



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

互联网应用中的均衡计算及收益优化

作者姓名: 荣江

指导教师: 安波 副研究员, 陈熙霖 研究员

中国科学院计算技术研究所

学位类别: 工学博士

学科专业: 计算机科学与技术

研究所: 中国科学院计算技术研究所

二〇一八年一〇月

Equilibrium Computation and Revenue Optimization in
Internet Applications

A Dissertation Submitted to
The University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in
Computer Science and Technology

by
Jiang Rong
Dissertation Supervisor: Professor Bo An, Xilin Chen

Institute of Computing Technology
Chinese Academy of Sciences
October, 2018

声 明

我声明本论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，本论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：

日期：

论文版权使用授权书

本人授权中国科学院计算技术研究所可以保留并向国家有关部门或机构送交本论文的复印件和电子文档，允许本论文被查阅和借阅，可以将本论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本论文。

（保密论文在解密后适用本授权书。）

作者签名：

导师签名：

日期：

摘 要

近十年来,随着互联网的飞速发展,基于互联网应用的经济活动已融入了人们的生活,其中的典型代表包括付费搜索和云计算服务,二者都在工业界取得了巨大的成功,并掀起了学术界的研究热潮。已经有大量的工作研究这些市场中的模型建立、均衡求解和收益优化问题,然而,现有的工作还存在一些不足,限制了其方法在真实市场中的应用。在市场建模方面,需要保证模型在能准确描述现实世界的同时又不会因过于复杂而无法得出有意义的解决方案,已有的模型仍不完善,如缺乏对有限理性、不完全信息和动态需求等的刻画。在均衡求解方面,已有工作集中在对市场中纳什均衡的研究,然而,纳什均衡假设参与者完全理性,且知道市场中的所有信息,包括对手的私人信息,这在现实世界中很难获得,因而纳什均衡的实用性十分受限。在收益优化方面,已有工作往往没有考虑不完全信息和市场竞争等因素,因此得出的策略并非最优。针对以上这些局限,本文将以付费搜索和云计算服务为具体对象,提出更贴合实际的市场模型、更合理的均衡策略以及更好的收益优化策略。本文的主要创新点如下。

1. 本文基于付费搜索广告业务研究了有限理性均衡。本文分析了纳什均衡在实际应用中的局限性,包括只能描述纯策略行为(真实数据显示广告商大多采用混合策略)以及要求市场的完全信息,并引入质反应均衡的概念来解决这些问题。质反应均衡弱化了纳什均衡对于完全理性的要求并能模拟广告商的混合策略行为,纳什均衡只是质反应均衡的一个特例。本文首先设计了一个基于同伦思想的高效算法来求解质反应均衡,然后通过分析付费搜索中收益函数的性质开发了专门的加速算法从而将复杂度从指数级降为线性级别。此外,本文提出了利用质反应均衡进行市场参数预测的算法。实验结果表明本文提出的加速算法性能远远优于普通算法,能被用于求解大规模付费搜索中的质反应均衡,且参数预测效果好于混合策略纳什均衡模型。

2. 本文基于付费搜索广告业务研究了信息或缺下的收益优化问题,Myerson 定理给出了广告价值分布已知时收益优化的基础理论。本文考虑更加现实的情况,即广告价值分布未知且无历史数据时如何制优化收益。未知信息通过采样来估测,本文首先将如何设置采样时间来最大化有限时间内搜索引擎的总收益的问题规划成了一个整数约束优化问题,且保证最优解不依赖于具体的价值分布形式,具有鲁棒性。然后,本文通过一系列等价问题转换将原问题变换成了一个连续约束优化问题。通过分析目标函数的性质,本文证明了最优采样时间的存在性和唯一性,并基于此提出了一个具有指数收敛速度的算法。基于模拟数据的实验结果表明,本文提出的方法给搜索引擎带来收益能达到理论上最大收益的 90% 以上。

3. 本文基于云计算服务研究了着眼长远利益的定价优化问题。通过充分考虑市场

的各种特性，包括市场的发展、市场中的不完全信息及服务商之间的竞争等，本文提出了一个更为准确的基于随机博弈的模型并将最优策略描述为一个马尔科夫完美均衡。传统的均衡求解方法需要多次迭代才能收敛到近似解且时间复杂度是关于子博弈个数的指数函数。本文将随机博弈分解为多个标准型子博弈，并设计了一个基于动态规划的算法，将时间复杂度降到了子博弈个数的线性级别。此外，本文还提出了计算期望即时收益的近似加速算法，在满足近似比的同时极大地提高了计算效率。实验结果表明本文提出的算法能高效地计算马尔科夫完美均衡策略，与文献中已有的最好的策略相比，本文提出的均衡策略能给服务商带来最高的收益。

4. 本文基于云计算服务研究了动态供需优化问题。通过对市场中的动态供需关系建模，本文提出了一个能够描述资源利用情况和市场竞争的模型。本文分析了纳什均衡在实际中无法达成的原因，并将定价策略定义成了一个现实中更加实用的近似均衡。本文为此开发了一个保证收敛的算法，还推导了均衡的很多重要的性质，包括动态价格关于资源利用率的单调性和收益函数的凹凸性等，并利用这些性质进一步提高了算法效率。本文通过实验表明所提的近似均衡策略优于其他已有的动态价格策略。

综上所述，本文提出了新的建模方式以弥补现有工作中的不足，并设计了高效的算法来解决对应的问题。实验证明本文的模型和算法取得了比现有研究成果更好的效果。

关键词： 互联网经济，广告竞标，云计算，均衡计算，收益优化

Abstract

Internet-based applications have become part of people's daily life due to rapid development of the Internet in the last decade, where sponsored search auction and cloud computing service are two typical representatives. Both of them have achieved great success in the industry and attracted much attention from the literature. There have been plenty of works investigating the modeling, equilibrium computation and revenue optimization problems in the market, which, however, have many shortcomings and the application of their methods to the real world is limited. First, the market modeling needs to trade off between the complexity and accuracy of the model, while existing models are not perfect, e.g., lack of consideration of bounded rationality, incomplete information and dynamic demand. Second, previous works on equilibrium analysis focus on Nash equilibrium, which, however, is impractical in the real market because it requires the players to be perfectly rational and to know all information in the market, including competitors' private information. Third, the works on revenue optimization usually ignore incomplete information, competition or dynamic demand in the market and hence their methods are not optimal. To address the above problems, this thesis proposes more comprehensive market models, more practical equilibrium concepts and better strategies based on sponsored search auctions and cloud computing service. The main contributions of this thesis are shown below.

1. This thesis investigates bounded rationality in sponsored search auctions (SSAs). The limitations of Nash equilibrium in the real world is analyzed, including the inability of modeling mixed strategy (real data indicate that most advertisers are using mixed strategies) and requirement of complete information, and the Quantal Response Equilibrium (QRE) is introduced to address these problems in this thesis. QRE does not assume perfect rationality and can model mixed strategy. Nash equilibrium is just a special case of QRE. This thesis first designs a homotopy-based algorithm to compute the QRE and then further reduces the complexity of the algorithm from exponential order to linearity by utilizing the structure property of revenue functions in SSAs. Besides, this thesis develops a parameter estimation algorithm through fitting the QRE model into real data. Experimental results show that the speed-up algorithm is much more efficient, which can be used to compute the QRE in large-size SSAs, and the accuracy of the estimation algorithm is higher than that of mixed-strategy Nash equilibrium.

2. This thesis studies the revenue optimization problem of SSAs with unknown information. Myerson theorem provides the foundation of revenue optimization with known advertise-

ment value distribution. This thesis considers more practical situations, i.e., the distribution is unknown and there is not historical data, to know which a sampling procedure is necessary. This thesis first formulates the problem of setting proper sample size to maximize search engine's overall revenue in finite horizon as a constrained integer optimization, the solution of which is robust and does not depend on the distribution. Then the original optimization problem is converted into an equivalent constrained continuous format. This thesis proves the existence and uniqueness of the optimal sample size through analyzing the property of the objective function, and designs an efficient algorithm with exponential convergence. The experiments show that the revenue by using methods proposed by this thesis can achieve 90% of the maximal revenue in theory.

3. This thesis investigates the long-term revenue optimization problem for cloud computing service. By considering characteristics of the cloud market, including the increasing number of users, incomplete information and competition between providers. This thesis proposes a more comprehensive stochastic-game-based model and formulates the optimal strategy as a Markov Perfect Equilibrium (MPE). The stochastic game can be divided into many normal-form sub-games. Traditional methods need to repeatedly update state values to converge to a near-optimal solution and are of exponential complexity with respect to the number of sub-games. This thesis designs a dynamic-programming-based algorithm, which reduces the complexity to linear order. Besides, an approximate algorithm is proposed to compute the expected immediate profit, which shows high approximation ratio and can significantly improve the efficiency of the algorithm. Experimental results indicate that the algorithms in this thesis can efficiently compute the MPE and compared with state-of-the-art methods, the algorithms can achieve the highest revenue for providers.

4. This thesis studies the dynamic demand problem of cloud computing service. Based on the analysis of supply and demand, this thesis proposes a model to describe resource utilization in competitive cloud market. The unreachable property of Nash equilibrium is demonstrated in this thesis and then the pricing strategy is defined as a more practical approximate equilibrium. This thesis designs an algorithm which is ensured to converge and derives many properties of the approximate equilibrium, including the monotonicity of dynamic pricing and convexity of the revenue function, and use them to further improve the efficiency of the algorithm. Experiments show that the proposed equilibrium strategy outperforms existing dynamic pricing strategies.

In summary, this thesis develops new models to cover the shortages of previous works and designs efficient algorithms to solve the corresponding problems. Experimental results indicate that the proposed methods get better performance than existing state-of-the-art achievements in the literature.

Keywords: Internet Economy; Sponsored Search Auctions; Cloud Computing; Equilibrium Computation; Revenue Optimization

目 录

摘 要	I
目 录	VII
图目录	XI
表目录	XIII
第一章 绪论	1
1.1 研究的背景及意义	1
1.2 研究的问题和挑战	3
1.3 研究现状	4
1.3.1 付费搜索广告竞标研究现状	4
1.3.2 云计算服务研究现状	7
1.4 本文主要贡献	8
1.5 本文组织结构	10
第二章 有限理性模型研究	11
2.1 引言	11
2.2 质反应均衡	12
2.2.1 研究动机	12
2.2.2 基本概念及假设	13
2.2.3 QRE 的定义	13
2.3 均衡计算	15
2.3.1 基于同伦变换的算法	15
2.3.2 针对 SSA 的高效算法	18
2.4 参数预测算法	21
2.5 实验评估	23
2.5.1 QRE 求解算法的性能评估	24
2.5.2 参数预测算法评估	25
2.5.3 QRE 模型对真实数据的解释能力评估	26

2.6 总结	29
第三章 基于保留价的收益优化	31
3.1 引言	31
3.2 背景知识	33
3.2.1 GSP 机制	33
3.2.2 权重价值已知时最优保留价的计算	34
3.3 问题定义	35
3.4 竞争比下界的推导	35
3.5 采样数优化	39
3.6 算法设计与分析	43
3.7 实验评估	44
3.8 总结	46
第四章 马尔科夫完美均衡和长远利益优化	49
4.1 引言	49
4.2 建模	50
4.2.1 不完全信息建模	50
4.2.2 市场发展建模	52
4.2.3 市场竞争建模	52
4.3 计算均衡策略	54
4.3.1 问题定义	55
4.3.2 下一状态已知时计算 $\Phi(s)$ 的 NE	56
4.3.3 计算整个随机博弈的 MPE	58
4.3.4 即时收益计算	61
4.3.5 近似算法	63
4.4 实验评估	65
4.4.1 算法评估	65
4.4.2 MPE 策略评估	67
4.5 总结	71

第五章 近似均衡和动态供需优化	73
5.1 引言	73
5.2 竞争性的市场和动态需求建模	74
5.2.1 研究动机	74
5.2.2 动态需求建模	74
5.2.3 多服务商模型	76
5.3 最优定价策略	77
5.3.1 均衡策略	77
5.3.2 均衡计算	80
5.3.3 计算最优响应	81
5.4 重要性质	82
5.5 实验评估	86
5.5.1 运行时间评估	87
5.5.2 资源总数对 AE 的影响	87
5.5.3 需求函数参数对 AE 的影响	88
5.5.4 忽视市场竞争的收益损失	88
5.5.5 收益的提升	89
5.5.6 AE 与纳什均衡的距离	90
5.6 总结	90
第六章 总结与展望	93
6.1 本文总结	93
6.2 下一步工作展望	94
参考文献	97
致 谢	107
作者简介	109

图目录

图 2.1	TM、AGG 和 AGGSU 的运行时间比较.....	24
图 2.2	参数估测算法性能评估.....	26
图 3.1	权重价值的概率密度函数及虚拟权重价值函数.....	45
图 4.1	策略向量图.....	56
图 4.2	用户数量增长曲线.....	65
图 4.3	近似公式评估.....	66
图 4.4	运行时间评估.....	67
图 4.5	MPE 策略.....	69
图 4.6	长期收益比较.....	70
图 5.1	资源总数对 AE 的影响.....	88
图 5.2	策略比较.....	89
图 5.3	不同参数下的 AE.....	90

表目录

表 2.1	QRE 求解算法性能评估	25
表 2.2	QRE 拟合精度	27
表 2.3	用 QRE 模型估测的参数	27
表 2.4	关键词 1 中广告商的质反应及收益	28
表 2.5	关键词 2 中广告商的质反应及收益	28
表 2.6	用 MSNE 模型估测的参数	29
表 3.1	针对对数正态分布的性能评估 ($N = 5$)。	45
表 3.2	针对对数正态分布的性能评估 ($N = 500$)。	46
表 3.3	针对指数分布的性能评估 ($N = 5$)。	46
表 3.4	针对指数分布的性能评估 ($N = 500$)。	47
表 4.1	长远利益符号总结	53
表 4.2	算法 4 的收敛性	67
表 4.3	评估用场景	68
表 4.4	服务商策略不同时的收益比较	70
表 5.1	动态供需符号总结	78
表 5.2	运行时间比较	87
表 5.3	需求函数参数对 AE 的影响	89
表 5.4	评估 ϵ	90

第一章 绪论

1.1 研究的背景及意义

近十年来，基于互联网的经济活动已慢慢融入人们的日常工作和生活，其中的典型代表包括还在快速发展期的付费搜索广告业务和云计算服务，它们已经成为了人们（尤其是科技工作者）重要的工具。基于付费搜索的广告竞标业务和基于云计算的服务为工业界创造了巨大的商业价值，同时也掀起了了学术界的研究热潮。

根据美国交互广告署（Interactive Advertising Bureau）的报道，2012 年全球的在线广告收入达到了 365 亿美元，这个数字相比于 2011 年的 314 亿美元上涨了 15%。作为在线广告的一种主要形式（其他形式的有如实时竞价广告（Real Time Bidding）），付费搜索广告业务在 2012 年的收入占到了全部在线广告业务收入的 46.3%（即 169 亿美元）[1]。而根据 Smart Insight 的报告，2016 年移动端和桌面端的付费搜索广告的总收益达到了约 2000 亿美元^①。鉴于商业搜索引擎（如谷歌（Google），必应（Bing），百度（Baidu））的巨大成功，以及搜索的庞大用户量，会有越来越多的公司使用付费搜索广告业务。这种广告服务通常由三部分组成：第一部分与广告商交互，收集他们要竞价的关键词组以及他们的出价等信息，为每个关键词组设置保留价格；第二部分则与用户交互，通过用户搜索的关键词组，选出一部分相关的参与竞价的广告，然后按照一定规则将它们排序，最终显示在搜索页面上；第三部分负责向广告商收取相关费用，常用的收费制度包括：1）按显示收费（Pay-Per-Impression, PPM），即每当广告商的广告被显示在搜索页面上则收取一定费用，2）按点击收费（Pay-Per-Click, PPC），即用户每次点击一个广告则向其投放者收取费用，3）按行为收费（Pay-Per-Action, PPA），即只有用户在通过广告链接跳转到指定页面并完成指定动作以后才向广告商收费。总体而言，PPM 模式对搜索引擎更有利，而 PPA 模式更受广告商青睐，作为折中，目前大部分商业搜索引擎使用 PPC 的收费模式。哪些广告被显示、如何对被选中的广告排序以及如何向广告商收费是工业界和学术界研究的核心问题，搜索引擎通常通过一个竞价机制完成这些任务。竞价机制也经历过一些变迁，最初 Overtrue 引入了第一价格竞标（Generalized First-Price Auction, GFP），即收取的费用等于广告商出的竞标价，这种机制随后被证明不能形成一个有效且稳定的纳什均衡，随后谷歌提出了广义第二价格竞标（Generalized Second-Price Auction, GSP），即对一个广告商收取的费用等于排在其后一位的广告商出的竞价。目前 GSP 已经主导了整个付费搜索市场，几乎所有的商业搜索引擎都在用 GSP。

^① <https://www.smartinsights.com/search-engine-marketing/search-engine-statistics/>

付费搜索是互联网经济的一个典型代表,近些年新兴的云计算服务则是另一个成功典范。谷歌和亚马逊大约于 2006 年相继推出了自己的云服务,随后云计算行业经历了飞速发展的十年。一个云平台通常包含能向用户租出的设备、平台和软件(应用)等。在工业界,云服务根据用户选择的服务形式具体可划分为设备即服务(Infrastructure as a Service, IaaS)、平台即服务(Platform as a Service, PaaS)和应用即服务(Software as a Service, SaaS) [2]。其中 IaaS 提供实体机器(如 CPU, GPU, 内存)、虚拟机(如 VMWare, Xen, HyperV)以及一些管理工具(如 IBM Tivoli, Racwise); PaaS 提供一些平台架构的服务,如 Java, .Net, Python 等,典型代表包括 Microsoft Azure 和 Google AppEngine; SaaS 则主要用于构建商业程序、网页应用和多媒体等,这方面 Google Apps、Facebook 和 Youtube 很具代表性。云计算服务使得人们从安装、配置、更新和维护各种软件和硬件的繁重体力劳动中解放了出来,能够将主要精力放在待解决的问题上,因而越来越多的人开始使用云服务。创建伊始,云计算的用户主要集中在企业级别,随着硬件成本的降低和服务的多样化,个人用户数量呈现出不断增长的趋势。国际数据公司(International Data Corporation)的一份市场研究报告指出,云服务的商业产值从 2008 年的 160 亿美元增长到了 2012 年的 420 亿美元 [3]。根据市场研究媒体(Market Research Media)的预测,云服务将保持 30% 的复合年增长率(Compound Annual Growth Rate, CAGR),其全球总产值将于 2020 年达到 2700 亿美元^①。云计算的蓬勃发展吸引着更多的公司加入这个行业,如今比较成功的有如谷歌的 GCP^②, 亚马逊的 AWS^③, 微软的 Azure^④ 等,市场竞争日趋激烈。

付费搜索广告和云计算在工业界取得巨大成功的同时,也引起了学术界的关注。广大学者对这两种经济模式进行了深入的研究 [4,5], 包括对参与者行为模型的研究 [6-8], 价格优化 [9-11], 机制设计 [12-14] 等。在这些研究当中,博弈论是最为常用的工具,因为其能够有效地分析、预测和控制利己主义个体的行为,从而为服务提供商和用户的决策提供有用的帮助。博弈论是一门用数学模型研究理性的智能体之间的对抗和协作问题的学科 [15,16], 和经济学一直有密不可分的关联 [17,18], 这种关联可以追溯到 1944 年《Theory of Games and Economic Behavior》 [19] 一书的出版,此后博弈论被广泛应用于研究各种经济市场的各种问题,例如竞拍 [20,21]、均衡分析 [22,23] 和公平分配 [24,25] 等。此外,博弈论也是安全领域 [26,27] 和计算机科学 [28,29] 的主要研究课题之一。博弈论模型在现实世界的应用很多,如美国海岸警卫队的 PROTECT 系统 [30] 和洛杉矶机场的 TRUSTS 系统 [31]。付费搜索广告业务和云计算服务中的参与者(包括服务提供商和用户)往往都是理性和利己主义的,即他们的目标是最大化自己的收益,因此博弈

① <https://www.marketresearchmedia.com/?p=839>

② <https://cloud.google.com/>

③ <https://aws.amazon.com/>

④ <https://azure.microsoft.com/>

论是最为有效的研究工具之一。然而，目前的研究成果还有很多不足之处，如对参与者个人的理性以及对对手的建模还不十分准确，对市场各种因素的考虑还不十分周全，从而导致服务商最终的定价并非最优，这将导致大量的收益损失。本文旨在用博弈论的方法，以付费搜索广告业务和云计算服务为研究模型，来解决互联网经济活动中一些关键问题，如有限理性、定价、均衡分析、机制优化等，提出更准确的市场模型，帮助服务商在激烈的市场竞争中寻找更优的策略，制定更合理的价格，从而进一步提高其收益。

付费搜索广告和云计算服务分别代表了互联网经济活动中的两类主要形式。其中付费搜索领域由于谷歌占了全球搜索市场 80% 左右的份额，该市场呈现出了垄断的形式，于是对付费搜索服务的研究主要集中在对某一特定搜索引擎的广告商行为的建模，以及对该搜索引擎的收益优化。而云计算领域由于不同服务商提供的云计算资源之间的差别较小，提供商之间的竞争比较激烈。此外，付费搜索和云计算服务的另一个区别在于付费搜索提供的服务量（即广告位的数量）通常远远小于用户需求，而在云计算平台经常会出现计算资源闲置的现象，因此优化库存与需求之间的匹配程度是云计算领域要考虑的一个问题。因此，付费搜索服务代表了垄断且供小于求的一类经济活动，如各种拍卖和竞标活动，而云计算服务代表了服务商之间竞争激烈且供需动态变化的一类经济活动，如租车和酒店预订服务等。因此对这两者进行研究得到的成果具有可拓展性，能够应用到其他相关领域。

1.2 研究的问题和挑战

本文以付费搜索广告和云计算作为具体研究对象，利用博弈论的方法来分析以它们为代表的互联网经济活动中的模型建立、均衡计算和收益优化问题。由于纳什均衡 [32-34] 是博弈论中最基本的研究问题，在各个经济活动领域中都有对纳什均衡的研究，主要包括纳什均衡的求解和性质分析。收益优化问题往往与均衡分析不可分割，因为收益的计算依赖于对模型中各个参与者行为的建模，如果参与者都是理性的，那么其策略很有可能会构成一个均衡。下面介绍在付费搜索和云计算服务的研究中要解决的主要问题和面临的挑战。

市场建模。在付费搜索服务中，目前主流的搜索引擎都在使用 GSP 机制，该机制本身并不复杂且已相对稳定，因此对市场建模的难点主要来自于对用户（即广告商）行为的预测。GSP 机制中的大部分参数对广告商来说都是未知的，如参与竞标的总人数和广告的点击率等。在现实世界中，广告商获取这些知识的能力是不同的，在博弈论中，这种能力被称为理性程度。因此，对广告商的建模需要考虑这些因素。在对广告商建立合理模型的基础上，搜索引擎才能通过调整机制中的参数来获取更多的收益。而现在已经有的工作为了易于分析，通常认为所有广告商都是完全理性的，如何在弱化这个假设的同时还能获得有意义的结果是本文面临的挑战。不同于付费搜索已经相对成熟的竞价机制，云计算平台目前尚未形成一套公认的定价体系。由于用户都是利益驱动型的，云计

算服务提供商之间必须通过设置合理的价格来争取更多的用户。制定价格时需要同时考虑用户黏性、需求对价格的敏感性、维护平台资源的费用以及当前价格对未来收益的影响等等。一方面，模型越能贴合现实，其复杂度往往会越高，从而求解最优定价策略就会更难；另一方面，模型越简单，其对真实世界的解释能力就越弱，但求解起来就会相对简单一些。因此，如何平衡云计算市场模型的复杂度和可解性是本文面临的另一挑战。

均衡求解。本文要研究的第二个问题是付费搜索市场和云计算市场中的均衡的求解与分析。目前已有很多相关工作研究两个市场中的纳什均衡。然而，达到纳什均衡的条件是参与者知道这个市场中的所有信息，包括对手的私有信息。这种条件在现实世界中是无法满足的。例如付费广告中对手广告的价格，云计算中用户的喜好和需求量以及对手的资源数量等。因此，纳什均衡通常只限于理论研究，而实际应用中参与者之间往往达不到纳什均衡。本文要深入研究付费搜索和云计算市场的特性，旨在找到具有实际指导意义的均衡概念，这当中面临的挑战是如何在不完全信息的条件下获得有意义的策略。此外，考虑到两个市场中的用户量都很大，本文还要设计高效的算法来以求解实际规模的均衡策略。

收益优化。收益的优化建立市场模型和均衡策略的基础上的。在付费搜索广告竞标中，搜索引擎可以通过调整保留价格来优化自己的收益。已有的工作提供了当广告的价值已知时求解最优保留价的方法，然而，在现实世界中广告的价值是广告商的私人信息，搜索引擎无法得知。因此在广告价值未知的情况下优化收益是本文面临的又一挑战。在云计算领域，计算资源的定价会影响用户的需求量，一般而言，价格越低越能吸引更多的用户，但是单位资源带来的收益会降低，反之，价格越高用户越少，但单位资源收益会越高，因此，收益优化的挑战在于如何根据资源的使用率动态调整价格，使得期望收益最大化。此外，考虑到不断增加的用户量和不断降低的硬件成本，现实世界中云资源的价格有不断降低的趋势 [35]，如何制定一个考虑长运利益的价格策略也是本文要研究的问题。

1.3 研究现状

第1.1章介绍了付费搜索广告和云计算服务在工业界取得的巨大经济效益，本章分为两个部分，分别阐述付费搜索广告竞标和云计算服务在学术界的研究现状。

1.3.1 付费搜索广告竞标研究现状

国内外的科研人员对付费搜索广告进行了多方面多角度的研究，研究课题的侧重点也有很大差别，总结起来主要的方向有关键词推荐 [36–38]，词条扩展和广告选择 [39–41]，点击率预测 [42–44]，有限预算 [45–47]，机制设计 [48–50] 等。Broder 等人在斯坦

福大学开设了一门课程—《Introduction to Computational Advertising》^①，专门研究付费搜索广告。

付费搜索广告竞标（Sponsored Search Auction）最早是 1997 年由 Overtrue 引入的，其当时使用的是 GFP[51–53]。随后，学者们发现 GFP 是不稳定的 [51,53]（即不存在纳什均衡），这对于广告商和搜索引擎来说都是不好的性质。之后，谷歌提出了 GSP，有效解决了这个问题。从此，GSP 在工业界成为了付费搜索广告领域的黄金标准，绝大多数的商业搜索引擎都采用 GSP 作为他们的竞价机制。同时，学术界的主要研究对象也集中在 GSP，来自不同行业的学者研究了 GSP 的各种性质以及如何进一步改善 GSP 以使得竞价的机制更优。从这些研究中诞生了很多关于机制优化和博弈分析的学术成果。本文的研究重点是基于 GSP 的均衡分析和机制设计，关于付费搜索广告的其他方向研究成果，读者可参考综述类文献 [1,4,54]。

1.3.1.1 均衡分析

广告商的广告每次被点击，其便有一定的（期望意义上的）利润，通常将这个利润称做广告的价值。如果一个广告商出的竞价等于他的广告的价值，则他的行为被称为真实出价。从博弈论的角度看，搜索引擎希望能有一个机制使得广告商的最优定价便是他们广告的价值，因为这样就避免了广告商干预竞标机制的可能性（理性的人总是选择最优的策略）。然而，GSP 并不能保证真实出价，也就是说，真实出价不是一个纳什均衡，出价低于广告价值对广告商来说是一个更好的策略 [20,55]。正因为如此，学术界提出了各种非真实出价的均衡的概念来描述广告商的行为，并深入研究了各种均衡的性质。Border 等人 [56] 证明了在 GSP 中可能存在多个纳什均衡。Edelman 等人 [20] 提出了局部无嫉妒均衡（Locally Envy-Free Equilibrium, LEFE）的概念，在这种均衡条件下，任何一个广告商都不能通过交换他和他前面一位的位置来获得更高的收益。LEFE 被 Edelman 证明了是一个稳定的分配。在所有 LEFE 中，搜索引擎的收益最低的被称作最低收益无嫉妒（Lowest-Revenue Envy-Free, LREF）均衡，Edelman 证明了当设定一个合适的保留价时，广告商不能达到一个比 LREF 收益更高的均衡。Cary 等人 [57] 证实了如果广告商的策略是采取当下最优出价（即贪婪策略），则整个系统最终会收敛到 LREF。Varian [55] 提出了对称纳什均衡（Symmetric Nash Equilibrium, SNE），该均衡的定义是任何一对广告商都不能通过交换位置来提高收益，同时，Varian 证明了在 GSP 中存在一个 SNE，其收益是所有可能的纳什均衡中最高的。可以验证得出 LEFE 和 SNE 本质上是相同的，二者均是完全信息下的纳什均衡的一个子集。在不完全信息的设定下，广告商的广告价值被假设成服从一个概率分布，且该分布是公开的信息。Gomes 和 Sweeney [58] 给出了在不完全信息的情况下，GSP 中有效的贝叶斯纳什均衡存在的充要条件。此外，一些研究人员还提出了前瞻纳什均衡（Forward Looking Nash Equilibrium,

^① <http://www.stanford.edu/class/msande239/>

FLNE) 的概念 [59,60]。除了研究均衡本身的性质, 还有很多研究工作关注广告商如何达到均衡状态。Cary 等人 [57] 提出了一种特殊的最优响应策略, 被称作平衡竞价策略 (Balanced Bidding Strategy), 该策略本质上等价于 [61] 中提出的 FLNE, 最终收敛到一个稳定点使得搜索引擎的收益等于 Vickrey-Clarke-Groves (VCG) 机制 [20,55] 的均衡, 作者还给出了收敛速度。Vorobeychik 和 Reeves 通过仿真实验说明了平衡竞价策略是非常稳定的, 即广告商不采取这种策略所能得到的额外收益很小 [62]。Nisan 等人 [63] 证明了在 GSP 中, 如果其他广告商都采取最优响应的策略, 那么自己的最好的策略也只能是最优响应, 从而解释了为什么最优响应是个合理的假设。

以上这些研究工作有个共同的不足是都假设广告商是完全理性的, 即他们知道所有信息并能够精确计算出每个竞价的收益, 从而总是能够找到最优响应的策略, 然而, 在现实世界中, 这种强假设很难成立, 由于广告商对收益的计算以及对对手策略的估计都是有误差的 [8,64], 因此广告商并不知道自己的最优响应策略是什么。本文将对 GSP 中非完全理性条件下的均衡进行深入研究。

1.3.1.2 机制设计

机制设计是搜索引擎关注的核心问题, 对于搜索引擎来说一个好的竞价机制应该具备以下性质: 1, 机制简单, 容易被广告商理解; 2, 广告商的最优策略形成一个均衡点从而保证他们的行为是可预测的; 3, 可以给搜索引擎带来较高的收益。VCG 的最好的性质是真实出价是纳什均衡, 然而 VCG 本身的收费机制过于复杂 (一个广告商支付的费用等于由于他的出现导致的其他广告商损失的利润), 而且当广告商出价相同时, GSP 带给搜索引擎的收益要高于 VCG 的, 因此现在搜索引擎很少再用 VCG。大部分研究工作集中在设计一个更好的 GSP 机制, 其中保留价的设定是一个重要的研究方向。在完全信息的条件下, Edelman 等人 [20] (Varian [55]) 研究了 GSP 在 LEFE (SNE) 下的收益, 他们指出 LEFE (SNE) 在最坏的情况下的收益也要高于 VCG 的均衡收益。Lucier 等人 [65] 研究了 GSP 中其他形式纳什均衡的收益, 分析指出搜索引擎使用 GSP 的收益可能比使用 VCG 低很多, 然而如果去除排在第一位的广告商的费用, 则 GSP 带来的收益最差情况下等于 VCG 的一半。Lucier 等人进一步证明了上述的界限是紧的, 而且不受保留价格的影响。不完全信息假设搜索引擎并不明确知道每个广告商的广告价值, 但知道它们的价值服从一定的分布。著名的关于最优竞价机制设计的 Myerson 定理 [66-68] 给出了如何在这种设定下优化保留价从而最大化搜索引擎的收益的基本思路: 当只有一个广告位且广告商的广告价值满足独立同分布时, 最优的竞价机制是将保留价格设置为虚拟价值函数零点的第二价格机制。Edelman 和 Schwarz [12] 将这个理论扩展到了更加通用的情况, 即搜索页面有多个广告位, 并证明了如果虚拟价格函数满足正则性质, 则使用最优保留价格的 GSP 是最优的机制。Thompson 和 Leyton-Brown [14] 研究了当 squashing 也是考虑因素时如何优化保留价。Sun 等人 [69] 经过研究发现, 当广告的权重

不同时，GSP 中的保留价依赖于具体的广告权重。

上述所有关于非完全信息下机制优化的工作都假设搜索引擎知道广告商的广告价值的分布，在此基础上计算最优保留价格。然而，在现实世界中，广告的价值是广告商的私人信息，搜索引擎必须通过一定的采样来收集分布信息，然后才能计算最优保留价。Ostrovsky 和 Schwarz [9] 利用雅虎 (Yahoo) 的付费搜索广告数据集做了如下实验: 他们假设广告的价值服从一个对数正态分布，然后用仿真的方式推测分布的参数，再用推测的参数计算最优保留价。实验结果表明由此得出的保留价可以有效提高搜索引擎的收益。Sun 等人 [70] 提出了用最大似然估计的方法来预估分布参数。Cole 和 Roughgarden [71] 则直接用采样的数据来计算最优保留价，而不用先估计分布参数，但他们的方法是基于一种比较简单的机制，即所有获得广告位的广告商支付同样的费用。以上工作研究了如何通过采样来计算最优保留价，却没有考虑由于采样过程使用的保留价并非最优而导致的收益损失。显然，当考虑有限时间段内的收益时，花越多的时间采样对最优保留价的估计越准，同时采样阶段的收益损失也越多，反之，采样时间越短带来的前期收益损失越小，但对最优保留价的估计也越不准，后期使用估测的保留价的收益损失就越大。因此，搜索引擎必须权衡两方面的收益，决定用多少时间来采样以最大化全局收益，即必须找一个最优采样数。Bulow 和 Klemperer [72] 提出了一些计算采样阶段收益损失的基本方法，但他们的模型与文献 [71,73] 中的一样过于简单，无法应用到 GSP 中。本文将深入研究 GSP 中采样数和保留价优化问题。

1.3.2 云计算服务研究现状

在如今的信息化和大数据时代，无论是工业界还是学术界都需要大量的计算资源来处理信息和数据，传统的本地式的计算资源已远远不能满足人们的需求，云计算服务应运而生。云计算不仅在工业界发展迅猛，也成为了学术界热门的研究课题，这些研究工作大体上可以分为以下几个大类: 1, 能源管理，设备是云计算服务的核心部件，一个云计算中心往往包含数百万的服务器、存储和网络设备，这将消耗大量的能源，例如亚马逊的云计算服务的总预算的 42% 是与能源相关的（这其中包括 19% 的供电设备和 23% 的冷却设备），关于这方面的研究工作读者可参考 [2,5]；2, 云计算性能，这部分工作主要考虑云计算的性能的评估和优化，例如规划工作流和平衡负载等 [74,75]；3, 数据管理，这个课题主要研究大规模和分布式数据在云计算中的管理，如数据冗余处理 [76]、数据挖掘 [77,78] 和分布式数据集成 [79] 等；4, 数据安全，这一直是云计算比较关心的问题，目前已经有很多相关的研究工作 [80,81]；5, 定价机制，该方向主要从经济学的角度来分析如何制定合理的价格（动态的或固定的）来使云服务提供商的收益最大化 [11,82]。本文着重于利用博弈论的相关理论来研究云计算中 IaaS 模式的定价问题，寻找在不断变化和竞争的市场中合理的价格策略，帮助服务提供商更高效地利用自己的资源及提高收益。

接下来本文将从两个角度来分析和研究最优定价的相关文献。第一，基于长期收益的价格优化。考虑到云计算市场是不断演化的，即用户量再不断增大，同时边际成本在不断减小（由于硬件和能源成本逐渐降低而导致），如何制定随时间变化的价格从而最大化长远收益是一个值得研究的问题。Feng 等人 [82] 提出了一个非协作的竞争模型，该模型计算服务商在不考虑长期利益下的均衡策略。有些工作研究了服务商随时间变化的最优价格 [83,84]，然而市场中只有一个服务商的假设与现实世界不符，这些模型的实用性不高。一些学者侧重于研究用户在不段演化的市场中的行为，却忽略了多个服务商之间的竞争行为 [10,85]。Xu 等人 [11] 提出的模型虽然考虑了用户的行为和市场竞争，但模型过于简单，即他们认为市场中只有一个自主优化价格的服务商，其他服务商只是简单地模仿那个自主定价服务商的策略，这种模型并不适用于今天的多巨头的云计算市场，因为每个大的厂商都不是简单地模仿他人的策略。文献 [86] 中的模型虽然考虑到了市场竞争，但他们的算法只能应用到含有几个用户的市场，无法应对现实世界中庞大的用户量。Truong-Huu 和 Tham [87] 用一个离散选择模型来描述用户的行为，但忽略了服务商并不知道用户的私人信息这个事实。本文将提出一个更加准确的模型来模拟真实的云计算市场，将市场的发展变化、市场竞争、用户的不完全信息这些因素都考虑进来，然后给出一个基于时间变化的最优价格策略来最大化服务商的长期收益。

第二，基于设备利用率的价格优化。在云计算市场发展的每一个阶段，服务商的设备利用率是变化的，由于用户往往都是对价格很敏感的，随着利用率的变化动态调整价格是一个很重要的策略。动态定价在传统的一些服务上起到了重要作用，如汽车租赁 [88] 和酒店预订 [88,89]。相比于传统服务业，云计算领域的竞争更加激烈。由于云计算属于在线服务，用户可以随时选择任何服务商，因而商家面对的对手更多，竞争更加激烈，而传统行业的用户一般就近选择服务商，因此竞争是制定最优动态定价时的一个重要考虑因素。学术界最常用的描述用户需求的模型是泊松过程 [84,90–92]，然而这些工作却没考虑服务商之间的竞争。Xu 和 Hoop[93] 则提出用几何布朗运动（Geometric Brownian Motion）来描述用户的需求并用完美贝叶斯均衡来模拟提供商的行为。此外，Levin 等人 [94] 研究了用子博弈完美均衡来表征动态定价。然而，以上研究工作的重点是一次性的库存补给问题，并不能应用到商品（设备）可重复利用的云计算市场。本文将深入研究云计算中设备利用率（用户需求）和价格之间的关系，综合考虑市场竞争等因素，建立合理的数学模型并求解服务商的最优动态定价。

1.4 本文主要贡献

第1.2章介绍了本文要研究的主要问题和面临的挑战，在解决这些具有挑战性的问题的过程中，本文主要做了四个研究工作，这些工作的主要贡献总结如下：

- 付费搜索广告中广告商的有限理性模型研究。本文弱化了广告商是完全理性的这个假设，引入了质反应均衡（Quantal Response Equilibrium, QRE）的概念。在这

种均衡下，期望收益越高的纯策略被选择的概率越高，但广告商无法获取市场上的所有信息，所以无法找到最优的纯策略。质反应均衡更符合真实世界中广告商的行为，纳什均衡只是它的一个特例。本文基于同伦（homotopy）转换的思想实现了一个高效算法来求解付费搜索广告竞标中的质反应均衡。算法中要频繁计算广告商采用不同纯策略的期望收益，为了加速计算这些数值的速度，本文深入研究了在 GSP 机制下广告商期望收益的特殊结构，提出了一个基于图的收益表示方法，极大地降低了计算期望收益的复杂度，提高了算法的性能。此外，本文用质反应模型去拟合真实数据，由此来估测模型中的未知参数，如广告价值和点击率等。

- 付费搜索广告中的收益优化。在假设广告商采用同一均衡策略的前提下，本文研究了如何通过设置保留价来优化搜索引擎的收益。Myerson 定理给出了在一些简单的机制中优化保留价的理论基础，其中最重要的条件是搜索引擎知道广告的价值分布。但在现实世界中，广告价值是广告商的私人信息。在广告价值分布未知的情况下，本文提出了一个采样-预测-实施的框架来帮助搜索引擎最大化全局收益。文中证明了采样阶段和实施阶段的收益的下界，给了利用样本预测最优保留价的方法及误差的理论保证。采样数直接影响着全局收益：采样时间越长预测的最优保留价越准，后期的实施阶段收益越高，但采样阶段的损失越大，反之则相反。本文证明了最优采样数存在且唯一，并设计了一个具有指数收敛速度的算法来计算最优采样数。
- 云计算服务中的马尔科夫完美均衡和长远利益优化。云计算市场正属于飞速发展的阶段，一方面用户数量在不断增加，另一方面随着技术的进步，服务商的边际成本在不断降低。本文研究了市场演化的因素和趋势，提出了一个基于长远利益的价格策略。市场的演化被模拟成一个离散时间的随机博弈（Stochastic Game）过程，优化长远利益的定价策略被规划成马尔科夫完美均衡（Markov Perfect Equilibrium, MPE）。经典的 Q-learning 之类的算法能用于求解随机博弈的最优策略，但其时间复杂度是子博弈个数的指数级函数。本文提出了一个基于动态规划的算法来求解最优定价，计算复杂度降到了子博弈个数的线性级别。
- 云计算服务中的近似均衡和动态定价优化。在云计算市场发展的每一个阶段，服务商平台上的计算资源的数量会随着用户的需求动态变化。当不同的服务商提供的计算资源差别不大时，价格就成了决定用户去留的关键因素。价格设置太高则会导致资源闲置的现象，设置太低则会出现供不应求。因此，服务商需要根据使用率动态调整资源价格使得期望收益最大化。本文综合考虑了服务商之间的市场竞争和动态供需关系，提出了一个更加贴合真实市场的供需关系模型。在此基础上，本文阐明了由于市场中存在不完全信息，纳什均衡无法达成，并提出了一个近似均衡的解的概念。本文分析了该近似均衡在现实中的可行性，同时设计了一

个算法来帮助服务商求解均衡策略。本文还推导了近似均衡的很多重要性质，如动态价格关于资源利用率的单调性。实验表明本文提出的均衡策略相比文献中已有的策略能给服务商带来更高的收益。

1.5 本文组织结构

第2章研究付费搜索广告竞标中广告商的有限理性模型，提出专门的加速算法，并利用该模型拟合真实数据。第3章研究搜索引擎的收益优化，解决当广告商的价值分布未知时权衡采样损失和保留价格精度的问题，设计相关算法计算最优采样数。第4章考虑云计算服务商的长远利益，研究在不断演化的市场中服务商如何设置长远的价格策略。第5章结合云计算中设备利用率动态变化和市场竞争激烈的特点，研究服务商的在市场发展的每一个阶段的最优动态定价策略，优化需求与供应关系从而最大化期望收益。第6章总结全文并讨论未来的工作计划。

第二章 有限理性模型研究

付费搜索广告竞标长期以来受到了学术界的广泛关注，学者们研究了不同的均衡概念来理解广告商的竞标策略。然而，对广告商完全理性的假设在现实世界中很难成立。本章引入质反应均衡的概念，该种均衡能有效地对有限理性建模。由于均衡的计算复杂度很高，现有的求解质反应均衡的方法在广告竞标中的性能非常差。通过研究广告竞标中收益函数的特殊结构，本章提出一种基于同伦的算法来求解付费搜索广告竞标中的质反应均衡。该算法有以下两个创新点：1) 本章将广告竞标表示成一个行为图模型，这样可以在多项式时间内计算广告商的期望收益，此外，本章进一步通过利用广告商的期望收益之间的相互关系来减少冗余计算从而极大地提高运行效率；2) 本章提出了一种利用质反应模型来估计广告竞标中重要参数的方法并用一个商业搜索引擎提供的数据做了测试。实验结果表明本章的算法能显著提高广告竞标中质反应均衡的求解效率且质反应模型能很准确地描述现实世界中广告商的竞标行为。

2.1 引言

付费搜索已经成为了商业搜索引擎（如谷歌，雅虎，必应等）的一个重要的收入来源，在实际应用中取得了巨大成功 [1,52,53]。当用户在搜索引擎中输入一个关键词时，除了相关的页面会被列出，一系列经过筛选的广告也会被显示在搜索结果页面上。为了让自己的广告显示出来，广告商（或称竞标者）需要对该关键词提交一个竞标价格（以下简称竞价）。大部分情况下，参与一个关键词竞标的广告商数量要远远大于搜索引擎结果页面能提供的广告位数量，因此搜索引擎需要一个机制来决定哪些广告被显示、如何对这些广告排序以及如何广告被用户点击后向广告商收费。

广义第二价格广告竞标（Generalized Second Price, GSP）是当今在付费搜索广告竞标（Sponsored Search Auction, SSA）中使用最广泛的机制，吸引了很多学术界的目光 [1,48,50,95]。在这些研究中，均衡分析是一个热门的话题，用以研究广告商的行为。文献 [55] 研究了 GSP 中的对称纳什均衡并证明了其存在性。文献 [96] 定义了一个纳什均衡的子集，称作局部无嫉妒均衡，该均衡概念实际上跟对称纳什均衡是等价的。文献 [97] 进一步证明了 GSP 机制中存在多个纳什均衡。前瞻纳什均衡在文献 [59,60] 中得到了深入研究。上述对均衡的研究工作的一个严重缺陷是它们都假设广告商是完全理性的，即广告商足够聪明，他们总是能找到最优的策略并采取最优的行动。然而，在现实中，一个广告商可能无法准确预测对手的竞价策略，因此无法找到自己最优响应的策略 [8,64]。由此，研究广告竞标中的有限理性均衡是很有必要的。本章将质反应均衡（Quantal Response Equilibrium, QRE） [98–101] 的概念引入到 SSA 中，该模型能描述有

限理性且在一般的标准型博弈中有很好的表现。具体而言，由于在现实世界中无法完全获取竞争对手和市场的信息，广告商不能准确计算自己每一个纯策略的期望收益，QRE 模型假设其中的误差服从一定的分布（例如极值分布 [100,101]），分布的参数通常用来表征广告商的理性程度。由于这些误差的干扰，广告商们在每次竞标中最大化的收益并非完全准确，这导致他们表现出来的策略会形成一个 QRE——一个混合策略均衡，其特点是期望收益越高的纯策略被选择的概率越大。一个广告商越理性，其选择能获得更高收益的纯策略的概率越大。

本章着重于设计一个高效的算法来计算 SSA 中的 QRE，同时研究 QRE 模型在现实世界中的适用性。经推导得出计算 QRE 等价于求解一组 Browder 方程的定点 [102–104]，其计算复杂度至少是 PPAD-完全的 [32,68,105]。这个问题可以进一步转换成求解一个连续的非线性方程组。传统的牛顿迭代类型的算法通常只在能提供一个好的初始迭代点的时候比较有效，然而，在 SSA 中这种初始点是很难寻找的。为了解决这个问题，本章利用同伦（homotopy）转换的思想 [106–108]，该思想已经被用于解一些简单的均衡问题 [109–112]。同伦算法的优势在于数值稳定性以及很大可能的全局收敛性。一个数学工具包 Gambit [113] 使用了类似的方法来求解标准型博弈中的 QRE，但它的算法是极其耗时的因而无法被直接用于 SSA。为克服这个难点，本章利用了 SSA 相比于普通的标准型博弈的一些很好的性质，包括独立于细节内容的结构以及广告商期望收益之间的一些特殊关系，从而优化了计算过程以致能极大地提高算法效率。实验结果表明本章改进的同伦算法能高效地计算 SSA 中的 QRE。此外，本章还研究了如何利用 QRE 模型来推测 SSA 中的参数，包括广告商的广告价值和理性程度，以及点击率。本章开发了一个基于最大似然估计 [114–116] 的预测算法。实验表明 QRE 模型能很好地拟合真实的数据。

总结起来，本章有两大突出贡献。第一，设计了一个高效的基于同伦和 SSA 特殊性质的算法来计算 SSA 中的 QRE。第二，将 QRE 模型应用到了真实的数据并做了大量的试验，验证了 QRE 能比纳什均衡更好地描述现实世界。

本章的剩余部分组织如下。第 2.2 节介绍 SSA 和 GSP 的基本机制，然后给出 QRE 的定义。基于同伦的算法在第 2.3 节中被提出，包括加速算法的设计。第 2.4 节给出参数估测算法。第 2.5 节做大量实验来评估算法性能。本章总结在最后一节。

2.2 质反应均衡

本节首先说明研究 SSA 中的 QRE 的动机，然后提出一些基本的概念和假设，最后给出 SSA 中 QRE 的定义。

2.2.1 研究动机

在 GSP 中，竞价等于自己的广告价值对于广告商而言并不是一个占优策略，因此纳什均衡的解的概念成为了一个重要的理解 SSA 中广告商行为的方法。学者们已经针

对 SSA 提出并研究了很多纳什均衡的概念，其中最被人熟知的是对称纳什均衡 [55] 和局部无嫉妒均衡 [96]，前者的定义为没有任何一对广告商愿意互换他们的位置，而后者则要求任何一个广告商都无法与排在其前一位的互换竞价来获得更高收益。尽管这些均衡概念有不错的性质，但是它们的一个共同的缺陷是假设广告商是完全理性的，即广告商可以完美地知道自己采取不同纯策略的期望收益，从而能够找出最优策略来最大化收益。然而，完全理性的假设在现实的 SSA 中很难成立。因此，一个自然的想法是如何在弱化完全理性假设的同时仍然得到一个有用的解的概念？本章从现实的 SSA 中观察到一个广告商通常对自己的期望收益抱有不确定性，他们有很大的可能选择期望收益大的纯策略，但并不总能找到最优的纯策略。为此，本章引入 QRE 模型来模拟广告商的这种行为。

2.2.2 基本概念及假设

本章主要研究广告竞标中的 GSP 机制。用 N 表示广告商的数量，用 K 表示广告位的数量，通常来说， $N \geq K$ 。符号 $[N]$ 表示集合 $\{1, 2, \dots, N\}$ 。令 v_i 表示广告商 i 的广告价值，等价于 i 愿意为每次广告被点击支付的最大价格。 $v = (v_1, v_2, \dots, v_N)$ 表示所有广告商的广告价值集合。用 b_i 表示 i 出的竞价。 θ_{ik} 是当 i 的广告被显示在位置 k 时被点击的概率，它通常被认为等于广告点击率 α_i 和位置点击率 β_k 的乘积 [1]。分别用 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)$ 和 $\beta = (\beta_1, \beta_2, \dots, \beta_K)$ 来表示广告和位置点击率的集合。一般而言，位置点击率满足

$$\beta_1 \geq \beta_2 \geq \dots \geq \beta_K. \quad (2-1)$$

在 GSP 机制中，搜索引擎按照广告商的分数的降序排列他们的广告，其中分数被定义为

$$s_i = \alpha_i b_i. \quad (2-2)$$

排在最前面的 $k \leq K$ 个广告有机会被展示出来，条件是对应的分数必须不低于一个保留价 r 。如果一个广告被用户点击，对应的广告商支付给搜索引擎的费用等于其维持自己当前位置需要的最低竞价。

2.2.3 QRE 的定义

令 B_i 表示广告商 i 的离散竞价策略空间， b_{ij} 表示 B_i 中第 j 小的值。将 i 的分数空间定义为

$$S_i = \{s_{ij} | s_{ij} = \alpha_i b_{ij}, j \in [|B_i|]\}, \quad (2-3)$$

其中 $|B_i|$ 是空间 B_i 的大小。于是有 $|S_i| = |B_i|, \forall i \in [N]$ 。此外定义除了 i 之外的其他广告商的分数的联合空间为

$$S_{-i} = \times_{l \in [N] \setminus \{i\}} S_l. \quad (2-4)$$

令 $q_{ij}(s_{-i})$ 和 $p_{ij}(s_{-i})$ 分别表示给定 $s_{-i} \in S_{-i}$ 和 $s_{ij} \in S_i$ 时 i 获得的位置以及支付的费用，则 i 的收益为

$$u_{ij}(s_{-i}) = \begin{cases} 0, & s_{ij} < r; \\ (v_i - p_{ij}(s_{-i}))\alpha_i \beta_{q_{ij}(s_{-i})}, & s_{ij} \geq r. \end{cases} \quad (2-5)$$

用 σ_i 表示 i 的基于空间 B_i 的混合策略，用 σ_{ij} 表示采取纯策略 b_{ij} 的概率。类似地，可以定义 $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ 以及 $\sigma_{-i} = (\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_N)$ 。当广告商 i 的分数为 s_{ij} 且 σ_{-i} 给定时其期望收益为

$$\bar{u}_{ij}(\sigma_{-i}) = \sum_{s_{-i} \in S_{-i}} P(s_{-i} | \sigma_{-i}) u_{ij}(s_{-i}), \quad (2-6)$$

其中 $P(s_{-i} | \sigma_{-i})$ 表示在 σ_{-i} 给定的前提下除了 i 之外的其他广告商的价格向量为 s_{-i} 的概率。广告商 i 对其他人的混合策略 σ_{-i} 的质反应被定义为^①

$$\pi_{ij}(\sigma_{-i}) = \frac{1}{\lambda_i} \cdot \frac{1}{|B_i|} + \left(1 - \frac{1}{\lambda_i}\right) \frac{e^{\bar{u}_{ij}(\sigma_{-i})\lambda_i}}{\sum_{k \in [B_i]} e^{\bar{u}_{ik}(\sigma_{-i})\lambda_i}} \quad (2-7)$$

$$= \frac{1}{\lambda_i |B_i|} + \left(1 - \frac{1}{\lambda_i}\right) \frac{1}{\sum_{k \in [B_i]} e^{(\bar{u}_{ik}(\sigma_{-i}) - \bar{u}_{ij}(\sigma_{-i}))\lambda_i}}, \quad (2-8)$$

其中 $\lambda_i \in [1, +\infty)$ 是广告商 i 的理性参数。很容易验证得到公式 (2-7) 与本章之前对 QRE 性质的描述是一致的，即期望收益越高的纯策略被选择的概率越高。当 $\lambda_i = 1$ 时，有 $\pi_{ij}(\sigma_{-i}) = \frac{1}{|B_i|}$ ，这意味着 i 随机从 B_i 中选择策略；当 $\lambda_i \mapsto +\infty$ 时，期望收益最高的策略被选择的概率接近 1。总结起来， λ_i 代表着广告商 i 的理性程度，它的值越大则 i 选择好的纯策略的概率越高。用向量 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$ 来表示所有广告商理性参数的集合。此时可以给出 QRE 的定义 [98,100]。

定义 2.1 一个以 λ 为理性参数的 QRE 是一个混合策略 σ ，它满足对所有的 $i \in [N]$ 和 $j \in [B_i]$,

$$\sigma_{ij} = \pi_{ij}(\sigma_{-i}). \quad (2-9)$$

定义 2.1 表明 QRE 本质上是一组 Browder 方程的定点 [102]，且纳什均衡是 QRE 在满足 $\lambda_i \rightarrow +\infty, \forall i \in [N]$ 时的一个特例。

① 在最常用的质反应定义中 [100,112]，对所有的 $i \in [N]$ ， α_i 和 λ_i 分别乘以和除以一个相同的数不会改变最终的 QRE，也就是说 α_i 和 λ_i 在这种形式的定义中实际上是作为一个整体起作用的。为了能够单独研究每个因子的作用，本章对该形式做了一点改动，但仍保证符合 QRE 的性质。

2.3 均衡计算

本节将设计一个基于同伦思想的算法来求解 SSA 中的 QRE，首先给出算法框架，然后讨论如何利用 SSA 的特殊性质来显著地加速算法。

2.3.1 基于同伦变换的算法

本节研究如何在给定 v , λ , α 和 β 的时候计算 QRE。定义

$$U = \sum_{i \in [N]} |B_i|, \quad (2-10)$$

从而根据定义 2.1 可以得到一个连续函数 $F : [0, 1]^U \mapsto [0, 1]^U$:

$$F_{ij}(\sigma) = \pi_{ij}(\sigma_{-i}) - \sigma_{ij}, \forall i \in [N], j \in [|B_i|]. \quad (2-11)$$

至此可见计算 SSA 的 QRE 等价于寻找非线性函数 $F(\sigma)$ 的零点。如果能提供一个好的初始点，可以直接利用牛顿迭代法来解这一问题。然而，对如何寻找这样的起始点通常没有任何先验知识。正如文献 [106,107] 中指出的，牛顿迭代法很容易不收敛，因为一个坏的起始点被选择的概率很高。

同伦算法的基本想法由两步组成：第一步，首先构造一个易解且解唯一的函数，然后建立一个连续的从构造的函数到原始要求解的函数的同伦变换；第二步，从求解构造的函数的零点开始，追踪这个同伦变换所经过的函数的零点构成的路径，直到最后得到原始要求解的函数的零点。为了设计一个同伦算法，首先要找到一个易解的原始问题的退化形式。在 SSA 中，可以通过令 $\bar{u}_{ij}(\sigma_{-i})$, $\forall i \in [N]$ 和 $j \in [|B_i|]$ ，都等于零来构造这样一个函数 G :

$$G_{ij}(\sigma) = \frac{1}{|B_i|} - \sigma_{ij}, i \in [N], j \in [|B_i|]. \quad (2-12)$$

显然， $G(\sigma)$ 有唯一的零点： $\sigma_{ij} = 1/|B_i|, \forall i \in [N], j \in [|B_i|]$ 。然后可定义基于 $F(\sigma)$ 和 $G(\sigma)$ 的同伦变换函数为 $H : [0, 1]^U \times [0, 1] \mapsto [0, 1]^U$ ，其满足

$$H_{ij}(\sigma, t) = \frac{1}{\lambda_i |B_i|} + (1 - \frac{1}{\lambda_i}) R(\sigma_{-i}, t)^{-1} - \sigma_{ij}, \quad (2-13)$$

$$R(\sigma_{-i}, t) = \sum_{k \in [|B_i|]} e^{(\bar{u}_{ik}(\sigma_{-i}) - \bar{u}_{ij}(\sigma_{-i})) \lambda_i t}, \quad (2-14)$$

很容易验证得到当 t 从 0 连续地增加到 1 的过程中，上式从 $H(\sigma, 0) = G(\sigma)$ 连续地变换到 $H(\sigma, 1) = F(\sigma)$ 。进一步将 $H(\sigma, t) = 0$ 的解集定义为

$$H^{-1}(0) = \{(\sigma, t) | H(\sigma, t) = 0\}. \quad (2-15)$$

根据 Browder 定点理论 [102]，对于一个给定的 $t \in [0, 1]$ ，一定存在一个 $\sigma(t)$ 满足 $H(\sigma(t), t) = 0$ 。从 $H(\cdot)$ 的定义可以知道 $\sigma(0)$ 和 $\sigma(1)$ 分别对应于 $G(\sigma)$ 和 $F(\sigma)$ 的零点。

剩下的问题是如何找到一条由 $(\sigma(t), t) \in H^{-1}(0)$ 构成的以 $(\sigma(0), 0)$ 为起始点以 $(\sigma(1), 1)$ 为终止点的路径。考虑到存在转折点 [117] 的可能性，即该路径先后通过的两个零点 $(\sigma(t), t)$ 和 $(\sigma(t'), t')$ 可能满足 $t > t'$ ，因此通过单调增加 t 来追踪这条路径是不安全的。为了避免转折点的影响，可以将 σ 和 t 同时看做一个变量 a 的隐函数，从而将原问题转换为寻找一个参数化路径

$$c(a) = (\sigma(t(a)), t(a)), \quad (2-16)$$

其满足

$$H(c(a)) = 0. \quad (2-17)$$

本章用于追踪这条参数化路径的方法被称为预测-更正法 [106,107]，它的基本思想是生成一系列满足 $\|H(c_i)\| \leq \varepsilon$ 的点 $c_i = (\sigma, t)_i, i = 1, 2, \dots$ 来追踪路径 $c(a) = (\sigma(t(a)), t(a))$ ，其中 $\varepsilon > 0$ 是一个很小的常数。具体地，给定路径 $c(a)$ 上一个已经找到的点 c_i ，本章用欧拉预测法来找寻 $c(a)$ 上的下一个点 c_{i+1} ，如下：

$$c_{i+1} = c_i + \Delta \cdot \frac{c'(a)|_{c(a)=c_i}}{\|c'(a)|_{c(a)=c_i}\|}, \quad (2-18)$$

其中 $c'(a)$ 是 $c(a)$ 对 a 的导数， $\Delta > 0$ 是步长。由于预测的点很可能不在 $c(a)$ 上，需要一个更正操作来使得 c_{i+1} 尽可能靠近路径 $c(a)$ 。于是本章采用了高斯-牛顿更正法，如下：

$$\widehat{c}_{i+1} = c_{i+1} - H'(c_{i+1})^+ H(c_{i+1}), \quad (2-19)$$

其中 $H'(c_{i+1})^+$ 是 $H(\cdot)$ 在点 c_{i+1} 的雅克比矩阵 $H'(c_{i+1})$ 的广义逆矩阵^①， \widehat{c}_{i+1} 是更正过的 c_{i+1} 。如果 $\|H(\widehat{c}_{i+1})\| > \varepsilon$ ，则将 \widehat{c}_{i+1} 代入到公式 (2-19) 来进一步更正它。修正步可能要执行多次，直到修正过的点达到了精度要求，然后将该点应用到欧拉预测步来预测再下一个点。这个预测-更正法从 $(\sigma(0), 0)$ 开始一步一步地执行上述操作直到到达 $(\sigma(1), 1)$ 。

现在讨论如何计算公式 (2-18) 中的导数 $c'(a)$ 以及公式 (2-19) 中的雅克比矩阵 $H'(c_{i+1})$ 。首先考虑 $c'(a)$ 的计算。通过对公式 (2-17) 求导可得到如下方程：

$$H'(c(a))c'(a) = 0. \quad (2-20)$$

公式 (2-20) 的解为

$$c'_d(a) = \mu \cdot (-1)^d \cdot \det(H'_{-d}(c(a))), \quad (2-21)$$

其中 $c'_d(a)$ ， $d = 1, \dots, U + 1$ ，表示 $c'(a)$ 的第 d 个元素^②； $H'_{-d}(c(a))$ 是将 $H'(c(a))$ 的第 d 列移除后的矩阵； $\det(\cdot)$ 是求行列式的操作； $\mu = \pm 1$ 是 $c'(a)$ 的符号。根

① 一个矩阵 A 的广义逆矩阵定义为 $A^+ = A^T(AA^T)^{-1}$ 。

② 假设 σ_{ij} 按照其下标的字典序对应 $c(a)$ 中的元素且 $c(a)$ 的最后一个元素对应 t 。

据公式 (2-21) 可以发现一旦知道 $H'(c(a))$ 的值, 则可以直接计算 $c'(a)$ 。此外公式 (2-19) 也涉及到计算 $H'(\cdot)$ 。因此接下来本节集中精力研究如何计算导数 $H'(\cdot)$ 。为了方便表示, 后文中用 \tilde{u}_{ij} 来替换 $\bar{u}_{ij}(\sigma_{-i})\lambda_i$ 。由于函数 H 是一个从 $[0, 1]^U \times [0, 1]$ 到 $[0, 1]^U$ 的映射, 它的雅克比矩阵包含 $U \cdot (U+1)$ 个偏导数, 它们可以按以下四种情况分类:

情况 1. $i \in [N]$ 且 $j \in [B_i]$:

$$\frac{\partial H_{ij}}{\partial \sigma_{ij}} = -1; \quad (2-22)$$

情况 2. $i \in [N]$, j 和 $k \in [B_i]$, $j \neq k$:

$$\frac{\partial H_{ij}}{\partial \sigma_{ik}} = 0; \quad (2-23)$$

情况 3. i 和 $l \in [N]$, $j \in [B_i]$, $m \in [B_l]$, $i \neq l$:

$$\frac{\partial H_{ij}}{\partial \sigma_{lm}} = -(1 - \frac{1}{\lambda_i})R(\sigma_{-i}, t)^{-2} \frac{\partial R(\sigma_{-i}, t)}{\partial \sigma_{lm}}, \quad (2-24)$$

$$\frac{\partial R(\sigma_{-i}, t)}{\partial \sigma_{lm}} = te^{-\tilde{u}_{ij}t} \sum_{k \in [B_i]} e^{\tilde{u}_{ik}t} \left(\frac{\partial \tilde{u}_{ik}}{\partial \sigma_{lm}} - \frac{\partial \tilde{u}_{ij}}{\partial \sigma_{lm}} \right); \quad (2-25)$$

情况 4. $i \in [N]$, $j \in [B_i]$:

$$\frac{\partial H_{ij}}{\partial t} = -(1 - \frac{1}{\lambda_i})R(\sigma_{-i}, t)^{-2} \frac{\partial R(\sigma_{-i}, t)}{\partial t}, \quad (2-26)$$

$$\frac{\partial R(\sigma_{-i}, t)}{\partial t} = e^{-\tilde{u}_{ij}t} \sum_{k \in [B_i]} e^{\tilde{u}_{ik}t} (\tilde{u}_{ik} - \tilde{u}_{ij}). \quad (2-27)$$

符号 μ 的选择要保证 t 对 a 的导数, 亦即 $c'(a)$ 的第 $(U+1)$ 个元素, 在点 $(\sigma(0), 0)$ 处是正的, 也就是说

$$\mu \cdot (-1)^{U+1} \det(H'_{-(U+1)}(\sigma(0), 0)) > 0. \quad (2-28)$$

将 $t = 0$ 代入到上面的不等式, 再结合情况 1-3, 可得到

$$\mu \cdot (-1)^{U+1} \cdot (-1)^U = (-1)^{2U+1} \mu > 0, \quad (2-29)$$

上式说明在点 $(\sigma(0), 0)$ 处 $\mu = -1$ 。

目前为止本节已经讨论了如何计算公式 (2-18) 和 (2-19)。这样便可以使用上述预测-更正法来一步步寻找 $(\sigma(1), 1) \in H^{-1}(0)$ 。求解 SSA 中的 QRE 的完整的过程被总结在算法 1 中。在第 1-2 行将 t 赋值为 0 并将初始点 c_1 设为 $(\sigma(0), 0)$ 。第 3 行初始化 Δ 和 ε 。第 4-9 行利用预测-更正法追踪参数化路径 $c(a)$ 上的一系列点 $c_i, i = 2, 3, \dots$ 直到最终找到 $(\sigma(1), 1)$ 。其中第 5 行根据公式 (2-18) 利用欧拉预测法和已知点 c_i 来预测下一个点 c_{i+1} , 然后高斯-牛顿更正法在第 6-7 行中不断提高第 5 行预测的新点的精度。 t 的值在

第 8 行被更新。本章在第 9 行中利用文献 [106,107] 中的渐近展开法来调整步长。最终结果在第 11 行被返回。

Algorithm 1: 计算 QRE

```

1  $t \leftarrow 0$ ;
2  $c \leftarrow (\sigma(0), 0)$ ;
3 初始化  $\Delta$  和  $\varepsilon$ ;
4 while  $t \neq 1$  do
5      $c \leftarrow c + \Delta \cdot \frac{c'(a)|_{c(a)=c}}{\|(c'(a)|_{c(a)=c})\|}$ ;
6     while  $\|H(c)\| > \varepsilon$  do
7          $c \leftarrow c - H'(c)^+ H(c)$ ;
8      $t \leftarrow c$  的最后一个元素;
9     调整步长  $\Delta$ ;
10  $(\sigma, t) \leftarrow c$ ;
11 return  $\sigma$ ;
```

2.3.2 针对 SSA 的高效算法

算法 1 表明在每一个预测和更正步都需要计算雅克比矩阵 $H'(\cdot)$ 。因此计算 $H'(\cdot)$ 的效率将极大地影响整个算法的性能。本节讨论如何通过利用 SSA 的一些特殊性质来高效地计算 $H'(\cdot)$ 进而降低算法 1 的复杂度。首先，本节将 SSA 表示成一个行为图模型 [118–120]，这样计算 $H'(\cdot)$ 的各个元素仅需要多项式时间，而将 SSA 看成一般的标准型博弈则需要指数时间。其次，本节通过利用广告商收益函数之间的特殊关系来进一步减少冗余计算来提升算法效率。

2.3.2.1 将 SSA 表示成行为图模型

正如第 2.3.1 节中提到的， $H'(\cdot)$ 的元素可以被分为四类，后两类涉及到计算 \bar{u}_{ij} 和 $\frac{\partial \bar{u}_{ij}}{\partial \sigma_{lm}}$ 。可以将公式 (2-6) 重写为

$$\begin{aligned}
 & \bar{u}_{ij}(\sigma_{-i}) \\
 = & \sum_{m \in [S_{-i}]} \sigma_{lm} \sum_{s_{-il} \in S_{-il}} P(s_{-il} | \sigma_{-il}) u_{ij}(s_{-il}, s_{lm}) \tag{2-30}
 \end{aligned}$$

$$= \sum_{m \in [S_{-i}]} \sigma_{lm} \frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}, \forall l \neq i, \tag{2-31}$$

其中

$$S_{-il} = \times_{i' \in [N] \setminus \{i, l\}} S_{i'}. \tag{2-32}$$

所以要计算 $H'(\cdot)$ 等同于计算一组偏导数，即

$$D = \left\{ \frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}} \mid i, l \in [N]; l \neq i; j \in [B_i]; m \in [B_l] \right\}. \quad (2-33)$$

如果将 SSA 看成是标准型博弈，则计算 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 需要遍历 S_{-il} 中的 $\prod_{i' \in [N] \setminus \{i, l\}} |B_{i'}|$ 个元素。显然，这种遍历的方法（Traversal Method, TM）的时间复杂度为 $O(M^N)$ ，即关于 N 是指数级别的，其中

$$M = \max\{|B_{i'}| \mid i' \in [N]\}. \quad (2-34)$$

幸运的是，采用 GSP 机制的 SSA 中广告商的期望收益有很多特殊的性质可以被用来降低时间复杂度。这里用分数为 s_{ij} 的广告商 i 做个例子来具体说明。定义

$$I_G = \{n \mid s_n > s_{ij}, n \neq i\}. \quad (2-35)$$

类似地，可以定义

$$I_E = \{n \mid s_n = s_{ij}, n \neq i\} \quad (2-36)$$

以及

$$I_L = \{n \mid s_n < s_{ij}, n \neq i\}. \quad (2-37)$$

假设分数出现平局时位置随机分配，则广告商 i 的收益具有如下性质：

1. 当 $|I_G| \geq K$ 时， i 的收益总是零。
2. 当 $|I_G| < K$ 且 $|I_G| + |I_E| \geq K$ 时，广告商 i 有 $\frac{1}{|I_E|+1}$ 的概率被分到从 $|I_G| + 1$ 到 K 的任何一个广告位，且其支付的费用为 $p_i = b_{ij}$ 。根据本节对平局的假设， i 在这种情形下的收益为

$$\frac{1}{|I_E| + 1} (v_i - b_{ij}) \sum_{k=|I_G|+1}^K \alpha_i \beta_k. \quad (2-38)$$

3. 当 $|I_G| + |I_E| < K$ 时， i 会以等概率 $\frac{1}{|I_E|+1}$ 被分至位置 $|I_G| + 1$ 到 $|I_G| + |I_E| + 1$ 。如果被分至位置 $|I_G| + 1$ 到 $|I_G| + |I_E|$ ，其需要支付的费用为 $p_i = b_{ij}$ ；如果被分至位置 $|I_G| + |I_E| + 1$ ，则其需要支付 $\frac{p_{max}}{\alpha_i}$ ，其中

$$p_{max} = \max\{r, s_n \mid n \in I_L\}. \quad (2-39)$$

广告商 i 在这种情形下的收益为

$$\frac{1}{|I_E| + 1} \left((v_i - b_{ij}) \sum_{k=|I_G|+1}^{|I_G|+|I_E|} \alpha_i \beta_k + (v_i - p_{max}) \alpha_i \beta_{|I_G|+|I_E|+1} \right). \quad (2-40)$$

上述性质表明，给定 i 的分数 $s_i = s_{ij}$ ，其收益仅依赖于 $|I_G|$ ， $|I_E|$ 和 p_{max} ，而不取决于 I_G 和 I_E 都具体包含哪些广告商以及他们的竞价是多少，也不关心 p_{max} 是谁出的竞价。也就是说，SSA 的收益有着很强的独立于具体细节的结构，因此可以被表示成一个行为图模型。行为图本质上是一个字典树^①，它的每个叶子结点对应于一个三元组 $(|I_G|, |I_E|, p_{max})$ 。具体地，当计算 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 时，因为 1) 仅考虑 $|I_G| < K$ 的情况，2) $|I_E| \leq N$ ，且 3) p_{max} 最多有 NM 种不同的值，行为图最多有 $O(KN^2M)$ 个叶子结点，因此可以在 $O(KN^2MN) = O(KN^3M)$ 时间内通过动态规划来创建 [118]。相比于将 SSA 看做一个标准型博弈并用 TM 方法来计算该偏导的 $O(M^N)$ 复杂度，这是一个很大的性能上的提升。

2.3.2.2 减少冗余计算

直观上看需要对 D 中的 $N(N-1)M^2$ 个偏导分别建立一个行为图以计算偏导的值。事实上，可以利用 SSA 的性质极大地减少冗余计算。定义 σ_{-il} 为除了广告商 i 和 j 以外的其他广告商的混合策略。于是有

$$\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}} = \sigma_{lm} \sum_{s_{-il} \in S_{-il}} P(s_{-il} | \sigma_{-il}) u_{ij}(s_{-il}, s_{lm}). \quad (2-41)$$

也就是说，可以将 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 看成是将 l 的策略固定为 m 后， i 采取策略 j 的期望收益再乘以一个权重 σ_{lm} 。因此在构建 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 的行为图时，可以利用 [118] 的算法先处理 S_{-il} 中的元素，然后再考虑 s_{lm} 。对于不同的 $m \in [|B_l|]$ ，处理 S_{-il} 的过程是一样的，于是只需要为它们建造一个公共的行为图，基于这个图可以很容易地推导出 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ ， $\forall m \in [|B_l|]$ 的行为图。这个性质预示着最多只需要构建 $N(N-1)M$ 个行为图便能计算 D 中的 $N(N-1)M^2$ 个偏导数。接下来基于以下四个命题做进一步优化。

命题 1 对于所有的 $l \neq i$ ，如果 $s_{ij} < r$ ，那么 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}} = 0$ 。

这个命题是很直观的，因为公式 (2-5) 表明当 $s_{ij} < r$ 时 $u_{ij}(s_{-il}, s_{lm}) = 0$ 。

命题 2 对于所有的 $l \neq i$ ，如果不存在一个 $s_{l'm'}$ ， $l' \neq i$ 且 $m' \in [|B_{l'}|]$ ，满足 $s_{ij} \leq s_{l'm'} \leq s_{ij+k}$ ，则有

$$\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}} = \frac{\partial \bar{u}_{ij+1}(\sigma_{-i})}{\partial \sigma_{lm}} = \dots = \frac{\partial \bar{u}_{ij+k}(\sigma_{-i})}{\partial \sigma_{lm}}. \quad (2-42)$$

证明 根据命题的假设， I_E 是空集。给定 b_{-il} ，除了 i 以外的广告商要么属于 I_G ，要么属于 I_L ，且他们的位置在 i 的分数从 s_{ij} 变到 s_{ij+n} ， $n \in [k]$ 的过程中不会改变。于是有，对于所有的 $s_{-il} \in S_{-il}$ ，

$$u_{ij}(s_{-il}, s_{lm}) = u_{ij+n}(s_{-il}, s_{lm}), \forall n \in [k], \quad (2-43)$$

因为它们对应相同的 I_G 、 I_E 和 I_L 。根据公式 (2-30)，对 s_{-il} 取期望便可完成命题的证明。□

① 字典树是一种数据结构，具体可参见 <https://en.wikipedia.org/wiki/Trie> 和 [121]。

命题 3 对于所有的 $l \neq i$, 如果 $S_i = S_l$, 则 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 和 $\frac{\partial \bar{u}_{lj}(\sigma_{-i})}{\partial \sigma_{im}}$ 的行为图是一样的。

证明 由于行为图本质上是一个字典树, 仅需证明它们有一样的叶子结点。因为 $s_{ij} = s_{lj}$ 、 $s_{im} = s_{lm}$ 且 $S_{-il} = S_{-li}$, 所以对于每一个 $s_{-i} = (s_{-il}, s_{lm})$, $s_{-il} \in S_{-il}$, 得到的三元组 $(|I_G|, |I_E|, p_{max})$, 一定可以找到一个 $s_{-l} = (s_{-il}, s_{im})$, $s_{-li} \in S_{-li}$, 对应同样的三元组; 反之也成立。由于三元组和行为图的叶子结点存在一一对应关系, 命题得证。 \square

命题 2 说明如果 S_i 的一个子集 $\{s_{ij}, s_{ij+1}, \dots, s_{ij+k}\}$ 满足命题中的条件, 那么 $\frac{\partial \bar{u}_{ij+1}(\sigma_{-i})}{\partial \sigma_{lm}}, \dots, \frac{\partial \bar{u}_{ij+k}(\sigma_{-i})}{\partial \sigma_{lm}}$ 可以直接从 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 直接推导出来, 而不用再额外创建行为图。命题 2 无法被应用到广告商拥有相同的分数空间的情况。与之相反地, 命题 3 尤其适用于这种情况, 它可将需要创建的行为图的数量降低到 $\frac{N(N-1)M}{2}$ 。

命题 4 对于所有的 $l \neq i$, 如果 1) $s_{lm_1}, s_{lm_2} \leq r$ 或者 2) $s_{lm_1}, s_{lm_2} > s_{ij}$, 则有

$$\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm_1}} = \frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm_2}}. \quad (2-44)$$

证明 与命题 2 的证明类似, 仅需要证明对于所有的 $s_{-il} \in S_{-il}$, $u_{ij}(s_{-il}, s_{lm_1}) = u_{ij}(s_{-il}, s_{lm_2})$, 该式是成立的, 因为 (s_{-il}, s_{lm_1}) 和 (s_{-il}, s_{lm_2}) 在命题的假设下对应相同的三元组 $(|I_G|, |I_E|, p_{max})$ 。 \square

此处用一个广告商拥有相同分数空间的例子来分析命题 3 和 4 的作用 (方便起见, 假设 $r = 0$)。给定 i 和 l , 仅需要计算 $j < M$ 和 $m \leq j + 1$ 时的 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$, 以及 $m \leq M$ 时的 $\frac{\partial \bar{u}_{im}(\sigma_{-i})}{\partial \sigma_{lm}}$, 因为根据命题 4 其他情况下的偏导等于 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{ij+1}}$ 。因此, 需要用行为图计算的 D 中的偏导的个数为

$$(2 + 3 + \dots + M + M) = \frac{M^2 + 3M - 2}{2}.$$

另一方面, 当 j 和 m 给定时, 命题 3 表明 $\frac{\partial \bar{u}_{ij}(\sigma_{-i})}{\partial \sigma_{lm}}$ 和 $\frac{\partial \bar{u}_{lj}(\sigma_{-i})}{\partial \sigma_{im}}$ 可以用一个行为图来计算。在 D 中一共有 $\frac{N(N-1)}{2}$ 对呈这种关系的偏导, 这导致优化后的计算量一共减少了

$$\begin{aligned} & N(N-1)M^2 - \frac{N(N-1)}{2} \frac{M^2 + 3M - 2}{2} \\ & > \frac{3}{4} N(N-1)M(M-1). \end{aligned} \quad (2-45)$$

由于 D 一共含有 $N(N-1)M^2$ 个偏导, 整体的计算效率被提高了大约 75%。

2.4 参数预测算法

本节提出一个基于 QRE 和最大似然估计的算法从真实数据中预测 SSA 的参数 v , α 和 β 。文献 [122] 指出搜索引擎很难精确地预测点击率, 50% 的误差是很常见的。本章的模型为搜索引擎提供了一种预测这些参数的方法。

假设一个 QRE 策略 σ 已经形成, 未知参数 v, λ, α 和 β 的对数似然为

$$L(v, \lambda, \alpha, \beta | \sigma) = \log \left(\prod_{i \in [N]} \prod_{j \in [B_i]} \pi_{ij}(\sigma_{-i})^{\sigma_{ij}} \right) \quad (2-46)$$

然后通过最大化似然函数来推测这些参数, 公式如下:

$$\max_{v, \lambda, \alpha, \beta} L(v, \lambda, \alpha, \beta | \sigma) \quad (2-47)$$

$$s.t. \begin{cases} v_i \geq 0, \forall i \in [N]; \\ \lambda_i \geq 1, \forall i \in [N]; \\ 1 > \beta_s \geq \beta_{s+1} > 0, \forall s \in [K-1]; \\ 0 < \alpha_i < 1, \forall i \in [N]. \end{cases} \quad (2-48)$$

然而, $q_{ij}(s_{-i})$ 并不是 $\alpha_i, \forall i \in [N]$ 的连续函数, 因此广告商的收益函数也是不连续的。这导致公式 (2-46) 中的似然函数关于 α 是不连续的。为了解决这个问题, 可将未知的参数分为两组, 然后分步优化两组参数。具体而言, 将 v, λ, β 分为一组, 将 α 分到另一组, 在算法的每一轮循环中, 先优化 v, λ, β (保持 α 不变) 再优化 α (保持 v, λ, β 不变), 当两组参数都无法再被优化时循环停止。

公式 (2-47) 中定义的似然函数 $L(v, \lambda, \alpha, \beta | \sigma)$ 相对于第一组参数是连续的。可以通过解下面的问题优化这些参数:

$$\max_{v, \lambda, \beta} L(v, \lambda, \alpha, \beta | \sigma) \quad (2-49)$$

$$s.t. \begin{cases} v_i \geq 0, \forall i \in [N]; \\ \lambda_i \geq 1, \forall i \in [N]; \\ 1 > \beta_s \geq \beta_{s+1} > 0, \forall s \in [K-1]. \end{cases} \quad (2-50)$$

由于上述优化问题的目标函数是非凸的, 因此很难找到其全局最优点。于是本节通过寻找一系列局部最优点然后选取最好的一个来提高找到全局最优的概率。

正如前面提到的, 似然函数相对于 α 是不连续的。以 α_i 为例, 当其他广告商的广告点击率 $\alpha_j, j \in [N] \setminus \{i\}$ 给定时, 可以发现似然函数具有以下不连续点:

$$\left\{ \frac{\alpha_j b_j}{b_i} \mid 0 < \frac{\alpha_j b_j}{b_i} < 1, j \in [N] \setminus \{i\}, b_i \in B_i, b_j \in B_j \right\}.$$

这些点可将 α_i 的取值空间划分成几个区间, 在各个区间上似然函数关于 α_i 是连续的。因此, 在优化 α 时采用逐个优化 $\alpha_i, i \in [N]$ 的方法, 即先固定 $\alpha_j, j \in [N] \setminus \{i\}$ 的值然后

在 α_i 使得似然函数连续的各个区间上分别进行优化，之后选取最优的那个值作为最终结果，形式化定义如下所示：

$$\max_{\alpha_i} L(v, \lambda, \alpha, \beta | \sigma) \quad (2-51)$$

$$s.t. \alpha_i \notin \left\{ \frac{\alpha_j b_j}{b_i} \mid 0 < \frac{\alpha_j b_j}{b_i} < 1, j \in [N] \setminus \{i\}, b_i \in B_i, b_j \in B_j \right\}. \quad (2-52)$$

由于上式的优化问题也很有可能是非凸的，为了避免陷入一个不好的局部最优点，也采用多个初始点进行多次求解然后选择最好的局部最优作为最终结果。完整的解法流程如算法2所示。第1行初始化似然值为负无穷并在第2行中随机生成广告点击率向量 α 。第3-13行不停地迭代更新两组参数的值。其中第4行固定 α 后更新 (v, λ, β) 。为防止算法陷入无限循环 [123]，第5行初始化集合 I 用于后面按随机的顺序更新 α 的元素，然后第6-8行则固定 (v, λ, β) 并更新 α 。第9-13行控制算法进程：如果能得到更好的参数，则继续优化，否则返回当前最好的参数。为了避免陷入局部最后，实验中会多次运行算法2，然后选取其中最好的结果。

Algorithm 2: 参数估测

```

1  $\hat{L}^* \leftarrow -\infty$ ;
2 随机生成广告点击率向量  $\alpha$ ;
3 while True do
4   固定  $\alpha$  并通过解公式 (2-49) 和 (2-50) 中的问题来更新  $v, \lambda, \beta$ ;
5    $I \leftarrow \{1, 2, \dots, N\}$ ;
6   while  $I \neq \emptyset$  do
7     随机从  $I$  中采样得到一个元素  $i$ ，并将  $i$  从  $I$  中移除;
8     固定  $\alpha_j, j \in [N] \setminus \{i\}$  以及  $v, \lambda, \beta$ ，在  $\alpha_i$  的每个连续区间上求解公式 (2-51)
       中的问题来更新  $\alpha_i$ ;
9    $\hat{L} \leftarrow L(v, \lambda, \alpha, \beta | \sigma)$ ;
10  if  $\hat{L} > \hat{L}^*$  then
11     $\hat{L}^* \leftarrow \hat{L}$ ;
12  else
13    return  $v, \lambda, \alpha, \beta$ ;

```

2.5 实验评估

本节通过大量的实验来评估 QRE 求解算法和参数预测算法。本节所有实验均由 C++ 实现，运行在一台搭载了四核 i7 处理器，运行内存 16G，操作系统为 64 位 Windows 10 的设备上。

2.5.1 QRE 求解算法的性能评估

首先比较三种计算 D 中所有偏导数的方法：遍历法 (TM)，行为图法 (AGG)，行为图结合上文提出的加速算法的方法 (AGGSU)。根据广告商是否有相同的分数空间可将实验分为两组。在每组实验中，用不同大小的 SSA 来测试算法性能，其中 $N = 5, 6, \dots, 20$ 且 $M = 5, 10, 15, 20$ 。在所有的实验中， α, β 和 $B_i, \forall i \in [N]$ ，分别从区间 $(0.1, 1)^N, (0.1, 1)^K$ 和 $(0, M)^M$ 中采样得到。令 $K = \lfloor N/2 \rfloor$ ，令 $v_i = \max_{b_i \in B_i} b_i, \forall i \in [N]$ 。每种方法在每种设定下的运行时间是经过 100 次实验的平均值。结果被显示在图 2.1 中，其中 y 轴使用的是对数坐标且超过 10^4 秒的数据没有被显示，“-S” 和 “-D” 分别表示相同和不同分数集的实验组。由于 TM 在两组实验中的运行速度几乎一样，图中仅画出了二者的平均值。

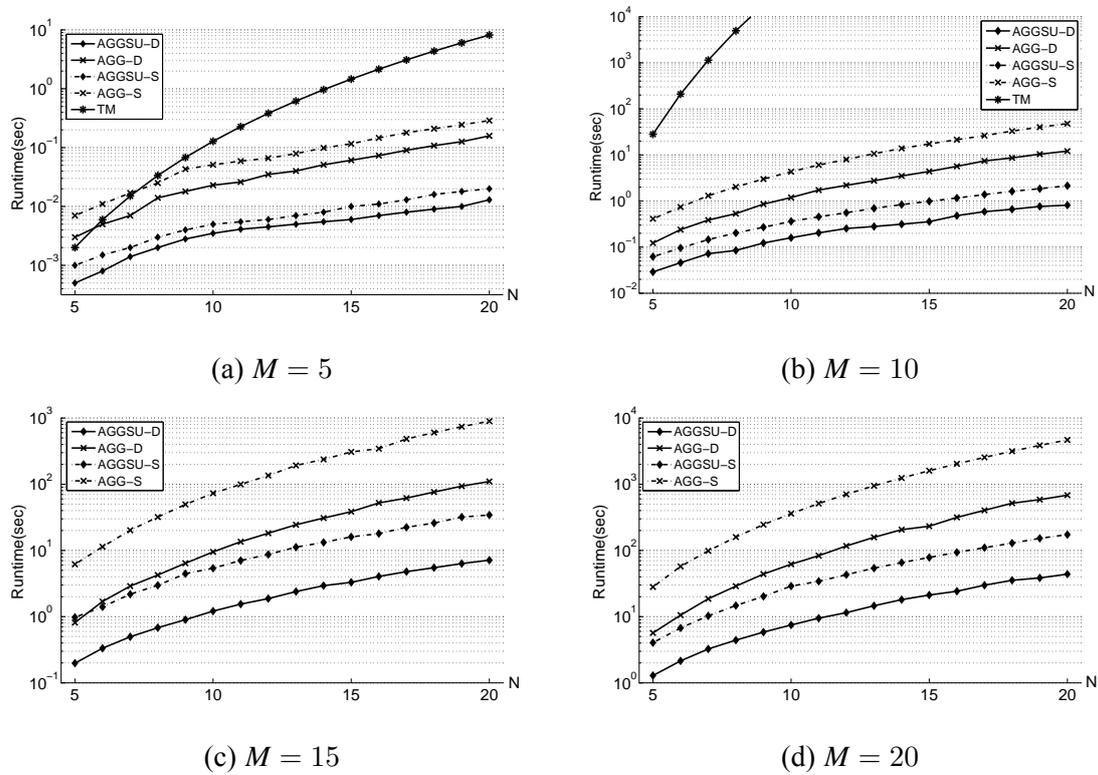


图 2.1 TM、AGG 和 AGGSU 的运行时间比较

从图 2.1 中可以看出遍历法的运行时间随着 N 的增加呈指数级增长。对于 8×10 (即 $N = 8, M = 10$) 或更大的 SSA，遍历法不能在 1 小时内给出结果。作为比较，AGG 和 AGGSU 比遍历法高效得多。可以观察到 AGGSU-D (AGGSU-S) 大约是 AGG-D (AGG-S) 效率的 10 倍，这验证了第 2.3.2.2 节提出的加速方法的有效性。进一步注意到当 SSA 大小相同时，AGGSU-S (AGG-S) 总是比 AGGSU-D (AGG-D) 计算得慢。这是因为对于 AGGSU-D (AGG-D) 来说 I_E 几乎总是空的，因此其行为图含有 $O(KNM)$ 个叶子结点，而 AGGSU-S (AGG-S) 的含有 $O(KN^2M)$ 个叶子结点。总体来说，本章提出的 AGGSU 方法在所有实验中性能都是最好的。

表 2.1 QRE 求解算法性能评估

Δ	$N = 10, M = 5$			$N = 10, M = 10$			$N = 15, M = 10$		
	时间	#E	#G	时间	#E	#G	时间	#E	#G
Dy	0.55	5.10	2.40	5.66	5.42	2.38	19.5	8.00	2.44
0.1	1.92	26.0	0.73	18.6	29.1	0.68	82.1	47.0	0.76
0.3	0.85	10.9	1.03	8.58	11.2	1.03	37.8	18.1	1.05
0.5	0.82	7.45	1.71	7.67	8.03	1.37	33.0	12.3	1.83
0.7	0.81	6.76	2.40	7.68	5.45	2.44	26.8	10.7	2.32
0.9	0.82	5.44	2.60	7.92	4.96	2.75	30.1	8.12	3.56
1.1	0.85	4.34	3.25	10.8	5.02	3.03	37.7	10.2	3.91

接下来评估算法1的性能。实验中 SSA 的参数如上面所述，此外 λ 从 $(0, 10)^N$ 中随机采样生成。此处不对广告商的分数空间做任何额外假设，因此在算法中用 AGGSU-D 来计算 D 中的偏导数。实验基于三种不同大小的 SSA。用动态步长 (Dy) 和几种固定步长来测试算法1。结果如表2.1所示，其中展示了算法的运行时间 (单位秒)，预测步的总步数 (#E)，以及一个预测步需要的平均更正次数 (#G)。从表2.1中可以看出步长较大时总的预测步数会降低，但是每步需要的更正步数会相应增加，这是由于步长越大预测的零点越不准。更正步在 $\Delta \geq 1.3$ 时经常不收敛，这说明直接用牛顿迭代法 (相当于只有更正步没有预测步) 求解 QRE 是不可行的，因为很可能采到一个不收敛的初始点。当步长 Δ 很小时，预测的零点通常已经足够准确，不需要额外的更正步，因此平均的更正步数有可能小于 1。可以验证当广告商数量 N 和策略空间大小 M 给定时，算法1的运行时间与预测和更正的总步数呈正相关关系。总体来说，动态步长要优于各个固定步长，且本章提出的算法能高效地求解 SSA 中 QRE。

2.5.2 参数预测算法评估

本节评估算法2在不同实验设定下的性能。令 $B_i = [M], \forall i \in [N]$ ，其他的 SSA 参数的设定同第2.5.1节，这样可以大体计算出 $\alpha_i, \forall i \in [N]$ 的不连续点的个数的下界，具体来说，因为 $0 < \alpha_j < 1$ ，所以 $\frac{\alpha_j b_{jm}}{b_{im}}, 1 \leq m \leq n$ ，一定是 α_i 的不连续点。于是对于每一个广告商来说，其广告点击率的不连续点的个数至少为 $1 + 2 + \dots + M = \frac{M(M+1)}{2}$ 。本节先用算法1来根据生成的参数计算得到 QRE 策略 σ ，然后基于 σ 用算法2来反推这些给定的参数 (N 和 M 除外)。当 σ 给定时，可以通过将公式 (2-46) 中的 $\pi_{ij}, \forall i \in N, j \in [M]$ ，替换为 σ_{ij} 来求得理论上最大的似然函数值 L^* 。本节用 Matlab 2016b 的 “fmincon” 函数 (内点法) 来解决算法2中第 4 和第 8 行中的优化问题，并将算法2 求解得到 \hat{L}^* 与 L^* 之间的差值 $L^* - \hat{L}^*$ 作为衡量算法估测的参数的精度的指标。实验结果显示在图2.2(a)中，其中每一个点都是 100 次实验的平均值，从中可以看出 SSA 的规模并没有显著影响算法的性能。在实际实现当中，全局最优值被初始化为负无穷，每次尝试新的初始迭代点时我们

都会将算法2的运行结果与当前的全局最优值做比较，如果算法2的结果大于当前全局最优值，则更新全局最优值，否则不更新；当连续5个初始迭代点都不能引起全局最优值的更新且迭代点总数大于等于10时，停止尝试新的初始迭代点并返回当前全局最优值对应的参数。图2.2(b)中展示了算法2的总循环（第3行）的平均次数随着SSA的大小的变化曲线，从中观察到虽然 α_i 的最少不连续点以 $O(M^2)$ 增加，但算法2的迭代次数却以逐渐放缓的速度增加，这说明算法在每一次迭代过程中对参数精度的提升能力并没有随着不连续点数的增加而线性地降低，这就保证算法可以应用到更大规模的SSA中。

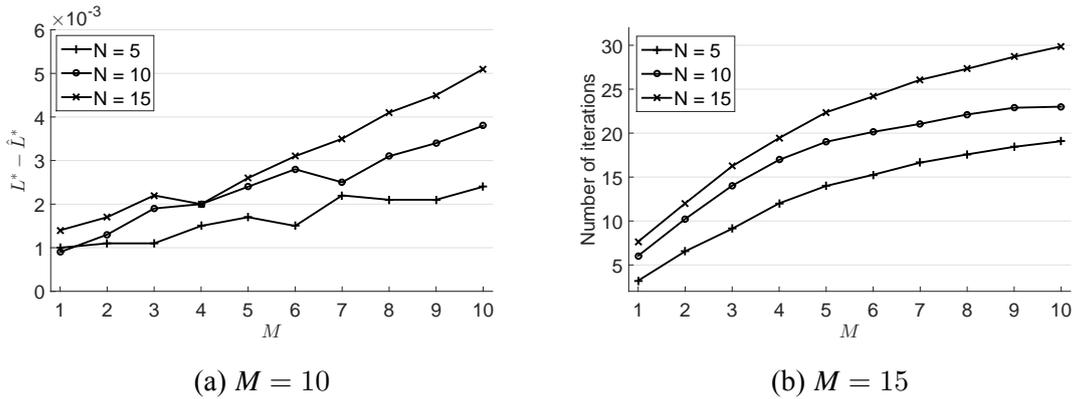


图 2.2 参数估测算法性能评估

2.5.3 QRE 模型对真实数据的解释能力评估

本节用雅虎公开的关于广告竞标的数据集 [51,124,125] 来做实验，这份数据集包含了超过四个月的广告商的竞价和排序信息。超过 89% 的关键词的广告商数量少于 5，数量的具体分布如表 2.2 所示。此外可发现，在很多关键词的竞标中广告商的信息是很不完整的，在四个月中只有几个记录。参照文献 [6] 中的做法，本节挑出 70 个广告商信息相对完整的竞价关键词（不包含仅有一个广告商的关键词），并且进一步删除了那些总是设置一个特别高或者特别低的竞价且从不改变竞价的广告商的信息，因为这些广告商似乎并没有采取某种策略来优化自己的竞价，这些数据对研究广告商行为没有任何启发意义。本节将 QRE 模型拟合到处理过的数据集中来估计 SSA 的参数。具体而言，对于每个关键词，首先用数据计算广告商的混合策略 σ ，然后用算法 2 来估测 v ， λ ， α 和 β ，接下来，通过这些参数代入到公式 (2-7) 中，可以计算出每个广告商 i 的质反应 π_i 。

本节的第一部分实验是检验对于每一个关键词，是否 π_i 等于 σ_i ， $\forall i \in [N]$ ，即检验从数据中得出的广告商的混合策略 σ 是否是一个 QRE。为此，对每一个关键词计算误差

$$\frac{1}{\sum_{i \in [N]} |B_i|} \sum_{i \in [N], j \in [B_i]} |\pi_{ij} - \sigma_{ij}|, \quad (2-53)$$

然后统计该误差的最大、最小和平均值。结果显示在表 2.2 中。从表 2.2 中可以看出，对

表 2.2 QRE 拟合精度

设定	广告商数量	数量占比	最大误差	最小误差	平均误差
1	3	77.14%	0.0732	0.0002	0.0373
2	4	15.71%	0.0813	0.0002	0.0388
3	5	4.29%	0.0922	0.0011	0.0432
4	≥ 6	2.86%	0.1078	0.0106	0.0592

表 2.3 用 QRE 模型估测的参数

关键词	i	v_i	λ_i	α_i	β	L^*
1	1	66.5	85	.04	.07	-3.2804
	2	43.0	1078	.03	.06	
	3	61.4	1602	.03	.05	
2	1	6.00	6251	.13	.15	-2.6098
	2	1.65	222	.65	.10	
	3	5.67	4700	.18	.05	

绝大部分关键词而言最大误差在 0.09 左右，最小误差的数量级为 10^{-4} ，且平均误差小于 0.06，这表明 QRE 模型能很好地拟合现实世界中广告商的行为。

第二部分实验仿照 [55] 的做法用两个具体的关键词的例子来详细分析用 QRE 模型估测的参数。由于 [55] 中提出的对称纳什均衡是一个纯策略均衡，无法解释广告商的混合策略行为，因此本节将 QRE 模型与混合策略纳什均衡 (Mixed Strategy Nash Equilibrium, MSNE) 进行比较来分析二者对真实数据的解释能力。由 QRE 模型估测的参数显示在表 2.3 中，通过将这些参数代入到公式 (2-7) 中，可计算出每一个广告商 i 的质反应 π_i 和期望收益 \bar{u}_i ，如表 2.4 和表 2.5 所示。从表中可以看出更好的纯策略被选择的概率更高，这与 QRE 的基本性质是一致的。表 2.3 表明广告商的理性参数彼此之间相差非常大。需要注意的是公式 (2-7) 表明除了理性参数 λ_i 的量级，广告商 i 采取不同纯策略得到的期望收益之间的差别的量级对其质反应也有很大的影响。以关键词 2 中的广告商 1 为例，从表 2.3 中看到 $\lambda_1 = 6251$ ，虽然这个值很大，但在广告商 1 的质反应 $\pi_1 = (.2682, .2682, .4635)$ 中期望收益最大的策略 3 被选择的概率并不是很大，这是因为广告商 1 的不同纯策略的期望收益之间的差别非常小 (如表 2.5 所示)，大约在 10^{-4} 左右，因此 λ_1 在公式 (2-7) 中的作用被很大程度地减弱了。

本节接下来用 MSNE 来拟合真实数据，并与 QRE 的结果做比较。在 MSNE 中，广告商采取不同纯策略得到的期望收益是相等的，由此可以通过用算法 2 解下面的优化问

表 2.4 关键词 1 中广告商的质反应及收益

σ_1	π_1	\bar{u}_1	σ_2	π_2	\bar{u}_2	σ_3	π_3	\bar{u}_3
.1261	.1239	.1202	.2975	.2979	.0479	.2609	.2607	.0906
.3025	.3563	.1328	.5950	.5948	.0485	.3565	.3697	.0908
.4202	.3714	.1333	.1074	.1073	.0469	.3826	.3697	.0908
.1513	.1484	.1224						

表 2.5 关键词 2 中广告商的质反应及收益

σ_1	π_1	\bar{u}_1	σ_2	π_2	\bar{u}_2	σ_3	π_3	\bar{u}_3
.2927	.2682	.0339	.6897	.6881	.0633	.0522	.0524	.0581
.2439	.2682	.0339	.1466	.1040	.0548	.1304	.1304	.0583
.4634	.4635	.0340	.1207	.1040	.0548	.8174	.8172	.0587
			.0431	.1040	.0548			

题来得到 MSNE 估测的 SSA 的参数:

$$\tilde{L}(v, \alpha, \beta | \sigma) = \log \left(\prod_{i \in [N]} \prod_{j \in [B_i]} \left(\frac{\bar{u}_{ij}(\sigma_{-i})}{\sum_{k \in [B_i]} \bar{u}_{ik}(\sigma_{-i})} \right)^{\frac{1}{|B_i|}} \right) \quad (2-54)$$

$$s.t. \begin{cases} v_i \geq 0, \forall i \in [N]; \\ 1 > \beta_s \geq \beta_{s+1} > 0, \forall s \in [K-1]; \\ 0 < \alpha_i < 1, \forall i \in [N]. \end{cases} \quad (2-55)$$

上面的目标函数的在下式成立时达到最大值:

$$\frac{\bar{u}_{ij}(\sigma_{-i})}{\sum_{k \in [B_i]} \bar{u}_{ik}(\sigma_{-i})} = \frac{1}{|B_i|}, \forall i \in [N], j \in [B_i],$$

上式等价于

$$\bar{u}_{ij}(\sigma_{-i}) = \bar{u}_{ik}(\sigma_{-i}), \forall i \in [N], j \in [B_i], k \in [B_i].$$

估测的参数显示在表 2.6 中。从似然函数的值来看，MSNE 的结果 (\tilde{L}^*) 要低于 QRE 的 (L^*)，这表明 QRE 能比 MSNE 更加准确地模拟广告商的竞价行为。从真实的数据中得到在关键词 1 的竞标中， $B_1 = \{10, 15, 20, 30\}$ ， $B_2 = \{10, 20, 25\}$ ， $B_3 = \{15, 25, 35\}$ 。由 QRE 估测出的广告价值大于广告商真实的竞价，这与平时的经验相一致，即广告商通常不会超限竞价 [64]。比较起来，MSNE 估测出来的参数非常的不合实际。此外，MSNE 估测的关键词 1 的点击率 θ_{ik} 过高，例如 $\alpha_1 \beta_1 = 0.42$ 且 $\alpha_2 \beta_1 = 0.76$ ，但在现实世界中点击率通常小于 10%。数据显示关键词 2 中广告商的最大竞价没有超过 10，然而 MSNE 估测出的广告价值却在 10^6 的量级，这很不可信。由 MSNE 预测的广告位点击率 β 看起来

表 2.6 用 MSNE 模型估测的参数

关键词	i	v_i	α_i	β	\tilde{L}^*
1	1	18.1	.51	.83	-3.5835
	2	12.8	.91	.02	
	3	112.2	.16	.02	
2	1	8.4×10^6	.02	.07	-4.7594
	2	4.8×10^6	.05	.07	
	3	3.08×10^6	.09	.07	

与现实不符，因为真实的广告位点击率通常是按照位置从高到低递减的。相比而言，由 QRE 估测的参数与人们对现实世界的理解始终是一致的。总的来讲，QRE 对真实数据的拟合能力要远远超过 MSNE。

2.6 总结

本章向付费搜索广告竞标中引入了质反应均衡的概念来模拟广告商的有限理性行为。本章主要有两大技术上的贡献。第一，本章开发了一个基于同伦思想的算法来求解广告竞标中的质反应均衡。通过利用广告商的收益函数的特殊结构，本章极大地提高了算法的效率。第二，本章提出了一个参数预测的算法并将质反应均衡模型拟合到一份真实的数据中来估测广告竞标中的参数，包括广告价值和点击率。此外，本章做了大量的实验来评估算法的性能，结果表明本章设计的算法能高效地求解广告竞标中的质反应均衡，且质反应均衡模型能比其他已有的均衡模型更好地拟合现实世界中广告商的行为。

第三章 基于保留价的收益优化

保留价是广告竞标中一个有效的调节收益的工具。最优的保留价依赖于广告商的广告价值的分布。然而，该分布对于搜索引擎来说通常是未知的。搜索引擎的常用做法是首先通过一个采样的过程来收集广告价值分布的信息，然后利用收集到的数据预估一个最优保留价并将其应用到后续的竞标活动中。为了最大化有限时间内的总收益，搜索引擎必须优化采样数来权衡采样阶段的损失和预估的保留价的质量。本章研究基于目前最为流行的 GSP 机制的广告竞标中采样数的优化问题，主要有以下三个创新点。第一，本章给出了采样前和采样后的竞标中收益的竞争比的下界。第二，本章将寻找最优采样数的问题转化成了一个非凸的混合整数优化问题，并证明了最优采样数的存在性和唯一性。第三，通过将上述整数优化问题转换为等价的连续变量优化形式，本章设计了一个高效的算法来求解。实验结果表明相比于基础方法，本章的方法能为搜索引擎带来更高的收益。

3.1 引言

正如前文所述，付费搜索广告竞标已经变成了一个重要的互联网经济活动 [1,50,53]。目前绝大多数搜索引擎使用 GSP 作为竞价机制 [20,55,69]。近年来有大量的工作研究该机制下的收益优化问题 [126–129]。在这些研究当中，保留价的优化是一个热门的话题，因为它能极大地影响搜索引擎的收益 [130–133]。著名的关于最优竞价机制设计的 Myerson 定理 [66–68] 给出了如何优化保留价从而最大化搜索引擎的收益的基本思路：当只有一个广告位且广告商的广告价值满足独立同分布时，最优的竞价机制是将保留价格设置为虚拟价值函数零点的第二价格机制。Edelman 和 Schwarz [12] 将这个理论扩展到了更加通用的情况，即搜索页面有多个广告位，并证明了如果虚拟价格函数满足正则 (regular) 性质，则使用最优保留价格的 GSP 是最优的机制。Thompson 和 Leyton-Brown [14] 研究了当 squashing 也是考虑因素时如何优化保留价。Sun 等人 [69] 经过研究发现，当广告的权重不同时，GSP 中的保留价依赖于具体的广告的权重。

上述所有工作都假设搜索引擎知道广告商的广告价值的分布，在此基础上计算最优保留价格。然而，在现实世界中，广告的价值是广告商的私人信息，搜索引擎必须通过一定的采样来收集分布信息，然后才能计算最优保留价。Ostrovsky 和 Schwarz [9] 利用雅虎的付费搜索广告数据集通过数值仿真得到的保留价格可以有效地提高搜索引擎的收益。Sun 等人 [70] 提出了用最大似然估计的方法来预估分布参数。Cole 和 Roughgarden [71] 则直接用采样的数据来计算最优保留价，而不用先估计分布参数，但他们的方法是基于一一种比较简单的机制，即所有被展示的广告商支付相同的费用，无论排在哪个位置。

Mohri 和 Medina[134] 提出了一个基于对称纳什均衡的算法来预估最优保留价。上述利用采样来计算最优保留价的工作有一个共同的问题，那就是没有考虑由于采样过程使用的保留价并非最优而导致的收益损失。近些年来，学者们提出了一些基于强化学习和数据驱动的方法来计算最优保留价 [135–137]，但这些方法没有提供理论上的保证。一个广告商愿意为自己的广告支付的最大费用通常只在一段时间内是保持不变的，如 iPhone X 刚上市时一个广告商愿意为此支付的广告费用与上市一年后会有明显不同。当考虑广告价值保持不变的有限时间段内搜索引擎的总收益时就要考虑采样数的设置问题。显然，花越多的时间采样对最优保留价的估计越准，在采样结束后的阶段用预估的保留价得到的收益也就越高，但采样阶段的收益损失也随之越多，反之，采样时间越短带来的前期收益损失越小，但对最优保留价的估计也越不准，后期使用估测的保留价的收益损失就越大。因此，搜索引擎必须权衡两方面的收益，决定用多少时间来采样以最大化全局收益，即必须找一个最优采样数。Bulow 和 Klemperer[72] 提出了一些计算采样阶段收益损失的基本方法，但他们的模型与文献 [71,73] 中的一样简化，无法应用到 GSP 中。

本章研究在有限次 GSP 竞标中采样阶段的收益损失和最优保留价计算精度之间的权衡问题。多臂赌博机 (Multi-Armed Bandit, MAB) 理论是用来解决探索和实施之间权衡问题的一个通用框架，该理论已被用于 Post-Price 竞标中 [138]。然而，MAB 理论的算法在 GSP 中的应用是很受限的，因为 1) 它们的算法需要频繁地更新保留价，由于一个关键词组可能在一天中就被搜索上万次，频繁改变保留价在现实世界中开销太大因而不可行；2) 该套理论需要对待优化的函数做一些特殊的假设 [139]，例如平滑和强凸性 [140]、单峰 [141] 或者满足局部 Lipschitz 条件 [142]。但在 GSP 中，搜索引擎的收益函数很难满足这些条件，它甚至都不是保留价的连续函数。

为了解决 GSP 中的采样数权衡问题，本章将有限次的广告竞标分成两个阶段：用于收集广告价值分布信息的采样阶段和利用所采样本计算最优保留价后的应用阶段。本章的目标是找到一个不依赖于广告价值的具体分布形式的最优的采样回合数，主要贡献有三点。第一，本章推导出了采样阶段的收益与理论上最优收益的竞争比。该竞争比是独立于广告价值分布的，仅与广告商和广告位的数量有关。文中假设（未知的）广告价值分布是正则的，但不对其具体是何种分布（如指数分布，正态分布，对数正态分布）做任何额外假设。此外，受文献 [73] 的启发，本章采用了利用采样数据来直接预估最优保留价的方法并证明了其在 GSP 中的收益的下界。第二，本章将这个权衡问题规划成了一个非凸的混合整数优化问题，其目标函数为搜索引擎在有限次竞标中全局收益的下界。此外，本章推导出该目标函数的许多性质，包括非凸性、单峰性和最优解的唯一性。第三，本章设计了一个基于以上性质的高效的算法（指数级收敛速度）。本章用数值仿真来评估算法的性能。实验结果表明，与基础算法相比，搜索引擎利用本章提出的算法可以极大地提高其全局收益。

本章的组织结构如下。第3.2节回顾 GSP 中的基本概念以及介绍保留价优化的基础

知识。问题的形式化定义在第3.3节中给出。第3.4节提出本章最重要的两个用于问题转化的定理。第3.5节对采样数优化问题进行分析，推导目标函数的重要性质。算法于第3.6节中给出，实验验证在第3.7节。本章总结放在最后一节。

3.2 背景知识

本节回顾 GSP 中的基本概念并介绍当广告价值分布已知时求解最优保留价的方法。

3.2.1 GSP 机制

沿用上一章的符号定义，用 N 来表示市场中的广告商的数量，用 K 表示搜索页面广告位的数量。一般来讲， N 远远大于 K 。每一个广告位有一个对应的点击率 β_j , $j \in [K]$ 。对于任意的 $j > K$ ，定义

$$\beta_j \equiv 0. \quad (3-1)$$

这样便可将点击率也表示为一个 N 维的向量 $\beta = (\beta_1, \beta_2, \dots, \beta_N)^T$ ，其中 T 表示矩阵的转置操作。通常来说，有 $\beta_1 \geq \beta_2 \geq \dots \geq \beta_K > 0$ 。广告商 $i \in [N]$ 的广告价值表示为 v_i ，他的出价则写做 b_i 。将所有广告的价值用一个向量 $v = (v_1, v_2, \dots, v_N) = (v_i, v_{-i})$, $\forall i \in [N]$ ，来表示，其中 v_{-i} 表示除了广告商 i 以外的所有广告的价值向量，同理有 $b = (b_1, b_2, \dots, b_N) = (b_i, b_{-i})$ 。不同的广告对用户的吸引力不同，因此每个广告对应一个点击率 α_i ，该因子也被称为广告的权重。当广告 $i \in [N]$ 被列在位置 $j \in [K]$ 时，它被用户点击的概率为 $\alpha_i \beta_j$ 。

GSP 的核心规则由两部分组成：排序规则和收费规则。本节给出两种规则的数学定义。首先，排序规则按照广告的分数的 $s_i = \alpha_i b_i$ 从大到小的将广告从高到低排列，用数学公式可以表示成一个映射： $x : \mathbb{R}^N \mapsto \beta^N$ 。具体而言，给定分数向量 $s = (s_1, s_2, \dots, s_N)$ ，则 $x(s) = (x_1(s), x_2(s), \dots, x_N(s))$ 是一个由 CTR 组成的向量，其中 $x_i(s) = \beta_k$ 当且仅当 s_i 是 s 中第 k 个大的出价（如果有相同的出价，则随机对它们排序）。当考虑保留价格时，排列规则可定义为映射 $x^r : \mathbb{R}^N \mapsto \beta^N$ ，该映射满足：

$$x_i^r(s) = \begin{cases} x_i(s) & \text{如果 } s_i \geq r; \\ 0 & \text{否则.} \end{cases} \quad (3-2)$$

假设广告的分数的满足 $s_i \geq s_{i+1}$ 且定义

$$s_{N+1} \equiv 0, \quad (3-3)$$

那么一旦第 i 个广告商的广告被用户点击，他就需要根据收费规则支付给搜索引擎 $\max\{\frac{s_{i+1}}{\alpha_i}, \frac{r}{\alpha_i}\}$ 。广告商 i 的期望收益可以写成

$$u_i(v_i, s) = v_i \alpha_i x_i^r(s) - p_i(s), \quad (3-4)$$

其中

$$p_i(s) = \alpha_i x_i^r(s) \cdot \max\left\{\frac{s_{i+1}}{\alpha_i}, \frac{r}{\alpha_i}\right\} \quad (3-5)$$

$$= x_i^r(s) \cdot \max\{s_{i+1}, r\} \quad (3-6)$$

是期望支付的费用。搜索引擎的收益等于 $\sum_{i=1}^N p_i(s)$ 。在现实世界中，参与每次竞标的广告商通常是不同的 [139]，同一个广告词组对于不同用户来说其价值通常也是不同的 [143]。这些不确定性可以通过假设广告的权重价值 $e = (\alpha_1 v_1, \alpha_2 v_2, \dots, \alpha_N v_N)$ 服从一个独立同分布来表征 [9,14,65]。此外，正如许多考虑实际情况的文献 [70,71,73] 所假设的那样，这个分布对于搜索引擎是未知的。

3.2.2 权重价值已知时最优保留价的计算

从上一章得知，搜索引擎的期望收益依赖于广告的分值和保留价格。接下来介绍当广告的权重价值分布已知时如何计算最优保留价。在收益优化的理论中，均衡策略的概念被广泛用于对广告商行为建模 [9,12]。Sun 等人 [69] 证明了如果所有广告商都采用同样的均衡策略 $\pi_r: \mathbb{R} \mapsto \mathbb{R}$ 且策略满足

$$s_i = \pi_r(e_i) \leq e_i, \quad (3-7)$$

其中 $e_i = \alpha_i v_i$ ，则搜索引擎在保留价为 r 时每次竞标活动中的期望收益可以写为

$$R(r) = \mathbf{E}_e \left\{ \sum_{i=1}^N \psi(e_i) x_i^r(e) \right\}, \quad (3-8)$$

其中

$$\psi(e_i) = e_i - \frac{1 - F(e_i)}{f(e_i)} \quad (3-9)$$

被称作虚拟价格函数 [67]， $f(\cdot)$ 和 $F(\cdot)$ 分别表示权重价值的概率密度函数和累积分布函数。按照大部分文献里假设的那样 [65,67,144]，假定虚拟价格函数是正则的，即

$$\psi'(x) > 0, \forall x \in (0, +\infty). \quad (3-10)$$

也就是说， $\psi(\cdot)$ 在 $(0, +\infty)$ 是一个严格单调增函数。给定权重价值的分布函数，当且仅当 $\psi(\cdot)$ 满足对任意的 e ， $\sum_{i=1}^N \psi(e_i) x_i^r(e)$ 都能达到最大值时，搜索引擎的收益 $R(r)$ 才能最大化。要达到这一要求，需要满足以下条件：(i) 广告位仅分配给虚拟价格函数为正的广告，(ii) 权重价格越高的广告排位越靠前（即获得越大的广告位点击率 β_j ， $j \in [K]$ ）。由于 $\psi(e_i)$ 和 $x_i^r(e)$ 都是 e_i 的非减函数，满足第二个条件。为了达成第一个要求，搜索引擎只需设置一个最优的保留价 r^* 使其满足

$$\psi(r^*) = 0, \quad (3-11)$$

因为当 $e_i < r^*$, $\forall i \in [N]$ 时, $x_i^{r^*}(e) = 0$ 。将拥有最优保留价的 GSP 称作 OGSP (Optimal GSP), 简称为 OGSP。由于广告的权重价值满足独立同分布, 于是有 $\forall i' \in [N]$,

$$\mathbf{E}_s \left\{ \sum_{i=1}^N \psi(s_i) x_i^{r^*}(e) \right\} = N \cdot \mathbf{E}_s \{ \psi(s_{i'}) x_{i'}^{r^*}(e) \}. \quad (3-12)$$

3.3 问题定义

现在考虑现实世界中广告的权重价值分布未知时搜索引擎的保留价和收益优化的问题。通常搜索引擎只关心有限时间内的收益, 用 M 来表示这段时间总的竞标次数。为了能够观察到完整的 (即没有被截断的) 权重价值分布, 将前 τ 次采样阶段的保留价设为零。注意搜索引擎只能观察到广告商出的价格, 而不是他们的权重价值, 但如果搜索引擎知道广告商的策略 (如 QRE 或 SNE 策略), 则可以反推 $s_i = \pi_r^{-1}(b_i)$ 。为了易于分析, 本章假设广告商策略已知。采样阶段过后, 可以利用已有的数据来计算得到一个预估的最优保留价, 记为 \bar{r} , 然后将它作为后面 $M - \tau$ 次竞标的保留价。搜索引擎的全局收益等于

$$\tau \cdot R(0) + (M - \tau) \cdot R(\bar{r}). \quad (3-13)$$

与理论上 (即当权重价值分布已知时) 能达到的最大收益 $M \cdot R(r^*)$ 相比的全局损失为

$$M \cdot R(r^*) - (\tau \cdot R(0) + (M - \tau) \cdot R(\bar{r})), \quad (3-14)$$

该式等价于

$$M \cdot R(r^*) \left(1 - \left(\frac{\tau}{M} \cdot \frac{R(0)}{R(r^*)} + \left(1 - \frac{\tau}{M} \right) \cdot \frac{R(\bar{r})}{R(r^*)} \right) \right). \quad (3-15)$$

于是, 最大化全局收益等价于最小化全局损失, 可以表示成如下优化问题:

$$\max_{\tau \in [M]} \frac{\tau}{M} \frac{R(0)}{R(r^*)} + \left(1 - \frac{\tau}{M} \right) \frac{R(\bar{r})}{R(r^*)}. \quad (3-16)$$

进一步的分析表明竞争比 $\frac{R(0)}{R(r^*)}$ 和 $\frac{R(\bar{r})}{R(r^*)}$ 都与概率密度函数 $f(\cdot)$ 有关, 因此公式 (3-16) 中定义的问题的解也是 $f(\cdot)$ 的函数, 而本章的目标是找到一个不依赖于未知分布 $f(\cdot)$ 的最优的采样数。为此, 第 3.4 节证明 $\frac{R(0)}{R(r^*)}$ 和 $\frac{R(\bar{r})}{R(r^*)}$ 都有一个与 $f(\cdot)$ 无关的下界。基于这个结果, 第 3.5 节将 $\frac{R(0)}{R(r^*)}$ 和 $\frac{R(\bar{r})}{R(r^*)}$ 换成其对应下界, 从而将原优化问题转换成最大化搜索引擎在最坏情况下的全局收益的问题, 这样得出的结果对不同的分布具有鲁棒性, 对搜索引擎来说是个更安全的策略。

3.4 竞争比下界的推导

为了求解公式 (3-16) 中定义的优化问题, 首先需要推导竞争比 $\frac{R(0)}{R(r^*)}$ 和 $\frac{R(\bar{r})}{R(r^*)}$ 的下界。定理 1 给出了采样阶段的界。

定理 1 对于有 N 个广告商和 K 个广告位的 GSP, $\frac{R(0)}{R(r^*)} \geq \frac{N-K}{N}$ 。

该定理的证明基于下面两个引理。

引理 1 有 N 个广告商、 K 个广告位和零保留价的 GSP 在每次竞标中的期望收益大于等于有 $N-K$ 个广告商和 K 个广告位的 OGSP 在每次竞标中的期望收益。

证明 有 $N-K$ 个广告商和 K 个广告位的 OGSP 在每次竞标中的期望收益等于

$$\mathbf{E}_{e^{N-K}} \left\{ \sum_{i=1}^{N-K} \psi(e_i) x_i^{r^*}(e^{N-K}) \right\}, \quad (3-17)$$

其中 e^{N-K} 是 $N-K$ 个广告商的权重价格向量。从第 3.2.2 节得知分数大于 r^* 的排名最靠前的 $k \leq \min\{K, N-K\}$ 个广告将会被按照分数从高到低的顺序被显示到前 k 个广告位。用函数 $\max_K: \mathbb{R}^N \mapsto \mathbb{R}^K$ 来表示含有 N 个元素的向量到它的前 K 个按从大到小排列的最大元素的子向量的映射。这样可以将最优的期望收益表示成

$$\mathbf{E}_{e^{N-K}} \{ \max_K(\psi(e_1), \psi(e_2), \dots, \psi(e_{N-K}), 0^K) \cdot \beta \}, \quad (3-18)$$

其中 0^K 是一个含有 K 个零的向量。当再有 K 个广告商, 以 $N-K+1, N-K+2, \dots, N$ 来标示, 加入到竞标中来时, 有

$$\mathbf{E}_{e^K} \{ \psi(e_i) \} = \begin{cases} \psi(e_i), & i = 1, 2, \dots, N-K; \\ \mathbf{E}_{e_{-i}^K} \{ \mathbf{E}_{e_i^K} \{ \psi(e_i) \} \} = \mathbf{E}_{e_{-i}^K} \{ 0 \} = 0, & i = N-K+1, N-K+2, \dots, N, \end{cases} \quad (3-19)$$

其中 e^K 表示新加入的 K 个广告商的权重价值分布。于是,

$$\begin{aligned} & \mathbf{E}_{e^{N-K}} \{ \max_K(\psi(e_1), \dots, \psi(e_{N-K}), 0^K) \cdot \beta \} \\ = & \mathbf{E}_{e^{N-K}} \{ \max_K(\mathbf{E}_{e^K} \{ \psi(e_1) \}, \dots, \mathbf{E}_{e^K} \{ \psi(e_N) \}) \cdot \beta \} \end{aligned} \quad (3-20)$$

$$\leq \mathbf{E}_{e^{N-K}} \{ \mathbf{E}_{e^K} \{ \max_K(\psi(e_1), \dots, \psi(e_N)) \cdot \beta \} \} \quad (3-21)$$

$$= \mathbf{E}_e \{ \max_K(\psi(e_1), \psi(e_2), \dots, \psi(e_N)) \cdot \beta \} \quad (3-22)$$

$$= \mathbf{E}_s \left\{ \sum_{i=1}^N \psi(e_i) x_i(e) \right\}, \quad (3-23)$$

其中最后一个方程等于有 N 个广告商和 K 个广告位的 GSP 在零保留价下每次竞标中的期望收益。□

引理 2 有 $N-K$ 个广告商和 K 个广告位的 OGSP 在每次竞标中的期望收益大于等于有 N 个广告商和 K 个广告位的 OGSP 的 $\frac{N-K}{N}$ 倍。

证明 在有 $N-K$ 广告商和 K 个广告位的 OGSP 中, 对于任意的广告商 $i \in [N]$ 来说, 新的广告商加入时不会提升 i 的排位, 也就是说, $x_i^{r^*}(e^{N-K}) \geq x_i^{r^*}(e^{N-K}, e^K)$ 。由此有以下公式:

$$\begin{aligned} & \mathbf{E}_{e^{N-K}}\{\psi(e_i)x_i^{r^*}(e^{N-K})\} \\ = & \mathbf{E}_{e^K}\{\mathbf{E}_{e^{N-K}}\{\psi(e_i)x_i^{r^*}(e^{N-K})\}\} \end{aligned} \quad (3-24)$$

$$\geq \mathbf{E}_{e^K}\{\mathbf{E}_{e^{N-K}}\{\psi(e_i)x_i^{r^*}(e^{N-K}, e^K)\}\} \quad (3-25)$$

$$= \mathbf{E}_s\{\psi(e_i)x_i^{r^*}(e)\}. \quad (3-26)$$

分别对第一个公式和最后一个公式乘以 $N-K$ 和 $\frac{N-K}{N} \cdot N$ 即可完成引理的证明。 \square

接下来首先介绍利用采样数据计算 \bar{r} 的方法, 然后在定理2中推导竞争比 $\frac{R(\bar{r})}{R(r^*)}$ 的下界。

最优保留价 r^* 满足 $\psi(r^*) = 0$, 等价于 $-\psi(r^*)f(r^*) = 0$ 。读者可以验证函数 $-\psi(r)f(r)$ 的原函数是 $P(r) = r(1-F(r))$ 。分布 $F(\cdot)$ 的正则性暗含着 $P(r)$ 在区间 $(0, r^*)$ 上随着 r 递增, 在区间 $(r^*, +\infty)$ 上随着 r 递减, 因此在 r^* 处达到最大值。给定从 $F(\cdot)$ 采的 m 个样本, 将其重命名以满足 $e_{(1)} \geq e_{(2)} \geq \dots \geq e_{(m)}$, 然后一个直观的计算最优保留价的想法就是将其看做对 $1-F(e_{(i)})$ 的一个近似, $i \in [m]$, 接着解以下问题: $\arg \max_{e_{(i)}} \frac{i}{m} e_{(i)}$, s.t. $i \in [m]$ 。这种方法有可能在 $F(\cdot)$ 是一个重尾 (heavy-tailed) 分布时得到一个过大的值。解决这个问题的常用做法是去除值过大的样本, 即利用下式估算最优保留价 [71,73]:

$$\bar{r} = \arg \max_{e_{(i)}} \frac{i}{m} \cdot e_{(i)}, \quad \text{s.t. } \delta m \leq i \leq m. \quad (3-27)$$

其中 $0 < \delta < 1$ 是一个精度参数。下面的引理给出了从 $P(r^*)$ 的视角看 \bar{r} 达到 $(1-\epsilon)$ 精度时的采样复杂度, 这将作为后面推导 $\frac{R(\bar{r})}{R(r^*)}$ 的下界的核心引理。

引理 3 ([73] 的引理 4.1^①) 对于任意的满足正则性的分布 $F(\cdot)$ 和 $0 < \epsilon < 1$, 下面的声明成立: 公式 (3-27) 在采样数

$$m = c\epsilon^{-3} \ln \epsilon^{-1} \quad (3-28)$$

时计算得到预估保留价 \bar{r} 的精度为 $1-\epsilon$ 的概率至少为 $1-\epsilon$, 即

$$P(\bar{r}) \geq (1-\epsilon)P(r^*) \quad (3-29)$$

的概率至少为 $1-\epsilon$, 其中 c 是一个正的常数。

基于引理3, 可以证明下面关于 $\frac{R(\bar{r})}{R(r^*)}$ 的下界的定义。

定理 2 对于 GSP 机制和由公式 (3-27) 在采样数为 $m = c\epsilon^{-3} \ln \epsilon^{-1}$ 时计算得到的保留价 \bar{r} , 有不等式 $\frac{R(\bar{r})}{R(r^*)} \geq (1-\epsilon)^2$ 对任意满足独立同分布的正则的权重价值总是成立的。

证明 给定除 i 之外其他广告的权重价值向量 e_{-i} , 令 t 表示在不考虑保留价的情况下广告商 i 的广告能被显示在页面上所需要的最低价格。在保留价分别为 r^* 和 \bar{r} 时, 广告商 i 给搜索引擎带来的期望收益分别为

$$\int_{\max\{r^*, t\}}^{+\infty} \psi(e_i) f(e_i) x_i(e) de_i \quad (3-30)$$

和

$$\int_{\max\{\bar{r}, t\}}^{+\infty} \psi(e_i) f(e_i) x_i(e) de_i. \quad (3-31)$$

根据 r^* , \bar{r} 和 t 之间的大小关系, 一共要考虑三种不同的情况。

情况 1. $\max\{r^*, \bar{r}, t\} = t$ 。

这种情况下搜索引擎使用保留价 r^* 和 \bar{r} 从广告商 i 那里得到的期望收益是相等的。

情况 2. $\max\{r^*, \bar{r}, t\} = r^*$ 。

假设除了广告商 i , 权重价值大于 r^* 的广告一共有 n 个。那么搜索引擎在保留价最优时从广告商 i 得到的期望收益, R^* , 满足

$$R^* = \beta_{n+1} \int_{r^*}^{e_{(n)}} \phi(e_i) + \beta_n \int_{e_{(n)}}^{e_{(n-1)}} \phi(e_i) + \dots + \beta_1 \int_{e_{(1)}}^{+\infty} \phi(e_i) \quad (3-32)$$

$$= \beta_{n+1}(P(r^*) - P(e_{(n)})) + \beta_n(P(e_{(n)}) - P(e_{(n-1)})) \\ + \dots + \beta_1(P(e_{(1)}) - 0) \quad (3-33)$$

$$= \beta_{n+1}P(r^*) + \sum_{j=1}^n (\beta_j - \beta_{j+1})P(e_{(j)}), \quad (3-34)$$

其中为了简化公式, 令

$$\phi(e_i) = \psi(e_i) f(e_i) de_i. \quad (3-35)$$

假设除了广告商 i , 有 k 个权重价值小于 r^* 但大于 $\max\{\bar{r}, t\}$ 的广告商。当 $k = 0$ 时, 搜索引擎从 i 得到的期望收益在保留价为 \bar{r} 和 r^* 时是相等的。当 $k > 0$ 时, 广告商 i 贡献给搜索引擎的期望收益在保留价为 \bar{r} 时记做 \bar{R} , 它满足

$$\bar{R} = \beta_{n+k+1} \int_{\max\{\bar{r}, t\}}^{e_{(n+k)}} \phi(e_i) + \beta_{n+k} \int_{e_{(n+k)}}^{e_{(n+k-1)}} \phi(e_i) + \dots + \beta_1 \int_{e_{(1)}}^{+\infty} \phi(e_i) \quad (3-36)$$

$$= \beta_{n+k+1}P(\max\{\bar{r}, t\}) + \sum_{j=1}^{n+k} (\beta_j - \beta_{j+1})P(e_{(j)}). \quad (3-37)$$

由于 $P(r)$ 在 $(0, r^*)$ 单调递增, 且

$$r^* > e_{(n+j)} > \max\{\bar{r}, t\} \geq \bar{r} \quad (3-38)$$

对任意的 $j \in [k]$ 都成立, 此外

$$P(\bar{r}) \geq (1 - \epsilon)P(r^*), \quad (3-39)$$

于是

$$P(e_{(n+j)}) > P(\max\{\bar{r}, t\}) \geq P(\bar{r}) \geq (1 - \epsilon)P(r^*) \quad (3-40)$$

对任意的 $j \in [k]$ 都成立。因此有

$$\begin{aligned} \bar{R} \geq & \beta_{n+k+1}(1 - \epsilon)P(r^*) + \sum_{j=n+1}^{n+k} (\beta_j - \beta_{j+1})(1 - \epsilon)P(r^*) \\ & + \sum_{j=1}^n (\beta_j - \beta_{j+1})P(e_{(j)}) \end{aligned} \quad (3-41)$$

$$= (1 - \epsilon)\beta_{n+1}P(r^*) + \sum_{j=1}^n (\beta_j - \beta_{j+1})P(e_{(j)}) \quad (3-42)$$

$$\geq (1 - \epsilon)\beta_{n+1}P(r^*) + (1 - \epsilon) \sum_{j=1}^n (\beta_j - \beta_{j+1})P(e_{(j)}) \quad (3-43)$$

$$= (1 - \epsilon)R^*. \quad (3-44)$$

对第三种情况 $\max\{r^*, \bar{r}, t\} = \bar{r}$ 的分析是类似的。总的来说，搜索引擎在保留价为 \bar{r} 时从广告商 i 处得到的期望收益至少是使用最优保留价 r^* 时的 $(1 - \epsilon)$ 倍。对 i 和 e_{-i} 取期望并考虑到 $(1 - \epsilon)$ 的概率，有 $\frac{R(\bar{r})}{R(r^*)} \geq (1 - \epsilon)^2$ 。□

3.5 采样数优化

本节将求解最优采样数的问题写成一个带约束的优化问题，然后证明解的存在性和唯一性。

考虑到定理1中的

$$\frac{R(0)}{R(r^*)} \geq \frac{N - K}{N} \quad (3-45)$$

和定理2中的

$$\frac{R(\bar{r})}{R(r^*)} \geq (1 - \epsilon)^2, \quad \tau N = c\epsilon^{-3} \ln \epsilon^{-1}, \quad (3-46)$$

最大化搜索引擎在最坏情况下的全局收益可以写为：

$$\max_{\tau} \widehat{\mathcal{R}}(\tau) = \tau \frac{N - K}{NM} + \left(1 - \frac{\tau}{M}\right)(1 - \epsilon)^2, \quad (3-47)$$

$$\text{s.t.} \begin{cases} \tau \in [M]; \\ \tau = \frac{c\epsilon^{-3} \ln \epsilon^{-1}}{N}. \end{cases} \quad (3-48)$$

由于本节推导出的界都与分布无关，使用采样数 τ^* 所得到的期望收益对于分布是鲁邦的，其中 τ^* 是上述优化问题的解。从公式 (3-47) 和 (3-48) 可以看出无法显式地将 ϵ 表

达成一个 τ 的函数，所以本节转面用 ϵ 来表示 τ ，从而得到下面等价的优化问题：

$$\max_{\epsilon} \mathcal{R}(\epsilon) = c\epsilon^{-3} \ln \epsilon^{-1} \frac{N-K}{N^2 M} + \left(1 - \frac{c\epsilon^{-3} \ln \epsilon^{-1}}{NM}\right) (1-\epsilon)^2, \quad (3-49)$$

$$\text{s.t.} \quad 1 \leq \frac{c\epsilon^{-3} \ln \epsilon^{-1}}{N} \leq M. \quad (3-50)$$

经过分析发现上述优化问题的目标函数有以下性质。

定理 3 ($\mathcal{R}(\epsilon)$ 的非凸性) 公式 (3-49) 中的目标函数 $\mathcal{R}(\epsilon)$ 在区间 $(0, 1)$ 上不是上凸(凹)的。

证明 只需证 $\mathcal{R}(\epsilon)$ 的二阶导函数在 $(0, 1)$ 上不是严格小于等于零的。用 $\mathcal{R}''(\epsilon)$ 表示 $\mathcal{R}(\epsilon)$ 的二阶导函数，它满足

$$\mathcal{R}''(\epsilon) = 2 - \frac{c}{NM} (\mathcal{R}'_1(\epsilon) + \mathcal{R}'_2(\epsilon)), \quad (3-51)$$

其中，

$$\mathcal{R}'_1(\epsilon) = \epsilon^{-5} \ln \epsilon^{-1} (2\epsilon^2 - 12\epsilon + 12\frac{K}{N}), \quad (3-52)$$

$$\mathcal{R}'_2(\epsilon) = \epsilon^{-5} (3\epsilon^2 - 10\epsilon + 7\frac{K}{N}). \quad (3-53)$$

很容易验证得到

$$\lim_{\epsilon \rightarrow 0^+} (\mathcal{R}'_1(\epsilon) + \mathcal{R}'_2(\epsilon)) \rightarrow +\infty. \quad (3-54)$$

凸函数 $2\epsilon^2 - 12\epsilon + 12\frac{K}{N}$ 和 $3\epsilon^2 - 10\epsilon + 7\frac{K}{N}$ 的零点分别是 $3 \pm \sqrt{9 - 6\frac{K}{N}}$ 和 $\frac{5 \pm \sqrt{25 - 21\frac{K}{N}}}{3}$ 。通过进一步推导可以得出

$$0 < \frac{5 - \sqrt{25 - 21\frac{K}{N}}}{3} < \eta \leq 1 < \frac{5 + \sqrt{25 - 21\frac{K}{N}}}{3} < 3 + \sqrt{9 - 6\frac{K}{N}}, \quad (3-55)$$

其中

$$\eta = \min\{3 - \sqrt{9 - 6\frac{K}{N}}, 1\}, \quad (3-56)$$

于是有 $\mathcal{R}'_2(\eta) < 0$ 。由于 $\mathcal{R}'_1(\eta) = 0$ ，于是有

$$\mathcal{R}'_1(\eta) + \mathcal{R}'_2(\eta) < 0. \quad (3-57)$$

因为 c 是一个正的常数，所以 $\mathcal{R}''(\eta) > 0$ 。 \square

定理 4 ($\mathcal{R}(\epsilon)$ 的单峰性) 在区间 $(0, 1)$ 上存在一个点 q 满足当 $\epsilon \in (0, q)$ 时 $\mathcal{R}'(\epsilon) > 0$ ， $\mathcal{R}'(q) = 0$ 以及当 $\epsilon \in (q, 1)$ 时 $\mathcal{R}'(\epsilon) < 0$ 。

证明 通过对 $\mathcal{R}(\epsilon)$ 求一阶导得到

$$\mathcal{R}'(\epsilon) = -2(1-\epsilon) - \frac{c \ln \epsilon^{-1}}{NM} \left(-3\frac{K}{N}\epsilon^{-4} + 4\epsilon^{-3} - \epsilon^{-2}\right) + \frac{c}{NM} \left(\frac{K}{N}\epsilon^{-4} - 2\epsilon^{-3} + \epsilon^{-2}\right). \quad (3-58)$$

可以验证 $\mathcal{R}'(\epsilon)$ 在边界的极限值满足

$$\lim_{\epsilon \rightarrow 0^+} \mathcal{R}'(\epsilon) \rightarrow +\infty > 0, \quad (3-59)$$

$$\lim_{\epsilon \rightarrow 1^-} \mathcal{R}'(\epsilon) \rightarrow \frac{c(\frac{K}{N} - 1)}{NM} < 0. \quad (3-60)$$

如果能证明存在 $(0, 1)$ 上存在一点 ϵ' 满足 $\mathcal{R}'(\epsilon)$ 在 $(0, \epsilon')$ 递减, 在 $(\epsilon', 1)$ 递增, 结合前面得到的结果 $\mathcal{R}'(0^+) > 0$ 和 $\mathcal{R}'(1^-) < 0$, 定理4得证。接下来就证明点 ϵ' 是存在的, 这等价于证明 $\mathcal{R}''(\epsilon)$ 在区间 $(0, \epsilon')$ 上为负, 在区间 $(\epsilon', 1)$ 上为正。

还沿用定理中3中的符号。首先能证明二阶导 $\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)$ 在区间 $(0, \eta)$ 上是凸函数 (见引理4)。结合在定理3的证明中得到的结果 $\lim_{\epsilon \rightarrow 0^+} (\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)) \rightarrow +\infty$ 和 $\mathcal{R}_1''(\eta) + \mathcal{R}_2''(\eta) < 0$, 可以推出在 $(0, \eta)$ 上存在一个点 ϵ'' , 满足 $\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)$ 随着 ϵ 从 0^+ 增加到 ϵ'' 而从正无穷单调递减到 0, 然后在 $\epsilon \in (\epsilon'', \eta)$ 时为负。因为 ϵ^{-5} 和 $\ln \epsilon^{-1}$ 在区间 $(0, 1)$ 上总是正的, 所以 $\mathcal{R}_1''(\epsilon)$ 和 $\mathcal{R}_2''(\epsilon)$ 在区间 $(\eta, 1)$ 上负的, 这表明 $\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)$ 在区间 $(\epsilon'', 1)$ 上是负的。

由于 $\mathcal{R}''(\epsilon)$ 和 $\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)$ 的单调性刚好相反, $\mathcal{R}''(\epsilon)$ 随着 ϵ 从 0^+ 增加到 ϵ'' 而从负无穷单调增加到 2, 然后在区间 $(\epsilon'', 1)$ 保持为正。所以, 在区间 $(0, \epsilon'') \subseteq (0, 1)$ 上存在一个点 ϵ' 使得 $\mathcal{R}'(\epsilon)$ 在 $(0, \epsilon')$ 上递减, 在 $(\epsilon', 1)$ 上递增。□

引理 4 ($\mathcal{R}''(\epsilon)$ 的上凸性) $\mathcal{R}(\epsilon)$ 的二阶导函数 $\mathcal{R}''(\epsilon)$ 在 $(0, \eta)$ 是上凸 (凹) 的, 其中

$$\eta = \min\left\{3 - \sqrt{9 - 6\frac{K}{N}}, 1\right\}. \quad (3-61)$$

证明 $\mathcal{R}(\epsilon)$ 的二阶导函数

$$\mathcal{R}''(\epsilon) = 2 - \frac{c}{NM} (\mathcal{R}_1''(\epsilon) + \mathcal{R}_2''(\epsilon)), \quad (3-62)$$

其中

$$\mathcal{R}_1''(\epsilon) = \epsilon^{-5} \ln \epsilon^{-1} \left(2\epsilon^2 - 12\epsilon + 12\frac{K}{N}\right), \quad (3-63)$$

$$\mathcal{R}_2''(\epsilon) = \epsilon^{-5} \left(3\epsilon^2 - 10\epsilon + 7\frac{K}{N}\right). \quad (3-64)$$

仅需证明 $\mathcal{R}_1''(\epsilon)$ 和 $\mathcal{R}_2''(\epsilon)$ 在 $(0, \eta)$ 是凸函数。

首先证明 $\mathcal{R}_1''(\epsilon)$ 的凸性, 它可以表示成以下两部分的乘积:

$$g_1(\epsilon) = \epsilon^{-5} \ln \epsilon^{-1} \quad (3-65)$$

和

$$g_2(\epsilon) = 2\epsilon^2 - 12\epsilon + 12\frac{K}{N}. \quad (3-66)$$

$\mathcal{R}'_1(\epsilon) = g_1(\epsilon)g_2(\epsilon)$ 是凸函数当且仅当其微分函数

$$\lim_{\Delta \rightarrow 0^+} \frac{g_1(\epsilon + \Delta)g_2(\epsilon + \Delta) - g_1(\epsilon)g_2(\epsilon)}{\Delta} \quad (3-67)$$

随着 ϵ 单调非减。令

$$\Delta_{g_i(\epsilon)} = g_i(\epsilon) - g_i(\epsilon + \Delta), \quad i = 1, 2. \quad (3-68)$$

给定 $\Delta > 0$, 有

$$\begin{aligned} & g_1(\epsilon + \Delta)g_2(\epsilon + \Delta) - g_1(\epsilon)g_2(\epsilon) \\ = & (g_1(\epsilon) - \Delta_{g_1(\epsilon)})(g_2(\epsilon) - \Delta_{g_2(\epsilon)}) - g_1(\epsilon)g_2(\epsilon) \end{aligned} \quad (3-69)$$

$$= -\Delta_{g_2(\epsilon)}(g_1(\epsilon) - \Delta_{g_1(\epsilon)}) - \Delta_{g_1(\epsilon)}g_2(\epsilon) \quad (3-70)$$

$$= -\Delta_{g_2(\epsilon)}g_1(\epsilon + \Delta) - \Delta_{g_1(\epsilon)}g_2(\epsilon) \quad (3-71)$$

假设 ϵ_1 和 ϵ_2 在 $(0, \eta)$ 上满足 $\epsilon_1 < \epsilon_2$ 。很容易验证 $g_1(\epsilon)$ 和 $g_2(\epsilon)$ 在 $(0, \eta)$ 是正的单调递减的凸函数, 于是有

$$g_1(\epsilon_1 + \Delta) > g_1(\epsilon_2 + \Delta) > 0, \quad (3-72)$$

$$g_2(\epsilon_1) > g_2(\epsilon_2) > 0, \quad (3-73)$$

$$\Delta_{g_i(\epsilon_1)} > \Delta_{g_i(\epsilon_2)} > 0. \quad (3-74)$$

由于不等式

$$-\Delta_{g_2(\epsilon_1)}g_1(\epsilon_1 + \Delta) - \Delta_{g_1(\epsilon_1)}g_2(\epsilon_1) < -\Delta_{g_2(\epsilon_2)}g_1(\epsilon_2 + \Delta) - \Delta_{g_1(\epsilon_2)}g_2(\epsilon_2) \quad (3-75)$$

在区间 $(0, \eta)$ 上对任意的 $\Delta > 0$ 和 $\epsilon_1 < \epsilon_2$ 都成立, 所以有 $-\Delta_{g_2(\epsilon)}g_1(\epsilon + \Delta) - \Delta_{g_1(\epsilon)}g_2(\epsilon)$ 在 $(0, \eta)$ 上随着 ϵ 单调非减, 故它的极限值也单调非减。由此, $\mathcal{R}'_1(\epsilon) = g_1(\epsilon)g_2(\epsilon)$ 在 $(0, \eta)$ 是凸函数。

接下来证明 $\mathcal{R}'_2(\epsilon)$ 的凸性。 $\mathcal{R}'_2(\epsilon)$ 的二阶导数为 $2\epsilon^{-7}(18\epsilon^2 - 100\epsilon + 105\frac{K}{N})$, 它在 $(0, \frac{50 - \sqrt{2500 - 1890\frac{K}{N}}}{18})$ 上是正的。此外, 可以验证

$$\eta \leq 3 - \sqrt{9 - 6\frac{K}{N}} < \frac{50 - \sqrt{2500 - 1890\frac{K}{N}}}{18}. \quad (3-76)$$

所以 $\mathcal{R}'_2(\epsilon)$ 在 $(0, \eta)$ 上是凸函数 (因为它的二阶导函数为正)。□

正如定理3所声明的, $\mathcal{R}(\epsilon)$ 在区间 $(0, 1)$ 上不是一个上凸函数。幸运的是, $\mathcal{R}(\epsilon)$ 被定理4证明具有单峰的性质。基于这个性质, 有如下推论。

推论 1 (ϵ^* 的存在性和唯一性) 公式 (3-49) 和 (3-50) 中定义的优化问题的最优解 ϵ^* 存在且唯一。

该推论很直观，此处省去证明。考虑到边界条件，最优解 ϵ^* 可用以下公式计算：

$$\epsilon^* = \begin{cases} q, & \text{如果 } g(q) \leq M; \\ g^{-1}(M), & \text{否则,} \end{cases} \quad (3-77)$$

其中函数 $g(\cdot)$ 的定义为

$$y = g(\epsilon) = \frac{c\epsilon^{-3} \ln \epsilon^{-1}}{N}, \quad (3-78)$$

$g^{-1}(\cdot)$ 为函数 $g(\cdot)$ 的逆函数。接下来我们给出如何利用 ϵ^* 来找到公式 (3-47) 和 (3-48) 中定义的优化问题的最优采样数解 τ^* 。因为 $g(\epsilon)$ 在区间 $[g^{-1}(M), 1) \subset (0, 1)$ 上是一个减函数且有 $g(g^{-1}(M)) = M$ 及 $g(1) = 0$ ，所以其逆函数 $g^{-1}(y)$ 在 $y \in (0, M]$ 也是减函数。由此， $\widehat{\mathcal{R}}(y) = \mathcal{R}(g^{-1}(y))$ 与 $\mathcal{R}(\epsilon)$ 在对应的区间上单调性相反。具体来说， $\widehat{\mathcal{R}}(y)$ 的单调增区间和减区间分别为 $(0, g(\epsilon^*)]$ 和 $[g(\epsilon^*), g(l)]$ ，这意味着 $\widehat{\mathcal{R}}(y)$ 在区间 $(0, M]$ 上也是单峰的，且唯一的最大点为 $y^* = g(\epsilon^*)$ 。于是 τ^* 可以用 y^* 来表示。

推论 2 最优采样数 τ^* 满足

$$\tau^* = \begin{cases} \lceil y^* \rceil, & \text{如果 } \widehat{\mathcal{R}}(\lceil y^* \rceil) > \widehat{\mathcal{R}}(\lfloor y^* \rfloor); \\ \lfloor y^* \rfloor, & \text{否则.} \end{cases} \quad (3-79)$$

证明 首先根据公式 (3-47) 和 (3-49) 有 $g(l) = M$ ，于是 $[N] \subset (0, g(l)]$ 。由于 $\lceil y^* \rceil \geq y^* \geq \lfloor y^* \rfloor$ ，根据 $\widehat{\mathcal{R}}(\cdot)$ 的单调性，有

$$\begin{cases} \widehat{\mathcal{R}}(\tau) < \widehat{\mathcal{R}}(\lceil y^* \rceil), & \text{如果 } \tau > \lceil y^* \rceil; \\ \widehat{\mathcal{R}}(\tau) < \widehat{\mathcal{R}}(\lfloor y^* \rfloor), & \text{如果 } \tau < \lfloor y^* \rfloor, \end{cases} \quad (3-80)$$

这意味着 $\widehat{\mathcal{R}}(\tau)$ 要么在 $\lceil y^* \rceil$ 达到最大值（当 $\widehat{\mathcal{R}}(\lceil y^* \rceil) > \widehat{\mathcal{R}}(\lfloor y^* \rfloor)$ 时），要么在 $\lfloor y^* \rfloor$ 达到最大值（当 $\widehat{\mathcal{R}}(\lfloor y^* \rfloor) > \widehat{\mathcal{R}}(\lceil y^* \rceil)$ 时）。□

定理 5 (最优采样数的存在性和唯一性) 最优采样数 τ^* 存在且唯一。

该定理的证明可由推论 1 和 2 直接得出。

3.6 算法设计与分析

基于定理 4 和推论 2，本节设计了算法 3 来计算最优采样数 τ^* 。算法第一步是基于 $\mathcal{R}(\epsilon)$ 的单峰性不停地迭代计算 q （第 4-10 行），即将搜索区间 (l, μ) 初始化为 $(0, 1)$ ，如果当

Algorithm 3: 计算最优采样数

```

1  $l \leftarrow 0$ ;
2  $\mu \leftarrow 1$ ;
3  $\epsilon \leftarrow 0$ ;
4 repeat
5    $\epsilon \leftarrow \frac{l+\mu}{2}$ ;
6   if  $\mathcal{R}'(\epsilon) > 0$  then
7      $l \leftarrow \epsilon$ ;
8   else
9      $\mu \leftarrow \epsilon$ ;
10 until  $\mu - l < \xi$ ;
11  $q \leftarrow \epsilon$ ;
12 基于公式 (3-77) 计算  $\epsilon^*$ ;
13  $y^* \leftarrow g(\epsilon^*)$ ;
14 利用公式 (3-79) 计算  $\tau^*$ ;
15 return  $\tau^*$ ;

```

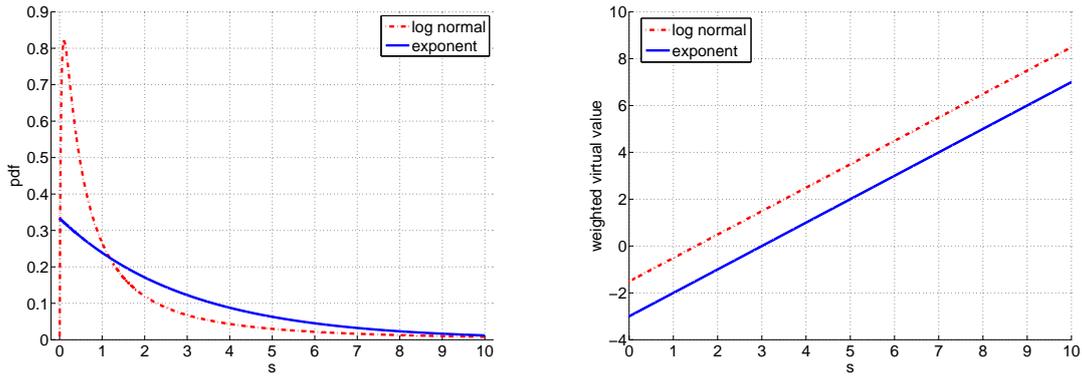
前点 ϵ 在函数 $\mathcal{R}(\cdot)$ 的增区间上, 则在下回迭代中搜索区间 (ϵ, μ) ; 否则, 在下回迭代中搜索区间 (l, ϵ) 。其中, ξ 是给定的精度参数。然后在第 12-14 行分别基于公式 (3-77) 和 (3-79) 计算 ϵ^* 和 τ^* 。下面分析算法的时间复杂度。

定理 6 算法3的时间复杂度为 $O(\log_2(\frac{1}{\xi}))$, 其中 ξ 是给定的精度参数。

证明 根据函数 $\mathcal{R}(\cdot)$ 的单峰性质可得, 在算法第 2-8 行的迭代中, q 始终在区间 (l, μ) 上。因为 ϵ 要么等于 l , 要么等于 μ , 所以当 $\mu - l < \xi$ 时, $|\epsilon - q| < \xi$ 。由于每次迭代都将区间 (l, μ) 缩小为原来的 $\frac{1}{2}$, 则经过 n 轮迭代后, $\mu - l = \frac{1}{2^n}$, 即迭代次数满足 $\frac{1}{2^n} < \xi$ 时算法终止。算法的其他部分只需要常数时间, 因此算法3的总体时间复杂度为 $O(\log_2(\frac{1}{\xi}))$ 。□

3.7 实验评估

本节用文献中最常用的对数正态分布和指数分布来评估算法3的性能。对数正态分布的参数设定为 $\mu = 0$, $\sigma = 1.5$, 对应的理论上的最优保留价格为 $r^* = 1.5105$ 。指数分布的参数设定为 $\lambda = 3$, 对应的最优保留价为 $r^* = 3.0005$ 。两个分布的概率密度函数及虚拟价格函数图3.1所示。本节所有实验均由 python 2.7 实现, 运行在一台搭载了四核 i7 处理器, 运行内存 16G, 操作系统为 64 位 Windows 10 的设备上。通过上一节的分析可知, 算法的时间复杂度不会随着广告商数量 N 、广告位数量 K 和广告竞标总次数 M 变化, 本文统一将 ξ 设置为 10^{-8} , 算法可在 1 秒内返回结果。



(a) 权重价值分布的概率密度函数

(b) 虚拟权重价值函数

图 3.1 权重价值的概率密度函数及虚拟权重价值函数

首先分析算法3的鲁棒性。 M 的设定一般考虑广告权重价值分布保持不变所能进行的竞标活动的总次数，搜索引擎可以通过专家经验或者历史数据来预估这个值，但预估的结果可能不准。令 $\tau^{(5)}$ 和 $\tau^{(10)}$ 分别来表示在对 M 的预估的误差为 5% 和 10% 的情况下由算法3算出的最优采样数。搜索引擎在采样数为 τ^* 、 $\tau^{(5)}$ 和 $\tau^{(10)}$ 时的全局期望收益展示在表3.1-3.4中，其中每个参数设定下的 $\widehat{\mathcal{R}}(\cdot)$ 都是 400 次实验的平均值。搜索引擎在知道广告权重价值分布时能达到的理论上的最大期望收益为 1。可以看到用 τ^* 回合来采样能得到非常高的全局收益，在所有的实验中， $\widehat{\mathcal{R}}(\tau^*) > 0.9$ 。通过比较不同参数下 $\widehat{\mathcal{R}}(\tau^*)$ 与 $\widehat{\mathcal{R}}(\tau^{(5)})$ 和 $\widehat{\mathcal{R}}(\tau^{(10)})$ 的值可得算法3对 M 的估计误差是鲁棒的，因为收益受该误差的影响很小。实验结果表明搜索引擎的收益随着 N 的增加而增加，随着 K 的增加而减少，这主要是因为采样阶段的收益损失与 N 呈负相关关系，而与 K 呈正相关关系（见定理1）。

表 3.1 针对对数正态分布的性能评估 ($N = 5$)。

M	K	$\widehat{\mathcal{R}}(\tau^*)$	$\widehat{\mathcal{R}}(\tau^{(5)})$	$\widehat{\mathcal{R}}(\tau^{(10)})$	$G(S_1)$	$G(S_2)$	$G(S_3)$
5e+6	2	.9199	.9174	.9148	13.2%	0.54%	16.2%
	3	.9079	.9047	.9015	6.65%	0.61%	45.5%
	4	.9005	.8969	.8933	3.14%	4.10%	101%
5e+7	2	.9614	.9601	.9587	1.67%	1.15%	20.2%
	3	.9513	.9496	.9478	0.28%	5.37%	50.8%
	4	.9449	.9428	.9408	0.09%	11.6%	114%

接下来将 τ^* 与三个基础策略做比较。前两个策略分别记为 S_1 和 S_2 ，它们分别用前 $0.1 \cdot M$ 回合和前 $0.3 \cdot M$ 回合来采样。第三个基础策略 S_3 作为一个基准，在整个 M 回合的竞标中一直将保留价设置为零。令 $G(S_j)$ 表示本章提出的策略与三个基础策略， S_j ,

表 3.2 针对对数正态分布的性能评估 ($N = 500$)。

M	K	$\widehat{\mathcal{R}}(\tau^*)$	$\widehat{\mathcal{R}}(\tau^{(5)})$	$\widehat{\mathcal{R}}(\tau^{(10)})$	$G(S_1)$	$G(S_2)$	$G(S_3)$
5e+6	2	.9678	.9657	.9642	6.07%	2.20%	12.1%
	3	.9662	.9650	.9641	5.14%	1.68%	13.2%
	4	.9655	.9601	.9589	3.23%	0.52%	14.3%
5e+7	2	.9752	.9728	.9701	5.25%	2.23%	15.6%
	3	.9701	.9688	.9656	4.94%	1.62%	17.2%
	4	.9662	.9652	.9643	4.05%	0.91%	20.1%

表 3.3 针对指数分布的性能评估 ($N = 5$)。

M	K	$\widehat{\mathcal{R}}(\tau^*)$	$\widehat{\mathcal{R}}(\tau^{(5)})$	$\widehat{\mathcal{R}}(\tau^{(10)})$	$G(S_1)$	$G(S_2)$	$G(S_3)$
5e+6	2	0.9677	0.9542	0.9520	11.2%	0.48%	15.3%
	3	0.9446	0.9401	0.9356	5.64%	0.56%	43.5%
	4	0.9186	0.9045	0.8993	3.10%	3.89%	89.1%
5e+7	2	0.9842	0.9837	0.9705	1.55%	1.03%	18.7%
	3	0.9700	0.9652	0.9595	0.22%	4.97%	47.6%
	4	0.9542	0.9450	0.9400	0.08%	11.2%	105%

$j = 1, 2, 3$, 相比全局收益增加的相对量, 定义如下:

$$G(S_j) = \frac{\mathcal{R}(\tau^*) - \mathcal{R}(S_j)}{\mathcal{R}(S_j)}, j = 1, 2, 3. \quad (3-81)$$

$G(S_j)$ 的值越大, 本章的策略与 S_j 相比就越好。从表3.1-3.4可以看出, 本章的方法得出的策略在所有实验设定下都要优于其他三个策略。可以发现当 N 比较小时, $G(S_1)$ 随着 K 递减, 而 $G(S_2)$ 随着 K 增加, 当 N 比较大时, $G(S_1)$ 和 $G(S_2)$ 都随着 K 而减小。这是由于在其他参数固定时, N 越大则最优采样数越大 (因为 N 越大采样阶段的损失越小), 因而 $N = 5$ 时最优采样数落在了 $0.1M$ 和 $0.3M$ 之间, $N = 500$ 时最优采样数大于 $0.3M$, 同时, 最优采样数随着 K 的增大而减小, S_j 越靠近最优采样数其对应的 $G(S_j)$ 越小。此外从表格中还发现, N 越小, 不设置保留价的损失越大。总体而言, 当广告商的权重价值服从指数分布时, 算法得到的收益略微高于使用对数正态分布, 但本文提出的算法不依赖具体的分布形式, 因而在两种分布中都比其他策略的性能好。与 S_3 的比较表明设置合适的保留价能极大地提高搜索引擎的收益。

3.8 总结

本章研究有限次的 GSP 广告竞标活动最优采样数的设置问题。本章分析了采样中和采样后搜索引擎的收益的竞争比, 并将原问题写成了一个带有约束的混合整数非凸优化问题。本章证明了最优采样数的存在性和唯一性。实验结果表明本章提出的方法优于

表 3.4 针对指数分布的性能评估 ($N = 500$)。

M	K	$\widehat{\mathcal{R}}(\tau^*)$	$\widehat{\mathcal{R}}(\tau^{(5)})$	$\widehat{\mathcal{R}}(\tau^{(10)})$	$G(S_1)$	$G(S_2)$	$G(S_3)$
5e+6	2	0.9857	0.9823	0.9786	5.89%	1.78%	11.7%
	3	0.9812	0.9794	0.9785	4.76%	1.56%	12.2%
	4	0.9762	0.9728	0.9711	3.33%	0.53%	13.8%
5e+7	2	0.9921	0.9896	0.9861	4.95%	1.79%	13.8%
	3	0.9878	0.9843	0.9814	4.33%	1.52%	16.6%
	4	0.9864	0.9820	0.9795	3.18%	0.84%	19.4%

基础方法。

第四章 马尔科夫完美均衡和长远利益优化

本章着重于研究云计算服务商基于长远利益的价格优化，主要考虑到市场的三个重要性质：1) 服务商仅知道市场中的不完全信息；2) 市场在不断发展中，因为用户数量在不断增加，但平台的边际成本在降低；3) 由于服务商和用户的逐利特性，市场充满竞争。本章首先提出一个随机博弈 (Stochastic Game) 模型来模拟云计算市场。为了便于分析，本章将时间离散化并假设在任意两个相邻时间点之间服务商的价格保持不变，下一章研究服务商在任意时间区间的动态定价策略的优化问题。本章用马尔科夫完美均衡 (Markov Perfect Equilibrium, MPE) 来定义服务商基于长远利益的最优价格策略，然后将计算 MPE 的问题分解为两个子任务：1) 将原始的随机博弈分解为多个标准形式博弈并计算每个标准型子博弈的均衡策略；2) 基于第一步得到的结果和一个简单的假设来计算原始的随机博弈的 MPE 策略。实验结果表明本章的算法对于求解 MPE 是十分高效的，且 MPE 策略比其他已有的基于长远利益的价格策略带来的收益更高。

4.1 引言

在过去的几年里，云计算服务在工业界取得了巨大的成功，因为它将人们从安装、配置、维护和更新一些列硬件和软件的体力劳动中解放了出来。越来越多的公司企业和个人和用户开始使用云计算服务，这导致市场的增长十分迅速。理性的用户总是会选择能最好地满足自己需求的服务商。尽管提高服务质量是重要的提高收益的途径，价格策略也是一项不可或缺的吸引用户的方法，尤其当市场中有多个竞争者时 [145–147]。考虑到市场的快速发展，如何优化定价策略来使得长远的总利益最大化对于服务商来说是一个重要且具有挑战性的问题。

本章着重于解决基于长远利益的价格策略优化问题，主要有以下两点贡献。第一，本章提出了一个实际的可以精确描述市场发展的模型。具体来说，本章的模型考虑到了云计算市场的如下性质：1) 服务商并不知道每个用户的需求量、收益函数和偏好，即用户信息对于服务商来说是不完全的；2) 市场是在不断发展的，即用户量在市场饱和前会持续增加，与此同时，随时技术的发展服务商的边际成本会持续降低；3) 市场中通常有多个服务商，会形成竞争。尽管已经有一些计算服务商基于长远利益的最优定价策略的工作，但没有一个考虑了上述所有特征，即不完全信息、市场的发展和市场竞争。文献 [148] 中提出了一个非合作竞争模型，其关注于求解一次性博弈中的均衡策略，并没有考虑长期收益。一些工作 [83,84] 假设市场中只有一个服务商，这与当今真实世界中的云市场相违背。其他一些工作分析用户对服务商价格的反应行为，但忽略了服务商之间的竞争 [10,85]。文献 [11] 中研究了竞争的市场，但只简单假设只有一个主动优化策

略的服务商，其他服务商都跟随前者相应调整价格，这种假设过强，现实世界中服务商（尤其大公司）往往会主动制定自己的价格策略。文献 [86] 考虑了竞争，但没有考虑市场的长远发展，且他们的算法只能被用于仅有几个用户的小规模市场。一个离散选择模型在 [87] 中被提出，用于描述用户的选择行为，但却忽略了用户信息对服务商来说是不完全的这一事实。本章将市场抽象成一个随机博弈 [149]，其中玩家独立且同时选择他们各自的策略。使用十分广泛的方案概念 MPE [150,151] 被用来模拟服务商的行为。由于市场中的不完全信息，庞大的用户数量和状态空间，求解 MPE 策略是一个很难的问题。

本章的第二个贡献在于设计高效的算法来计算 MPE。原始的随机博弈被分解成许多标准形式子博弈，然后原问题可以分两步解决。第一步，计算每个标准型子博弈在不完全信息下的纳什均衡 (Nash Equilibrium under Incomplete Information)，以下简称 NE；第二步，将这些 NE 整合起来构建完整随机博弈的 MPE。最常用的均衡计算方法为最优响应迭代 [68,86,123,152]，该算法被证明在拥塞博弈和势博弈中一定收敛 [68,152]。在标准型子博弈中，这类算法可能陷入无限循环无法收敛 [123]。本章设计了一个计算 NE 的算法，如果 NE 存在，则算法一定会收敛。此外，本章还讨论如何扩展算法来解决 NE 不存在的问题。整合这些 NE 来计算原始完整的随机博弈的 MPE 可以使用 Q-learning 类型的算法 [153,154]，然而，这类算法效率很低，因为它们需要不停地迭代更新每个状态的值直到最后收敛。本章提出了一个基于动态规划的算法，无需迭代求解。该算法建立在一个简单的假设上，即服务商在下一个时间点的价格不会高于前一个时间点，亦即广告商不会提高价格 [11,35,90]。

本章做了大量实验来评估算法的有效性，并展示了服务商如何根据 MPE 来设置价格。实验结果表明，与已有的基于长远利益的策略相比，MPE 策略可以为服务商带来极大的性能提升。

本章的剩余部分组织如下。第 4.2 节首先提出一个考虑了不完全信息、发展和竞争的市场模型，并将其规划成一个随机博弈，然后给出 MPE 的定义。第 4.3 节介绍如何分解求解 MPE 的问题，并设计相应的算法。第 4.4 节做大量的实验来评估算法性能。总结在最后一节给出。

4.2 建模

本节首先提出涵盖不完全信息、发展和竞争的市场模型，然后将其规划成一个随机博弈并给出基于长远利益的均衡策略的定义。

4.2.1 不完全信息建模

本章考虑时间离散化的云计算服务市场。令 $t \in \mathbb{N}^+$ 表示离散的时间点（阶段）， $N(t) \in \mathbb{N}^+$ 表示阶段 t 时用户的数量。假设市场中一共有 $K \in \mathbb{N}^+$ 个服务商。用 $[K]$ 代表

集合 $\{1, 2, \dots, K\}$ 。对每个用户 $j \in [N(t)]$ ，他选择服务商 i 能获得的收益为

$$u_{j,i}^t(p_{i,t}) = d_{j,t}(\theta_j - p_{i,t} + \phi_{j,i}), \quad (4-1)$$

其中 $d_{j,t}$ 是用户 j 在阶段 t 的需求量， θ_j 是用户 j 的边际收益， $p_{i,t}$ 是服务商 i 在阶段 t 设定的价格， $\phi_{j,i}$ 代表用户 j 对服务商 i 的偏好。用户需求可以被看成是其在阶段 t 需要的虚拟机（或其他设备）的数量，偏好用于表征用户对于一个服务商的忠诚度 [86,148]。理性的用户总会选择能给自己带来最大收益的服务商。令 $N_i(t, p_t)$ 表示在阶段 t 选择服务商 i 的用户的集合，则有^①

$$N_i(t, p_t) = \{j \in [N(t)] | u_{j,i}^t(p_{i,t}) > \max\{0, u_{j,i'}^t(p_{i',t})\}; \forall i' \neq i\}. \quad (4-2)$$

其中 $p_t = (p_{1,t}, p_{2,t}, \dots, p_{K,t})$ 。

已有的工作假设服务商精确地知道每个用户的私人信息，包括他们的需求、收益函数和用户偏好 [86,148]，但在现实世界中，这种假设是不成立的。尽管文献 [11] 中假设这些信息服从一定的分布，但其算法本质上还是依赖于用户的准确信息，而非分布信息。本章从现实角度出发，假设服务商仅知道用户私人信息的概率分布。为了便于跟已有的工作比较 [11,155,156]，本章采用如下假设^②：

- 用户在阶段 t 的需求是独立同分布的，且服从一个指数分布，该分布的概率密度函数为

$$f(x | \lambda(p_{t-1})) = \lambda(p_{t-1}) e^{-\lambda(p_{t-1})x}, \quad (4-3)$$

其中

$$\lambda(p_{t-1}) = \frac{\Lambda}{K} \sum_{i=1}^K p_{i,t-1}, \quad (4-4)$$

Λ 是分布的参数。这个公式与经验相符，即降低价格会导致用户需求增加。

- 用户的边际收益也满足独立同分布，服从一个均衡分布，其累积分布函数为 $U(x)$ ，其中 $x \in [U_l, U_u]$ 。
- 用户的对服务商 i 的偏好同样满足独立同分布，服从一个伽马分布，其概率密度函数为

$$g_i(x) = \frac{1}{b_i^{q_i} \Gamma(q_i)} x^{q_i-1} e^{-\frac{x}{b_i}}, \quad (4-5)$$

其中 $\Gamma(\cdot)$ 是伽马函数， b_i 和 q_i 是分布的参数。该分布的累积分布函数记为 $G_i(x)$ 。

① 由于用户需求、收益和偏好都被假定为连续变量，平局出现的概率为零，因此本章没有考虑平局处理的规则。

② 本章中的结果可以被直接应用到其他分布形式，因为本章的算法于分布无关。

4.2.2 市场发展建模

本节介绍市场的发展特性，这主要是因为增长的用户数量的降低的边际成本。

当今的云计算市场正在快速发展，越来越多的用户开始使用云服务。本章用经典的逻辑增长函数 [157,158] 来表征增长的用户数量：

$$N(t) = \lceil \frac{N_0 N_\infty}{N_0 + (N_\infty - N_0)e^{-\kappa t}} \rceil, \quad (4-6)$$

其中 N_0 是最初的用户数量， $N(t)$ 表示在阶段 t 市场中的用户数量， N_∞ 是市场饱和后的用户数量， $\kappa > 0$ 是表征市场发展速度的参数。当 t 足够大时， $N(t)$ 收敛到 N_∞ 。用 T_N 来表示市场达到饱和的时间，其满足当 $t \geq T_N$ 时 $N(t) = N_\infty$ ，否则 $N(t) < N_\infty$ 。

给定 $d_{j,t}$ 、 θ_j 和 $\phi_{j,i}$ 的值， $\forall j \in [N(t)]$ 和 $i \in [K]$ ，可以很容地计算 $N_i(t, p_t)$ ，于是服务商 i 的边际成本可以表示为 [159]

$$c_i(t, p_t) = c_{i,0} \left(\sum_{j \in N_i(t, p_t)} d_{j,t} \right)^{-\beta} \delta(t), \quad (4-7)$$

其中 $c_{i,0}$ 是服务商 i 最开始的边际成本， β 是一个很小的正的常数（通常 $\ll 1$ ），用于表征由于经济效应，边际成本会随着总的用户需求增加而递减。将 $\delta(t)$ 称作时间因子，定义如下：

$$\delta(t) = \begin{cases} e^{-\eta t} & \text{if } t < T_c; \\ e^{-\eta T_c} & \text{否则,} \end{cases} \quad (4-8)$$

其中 T_c 是一个充分大的数， $0 < \eta < 1$ 是一个参数。这个因子表明，随着技术的进步和硬件的发展，边际成本会逐渐降低，但最低有个下界 T_c ，因为成本不可能降为零。

至此为止本节已定义的符号总结在表4.1中。

4.2.3 市场竞争建模

随机博弈模拟多智能体交互系统，其中智能体之间相互独立且无协作地同时行动 [149]。

定义 4.1 一个拥有 K 个玩家的随机博弈是一个元组 $\langle S, A_1, A_2, \dots, A_K, r_1, r_2, \dots, r_K, \gamma, \varphi \rangle$ ，其中

- S 是状态空间；
- A_i 是玩家 i 的纯策略空间；
- $r_i : S \times A \rightarrow \mathbb{R}$ ， $A = \times_{i=1}^K A_i$ ，是玩家 i 的即时收益函数；
- γ 是折损因子；
- $\varphi : S \times A \times S \rightarrow [0, 1]$ 是状态转移概率函数。

表 4.1 长远利益符号总结

t	离散时间点（阶段）
$N(t), N_0, N_\infty$	阶段 t 、初始阶段和饱和阶段的用户数量
K	服务商数量
$[K]$	集合 $\{1, 2, \dots, K\}$
κ	市场发展速度参数
β	经济效应参数
T_c	市场饱和时间点
η	边际成本参数
$p_{i,t}$	服务商 i 在阶段 t 设置的价格
p_t	服务商的价格向量 $(p_{1,t}, p_{2,t}, \dots, p_{K,t})$
$N_i(\cdot)$	服务商 i 的用户集合
$c_{i,0}, c_i(t, \cdot)$	服务商 i 在初始和阶段 t 的边际成本
$d_{j,t}$	用户 j 在时刻 t 的需求量
θ_j	用户 j 的边际收益
$\phi_{j,i}$	用户 j 对服务商 i 的偏好
$u_{j,i}^t(\cdot)$	用户 j 在时刻 t 选择服务商 i 能获得的收益
$f(\cdot), \Delta$	用户需求的概率密度函数和分布参数
$U(\cdot), U_l, U_u$	用户边际收益的累积分布函数和分布参数
$g_i(\cdot), G_i(\cdot), b_i, q_i$	用户对服务商 i 的偏好的概率密度函数、累积分布函数和分布参数

在一个随机博弈中，给定当前的状态 $s \in S$ ，每个玩家 $i \in [K]$ 采取一个纯策略 $a_i \in A_i(s)$ 并得到一个即时收益 $r_i(s, a)$ ，其中 $A_i(s) \subset A_i$ 是玩家 i 在状态 s 的可行集（可行纯策略集合）且 $a = (a_1, a_2, \dots, a_n)$ 。博弈的状态以概率 $\varphi(s, a, s')$ 从 s 转移到 $s' \in S$ 。令

$$\pi_i : S \rightarrow A_i \quad (4-9)$$

表示玩家 i 的策略，玩家 i 基于该策略的长期收益（未来折损收益总和）被定义为

$$V_i(s, \pi(s)) = r_i(s, \pi(s)) + \gamma \sum_{s' \in S} \varphi(s, \pi(s), s') V_i(s', \pi(s')), \quad (4-10)$$

其中 $\pi(\cdot) = (\pi_1(\cdot), \pi_2(\cdot), \dots, \pi_n(\cdot))$ 是所有玩家的策略向量。在云服务市场，纯策略空间 $A_i, \forall i \in [K]$ ，包含有限个元素，因为纯策略对应于离散的价且真实世界中的服务商不能设置无限大的价格。给定 p_{t-1} 、 t 和 p_t ，服务商 i 的期望即时收益可以被写为

$$r_i(p_{t-1}, t, p_t) = \mathbf{E}_i^{p_{t-1}} \left\{ \sum_{j \in N_i(t, p_t)} d_{j,t} (p_{i,t} - c_i(t, p_t)) \right\}, \quad (4-11)$$

其中 $\mathbf{E}_i^{p_{t-1}}\{\cdot\}$ 是对分布 $f(\cdot | \lambda(p_{t-1}))$ 、 $U(\cdot)$ 和 $g_i(\cdot)$ 取的期望。上式的详细计算过程将在第4.3.4节给出。上式表明，一个服务商的收益不仅与自己设定的价格有关，还依赖于其他服务商的定价。因此，服务商需要通过制定合适的价格策略与他人竞争从而获得更高

的收益。云市场中的状态 s 可以被表示为一个元组 (p_{t-1}, t) ，纯策略向量 a 对应于服务商的价格 p_t 。于是有

$$S = A \times \mathbb{N}^+. \quad (4-12)$$

本章的模型中的状态转移概率 $\varphi(\cdot)$ 的计算非常直接。给定当前状态 $s = (p_{t-1}, t)$ 和价格（纯策略）向量 $p_t \in A$ ，状态以概率 1 转移到 $(p_t, t+1)$ ，即，

$$\varphi(s, p_t, s') = \begin{cases} 1 & \text{if } s' = (p_t, t+1); \\ 0 & \text{否则.} \end{cases} \quad (4-13)$$

用符号

$$\varphi^{-1}(s, p_t) = (p_t, t+1) \quad (4-14)$$

来表示确定性的下一个状态 $(p_t, t+1)$ 。

博弈论在研究定价策略的问题中被广泛使用 [160,161]。由于每个服务商在考虑竞争对手的策略的同时都想最大化自己的长期收益，均衡的方案概念适用于对服务商行为进行建模。在随机博弈中最常用的纯策略均衡是马尔科夫完美均衡，MPE，该均衡是子博弈完美纳什均衡的一个改进形式，已经被用到了云计算市场来描述服务商的行为 [86,87]。令 $\pi_{-i}(\cdot)$ 表示除了 i 其他服务商的策略向量，则可以如下定义 MPE：

定义 4.2 云服务市场的一个 MPE 是一个策略向量 $\pi^*(\cdot)$ ，它满足 $\forall s \in S, i \in [K]$,

$$V_i(s, \pi^*(s)) \geq V_i(s, (\pi_i(s), \pi_{-i}^*(s))), \forall \pi_i(\cdot). \quad (4-15)$$

4.3 计算均衡策略

本节设计一个算法来求解 MPE 策略。正如文献 [11,90] 中指出的以及从现实世界中观察到的那样 [35]，从长远来看，当今市场中的服务商很少会逐步提高价格，因为用户数量随着时间增长的同时边际成本在降低。因此，本章很现实地假设服务商在当前阶段设置的价格不会高于上一个阶段，公式化的定义如下：

假设 4.1 $\forall i \in [K]$ 和 $(p_{t-1}, t) \in S$ ， $A_i(p_{t-1}, t) = \{a_i | a_i \in A_i \wedge a_i \leq p_{i,t-1}\}$ 。

根据定义 4.2，可以将 MPE 策略重写为 $\forall i \in [K]$ 和 $s \in S$ ，

$$\pi_i^*(s) = \arg \max_{a_i \in A_i(s)} \{r_i(s, (a_i, \pi_{-i}^*(s))) + \gamma V_i^*(s')\}, \quad (4-16)$$

其中

$$s' = \varphi^{-1}(s, (a_i, \pi_{-i}^*(s))) \quad (4-17)$$

是下一个状态,

$$V_i^*(s') = V_i(s', \pi^*(s')) \quad (4-18)$$

被称作服务商 i 的状态值。为了方便表示, 定义符号 $V^*(s') = (V_1^*(s'), V_2^*(s'), \dots, V_K^*(s'))$ 。

Q-learning 类型的算法可被用来计算随机博弈中的 $\pi_i^*(s)$ [86,162], 然而, 这类算法需要重复更新所有 $i \in [K]$ 和 $s \in S$ 的状态值 $V_i^*(s)$ 直到收敛到一个近乎最优的解。Q-learning 算法中的 Q-value 定义如下:

$$Q_i^*(s, a_i) = r_i(s, a_i, \pi_{-i}^*(s)) + \gamma V_i^*(s'), \quad a_i \in A_i(s_i). \quad (4-19)$$

上式中, $Q_i^*(s, a_i)$ 是 i 在状态 s 采取纯策略 a 并在之后采用 MPE 策略能得到的长期收益。公式 (4-16) 也可以写为

$$\pi^*(s) = \arg \max_{a_i \in A_i(s_i)} Q_i^*(s, a_i), \quad (4-20)$$

这表明一旦知道了所有状态和纯策略的 $Q_i^*(s, a_i)$, 可以直接计算 $\pi_i^*(\cdot)$ 。Q-learning 算法随机初始化所有 $i \in [K]$ 、 $s \in S$ 和 $a_i \in A_i$ 的 $Q_i(s, a_i)$, 然后基于下面的公式更新 Q-value:

$$Q_i(s, a_i) \leftarrow r_i(s, a_i, \pi_{-i}(s)) + \gamma \max_{a'_i \in A_i(s'_i)} Q_i(s', a'_i), \quad (4-21)$$

其中 $\pi_i(s) = \arg \max_{a_i \in A_i(s_i)} Q_i(s, a_i)$ 。理论上讲, 每一个 Q-value 需要被更新无限次以保证它能收敛到精确的 $Q_i^*(s, a_i)$, 当状态空间很大时, 这种操作是非常耗时的。在本章的模型中, 状态空间的大小为

$$|S| = T \prod_{i \in [K]} |A_i|, \quad (4-22)$$

是非常大的, 因此用 Q-learning 算法的效率很低。下一节将设计一个基于假设 4.1 的高效的算法, 每个状态的值最多被更新一次且算法可以得到精确的 MPE 策略。

4.3.1 问题定义

从公式 (4-16) 可以看出, 如果给定所有的 $a \in A(s)$ 的状态值 $V^*(\varphi^{-1}(s, a))$, 状态 s 的 MPE 策略 $\pi^*(s)$ 等价于一个有 K 个玩家的标准型子博弈 [68], $\Phi(s)$, 的 NE 策略。该标准型子博弈中玩家 i 的长期收益函数为

$$r_i(s, a) + \gamma V_i^*(\varphi^{-1}(s, a)). \quad (4-23)$$

基于此, 可以将计算完整的随机博弈的 MPE 策略的问题分解两个子问题, 第一个是计算一个给定状态 s 对应的标准型子博弈 $\Phi(s)$ 在所有的 $\varphi^{-1}(s, a)$, $\forall a \in A(s)$, 已知时的 NE, 第二个是考虑到原问题中状态值都是未知情况如何利用第一个子问题的解来构建 $\pi^*(\cdot)$ 。

4.3.2 下一状态已知时计算 $\Phi(s)$ 的 NE

本节中假设对于所有的 $a \in A(s)$, $V^*(\varphi^{-1}(s, a))$ 的值都是已知的。标准型子博弈 $\Phi(s)$ 的 NE 意味着每个服务商的策略都是对对手策略的最优响应。受这个观察的启发, 已有的工作引入了最优响应迭代的算法来计算 NE [68,123,152], 该算法首先随机初始化 $\pi^*(s)$, 然后在每轮迭代中随机选择某个 $i \in [K]$ 并用 i 对 $\pi_{-i}^*(s)$ 的最优响应来更新 $\pi_i^*(s)$, 直到每个人的最优响应都等于自己当前的策略。这个算法被证明在势博弈和拥塞博弈中一定收敛 [152], 但在一般的博弈中没有理论保证。为了说明原因, 本节引入策略向量图 (Strategy Profile Graph, SPG) 的概念 [123], 图中的结点对应于一个纯策略向量 $a \in A(s)$, 如果 a 能通过某个玩家的最优响应行为变换到 $a' \in A(s)$, 那么图中就有一条从 a 指向 a' 的有向边。一旦最优响应算法陷入到 SPG 的一个没有出边的强连通子图, 或称 Sink Equilibrium, 算法将进入无限循环无法收敛, 如图4.1所示。为了解决这个问题,

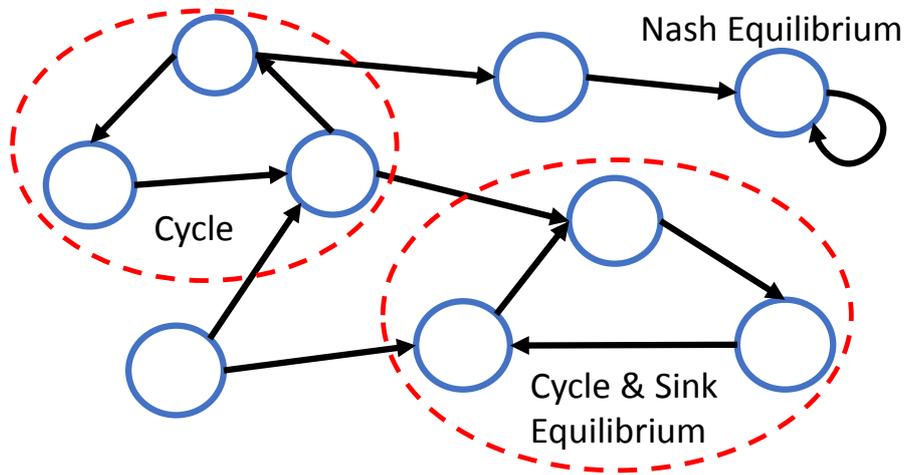


图 4.1 策略向量图

本节设计了算法4, 如果 NE 存在, 则该算法保证能收敛到 $\Phi(s)$ 的一个 NE。令 $C_s(a)$ 表示 SPG 中结点 $a \in A(s)$ 的子结点, 于是有当且仅当 $a \in C_s(a)$ 时 a 是一个 NE。算法4的基本思想是从一个结点开始, 沿着最优响应边搜索 NE, 过程中删除每个被访问结点的入边来避免陷入环路。

命题 4.1 子程序 $findNE(a)$ 一定会终止。

证明 由于 a 被从所有 $a' \in A(s)$ 的 $C_s(a')$ 中删除了, 这等价于删除结点 a 的所有入边, 当 $findNE(\cdot)$ 被调用时, 图中的每个顶点只会在 $findNE(\cdot)$ 的递归中被处理最多一次, 即, $findNE(\cdot)$ 不会陷入无限递归。由于 $A(s)$ 是有限的, $findNE(a)$ 一定会在 NE 已经被找到或者 $C_s(a)$ 为空时终止。 \square

定理 4.1 如果子博弈 $\Phi(s)$ 有 NE, 算法4一定会返回一个 NE 策略 $\pi^*(s)$ 。

Algorithm 4: 计算 $\Phi(s)$ 的 NE

```

// 主程序
1 初始化所有  $a \in A(s)$  的  $A(s)$  和  $C_s(a)$ ;
2 while True do
3   随机选择一个初始纯策略向量  $a \in A(s)$ ;
4   if ( $result \leftarrow findNE(a) \neq NONE$ ) then
5     return  $result$ ;

// 子程序  $findNE(a)$ 
6 if  $a \in C_s(a)$  then
7   return  $a$ ;
8 从  $C_s(a')$ ,  $\forall a' \in A(s)$ , 和  $A(s)$  中删除  $a$ ;
9 while  $C_s(a)$  非空 do
10  随机从  $C_s(a)$  中选择一个  $a'$ ;
11  if ( $result \leftarrow findNE(a') \neq NONE$ ) then
12  返回  $result$ ;
13 return NONE;

```

证明 令 a_0 表示算法使用的第一个初始结点并考虑以下两种情况。

情况 1: 存在一条连接 a_0 和一个 NE 的路径。如果 a_0 是一个 NE (即 $C_s(a_0) = \{a_0\}$), 那么 $findNE$ 直接终止并返回 a_0 。否则, $findNE(a_0)$ 会删除所有指向 a_0 的边。注意这个操作不影响 a_0 和 NE 的连接性。接下来, $findNE(a_0)$ 随机选择 a_0 的一个子结点 a' 并调用 $findNE(a')$ 。基于 a' 是否连接到一个 NE, 有两种场景需要考虑。

第一, 如果 a' 没有连接到一个 NE, 那么根据命题 4.1, $findNE(a')$ 会返回 *NONE*。此外, $findNE(a')$ 只能删除 a' 的入边以及 a' 能跳转到的结点。由于 a' 没有连接到一个 NE, 它能跳转到的结点也一定没有连接到一个 NE。因此, $findNE(a')$ 删除的边没有改变 a_0 和 NE 的连接性。在 $findNE(a')$ 返回 *NONE* 之后, $findNE(a_0)$ 会从 $C_s(a_0)$ 中随机选择另一个子结点并调用 $findNE$ 。值得注意的是, 根据情况 1 的假设, 至少有一个 a_0 的子结点会连接到一个 NE, 这类子结点一定会被 $findNE(a')$ 选择到, 从而触发以下的第二种场景。

第二, 如果 a' 连接着一个 NE, $findNE(a')$ 删除 a' 的入边没有破坏 a' 与 NE 的连接性。此时可以将 a' 视为一个新的初始点。根据上一段的分析, a' 的一个连接 NE 的子结点, 假设为 a'' , 会被 $findNE(a')$ 选择到, 然后 $findNE(a'')$ 被调用。由于递归函数 $findNE$ 一定会终止 (命题 4.1), 一个连接 NE 的结点 a^* 最终会被选择到并由 $findNE(a_0)$

返回。

情况 2: a_0 没有连接到任何 NE。当 $findNE(a_0)$ 被调用时, a_0 可以跳转到的结点将会被从 $A(s)$ 中删除。由于 NE 存在, 至少有一个结点 $a_1 \in A(s)$ 连接到 NE 或者 a_1 本身就是 NE。因为 a_0 没有连接到 NE, a_0 能跳转到的结点也不会连接到。由此, $findNE(a_0)$ 删除这些结点的入边不会改变 a_1 和 NE 的连接性。每次尝试一个新的初始点, 算法 4 要么返回一个 NE, 要么从 $A(s)$ 中删除一些结点。由于 $|A(s)|$ 是有限的, 算法 4 一定会返回一个 NE。 \square

考虑到子博弈 $\Phi(s)$ 中 NE 可能不存在, 本节引入 ξ -approximate NE (ANE) [68], 它被定义为一个策略向量, 满足玩家通过单独偏离它能额外获得的收益不会超过 ξ 的条件。ANE 总是存在的, 因为 ξ 可以设置为无限大。这样便可以将算法 4 扩展到计算 ξ -ANE。具体来说, 首先要定义玩家的 ξ -最优响应为一个纯策略集合, 采用里面的策略带来的收益比采取最优响应得到的收益少不超过 ξ , 然后通过用 ξ -最优响应边替换 SPG 中原有的边来构建一个新的图并将算法 4 用到新的图中。定理 4.1 意味着算法 4 会返回 $\Phi(s)$ 的一个 ξ -ANE。本章的实验结果表明在绝大部分设定下 ξ 都非常小。

4.3.3 计算整个随机博弈的 MPE

首先定义市场进入稳态的时间为

$$T = \max\{T_N, T_c\}. \quad (4-24)$$

从第 4.3.2 节可以得出, 为了用算法 4 计算 $\pi^*(s)$ (其中 $s = (p_{t-1}, t)$, $t < T$), 需要知道所有 $a \in A(s)$ 的 $V^*(a, t+1)$, \rightarrow 为了计算所有 $a \in A(s)$ 的 $V^*(a, t+1)$, 需要知道所有 $a' \in A(s)$ 的 $V^*(a', t+2)$ 和 $\pi^*(a', t+2)$, $\rightarrow \dots, \rightarrow$ 需要知道所有 $a'' \in A(s)$ 的 $V^*(a'', T)$ 和 $\pi^*(a'', T)$ 。此外还有如下定理:

定理 4.2 $r_i(a, t, a') = r_i(a, T, a')$, $\forall i \in [K]$, $a \in A$, $a' \in A$ 和 $t \geq T$ 。

证明 考虑两个分别以 (a, t) 和 (a, T) 为初始状态的原始随机博弈的子博弈, 其中 $t > T$ 。由于当 $t > T$ 时用户数量 $N(t)$ 和边际成本 $c_i(t, \cdot)$ 不在随着 t 改变, 下式成立:

$$r_i(a, t, a') = r_i(a, T, a'), \forall t > T. \quad (4-25)$$

也就是说, 可以忽略参数 t 并将两个子博弈看成是等价的。 \square

于是可以将 (a, t) 和 (a, T) 看成是一样的状态, 因而有

$$\pi^*(a, t) = \pi^*(a, T), \forall a \in A, t \geq T. \quad (4-26)$$

基于上述结果，计算完整的随机博弈的 MPE 的关键就在于计算所有 $a \in A$ 和 $t \geq T$ 的状态值 $V^*(a, t)$ 以及 MPE 策略 $\pi^*(a, t)$ 。本节将设计一个基于动态规划的算法来计算集合 $\{(a, t) | a \in A, t < T\}$ 中所有状态的 MPE 策略。

令 $s^0 = (a^{min}, t)$, $\forall t \geq T$, 其中 $a^{min} = (\min(A_1), \min(A_2), \dots, \min(A_K))$ 。从假设 4.1 可得 $A(s^0) = \{a^{min}\}$, 这意味着

$$\pi^*(s^0) = a^{min}. \quad (4-27)$$

由于长期收益 $V_i(s, \pi(s))$ 是未来无限时间的折现收益的总和，结合假设 4.1，有 $\forall i \in [K]$,

$$V_i^*(s^0) = \sum_{m=0}^{\infty} \gamma^m r_i(s^0, a^{min}) = \frac{r_i(s^0, a^{min})}{1 - \gamma}, \quad (4-28)$$

其中 $r_i(s^0, a^{min})$ 可以用公式 (4-11) 来计算。为了计算所有 $a \neq a^{min}$ 和 $t \geq T$ 的 $V^*(a, t)$ 和 $\pi^*(a, t)$, 引入符号 $A_{i,j}$ 来表示 A_i 中第 j 小的元素，然后定义函数 $h: A \rightarrow \mathbb{N}$ 来将一个纯策略向量映射到一个整数：

$$h(A_{1,j_1}, \dots, A_{K,j_K}) = (j_1 - 1) + \sum_{i=2}^K ((j_i - 1) \prod_{l=1}^{i-1} |A_l|). \quad (4-29)$$

也就是说， $h(\cdot)$ 建立了一个从有 $\prod_{i=1}^K |A_i|$ 个元素的空间 A 到集合 $\{0, 1, \dots, \prod_{i=1}^K |A_i| - 1\}$ 的一对一映射。例如，假设市场中有 $K = 3$ 个服务商，纯策略空间分别为 $A_1 = \{0.1, 0.2, 0.3\}$, $A_2 = \{0.25, 0.35\}$ 和 $A_3 = \{0.15, 0.4, 0.5\}$, 那么有 $h(0.1, 0.25, 0.15) = 0$ 和 $h(0.2, 0.35, 0.4) = 10$ 。用 $h^{-1}(n)$ 表示映射 h 中 n 的原像并用 s^n 来表示状态 $(h^{-1}(n), t)$, $n = 0, 1, \dots, \prod_{i=1}^K |A_i| - 1$, $t \geq T$ 。下面的命题可将状态的映射和假设 4.1 联系起来。

命题 4.2 对于所有的 $t \geq T$, $a' \in A(a, t)$ 意味着 $h(a') \leq h(a)$ 。

根据公式 (4-29) 可以很容易地证明该命题。命题 4.2 表明 $\pi^*(s^n)$ 仅与所有 $m \leq n$ 的状态值 $V^*(s^m)$ 有关。由于已经知道了 $V^*(s^0)$, 如果 $V^*(s^1)$ 也已知，则可以用算法 4 来计算 $\pi^*(s^1)$ 。然而， $V^*(s^1)$ 依赖于 $\pi^*(s^1)$, 但 $\pi^*(s^1)$ 是未知的。为了解决这个问题，首先需要验证 $h^{-1}(1)$ 是否是子博弈 $\Phi(s^1)$ 的一个 NE，如果是，那么 $\pi^*(s^1) = h^{-1}(1)$, 否则，从 $\Phi(s^1)$ 的 SPG 中删除 $h^{-1}(1)$ 并用算法 4 来计算 $\pi^*(s^1)$, 此时无需知道 $V^*(s^1)$ 的值。 $\pi^*(s^1)$ 求得之后可用公式 (4-10) 结合 $\pi^*(s^1)$ 和 $V^*(s^0)$ 来计算 $V^*(s^1)$ 。采用的同样的方法，可以依次基于 $V^*(s^0)$ 和 $V^*(s^1)$ 计算 $\pi^*(s^2)$, 基于 $V^*(s^0)$ 、 $V^*(s^1)$ 和 $V^*(s^2)$ 计算 $\pi^*(s^3)$, 等等，如算法 5 的第 1-8 行所示。验证 $h^{-1}(n)$ 是否是子博弈 $\Phi(s^n)$ 的一个 NE 需要以下定理：

定理 4.3 当且仅当 $\forall i \in [K]$, 服务商 i 无法通过单独偏离 $h^{-1}(n)$ 而获得比 $\frac{r_i(s^n, h^{-1}(n))}{1 - \gamma}$ 更高的收益时， $h^{-1}(n)$ 是子博弈 $\Phi(s^n)$ 的一个 NE。

证明 根据公式 (4-10), 如果 $h^{-1}(n)$ 是一个 NE, 那么

$$\begin{aligned} & V_i^*(s^n) \\ = & V_i(s^n, \pi^*(s^n)) \end{aligned} \quad (4-30)$$

$$= V_i(s^n, h^{-1}(n)) \quad (4-31)$$

$$= r_i(s^n, h^{-1}(n)) + \gamma V_i^*(s^n), \forall i \in [K]. \quad (4-32)$$

上式表明

$$V_i^*(s^n) = \frac{r_i(s^n, h^{-1}(n))}{1 - \gamma}. \quad (4-33)$$

结合 NE 的定义, 证明完成。 \square

Algorithm 5: 计算整个随机博弈的 MPE 策略

```

// t ≥ T
1 π*(s0) ← amin, Vi*(s0) ←  $\frac{r_i(s^0, a^{\min})}{1-\gamma}$ , ∀i ∈ [K];
2 for n = 1, 2, ..., ∏i=1K |Ai| - 1 do
3   if h-1(n) 是子博弈 Φ(sn) 的 NE then
4     π*(sn) ← h-1(n);
5   else
6     从子博弈 Φ(sn) 的 SPG 中移除 h-1(n);
7     用算法4计算 π*(sn);
8     基于公式 (4-10) 计算 V*(sn);
// t < T
9 for t = T - 1, T - 2, ..., 1 do
10  for a ∈ A do
11    s ← (a, t);
12    用算法4计算 π*(s);
13    基于公式 (4-10) 计算 V*(s);
14 return MPE 策略 π*(·);
    
```

根据定理4.3, 算法需要遍历每个服务商在状态 s^n 时的可行集并检测是否存在某个服务商 i 能得到比 $\frac{r_i(s^n, h^{-1}(n))}{1-\gamma}$ 更高的收益, 这个操作可在 $\sum_{i=1}^K |A_i|$ 的线性时间内完成。在所有 $n \in \{0, 1, \dots, \prod_{i=1}^K |A_i| - 1\}$ 的状态值 $V^*(s^n)$ 得到后, 可以用动态规划来计算所有 $a \in A$ 和 $t < T$ 的 $\pi^*(a, t)$, 该过程总结在算法5的第 9-13 行。

4.3.4 即时收益计算

算法4和5都涉及到公式4-11中定义的即时收益 $r(\cdot)$ 。本节研究如何计算 $r(\cdot)$ ，难点主要集中在计算 $\mathbf{E}_i^{p_{t-1}}(\sum_{j \in N_i(t, p_t)} d_{j,t})$ 和 $\mathbf{E}_i^{p_{t-1}}((\sum_{j \in N_i(t, p_t)} d_{j,t})^{1-\beta})$ 。首先给出下面两个命题。

命题 4.3

$$\mathbf{E}_i^{p_{t-1}}\left(\sum_{j \in N_i(t, p_t)} d_{j,t}\right) = \frac{N(t)P_i(p_t)}{\lambda_t(p_{t-1})}, \quad (4-34)$$

其中 $P_i(p_t)$ 是用户在服务商的价格向量为 p_t 时选择 i 的服务的概率，它等于

$$\int_0^\infty [(1 - U(p_{i,t} - x)) \prod_{i' \neq i} G_{i'}(x - p_{i,t} + p_{i',t})] g_i(x) dx. \quad (4-35)$$

证明 $\mathbf{E}_i^{p_{t-1}}(\sum_{j \in N_i(t, p_t)} d_{j,t})$ 为阶段 t 对服务商 i 的期望总需求，它等于每个用户期望的需求量和选择服务商 i 的期望用户数的乘积。因为用户的需求是独立同分布的，于是有 $\forall j' \in [N(t)]$,

$$\begin{aligned} & \mathbf{E}_i^{p_{t-1}}\left(\sum_{j \in N_i(t, p_t)} d_{j,t}\right) \\ &= \mathbf{E}_i^{p_{t-1}}(d_{j,t}) \mathbf{E}_i^{p_{t-1}}(N_i(t, p_t)) \end{aligned} \quad (4-36)$$

$$= \frac{1}{\lambda_t(p_{t-1})} \mathbf{E}_i^{p_{t-1}}(N_i(t, p_t)). \quad (4-37)$$

因为用户的边际收益和偏好分别是独立同分布的，所以

$$\mathbf{E}_i^{p_{t-1}}(N_i(t, p_t)) = N(t)P_i(p_t). \quad (4-38)$$

给定 p_t 和 $\phi_{j,i}$ ，用户 j 选择服务商 i 的概率为

$$P_i(p_t, \phi_{j,i}) = (1 - U(p_{i,t} - \phi_{j,i})) \prod_{i' \neq i} G_{i'}(\phi_{j,i} - p_{i,t} + p_{i',t}) \quad (4-39)$$

对 $\phi_{j,i}$ 取期望有

$$P_i(p_t) = \int_0^\infty g_i(x) P_i(p_t, x) dx. \quad (4-40)$$

□

命题 4.4

$$\begin{aligned} & \mathbf{E}_i^{p_{t-1}}\left(\left(\sum_{j \in N_i(t, p_t)} d_{j,t}\right)^{1-\beta}\right) \\ &= \int_0^{+\infty} z^{-\beta} e^{-\lambda_t(p_{t-1})z} dz \cdot \sum_{m=1}^{N(t)} (b(m; N(t), P_i(p_t)) \frac{\prod_{k=1}^m (k - \beta)}{(m - 1)!}), \end{aligned} \quad (4-41)$$

其中 $b(\cdot)$ 表示二项分布。

证明 首先有

$$\begin{aligned} & \mathbf{E}_i^{p_{t-1}} \left(\left(\sum_{j \in N_i(t, p_t)} d_{j,t} \right)^{1-\beta} \right) \\ = & \sum_{m=1}^{N(t)} P(|N_i(t, p_t)| = m) \cdot \mathbf{E}_i^{p_{t-1}} \left(\left(\sum_{j \in N_i(t, p_t)} d_{j,t} \right)^{1-\beta} \middle| |N_i(t, p_t)| = m \right), \end{aligned} \quad (4-42)$$

其中 $P(|N_i(t, p_t)| = m)$ 是在阶段 t 有 m 个用户选择服务商 i 的概率，它服从二项分布 $b(m; N(t), P_i(p_t))$ 。令 \mathcal{N}_m 表示所有包含 m 个不同用户的集合构成的集合，令

$$z_{N_i(t, p_t)} = \sum_{j \in N_i(t, p_t)} d_{j,t}, \quad (4-43)$$

然后有

$$\begin{aligned} & \mathbf{E}_i^{p_{t-1}} \left(\left(\sum_{j \in N_i(t, p_t)} d_{j,t} \right)^{1-\beta} \middle| |N_i(t, p_t)| = m \right) \\ = & \mathbf{E}_{N_i(t, p_t) \in \mathcal{N}_m} \left(\mathbf{E}_{z_{N_i(t, p_t)}} \left((z_{N_i(t, p_t)})^{1-\beta} \right) \right) \end{aligned} \quad (4-44)$$

$$= \mathbf{E}_{z_{N_i(t, p_t)}} \left((z_{N_i(t, p_t)})^{1-\beta} \right), \forall N_i(t, p_t) \in \mathcal{N}_m, \quad (4-45)$$

其中第二个等式由用户需求是独立同分布的假设得来。令 $\omega_m(z|\lambda_t(p_{t-1}))$, $z \in (0, +\infty)$, 表示 $N_i(t, p_t) \in \mathcal{N}_m$ 时 $z_{N_i(t, p_t)}$ 的概率密度函数。因为 $d_{j,t}$, $\forall j \in [N(t)]$, 服从同样的概率密度函数 $f(\cdot)$, 所以有 $z_{N_i(t, p_t)}$, $\forall N_i(t, p_t) \in \mathcal{N}_2$, 的概率密度函数满足下式:

$$\begin{aligned} & \omega_2(z|\lambda_t(p_{t-1})) \\ = & \int_0^z f(z-x|\lambda_t(p_{t-1}))f(x|\lambda_t(p_{t-1}))dx \end{aligned} \quad (4-46)$$

$$= \int_0^z \lambda_t(p_{t-1})e^{-\lambda_t(p_{t-1})(z-x)} \lambda_t(p_{t-1})e^{-\lambda_t(p_{t-1})x} dx \quad (4-47)$$

$$= \lambda_t(p_{t-1})^2 z e^{-\lambda_t(p_{t-1})z}. \quad (4-48)$$

类似地，可以基于 $\omega_2(\cdot)$ 得到 $z_{N_i(t, p_t)}$, $\forall N_i(t, p_t) \in \mathcal{N}_3$, 的概率密度函数:

$$\begin{aligned} & \omega_3(z|\lambda_t(p_{t-1})) \\ = & \int_0^z \omega_2(z-x|\lambda_t(p_{t-1}))f(x|\lambda_t(p_{t-1}))dx \end{aligned} \quad (4-49)$$

$$= \frac{\lambda_t(p_{t-1})^3}{2} z^2 e^{-\lambda_t(p_{t-1})z}. \quad (4-50)$$

通过一步步归纳推导可以得出 $\omega_m(z|\lambda_t(p_{t-1}))$ 的通用概率密度函数，即

$$\omega_m(z|\lambda_t(p_{t-1})) = \frac{\lambda_t(p_{t-1})^m}{(m-1)!} z^{m-1} e^{-\lambda_t(p_{t-1})z}, \quad (4-51)$$

其中 $z \in (0, +\infty)$ 。然后公式 (4-44) 可以显式地重写为

$$\mathbf{E}_{z_{N_i(t, p_t)}}((z_{N_i(t, p_t)})^{1-\beta}) = \int_0^{+\infty} z^{1-\beta} \omega_m(z | \lambda_t(p_{t-1})) dz. \quad (4-52)$$

于是有

$$\mathbf{E}_i^{p_{t-1}}\left(\left(\sum_{j \in N_i(t, p_t)} d_{j,t}\right)^{1-\beta}\right) = \sum_{m=1}^{N(t)} \left(\frac{b(m; N(t), P_i(p_t)) \lambda_t(p_{t-1})^m}{(m-1)!}\right) \int_0^{+\infty} z^{m-\beta} e^{-\lambda_t(p_{t-1})z} dz, \quad (4-53)$$

通过分部积分可以得到下面的公式：

$$\int_0^{+\infty} z^{m-\beta} e^{-\lambda_t(p_{t-1})z} dz = -\frac{e^{-\lambda_t(p_{t-1})z}}{\lambda_t(p_{t-1})} z^{m-\beta} \Big|_0^{+\infty} + \frac{m-\beta}{\lambda_t(p_{t-1})} \int_0^{+\infty} z^{m-\beta-1} \cdot e^{-\lambda_t(p_{t-1})z} dz. \quad (4-54)$$

由于 $m-\beta > 0$ ，有

$$-\frac{e^{-\lambda_t(p_{t-1})z}}{\lambda_t(p_{t-1})} z^{m-\beta} \Big|_{z=0}^{+\infty} = 0. \quad (4-55)$$

很容易验证

$$\lim_{z \rightarrow +\infty} -\frac{e^{-\lambda_t(p_{t-1})z}}{\lambda_t(p_{t-1})} z^{m-\beta} = 0. \quad (4-56)$$

于是有

$$\int_0^{+\infty} z^{m-\beta} e^{-\lambda_t(p_{t-1})z} dz = \frac{m-\beta}{\lambda_t(p_{t-1})} \int_0^{+\infty} z^{m-\beta-1} e^{-\lambda_t(p_{t-1})z} dz. \quad (4-57)$$

通过不停迭代调用上面的推导可以得到

$$\int_0^{+\infty} z^{m-\beta} e^{-\lambda_t(p_{t-1})z} dz = \frac{\prod_{k=1}^m (k-\beta)}{\lambda_t(p_{t-1})^m} \int_0^{+\infty} z^{-\beta} e^{-\lambda_t(p_{t-1})z} dz, \quad (4-58)$$

其中 $\int_0^{+\infty} z^{-\beta} e^{-\lambda_t(p_{t-1})z} dz$ 不依赖于 m ，可用数值计算的方法来求得。 \square

此时可以将公式 (4-11) 中定义的期望即时收益显式地写为

$$\begin{aligned} & r_i(p_{t-1}, t, p_t) \\ = & p_{i,t} \frac{N(t)P_i(p_t)}{\lambda_t(p_{t-1})} - c_{i,0}\delta(t) \int_0^{+\infty} z^{-\beta} e^{-\lambda_t(p_{t-1})z} dz \\ & \cdot \sum_{m=1}^{N(t)} (b(m; N(t), P_i(p_t)) \frac{\prod_{k=1}^m (k-\beta)}{(m-1)!}). \end{aligned} \quad (4-59)$$

4.3.5 近似算法

从第4.3.1节可以看出在算法4和5的每一轮循环中都要计算所有 $i \in [K]$ 的即时收益 $r_i(\cdot)$ ，其他的变量都依赖于 $r_i(\cdot)$ 的值，因此计算 $r_i(\cdot)$ 的效率会极大地影响算法的性能。本节讨论如何高效地求得一个近似 $r_i(\cdot)$ 。

根据公式 (4-59), 即时收益函数涉及计算求和公式

$$\Upsilon = \sum_{m=1}^{N(t)} (b(m; N(t), P_i(p_t)) \frac{\prod_{k=1}^m (k - \beta)}{(m - 1)!}), \quad (4-60)$$

其复杂度为 $O(N(t)^2)$ 。当 $N(t)$ 很大时 (例如 $N(t) = 10^6$), 算法5的第 9-13 行尤为耗时, 因为它需要遍历集合 $\{N(t-1), \dots, N(1)\}$ 。本节提出一个基于定理4.4的方法近似计算 Υ , 从而使得算法更高效。

定理 4.4

$$\frac{\prod_{k=1}^{N(t)} (k - \beta)}{N(t)!} \leq \frac{\Upsilon}{N(t)P_i(p_t)} \leq (1 - \beta). \quad (4-61)$$

证明 很容易验证

$$\frac{1}{1 - \beta} \leq \frac{m!}{\prod_{k=1}^m (k - \beta)} \leq \frac{N(t)!}{\prod_{k=1}^{N(t)} (k - \beta)}, \quad (4-62)$$

其中 $m \in [N(t)]$ 。给 Υ 的第 m 个元素, $m = 1, 2, \dots, N(t)$, 分别乘以 $\frac{1}{1 - \beta}$, $\frac{m!}{\prod_{k=1}^m (k - \beta)}$ 和 $\frac{N(t)!}{\prod_{k=1}^{N(t)} (k - \beta)}$ 可以得到下面的公式:

$$\frac{1}{1 - \beta} \Upsilon \leq \sum_{m=1}^{N(t)} m \cdot b(m; N(t), P_i(p_t)) \leq \frac{N(t)!}{\prod_{k=1}^{N(t)} (k - \beta)} \Upsilon, \quad (4-63)$$

其中中间的公式等于二项分布 $b(m; N(t), P_i(p_t))$ 的期望, 等于 $N(t)P_i(p_t)$ 。于是上面的不等式可以重写为

$$\frac{\prod_{k=1}^{N(t)} (k - \beta)}{N(t)!} \leq \frac{\Upsilon}{N(t)P_i(p_t)} \leq (1 - \beta). \quad (4-64)$$

□

由于 β 通常很小, $\frac{\prod_{k=1}^{N(t)} (k - \beta)}{N(t)!}$ 和 $(1 - \beta)$ 会很接近于 1。因此可以用 $\alpha N(t)P_i(p_t)$, $\alpha \in [\frac{\prod_{k=1}^{N(t)} (k - \beta)}{N(t)!}, 1 - \beta]$, 来近似 Υ 的值。值得注意的是 $m \cdot b(m; N(t), P_i(p_t))$ 在 $m = \lfloor N(t)P_i(p_t) \rfloor$ 时达到最大值, 且当 $N(t)$ 很大时, 根据二项分布的性质, $m \cdot b(m; N(t), P_i(p_t))$ 的值会随着 m 从 $\lfloor N(t)P_i(p_t) \rfloor$ 减少 (增加) 到 1 ($N(t)$) 而迅速减小到 0。因此可用

$$\widehat{\Upsilon} = \frac{\prod_{k=1}^{\lfloor N(t)P_i(p_t) \rfloor} (k - \beta)}{\lfloor N(t)P_i(p_t) \rfloor!} N(t)P_i(p_t) \quad (4-65)$$

来更加精确地近似 Υ , 上式的计算复杂度为 $O(N(t)P_i(p_t))$ 。通过将算法4和5中的即时收益函数 $r_i(\cdot)$ 中的 Υ 替换为 $\widehat{\Upsilon}$, 可以得到相应的近似算法。

4.4 实验评估

本节做大量实验来评估所提的方法，首先测试算法的效率，然后阐释服务商在不同参数设定下如何制定基于长远利益的价格策略，最后比较 MPE 和其他已有的策略。本章的算法由 Python 2.7.13 实现，所有实验运行在一台装备了 64 位 windows 系统、拥有 64GB 内存和 3.4GHz 四核处理器的计算机上。

4.4.1 算法评估

本节首先评估近似公式 $\widehat{\Upsilon}$ 的性能，包括近似程度和计算效率。定义近似比为

$$\frac{\widehat{\Upsilon}}{\Upsilon}, \quad (4-66)$$

该比值衡量着近似公式 $\widehat{\Upsilon}$ 与 Υ 的接近程度。实验中部分参数设定如下：

- $K = 3$,
- $N_0 = 1000$,
- $N_\infty = 50000$,
- $\kappa = 0.15$,
- $\beta = 0.01$,

根据 N_0 、 N_∞ 和 κ 可得 $T_N = 99$ 。因此用户数量从阶段 1 到 15 以比较缓慢的速率增长，之后从阶段 15 到 50 进入快速增长期，再之后逐渐进入稳定期，如图 4.2 所示。

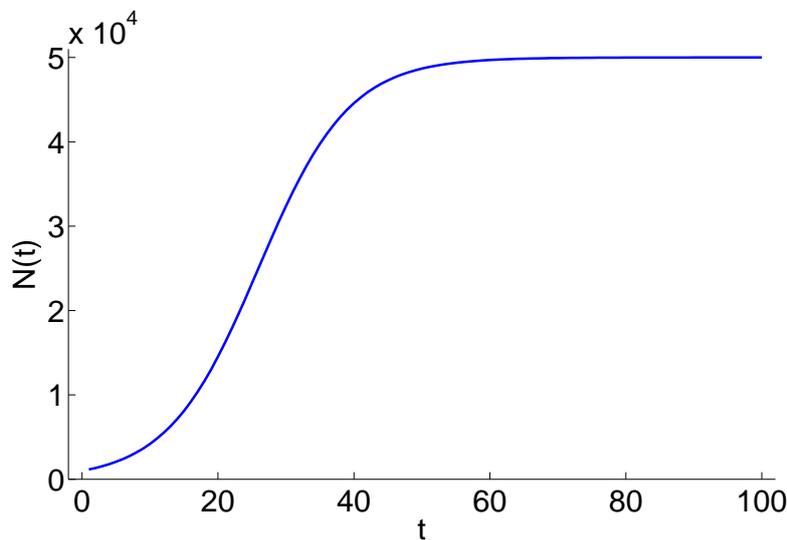


图 4.2 用户数量增长曲线

令 $T_c = 100$ 并令 $|A_i| = 10 + i$, $A_{i,j} = j/10$, $\forall i \in [K] j \in [|A_i|]$ 。其余参数由随机采样生成，取值空间如下所示：

- $\eta \in [0.01, 0.1]$,
- $\gamma \in [0.1, 1]$,

- $[U_l, U_u] = [1, 2]$,
- $\Lambda \in [1, 10]$, $b_i \in [0.1, 1]$,
- $q_i \in [0.1, 1]$ 。

通过随机生成参数运行 100 次实验得到平均近似比和运行时间分别显示在图4.3(a)和4.3(b)中。从图4.3(a)中可以看出近似比非常接近 1，这表明 $\widehat{\Upsilon}$ 与 Υ 非常接近。不同用户数量时计算 Υ 和 $\widehat{\Upsilon}$ 的时间以对数纵坐标展示在图4.3(b)中，可以看出当 t 很大 ($N(t)$ 很大) 时，计算 Υ 非常耗时。与之比较，计算 $\widehat{\Upsilon}$ 要高效得多。所以，近似公式在满足高精度的同时计算效率也非常高。后文对算法的性能评估都基于该近似公式。

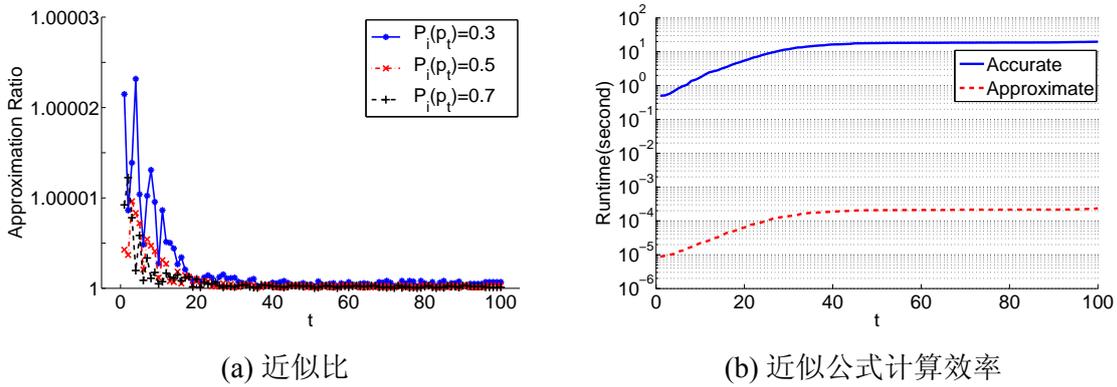


图 4.3 近似公式评估

接下来评估所提算法的收敛性和执行效率。实验中的前 5 个参数也用随机采样的方法生成，采样空间如下所示：

- $K \in \{2, 3, 4, 5\}$,
- $N_0 \in \{1000, 2000, \dots, 10000\}$,
- $N_\infty \in \{50000, 60000, \dots, 100000\}$,
- $\kappa \in [0.1, 1]$,
- $\beta \in [0.01, 0.05]$,

其他参数的取值空间与上文保持一致。为了评估算法4的收敛性，对于每一个 $K \in \{2, 3, 4, 5\}$ ，进行 1000 次实验，每次实验中的其他参数由上述基本设定随机生成。用算法5计算 MPE 时，首先设定 $\xi = 0$ 。如果算法4在某个时间点 $t \in \{1, 2, \dots, T\}$ 返回 NONE，则将 ξ 增大一个定值（本节的实验中取 0.1），直到算法能收敛到一个 ξ -ANE。在每个 K 的取值下， ξ 在 1000 次实验中的统计结果列在表4.2中。从表中可以观察到对于 $K = 2, 3$ ，算法4能收敛到每个子博弈的 NE，因此算法可以计算出整个随机博弈的精确的 MPE 策略。当 $K \geq 4$ 时，大部分实验中（超过 99%）算法都找到了精确的 NE 和 MPE。在本节的设定中，每个服务商的长期收益通常不会小于 100（见图4.6中的例子），因此 ξ 相较于收益而言是非常小的。也就是说，尽管个别实验中的随机博弈没有精确的 MPE，服务商却能达到一个距离 MPE 非常接近的近似 MPE。

表 4.2 算法4的收敛性

K	$\max \xi$	$\min \xi$	$\text{avg } \xi$	$\text{Prob}(\xi = 0)$
2	0	0	0	1
3	0	0	0	1
4	0.3	0	1e-3	0.993
5	0.2	0	7e-4	0.996

为了评估算法5的效率，令 $N_0 = 10000$ 且 $\kappa = 0.1$ ， T_N 选自 $\{100, 110, \dots, 150\}$ 。然后有

$$T = \max\{T_N, T_c\} = T_N \in \{100, 110, \dots, 150\}. \quad (4-67)$$

N_∞ 可以通过 N_0 、 κ 和 T_N 计算出来。本节假设服务商拥有相同的纯策略空间并研究算法的运行时间如何随着 K 、 $|A_i|$ 和 T 变化，因为全局的状态空间高度依赖于这三个变量。给定 K 、 $|A_i|$ 和 T ，评估中计算 100 次实验的平均运行时间，其他参数根据上述基本设定随机采样生成。实验结果展示在图4.4中，结果表明运行时间几乎是 T 的线性函数。此外，运行时间的增长率通常远远小于总纯策略空间大小的增长率。例如，当 K 从 4 增加到 5 且每个服务商 $i \in [K]$ 的纯策略空间都为 $|A_i| = 10$ 时，总的纯策略空间增长了 10 倍，但算法5的运行时间仅增加了大约 2 倍。由于现实世界中的市场通常有 3-4 个大公司，本章提出的算法可以有效地应用于真实的云计算服务领域。

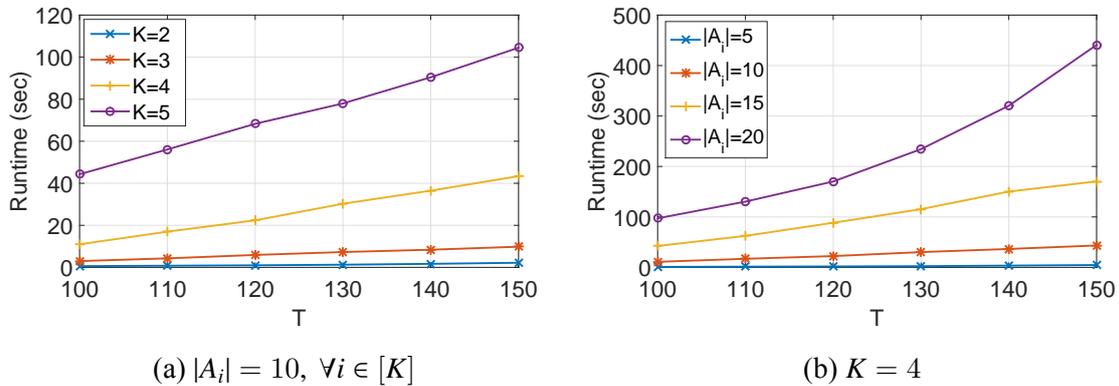


图 4.4 运行时间评估

4.4.2 MPE 策略评估

本节详细研究 MPE 策略，首先给出参数设置规则，然后展示服务商使用 MPE 策略的价格变化曲线。对于每一种参数设定，评估结果都是 100 次使用不同初始点（算法4的第 3 行）的实验的平均值。本节最后以实验论证服务商没有单独偏离 MPE 策略的动机，且 MPE 策略在优化长远利益方面优于其他已有策略。

4.4.2.1 参数设定

本节的实验中，令

- $N_0 = 1000$,
- $N_\infty = 50000$,
- $\beta = 0.01$,
- $\kappa = 0.15$,
- $\eta = 0.02$,
- $T = T_c = 100$,
- $\gamma = 0.8$,
- $\Lambda = 10$,
- $b_i = 1, \forall i \in [K]$,

于是有 q_i 越大则说明在服务商设定相同的价格时 i 的服务被用户选择的几率越大。在现实应用中，这些参数可由领域专家来设置或者通过历史数据来估测，例如，估测用户需求 [163] 和偏好 [164]。

算法的评估基于表4.3中的四种场景。令前两个场景中 $K = 4$ ，后两种场景中 $K = 3$ ，因为现实中大约有 3-4 个大的云服务商。在所有场景中，令 $|A_i| = 10 + i$ 且 $A_{i,j} = j/10$ ， $\forall i \in [K]$ ， $j \in [|A_i|]$ 。假设服务商 i 的初始价格为 $p_{i,0} = \max_j A_{i,j}$ ， $\forall i \in [K]$ 。在场景 1 和 3 中，初始边际成本 ($c_{i,0}$) 高的服务商的用户偏好 (q_i) 也高，在场景 2 和 4 中的设定刚好相反。

表 4.3 评估用场景

Scenario	K	$c_{i,0}$	q_i
1	4	(0.45, 0.5, 0.55, 0.6)	(0.25, 0.3, 0.35, 0.4)
2	4	(0.6, 0.55, 0.5, 0.45)	(0.25, 0.3, 0.35, 0.4)
3	3	(0.5, 0.55, 0.6)	(0.3, 0.35, 0.4)
4	3	(0.6, 0.55, 0.5)	(0.3, 0.35, 0.4)

4.4.2.2 MPE 策略

本节展示云计算服务商在不同的场景下如何按照 MPE 策略来调整价格，见图4.5。从图中可以看出在所有场景中，市场饱和之前的价格竞争非常激烈。由于用户数量的增长和边际成本的降低，所有服务商趋向于逐渐降低自己的价格。在场景 1 和 3 中，边际成本较高的服务商通常会设置较高的价格，主要原因有两个：1) 用户对他们的偏好较高，因此即使价格比用户偏好低的服务商稍微高一点，他们依然能吸引到用户；2) 由

于边际成本高，他们无法设置过低的价格，否则会得到负的收益，而用户偏好和边际成本都低的服务商则倾向于用价格战争来获取更多利益。在场景 2 和 4 中，边际成本低的服务商在最初的阶段首先降价，边际成本高的服务商不得不动跟着降价，随着时间的推移，所有服务商的边际成本都因为时间因子 $\delta(t)$ 逐渐变得很小而降得很低，这时用户偏好低的服务商主动降价来获得更多用户的动机变大。在所有场景中，服务商的价格最终收敛到一个稳定状态。

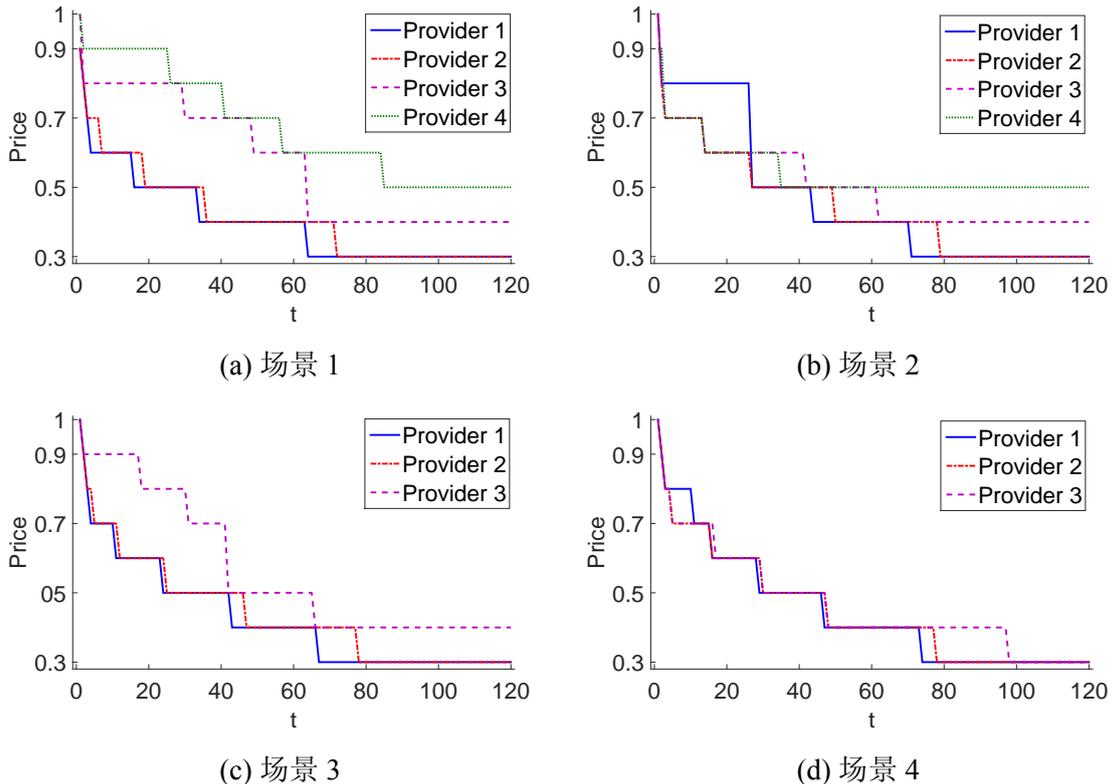


图 4.5 MPE 策略

4.4.2.3 长期收益比较

为了评估 MPE 策略的优势，本节将其与两个已有的最好的基于长远利益的定价策略 [11] 进行比较。具体而言，使用第一个定价策略 (P1) 的服务商在最大化自己长期收益的同时假设其他服务商会跟随自己的绝对价格变动来调整价格 (例如自己降价 10 元，则假设对手的反应为也降价 10 元)，使用第二个定价策略 (P2) 的服务商假设对手会仿照自己的相对价格变动来调整价格 (例如自己降价 10%，则假设对手也会降价 10%)。

对于每个场景，本节首先计算服务商全程采用 MPE 策略的长期收益，然后检查服务商采取 P1 或 P2 能否获得更高的收益。比较的结果展示在图 4.6 中，其中与服务商 i 关联的三个柱状图分别对应以下三种情况下的长期收益：1) 所有服务商都采用 MPE 策略；2) i 单独使用策略 P1；3) i 单独采用 P2。可以看出在所有场景中，服务商在全部采用 MPE 策略时得到的长期收益最高。另一个发现是用户偏好越高的服务商通常得到

的长期收益也越高，无论他采用什么策略，因为他能在服务商价格相近时吸引更多的用户。

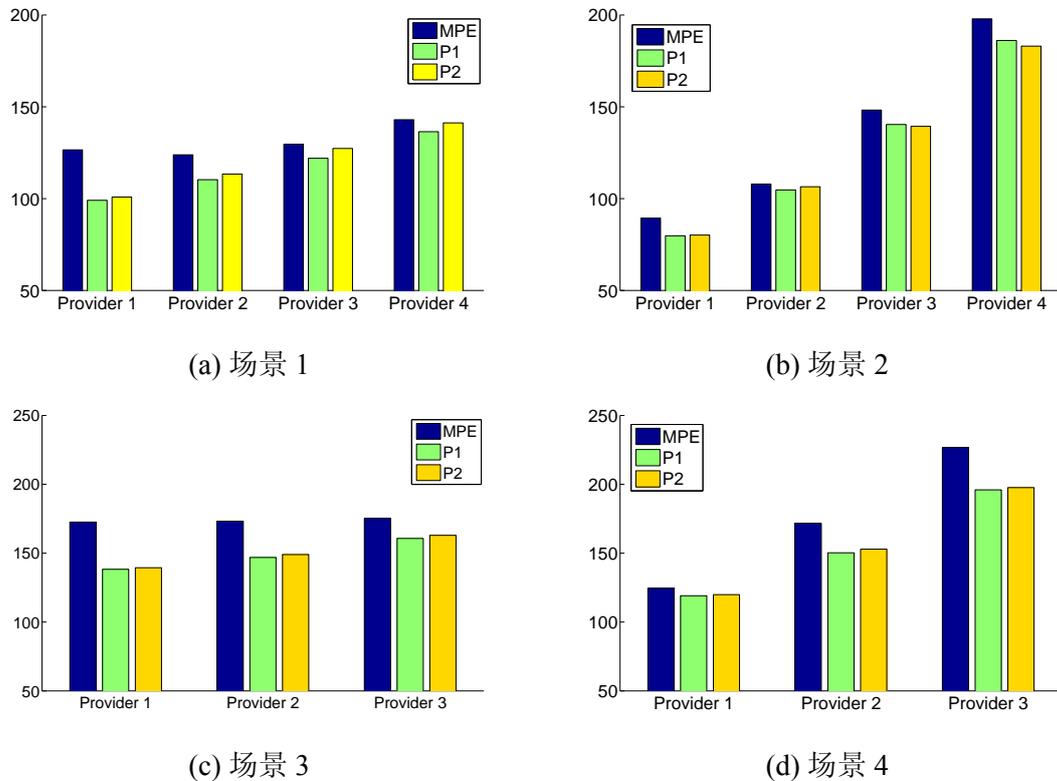


图 4.6 长期收益比较

表 4.4 服务商策略不同时的收益比较

i	Strategy	$c_{i,0}$	q_i	A_i	Profit
1	MPE	0.5	0.3	{0.1, 0.2, ..., 1}	162
2	P1	0.5	0.3	{0.1, 0.2, ..., 1}	151
3	P2	0.5	0.3	{0.1, 0.2, ..., 1}	147
1	MPE	0.6	0.3	{0.1, 0.2, ..., 1}	112
2	P1	0.6	0.3	{0.1, 0.2, ..., 1}	98
3	P2	0.6	0.3	{0.1, 0.2, ..., 1}	101
1	MPE	0.5	0.4	{0.1, 0.2, ..., 1}	216
2	P1	0.5	0.4	{0.1, 0.2, ..., 1}	207
3	P2	0.5	0.4	{0.1, 0.2, ..., 1}	201

接下来研究当服务商分别采用完全不同的策略时 MPE 策略是否还是最优的。令 $K = 3$ 并令服务商拥有相同的 $c_{i,0}$ 、 q_i 和 A_i ，如表 4.4 所示，其他参数设定同第 4.4.2.1 节。令服务商 1 使用 MPE 策略，服务商 2 采用 P1 策略，服务商 3 利用 P2 策略，然后计算他们的长期收益。实验结果表明尽管服务商 1 的收益没有全部服务商都采用 MPE 策略

时的高，但它仍高于其他两个策略的收益，也就是说，即使服务商采用完全不同的策略，MPE 策略仍优于其他两个策略。

4.5 总结

本章研究了云计算市场中基于长远利益的定价优化问题，主要做了以下两个贡献。第一，本章提出了一个更现实的市场模型，考虑了市场的发展、竞争和不完全信息，此外还将市场模型规划成了一个随机博弈。第二，本章设计了一个基于动态规划的高效算法来计算 MPE 策略。本章做了大量而充分的实验，深入讨论了服务商如何在不断发展的市场中制定价格，结果表明与已有的策略相比，MPE 策略能给服务商带来更高的长远利益。

第五章 近似均衡和动态供需优化

互联网的普及使得资源可重复利用的互联网服务获得了长足的发展，如网上租车、预订酒店和云计算服务，考虑到激烈的市场竞争和对价格敏感的用户，服务商必须制定合理的价格策略优化自己的收益。上一章提出了一个基于长远利益的价格策略，其假设在每个阶段服务商保持价格不变。本章弱化这一假设，充分考虑设备利用率的动态特性，研究在一个特定阶段服务商基于设备利用率的动态定价策略，主要有三个贡献。第一，本章提出了一个模型来模拟在服务商价格的相互影响下的用户需求变化。第二，本章提出一个近似均衡的概念来描述服务商的行为，该均衡比纳什均衡的实用性更强。第三，本章在没有对用户需求函数做任何限制的情况下推导了该模型的很多性质，可以利用这些性质来降低算法的搜索空间。最后，本章基于不同的参数设定做了大量实验，结果表明本章提出的近似均衡策略很接近纳什均衡，且与已有的策略相比能给服务商带来更高的收益。

5.1 引言

在现实世界的许多应用中，服务商提供的资源（或商品）是可重复利用的，本章以云计算服务为具体研究对象。平台中的设备的利用率会动态变化，因为用户可以随时开始或结束使用一个设备。面对价格敏感的用户，动态定价对于创造更多利润有着重要作用，其已在工业界取得了一些成功，例如租车 [88]，酒店预订 [89,165]，网络连接服务 [92] 以及云计算 [83,84] 等，并引发了广泛的学术关注 [11,90,94,166,167]。然而，现有的工作却未能同时考虑动态供需和市场竞争，因此已有的价格策略可能会在现实应用中得到很差的收益。基于这个背景，本章研究如何利用价格来优化供需关系，从而最大化处于市场竞争中的服务商的期望收益。本章给出了许多理论和实验结果，总结起来有三大贡献。

第一，本章建立了一个可以描述动态供需和市场竞争的模型，其中服务商提供的资源是可重复使用的。目前还没有相关工作在考虑了动态特性的同时刻画了市场竞争。文献 [88,89] 中研究了一种需求预测模型，但使用最广泛的描述用户动态需求的模型是泊松过程 [84,90–92]。然而，这些工作没有考虑到市场竞争对用户需求的影响。Xu 和 Hopp [93] 假设用户的需求增加率服从一个几何布朗运动过程，并用一个完美贝叶斯均衡来模拟服务商的行为。Levin 等人 [94] 考虑了用户的策略并提出一个子博弈完美均衡的框架。然而，这些工作只研究了一次性的库存补充模型中的动态定价问题，这些方法并不能解决资源可重复利用的市场的定价问题。本章同时考虑动态的用户需求和市场竞争。本章采用文献中常用的假设，即需求满足泊松过程，基于此本章将动态和充满竞争的市场抽

象成一个连续时间马尔科夫过程 (Continuous-Time Markov Process) [168–170]。

由于每个服务商的目的都是最大化自己的期望收益，最优策略应该会形成一个均衡。本章阐明真实世界中的服务商的策略很难达到纳什均衡，因为服务商无法获取竞争对手的全部信息，且其自己的收益并非其价格策略的显函数。本章的第二个主要的贡献是提出一个近似均衡 (Approximate Equilibrium, AE) 的方案概念 [68,171] 并提供了一个基于最优响应的高效算法。

第三，本章推导出所提模型的很多性质。文献 [92] 分析了服务商不考虑用户使用服务时间的长短，只一次性收费这种设定。Xu 等人 [10] 研究了云计算市场动态定价的单调性，但他们的结果需要需求函数满足特定的条件，而本章的结果不需要对需求函数做任何特殊的假设。本章发现的性质包括期望全局收益的单调性和凹凸性，以及 AE 策略关于资源利用率的凹凸性。这些结果可被用于缩减计算 AE 策略的算法的搜索空间。

本章做了大量的实验来评估算法的性能，讨论了算法的收敛性。实验结果表明本章的动态定价策略比已有的策略更优，且 AE 非常接近于一个纳什均衡。

本章的剩余部分组织如下。在第 5.2 节，首先介绍用于模拟动态需求的马尔科夫过程，然后给出多个服务商的市场竞争模型。第 5.3 节给出最优动态定价的方案概念，并提出一个有效的算法来求解该策略。第 5.4 节推导模型的许多重要的性质，它们可以被用来提高算法效率。第 5.5 节通过实验来评估提出的方法。结论放在最后一节。

5.2 竞争性的市场和动态需求建模

5.2.1 研究动机

竞争是当今商业服务行业的一个重要特征，例如云计算市场，其中有众多公司通过互联网为用户提供相似的商品或服务 [11,167]。最常用的云计算服务有亚马逊的 AWS 和微软的 Azure。市场的另一个重要特征是用户的需求和服务商的库存是动态变化的，即，一个云服务平台上被用户占用的虚拟机器数量是不断变化的。由于用户通常是对价格敏感的，服务商需要以一定的策略来设定价格来影响用户对其资源的需求，从而更好地利用自己的资源。事实上，亚马逊的 EC2 引入了“spot pricing”策略来动态更新虚拟机的价格。本章研究服务商的最优动态价格策略，从而帮助其最大化自己的期望收益。接下来两节先给出本章的市场模型。

5.2.2 动态需求建模

沿用上一章的符号定义，用 K 表示服务商的数量，用 $[K]$ 表示集合 $\{1, 2, \dots, K\}$ 。按照文献中常用的做法 [10,90,92,93]，假设用户对服务商 $k \in [K]$ 的资源的需求由两个独立的泊松过程决定，即需求增加过程和需求减少过程。前者模拟新需求的到来，后者描述已有需求的结束。两个过程都不仅与服务商自己的价格 p_k 有关，考虑到市场竞争，还与其他服务商的价格 p_{-k} 有关。具体地，用 $\lambda_k(\cdot)$ 来表示服务商 k 的需求增加泊松过程

(单位时间内新增的需求数), 它满足下式 [172]:

$$\lambda_k(p) \geq 0; \quad (5-1)$$

$$\frac{\partial \lambda_k(p)}{\partial p_k} < 0; \quad (5-2)$$

$$\frac{\partial \lambda_k(p)}{\partial p_{k'}} > 0, \forall k \neq k', \quad (5-3)$$

其中, $p = (p_1, p_2, \dots, p_K)$ 。上式跟实际情况是一致的, 即 1) 需求增加率不可能是负的; 2) 降低 k 的价格能吸引更多用户使用 k 的服务; 3) 当其他服务商的价格增加时, 原本使用其他服务商资源的用户可能转而使用 k 的, 导致 k 的需求增加率提高。类似地, 泊松需求减少过程表示为 $\mu_k(\cdot)$, 其满足

$$\mu_k(p) \geq 0; \quad (5-4)$$

$$\frac{\partial \mu_k(p)}{\partial p_k} > 0; \quad (5-5)$$

$$\frac{\partial \mu_k(p)}{\partial p_{k'}} < 0, \forall k \neq k'. \quad (5-6)$$

用符号 $(p_k, p_{-k}) = p$ 并用 δ_k 来表示服务商 k 的离散价格策略^①。 δ_k 中最小的元素 p_k^{min} 和最大的元素 p_k^{max} 满足

$$\mu_k(p_k^{min}, p_{-k}) = 0, \forall p_{-k} \in \delta_{-k} \quad (5-7)$$

$$\lambda_k(p_k^{max}, p_{-k}) = 0, \forall p_{-k} \in \delta_{-k} \quad (5-8)$$

其中

$$\delta_{-k} = \times_{i \in C(k)} \delta_i, \quad (5-9)$$

$$C(k) = [K] \setminus \{k\}. \quad (5-10)$$

本章假设 $\lambda_k(p_k, p_{-k})$ 和 $\mu_k(p_k, p_{-k})$ 都是有界的, 因为现实世界中服务商不可能得到无限的需求增加和减少率, 二者被统称为需求函数。令 N_k 表示服务商 k 的最大资源数量, 令 $[N_k]^+$ 表示集合 $\{0, 1, \dots, N_k\}$ 。由于需求增加和减少都是随机过程, 用户占用的机器数量可以被表示成一个连续时间马尔科夫过程, 其中服务商 k 的状态 $n \in [N_k]^+$ 对应于其平台上已经被用户占用的机器数量。服务商 k 的动态价格策略定义为

$$P_k = (p_{k,0}, p_{k,1}, \dots, p_{k,N_k}), \quad (5-11)$$

其中 $p_{k,n}$ 是为状态 n 设定的价格。此外, 有 $P_k \in \Delta_k$, 其中

$$\Delta_k = \times_{n=0}^{N_k} \delta_k. \quad (5-12)$$

① 价格对应于现实世界中的金钱, 是一个离散变量。

定义

$$P_{-k} = \times_{i \in C(k)} P_i \quad (5-13)$$

为除了服务商 k 外的其他服务商的价格策略向量，此外定义

$$\Delta_{-k} = \times_{i \in C(k)} \Delta_i. \quad (5-14)$$

令 $P = (P_k, P_{-k})$, $\forall k \in [K]$ 。

5.2.3 多服务商模型

本节首先介绍单服务商模型。假设服务商 1 在市场中是唯一的，那么其马尔科夫过程的状态转移矩阵 [169,173] 可以被写作 $Q_1(P_1) = (q_{i,j}^1(P_1))_{i,j}$, $i, j \in [N_1]^+$:

$$q_{i,j}^1(P_1) = \begin{cases} \lambda_1(p_{1,i}), & \text{if } j = i + 1; \\ \mu_1(p_{1,i}), & \text{if } j = i - 1; \\ -\sum_{l \neq i} q_{i,l}^1(P_1), & \text{if } j = i; \\ 0, & \text{否则,} \end{cases} \quad (5-15)$$

其中 $q_{i,j}^1(P_1)$ 表示状态从 i 跳转到 j 的速率。状态 n 在该马尔科夫过程中出现的概率， $\varphi_{1,n}(P_1)$, $n \in [N_1]$ ，满足

$$\sum_{n \in [N_1]} \varphi_{1,n}(P_1) = 1; \quad (5-16)$$

$$\varphi_1(P_1) \cdot Q_1(P_1) = 0, \quad (5-17)$$

上式中 $\varphi_1(P_1) = (\varphi_{1,0}(P_1), \varphi_{1,1}(P_1), \dots, \varphi_{1,N_1}(P_1))$ 被称作稳态概率分布 (Steady-State Probability)。

当市场中有多个服务商的时候，他们的价格策略会影响彼此的需求增加和减少过程，如公式 (5-2)-(5-5) 所示，因此任意服务商 k 的稳态概率分布是 P (不仅是 P_k) 的函数。令

$$\varphi_k(P) = (\varphi_{k,0}(P), \varphi_{k,1}(P), \dots, \varphi_{k,N_k}(P)), \quad (5-18)$$

$$\varphi_{-k}(P) = \times_{i \in C(k)} \varphi_i(P), \quad (5-19)$$

$$\varphi(P) = (\varphi_k(P), \varphi_{-k}(P)). \quad (5-20)$$

由于 $p_{k,n}$ 是 k 在状态 n 时使用的价格， $\varphi_{k,n}(P)$ 也可以看成 k 的定价为 $p_{k,n}$ 的概率。此时

服务商 k 的状态转移矩阵写作 $Q_k(P) = (q_{i,j}^k(P))_{i,j}$, $i, j \in [N_k]^+$, 其满足

$$q_{i,j}^k(P) = \begin{cases} \mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}(P)} \{\lambda_k(p_{k,i}, p_{-k})\}, & \text{if } j = i + 1; \\ \mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}(P)} \{\mu_k(p_{k,i}, p_{-k})\}, & \text{if } j = i - 1; \\ -\sum_{l \neq i} q_{i,l}^k(P), & \text{if } j = i; \\ 0, & \text{否则,} \end{cases} \quad (5-21)$$

其中

$$\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}(P)} \{f(p_{-k})\} = \sum_{p_{-k} \in P_{-k}} f(p_{-k}) Pr(p_{-k} | \varphi_{-k}(P)), \quad (5-22)$$

$Pr(p_{-k} | \varphi_{-k}(P))$ 是给定 $\varphi_{-k}(P)$ 后价格向量 p_{-k} 出现的概率。类似于单服务商场景, 有

$$\sum_{n \in [N_k]^+} \varphi_{k,n}(P) = 1, \quad (5-23)$$

$$\varphi_k(P) \cdot Q_k(P) = 0. \quad (5-24)$$

当 n 个资源正被用户使用, 服务商 k 能在单位时间内得到 $n \cdot p_{k,n}$ 的收益。于是, 服务商 k 的期望收益率为

$$J_k(P_k, P_{-k}) = \sum_{n=0}^{N_k} \varphi_{k,n}(P) \cdot n \cdot p_{k,n}. \quad (5-25)$$

为了方便阅读, 此处将上节中定义的所有符号总结在表5.1中。

5.3 最优定价策略

由于每个服务商都在考虑其他竞争者的同时最大化自己的收益率, 本章需要研究均衡价格策略。首先, 本节阐明服务商在现实世界中很难达到纳什均衡。为了解决这个问题, 本节提出了一个 AE 的方案概念, 并设计了一个高效的算法来求解该均衡策略。

5.3.1 均衡策略

定义 1 (纳什均衡) 纳什均衡是一个价格策略向量 $P^* = \times_{k \in [K]} P_k^*$, 其满足 $\forall k \in [K]$,

$$J_k(P_k^*, P_{-k}^*) \geq J_k(P_k, P_{-k}^*), \forall P_k \in \Delta_k. \quad (5-26)$$

也就是说, 没有人可以通过单独改变自己的策略而获得更高的收益。然而, $J_k(P_k, P_{-k})$ 并不是关于 $P = (P_k, P_{-k})$ 的显函数。具体而言, 为了求解方程式 (5-23)-(5-24), 有

$$\varphi_{k,n}(P) \cdot q_{n,n+1}^k(P) = \varphi_{k,n+1}(P) \cdot q_{n+1,n}^k(P), \quad (5-27)$$

由此可以推导出 $\forall m, n \in [N_k]^+$ 和 $m > n$,

$$\varphi_{k,m}(P) = \varphi_{k,n}(P) \prod_{j=n}^{m-1} \frac{q_{j,j+1}^k(P)}{q_{j+1,j}^k(P)}. \quad (5-28)$$

表 5.1 动态供需符号总结

K	服务商的数量
$[K]$	集合 $\{1, 2, \dots, K\}$
$\lambda_k(\cdot)$	服务商 $k \in [K]$ 的需求增加函数
$\mu_k(\cdot)$	服务商 $k \in [K]$ 的需求减少函数
p_k	服务商 $k \in [K]$ 的价格
p_{-k}	除了服务商 $k \in [K]$ 其他人的价格向量
p	所有服务商的价格向量 $(p_1, p_2, \dots, p_K) = (p_k, p_{-k})$
δ_k	服务商 k 的价格空间
$C(k)$	集合 $[K] \setminus \{k\}$
δ_{-k}	价格空间 $\times_{i \in C(k)} \delta_i$
N_k	服务商 k 的资源总数
$[N_k]^+$	集合 $\{0, 1, \dots, N_k\}$
n	被使用的资源数
$p_{k,n}$	服务商 k 为状态 $n \in [N_k]^+$ 设置的价格
P_k	服务商 k 的价格策略
P_{-k}	除了服务商 k 其他人的价格策略向量 $\times_{i \in C(k)} P_i$
P	所有服务商的价格策略向量 (P_k, P_{-k})
Δ_k	空间 $\times_{n=0}^{N_k} \delta_k$
Δ_{-k}	空间 $\times_{i \in C(k)} \Delta_i$
$Q_k(\cdot)$	服务商 k 的状态转移矩阵
$\varphi_k(\cdot)$	服务商 k 的马尔科夫过程的稳态概率分布
$\varphi_{-k}(\cdot)$	除了服务商 k 其他人的稳态概率分布向量
$\varphi(\cdot)$	所有服务商的稳态概率分布向量 $(\varphi_k(\cdot), \varphi_{-k}(\cdot))$
$J_k(\cdot)$	服务商 k 的期望收益率

通过进一步推导可得到

$$\varphi_{k,n}(P) \left(\sum_{m=0}^{n-1} \frac{\varphi_{k,m}(P)}{\varphi_{k,n}(P)} + 1 + \sum_{m=n+1}^{N_k} \frac{\varphi_{k,m}(P)}{\varphi_{k,n}(P)} \right) = 1, \quad \forall n \in [N_k]^+, \quad (5-29)$$

因此有

$$\varphi_{k,n}(P) = \frac{1}{\left(\sum_{m=0}^{n-1} \prod_{j=m}^{n-1} \frac{q_{j+1,j}^k(P)}{q_{j,j+1}^k(P)} + 1 + \sum_{m=n+1}^{N_k} \prod_{j=n}^{m-1} \frac{q_{j,j+1}^k(P)}{q_{j+1,j}^k(P)} \right)}, \quad \forall n \in [N_k]^+. \quad (5-30)$$

以上公式表明 $\varphi_{k,n}(P)$ 无法被显式地表示成 P 的函数, 因为它依赖于 $q_{i,j}^k(P)$, 而求解 $q_{i,j}^k(P)$ 需要知道 $\varphi_{-k}(P)$ 。因此, 期望收益率 $J_k(P_k, P_{-k})$ 是关于 P 的隐函数, 这导致求解纳什均衡是非常难的, 因为它等价于优化一系列带有非线性约束的非线性目标函数。已有的求解纳什均衡的方法 (包括最优响应迭代 [86,123] 和质反应映射 [112,129]) 都需要收益函数的显式表达式。此外, 计算纳什均衡需要知道竞争对手的全部信息 (包括他们的需求函数), 然而, 这些信息在现实世界中通常是无法获取的。因此, 广告商很难达到纳什

均衡。为了解决这些问题，当优化服务商 k 的策略时，先将其他服务商的稳态概率分布向量 $\varphi_{-k} = (\varphi_1, \dots, \varphi_{k-1}, \varphi_{k+1}, \dots, \varphi_K)$ 看成是固定的（即不随着 P 而改变）。在这种假设下，服务商 k 的稳态概率分布， $\widehat{\varphi}_k(P|\varphi_{-k}) = (\widehat{\varphi}_{k,0}(P|\varphi_{-k}), \widehat{\varphi}_{k,1}(P|\varphi_{-k}), \dots, \widehat{\varphi}_{k,N_k}(P|\varphi_{-k}))$ ，可以基于下面的线性方程来计算：

$$\sum_{n \in [N_k]^+} \widehat{\varphi}_{k,n}(P|\varphi_{-k}) = 1, \quad (5-31)$$

$$\widehat{\varphi}_k(P|\varphi_{-k}) \cdot Q_k(P|\varphi_{-k}) = 0, \quad (5-32)$$

其中 $Q_k(P|\varphi_{-k})$ 和 $Q_k(P)$ 是一样的，除了将定义（公式 (5-21)）中的 $\varphi_{-k}(P)$ 替换成了 φ_{-k} 。对应的期望收益率函数为

$$\widehat{J}_k(P_k, P_{-k}|\varphi_{-k}) = \sum_{n=0}^{N_k} \widehat{\varphi}_{k,n}(P|\varphi_{-k}) \cdot n \cdot p_{k,n}. \quad (5-33)$$

公式 (5-33) 中定义的收益率是 P 的显函数。如果所有服务商的目标都是最大化上述收益率函数，最终的定价策略向量， $\widehat{P}^* = \times_{k \in [K]} \widehat{P}_k^*$ ，满足 $\forall k \in [K]$ 和 $P_k \in \Delta_k$ ，

$$\widehat{J}_k(\widehat{P}_k^*, \widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*)) \geq \widehat{J}_k(P_k, \widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*)), \quad (5-34)$$

根据定义1，价格策略向量 \widehat{P}^* 不是一个纳什均衡，而是一个 AE，定义如下。

定义2 (近似均衡) 一个 ϵ -AE 是一个策略向量 \widehat{P}^* ，满足 $\forall k \in [K]$ ，

$$J_k(\widehat{P}_k^*, \widehat{P}_{-k}^*) + \epsilon_k \geq J_k(P_k, \widehat{P}_{-k}^*), \forall P_k \in \Delta_k, \quad (5-35)$$

其中 $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_K)$ 。

可以看到 ϵ 是衡量 \widehat{P}^* 和 P^* 之间距离的指标（即 \widehat{P}^* 与纳什均衡 P^* 有多接近），如果对任意的 $k \in [K]$ ， $\epsilon_k = 0$ 成立，则 $P^* = \widehat{P}^*$ 。由于

$$\widehat{J}_k(\widehat{P}_k^*, \widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*)) = J_k(\widehat{P}_k^*, \widehat{P}_{-k}^*) \quad (5-36)$$

于是有

$$\epsilon_k = \max_{P_k \in \Delta_k} J_k(P_k, \widehat{P}_{-k}^*) - \widehat{J}_k(\widehat{P}_k^*, \widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*)). \quad (5-37)$$

也就是说， ϵ_k 可以被看做服务商 k 可以通过单独偏离 \widehat{P}^* 能获得的最大额外收益。需要注意的是 \widehat{P}^* 是方程 (5-34) 的解，而不是 ϵ 的函数。亦即，给定模型的其他参数， \widehat{P}^* 和 ϵ 都是确定的。后文通过实验说明 ϵ 是很小的。

AE 策略 \widehat{P}^* 比纳什均衡 P^* 更具有现实意义，因为服务商 k 通常不知道其他竞争对手的需求函数 ($\lambda_{k'}(\cdot)$ 和 $\mu_{k'}(\cdot), \forall k' \neq k$)，因此无法计算 P^* 。然而，每个服务商 k 可以观察到其他人的定价策略 P_{-k} 以及他们的稳态概率分布 φ_{-k} ，因而可以基于此优化自己的策略，这将导致最终的价格策略收敛到 AE 策略 \widehat{P}^* ，细节将在下节讨论。

5.3.2 均衡计算

本节利用第4章中提出的基于最优响应迭代的算法思想来计算 AE。在本章的模型中， \widehat{P}_k^* 可以被看做对其他人策略向量 \widehat{P}_{-k}^* 和稳态概率分布向量 $\varphi_{-k}(\widehat{P}^*)$ 的最优响应。此处正式定义服务商 k 对 P_{-k} 和 φ_{-k} 的最优响应为：

$$\widehat{B}_k(P_{-k}|\varphi_{-k}) = \arg \max_{P'_k \in \Delta_k} \widehat{J}_k(P'_k, P_{-k}|\varphi_{-k}). \quad (5-38)$$

计算 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的细节将在下一章节给出。然后可以按如下方式构建一个 SPG：图中的结点对应于一个策略向量 $P \in \Delta$ ，结点 P 的出边集合记为 $\varphi_P = \{\widehat{B}_k(P_{-k}|\varphi_{-k})|k \in [K]\}$ ，其中 $\Delta = \times_{k \in [K]} \Delta_k$ 且 $\varphi_{-k} = \varphi_{-k}(P)$ 。基于 SPG 和最优响应的计算 AE 的完整过程显示在算法6中。

Algorithm 6: AE 计算

```

1 while True do
2   随机选择一个起始点  $P \in \Delta$ ;
3   if 如果  $P$  满足公式 (5-34) then
4     return  $P$ ;
5   将路径  $A$  初始化为空集;
6   while True do
7     从  $\Delta$  和所有  $\varphi_{P'}$ ,  $P' \in \Delta$ , 中删除  $P$ ;
8     if  $\varphi_P$  非空 then
9       随机从  $\varphi_P$  中选择一个  $P'$ ;
10      if  $P'$  满足公式 (5-34) then
11        return  $P'$ ;
12      将  $P$  添加到  $A$ ;
13       $P \leftarrow P'$ ;
14    else
15      if  $A$  是空集 then
16        Break;
17       $P \leftarrow A$  的最后一个元素;
18      从  $A$  中删除  $P$ ;

```

定理 7 如果 \widehat{P}^* 存在，算法6保证收敛到 \widehat{P}^* 。

证明 首先证明第 6 行中的循环一定会终止。从初始点 P^0 开始, 算法 6 在第 6-18 行搜索 \widehat{P}^* , 具体规则为 1) 如果结点 P 有一个子结点 P' (第 8 行), 则将 P 添加到路径 A (第 12 行) 并跳转到 P' (第 13 行), 否则 (第 14 行), 2) 返回到 P 的父结点 (第 17 行), 从 A 中删除它的父结点 (第 18 行), 然后在下一轮迭代中将上述规则应用到它的父结点。第 2 层循环在找到 \widehat{P}^* (第 10 行) 或者 A 是空集 (第 15 行) 的时候终止, 否则, A 的长度要么增加 (第 12 行) 要么减少 (第 18 行)。因此, 为了证明第 2 层循环总会终止, 只需要证明 A 的长度会在某一时刻停止改变。由于算法在访问结点 P 后会从 SPG 中删除它 (第 7 行, 因此每个结点最多被访问一次) 且 $|\Delta|$ 是有限集合, A 的长度要么因为 \widehat{P}^* 被找到 (第 10 行), 要么因为当前结点没有子结点而停止增加。从那以后, A 的长度只能减少, 直到 A 变成空集 (第 15 行)。因此第 2 层循环一定会终止。

如果第 2 层循环中找到了 \widehat{P}^* , 算法终止。否则 (如果 A 是空集), P^0 能跳转到的所有结点都没有连接到一个 \widehat{P}^* 。由此, 如果初始的 SPG 中有连接到 \widehat{P}^* 的路径, 则已经删除的结点不会切换这些路径。因为 $|\Delta|$ 是有限的集合且随着算法运行越变越小, 连接 \widehat{P}^* 的初始点一定会被从 Δ 中采到 (第 2 行, 如果 \widehat{P}^* 存在), 即, 第一层循环总会终止且返回 \widehat{P}^* 。 \square

5.3.3 计算最优响应

算法 6 剩下的问题是如何计算 $\varphi(P)$ 以及对 P_{-k} 和 $\varphi_{-k} = \varphi_{-k}(P)$ 的最优响应 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 。前者可以用标准的牛顿迭代法来解决。接下来主要研究 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的计算。使用最优响应能得到的期望收益率为

$$\widehat{J}_k^*(P_{-k}|\varphi_{-k}) = \widehat{J}_k(\widehat{B}_k(P_{-k}|\varphi_{-k}), P_{-k}|\varphi_{-k}). \quad (5-39)$$

由于 $\lambda_k(p_k, p_{-k})$ 和 $\mu_k(p_k, p_{-k})$ 都是有界的, 服务商 k 的连续时间马尔科夫过程可被一致化为一个离散时间马尔科夫链, 其转移概率矩阵为 [174-176]

$$T_k(P|\varphi_{-k}) = (t_{i,j}^k(P|\varphi_{-k}))_{i,j} = I + \frac{Q_k(P|\varphi_{-k})}{v_k}, \quad (5-40)$$

其中 I 是单位矩阵,

$$v_k \geq \max_{p_k \in \delta_k, p_{-k} \in \delta_{-k}} \{\lambda_k(p_k, p_{-k}) + \mu_k(p_k, p_{-k})\} \quad (5-41)$$

是一致化参数。然后, 给定 P 和 φ_{-k} , 服务商 k 的状态在任意时间点从 n 跳转到 $n+1, n-1$ 和 n 的概率分别为 $t_{n,n+1}^k(P|\varphi_{-k})$, $t_{n,n-1}^k(P|\varphi_{-k})$ 和 $t_{n,n}^k(P|\varphi_{-k}) = 1 - t_{n,n+1}^k(P|\varphi_{-k}) - t_{n,n-1}^k(P|\varphi_{-k})$, 其中

$$t_{n,n+1}^k(P|\varphi_{-k}) = \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_{k,n}, p_{-k})\}}{v_k}, \quad (5-42)$$

$$t_{n,n-1}^k(P|\varphi_{-k}) = \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(p_{k,n}, p_{-k})\}}{v_k}. \quad (5-43)$$

自然地，定义

$$t_{N_k, N_k+1}^k(P|\varphi_{-k}) = t_{0,-1}^k(P|\varphi_{-k}) = 0. \quad (5-44)$$

令 $M_{k,n}(P_{-k}|\varphi_{-k})$ 表示服务商 k 以 n 为初始状态并以最优响应 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 为定价策略的马尔科夫过程。令 $R_{k,n}(P_{-k}|\varphi_{-k})$ 表示上述过程的期望收益。假设 $M_{k,n}(P_{-k}|\varphi_{-k})$ 在一开始的 $m > 0$ 个离散时间点内保持在状态 n ，然后在时刻 $m + 1$ 跳转到状态 $n + 1$ ，该事件发生的概率为 $t_{n,n}^k(\widehat{B}_k(P_{-k}|\varphi_{-k}), P_{-k}|\varphi_{-k})^m \cdot t_{n,n+1}^k(\widehat{B}_k(P_{-k}|\varphi_{-k}), P_{-k}|\varphi_{-k})$ 。这之后，期望上 $M_{k,n}$ 在时刻 i 的状态与 $M_{k,n+1}(P_{-k}|\varphi_{-k})$ 在时刻 $i - m$ 的状态是一样的，因为他们使用了相同的策略 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 。 $M_{k,n+1}(P_{-k}|\varphi_{-k})$ 和 $M_{k,n}(P_{-k}|\varphi_{-k})$ 在上述假设下的期望收益的差值等于

$$m(\widehat{J}_k^*(P_{-k}|\varphi_{-k}) - n\widehat{b}_k(P_{-k}|\varphi_{-k})), \quad (5-45)$$

其中 $\widehat{b}_{k,n}(P_{-k}|\varphi_{-k})$ 是 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的第 n 个元素，对应于服务商 k 对状态 n 设置的最优响应价格。如果 $M_{k,n}(P_{-k}|\varphi_{-k})$ 在时刻 $m + 1$ 跳转到了状态 $n - 1$ ，也可以推导出类似的结果。于是有

$$\begin{aligned} & R_{k,n+1} + R_{k,n-1} - R_{k,n} \\ = & \sum_{m=1}^{\infty} (t_{n,n}^k)^m [t_{n,n+1}^k (m(\widehat{J}_k^* - n\widehat{b}_k) + R_{k,n-1}) + t_{n,n-1}^k (m(\widehat{J}_k^* - n\widehat{b}_k) + R_{k,n+1})], \end{aligned} \quad (5-46)$$

其中为了表达式清晰，隐去了 $(\widehat{B}_k(P_{-k}|\varphi_{-k}), P_{-k}|\varphi_{-k})$ 和 $(P_{-k}|\varphi_{-k})$ 。从以上公式可以得出， $\forall k \in [K]$ 和 $n \in [N_k]^+$,

$$\widehat{J}_k^* = n\widehat{b}_{k,n} + t_{n,n+1}^k(R_{k,n+1} - R_{k,n}) + t_{n,n-1}^k(R_{k,n-1} - R_{k,n}). \quad (5-47)$$

最优响应价格 $\widehat{b}_{k,n}$ 于是可以用标准的动态规划算法高效地求解，例如策略迭代 [177]，其基本思想是首先为所有 $n \in [N_k]^+$ 随机初始化 $R_{k,n}$ ，然后重复以下三个步骤直到 $\widehat{b}_{k,n}$ ， $\forall n \in [N_k]^+$ ，不再改变：

- 1) 通过最大化公式 (5-47) 右侧来计算 $\widehat{b}_{k,n}$ ， $\forall n \in [N_k]^+$ ；
- 2) 基于公式 (5-39) 计算 \widehat{J}_k^* ；
- 3) 为所有 $n \in [N_k]^+$ 更新 $R_{k,n}$ 。

5.4 重要性质

本节研究 AE 策略和收益率的重要性质。第一个是关于 $R_{k,n}(P_{-k}|\varphi_{-k})$ 的单调性。

定理 8 ($R_{k,n}(P_{-k}|\varphi_{-k})$ 的单调性) 对所有的 $k \in [K]$ 、 $P_{-k} \in \Delta_{-k}$ 和 $\varphi_{-k} \in \times_{i \in C(k)} [0, 1]^{N_i+1}$ ，有 $\forall n \in [N_k - 1]$,

$$R_{k,n+1}(P_{-k}|\varphi_{-k}) \geq R_{k,n}(P_{-k}|\varphi_{-k}). \quad (5-48)$$

证明 给定 P_{-k} 和 φ_{-k} , 考虑 k 的系统的两个副本, 第一个, 用系统 A 标识, 以 $n+1$ 为起始状态, 第二个, 用系统 B 标识, 以状态 n 为起始状态。令 B 使用最优响应策略, 令 A 在任意时刻都设定跟 B 一样的价格。由此, B 的期望总收益为 $R_{k,n}(P_{-k}|\varphi_{-k})$ 。由于 A 和 B 在任意时刻的价格都一样, 可以期望他们经历了相同的需求增加和减少过程。有两个特殊情况需要处理。第一个是 A 在状态 N_k 且 B 在状态 $N_k - 1$, 此时一个新的需求到来, 在下个时刻, A 继续停留在 N_k , B 则跳转到 N_k , 从此之后, A 和 B 将总是保持相同的状态。对第二种情况, A 在状态 1 且 B 在状态 0 且此时需求减少, 可以进行类似的分析。

总之, A 中被用户利用的资源数量总是不会少于 B 的。因此, A 的期望总收益不会少于 $R_{k,n}(P_{-k}|\varphi_{-k})$ 。如果令 A 使用最优响应策略 $\widehat{B}_k(P_{-k}|\varphi_{-k})$, 则 A 会得到更高的期望总收益, 即, $R_{k,n+1}(P_{-k}|\varphi_{-k}) \geq R_{k,n}(P_{-k}|\varphi_{-k})$ 。 \square

定理8说明最大期望收益, $R_{k,n}(P_{-k}|\varphi_{-k})$, 随着系统中资源的利用率的提高而增加。下面的定理进一步证明了 $R_{k,n}(P_{-k}|\varphi_{-k})$ 的上凸性, 即

$$U_{k,n}(P_{-k}|\varphi_{-k}) \geq U_{k,n+1}(P_{-k}|\varphi_{-k}) \geq 0, \quad (5-49)$$

其中

$$U_{k,n}(P_{-k}|\varphi_{-k}) = R_{k,n}(P_{-k}|\varphi_{-k}) - R_{k,n-1}(P_{-k}|\varphi_{-k}). \quad (5-50)$$

定理 9 ($R_{k,n}(P_{-k}|\varphi_{-k})$ 是凹函数) 对所有的 $k \in [K]$ 、 $P_{-k} \in \Delta_{-k}$ 和 $\varphi_{-k} \in \times_{i \in C(k)} [0, 1]^{N_i+1}$, 有 $\forall n \in [N_k - 1]$,

$$U_{k,n}(P_{-k}|\varphi_{-k}) \geq U_{k,n+1}(P_{-k}|\varphi_{-k}). \quad (5-51)$$

证明 本节用数学归纳法来证明上述定理, 该方法包含两个主要步骤。第一, 证明

$$U_{k,1}(P_{-k}|\varphi_{-k}) \geq U_{k,2}(P_{-k}|\varphi_{-k}). \quad (5-52)$$

可以将公式 (5-47) 重写为

$$\begin{aligned} & \widehat{J}_k^*(P_{-k}|\varphi_{-k}) \\ = & \max_{p_{k,n}} \{np_{k,n} + t_{n,n+1}^k(P|\varphi_{-k}) \cdot U_{k,n+1}(P_{-k}|\varphi_{-k}) - t_{n,n-1}^k(P|\varphi_{-k})U_{k,n}(P_{-k}|\varphi_{-k})\}. \end{aligned} \quad (5-53)$$

为了公式简单令 $\widehat{b}_{k,n}$ 表示 $\widehat{b}_{k,n}(P_{-k}|\varphi_{-k})$ 并定义

$$g_k(p_k, P_{-k}, \varphi_{-k}) = \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k^{min}, p_{-k})\}}{v_k} - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k, p_{-k})\}}{v_k}. \quad (5-54)$$

于是有

$$\widehat{J}_k^*(P_{-k}|\varphi_{-k}) = 0 \cdot \widehat{b}_{k,0} + \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,0}, p_{-k})\}}{v_k} \cdot U_{k,1}(P_{-k}|\varphi_{-k}) \quad (5-55)$$

$$= U_{k,1}(P_{-k}|\varphi_{-k}) \left(\frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k^{\min}, p_{-k})\}}{v_k} - g_k(\widehat{b}_{k,0}, P_{-k}, \varphi_{-k}) \right) \quad (5-56)$$

$$= 1 \cdot \widehat{b}_{k,1} + \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,1}, p_{-k})\}}{v_k} \cdot U_{k,2}(P_{-k}|\varphi_{-k}) - \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,1}, p_{-k})\}}{v_k} U_{k,1}(P_{-k}|\varphi_{-k}) \quad (5-57)$$

$$\geq p_k^{\min} + U_{k,2}(P_{-k}|\varphi_{-k}) \left(\frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k^{\min}, p_{-k})\}}{v_k} - g_k(p_k^{\min}, P_{-k}, \varphi_{-k}) \right) - U_{k,1}(P_{-k}|\varphi_{-k}) \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(p_k^{\min}, p_{-k})\}}{v_k} \quad (5-58)$$

$$= p_k^{\min} + U_{k,2}(P_{-k}|\varphi_{-k}) \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k^{\min}, p_{-k})\}}{v_k}. \quad (5-59)$$

综合考虑第三个和最后一个公式，有

$$\frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(p_k^{\min}, p_{-k})\}}{v_k} (U_{k,1}(P_{-k}|\varphi_{-k}) - U_{k,2}(P_{-k}|\varphi_{-k})) \geq p_k^{\min} + U_{k,1}(P_{-k}|\varphi_{-k}) g_k(\widehat{b}_{k,0}, P_{-k}, \varphi_{-k}) \quad (5-60)$$

$$\geq 0, \quad (5-61)$$

上式表明 $U_{k,1}(P_{-k}|\varphi_{-k}) \geq U_{k,2}(P_{-k}|\varphi_{-k})$ 。

第二步是要证明 $\forall n \in \{1, 2, \dots, N_k - 2\}$ ，如果

$$U_{k,n}(P_{-k}|\varphi_{-k}) \geq U_{k,n+1}(P_{-k}|\varphi_{-k}), \quad (5-62)$$

那么

$$U_{k,n+1}(P_{-k}|\varphi_{-k}) \geq U_{k,n+2}(P_{-k}|\varphi_{-k}). \quad (5-63)$$

首先假设 $U_{k,n+1}(P_{-k}|\varphi_{-k}) < U_{k,n+2}(P_{-k}|\varphi_{-k})$ ，据此有 $\forall n \in \{1, 2, \dots, N_k - 2\}$ ，

$$\widehat{J}_k^*(P_{-k}|\varphi_{-k}) = n \cdot \widehat{b}_{k,n} + \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n}, p_{-k})\}}{v_k} \cdot U_{k,n+1}(P_{-k}|\varphi_{-k}) - \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n}, p_{-k})\}}{v_k} U_{k,n}(P_{-k}|\varphi_{-k}) \quad (5-64)$$

$$< (n+1) \widehat{b}_{k,n} + \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n}, p_{-k})\}}{v_k} U_{k,n+2}(P_{-k}|\varphi_{-k}) - \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n}, p_{-k})\}}{v_k} U_{k,n+1}(P_{-k}|\varphi_{-k}) \leq (n+1) \widehat{b}_{k,n+1} + \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n+1}, p_{-k})\}}{v_k} U_{k,n+2}(P_{-k}|\varphi_{-k}) - \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n+1}, p_{-k})\}}{v_k} U_{k,n+1}(P_{-k}|\varphi_{-k}) \quad (5-65)$$

$$= \widehat{J}_k^*(P_{-k}|\varphi_{-k}), \quad (5-66)$$

即

$$\widehat{J}_k^*(P_{-k}|\varphi_{-k}) < \widehat{J}_k^*(P_{-k}|\varphi_{-k}), \quad (5-67)$$

这不成立。因此，如果 $U_{k,n}(P_{-k}|\varphi_{-k}) \geq U_{k,n+1}(P_{-k}|\varphi_{-k})$ ，那么有 $U_{k,n+1}(P_{-k}|\varphi_{-k}) \geq U_{k,n+2}(P_{-k}|\varphi_{-k})$ 。□

基于定理8和9，可以推导出最优响应策略 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的单调性。

定理 10 ($\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的单调性) 对所有的 $k \in [K]$ 、 $P_{-k} \in \Delta_{-k}$ 和 $\varphi_{-k} \in \times_{i \in C(k)} [0, 1]^{N_i+1}$ ，有 $\forall n \in [N_k - 1]$ ，

$$\widehat{b}_{k,n}(P_{-k}|\varphi_{-k}) \leq \widehat{b}_{k,n+1}(P_{-k}|\varphi_{-k}). \quad (5-68)$$

证明 为了公式简洁，用 $\widehat{b}_{k,n}$ 来表示 $\widehat{b}_{k,n}(P_{-k}|\varphi_{-k})$ 。从公式(5-47)可以得出

$$\begin{aligned} \widehat{J}_k^*(P_{-k}|\varphi_{-k}) &= n \cdot \widehat{b}_{k,n} + \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n}, P_{-k})\}}{v_k} \\ &\quad \cdot U_{k,n+1}(P_{-k}|\varphi_{-k}) - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n}, P_{-k})\}}{v_k} U_{k,n}(P_{-k}|\varphi_{-k}) \end{aligned} \quad (5-69)$$

$$\begin{aligned} &\geq n \cdot \widehat{b}_{k,n-1} + \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n-1}, P_{-k})\}}{v_k} U_{k,n+1}(P_{-k}|\varphi_{-k}) \\ &\quad - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n-1}, P_{-k})\}}{v_k} U_{k,n}(P_{-k}|\varphi_{-k}). \end{aligned} \quad (5-70)$$

类似地，有

$$\begin{aligned} &(n-1) \cdot \widehat{b}_{k,n-1} + \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n-1}, P_{-k})\}}{v_k} U_{k,n}(P_{-k}|\varphi_{-k}) \\ &\quad - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n-1}, P_{-k})\}}{v_k} U_{k,n-1}(P_{-k}|\varphi_{-k}) \end{aligned} \quad (5-71)$$

$$\begin{aligned} &\geq (n-1) \cdot \widehat{b}_{k,n} + \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n}, P_{-k})\}}{v_k} U_{k,n}(P_{-k}|\varphi_{-k}) \\ &\quad - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n}, P_{-k})\}}{v_k} U_{k,n-1}(P_{-k}|\varphi_{-k}). \end{aligned} \quad (5-72)$$

由上面两个不等式可得

$$\begin{aligned} \widehat{b}_{k,n} - \widehat{b}_{k,n-1} &\geq (U_{k,n}(P_{-k}|\varphi_{-k}) - U_{k,n+1}(P_{-k}|\varphi_{-k})) \\ &\quad \cdot \left(\frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n}, P_{-k})\}}{v_k} - \frac{\mathbf{E}_{P_{-k} \in P_{-k}}^{\varphi_{-k}} \{\lambda_k(\widehat{b}_{k,n-1}, P_{-k})\}}{v_k} \right) \\ &\quad + (U_{k,n-1}(P_{-k}|\varphi_{-k}) - U_{k,n}(P_{-k}|\varphi_{-k})) \end{aligned}$$

$$\cdot \left(\frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n-1}, p_{-k})\}}{v_k} - \frac{\mathbf{E}_{p_{-k} \in P_{-k}}^{\varphi_{-k}} \{\mu_k(\widehat{b}_{k,n}, p_{-k})\}}{v_k} \right). \quad (5-73)$$

如果 $\widehat{b}_{k,n} < \widehat{b}_{k,n-1}$, 则根据定理9和公式(5-2)及(5-5), 上面不等式的右侧不会小于零, 于是有

$$0 > \widehat{b}_{k,n} - \widehat{b}_{k,n-1} \geq 0. \quad (5-74)$$

上式不成立。因此 $\widehat{b}_{k,n} \geq \widehat{b}_{k,n-1}$ 。□

定理10有现实的意义。一方面, 当系统负载很重, 服务商趋向于设定更高的价格从而获得更高的收益, 同时防止需求继续增加造成平台过载。另一方面, 当系统有很多闲置资源时, 服务商通过降低价格来吸引用户。此外, 定理可以被用来缩减计算最优响应 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的算法的搜索空间。具体而言, $\widehat{b}_{k,n}(P_{-k}|\varphi_{-k})$ 的搜索空间由 $[p_k^{\min}, p_k^{\max}]$ 被降低到 $[\widehat{b}_{k,n-1}(P_{-k}|\varphi_{-k}), p_k^{\max}]$, 这让 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 的计算更加高效。

推论3 (\widehat{P}_k^* 的单调性) 令 \widehat{P}_k^* 表示 $(\widehat{p}_{k,0}^*, \widehat{p}_{k,1}^*, \dots, \widehat{p}_{k,N_k}^*)$, 则有对所有的 $k \in [K]$ 和 $n \in [N_k - 1]$,

$$\widehat{p}_{k,n}^* \leq \widehat{p}_{k,n+1}^*. \quad (5-75)$$

证明 这是定理10的一个直接的推论, 因为 $\widehat{P}_k^* = \widehat{B}_k(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 。□

5.5 实验评估

本节在实验中使用如下的需求增加和减少函数, 所有的评估方法可以直接被用于其他任何形式的满足公式(5-2)-(5-5)的需求函数。具体而言, 定义

$$\lambda_k(p_1, p_2, \dots, p_K) = l_k(1 - p_k^2) \frac{\sum_{i \in C(k)} p_i^2}{K - 1}, \quad (5-76)$$

$$\mu_k(p_1, p_2, \dots, p_K) = u_k p_k^2 \frac{\sum_{i \in C(k)} (1 - p_i^2)}{K - 1}, \quad (5-77)$$

其中 l_k 和 u_k 是参数。令 $\delta_k = \{0, 0.001, 0.002, \dots, 1\}$, $\forall k \in [K]$, 并在实验中令 $K = 3$, 因为现实世界中大约有 3~4 个大的云服务公司, 例如微软, 亚马逊, IBM, 阿里巴巴等。注意算法6中有些随机性, 包括第 2 行中结点的选择以及计算 $\varphi(P)$ 时使用的牛顿迭代法的随机初始点。下面实验章节中所有的结果都是 100 次实验的平均数据。算法由 Python 2.7.13 实现, 实验平台为一台装有 64 位 Windows 系统并搭载有 64GB 内存和 4 个 3.4GHz 处理器的计算机。结果表明不同批次实验中 $\widehat{p}_{k,n}^*$, $\forall k \in [K]$ 和 $n \in [N_k]^+$, 之间的差值小于 $\pm 10^{-5}$ 。因此, 假定 \widehat{P}^* 是唯一的。

5.5.1 运行时间评估

本章的算法通过策略迭代法求解方程 (5-47) 来计算 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ ，策略搜索的空间为 δ_k 。第5.4节中的定理10可以极大地缩减该空间。为了展示缩减的效用，评估算法6中使用原始的搜索空间 (R_o) 和缩减后的搜索空间 (R_r) 来计算最优响应 ($\widehat{B}_k(P_{-k}|\varphi_{-k}), \forall k \in [K]$) 的时间。实验结果显示在表5.2中，其中 $N_1 = N_2 = N_3 = N$ ，性能提升比根据

$$\frac{|R_o - R_r|}{R_o} \quad (5-78)$$

来计算。

可以看到性能的提升随着资源总数而提高，这与期望的结果是一致的，因为 N 越大计算 $\widehat{B}_k(P_{-k}|\varphi_{-k})$ 时减少的冗余计算就越多。此外，提升比的增长率随着 N 的增加而减小，这是因为 $\widehat{b}_{k,n}(P_{-k}|\varphi_{-k})$ 随着 N_k 的增加而减小（类似的现象可从图5.1(a)和5.3(b)中观察到），因此缩减的搜索空间随着 N 增加而减小。总的来说，可以利用定理10极大地提高算法6的性能。

表 5.2 运行时间比较

N	R_o (seconds)	R_r (seconds)	Ratio (%)
10	2.19	1.68	23.1
20	6.52	4.62	29.0
30	8.99	5.94	34.2
40	11.36	7.12	37.3
50	14.51	9.08	39.1
60	17.18	10.27	40.2

5.5.2 资源总数对 AE 的影响

本节首先研究 AE 策略在资源总数不同时的性质。本节中令 $l_1 = l_2 = l_3 = 1.4$ ， $u_1 = u_2 = u_3 = 1$ 。服务商的资源总数分别设定为 $N_1 = 10$ ， $N_2 = 15$ ， $N_3 = 20$ ^①。本节将 AE 策略 \widehat{P}^* 画在图5.1(a)中并将 $U_{k,n}(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 展示在图5.1(b)中。实验结果与给出的定理相吻合。具体而言，图5.1(b)表明 $U_{k,n}(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*)) \geq 0$ ，这验证了 $R_{k,n}(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 的单调性（定理8）；此外， $U_{k,n}(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 随着 n 的增加而递减，这表明 $R_{k,n}(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 是凹函数（定理9）；进一步地， $\widehat{p}_{k,n}^*$ 随着 n 的增加而增加，正如定理10所陈述的。可以观察到当 $N_k > N_{k'}$ 时 $\widehat{p}_{k,n}^* > \widehat{p}_{k',n}^*$ ，这是合理的，因为当平台上有很多资源未被利用时，服务商需要设置很低的价格来吸引用户。三个服务商的最大收益率分别为 7.667，11.647 和

① 使用其他数据得到的实验结果是相似的。

15.661。一个有趣的发现是

$$\frac{\widehat{J}_k^*(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))}{\widehat{J}_{k'}^*(\widehat{P}_{-k'}^*|\varphi_{-k'}(\widehat{P}^*))} \approx \frac{N_k}{N_{k'}}, \quad (5-79)$$

即，当服务商的需求增加和减少函数一样时， $\widehat{J}_k^*(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ 是 N_k 的线性函数。接下来，本节通过实验研究需求函数的参数是如何影响服务商的定价策略和收益率的。

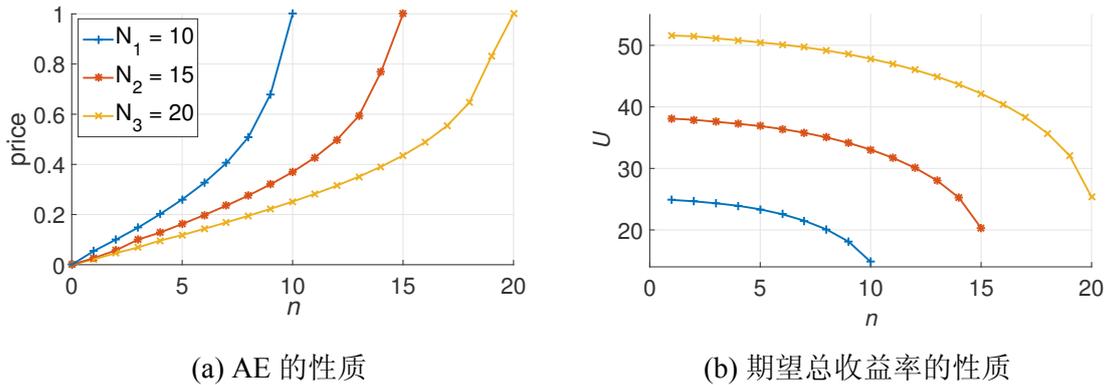


图 5.1 资源总数对 AE 的影响

5.5.3 需求函数参数对 AE 的影响

不同服务商的需求函数通常是不一样的，因为他们能提供的服务的质量不同。在实验中，需求函数之间的差别由参数 l_k 和 u_k 来表。本节研究服务商的 AE 策略和收益率是如何随着这些参数改变的。

可以基于公式 (5-47) 验证得到 $\widehat{b}_{k,0}(P_{-k}|\varphi_{-k}) = p_k^{\min}$ 。因此有 $\widehat{p}_{k,0}^* = p_k^{\min}, \forall k \in [K]$ 。这部分的实验包括两个场景，如表 5.3 所示，其中 $\widehat{p}_{k,0}^* = 0$ 和 $\widehat{p}_{k,6}^* = 1$ 被缺省了。为了方便令 $N_1 = N_2 = N_3 = 6$ 。在第一个场景中，假设服务商拥有相同的需求减少函数和不同的需求增加函数。可以看到较大的 l_k 能带来较高的价格和 $\widehat{J}_k^*(\widehat{P}_{-k}^*|\varphi_{-k}(\widehat{P}^*))$ ，如果令两个服务商参数相同，则他们的 AE 策略和得到的收益率也相同。较大的 l_k 意味着对应的服务商在设定与竞争对手相同的价格时对用户的吸引力更大，即该服务商在市场中更具有竞争优势，因此可以设定相对较高的价格来获得更高的收益。对于第二种场景，即服务商有不同的需求减少函数和相同的需求增加函数，可以得到类似的实验结果。

5.5.4 忽视市场竞争的收益损失

为了评估本章提出的 \widehat{P}^* 的优势，将它与文献中已有的最好的非竞争定价策略 [84, 90, 92] 进行比较。非竞争定价策略最大化每个服务商 k 的 $\sum_{n=0}^{N_k} \varphi_k(P_k) n p_{k,n}$ ，但不考虑其他服务商的策略 P_{-k} 。用表 5.3 中两个场景的设定，并计算服务商 k 使用非竞争定价策略的收益率，即 k 使用

$$\lambda_k(p_k) = l_k(1 - p_k)^2 \quad (5-80)$$

表 5.3 需求函数参数对 AE 的影响

k	l_k	u_k	$\widehat{p}_{k,1}^*$	$\widehat{p}_{k,2}^*$	$\widehat{p}_{k,3}^*$	$\widehat{p}_{k,4}^*$	$\widehat{p}_{k,5}^*$	\widehat{J}_k^*
1	2	1	.081	.171	.270	.392	.577	4.639
2	1.6	1	.073	.168	.265	.385	.565	4.519
3	1.2	1	.067	.164	.259	.375	.550	4.350
1	1.6	0.8	.079	.172	.271	.394	.581	4.675
2	1.6	1	.074	.169	.267	.387	.569	4.560
3	1.6	1.2	.068	.167	.263	.381	.560	4.458

和

$$\mu_k(p_k) = u_k p_k^2 \tag{5-81}$$

为其需求增加和减少函数，而其他服务商使用算法6来计算自己的定价策略。比较结果显示在图5.2中，其中的每个服务商 k 的两个直方图表示 k 在以下两种情况下的最大收益率：1) 所有服务商都用 AE 策略 \widehat{P}^* ；2) 服务商 k 采用上述非竞争策略。可以看到在两个场景中，服务商使用 AE 策略 \widehat{P}^* 能获得更高的收益，非竞争策略会导致大约 10% 的收益损失。实验结果表明本章提出的策略在现实世界中优于已有的策略。

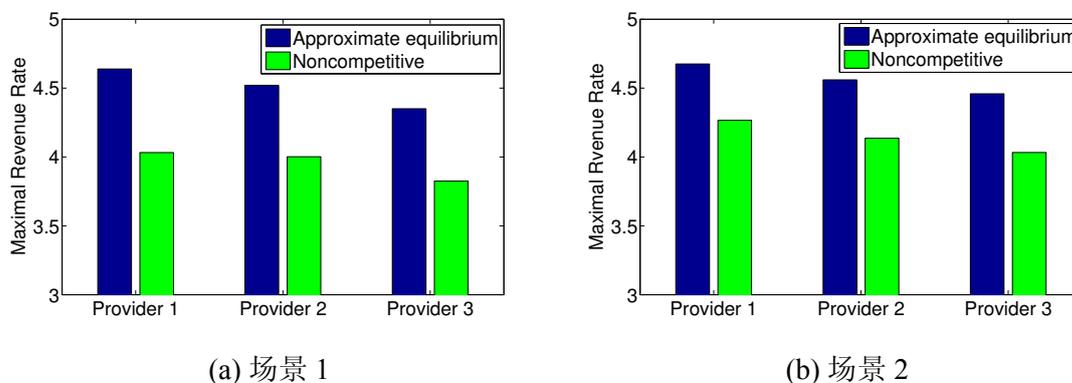


图 5.2 策略比较

5.5.5 收益的提升

本节用实验评估 AE 策略下服务商的收益率随着参数的变化规律，基础的设定同表5.3中的第一个场景。本节测试服务商的收益率如何随着需求到达参数和资源总数变化的。以服务商 3 为参考，用算法6计算每个服务商在不同的 l_3 和 N_3 下的 AE 策略。图5.3(a)中展示了三个服务商的收益率，可以看出它们都随着 l_3 的增加而增加。这个观察结果是合理的，因此较大的 l_3 表明服务商 3 的吸引力更大，因此可以为每个状态设定较高的价格，这反过来提高（降低）了其他服务商的需求增加（减少）率。从图中可以发现收益率是 l_3 的凹函数，且服务商在参数相同的情况下收益率也相同。图5.3(b)画了

服务商 3 的 AE 策略和稳态概率分布随着资源总数 ($N_3 = 6, 8, 10$) 变化的曲线。该图表明当资源总数增加时, 服务商总是倾向于降低价格以提高平台资源的利用率。

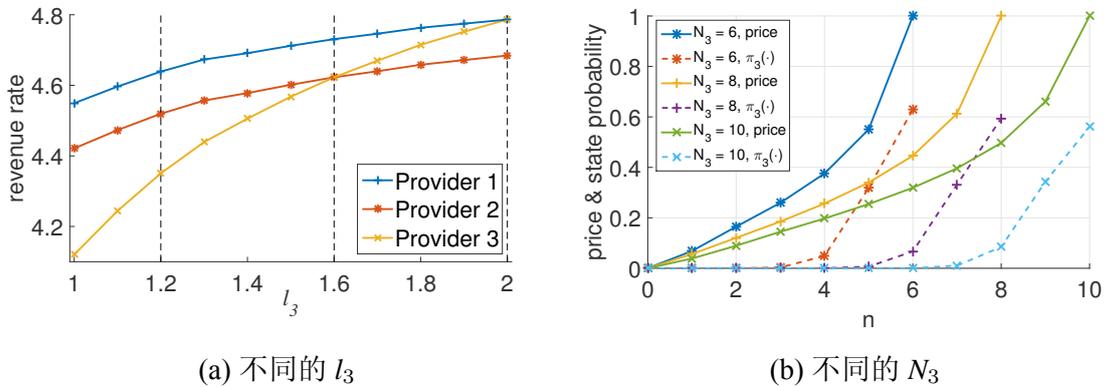


图 5.3 不同参数下的 AE

5.5.6 AE 与纳什均衡的距离

对于每个参数设定, 本节用 Matlab 2016b 的 “fmincon” 函数 (内点法) 计算所有 $k \in [K]$ 的 ϵ_k 。事实上, ϵ_k 可以看成服务商 k 能通过单独偏离 AE 策略 \widehat{P}_k^* 而额外获得的最大收益率。结果显示在表 5.4 中。可以观察到在所有情况下, ϵ_k 与 $\widehat{J}_k^*(\widehat{P}_{-k}^* | \varphi_{-k}(\widehat{P}^*))$ 相比是很小的。表中的比率说明收益最多只增加了 1.66%, 这说明偏离 \widehat{P}_k^* 能得到的好处是非常有限的。因此, 假设服务商采用 \widehat{P}_k^* 是非常合理的。该 AE 在完全信息和不完全信息两种条件下均可计算, 因此是一个具有现实意义的定价策略。

表 5.4 评估 ϵ

k	l_k	u_k	N_k	$\widehat{J}_k^*(\cdot)$	ϵ_k	$\epsilon_k / \widehat{J}_k^*(\cdot)$
1	2	1	6	4.6390	.0620	1.33%
2	1.6	1	6	4.5194	.0553	1.22%
3	1.2	1	6	4.3506	.0463	1.06%
1	1.6	0.8	6	4.6753	.0626	1.34%
2	1.6	1	6	4.5603	.0572	1.25%
3	1.6	1.2	6	4.4589	.0396	0.89%
1	1.4	1	10	7.6678	.1140	1.49%
2	1.4	1	15	11.647	.1869	1.60%
3	1.4	1	20	15.661	.2599	1.66%

5.6 总结

本章研究了提供可重复利用资源的服务市场的动态定价优化问题, 主要做了三点贡献。第一, 本章提出了一个综合性的模型来描述市场的动态和竞争特性。第二, 本章将

服务商的最优定价策略规划成了一个 AE 并设计了求解算法。第三，本章推导了模型的很多有用的性质，同时未对需求函数做过多假设。实验结果表明本章提出的定价策略比文献中已有的策略性能更好。

本章的算法也可以被扩展来处理 \widehat{P}^* 不存在的情况。具体而言，可以将结点 P 的最优响应边的集合扩展为

$$\{P_k | P_k \geq \widehat{B}_k(P_{-k} | \varphi_{-k}(P)) - \xi, k \in [K]\}. \quad (5-82)$$

相应地，将算法6中的终止条件由公式 (5-34) 修改为 $\forall k \in [K]$ 和 $P_k \in \Delta_k$,

$$\widehat{J}_k(\widehat{P}_k, \widehat{P}_{-k} | \varphi_{-k}(\widehat{P}^*)) + \xi \geq \widehat{J}_k(P_k, \widehat{P}_{-k} | \varphi_{-k}(\widehat{P}^*)). \quad (5-83)$$

这样可以通过设定合适的 ξ 使得算法6总能收敛到一个 $\epsilon + \xi$ (近似) 纳什均衡。

第六章 总结与展望

本章首先总结本文的主要工作，然后给出下一步的工作安排。

6.1 本文总结

本文以付费搜索广告竞标和云计算服务为具体对象，研究了互联网应用中的模型建立、均衡计算和收益优化问题，具体包括以下四个工作：1. 有限理性模型研究，在博弈参与方无法精确获知环境和对手信息时，有限理性模型能更好地模拟不确定性；2. 基于保留价的收益优化，其中用于计算保留价的信息未知，考虑到有限时间内的总收益，获取信息的采样数影响着全局收益；3. 基于长远利益的定价策略优化，市场在不断发展且竞争激烈，参与者需要制定长远的价格策略来最大化全局收益；4. 基于资源利用率的定价策略优化，在市场发展的每一个阶段，参与者需要动态调整供需关系和资源的利用率来最大化收益。本文针对每个问题进行了建模分析，提出了相应的高效的算法并进行了大量实验评估所提方法的性能。

针对付费搜索中的有限理性的建模问题，本文分析了纳什均衡在实际应用中的局限性，包括只能描述纯策略行为（真实数据显示广告商大多采用混合策略）以及假设广告商知道市场和对手的所有信息，然后引入了质反应均衡的概念来解决这些问题。质反应均衡弱化了纳什均衡对于完全理性的假设，在这种均衡下，期望收益越高的纯策略被采取的概率越高，质反应均衡本身是一个混合策略。本文主要研究了付费搜索中质反应均衡的计算问题和参数预测问题。本文设计了一个基于同伦思想的高效算法来求解该均衡，通过深入分析付费搜索中收益函数的性质，本文开发了专门的加速算法，将复杂度从指数级降为了线性级别。此外，本文还提出了利用质反应均衡的拟合功能来推测付费搜索中的参数的方法，包括广告价值、点击率和广告商的理性程度等。本文在模拟数据集和真实数据集上做了大量实验，结果表明本文提出的加速算法的性能远远优于普通算法，能被用于求解大规模付费搜索中的质反应均衡，此外，质反应均衡对真实数据的拟合能力很强，参数预测效果好于混合策略纳什均衡模型。

针对付费搜索中基于保留价的收益优化问题，本文考虑了更加现实的情况，即广告价值分布未知且无历史数据时如何制定保留价格。Myerson 定理给出了广告价值分布已知时计算最优保留价格的基础理论。对于分布未知的情况，需要先通过采样来获取足够信息。由于广告价值分布不会一直不变，本文研究如何设置采样时间来最大化有限时间内搜索引擎的总收益，并将该问题被规划成了一个整数约束优化问题，最优解不依赖于具体的价值分布形式，具有鲁棒性。为了求解最优采样时间，本文通过一系列等价问题转换将原问题变换成了一个连续约束优化问题。通过分析目标函数的性质，本文证明了

最优采样时间的存在性和唯一性，并基于此提出了一个类似于二分查找的算法，该方法具有指数级的收敛速度。基于模拟数据的实验结果表明，本文提出的方法给搜索引擎带来的收益与知道广告价值分布时的理论最大收益非常接近，前者与后者的比值超过 0.9。

针对云计算中基于长远利益的定价优化问题，本文充分考虑了市场的各种特性，包括用户数量的增长、市场中的不完全信息及服务商之间的竞争等，提出了一个具有现实意义的模型。为了求解最优的定价策略，本文将模型规划成一个随机博弈，其中每个服务商独立且同时制定价格，当前时刻的定价会考虑到未来的长远利益。本文将最优策略描述为一个马尔科夫完美均衡，为了计算该均衡，本文将随机博弈按照时间线展开成多个标准型子博弈，并设计了一个基于动态规划的算法，该算法的时间复杂度是子博弈个数的线性函数，而传统的 Q-learning 类型的算法需要多次迭代才能收敛到近似解且时间复杂度是关于子博弈个数的指数函数。此外，为了进一步提升效率，本文还提出了计算期望即时收益的加速算法。实验结果表明本文提出的算法能高效地计算马尔科夫完美均衡策略，与文献中已有的最好的策略相比，本文提出的方法能给服务商带来最高的收益。

针对云计算中基于资源利用率的定价优化问题，本文研究了市场中的动态供需关系，提出了一个能够综合考虑资源利用情况和市场竞争的模型。本文阐明了市场中存在不完全信息因而纳什均衡无法达到，并将定价策略定义成了一个近似均衡，该均衡概念在现实中更加实用。本文开发了一个保证收敛的算法来求解近似均衡，此外，还推导了均衡的很多重要的性质，包括动态价格关于资源利用率的单调性和收益函数的凹凸性等，这些性质被进一步利用来提高算法效率。本文通过实验验证了所提的近似均衡策略优于其他已有的策略，且近似均衡很接近纳什均衡。

6.2 下一步工作展望

本文以付费搜索广告和云计算作为具体研究对象，利用博弈论的方法来研究以它们为代表的互联网应用中的均衡计算和收益优化问题。今后的工作主要沿着以下三个方向开展。

第一，本文给出了一个基于同伦的计算质反应均衡的算法，这种方法只能求得均衡的数值解，今后的工作会深入研究质反应均衡的解析解，可以先从简单的质反应函数开始，然后逐步扩展到更为复杂的定义。此外，质反应均衡的存在性证明也是今后工作的重点之一。

第二，考虑广告商策略性反应的保留价优化问题。本文的第二个工作假设广告商采用固定的均衡策略，不会对搜索引擎设置不同的保留价有任何反应。这个假设可以被弱化，即如果广告商知道搜索引擎会通过采用预估广告价值分布，那么广告商可以策略性地在搜索引擎采样期间降低自己的竞价，从而使得搜索引擎计算出一个较低的保留价，这样在采样阶段结束后广告商就能获得更高的收益。这样搜索引擎和广告商之间会形

成一种博弈，保留价优化问题会因此变得更为复杂，但毫无疑问这种假设更符合现实情况，是今后要深入研究的一个课题。

第三，本文从两个维度对云计算中的定价优化问题进行了研究，即基于长远利益和基于资源利用率的策略。其中基于长远利益的定价更加宏观，考虑的是在市场的长远发展过程中如何为不同阶段设置合适的价格，而基于资源利用率的定价则更关注于一个具体的发展阶段。一个值得研究的问题是如何将两种定价策略整合成一个统一的策略，同时考虑长远利益和资源利用率。一种可行的方案包为首先通过长远利益优化得出每个阶段价格的平均值，然后对每个阶段基于资源利用率的价格优化设置额外的限制条件—动态价格的均值要与前者吻合。今后将深入研究该方案。

参考文献

- [1] QIN T, CHEN W, LIU T-Y. Sponsored search auctions: Recent advances and future directions[J]. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2015, 5(4): 60.
- [2] JING S-Y, ALI S, SHE K, et al. State-of-the-art research study for green cloud computing[J]. *The Journal of Supercomputing*, 2013: 1–24.
- [3] YANG H, TATE M. A descriptive literature review and classification of cloud computing research[J]. *Communications of the Association for Information Systems*, 2012, 31(2): 35–60.
- [4] MAILLÉ P, MARKAKIS E, NALDI M, et al. Sponsored search auctions: an overview of research with emphasis on game theoretic aspects[J]. *Electronic Commerce Research*, 2012, 12(3): 265–300.
- [5] SAKELLARI G, LOUKAS G. A survey of mathematical models, simulation approaches and testbeds used for research in cloud computing[J]. *Simulation Modelling Practice and Theory*, 2013, 39: 92–103.
- [6] DUONG Q, LAHAIE S. Discrete choice models of bidder behavior in sponsored search[C] // *International Workshop on Internet and Network Economics*. 2011: 134–145.
- [7] TIAN L-Q, LIN C, NI Y. Evaluation of user behavior trust in cloud computing[C] // *Computer Application and System Modeling (ICCSM)*, 2010 International Conference on: Vol 7. 2010: V7–567.
- [8] XU H, GAO B, YANG D, et al. Predicting advertiser bidding behaviors in sponsored search by rationality modeling[C] // *Proceedings of the 22nd international conference on World Wide Web*. 2013: 1433–1444.
- [9] OSTROVSKY M, SCHWARZ M. Reserve prices in internet advertising auctions: a field experiment[C] // *Proceedings of the 12th ACM conference on Electronic commerce*. 2011: 59–60.
- [10] XU H, LI B. Dynamic cloud pricing for revenue maximization[J]. *IEEE Transactions on Cloud Computing*, 2013, 1(2): 158–171.
- [11] XU B, QIN T, QIU G, et al. Optimal Pricing for the Competitive and Evolutionary Cloud Market.[C] // *IJCAI*. 2015: 139–145.
- [12] EDELMAN B, SCHWARZ M. Optimal auction design and equilibrium selection in sponsored search auctions[J]. *The American Economic Review*, 2010: 597–602.
- [13] PRASAD A S, RAO S. A mechanism design approach to resource procurement in cloud computing[J]. *IEEE Transactions on Computers*, 2014, 63(1): 17–30.
- [14] THOMPSON D R, LEYTON-BROWN K. Revenue optimization in the generalized second-price auction[C] // *Proceedings of the fourteenth ACM conference on Electronic commerce*. 2013: 837–852.
- [15] HARSANYI J C, SELTEN R, OTHERS. *A general theory of equilibrium selection in games*[J]. MIT Press Books, 1988, 1.
- [16] MYERSON R B. *Game theory*[M]. [S.l.]: Harvard university press, 2013.
- [17] FRIEDMAN J W. *Game theory with applications to economics*[M]. [S.l.]: Oxford University Press New York, 1990.
- [18] RABIN M. Incorporating fairness into game theory and economics[J]. *The American economic review*, 1993: 1281–1302.

- [19] NEUMANN J V, MORGENSTERN O, OTHERS. Theory of games and economic behavior[M]. [S.l.]: Princeton university press Princeton, 1944.
- [20] EDELMAN B, OSTROVSKY M, SCHWARZ M. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords[J]. The American economic review, 2007, 97(1): 242–259.
- [21] JEHIEL P, MOLDOVANU B, STACCHETTI E. Multidimensional mechanism design for auctions with externalities[J]. Journal of economic theory, 1999, 85(2): 258–293.
- [22] KALAI E, LEHRER E. Rational learning leads to Nash equilibrium[J]. Econometrica: Journal of the Econometric Society, 1993: 1019–1045.
- [23] MASKIN E. Nash equilibrium and welfare optimality[J]. The Review of Economic Studies, 1999, 66(1): 23–38.
- [24] BRAMS S J, TAYLOR A D. Fair Division: From cake-cutting to dispute resolution[M]. [S.l.]: Cambridge University Press, 1996.
- [25] MIRCHANDANI R S. Superadditivity and subadditivity in fair division[J]. Journal of Mathematics Research, 2013, 5(3): 78.
- [26] WANG Z, YIN Y, AN B. Computing Optimal Monitoring Strategy for Detecting Terrorist Plots.[C] // AAAI. 2016: 637–643.
- [27] YIN Y, AN B. Efficient Resource Allocation for Protecting Coral Reef Ecosystems.[C] // IJCAI. 2016: 531–537.
- [28] SHOHAM Y. Computer science and game theory[J]. Communications of the ACM, 2008, 51(8): 74–79.
- [29] YANG S, GAO Y, AN B, et al. Efficient Average Reward Reinforcement Learning Using Constant Shifting Values.[C] // AAAI. 2016: 2258–2264.
- [30] SHIEH E, AN B, YANG R, et al. Protect: A deployed game theoretic system to protect the ports of the united states[C] // Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1. 2012: 13–20.
- [31] YIN Z, JIANG A X, JOHNSON M P, et al. TRUSTS: Scheduling Randomized Patrols for Fare Inspection in Transit Systems.[C] // IAAI. 2012.
- [32] DASKALAKIS C, GOLDBERG P W, PAPADIMITRIOU C H. The complexity of computing a Nash equilibrium[J]. SIAM Journal on Computing, 2009, 39(1): 195–259.
- [33] GIBBONS R. A primer in game theory[M]. [S.l.]: Harvester Wheatsheaf, 1992.
- [34] NASH J. Non-cooperative games[J]. Annals of mathematics, 1951: 286–295.
- [35] REPORT. Analyzing AWS EC2 Price Drops over the Past 5 Years[J]. <https://www.cloudyn.com/blog/analyzing-aws-ec2-price-drops-over-the-past-5-years/>, 2015.
- [36] ABHISHEK V, HOSANAGAR K. Keyword generation for search engine advertising using semantic similarity between terms[C] // Proceedings of the ninth international conference on Electronic commerce. 2007: 89–94.
- [37] JOSHI A, MOTWANI R. Keyword generation for search engine advertising[C] // Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on. 2006: 490–496.
- [38] RAVI S, BRODER A, GABRILOVICH E, et al. Automatic generation of bid phrases for online advertising[C] // Proceedings of the third ACM international conference on Web search and data mining. 2010: 341–350.

- [39] CHOI Y, FONTOURA M, GABRILOVICH E, et al. Using landing pages for sponsored search ad selection[C] // Proceedings of the 19th international conference on World wide web. 2010 : 251 – 260.
- [40] RADLINSKI F, BRODER A, CICCULO P, et al. Optimizing relevance and revenue in ad search: a query substitution approach[C] // Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. 2008 : 403 – 410.
- [41] WANG H, LIANG Y, FU L, et al. Efficient query expansion for advertisement search[C] // Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. 2009 : 51 – 58.
- [42] KIM S, QIN T, LIU T-Y, et al. Advertiser-centric approach to understand user click behavior in sponsored search[J]. Information Sciences, 2014, 276 : 242 – 254.
- [43] SCULLEY D, MALKIN R G, BASU S, et al. Predicting bounce rates in sponsored search advertisements[C] // Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. 2009 : 1325 – 1334.
- [44] XIONG C, WANG T, DING W, et al. Relational click prediction for sponsored search[C] // Proceedings of the fifth ACM international conference on Web search and data mining. 2012 : 493 – 502.
- [45] DING W, QIN T, ZHANG X-D, et al. Multi-Armed Bandit with Budget Constraint and Variable Costs.[C] // AAAI. 2013.
- [46] MUTHUKRISHNAN S, PÁL M, SVITKINA Z. Stochastic models for budget optimization in search-based advertising[C] // International Workshop on Web and Internet Economics. 2007 : 131 – 142.
- [47] ZHANG W, ZHANG Y, GAO B, et al. Joint optimization of bid and budget allocation in sponsored search[C] // Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012 : 1177 – 1185.
- [48] GATTI N, LAZARIC A, TROVÒ F. A truthful learning mechanism for multi-slot sponsored search auctions with externalities[C] // Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3. 2012 : 1325 – 1326.
- [49] GHOSH A, SAYEDI A. Expressive auctions for externalities in online advertising[C] // Proceedings of the 19th international conference on World wide web. 2010 : 371 – 380.
- [50] MILGROM P. Simplified mechanisms with an application to sponsored-search auctions[J]. Games and Economic Behavior, 2010, 70(1) : 62 – 70.
- [51] EDELMAN B, OSTROVSKY M. Strategic bidder behavior in sponsored search auctions[J]. Decision support systems, 2007, 43(1) : 192 – 198.
- [52] LAHAIE S, PENNOCK D M, SABERI A, et al. Sponsored search auctions[J]. Algorithmic game theory, 2007 : 699 – 716.
- [53] ANON. [J], .
- [54] YAO S, MELA C F, OTHERS. Sponsored search auctions: Research opportunities in marketing[J]. Foundations and Trends® in Marketing, 2009, 3(2) : 75 – 126.
- [55] VARIAN H R. Position auctions[J]. International Journal of Industrial Organization, 2007, 25(6) : 1163 – 1178.
- [56] BORGERS T, COX I, PESENDORFER M, et al. Equilibrium bids in sponsored search auctions: Theory and evidence[J], 2007.

- [57] CARY M, DAS A, EDELMAN B, et al. Greedy bidding strategies for keyword auctions[C] // Proceedings of the 8th ACM conference on Electronic commerce. 2007 : 262–271.
- [58] GOMES R, SWEENEY K. Bayes–nash equilibria of the generalized second-price auction[J]. Games and Economic Behavior, 2014, 86 : 421–437.
- [59] BU T-M, DENG X, QI Q. Forward looking Nash equilibrium for keyword auction[J]. Information Processing Letters, 2008, 105(2) : 41–46.
- [60] BU T-M, LIANG L, QI Q. On robustness of forward-looking in sponsored search auction[J]. Algorithmica, 2010, 58(4) : 970–989.
- [61] BU T-M, DENG X, QI Q. Dynamics of strategic manipulation in ad-words auction[C] // Proc. of the 3rd Workshop on Ad Auctions, Banff, Canada. 2007.
- [62] VOROBAYCHIK Y, REEVES D M. Equilibrium analysis of dynamic bidding in sponsored search auctions[J]. International Journal of Electronic Business, 2008, 6(2) : 172–193.
- [63] NISAN N, SCHAPIRA M, VALIANT G, et al. Best-response auctions[C] // Proceedings of the 12th ACM conference on Electronic commerce. 2011 : 351–360.
- [64] HE D, CHEN W, WANG L, et al. A game-theoretic machine learning approach for revenue maximization in sponsored search[C] // Twenty-Third International Joint Conference on Artificial Intelligence. 2013.
- [65] LUCIER B, PAES LEME R, TARDOS E. On revenue in the generalized second price auction[C] // Proceedings of the 21st international conference on World Wide Web. 2012 : 361–370.
- [66] HARTLINE J D, ROUGHGARDEN T. Simple versus optimal mechanisms[C] // Proceedings of the 10th ACM conference on Electronic commerce. 2009 : 225–234.
- [67] MYERSON R B. Optimal auction design[J]. Mathematics of operations research, 1981, 6(1) : 58–73.
- [68] NISAN N, ROUGHGARDEN T, TARDOS E, et al. Algorithmic game theory : Vol 1[M]. [S.l.] : Cambridge University Press Cambridge, 2007.
- [69] SUN Y, ZHOU Y, DENG X. Optimal reserve prices in weighted GSP auctions[J]. Electronic Commerce Research and Applications, 2014, 13(3) : 178–187.
- [70] SUN Y, ZHOU Y, YIN M, et al. On the convergence and robustness of reserve pricing in keyword auctions[C] // Proceedings of the 14th Annual International Conference on Electronic Commerce. 2012 : 113–120.
- [71] COLE R, ROUGHGARDEN T. The sample complexity of revenue maximization[C] // Proceedings of the 46th Annual ACM Symposium on Theory of Computing. 2014 : 243–252.
- [72] BULOW J, KLEMPERER P. Auctions vs. negotiations[R]. [S.l.] : National Bureau of Economic Research, 1994.
- [73] DHANGWATNOTAI P, ROUGHGARDEN T, YAN Q. Revenue maximization with a single sample[J]. Games and Economic Behavior, 2015, 91 : 318–333.
- [74] BYUN E-K, KEE Y-S, KIM J-S, et al. Cost optimized provisioning of elastic resources for application workflows[J]. Future Generation Computer Systems, 2011, 27(8) : 1011–1026.
- [75] KONG X, LIN C, JIANG Y, et al. Efficient dynamic task scheduling in virtualized data centers with fuzzy prediction[J]. Journal of network and Computer Applications, 2011, 34(4) : 1068–1077.
- [76] PAMIES-JUAREZ L, GARCÍA-LÓPEZ P, SÁNCHEZ-ARTIGAS M, et al. Towards the design of optimal data redundancy schemes for heterogeneous cloud storage infrastructures[J]. Computer Networks, 2011, 55(5) : 1100–1113.

- [77] JOHNSON J L. SQL in the Clouds[J]. *Computing in Science & Engineering*, 2009, 11(4) : 12–28.
- [78] LIN K W, DENG D-J. A novel parallel algorithm for frequent pattern mining with privacy preserved in cloud computing environments[J]. *International Journal of Ad Hoc and Ubiquitous Computing*, 2010, 6(4) : 205–215.
- [79] CHEN G, WU Y, LIU J, et al. Optimization of sub-query processing in distributed data integration systems[J]. *Journal of Network and Computer Applications*, 2011, 34(4) : 1035–1042.
- [80] LI J, LI B, WO T, et al. CyberGuarder: A virtualization security assurance architecture for green cloud computing[J]. *Future Generation Computer Systems*, 2012, 28(2) : 379–390.
- [81] CHONKA A, XIANG Y, ZHOU W, et al. Cloud security defence to protect cloud computing against HTTP-DoS and XML-DoS attacks[J]. *Journal of Network and Computer Applications*, 2011, 34(4) : 1097–1107.
- [82] FENG Y, LI B, LI B. Price competition in an oligopoly market with multiple iaas cloud providers[J]. *IEEE Transactions on Computers*, 2014, 63(1) : 59–73.
- [83] KANTERE V, DASH D, FRANCOIS G, et al. Optimal service pricing for a cloud cache[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2011, 23(9) : 1345–1358.
- [84] XU H, LI B. Maximizing revenue with dynamic cloud pricing: The infinite horizon case[C] // *Communications (ICC), 2012 IEEE International Conference on*. 2012 : 2929–2933.
- [85] VENGEROV D. A gradient-based reinforcement learning approach to dynamic pricing in partially-observable environments[J]. *Future Generation Computer Systems*, 2008, 24(7) : 687–693.
- [86] TRUONG-HUU T, THAM C-K. A game-theoretic model for dynamic pricing and competition among cloud providers[C] // *Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*. 2013 : 235–238.
- [87] TRUONG-HUU T, THAM C-K. A novel model for competition and cooperation among cloud providers[J]. *IEEE Transactions on Cloud Computing*, 2014, 2(3) : 251–265.
- [88] GERAGHTY M K, JOHNSON E. Revenue management saves national car rental[J]. *Interfaces*, 1997, 27(1) : 107–127.
- [89] WEATHERFORD L R, KIMES S E. A comparison of forecasting methods for hotel revenue management[J]. *International journal of forecasting*, 2003, 19(3) : 401–415.
- [90] ELMAGHRABY W, KESKINOCAK P. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions[J]. *Management Science*, 2003, 49(10) : 1287–1309.
- [91] GALLEGO G, VAN RYZIN G. Optimal dynamic pricing of inventories with stochastic demand over finite horizons[J]. *Management Science*, 1994, 40(8) : 999–1020.
- [92] PASCHALIDIS I C, TSITSIKLIS J N. Congestion-dependent pricing of network services[J]. *IEEE/ACM Transactions on Networking*, 2000, 8(2) : 171–184.
- [93] XU X, HOPP W J. A monopolistic and oligopolistic stochastic flow revenue management model[J]. *Operations Research*, 2006, 54(6) : 1098–1109.
- [94] LEVIN Y, MCGILL J, NEDIAK M. Dynamic pricing in the presence of strategic consumers and oligopolistic competition[J]. *Management science*, 2009, 55(1) : 32–46.
- [95] RONG J, QIN T, AN B, et al. Optimal Sample Size for Adword Auctions[C] // *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. 2016 : 1459–1460.

- [96] EDELMAN B, OSTROVSKY M, SCHWARZ M. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords[R]. [S.I.]: National Bureau of Economic Research, 2005.
- [97] BÖRGERS T, COX I, PESENDORFER M, et al. Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence[J]. *American economic Journal: microeconomics*, 2013, 5(4): 163–187.
- [98] GOEREE J K, HOLT C A, PALFREY T R. Regular quantal response equilibrium[J]. *Experimental Economics*, 2005, 8(4): 347–367.
- [99] HAILE P A, HORTAÇSU A, KOSENOK G. On the empirical content of quantal response equilibrium[J]. *The American Economic Review*, 2008, 98(1): 180–200.
- [100] MCKELVEY R D, PALFREY T R. Quantal response equilibria for normal form games[J]. *Games and economic behavior*, 1995, 10(1): 6–38.
- [101] MCKELVEY R D, PALFREY T R. Quantal response equilibria for extensive form games[J]. *Experimental economics*, 1998, 1(1): 9–41.
- [102] BROWDER F E. On continuity of fixed points under deformations of continuous mappings[J]. *Summa Brasiliensis Mathematicae*, 1960, 4: 183–191.
- [103] KELLOGG R B, LI T-Y, YORKE J. A constructive proof of the Brouwer fixed-point theorem and computational results[J]. *SIAM Journal on Numerical Analysis*, 1976, 13(4): 473–483.
- [104] SMART D R. Fixed point theorems: Vol 66[M]. [S.I.]: CUP Archive, 1980.
- [105] DASKALAKIS C, MEHTA A, PAPADIMITRIOU C. Progress in approximate Nash equilibria[C] // *Proceedings of the 8th ACM conference on Electronic commerce*. 2007: 355–358.
- [106] ALLGOWER E L, GEORG K. Introduction to numerical continuation methods: Vol 45[M]. [S.I.]: SIAM, 2003.
- [107] ALLGOWER E L, GEORG K. Numerical continuation methods: an introduction: Vol 13[M]. [S.I.]: Springer Science & Business Media, 2012.
- [108] LIAO S. On the homotopy analysis method for nonlinear problems[J]. *Applied Mathematics and Computation*, 2004, 147(2): 499–513.
- [109] GOLDBERG P W, PAPADIMITRIOU C H, SAVANI R. The complexity of the homotopy method, equilibrium selection, and Lemke-Howson solutions[J]. *ACM Transactions on Economics and Computation*, 2013, 1(2): 9.
- [110] HERINGS P J-J, PEETERS R. Homotopy methods to compute equilibria in game theory[J]. *Economic Theory*, 2010, 42(1): 119–156.
- [111] RONG J, QIN T, AN B. Computing quantal response equilibrium for sponsored search auctions[C] // *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. 2015: 1803–1804.
- [112] TUROCY T L. A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence[J]. *Games and Economic Behavior*, 2005, 51(2): 243–263.
- [113] MCKELVEY R D, MCLENNAN A M, TUROCY T L. Gambit: Software Tools for Game Theory, Version 14.1.0. <http://www.gambit-project.org>. [J], 2014.
- [114] AKAIKE H. Information theory and an extension of the maximum likelihood principle[G] // *Selected Papers of Hirotugu Akaike*. 1998: 199–213.

- [115] JOHANSEN S, JUSELIUS K. Maximum likelihood estimation and inference on cointegration with applications to the demand for money[J]. *Oxford Bulletin of Economics and statistics*, 1990, 52(2): 169–210.
- [116] MURSHUDOV G N, VAGIN A A, DODSON E J. Refinement of macromolecular structures by the maximum-likelihood method[J]. *Acta Crystallographica Section D: Biological Crystallography*, 1997, 53(3): 240–255.
- [117] JUDD K L. *Numerical methods in economics*[M]. [S.l.]: MIT press, 1998.
- [118] JIANG A X, LEYTON-BROWN K. A polynomial-time algorithm for Action-Graph Games[C] // *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE: Vol 21*. 2006: 679.
- [119] JIANG A X, LEYTON-BROWN K, BHAT N A. Action-graph games[J]. *Games and Economic Behavior*, 2011, 71(1): 141–173.
- [120] THOMPSON D R M, LEYTON-BROWN K. Computational analysis of perfect-information position auctions[C] // *Proceedings of the 10th ACM conference on Electronic commerce*. 2009: 51–60.
- [121] BLACK P E. *Dictionary of algorithms and data structures*[M]. [S.l.]: National Institute of Standards and Technology, 2004.
- [122] GRAEPEL T, CANDELA J Q, BORCHERT T, et al. Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft’s bing search engine[C] // *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 2010: 13–20.
- [123] GOEMANS M, MIRROKNI V, VETTA A. Sink equilibria and convergence[C] // *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*. 2005: 142–151.
- [124] SHI S W, DONG X. The Effects of Bid Pulsing on Keyword Performance in Search Engines[J]. *International Journal of Electronic Commerce*, 2015, 19(2): 3–38.
- [125] YAHOO. Webscope for Advertising and Market Data. A3: ydata-ysm-keyphrase-bid-imp-click-v1_0. <http://webscope.sandbox.yahoo.com/>[C] //. 2013.
- [126] CARAGIANNIS I, KAKLAMANIS C, KANELLOPOULOS P, et al. On the efficiency of equilibria in generalized second price auctions[C] // *Proceedings of the 12th ACM conference on Electronic commerce*. 2011: 81–90.
- [127] CHEN W, HE D, LIU T-Y, et al. Generalized second price auction with probabilistic broad match[C] // *Proceedings of the fifteenth ACM conference on Economics and computation*. 2014: 39–56.
- [128] MA W, WU T, QIN T, et al. Generalized second price auctions with value externalities[C] // *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 2014: 1457–1458.
- [129] RONG J, QIN T, AN B, et al. Modeling Bounded Rationality for Sponsored Search Auctions[C] // *Proceedings of the 22nd European Conference on Artificial Intelligence*. 2016: 515–523.
- [130] XIAO B, YANG W, LI J. Optimal reserve price for the generalized second-price auction in sponsored search advertising[J]. *Journal of Electronic Commerce Research*, 2009, 10(3): 114.
- [131] RADOVANOVIC A, HEAVLIN W D. Risk-aware revenue maximization in display advertising[C] // *Proceedings of the 21st international conference on World Wide Web*. 2012: 91–100.
- [132] CARAGIANNIS I, KAKLAMANIS C, KANELLOPOULOS P, et al. Revenue guarantees in the generalized second price auction[J]. *ACM Transactions on Internet Technology (TOIT)*, 2014, 14(2-3): 17.

- [133] YUAN S, WANG J, CHEN B, et al. An empirical study of reserve price optimization in real-time bidding[C] // Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014 : 1897 – 1906.
- [134] MOHRI M, MEDINA A M. Non-parametric Revenue Optimization for Generalized Second Price Auctions[C] // Proceedings of the 31st conference on Uncertainty in Artificial Intelligence. 2015.
- [135] CUI T, PENG L, PARDOE D, et al. Data-Driven Reserve Prices for Social Advertising Auctions at LinkedIn[C] // Proceedings of the ADKDD'17. 2017 : 5.
- [136] HE L, WANG L, LIU K, et al. Optimizing Sponsored Search Ranking Strategy by Deep Reinforcement Learning[J]. arXiv preprint arXiv:1803.07347, 2018.
- [137] SHEN W, PENG B, LIU H, et al. Reinforcement mechanism design, with applications to dynamic pricing in sponsored search auctions[J]. arXiv preprint arXiv:1711.10279, 2017.
- [138] KLEINBERG R, LEIGHTON T. The value of knowing a demand curve: Bounds on regret for on-line posted-price auctions[C] // Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science. 2003 : 594 – 605.
- [139] CESA-BIANCHI N, GENTILE C, MANSOUR Y. Regret minimization for reserve prices in second-price auctions[C] // Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms. 2013 : 1190 – 1204.
- [140] COPE E W. Regret and convergence bounds for a class of continuum-armed bandit problems[J]. IEEE Transactions on Automatic Control, 2009, 54(6) : 1243 – 1253.
- [141] JIA Y Y, MANNOR S. Unimodal bandits[C] // Proceedings of the 28th International Conference on Machine Learning. 2011 : 41 – 48.
- [142] BUBECK S, MUNOS R, STOLTZ G, et al. X-armed bandits[J]. Journal of Machine Learning Research, 2011, 12(May) : 1655 – 1695.
- [143] ABRAMS Z, SCHWARZ M. Ad auction design and user experience[G] // Internet and Network Economics. 2007 : 529 – 534.
- [144] HARTLINE J. Lectures on optimal mechanism design[J]. Lecture notes, 2006.
- [145] LAATIKAINEN G, OJALA A, MAZHELIS O. Cloud services pricing models[G] // Software Business. From Physical Products to Software Services and Solutions. 2013 : 117 – 129.
- [146] SHARMA B, THULASIRAM R K, THULASIRAMAN P, et al. Pricing cloud compute commodities: a novel financial economic model[C] // Proceedings of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. 2012 : 451 – 457.
- [147] WANG H, JING Q, CHEN R, et al. Distributed systems meet economics: pricing in the cloud[C] // Proceedings of the 2nd USENIX conference on Hot topics in cloud computing. 2010 : 6 – 6.
- [148] FENG Y, LI B, LI B. Price Competition in an Oligopoly Market with Multiple IaaS Cloud Providers[J]. IEEE Transactions on Computers, 2014, 63(1) : 59 – 73.
- [149] THUIJSMAN F. Optimality and equilibria in stochastic games : Vol 82[M]. [S.l.] : CWI, 1992.
- [150] DORASZELSKI U, ESCOBAR J F. A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification[J]. Theoretical Economics, 2010, 5(3) : 369 – 402.
- [151] MASKIN E, TIROLE J. Markov perfect equilibrium: I. Observable actions[J]. Journal of Economic Theory, 2001, 100(2) : 191 – 219.

- [152] CHIEN S, SINCLAIR A. Convergence to approximate Nash equilibria in congestion games[C] // Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. 2007 : 169–178.
- [153] BOWLING M, VELOSO M. Multiagent learning using a variable learning rate[J]. Artificial Intelligence, 2002, 136(2) : 215–250.
- [154] BUSONI L, BABUSKA R, DE SCHUTTER B. A comprehensive survey of multiagent reinforcement learning[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2008, 38(2) : 156–172.
- [155] JIANG B, CHEN P-Y, MUKHOPADHYAY T. Software licensing: Pay-per-use versus perpetual[J], 2007.
- [156] MA H, LIU C, KING I, et al. Probabilistic factor models for web site recommendation[C] // Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011 : 265–274.
- [157] OLIVER F. Methods of estimating the logistic growth function[J]. Applied statistics, 1964 : 57–66.
- [158] PEARL R, REED L J. On the rate of growth of the population of the United States since 1790 and its mathematical representation[J]. Proceedings of the National Academy of Sciences of the United States of America, 1920, 6(6) : 275–288.
- [159] JUNG H, KLEIN C M. Optimal inventory policies under decreasing cost functions via geometric programming[J]. European Journal of Operational Research, 2001, 132(3) : 628–642.
- [160] ALLON G, GURVICH I. Pricing and dimensioning competing large-scale service providers[J]. Manufacturing & Service Operations Management, 2010, 12(3) : 449–469.
- [161] LIN K Y, SIBDARI S Y. Dynamic price competition with discrete customer choices[J]. European Journal of Operational Research, 2009, 197(3) : 969–980.
- [162] HU J, WELLMAN M P. Multiagent reinforcement learning: theoretical framework and an algorithm.[C] // Proceedings of the 15th International Conference on Machine Learning : Vol 98. 1998 : 242–250.
- [163] SCHNEIDER M L, WHITLATCH E E. User-specific water demand elasticities[J]. Journal of Water Resources Planning and Management, 1991, 117(1) : 52–73.
- [164] PARK S-T, CHU W. Pairwise preference regression for cold-start recommendation[C] // Proceedings of the third ACM conference on Recommender systems. 2009 : 21–28.
- [165] CHOI S, MATTILA A S. Hotel revenue management and its impact on customers' perceptions of fairness[J]. Journal of Revenue and Pricing Management, 2004, 2(4) : 303–314.
- [166] ŞEN A. A comparison of fixed and dynamic pricing policies in revenue management[J]. Omega, 2013, 41(3) : 586–597.
- [167] WANG C, MA W, QIN T, et al. Selling reserved instances in cloud computing[C] // Proceedings of the 24th International Conference on Artificial Intelligence. 2015 : 224–230.
- [168] GOPALRATNAM K, KAUTZ H, WELD D S. Extending continuous time Bayesian networks[C] // Proceedings of the National Conference on Artificial Intelligence : Vol 20. 2005 : 981.
- [169] NORRIS J R. Markov chains[M]. [S.l.] : Cambridge University Press, 1998.
- [170] SIMMONS R G, YOUNES H L. Solving Generalized Semi-Markov Decision Processes using Continuous Phase-type Distributions[C] // Proceedings of the Nineteenth National Conference of Artificial Intelligence. 2004 : 742–747.

- [171] TSAKNAKIS H, SPIRAKIS P G. An optimization approach for approximate Nash equilibria[J]. *Internet Mathematics*, 2008, 5(4): 365–382.
- [172] DOCKNER E, JØRGENSEN S. Optimal pricing strategies for new products in dynamic oligopolies[J]. *Marketing Science*, 1988, 7(4): 315–334.
- [173] GUO X, HERNÁNDEZ-LERMA O. Continuous-time Markov decision processes[G] // *Continuous-Time Markov Decision Processes*. [S.l.]: Springer, 2009: 9–18.
- [174] CASSANDRAS C G, LAFORTUNE S. Introduction to discrete event systems[M]. [S.l.]: Springer Science & Business Media, 2009.
- [175] GROSS D, MILLER D R. The randomization technique as a modeling tool and solution procedure for transient Markov processes[J]. *Operations Research*, 1984, 32(2): 343–361.
- [176] STEWART W J. Probability, Markov chains, queues, and simulation: The mathematical basis of performance modeling[M]. [S.l.]: Princeton University Press, 2009.
- [177] BERTSEKAS D P, BERTSEKAS D P, BERTSEKAS D P, et al. Dynamic programming and optimal control: Vol 1[M]. [S.l.]: Athena Scientific Belmont, MA, 1995.

致 谢

博士生涯即将进入尾声，我能顺利完成博士的学业离不开老师、同学、家人和朋友的支持与帮助。在此我要由衷地感谢他们。

我要感谢我的导师安波老师。作为安老师在计算所招收的第一个也是最后一个直博士生，这五年来受到了安老师无数的教导、支持与帮助。犹记得读博初期不在状态，安老师不厌其烦地找我沟通，不停地激励我，一步步带我走入科研正轨。安老师是一个对学生极度负责的好导师，对学生写的论文，他会一字一句地帮修改，大到论文架构，小到标点符号。如果没有安老师的帮助，我不可能在 AAI 等人工智能盛会上发表论文。安老师总是尽可能地为学生提供与国际顶级学者当面交流的机会，让我们从中得到锻炼和拓宽视野。同时，还要感谢安老师在我家中遇到变故的那段时间对我的关怀和鼓励，帮我从阴影中尽快走出来。我能回报的就是更加努力，做出更好的成绩，不忘师恩，不枉费安老师的栽培。

我要感谢微软亚洲研究院的秦涛老师。秦老师对学生总是和蔼可亲，在秦老师的悉心指导下，我从一个科研小白逐渐成长为一个有独立研究能力的博士生。在微软亚洲研究院实习的四年间，我从秦老师身上学到了太多太多东西，包括专业知识、待人处事的方式以及对科研的态度等等。感谢秦老师一直以来对我的帮助，以及在论文每一次被拒时的鼓励，如果没有秦老师的帮助与鼓励，我不可能在历次失败中重拾信心，也就没有现在的研究成果。能在读博期间到微软亚洲研究院实习是幸运的，更幸运的是能遇到秦老师这样的好导师。在今后的学习工作中我会继续努力，期待能再次跟秦老师合作。

我要感谢我的导师陈熙霖老师。是陈老师的支持与帮助，让我能够顺利度过博士生涯中的种种关卡；是陈老师的鼓励与信任，让我有机会申请种种奖学金，让我变得更自信。陈老师对待工作的严谨态度一直潜移默化地影响着我，使我受益匪浅。陈老师是一位德高望重的学术巨人，但待人接物却总是那么平易近人，令人敬佩。在以后的工作生活中，我要以陈老师为标杆，向他学习，提高自己。

我要感谢微软亚洲研究院的刘铁岩老师。刘老师一直对科研抱有极大的热情，这深深感染了我。刘老师渊博的学识和敏捷的思维让人肃然起敬，这也激励我要不断进步，提高自己的专业水平。刘老师总是鼓励我们要培养“争论”的文化，面对任何事情要勇于提出自己的不同看法，勇于质疑前辈们的观点，不要盲从，令我受益颇多，今后我也会继续这样做。

我要感谢尊敬的史忠植老师。史老师在我心中一直是高山仰止的前辈，他对科学研究的锲而不舍的追求精神使我备受鼓舞。感谢史老师在我读博期间对我学术上的指导和生活中的帮助。史老师严谨的治学精神和平易的处事态度都是我要学习的，今后我要以

他为榜样，不断前行。

我要感谢我的师兄甘家瑞和师姐殷越，他们一直是我学习的榜样，每次遇到问题，我第一个想到的都是找他们借鉴些经验，感谢他们总是不厌其烦地帮助与指导我，感谢他们在我遇到挫折时给我的鼓励。我还要感谢胡兰萍老师对我学习生活中的帮助，感谢胡老师耐心提醒我申请各种奖学金，感谢胡老师多年来为课题组的默默付出。此外，我要感谢研究生部的李琳老师、周世佳老师、张平老师、李丹老师、宋守礼老师和冯钢老师等各位老师在我的学习过程中给予的支持和帮助。

最后，我要感谢我的爱人李莹姝、家人和朋友。感谢你们在我读博这五年给我的支持与鼓励，没有你们我不可能完成学业，感谢你们在我背后的无限付出，感谢你们的陪伴，今后我会让自己变得更强大，更好地照顾你们。

作者简介

姓名：荣江 性别：男 出生日期：1990.06.08 籍贯：河北

2013.9 – 现在 中国科学院计算技术研究所直博生

2009.9 – 2013.7 华南理工大学本科生

【攻读博士学位期间发表的论文】

- [1] 荣江, 秦涛, 安波. 广告竞标中的质反应均衡计算及参数预测. 软件学报, 2018 (已接收).
- [2] **Jiang Rong**, Tao Qin and Bo An. Dynamic pricing for reusable resources in competitive market with stochastic demand. *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI'18)*, pp.4718-4726, 2018.
- [3] **Jiang Rong**, Tao Qin, Bo An and Tie-Yan Liu. Revenue maximization for finitely repeated ad auctions. *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI'17)*, pp.663-669, 2017.
- [4] **Jiang Rong**, Tao Qin, Bo An and Tie-Yan Liu. Pricing optimization for selling reusable resources. *Proceedings of the 16th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'17)*, pp.1719-1721, 2017.
- [5] **Jiang Rong**, Tao Qin, Bo An, Tie-Yan Liu. Modeling bounded rationality for sponsored search auctions. *Proceedings of the 22nd European Conference on Artificial Intelligence (ECAI'16)*, pp.515-523, 2016.
- [6] **Jiang Rong**, Tao Qin, Bo An, Tie-Yan Liu. Optimal sample size for adword auctions. *Proceedings of the 15th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'16)*, pp.1459-1460, 2016.
- [7] **Jiang Rong**, Tao Qin, Bo An. Computing quantal response equilibrium for sponsored search auctions. *Proceedings of the 14th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'15)*, pp.1803-1804, 2015.
- [8] **Jiang Rong**, Tao Qin, Bo An. Quantal response equilibrium for sponsored search auctions: Computation and inference. *Proceedings of The 10th workshop on Ad Auctions, in conjunction with the 15th ACM Conference on Electronic Commerce (EC'14)*, 2014.

【攻读博士学位期间参加的科研项目】

- [1] 微软亚洲研究院合作研究项目“云计算中的定价优化”. 2015年7月 – 2018年7月

【攻读博士学位期间的获奖情况】

- | | |
|--------------------------------|------|
| [1] 中科院计算所三好学生标兵 ((获奖比例 < 1%). | 2018 |
| [2] 中科院计算所三好学生 ((获奖比例 < 13%). | 2018 |
| [3] 中科院计算所北纬通信奖 ((获奖比例 < 5%). | 2017 |