

随机过程及应用项目报告

吴苛铭 龚欣怡 方泽宇
522120910244 522120910203 521120910079

日期：2025 年 6 月 19 日

1 研究主题介绍

1.1 相关论文

Agrawal, Narendra, Sami Najafi-Asadolahi, and Stephen A. Smith. “A Markov Decision Model for Managing Display-Advertising Campaigns.” *Manufacturing & service operations management* 25.2 (2023): 489–507.[1]

1.2 研究问题

在现实的广告投放过程中，广告代理商需要根据广告方的需求向特定的观看人群投放相关广告。为此，广告代理商需要制定一个竞价策略，从而在互联网广告交易平台上争取到相应的观众。在现实的复杂情况下确定一个最优的竞价策略是十分困难的，因此文章在简化的条件下，使用马尔科夫决策过程对广告投放问题进行建模，得到了一个基于当前系统状态的动态策略。

1.3 数据来源

我们和论文中一样使用 iPinYou Dataset，这是一个基于现实世界数据的数据集，可以用于 Real-Time Bidding Benchmarking，部分数据描述如下：[2]

表 1: 部分数据字段

| 列编号 | 数据字段 | 例子 |
|------|--------------------|-------------------------|
| *1 | Bid ID | 015300008...3f5a4f5121 |
| 2 | Timestamp | 20130218001203638 |
| *4 | iPinYou ID | 35605620124122340227135 |
| *9 | Ad exchange | 2 |
| 13 | Ad slot ID | 2147689_8764813 |
| 14 | Ad slot width | 300 |
| 15 | Ad slot height | 250 |
| 16 | Ad slot visibility | SecondView |
| 17 | Ad slot format | Fixed |
| *20 | Bidding price | 753 |
| *†21 | Paying price | 15 |

2 模型分析

文章提出一种基于马尔可夫决策过程 (Markov Decision Process) 的建模框架, 用于动态管理多个并发的展示类广告投放活动。广告代理在有限的资源和观众面前, 需要在不同广告活动之间做出最优投放选择, 以最大化长期收益。该问题本质上是一个具有多个广告类型、多个观众类型与有限服务容量的多维队列控制问题。

首先作者对问题进行简化, 证明了该问题可以根据不同的观众类型分隔为多个独立的子问题, 每个子问题具有多个广告类型与单个观众类型。接下来, 作者对每个子问题进行建模。

在精确建模方面, 作者将系统状态定义为每种广告类型当前的排队人数, 并设定奖励函数、服务率 (代表观众到达率)、以及惩罚函数 (如广告未投放带来的损失), 构建出一套完整的有限期动态规划模型。考虑一个具有 N 类广告活动的系统, 状态 $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \dots, x_{N,t})$ 表示各类广告在时间 t 时的排队请求数量, 单位时间内观众以速率 μ 到达, 广告活动以速率 λ 到达, 每次请求引入 s 个待服务广告。定义 $V_t(\mathbf{x})$ 为状态 \mathbf{x} 下在剩余 t 步内的最大期望收益, 其贝尔曼方程可表示为:

$$V_t(\mathbf{x}) = \max_{\mathbf{w} \in [0,1]^N} \left\{ \frac{\mu}{\lambda N + \mu} \cdot \sum_{i=1}^N [w_i (r - \pi(w_i) + V_{t-1}(\mathbf{x} - \mathbf{e}_i)) + (1 - w_i) V_{t-1}(\mathbf{x})] + \frac{\lambda}{\lambda N + \mu} \cdot \sum_{i=1}^N V_{t-1}(\mathbf{x} + s \cdot \mathbf{e}_i) - c \cdot \|\mathbf{x}\|_1 \right\} \quad (1)$$

其中, $\pi(w) = -\log(1 - w)/\gamma$ 为投放概率 w 下的竞价代价函数, r 为单位广告服务的收益, c 为等待成本, \mathbf{e}_i 表示第 i 类广告的单位向量。

在此框架下, 每个时间点的最优策略需基于整个系统的当前状态, 计算在不同投放选择下的期望收益与未来价值, 从而逐步向后迭代得到最优策略。这种方法在理论上可以得到全局最优解, 但其计算复杂度随广告类型数指数增长, 限制了在大规模系统中的实际应用。

为了应对高维状态空间带来的计算瓶颈, 作者提出一种启发式近似解法。该方法将整体系统近似拆分为多个独立的单广告类型子系统, 分别求解每个广告类型的最优竞价策略, 并假设当前观众将被分配给出价最高的广告类型。在该简化模型中, 系统状态退化为单维 x , 则最优竞价概率 w^* 可通过如下一阶条件求解:

$$r - \pi(w) - w \cdot \pi'(w) = V_{t-1}(x) - V_{t-1}(x - 1) \quad (2)$$

这种策略无需遍历整个多维状态空间, 仅需处理单一维度, 计算效率显著提升。虽然此方法牺牲了一定的最优性, 但在实证验证中表现出良好的效果, 其收益在绝大多数情境下接近最优解。

3 数值计算与模拟

3.1 数据分析与假设验证

论文在建模过程中使用到了两个关键假设: 广告投放需求与广告观众以泊松到达的形式来到系统 (即相邻到达的间隔符合指数分布); 并且出价的价格同样符合指数分布。首先, 我们基

于现实的数据对这两个假设进行检验。

我们首先采用四分位数范围法对时间间隔和出价价格变量进行处理，仅保留位于统计正常范围内的观测值。针对清洗后的数据，我们使用最大似然估计法拟合相应分布，并生成相应的累计分布函数 (CDF)，以用于与实际观测数据进行比较。与此同时，我们采用 Kolmogorov–Smirnov (KS) 检验衡量理论分布与经验分布之间的差异程度，判断变量是否服从特定分布。

观众到达时间间隔分布：通过计算连续观众之间的到达时间差，我们构造时间间隔变量，拟合相应的指数分布。检验结果显示，KS 检验统计量为 0.1727，p 值远小于 0.05，这从统计上拒绝了“观众到达时间间隔服从指数分布”的原假设。

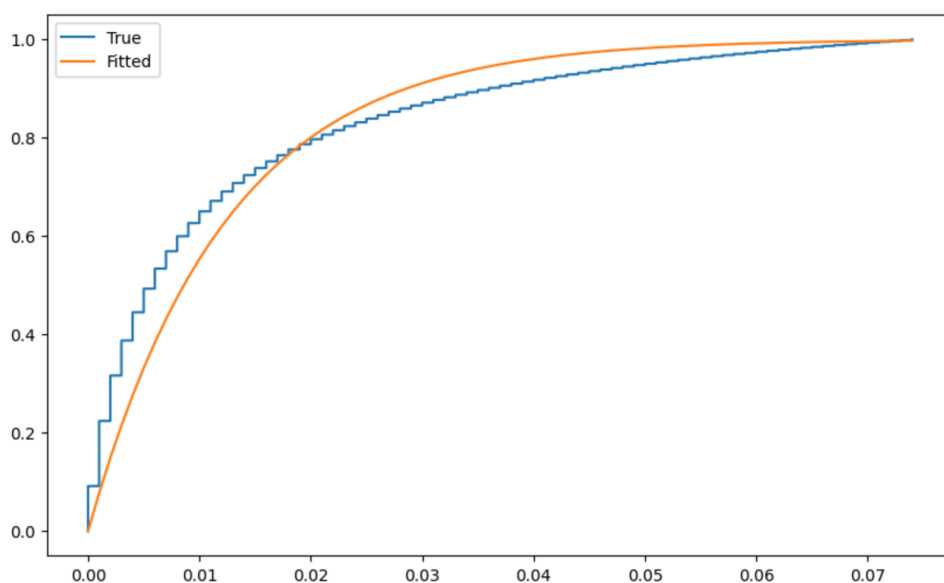


图 1: 到达时间间隔经验分布与拟合指数分布 CDF

实际成交出价分布：我们以实际成交出价为对象拟合相应的指数分布。检验结果显示，KS 检验统计量为 0.1430，p 值远小于 0.05，统计检验显著拒绝了“出价价格服从指数分布的假设”。

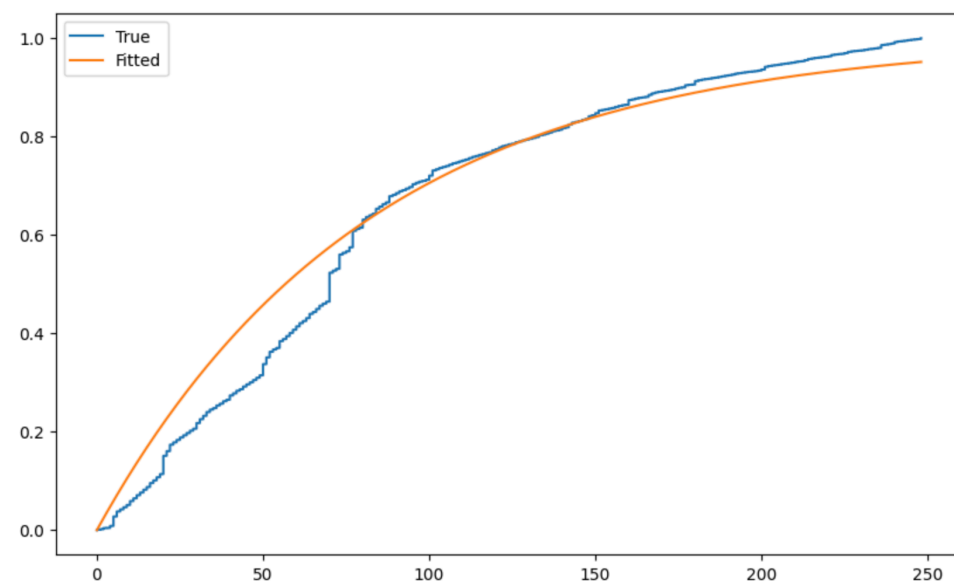


图 2: 经验分布与拟合指数分布 CDF

统计检验的结果表明在现实中观众的到达过程以及真实的出价行为具有更加复杂的统计特征（包括异常值）。然而，观察1与2中的累积分布函数对比我们可以发现指数分布确实可以描述两组数据的一些统计特征，因此我们认为论文中的假设有一定的合理性。

3.2 Base Case 稳态求解

在数值计算部分，我们尝试复现文章中的 Base Case 的稳态概率与最优策略的求解，具体实验参数设置如下：

表 2: Base Case 参数设定

| 参数 | 取值 |
|--------------------------|-----|
| 广告类型数 (N) | 3 |
| 一个广告需求的观众数量 (s) | 2 |
| 单次投放收益 (r) | 5 |
| 服务率（观众到达率）(μ) | 1 |
| 等待成本（单位时间）(c) | 0.2 |
| 广告活动到达率 (λ) | 0.2 |
| 状态容量（最大排队数）(A) | 15 |
| 到达率与服务率缩放系数 (γ) | 1 |

实验结果显示，在稳态情况下系统为空的时间占比约为 27.40%，平均利润率在 0.59 左右，结果与文献中完全一致。分析结果图可以发现，我们得到的动态策略基本上会随着系统队列长度增加相应的出价，这同样与现实逻辑相符：随着未完成的广告需求增加，系统需要更快地争取到观众来减少队列中的堆积。

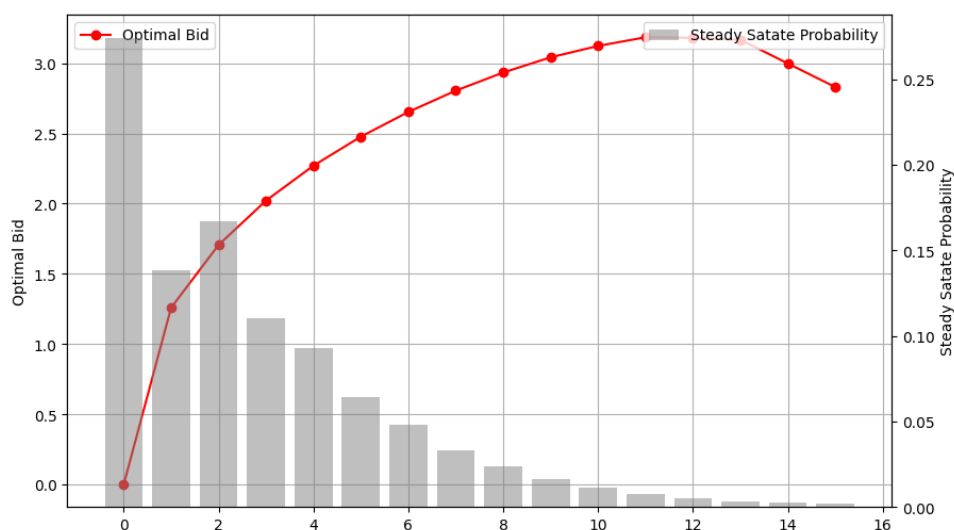


图 3: Base Case 优化求解结果

3.3 动态规划精确解与启发式近似解的对比

在论文提出的广告投放管理模型中，广告代理机构面临的关键决策包括对观众的竞价与将观众分配至不同广告活动的选择。由于全局最优策略依赖于多维状态空间与动态变化的系统参数，其数值求解复杂度极高。因此，作者提出一种启发式解法（Heuristic Policy），通过对每一广告类型独立求解简化模型，降低计算复杂度。本实验旨在量化启发式方法与一般的动态规划（Dynamic Programming, DP）最优解之间的差距，验证该近似策略在实际应用中的可行性与效率。具体来说，本实验包含以下步骤：

1. 全局最优解（General DP）求解：考虑系统中存在 N 种广告类型，每种类型具有相同的结构与参数。状态空间为所有广告类型对应的待投放广告数量（最大为 A ），在 T 的时间周期内，采用完全动态规划求解最优策略。

2. 启发式解法（Heuristic DP）求解：将每种广告类型视为独立子系统，分别求解单类型下的最优策略，并假设系统会将下一位观众分配给当前出价最高的广告类型。该方法显著降低了维度与运算量。

3. 稳态分布计算（Steady-State Distribution）：在系统稳定运行时，推导每种广告类型队列长度的长期概率分布，并据此评估上述策略在长期运行下的总期望收益。

表 3: 动态规划精确解与启发式近似解对比实验参数设定

| 参数 | 取值 |
|---------------------------|-----|
| 广告类型数 (N) | 3 |
| 一个广告需求的观众数量 (s) | 2 |
| 单次投放收益 (r) | 5 |
| 服务率（观众到达率）(μ) | 3 |
| 惩罚成本（末期未完成广告投放的成本）(C) | 1 |
| 等待成本（单位时间）(c) | 0.2 |
| 广告活动到达率 (λ) | 0.2 |
| 时间周期数 (T) | 300 |
| 状态容量（最大排队数）(A) | 5 |
| 到达率与服务率缩放系数 (γ) | 1 |

我们在3中展示实验中的相应参数。通过实验模拟计算我们得到：**全局最优解的期望收益为 377.61；启发式解法的期望收益为 334.46**。可见启发式方法获得的总期望收益较全局最优解低约 11.4%，即启发式解法在当前参数设定下损失一定收益。该差距来自于启发式方法未能动态感知整体系统状态，例如某一广告类型已临近最大容量但仍被分配观众，或系统中部分类型更迫切但未被优先投放。尽管如此，启发式策略无需在三维状态空间中进行遍历与存储，极大地减少了计算复杂度与内存需求，使其在实际工程部署中更具操作性。

接下来，我们尝试将广告类型数 (N) 由 3 更改为 2。实验结果显示：**全局最优解的期望收益为 418.64；启发式解法的期望收益为 402.37**，启发式方法的总期望损失大幅下降至 3.89%。由此可见，启发式解法相较于全局最优解的差距与广告类型数（系统复杂度）正相关。

3.4 参数敏感性分析

接下来，我们对模型进行参数敏感性分析。对于每个单独的参数，我们尝试更改其取值（未作特殊说明则其余参数取值与 Base Case 中相同），观察不同取值时的如下指标：

- 队列长度保持为 15 时，系统不同状态下，动态策略的出价
- 队列长度保持为 15 时，策略的利润率
- 搜索到最优的队列长度时，策略的最优利润率

3.4.1 参数 γ 敏感性分析

首先，我们分析缩放系数 γ 的影响，它会同时改变观众到达率与广告活动到达率。如4所示，随着 γ 从 0.5 增加至 2.5，所有状态下的最优出价均呈下降趋势，这与论文中的分析结果相符。更高的缩放系数意味着系统整体工作强度增大，广告代理商可以通过较低的出价获得更高的利润。如4中右图所示，在使用最优队列容量 A^* 时，当 γ 扩大五倍（从 0.5 到 2.5）时，单位时间的利润增长了近十倍，这体现了规模经济效应。

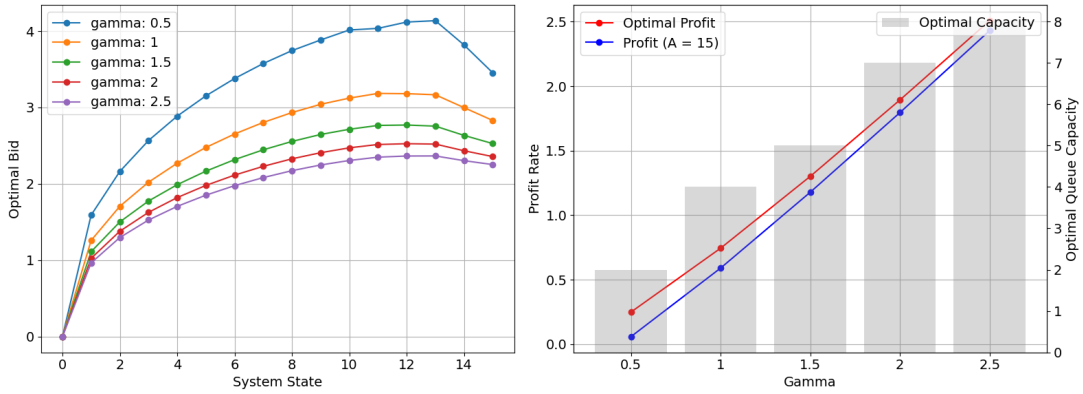


图 4: 参数 γ 敏感性分析结果

3.4.2 参数 μ 敏感性分析

接下来我们单独分析观众到达率 μ 的变化。如5所示，当观众到达率 μ 从 0.5 增加至 4 时，最优出价随之降低，这是因为观众供应增加，竞价成功的机会更多。同时，更高的观众到达率也带来了更高的最优利润率与更优的队列容量。当观众到达率较低时，主动降低队列容量可以有效防止广告需求积压，从而显著提升利润率。

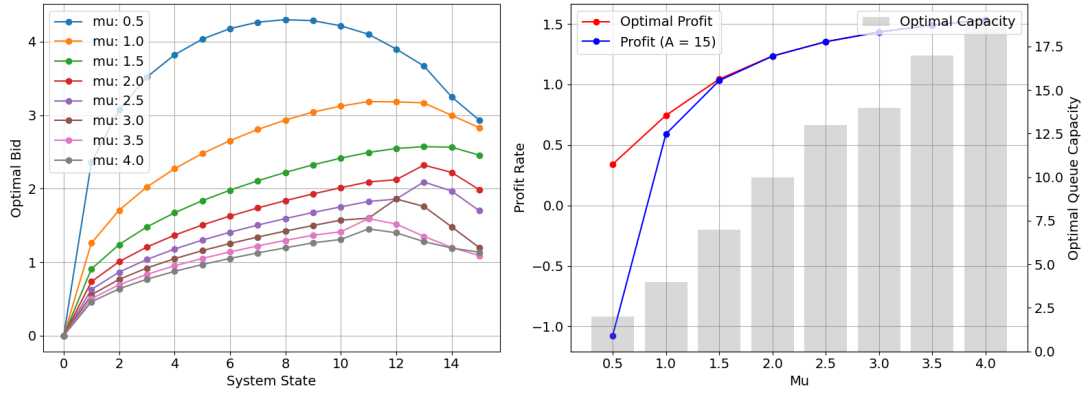


图 5: 参数 μ 敏感性分析结果

3.4.3 参数 λ 敏感性分析

广告活动到达率 λ 的变化对最优策略的影响较为复杂。如图6所示，最优出价关于队列长度的曲线依然是凹函数，但随着 λ 的增加，曲线变得更为凹陷，并出现交叉现象。这表明在广告需求旺盛且队列接近容量时，适当降低出价以利用队列容量截断新到来的广告活动，反而成为更优的策略。此外，在固定队列容量下，利润率先增后减，因为过高的需求会导致等待成本急剧上升。通过动态优化队列容量，可以显著改善盈利能力，例如当 $\lambda = 0.5$ 时，利润率提升了近 41%。

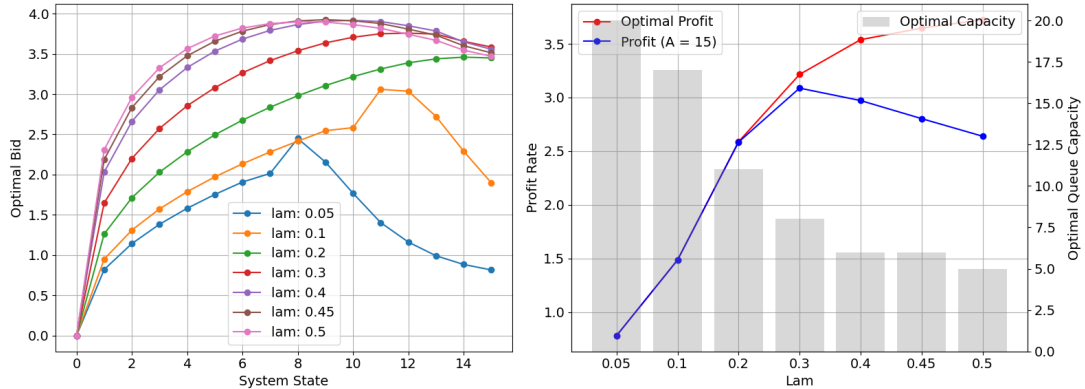


图 6: 参数 λ 敏感性分析结果

3.4.4 参数 r 敏感性分析

我们接着分析两个财务参数的影响。首先是单位观众收益 r 。如图7所示，最优出价随 r 的增加而增加，这符合理论预期，即更高的收益使得更高的出价在经济上更为合理。同时，最优利润率与最优队列容量也随 r 的增加而同步提升。

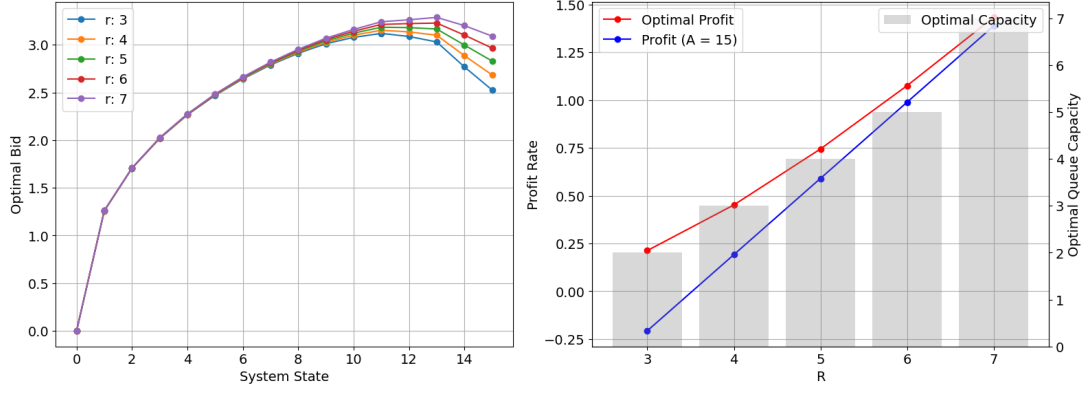


图 7: 参数 r 敏感性分析结果

3.4.5 参数 c 敏感性分析

其次是单位时间等待成本 c 。如8所示，随着 c 的增加，广告代理商会提高出价以加快广告投放，从而减少队列长度与等待成本。相应地，最优队列容量会降低，以截断或拒绝更多的新广告活动，防止成本过高。尽管总利润会因成本上升而下降，但通过优化队列容量可以显著缓解高等待成本带来的负面影响。

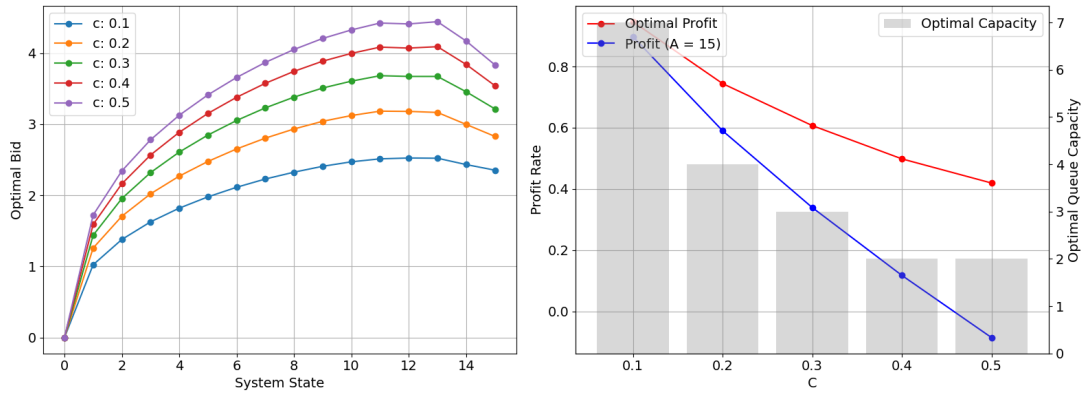


图 8: 参数 c 敏感性分析结果

3.4.6 参数 s 敏感性分析

在之前的所有分析中，我们都将 s 作为一个给定的常数，即每个广告所需的观众数量是相同的。然而，论文中的模型实际上可以处理 s 作为一个给定分布的变量的情形。这里我们尝试让 s 服从一些常见的分布（经过离散化），其余参数与 Base Case 相同，观察策略的表现与系统状态。

首先，我们假设 s 在 $1-A$ 之间等可能地取值。可以发现 s 增加（或 s 取大值的概率增加）的效果与广告活动到达率 λ 增加类似，因为两者都会提高系统的总需求率。此时队列更容易达到容量上限，导致更多的广告活动被截断，因此需要结合队列容量进行管理。

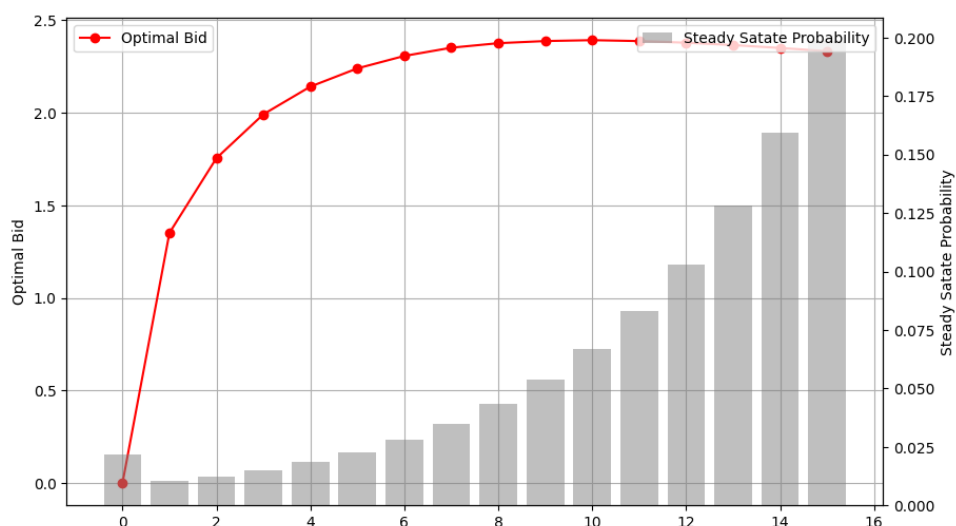


图 9: 参数 s 为均匀分布时求解结果

接下来我们假设 s 服从离散化的类正态分布，均值在 A 的一半。我们同样可以发现系统经常达到容量上限，策略的出价随队列长度增长也迅速变高。

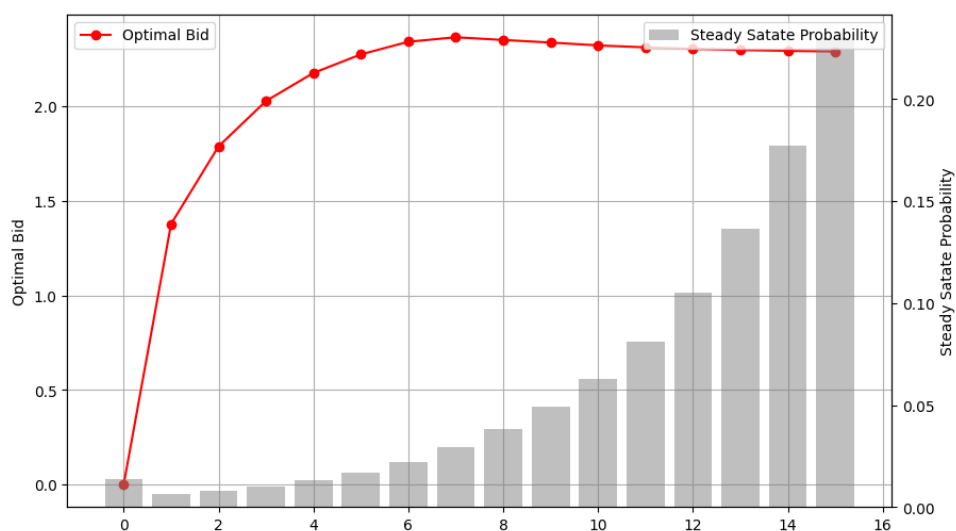


图 10: 参数 s 为类正态分布时求解结果

在此基础上，我们尝试降低其均值，发现结果发现较大改变。这说明在现实中，合理估计与考虑 s 的分布是十分重要的，因为其会在很大程度上影响系统状态分布与出价策略，从而直接影响到策略的利润率。

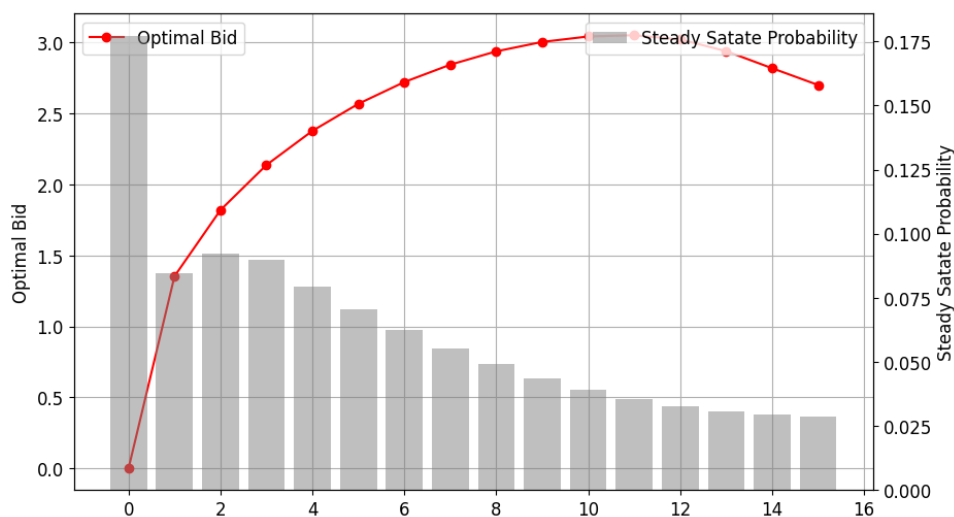


图 11: 参数 s 为类正态分布（均值较小）时求解结果

参考文献

- [1] Narendra Agrawal, Sami Najafi-Asadolahi, and Stephen A. Smith. “A Markov Decision Model for Managing Display-Advertising Campaigns”. In: *Manufacturing & Service Operations Management* 25.2 (Mar. 2023), pp. 489–507. issn: 1526-5498. DOI: [10.1287/msom.2022.1142](https://doi.org/10.1287/msom.2022.1142). URL: <https://doi.org/10.1287/msom.2022.1142>.
- [2] Weinan Zhang, Shuai Yuan, and Jun Wang. “Real-Time Bidding Benchmarking with iPinYou Dataset”. In: *CoRR* abs/1407.7073 (2014). arXiv: [1407.7073](https://arxiv.org/abs/1407.7073). URL: <http://arxiv.org/abs/1407.7073>.

附录

A 项目分工

- 吴苛铭：代码框架实现
- 龚欣怡：模型与数据分析
- 方泽宇：实验结果分析

B 项目代码

项目所用到的全部代码与数据位于该 [Github 仓库](#) 中，数据文件由于大小较大需要使用 Git LFS 拉取。