，我们假设 (13) 对所有 $t' > t$ 和 $\delta > \theta_l$均成立。由此，根据 (9)，我们有

$$\begin{aligned}
\tilde{R}_{u,t}^{\delta} =& \sum_{l\in\mathcal{L}} p'_{u,l,t} \max\{\tilde{Q}_{u,l,t}^{\delta}, \tilde{R}_{u,t+1}^{\delta}\} + \Big(1 - \sum_{l\in\mathcal{L}} p'_{u,l,t}\Big)\tilde{R}_{u,t+1}^{\delta} \\
=& \sum_{l\in\mathcal{L}} \max\{p'_{u,l,t}\tilde{Q}_{u,l,t}^{\delta} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta}, \frac{1}{L}\tilde{R}_{u,t+1}^{\delta}\} \\
\ge& \sum_{l\in\mathcal{L}} \max\{\frac{\delta}{\delta+\theta_l}[p'_{u,l,t}\tilde{Q}_{u,l,t}^{\delta} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta}] + \frac{\theta_l}{\delta+\theta_l}\cdot\frac{1}{L}\tilde{R}_{u,t+1}^{\delta}, \frac{1}{L}\tilde{R}_{u,t+1}^{\delta}\}. \qquad (49)
\end{aligned}$$

根据 (10) 和归纳假设，我们有

$$\tilde{Q}_{u,l,t}^{\delta} \ge q'_{u,l,t}\Big(r'_{u,l,t} + \sum_{d'\in[T-t]} \Pr\{d_l = d'\}\tilde{R}_{u,t+d'}^{\delta}\Big) + (1 - q'_{u,l,t})\frac{\delta-\theta_l}{\delta}\tilde{R}_{u,t+1}^{\delta}, \qquad (50)$$

那么在 (49) 中，我们有

$$\begin{aligned}
&\frac{\delta}{\delta+\theta_l}[p'_{u,l,t}\tilde{Q}_{u,l,t}^{\delta} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta}] + \frac{\theta_l}{\delta+\theta_l}\cdot\frac{1}{L}\tilde{R}_{u,t+1}^{\delta} \\
\ge& \frac{\delta}{\delta+\theta_l}\left[p'_{u,l,t}\left(q'_{u,l,t}\Big(r'_{u,l,t} + \sum_{d'=1}^{T-t}\Pr\{d_l = d'\}\tilde{R}_{u,t+d'}^{\delta}\Big) + (1 - q'_{u,l,t})\frac{\delta-\theta_l}{\delta}\tilde{R}_{u,t+1}^{\delta}\right) + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta}\right] \\
&+ \frac{\theta_l}{\delta+\theta_l}\cdot\frac{1}{L}\tilde{R}_{u,t+1}^{\delta} \\
=& \frac{\delta}{\delta+\theta_l}p'_{u,l,t}q'_{u,l,t}\sum_{d'=1}^{T-t}\Pr\{d_l = d'\}\tilde{R}_{u,t+d'}^{\delta} + \frac{\delta}{\delta+\theta_l}p'_{u,l,t}(1 - q'_{u,l,t})\frac{\delta-\theta_l}{\delta}\tilde{R}_{u,t+1}^{\delta} + \frac{\delta}{\delta+\theta_l}(\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta} \\
&+ \frac{\delta}{\delta+\theta_l}p'_{u,l,t}q'_{u,l,t}r'_{u,l,t} + \frac{\theta_l}{\delta+\theta_l}p'_{u,l,t}q'_{u,l,t}\sum_{t'\ge1}\Pr\{d_l = t'\}\tilde{R}_{u,t+1}^{\delta} + \frac{\theta_l}{\delta+\theta_l}p'_{u,l,t}(1 - q'_{u,l,t})\tilde{R}_{u,t+1}^{\delta} \\
&+ \frac{\theta_l}{\delta+\theta_l}(\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta} \\
=& \frac{\delta}{\delta+\theta_l}p'_{u,l,t}q'_{u,l,t}r'_{u,l,t} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta} + p'_{u,l,t}q'_{u,l,t}\sum_{t'\ge1}\Pr\{d_l = t'\}\Big(\frac{\theta_l}{\delta+\theta_l}\tilde{R}_{u,t+1}^{\delta} + \frac{\delta}{\delta+\theta_l}\tilde{R}_{u,t+t'}^{\delta}\Big) \\
&+ p'_{u,l,t}(1 - q'_{u,l,t})\frac{\delta}{\delta+\theta_l}\tilde{R}_{u,t+1}^{\delta}. \qquad (51)
\end{aligned}$$

根据 (9) ，当 $t' \ge 1$ 时，我们有 $\tilde{R}_{u,t+1}^{\delta} \ge \tilde{R}_{u,t+t'}^{\delta}$。因此，我们从式（51）的最后一个方程继续推导。

$$\begin{aligned}
&\frac{\delta}{\delta+\theta_l}[p'_{u,l,t}\tilde{Q}_{u,l,t}^{\delta} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta}] + \frac{\theta_l}{\delta+\theta_l}\cdot\frac{1}{L}\tilde{R}_{u,t+1}^{\delta} \\
\ge& \frac{\delta}{\delta+\theta_l}p'_{u,l,t}q'_{u,l,t}r'_{u,l,t} + (\frac{1}{L} - p'_{u,l,t})\tilde{R}_{u,t+1}^{\delta} + p'_{u,l,t}q'_{u,l,t}\sum_{t'\ge1}\Pr\{d_l = t'\}\Big(\frac{\theta_l}{\delta+\theta_l}\tilde{R}_{u,t+1}^{\delta} + \frac{\delta}{\delta+\theta_l}\tilde{R}_{u,t+t'}^{\delta}\Big) \\
&+ p'_{u,l,t}(1 - q'_{u,l,t})\frac{\delta}{\delta+\theta_l}\tilde{R}_{u,t+1}^{\delta}
\end{aligned}$$

$$
\begin{aligned}
&\geq \frac{\delta}{\delta+\theta_l} p'_{u,l,t} q'_{u,l,t} r'_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1} + p'_{u,l,t} q'_{u,l,t} \sum_{t' \geq 1} \Pr\{d_l = t'\} \tilde{R}^{\delta}_{u,t+t'} \\
&\quad + p'_{u,l,t}(1 - q'_{u,l,t}) \frac{\delta}{\delta+\theta_l} \tilde{R}^{\delta}_{u,t+1} \\
&\geq \frac{\delta}{\delta+\theta_l} \Big[p'_{u,l,t} q'_{u,l,t} r'_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta+\theta_l}_{u,t+1} + p'_{u,l,t} q'_{u,l,t} \sum_{t' \geq 1} \Pr\{d_l = t'\} \tilde{R}^{\delta+\theta_l}_{u,t+t'} + p'_{u,l,t}(1 - q'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1} \Big] \\
&= \frac{\delta}{\delta+\theta_l} \Big[p'_{u,l,t} \Big(q'_{u,l,t} \big(r'_{u,l,t} + \sum_{t' \geq 1} \Pr\{d_l = t'\} \tilde{R}^{\delta+\theta_l}_{u,t+t'} \big) + (1 - q'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1} \Big) + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta+\theta_l}_{u,t+1} \Big] \\
&= \frac{\delta}{\delta+\theta_l} \Big[p'_{u,l,t} \tilde{Q}^{\delta+\theta_l}_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta+\theta_l}_{u,t+1} \Big]. \qquad (52)
\end{aligned}
$$

式（52）的最后一个等式由式（10）得出。根据式（52）和式（49），我们有

$$
\begin{aligned}
\tilde{R}^{\delta}_{u,t} &\geq \sum_{l \in \mathcal{L}} \max \Big\{ \frac{\delta}{\delta+\theta_l} \Big[p'_{u,l,t} \tilde{Q}^{\delta+\theta_l}_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta+\theta_l}_{u,t+1} \Big], \frac{1}{L} \tilde{R}^{\delta}_{u,t+1} \Big\} \\
&\geq \frac{\delta}{\delta+\theta_l} \sum_{l \in \mathcal{L}} \max \Big\{ p'_{u,l,t} \tilde{Q}^{\delta+\theta_l}_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta+\theta_l}_{u,t+1}, \frac{1}{L} \tilde{R}^{\delta+\theta_l}_{u,t+1} \Big\} \\
&= \frac{\delta}{\delta+\theta_l} \tilde{R}^{\delta+\theta_l}_{u,t}. \qquad (53)
\end{aligned}
$$

根据 (53)、归纳假设以及 $t = T$ 处的初始条件，我们证明了当 $\delta > \theta_l$ 时，(13) 对所有 $t \in [T]$ 也成立。因此，引理 5 得证。

E.7 引理 6 的证明

为证明引理 6，我们首先推导关于 $\tilde{R}^{\delta}_{u,\,t}$ 的不等式 (54)。接着，我们构造线性规划 (57)，在满足 (54) 中条件的前提下最小化 $\tilde{R}^{\Delta_u}_{u,\,1}$，从而通过求解 (57) 的对偶线性规划 (58) 并找到一个可行解，得到 $\tilde{R}^{\Delta_u}_{u,\,1}$ 的下界。利用这个对偶线性规划的可行解，我们可以得到 $\tilde{R}^{\Delta_u}_{u\,1,}$ 的下界。下面逐步给出该证明的详细过程。

1）第一步，我们推导关于 $\tilde{R}^{\delta}_{u,\,t}$ 的不等式 (54)。根据 (9) 和引理 5，对于 $\tilde{R}^{\delta}_{u,\,t}$，有以下不等式：

$$
\begin{aligned}
\tilde{R}^{\delta}_{u,t} &\geq \sum_{l \in \mathcal{L}} \max \Big\{ p'_{u,l,t} \hat{Q}^{\delta}_{u,l,t} + (\frac{1}{L} - p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1}, \frac{1}{L} \tilde{R}^{\delta}_{u,t+1} \Big\} \\
&\geq \max \Big\{ \sum_{l \in \mathcal{L}} p'_{u,l,t} \hat{Q}^{\delta}_{u,l,t} + (1 - \sum_{l \in \mathcal{L}} p'_{u,l,t}) \tilde{R}^{\delta}_{u,t+1}, \tilde{R}^{\delta}_{u,t+1} \Big\}, \qquad (54)
\end{aligned}
$$

其中

$$
\hat{Q}^{\delta}_{u,l,t} = q'_{u,l,t} (r'_{u,l,t} + \sum_{d' \in [T-t']} \Pr\{d_l = d'\} \tilde{R}^{\delta}_{u,t+d'}) + (1 - q'_{u,l,t}) \frac{\delta - \theta_l}{\delta} \tilde{R}^{\delta}_{u,t+1}. \qquad (55)
$$

为方便记号，我们定义 $B_{u,l,t} \text{ def}= p'_{u,l,t} q'_{u,l,t}$ 以及 $A_{u,l,t} \text{ def}= p'_{u,l,t} q'_{u,l,t} \Pr\{d_l = 1\} + p'_{u,l,t}(1 - q'_{u,l,t}) \frac{\Delta_u - \theta_l}{\Delta_u} + (1/L - p'_{u,l,t})$，对于每个 $t \in [T]$，从而对于 $\tilde{R}^{\Delta_u}_{u,t}$，我们有

$$
\begin{aligned}
\tilde{R}^{\Delta_u}_{u,t} &\geq \max \Big\{ \sum_{l \in \mathcal{L}} p'_{u,l,t} \hat{Q}^{\Delta_u}_{u,l,t} + (1 - \sum_{l \in \mathcal{L}} p'_{u,l,t}) \tilde{R}^{\Delta_u}_{u,t+1}, \tilde{R}^{\Delta_u}_{u,t+1} \Big\} \\
&= \max \Big\{ \sum_{l \in \mathcal{L}} B_{u,l,t} r'_{u,l,t} + \sum_{l \in \mathcal{L}} A_{u,l,t} \tilde{R}^{\Delta_u}_{u,t+1} + \sum_{l \in \mathcal{L}} B_{u,l,t} \sum_{d'=2}^{T-t} \Pr\{d_l = d'\} \tilde{R}^{\Delta_u}_{u,t+d'}, \tilde{R}^{\Delta_u}_{u,t+1} \Big\}. \qquad (56)
\end{aligned}
$$

2) 在第二步中，我们构造一个 LP (57)，以最小化 $\tilde{R}_{u,1}^{\Delta_u}$，并满足如下所示的条件 (54)。

$$
\begin{aligned}
\text{最小化 } & \tilde{R}_{u,1}^{\Delta_u} \\
\text{s.t} \quad & \tilde{R}_{u,t}^{\Delta_u} \geq \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t} + \sum_{l\in\mathcal{L}} A_{u,l,t} \tilde{R}_{u,t+1}^{\Delta_u} + \sum_{l\in\mathcal{L}} B_{u,l,t} \sum_{d'=2}^{T-t} \Pr\{d_l = d'\} \tilde{R}_{u,t+d'}^{\Delta_u}, \ (t \in [T-1]), \\
& \tilde{R}_{u,t}^{\Delta_u} \geq \tilde{R}_{u,t+1}^{\Delta_u}, \quad (t \in [T-1]), \\
& \tilde{R}_{u,T}^{\Delta_u} \geq \sum_{l\in\mathcal{L}} B_{u,l,T} r'_{u,l,T}, \\
& \tilde{R}_{u,t}^{\Delta_u} \geq 0, \quad (t \in [T]).
\end{aligned}
\tag{57}
$$

请注意，我们无需计算（57）的具体解。我们只需使用上述线性规划进行分析。（57）中的最小化问题是一个线性规划问题，其对偶线性规划为

$$
\begin{aligned}
\max \quad & \sum_{t\in[T]} \alpha_{u,t} \sum_{l\in\mathcal{L}} B_{u,l,t} r'_{u,l,t} \\
\text{s.t.} \quad & \alpha_{u,1} + \beta_{u,1} \leq 1, \\
& \alpha_{u,2} + \beta_{u,2} \leq \alpha_{u,1} \sum_{l\in\mathcal{L}} A_{u,l,1} + \beta_{u,1}, \\
& \alpha_{u,t} + \beta_{u,t} \leq \sum_{t'\in[t-2]} \alpha_{u,t'} \sum_{l\in\mathcal{L}} B_{u,l,t'} \Pr\{d_l = t - t'\} + \beta_{u,t-1} + \alpha_{u,t-1} \sum_{l\in\mathcal{L}} A_{u,l,t-1}, \ (t \in [3, T-1]), \\
& \alpha_{u,T} \leq \sum_{t'\in[T-2]} \alpha_{u,t'} \sum_{l\in\mathcal{L}} B_{u,l,t'} \Pr\{d_l = T - t'\} + \beta_{u,T-1} + \alpha_{u,T-1} \sum_{l\in\mathcal{L}} A_{u,l,T-1}, \\
& \alpha_{u,t} \geq 0, \quad (t \in [T]), \\
& \beta_{u,t} \geq 0, \quad (t \in [T-1]).
\end{aligned}
\tag{58}
$$

3）在第三步中，为了求得（57）的下界，我们构造其对偶线性规划（58）的一个可行解。令

$$
\begin{cases}
\alpha_{u,t} = \gamma, & t \in [T], \\
\beta_{u,1} = 1 - \gamma, & \\
\beta_{u,2} = 1 - \gamma - \gamma(1 - \sum_{l\in\mathcal{L}} A_{u,l,1}), & \\
\beta_{u,t} = 1 - \gamma - \gamma \sum_{l\in\mathcal{L}} \sum_{t'<t} p'_{u,l,t'} q'_{u,l,t'} \Pr\{d_l \geq t - t' + 1\} - \gamma \sum_{l\in\mathcal{L}} \dfrac{\theta_l}{\Delta_u} \sum_{t'<t} p'_{u,l,t'} (1 - q'_{u,l,t'}), & t \in [3, T-1],
\end{cases}
\tag{59}
$$

并且我们证明，当 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$ 时，这是（58）的一个可行解。

根据 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$ 和（59），很容易看出，$\alpha_{u,t}$ 和 $\beta_{u,t}$ 满足（58）中的第一、第二和第五个约束。

对于（58）中的第三个约束，我们首先证明当 $t = 3$ 时该约束成立。当 $t = 3$ 时，（58）中的第三个约束等价于

$$
\gamma + \beta_{u,3} \leq \beta_{u,2} + \gamma \sum_{l\in\mathcal{L}} A_{u,l,2} + \gamma \sum_{t'\in[1]} \sum_{l\in\mathcal{L}} B_{u,l,t'} \Pr\{d_l = 3 - t'\}, \tag{60}
$$

并且我们有（60）左端项（LHS）和右端项的以下等式

$$\text{LHS of (60)} = 1-\gamma\sum_{l\in\mathcal{L}}\frac{\theta_l}{\Delta_u}\sum_{t'<3}p'_{u,l,t'}(1-q'_{u,l,t'})-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<3}p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l\geq 4-t'\}, \tag{61}$$

with

$$\begin{aligned}\text{RHS of (60)} =& 1-2\gamma+\gamma\sum_{l\in\mathcal{L}}\sum_{t'<3}\Big[p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l=1\}+p'_{u,l,t'}(1-q'_{u,l,t'})\frac{\Delta_u-\theta_l}{\Delta_u}+(\frac{1}{L}-p'_{u,l,t'})\Big]\\ &+\gamma\sum_{l\in\mathcal{L}}p'_{u,l,1}q'_{u,l,1}\Pr\{d_l=2\}.\end{aligned} \tag{62}$$

根据 (61) 和 (62)，若.....则 (60) 成立。

$$\gamma\sum_{l\in\mathcal{L}}\sum_{t'<3}p'_{u,l,t'}q'_{u,l,t'}(\Pr\{d_l\geq 1\}-1)\geq 0. \tag{63}$$

对于所有 l，我们有 $\Pr\{d_l\geq 1\}=1$，因此（63）得到满足，而与 γ 的取值无关。因此，当 $t=3$ 时，（58）中的第三个约束得到满足，而与 γ 的取值无关。

当 $t\in[4,T-\quad 1]$, 我们有以下等式，分别对应第三个约束的左端项和右端项。 (58)

$$\text{LHS} = 1-\gamma\sum_{l\in\mathcal{L}}\frac{\theta_l}{\Delta_u}\sum_{t'<t}p'_{u,l,t'}(1-q'_{u,l,t'})-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t}p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l\geq t-t'+1\}, \tag{64}$$

$$\begin{aligned}\text{RHS} =& 1-\gamma-\gamma\sum_{l\in\mathcal{L}}\frac{\theta_l}{\Delta_u}\sum_{t'<t-1}p'_{u,l,t'}(1-q'_{u,l,t'})-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<t-1}p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l\geq t-t'\}\\ &+\gamma\sum_{t'\in[t-2]}\sum_{l\in\mathcal{L}}p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l=t-t'\}+\gamma\sum_{l\in\mathcal{L}}A_{u,l,t-1}.\end{aligned} \tag{65}$$

根据（64）和（65），若满足以下条件，则（58）中的第三个约束成立。

$$0\leq\gamma\sum_{l\in\mathcal{L}}p'_{u,l,t-1}q'_{u,l,t-1}(\Pr\{d_l\geq 1\}-1). \tag{66}$$

因为对于每个 l，$\Pr\{d_l\geq 1\}=1$，当 $\gamma\in[0,\Delta_u/(3\Delta_u-1)]$, 时，(66) 成立，因此当 $\gamma\in[0,\Delta_u/(3\Delta_u-1)]$ 时，第三个约束成立。

为了证明第四个约束得到满足，我们需要证明对于 $t\in[T]$，(58) 中的最后一个约束得到满足，其中 $\beta_{u,\ T}$ 的设置如下所示

$$\beta_{u,T} = 1-\gamma-\gamma\sum_{l\in\mathcal{L}}\sum_{t'<T}p'_{u,l,t'}q'_{u,l,t'}\Pr\{d_l\geq T-t'+1\}-\gamma\sum_{l\in\mathcal{L}}\frac{\theta_l}{\Delta_u}\sum_{t'<T}p'_{u,l,t'}(1-q'_{u,l,t'}). \tag{67}$$

对于 $t=1$，我们有 $\beta_{u,\ 1}=1-\gamma\geq 0$ 当 $\gamma\in[0,\Delta_u/(3\Delta_u-1)]$ 时。对于 $t=2$，我们有 $\beta_{u,\ 2}=1-\gamma-(1-\sum_{l\in\mathcal{L}}A_{u,\ l,\ 1})\gamma$，其中

$$\sum_{l\in\mathcal{L}}A_{u,l,1}=\sum_{l\in\mathcal{L}}p'_{u,l,1}q'_{u,l,1}\Pr\{d_l=1\}+(1-\sum_{l\in\mathcal{L}}p'_{u,l,1})+\sum_{l\in\mathcal{L}}p'_{u,l,1}(1-q'_{u,l,1})\frac{\Delta_u-\theta_l}{\Delta_u}. \tag{68}$$

由于 $q'_{u,l,1}\leq 1$ 以及 $\sum_{l\in\mathcal{L}}p'_{u,l,1}\leq 1$，我们有 $\sum_{l\in\mathcal{L}}A_{u,l,1}\geq 0$，因此 $\beta_{u,2}\geq 1-2\gamma$。当 $\gamma\in[0,\Delta_u/(3\Delta_u-1)]$ 时，我们有 $\beta_{u,2}\geq 0$。

对于 $t \in [3, T]$，我们首先将 $\lambda_{u,t}$ 定义为

$$\lambda_{u,t} \stackrel{\text{def}}{=} \sum_{l \in \mathcal{L}} \frac{1}{\Delta_u} \sum_{t'<t} p'_{u,l,t'}(1 - q'_{u,l,t'})\theta_l + \sum_{l \in \mathcal{L}} \sum_{t'<t} p'_{u,l,t'} q'_{u,l,t'} \Pr\{d_l \geq t - t' + 1\}, \tag{69}$$

然后我们有 $\beta_{u,\,t} = 1 - \gamma - \gamma\lambda_{u,\,t}$。对于 $p'_{u,\,l,\,t} = \sum_{e \in\, E_u} x^*_{e,\,l,\,t}$ 和 $p'_{u,\,l,\,t} q'_{u,\,l,\,t} = \sum_{e \in\, E_u} q_e x^*_{e,\,l,\,t}$，我们有

$$\begin{aligned}\lambda_{u,t} =& \frac{1}{\Delta_u} \sum_{l \in \mathcal{L}} \sum_{t'<t} \sum_{e \in E_u} x^*_{e,l,t} \left[(1 - q_e)\theta_l + \theta q_e \Pr\{d_l \geq t - t' + 1\}\right] \\ &+ \frac{\Delta_u - \theta}{\Delta_u} \sum_{l \in \mathcal{L}} \sum_{t'<t} \sum_{e \in E_u} x^*_{e,l,t} \Pr\{d_l \geq t - t' + 1\}. \end{aligned} \tag{70}$$

根据 (1) 和 (2)，对于所有 $t \in [T]$，有 $\lambda_{u,t} \leq 1 + (\Delta_u - 1)/\Delta_u$。当 $\beta_{u,t} = 1 - \gamma - \gamma\lambda_{u,t} \geq 0$ 时，我们需要 $1 - (\lambda_{u,t} + 1)\gamma \geq 0$。显然，$\lambda_{u,t} \geq 0$，因此上述条件等价于 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$。换言之，当 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$, 时，(58) 中最后一个约束对所有 $t \in [T]$ 均成立。

对于（58）中的第四个约束，我们利用以下结论：当 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$ 时，$\beta_{u,\,T} \geq 0$。然后根据对第三个约束的分析，以下不等式成立：

$$\alpha_{u,T} + \beta_{u,T} \leq \sum_{t' \in [T-2]} \alpha_{u,t'} \sum_{l \in \mathcal{L}} B_{u,l,t'} \Pr\{d_l = T - t'\} + \beta_{u,T-1} + \alpha_{u,T-1} \sum_{l \in \mathcal{L}} A_{u,l,T-1}. \tag{71}$$

令 $\beta_{u,\,T} \geq 0$，则有 $\alpha_{u,\,T} \leq \alpha_{u,\,T} + \beta_{u,\,T}$。当 $\gamma \in [0, \Delta_u/(3\Delta_u - 1)]$ 时，第四个约束得到满足。$\alpha_{u,\,t}$ 以及在（59）中定义的 $\beta_{u,\,t}$ 是（58）的可行解。

4）在第四步中，我们为 $\tilde{R}^{\Delta_u}_{u,\,1}$ 提供一个有效的下界，其中 $\alpha_{u,\,t}$ 和在（59）中定义的 $\beta_{u,\,t}$。对于 $\gamma = \Delta_u/(3\Delta_u - 1)$，在（59）中定义的 $\alpha_{u,\,t}$ 和 $\beta_{u,\,t}$ 是（58）的可行解，并且（58）的目标函数结果为

$$\frac{\Delta_u}{3\Delta_u - 1} \sum_{t \in [T]} \sum_{l \in \mathcal{L}} B_{u,l,t} r'_{u,l,t} = \frac{\Delta_u}{3\Delta_u - 1} \sum_{t \in [T]} \sum_{l \in \mathcal{L}} \sum_{e \in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{72}$$

根据线性规划对偶定理，我们有

$$\min \tilde{R}^{\Delta_u}_{u,1} = \max \sum_{t \in [T]} \alpha_{u,t} \sum_{l \in \mathcal{L}} B_{u,l,t} r'_{u,l,t}, \tag{73}$$

and

$$\min \tilde{R}^{\Delta_u}_{u,1} \geq \frac{\Delta_u}{3\Delta_u - 1} \sum_{t \in [T]} \sum_{l \in \mathcal{L}} \sum_{e \in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{74}$$

因此，引理 6 得证。

E.8 引理 7 的证明

由引理 1 可知，蕴含

$$\mathrm{LP}(\texttt{Off}) = \sum_{t \in [T]} \sum_{e \in E} q_e \sum_{l \in \mathcal{L}} r_{u,v,l} x^*_{e,l,t} \geq \mathbb{E}_{I \sim \mathcal{I}}[\mathrm{OPT}(I)]. \tag{75}$$

根据引理 2 和引理 6，我们有

$$R^{\Delta_u}_{u,1} \geq \tilde{R}^{\Delta_u}_{u,1} \geq \frac{\Delta_u}{3\Delta_u - 1} \sum_{t \in [T]} \sum_{l \in \mathcal{L}} \sum_{e \in E_u} q_e r_{u,v,l} x^*_{e,l,t}. \tag{76}$$

根据激活值的定义，OMLA 在原始系统中的期望总奖励$\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]$ 等于 $\sum_{u\in U} R_{u,1}^{\Delta_u}$（对于有限拒绝情形）。然后根据（75）和（76），我们有

$$\begin{aligned}\frac{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]} \geq & \frac{\sum_{u\in U} R_{u,1}^{\Delta_u}}{\mathrm{LP}(\mathtt{Off})} \\ \geq & \frac{\sum_{u\in U}\frac{\Delta_u}{3\Delta_u-1}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\mathrm{LP}(\mathtt{Off})}. \end{aligned} \tag{77}$$

引理 7 得证。

E.9 定理 2 的证明

由引理 7 可得

$$\begin{aligned}\frac{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]} \geq & \frac{\sum_{u\in U}\frac{\Delta_u}{3\Delta_u-1}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\sum_{u\in U}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}} \\ \geq & \min_{u\in U}\frac{\frac{\Delta_u}{3\Delta_u-1}\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}}{\sum_{t\in[T]}\sum_{l\in\mathcal{L}}\sum_{e\in E_u} q_e r_{u,v,l} x^*_{e,l,t}} \\ = & \min_{u\in U}\frac{\Delta_u}{3\Delta_u-1}. \end{aligned} \tag{78}$$

我们定义 $\Delta=\max_{u\in U}\Delta_u$，则有

$$\begin{aligned}\frac{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{ALG}(I)]}{\mathbb{E}_{I\sim\mathcal{I}}[\mathrm{OPT}(I)]} \geq & \min_{u\in U}\frac{\Delta_u}{3\Delta_u-1} \\ = & \frac{\Delta}{3\Delta-1}. \end{aligned} \tag{79}$$

由式 (79)，定理 2 得证。