

# 互联网广告机制设计的若干前沿问题

唐平中<sup>1</sup> 王子贺<sup>2</sup> 沈蔚然<sup>3</sup> 等

<sup>1</sup> 清华大学

<sup>2</sup> 上海财经大学

<sup>3</sup> 卡耐基梅隆大学

关键词：互联网广告 计算经济学 动态保留价

广告位通常以竞价拍卖的方式在互联网平台（卖家）上分配给广告主（买家）。每位买家向卖家提交一个竞价数值，卖家根据所有买家的竞价，按照预先公布的拍卖机制，将广告位售予某个买家，并收取其相应的费用。广告收入是几乎所有主流互联网公司的主要收入来源，例如，谷歌公司在2019年的广告收入高达1348亿美元<sup>1</sup>。设计广告机制，提高流量的分配效率以及买卖双方的利益，是互联网行业最重要的经济问题，是国家层面互联网的核心竞争力，也是计算与经济交叉学科当前的研究热点。本文从机制设计的角度，介绍互联网广告机制设计中的科学问题与挑战。

## 动态保留价与强化机制设计

保留价 (reserve pricing) 是指买家想要赢得拍卖所需付出的最低价格，一般也是买家参加拍卖所需的最低报价。根据不同的设计形式，我们可以对保留价做进一步的分类，例如根据是否与买家身份相关，保留价可分为匿名保留价 (anonymous reserve) 和个性化保留价 (personalized reserve)；根据是否随

时间变化，保留价可分为静态保留价 (static reserve) 和动态保留价 (dynamic reserve)。保留价机制规则简单，与多种拍卖形式兼容，并且对收益影响效果明显，这些优点使其在许多现行拍卖机制中得到广泛应用。

如何设置合理的保留价需要卖家掌握市场上买家对拍卖商品的估值信息。雅虎 (Yahoo!) 于2009年报道了其对保留价机制进行的大规模实验<sup>[1]</sup>，这是保留价机制在互联网广告拍卖领域的第一次有较大影响力的应用与实验。近年来，随着人工智能技术的快速发展，利用机器学习、神经网络等算法对保留价进行自动化设计的方法得到了各大互联网广告拍卖平台（如百度、谷歌等）大规模的探索和应用。百度公司于2017年开始探索利用强化学习进行保留价设计（即“强化机制设计”）<sup>[2]</sup>。该研究通过实验发现，静态保留价设计容易出现“二阶效应”问题，即买家的竞价行为会随着保留价机制的变化而变化，这使得根据历史数据计算得到的保留价在实际中被应用在一个新环境下时，导致性能表现不如预期。

二阶效应的深层原因来自买家的非理性行为。

<sup>1</sup> <https://www.statista.com/statistics/266249/advertising-revenue-of-google/>。

机制设计理论基于理性人假设,即能够掌握足够的信息,并且拥有足够的计算能力,在竞价前计算出拍卖博弈的某种均衡策略。而这种假设在现实中并不成立。实际上,很多买家拥有复杂的竞价偏好和效益函数,对市场信息的掌握也不完整。因此,要对一套保留价机制的表现进行预估,首先需要对该买家的策略和行为进行建模。百度采用循环神经网络(RNN)来描述买家的策略,即通过买家前  $m$  个时刻的竞价策略和关键绩效指标(Key Performance Indicator, KPI)值,来预测下一个时刻该买家的竞价策略  $s^{(t+1)} = g(s^{(t-m+1:t)}, h^{(t-m+1:t)})$ , 其中  $s^{(t+1)}$  为所有买家在  $t+1$  时刻的竞价策略集合,  $s^{(t-m+1:t)}$  与  $h^{(t-m+1:t)}$  分别为所有买家在前  $m$  时刻的竞价策略集合和 KPI 集合,  $g$  为循环神经网络。买家在  $t+1$  时刻 KPI 集合  $h^{(t+1)}$  可以表示为以  $\gamma^{(t+1)}$  作为保留价时拍卖机制运行的结果。见图 1。

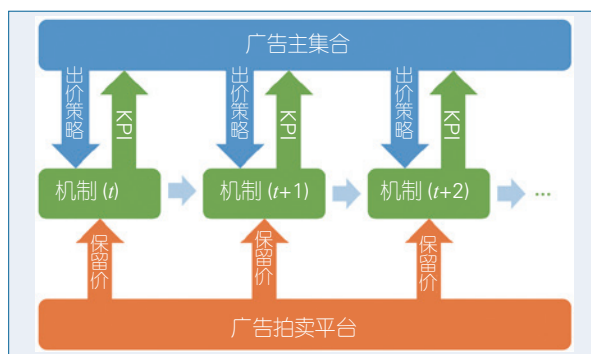


图 1 设计动态保留价的过程

至此,学习保留价问题被完整建模为一个马尔可夫决策过程(Markov Decision Process, MDP)。其中,买家集合是与广告拍卖平台进行交互的环境,在  $t$  时刻,竞价策略集合  $s^t$  为环境的状态,拍卖平台的行为是给出一个保留价集合  $\gamma^t$ ,而状态转移函数则由神经网络  $g$  决定。因此,求解该决策过程就是为平台设计动态保留价的过程。

上述研究采用蒙特卡洛树搜索(Monte-Carlo tree search)的强化学习方法来求解该马尔可夫决策过程。蒙特卡洛树搜索算法的有效性已在多种复杂场景下得到证明(例如谷歌公司的围棋人工智能AlphaGo<sup>[3]</sup>)。在一些简化场景的模拟实验中,强化

机制设计方法,以及其他基于买家马尔可夫行为模型的动态保留价设计方法,可以较好地解决二阶效应问题,与传统静态保留价相比可以大幅度提升平台收益。百度在线广告拍卖系统已实现该强化机制设计算法框架并正式上线运营,并在其 2018 年第一季度财务报表中予以强调。

谷歌近年来也一直积极探索人工智能算法在拍卖设计上的应用,其中也包括利用机器学习方法来进行保留价的设置。此类研究通常需要首先将定价问题描述为一个机器学习问题,例如,最近一项来自谷歌的研究<sup>[4]</sup>考虑第二价格拍卖(second-price auction)中的静态定价问题,并采用监督学习的框架来看待该问题:将每次拍卖数据看成一个数据点  $(x, y)$ , 其中  $x$  为与此次拍卖相关的特征(例如用户的设备类型、操作系统、国家地区等),  $y$  则为此次拍卖中买家的竞价集合。在该问题中,广告拍卖平台需要学习一个定价函数  $p(x)$ , 以使某损失函数  $l(p(x), y)$  最小化。损失函数  $l(p(x), y)$  的选择取决于拍卖平台期望定价函数  $p(x)$  所能达到的目标,如收益、社会福利等。由于收益一般是平台最重要的目标,因此一种简单的选择是直接损失函数定为负的收益函数。然而,在第二价格拍卖中,收益函数关于价格  $p(x)$  是非凸的,甚至并不连续,导致在对应的机器学习问题中,直接优化收益函数往往无法达到理想的效果。

一种可行的方法是将拍卖过程看作一个市场清算过程<sup>[4]</sup>,而不是直接优化收益。市场清算是一个经典的经济学问题,其目标是求解一个合理的价格,使得市场中的供给与需求达到平衡。在第二价格拍卖中,市场清算价格是什么任何介于最高竞价和次高竞价之间的价格。然而直接求解该价格对定价函数  $p(x)$  的指导意义并不明确。为此,我们可以先考虑与之对应的物品分配问题:

$$\begin{aligned} \max_{0 \leq s_i \leq 1, 0 \leq t \leq \lambda} & \sum_{i=1}^n y_i s_i - ct \\ \text{s.t.} & \sum_{i=1}^n s_i = t \end{aligned}$$

其中  $s_i$  是买家  $i$  获得物品的数量,  $\lambda$  是卖家可

供出售的物品数量,  $t$  是卖家实际售出的物品数量, 而  $c$  为卖家产出单位数量物品的成本。在第二价格拍卖中, 一般  $\lambda=1$ , 而在互联网广告拍卖中, 我们通常假设  $c=0$  (即展现一次广告所需的成本可以忽略不计)。而以上物品分配问题的对偶问题实际上是一个定价问题:

$$\min_{p(x)} \sum_{i=1}^n [y_i - p(x)]_+ + \lambda[p(x) - c]_+$$

在市场清算问题的视角指导下, 我们可以将该定价问题数学规划中的目标函数作为前述机器学习问题的损失函数:

$$\ell(p(x), y) = \sum_{i=1}^n [y_i - p(x)]_+ + \lambda[p(x)]_+$$

其中  $[\cdot]_+$  代表函数  $\max\{0, \cdot\}$ 。一般  $\lambda=1$ , 但是  $\lambda$  可以当作一个可以调节的参数: 适当松弛或收紧  $\lambda$  的值可以调节以上问题的最优解是更接近最高价, 还是更接近次高价; 作为一个机器学习问题的损失函数, 上式右边第二项可自然地看作是一个正则项。这个损失函数存在一个天然的优势, 即该函数是一个凸函数, 这意味着它可以在任何优化算法下快速收敛。并且, 通过调节  $\lambda$  的值来调节平台收益和社会福利的能力。实验表明, 该损失函数在以上两个指标的表现上, 相较于其他经典的损失函数 (如最高价或次高价的均方误差等) 是帕累托最优 (Pareto optimality) 的。

保留价作为简单而有效的机制, 已经在大量现实拍卖中得到广泛应用, 尤其在互联网广告拍卖领域。近年来人工智能技术的发展和互联网企业本身拥有的海量数据, 使得探索高效而智能的定价方法逐渐成为该行业的共识。

## 先验分布与广告机制设计

拍卖机制可以分为两类: (1) 先验相关拍卖 (prior-dependent auction), 即拍卖机制用到先验分布的信息, 如 Myerson 拍卖; (2) 先验无关拍卖 (prior-free auction), 即拍卖机制没有用到先验分布的任何信息。广告位拍卖机制的很多参数都是基于先验分布进行

设计。给定先验分布, 卖家就可以基于这个分布进行机制设计来满足各种性质和需求。

然而 Wilson<sup>[6]</sup> 提出拍卖机制的设计不应过度依赖先验分布, 即 Wilson's Principle。广告位买家的信息通过学习得到, 不能保证真实性和准确性。比如买家知道卖家是根据他过去 90 天的报价历史记录来学习他的先验分布 (广告位的估值分布), 那么, 买家一开始就可以按照另外一个“伪造”的分布进行报价, 让卖家认为他的先验分布是这个“伪造”的分布。在之后的时间里, 买家也会按照这个“伪造”分布进行报价。显而易见, 选择伪造分布对买家必定是有利的: 以只有一个买家为例, 他每天的报价都可以是一个接近 0 的数  $\epsilon$ , 90 天以后, 卖家就会认为买家的先验分布始终等于  $\epsilon$ 。如果仍按照收益最大化来定价, 那么定价也会在  $\epsilon$  左右。这相当于买家每天都能以接近 0 的价格购买广告位。

在买家可以伪造先验分布的情况下, 卖家收益利用先验分布信息的效果就会下降<sup>[6]</sup>。如果卖家选择 Myerson 拍卖机制<sup>[7]</sup>, 并且买家可以伪造分布, 那么卖家的收益等于卖家选择一价拍卖机制时的收益。比如有 2 个买家参与拍卖, 他们的真实先验分布都是  $[0, 1]$  均匀分布, 卖家采用 Myerson 拍卖机制, 这时买家会伪装成估值服从  $[0.25, 0.5]$  均匀分布, 此时卖家的收益就变为  $1/3$ , 与选择一价拍卖机制的收益相同。

在二价拍卖中, 如果所有买家拥有独立且相同的先验分布, 卖家采用 Myerson 保留价或者采用一些随机取样的保留价, 则买家会伪装分布从而影响保留价的设置, 导致卖家的收益不会超过不使用保留价时的收益。

这些证据表明在没有掌握买家真实信息的情况下, 买家伪造先验分布的行为可以导致先验相关拍卖失去效果, 表现不如先验分布无关的拍卖机制。

买家不仅可以伪造先验分布, 还可以伪造身份, 即在多轮拍卖中以不同的身份参与拍卖, 那么买家同样可以增加自己的收益。假设某一位买家在估值满足一定条件时, 以一个新身份参与拍卖, 那么卖家针对新身份的买家设置新的拍卖价格。伪造身份



的策略会对买家和卖家的收益产生影响,所有可能发生的收益影响已经得到刻画,并可以找到买家最优的伪造身份策略<sup>[8,9]</sup>。

## 智能实时竞价系统

广告位拍卖通常发生在每个用户请求到来时,这意味着竞价拍卖将以极高的频率重复进行,每次向用户展示一个或多个广告位,并向买家收取相应的费用。由于拍卖发生得过于频繁,买家难以对其参加的每一次拍卖分别给出竞价。买家对每轮拍卖中广告位的价值估计都有所不同,并且每轮拍卖的竞争对手也不相同,因此买家无法通过选取单个固定竞价取得多轮拍卖的最优解。另外,买家受到信息获取能力、计算能力以及广告预算等的限制,无法给出最优竞价。在此背景下,平台提供智能实时竞价服务代替买家给出参考竞价,使得广告位的分配更加满足买家的需求。与直接采用买家提供的竞价数值相比,智能实时竞价能够给买家带来更好的收益,因此以代理竞价的形式被买家接受。

在买家的界面中,平台的实时竞价机制通常被呈现为可供选择的预算管理服务。买家创建广告计划,并为广告计划设置相应的预算。预算一旦耗尽,广告计划将无法继续参与拍卖,直到买家补充新的预算。通过控制一个广告计划的拍卖竞价,平台可以控制该广告计划在拍卖中的成交概率和付费金额,从而控制该广告计划的预算消耗速率。作为智能实时竞价的实现之一,脸书(Facebook)提供的“预算使用速率调整功能”可以控制广告的预算消耗速度<sup>[10]</sup>。对限定总预算、采用“最低费用竞价策略”的广告计划(买家购买广告的计划)而言,智能实时竞价使得广告的预算在指定的周期内以有限的速率消耗。如果周期内早期广告拍卖的成交价格较高,广告计划的预算就可以避免过快耗尽,在周期内后期获得更多低价广告位的可能性就更大。

为了限制广告计划的预算消耗速率, PTR (Pass Through Rate) 模型对每个广告计划  $i$  维护数值  $p_i \in [0,1]$ , 代表该广告在竞价拍卖中的通过率。当

一个涉及该广告计划的竞价拍卖发生时,该广告以概率  $p_i$  参与此次拍卖,否则跳过此次拍卖。领英(LinkedIn)的智能实时竞价系统,使用基于历史数据的统计方法,平台在每日开始时为每个广告计划预测目标花费曲线,即理想的预算消耗量与时间的关系。以1分钟为周期,对全部的广告计划更新通过率,从而使系统快速达到稳态。初始的通过率被统一设为0.1,这一较低的通过率使得每个广告计划均有一定的机会赢得拍卖,从而给启发式算法充足的调整通过率的机会。在2周的线上测试中,对于领英的2种广告产品,引入PTR模型的智能实时竞价均能增大广告计划的生命周期(耗尽95%的预算所花费的时间)中位数,增加单位广告预算带来的平均浏览用户数量、服务的广告计划总数量和广告位的平均利润<sup>[11]</sup>。

竞价系数模型是另外一种控制预算消耗的方法。对广告计划  $i$  和给定时间段内的用户请求  $j$ ,  $v_{ij}$  表示向  $j$  展示  $i$  所能产生的价值,  $b_i$  表示每个广告计划  $i$  的预算。平台拥有全部  $v_{ij}$  和  $b_i$  的信息。与PTR模型类似,对每个广告计划  $i$ , 平台计算单个数值  $\alpha_i \in [0,1]$ 。为将用户的广告位分配给广告计划,对每个用户请求  $j$  进行一次单物品竞价拍卖,其中广告计划  $i$  的智能实时竞价为  $\alpha_i v_{ij}$ 。同样限制了广告计划  $i$  赢得广告位的概率,从而间接限制了  $i$  预算的消耗速率。当使用一价拍卖机制时,可以证明存在唯一的参数  $\{\alpha_i\}$ , 使得每个广告计划  $i$  满足以下二者之一: (1) 买家  $i$  的预算  $b_i$  恰好被给定时间段内的全部拍卖耗尽; (2)  $\alpha_i = 1$ <sup>[12]</sup>。当使用二价拍卖时,至少存在一组参数  $\{\alpha_i\}$ , 使得每个  $\alpha_i$  都能够在不超出买家  $i$  预算的前提下,将买家  $i$  的收益最大化<sup>[13]</sup>。

上述两个模型包含对平台预测能力的假设: 平台能预测每个广告计划理想的预算消耗随时间的变化,或是能够提前获知全部用户请求的广告位价值。当这一假设不成立时,平台需要应对来自未来竞价拍卖参数(即每个广告计划在未来用户请求的广告位上的价值)的不确定性,需要使用在线算法给出智能实时竞价。在线算法能给出的竞价通常是近似最优解。

通过一个简化假设,可以将广告位的在线分配问题简化为在线二分图匹配问题的一个推广:假设赢得用户请求 $j$ 的广告位的广告计划 $i$ 付出的费用恒为 $v_{ij}$ 。此时,对于顺序到来的用户请求,在线算法需要为每个请求分配一个广告计划,并从其预算中扣除相应的金额。当一个广告计划的剩余预算不足时,将无法被分配。因此,如果部分广告计划的预算被过早地耗尽,导致其无法参与后续的分配,后续的请求可能会被分配给 $v_{ij}$ 非常低的广告计划,以至于平台和买家的整体利润受损。当平台没有关于未来 $v_{ij}$ 的任何信息时,若满足 $\max_{i,j} v_{i,j}$ 相对于 $b_{i,j}$ 充分小,则一个简单的算法可以达到最优利润的 $1-1/e$ 近似:将当前用户请求 $j$ 分配给使得 $[v_{i,j} \cdot (1-e^{-(1-m_i/b_j)})]$ 最小的广告计划 $i$ ,其中 $m_i$ 是 $j$ 到来时,广告计划 $i$ 已经被花费的预算量<sup>[14]</sup>。

假设每个广告计划 $i$ 在用户请求上的价值服从固定的概率分布 $F_i$ ,且每个广告计划在每个用户请求上的价值彼此独立。当平台没有关于 $F_i$ 的全部信息时,智能实时竞价问题是在线问题:平台不断从新的用户请求中获取每个广告计划的价值,即来自每个 $F_i$ 的抽样,从而积累关于概率分布的信息,并以此为依据更新当前的智能实时竞价。特别地,当平台预先没有关于 $F_i$ 的任何信息时,全部的信息都来自在线积累。这种情境下的理论结果得出,对全体广告计划应用特定的在线竞价算法时,智能实时竞价将收敛至为每个广告计划分别设置一个 $[0,1]$ 之间的数值,广告计划对当前请求的竞价由其价值乘以该数值给出。与此同时,当买家同样预先没有关于 $F_i$ 的任何信息时,随着市场规模的扩大,任何买家通过采取其他在线算法能够获得的收益增量将趋于0<sup>[15]</sup>。这意味着智能实时竞价给出了对任何买家均近似最优的竞价调整方案。

智能实时竞价出现的根源是平台与买家关于拍卖市场的信息不对称,以及平台实时、高频的竞价拍卖超出了买家的竞价能力。为了适应广告拍卖的需求,工业界对智能实时竞价的算法与实现做了很多探索,相应的理论工作也不断出现。除预算限制以外,代表工业界需求的其他限制条件同样带来了

新的拍卖设计问题<sup>[16,17]</sup>,以及相应的智能实时竞价算法问题。

## 总结

本文从三个角度介绍了当前互联网广告的拍卖设计。保留价机制描述简单,在实际拍卖中被很多公司广泛用来增加收入。但是,如何找到最优的保留价是一个难点,最优的保留价应该和买家先验信息有关。从理论方面分析,利用买家的先验信息进行机制设计可能是无效的。由于买家进行不理性价,广告平台会提供智能实时竞价服务,按照买家的需求进行更合理的竞价。这些问题的本质就是买家和卖家之间信息不对称,以及买家的有限理性。如何更好地解决这些困难,实现买家和卖家的双赢,仍然需要进一步研究。 ■



唐平中

CCF 专业会员。清华大学副教授。主要研究方向为人工智能、多智能体系统、计算经济学、电子商务、机制设计。  
kenshinping@gmail.com



王子贺

CCF 专业会员。上海财经大学助理教授。主要研究方向为博弈论、机制设计。  
wang.zihe@mail.shufe.edu.cn



沈蔚然

卡耐基梅隆大学博士后研究员。主要研究方向为博弈论、机制设计、多智能体系统、机器学习。

其他作者: 曾驭龙 倪博楠

## 参考文献

- [1] Ostrovsky M, Schwarz M. Reserve Prices in Internet Advertising Auctions: A Field Experiment[J]. *Social Science Electronic Publishing*, 2011.
- [2] Shen W, Peng B, Liu H, et al. Reinforcement Mechanism

- Design: With Applications to Dynamic Pricing in Sponsored Search Auctions[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020.
- [3] David S, Huang A, Chris J et al. Mastering the game of Go with deep neural networks and tree search[J]. *Nature* 529, no. 7587 (2016): 484.
- [4] Shen W, Sébastien L and Renato P. Learning to Clear the Market[C]// Thirty-sixth International Conference on Machine Learning, 2019.
- [5] Medina A M and Mehryar M. Learning theory and algorithms for revenue optimization in second price auctions with reserve[C]//Proceedings of the 31st International Conference on Machine Learning (ICML-14), 2014: 262-270.
- [6] Tang P and Zeng Y. The price of prior dependence in auctions[C]// Proceedings of the 2018 ACM Conference on Economics and Computation, 2018:485-502.
- [7] Myerson R B. Optimal auction design[M]. *Mathematics of operations research*, 1981,6(1): 58-73.
- [8] Bergemann D, Benjamin B and Stephen M. The limits of price discrimination[M]. *American Economic Review*, 2015,105(4): 921-57.
- [9] Shen W, Tang P and Zeng Y. A closed-form characterization of buyer signaling schemes in monopoly pricing[C]// Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, 2018:1531-1539.
- [10] Facebook. About pacing. <https://www.facebook.com/business/help/1754368491258883>.
- [11] Deepak Agarwal, Souvik Ghosh, KaiWei, and Siyu You. Budget pacing for targeted online advertisements at linkedin. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 08 2014.
- [12] Vincent Conitzer, Christian Kroer, Debmalaya Panigrahi, Okke Schrijvers, Eric Sodomka, Nicolas Stier-Moses, and Chris Wilkens. Pacing equilibrium in  $\_rst$ -price auction markets. pages 587{587, 06 2019.
- [13] Vincent Conitzer, Christian Kroer, Eric Sodomka, and Nicolas E. Stier-Moses. Multiplicative pacing equilibria in auction markets. In *WINE*, 2018.
- [14] Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *J. ACM*, 54(5):22{es, October 2007.
- [15] Santiago Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65:3952{3968, 09 2019.
- [16] Negin Golrezaei, Ilan Lobel, and Renato Paes Leme. Auction design for roi-constrained buyers. 2018.
- [17] Benamin Heymann. Roi constrained auctions. *ArXiv*, abs/1809.08837, 2018. 3