

实时竞价在展示广告中的应用研究及进展

刘梦娟 岳威 仇笠舟 李家兴 秦志光

(电子科技大学网络与数据安全四川省重点实验室 成都 610054)

摘要 随着在线广告在产业界取得巨大成功,其在学术界特别是数据挖掘和机器学习领域的研究也吸引了大量学者的关注. 本论文围绕实时竞价机制在展示广告投放中的关键问题展开研究. 首先介绍了实时竞价的基本流程、主要参与者的功能、定价模型和交易机制;然后分别从需求方、供应方和交易中心的角度,介绍了实时竞价中存在的
关键问题,以及目前的研究方法、理论和模型,具体包括:用户响应预测、出价策略、预算与步进管理、保留价优化、库存分配、拍卖机制等,特别针对用户响应预测和出价策略两个研究热点展开了广泛讨论,并对其中的代表性方法进行了量化对比;在此基础上对主要的广告欺诈方式和检测手段进行了整理;最后对该方向未来的研究趋势进行展望.

关键词 展示广告;实时竞价;点击率预测;出价策略;广义第二价格拍卖;广告欺诈

中图法分类号 TP311 **DOI号** 10.11897/SP.J.1016.2020.01810

Research Progress of Real-Time Bidding for Display Advertising

LIU Meng-Juan YUE Wei QIU Li-Zhou LI Jia-Xing QIN Zhi-Guang

(Network and Data Security Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China, Chengdu 610054)

Abstract During the past several years, online advertising has achieved huge success in the industry. As one of the most exciting advances in online advertising, real-time bidding (RTB) has attracted a lot of attentions from both academia and industry, especially in the fields of data mining and machine learning. Different from the traditional contract-based ad delivery, in RTB the publishers (such as websites and mobile apps) can sell their ad impressions through public auctions and advertisers can evaluate the auctioned impressions and bid for them. Therefore, in RTB publishers can sell more impressions and make more money and advertisers can optimize their budget efficiency by allocating the budget to all of the available impressions based on their values. This paper focuses on the research progress of RTB for display advertising. Firstly, we briefly introduce the ad delivery process, the functions of main participants, the pricing models and auction mechanisms in RTB. Secondly, from the perspectives of demand side platform (DSP), supply side platform (SSP) and ad exchange (ADX), we detail the key problems in RTB, as well as the existing research methods, theories and models. Specially, DSP plays a critical role as the agent for advertisers, where we introduce three topics, i. e. user response prediction, bidding strategy and budget (pacing) management. Among them the first two topics are the most popular research areas. For user response prediction, there are three structures that are widely used in prediction models: the first one is shallow structure, the second one is deep neural network, and the last one is the hybrid structure. We introduce the representative models

收稿日期: 2018-12-04; 在线发布日期: 2019-11-12. 本课题得到国家自然科学基金(61202445, 61502087)、中央高校基本业务费(ZYGX2016J096)资助. 刘梦娟, 博士, 副教授, 主要研究方向为机器学习、计算广告. E-mail: mjliu@uestc.edu.cn. 岳威, 硕士研究生, 主要研究方向为机器学习、计算广告. 仇笠舟, 硕士研究生, 主要研究方向为机器学习、计算广告. 李家兴, 硕士研究生, 主要研究方向为机器学习、计算广告. 秦志光, 教授, 博士生导师, 主要研究领域为网络安全、机器学习.

based on the above structures and evaluate their prediction performance based on a real-world dataset. Bidding strategy is used to decide the bid price for an impression on behalf of the advertiser. In this paper, we introduce the widely used linear methods and the latest nonlinear and reinforcement learning based methods. Empirically, we compare the characteristics and performance of some typical bidding strategies. The goal of SSP is to provide services for publishers. Besides reserve price optimization, a new task has emerged, that is, how to dynamically allocate ad inventory between the offline and online (i. e. contract-based and RTB-based) sales channels to maximize the publishers' total profits. Briefly, we introduce the latest research progress on this issue. Auction mechanism is one of the most important components in ADX and the sealed-bid second price auction is the most widely used auction mechanism in RTB systems. In this paper, we introduce three typical second price auctions, i. e. Vickrey auction, generalized second price (GSP) auction and Vickrey-Clarke-Groves (VCG) auction, where GSP is adopted by many famous RTB systems, such as Google and Baidu. However, second price auction has been questioned by publishers and advertisers because it provides spaces for ADX cheating. So we supplement the latest progress on the first price auctions and header auctions. On the other hand, ad fraud has become an inevitable problem in RTB, so we summarize the main advertising fraud types and detection methods in this paper. Finally, the future research trend of this direction is prospected.

Keywords display advertising; real-time bidding; click-through rate prediction; bidding strategy; generalized second price; ad fraud

1 引言

近十年, 在线广告(online advertising)获得了巨大的成功, 已经成为最主要的广告投放渠道. 与传统的电视、报纸等投放渠道相比, 在线广告不仅为互联网媒体提供商找到了规模化变现的手段, 而且提供了一种以准确接触目标受众为方法论的全新广告营销渠道^[1]. 作为大数据技术形成规模化营收的重要落地应用, 在线广告不仅在产业界取得了巨大成功, 同时也在学术界特别是数据挖掘和机器学习领域吸引了大量研究人员的关注, 自 2007 年 SIGKDD 开始同时举办面向在线广告的数据挖掘技术讨论会(ADKDD), 同时 WWW、ICDM 等会议上也出现了大量与该主题相关的论文. 目前在线广告相关技术已经发展成一门跨越信息检索、数据挖掘、机器学习、推荐系统和交易理论等多领域的应用型交叉学科.

在线广告分为赞助搜索(sponsored search)^[2]和展示广告(display advertising)^[3]两大类. 其中赞助搜索^[4-6]主要针对搜索平台, 由广告主根据潜在用户的特点设置竞标的关键词及价格, 每当有用户输

入搜索关键词时, 平台根据用户搜索的关键词与广告主竞标的关键词进行相关性匹配, 并结合竞标价格决定最终的广告显示排位, 排位最高的若干广告将与搜索结果一起展示给用户. 目前主流的搜索平台和电商平台都提供赞助搜索广告. 展示广告通常是以图像或者视频的形式展示在网站提供的广告位上, 当用户浏览网页或打开 app 时, 广告将被自动加载或播放. 展示广告的主要投放方式有两种: 担保模式(Guaranteed Delivery, GD)和实时竞价模式(Real Time Biding, RTB). 在 GD 中, 广告主通过签订私有合同的方式向网站购买指定广告位的广告展示机会(ad impression), 通常按照展示次数或展示时间结算. 在采用展示时间作为结算依据的模式中, 通常按照展示天数计费(Cost-Per-Day, CPD), 即广告主购买特定广告位的投放时段. 这种模式的优点是支持长期在特定位置进行广告投放, 能充分发挥橱窗效应, 其缺点是对所有浏览包含该广告位的页面的用户均投放广告, 不具有针对性. 在采用展示次数作为结算依据的模式中, 网站保证按照广告主的投放规则(target rules)在指定广告位向用户展示广告的次数, 因此该模式可以支持有针对性的定向广告投放. 通常来说, 当广告主以进行品牌宣传扩大影响力

为目标时,多采用 GD 模式投放广告。

RTB 模式^[7]则是网站将广告位对应的广告展示机会通过实时竞价系统进行公开拍卖,广告主通过竞标获得广告展示机会。RTB 模式支持广告主对每个广告展示机会进行评估并出价,因此能够帮助广告主优化其预算使用效率。早期展示广告的投放以担保模式为主,但是随着实时竞价模式近年来的飞速发展,通过 RTB 投放的展示广告量也已经达到了非常庞大的规模。移动广告平台 Fikisu 宣称其需求方平台每天处理的广告展示机会数(ad impressions)达到了 32 billion^[8];著名在线广告公司 Turn 的需求方平台在峰值时间大约每秒钟需要处理的展示机会数达到了 2.5 million^[9];国内 iPinYou 公司的需求方平台每天处理的展示机会数达到了 18 billion^[10]。

与传统的合约式定向投放相比,RTB 具有如下优势:对于广告发布媒体,可以提升广告库存的销售比例,同时拍卖竞价机制使得广告库存的销售也更加公平,提升资源配置效率;对于广告主,可以实现精准的受众定向,而且允许广告主针对每个广告展示机会动态出价,使得预算可以集中花费到那些更能产生收益的曝光机会上,优化了广告主的预算使用效率;最后,对于用户,RTB 通过用户响应行为预测支持更为细粒度地精准投放,使得广告与用户兴趣(需求)更为匹配,从而提升了用户体验。然而遗憾的是,目前在线广告的研究工作主要集中在赞助搜索和基于担保模式的定向广告方面,作为在线广告最新的投放方式,RTB 的相关研究仍处于起步阶段。研究主要集中在用户响应行为预测、出价策略等几个关键机制上,除了文献[3,11-12],很少有论文对整个 RTB 的交易流程和关键问题的最新进展进

行梳理。然而近年来随着深度神经网络、强化学习、Header 拍卖等新技术的出现,RTB 中的若干关键问题,例如用户响应预测、出价策略、拍卖机制等,出现了新的解决思路 and 方案。

为了帮助该领域的研究者和算法工程师更好地理解目前 RTB 的核心问题和解决方案,特别是在工业界的应用,本论文对 RTB 的最新研究进行了调研,并分别从需求方平台、提供方平台、交易中心三个主要参与平台的角度,梳理其中的关键问题,以及目前已有的研究方法、理论和模型,特别是对目前热度最高的用户响应预测及出价策略的最新研究进行了详细地介绍和分析。另一方面,随着在线广告产业规模的不断扩大,欺诈活动已经成为其中不可避免的问题,特别是在 RTB 中,广告主与广告展示机会的提供方不直接交易,因此更容易滋生展示机会欺诈、点击欺诈、谎报展示机会价值等欺诈行为,因此本文整理了目前主要的广告欺诈方式和检测手段作为关键问题介绍的补充。最后论文对 RTB 在展示广告投放中的应用前景和研究趋势进行了展望。

2 基于实时竞价的展示广告投放系统

RTB 的基本竞价过程如图 1 所示^[3]:(1)当用户浏览发布媒体网页时,网页上嵌入的广告位的脚本代码会向实时竞价系统中的提供方平台(Service Side Platform, SSP)发起一次广告展示机会的请求;(2)SSP 收到这个广告展示机会的请求时,会针对该请求向 RTB 的交易中心(AD eXchange, ADX)发起一次广告展示机会的竞价请求,该请求携带了用户的 cookie 和网页的上下文信息;(3)ADX 收到

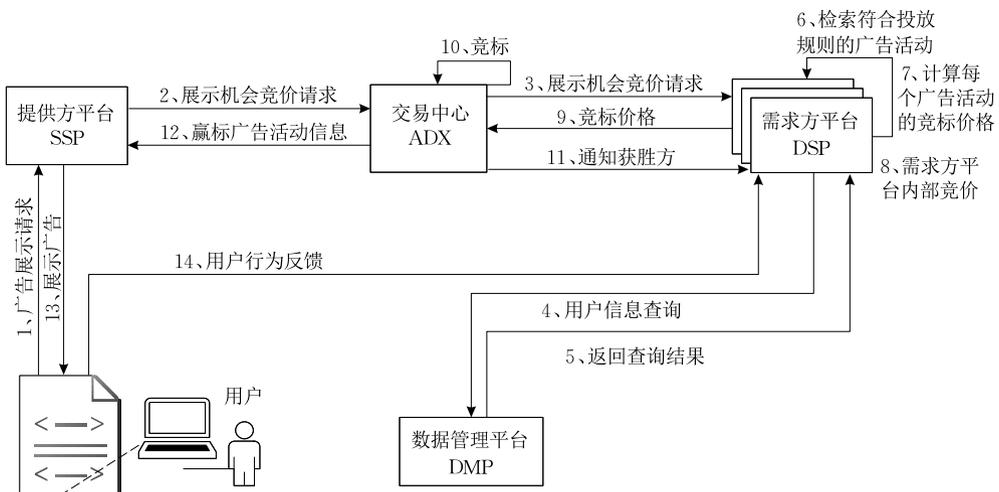


图 1 RTB 的基本竞价过程

竞价请求后,将竞价请求向连入 RTB 的需求方平台(Demand Side Platform,DSP)发布;(4) DSP 收到广告展示机会的竞价请求后,首先通过 cookie 映射技术以及查询相关的数据管理平台(Data Management Platform,DMP),获取用户的特征标签,例如性别、年龄、购物兴趣等;(5) DMP 将该用户的特征标签返回 DSP;(6) DSP 根据每个广告活动的投放规则,检索该次广告展示机会符合的广告活动;(7) DSP 根据每个广告活动的预算,为每个符合投放规则的广告活动计算竞标价格;(8) DSP 首先在平台内部进行一轮竞价,最高竞标价格的广告活动赢得平台内部的竞拍;(9) 最高竞标价格将作为 DSP 的竞标价格返回给 ADX;(10) ADX 对收到的所有 DSP 返回的竞标价格,按照广义第二高价机制(Generalized Second Price,GSP)进行拍卖^[12],即由最高竞标价格的广告活动赢得展示机会,但是实际成交价格是由该次竞拍中的第二高竞标价格决定;(11) ADX 通知获胜方,并按照实际成交价格进行扣费;(12) ADX 将获胜的广告活动的链接返回给 SSP;(13) 由 SSP 将广告链接返回给用户浏览的页面;(14) 通常 RTB 会跟踪用户对投放的广告的响应行为,将点击或者转化行为向 DSP 报告,以便于 DSP 进一步优化用户响应预测模型和出价方法.在实际应用系统中,上述竞价过程通常要求在 10~100 ms 内完成.

如图 1 所示,RTB 系统的主要参与者包括发布媒体(publisher)、广告主(advertiser)和四个核心平台 SSP、ADX、DSP、DMP.其中发布媒体是提供广告展示机会的在线网站,每当用户访问这些网站的网页时,网页上嵌入的脚本代码向 RTB 的 SSP 发起一次广告展示机会(曝光机会)的竞价请求,即发布媒体将自己的广告展示机会托管给 RTB 进行售卖;广告主在 DSP 上注册广告营销活动,同时设置这次广告活动的投放周期、创意、预算和定向投放规则,即广告主是广告展示机会的购买者;SSP 的主要功能是登记发布媒体的广告库存(ad inventory),然后向接入的 ADX 发起拍卖请求,因此 SSP 是服务于发布媒体的,目标是最大化发布媒体的收益;ADX 是 RTB 的交易中心,每当收到拍卖请求后,会向接入 ADX 的所有 DSP 发送出价请求,并根据各个广告活动的出价和拍卖机制决定最终的获胜者;DSP 是广告主代理平台,帮助管理广告主在平台上注册的各种广告活动进行定向投放,主要功能包括用户响应预测、出价策略、预算和步进管理等,因此 DSP 是服务于广告商的,目标是最大化广告主的收

益;DMP 作为数据管理平台主要为 DSP、SSP、ADX 实时提供用户的历史行为数据.

RTB 系统中最早采用的拍卖机制是第一价格密封拍卖,例如 Open Text 和 Goto.com.但是自 2000 年,Google AdWords 推出广义第二高价拍卖(Generalized Second Price,GSP)机制以来,GSP 目前已经被绝大多数的 RTB 系统采用^[13].RTB 的计费模式包括 CPM(Cost-Per-Mile)、CPC(Cost-Per-Click)、CPA(Cost-Per-Action)、dCPM(dynamic CPM)等.其中 CPM 表示每千次展示的价格,CPM 计费表示广告主需要为每次广告展示付费;CPC 表示每次点击的价格,CPC 计费表示只有当广告被展示且被用户点击后才需要广告主付费;CPA 表示每次转化的价格,CPA 计费表示只有当用户点击广告且成功发生转化才需要广告主付费,这里转化通常表示点击广告后的进一步行为,例如下载软件、完成注册、购买商品等;dCPM 计费表示广告主需要为广告创意的每次展示付费,费用根据每次拍卖的情况动态变化.目前大多数的 RTB 系统中都采用的是 dCPM 计费模式.

3 需求方平台的研究进展

目前针对 RTB 的研究中最为重要、活跃的研究点主要集中在需求方平台,具体包括用户响应预测、出价策略、预算及步进管理等.在详细介绍之前,首先对 DSP 在实时竞价过程中的基本工作流程进行描述.一个广告商如果希望通过 RTB 系统来进行广告投放,需要在 DSP 上进行注册,并发起一个给定预算的广告活动,同时设置其广告的定向投放规则,例如广告目标用户的年龄范围、性别,投放页面的类型等,广告活动应只对满足其投放规则的展示机会进行竞价.因此,当一个广告展示机会的竞价请求到达某 DSP 时,DSP 首先筛选出广告展示机会符合投放规则的广告活动的集合,如图 2 所示,然后为每个广告活动确定其竞标价格,DSP 将最高竞标价格反馈给 ADX 作为本平台的出价.该价格将会与 ADX 收到的来自其他 DSP 的出价再次进行竞争(通常是第二价格机制),若该出价仍为最高价格,则其对应的广告活动赢得广告展示机会,并支付相应的费用.因此 DSP 的基本功能是,在预算范围内针对符合定向规则的展示机会为每个广告活动计算出最优的竞标价格,以帮助广告商实现广告收益(利润)最大化等目标.

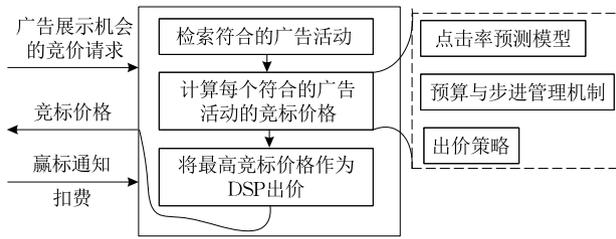


图 2 DSP 的工作流程及关键问题

每个广告活动的竞标价格由 DSP 的出价策略决定。目前主要的出价策略包括：固定出价以及基于关键性能指标的出价^[14-16]。这里关键性能指标通常是指广告投放到展示机会后，用户产生点击（转化）行为的概率，一般认为点击率（转化率）越高，广告展示机会对于广告活动的价值越高，因此出价也应该更高。文献[14]通过实验证明，基于关键性能指标的出价策略明显优于固定出价，是目前研究和应用最为广泛的出价策略。基于关键性能指标出价的难点在于关键性能指标的计算，目前通常采用预测点击率(predicted Click-Through Rate, pCTR)或预测转化率(predicted Conversion Rate, pCVR)作为关键性能指标。此外，从广告投放效果的角度来看，广告商更希望在整个投放周期均匀平滑地花费预算，而不是将预算集中花费到某几个集中的时段上，为此，需要在出价策略的基础上增加对每个时段预算和广告投放频率的管理机制，以提升广告预算的使用效率。本节将分别围绕用户响应预测、出价策略、预算与步进管理这三个关键问题展开介绍。

3.1 用户响应预测

作为需求方平台最为关键的问题，用户响应预测的研究吸引了大量来自工业界和学术界的关注^[17-22]。用户响应是指广告被展示给用户后，用户产生的反馈行为。如果用户对广告内容感兴趣，会点击广告进入广告页面，即产生点击行为；进一步地，如果用户在广告页面进行了购买、下载等广告主预定义的操作，即产生转化行为。通常将广告点击次数与展示次数的比值定义为点击率(Click-Through Rate, CTR)，将转化次数与点击次数的比值定义为转化率(Conversion Rate, CVR)。转化率能够更为准确地衡量广告展示机会对于广告活动的价值，然而建立准确的转化率预测模型是非常困难的，这是因为转化行为的发生频率远远低于点击行为的发生频率，并且转化与点击之间可能会有较长的延迟(数天甚至数周)^[23]。因此，无论是工业界还是学术界都更多地将预测点击率(pCTR)作为 RTB 中出价的依

据，本文也主要围绕 CTR 预测方案展开介绍，CVR 的预测方案^[4, 24-25]与 CTR 预测方案是相似的。

3.1.1 问题描述

CTR 预测是指对广告投放到一个广告展示机会后用户点击该广告的概率进行预测，可形式化描述为一个以实现二分类(点击/不点击)为目的的回归问题，通常利用有监督的机器学习模型来解决。首先利用历史数据作为训练样本集为每个广告活动学习一个 CTR 预测模型，对于新到达的广告展示机会，将其特征向量输入每个广告活动的预测模型中，从而得到广告投放到展示机会的预测 CTR。假设训练集为 $T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$ ，其中广告展示机会的特征向量记为 $\mathbf{x}_i (i = 1, \dots, N)$ ， y_i 是广告是否被点击的标签， $y_i = 1$ 表示点击， $y_i = 0$ 表示没有点击；CTR 预测模型为 $Y = f(\mathbf{x})$ ， θ 是函数 $f(\mathbf{x})$ 的参数向量； $L(\cdot)$ 为损失函数，用于衡量预测模型的错误程度，则 CTR 预测模型可通过式(1)学习，即学习在训练集上使损失函数最小的最优模型参数 θ 。

$$\min_{\theta} L(f(\mathbf{x}_i), y_i) \quad (1)$$

其中 $f(\mathbf{x}) = P(y=1|\mathbf{x})$ ， $1 - f(\mathbf{x}) = P(y=0|\mathbf{x})$ 。

在有监督学习中，需要确定 $f(\mathbf{x})$ 和 $L(\cdot)$ 的形式，然后利用优化算法求解模型参数 θ 。在已有的 CTR 预测方案中， $f(\mathbf{x})$ 有多种形式，既可以采用浅层模型，也可以采用深度学习模型，或者两者的融合模型；而损失函数 $L(\cdot)$ 通常采用对数似然损失函数(Log-Loss)，如式(2)所示；求解最优模型参数的算法通常采用梯度下降(Gradient Descent)^[26]的相关方法。值得注意的是， $f(\mathbf{x})$ 的形式对模型的预测性能有较大的影响，尤其在 RTB 中存在大量分类特征的场景下，学习能力强的模型结构往往可以产生更好的结果。因此，目前 CTR 预测方案的研究主要集中在 $f(\mathbf{x})$ 的设计上，本文根据 $f(\mathbf{x})$ 的结构差异将已有方案分为三类：基于浅层模型方案、基于深度学习模型方案和基于融合模型方案。

$$L(\theta) = -\frac{1}{N} \sum_{i=1}^N [y_i \log f(\mathbf{x}_i) + (1 - y_i) \log(1 - f(\mathbf{x}_i))] \quad (2)$$

与传统预测任务相比，CTR 预测存在如下挑战：(1) 广告展示机会的特征中有大量的分类特征，这些离散特征值不能直接输入到机器学习模型中，需要使用独热编码(one-hot)将其转化为二值(0/1)特征向量后再输入，由于在样本集中一个分类特征通常包含大量特征值，并且一个样本中每个分类特

征只取一个特征值,因此独热编码后的特征向量是一个高维稀疏的二值向量^[27],它不仅增加了模型的参数复杂度,而且会降低模型对特征的学习能力;(2)RTB 日志中,真实点击率非常低,例如 iPinYou 公开数据集^①中,4 个广告活动的平均点击率均小于 0.001^[27],这就意味着原始的训练样本集中存在严重的正负样本比例不平衡问题,导致模型对正样本学习不充分,容易产生过拟合。

3.1.2 基于浅层模型的方案

早期 CTR 预测任务通常使用逻辑回归(Logistic Regression, LR)模型^[17]来完成。LR 是一种广义线性模型,易于实现大规模实时并行处理,因此在工业界获得了广泛应用。然而 LR 在实际应用中的预测性能仍有较大的提升空间,这是因为 LR 本身的学习能力有限,只能学习一阶特征携带的信息,且对于稀疏特征的学习效果较差。为此,文献[28]在基本 LR 的基础上增加了一种自动的特征筛选方法;文献[29]则将自然语言处理中的 Neural Word Embedding 的思想引入到 CTR 预测中,通过将每个特征转化为一个嵌入向量,将样本中包含的特征的嵌入向量拼接作为 LR 模型的输入,从而有效解决了原有输入特征向量高维稀疏的问题,提升了模型对稀疏特征的学习能力。

为了增加 CTR 预测模型能利用的特征信息,直观的做法是在输入特征中同时包含一阶和二阶特征^[14],如图 3 所示。然而学习二阶特征(任意特征两两组合)的权重将导致极大的计算复杂度,因此通常采用 Feature Engineering 技术来辅助实现特征组合(Feature combination)。为了降低二阶组合特征权重学习的复杂度,文献[18]提出了因子分解机模型(Factorization Machine, FM)。在 FM 中,作者利用矩阵分解技术^[30]来学习每个一阶特征的隐含向量,然后利用两个组合特征的隐含向量进行内积来计算二阶组合特征的权重。FM 非常适合于点击率预测这样的稀疏数据集,由于每个一阶特征都可以

表示为一个固定维度的隐含向量,因此当测试集中出现了训练集中没有的组合特征时,仍然可以通过隐含向量内积来计算它们之间的关系,因此能够有效克服输入特征高维稀疏的问题。

FM 是第一个结合一阶和二阶特征进行点击率预测的高效模型。在它被提出后,受到了来自工业界和学术界的广泛关注。文献[19]的作者在 FM 的基础上引入特征域(field)的概念,提出将独热编码后的特征按照其语义属性划分为若干个 fields,例如 Male 和 Female 特征属于同一个性别 field,然后每个特征将针对每个 field 学习单独的隐含向量,在两个特征组合为二阶特征时,使用组合的特征对应的 fields 对应的隐含向量的内积来计算权重。这个改进后的模型被称为 FFM(Field-aware Factorization Machine)。实验证明 FFM 能够取得非常优异的预测性能,但是需要学习的隐含向量数量大幅增加,导致在实际生产系统中难以部署。为了减少需要学习的隐含向量数量,又体现不同特征域特征两两组合的权重差异,文献[31]引入了一个关于特征域对的权重矩阵,用以捕捉特征域之间的交互关系,即每个特征只需要学习一个隐含向量,两两特征组合时,只需要在隐含向量内积的基础上乘以两个特征所属的特征域对的权重。文献[32]的思想与文献[31]相似,但是它通过 Attention 机制来学习不同特征域组合的权重,这种机制被称为基于注意力的因子分解机模型(Attentional Factorization Machine, AFM)。

图 3 展示了考虑一阶和二阶组合特征进行 CTR 预测的通用模型结构。其共同特点是都只考虑了低阶特征,因此将其称为浅层结构模型。这类基于浅层结构的 CTR 模型,除了上述典型方案,还包括采用梯度提升决策树(Gradient Boost Decision Tree, GBDT)与 LR 级联的预测方案(GBDT+LR)^[33]以及线性模型的预测方案^[34-35]、有序回归模型的预测方案^[36]等。此外,还有一类基于用户行为协同的方案^[37-38]。这类基于矩阵分解和张量分解的方案,首先被广泛应用在推荐系统中,取得了极大的成功^[39]。矩阵分解的基本思想是首先将已知的用户-物品评分矩阵分解为用户的隐语义矩阵和物品的隐语义矩阵,然后通过这两个矩阵相乘可以得到一个新的用户-物品评分矩阵,在这个新矩阵中,原评分矩阵中缺失的项被填充,这个值即可作为用户对物品的评价。

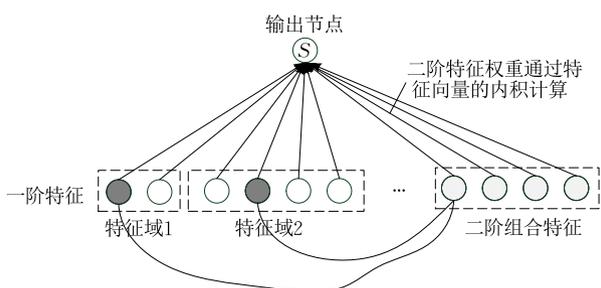


图 3 考虑一阶和二阶组合特征模型的示意图

① iPinYou 公开数据集. <http://data.computational-advertising.org/>

分预测. 矩阵分解通常是针对二元关系, 例如用户与广告、网页与广告等, 为此研究者进一步在矩阵分解的基础上提出可以包含“用户-上下文-广告”三元关系的张量分解方案^[40]. 矩阵分解和张量分解方案由于需要用户、上下文和广告的具体信息, 因此通常只适合于对那些发布媒体、交易中心、提供方平台一体的或信息相互透明的平台.

综上所述, 基于浅层结构的 CTR 预测模型简单、有一定效果, 因此在工业界得到了大规模的部署应用; 但是由于现实应用中点击事件非常稀疏, 导致预测性能仍然有较大的提升空间, 因此最新的研究集中在设计具有更强特征探索能力的模型结构上, 以提升预测性能.

3.1.3 基于深度神经网络的方案

近年来, 深度学习 (deep learning) 的研究工作取得了极大的进展^[41-43], 特别是在探索高阶特征组合方面的能力受到了大量的关注. 目前也有研究人员将深度神经网络 (Deep Neural Network, DNN) 引入到 CTR 预测模型中. 较早基于 DNN 的 CTR 预测方案是文献^[20]提出的基于因子分解机的神经网络模型 FNN (Factorization machine supported Neural Network), 如图 4 所示. 在 FNN 中, 首先利用 FM 模型进行预训练得到每个特征的嵌入向量, 然后将样本中包含的特征对应的嵌入向量拼接 (concat) 起来, 输入到一个包含若干隐层的前馈神经网络中, 最后利用 sigmoid^[44] 函数计算预测 CTR. FNN 的结

构比较简单, 只输入了一阶特征的信息, 文献^[21]通过在 FNN 的嵌入层中引入二阶特征组合的权重节点来提升输入隐层的信息含量, 这种方案被称为 PNN (Product-based Neural Network).

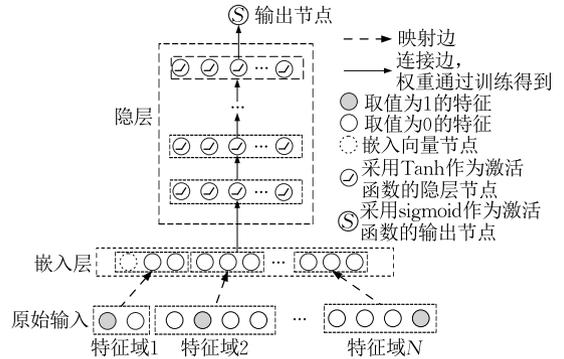


图 4 FNN 模型

文献^[45]在 FNN 的基础上, 提出一种针对电子商城场景的 CTR 预测模型——深度兴趣网络 (Deep Interest Network, DIN). 其特点是样本的特征向量不仅包括广告展示机会和广告创意的特征信息, 还包括用户的历史行为信息. 然而用户的历史行为通常是十分广泛的, 例如可以购买食品、衣服、书籍等多种商品, 这种多值特征不适合使用独热编码, 因此文献^[45]采用了如式 (3) 所示的 multi-hot 编码. 例子中用户购买了爱他美奶粉和机器学习书籍, 因此商品特征编码后有 2 个元素为 1, 其余元素为 0; 进一步地看, 对于多值特征向量作者提出加权求和的池化 (SUM Pooling) 方法, 如图 5 所示. 然后针对

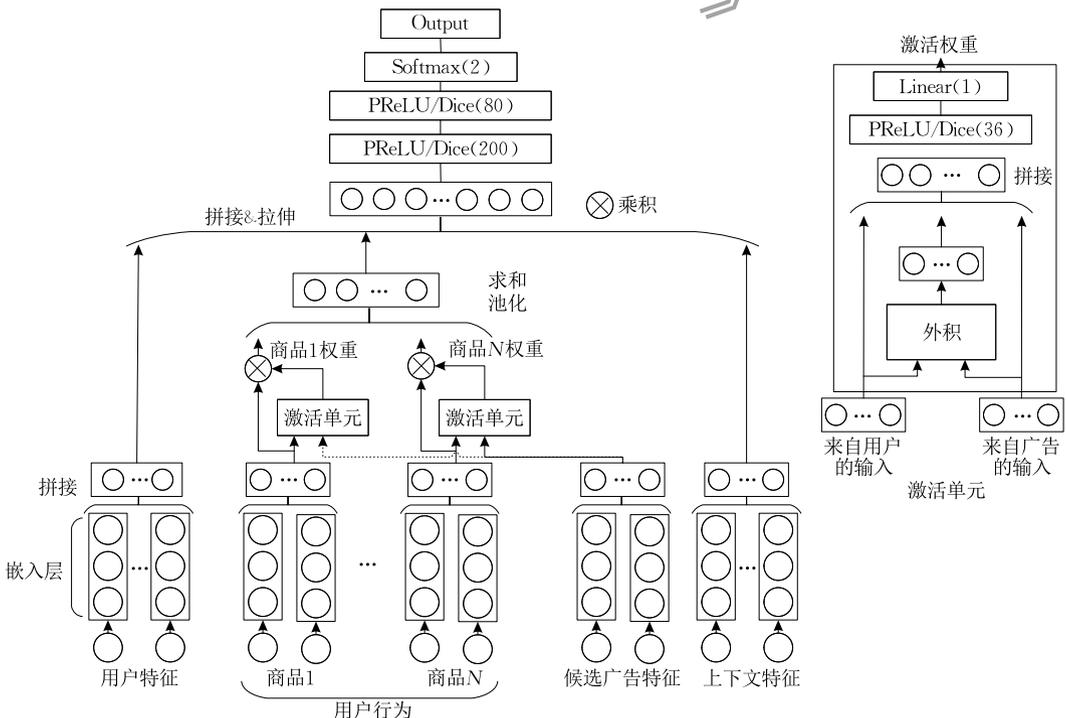


图 5 深度兴趣网络 DIN

用户行为与候选广告的关联性差异,提出用户兴趣的特征向量应该随着候选广告的不同而变化,为此引入注意力机制,提出了权重激活单元.这里的嵌入向量是随机初始化的,嵌入向量作为模型参数的一部分,在模型训练过程中不断更新. DIN 模型为多值特征在深度神经网络中的应用提供了一种可行的解决思路. 综上所述,基于单一深度学习模型方案由于能探索高阶组合特征的信息,因此往往获得优于浅层模型(LR, FM)的预测性能.

$$[0, \dots, 1, \dots, 1, \dots, 0] \quad (3)$$

goods_id = {爱他美奶粉, 机器学习书籍}

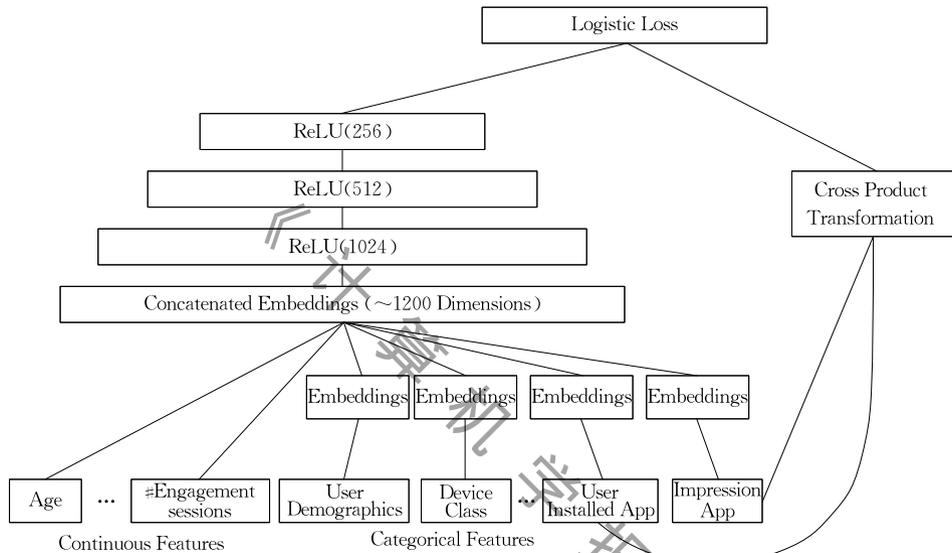


图 6 Wide&Deep 模型

Wide&Deep 虽然取得了令人满意的预测性能,但它仍然依赖于手动的特征工程. 为了实现低阶特征的自动组合,文献[44]的作者提出了 DeepFM 模型,其与 Wide&Deep 的差异主要在将模型中的 wide 部分替换为 FM 模型. 另一种基于 Wide&Deep 的改进是 Deep&Cross 模型^[46],它通过将 wide 部分替换为 Cross 网络,以实现自动、显式、有限度的特征交叉,从而避免特征过度组合导致的预测性能下降以及计算开销增大. 文献[47]对 Deep&Cross 模型进行了改进,作者首先指出 Cross 网络的输出受限于一中特定的形式,即每个隐藏层的输出都是输入的嵌入向量的标量倍数;并且特征组合是在比特级别(bit-wise level)上进行的,也即同一特征域的嵌入向量中的不同元素也会相互组合. 为此作者提出一种新的压缩交互网络(Compressed Interaction Network, CIN)来代替 Cross 网络,它以显式的方式在向量级别上进行特征组合;作者进一步将 CIN 和 DNN 组合,得到一个新的融合模型称为 xDeepFM.

3.1.4 基于融合模型的方案

基于 DNN 的 CTR 预测方案主要利用 DNN 对高阶组合特征的挖掘能力,然而 Google 的研究人员通过实验发现在 CTR 预测中,低阶特征也是非常重要的^[22]. 为此,他们提出了一种能够融合低阶和高阶特征的融合结构^[22],如图 6 所示. 融合结构的左边(deep 部分)是一个包含 3 个隐层的 DNN,右边(wide 部分)是通过手动选择的原始特征进行“叉积”后得到的低阶组合特征. 这种同时考虑了低阶和高阶特征的模型被称为 Wide&Deep.

xDeepFM 不需要特征工程,既可以显式地学习一定程度的特征交互,也可以隐式地学习低阶和高阶特征的交互.

3.1.5 典型方案的特点对比

表 1 分别从模型结构、特征探索能力、特征输入形式、是否需要手动的特征工程这几个方面对已有的典型方案进行总结和对比. 可以发现 CTR 预测的研究热点主要集中在特征预处理和模型设计两个方面. 对于 CTR 预测,训练样本的特征包括数值特征和分类特征(以分类特征为主),而分类特征是不能直接用于回归预测的,因此采用独热编码(one-hot)对原始特征进行处理,转化为高维稀疏的二值(0/1)特征向量. 对于浅层模型,参数复杂度较低(FM 除外),通常将高维稀疏特征向量直接输入到模型中;而对于深度学习模型,如果将高维稀疏特征向量直接输入到 DNN 中将使得参数复杂度非常高,为此需要将每个特征映射为一个固定维度的稠密实数值的嵌入向量,只需要将样本中出现特征的嵌入向量

表 1 典型 CTR 预测方案的特点

	模型结构	特征学习能力	特征输入形式	手动特征工程
LR	LR	1 阶	one-hot	需要
FM	FM	1 阶+2 阶	one-hot	不需要
FFM	FFM	1 阶+2 阶	one-hot	不需要
AFM	FM+Attention	1 阶+2 阶	one-hot	不需要
GBDT+LR	GBDT+LR	低阶+高阶	one-hot	不需要
FNN	DNN	高阶	嵌入向量拼接(FM 预训练)	不需要
PNN	DNN(Product layer)	高阶	嵌入向量拼接(FM 预训练)	不需要
DIN	DNN(Attention)	高阶	嵌入向量拼接(联合训练)	不需要
Wide&Deep	LR+DNN	低阶+高阶	嵌入向量拼接(联合训练)+one-hot	需要
DeepFM	FM+DNN	低阶+高阶	嵌入向量拼接(联合训练)+one-hot	不需要
Deep&Cross	Cross+DNN	低阶+高阶	嵌入向量拼接(联合训练)	不需要
xDeepFM	CIN+DNN	低阶+高阶	嵌入向量拼接(联合训练)+one-hot	不需要

拼接起来作为 DNN 的输入. 近年来, 也有研究者尝试将嵌入向量拼接作为输入的方式引入到浅层模型中, 例如文献[29]. 进一步来看, 特征嵌入向量的学习方式又分为两种: 一种是通过 FM 提前预训练; 另一种是随机初始化, 与模型参数一起通过联合学习最优化.

模型设计作为 CTR 预测中最重要的研究点, 一直是研究者和广告算法工程师关注的焦点, 从早期的浅层模型到深度学习模型, 到目前浅层与深度的融合模型. 可以发现引入不同模型的目的是为了对样本的特征及特征间的交互(feature interactions)信息进行学习, 模型对特征组合的学习能力越强, 最终的预测性能越好. LR 只能学习一阶特征的信息, 因此通常需要通过手动的特征工程来帮助提取有效的组合特征; FM、FFM、AFM 模型都是为了学习二阶组合特征的信息, GBDT+LR 利用 GBDT 来筛选有效特征及组合特征; FNN、PNN、DIN 是利用深度学习来学习高阶组合特征的信息; Wide&Deep、DeepFM 利用融合模型同时学习低阶和高阶特征携带的信息; Deep&Cross、xDeepFM 不仅同时学习低阶和高阶特征的信息, 而且通过新的结构学习按向量组合的高阶特征信息. 因此, 本文将已有的 CTR 预测模型归纳为如图 7 所示的结构, 即目前的研究均是通过模型结构的设计, 实现更为高效的自动特征工程. 最后, 图 8 展示了典型 CTR 预测方案在 Criteo 公开数据集^①上的性能对比, 这里采用 AUC 作为评价指标, AUC 值越大说明预测性能越好, W&D 表示 Wide&Deep 模型, D&C 表示 Deep&Cross 模型. 实验结果显示在单一结构下, 深度学习模型的预测性能确实优于浅层结构的 LR 和 FM 方案, 但是弱于 FFM, 这是因为 FFM 能够学习不同特征域的特征进行二阶特征组合时的嵌入向量差异, 具有较强的学习能力, 但是

FFM 的参数复杂度也非常高; 基于融合结构的方案的预测性能确实略优于单一结构模型的性能, 例如 Wide&Deep 和 DeepFM. 需要说明的是为了便于比较, 所有包含深度神经网络的方案中, 都采用了相同的神经网络结构(隐层数和隐层节点数相同), 可能对于某些方案得到的结果是非最优的, 关于点击率预测更详细的量化分析可参考前期工作^[27,48].

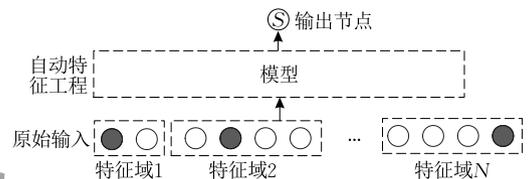


图 7 CTR 预测模型结构

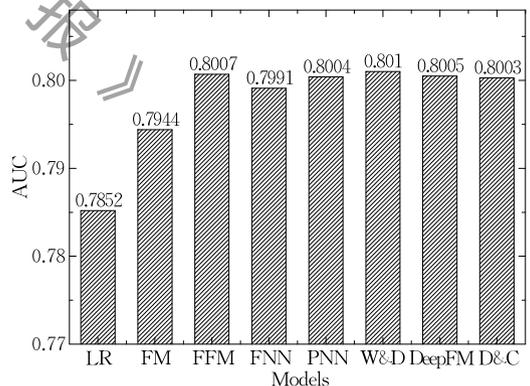


图 8 部分典型 CTR 预测方案的性能对比

3.2 出价策略

在 RTB 拍卖中, 需求方平台需要为广告展示机会符合投放规则的每个广告活动计算竞标价格, 由于竞标价格关系到广告活动是否能赢得展示, 以及广告预算的使用效率, 因此出价策略是需求方平台的核心功能. 在第二价格拍卖中, 已经证明每个竞标者按照拍卖物品对自己的真实价值出价, 将获得最

① Criteo 公开数据集. <http://labs.criteo.com/2014/02/download-kaggle-display-advertising-challenge-dataset/>

优的收益^[3],即按照广告展示机会 i 对广告活动 j 的真实价值出价,如式(4)所示,这里 $eCPI$ 表示期望的展示价值, CTR 表示点击率, CPC 表示点击价值,通常假设对于同一个广告活动,点击产生的价值是相同的,因此出价应该正比于该广告展示机会的期望点击率. 线性出价策略即是根据这一思想进行设计.

$$eCPI = CTR \times CPC \quad (4)$$

然而在 RTB 场景下,由于在需求方和供应方都存在约束,导致式(4)得到的出价方案并非最优的,例如每个广告活动都有预算和生存周期的限制,而每天到达需求方平台的广告展示机会数量也是有限的. 目前研究者已经针对不同优化目标和应用场景提出了各种出价策略. 本节对其中的典型策略展开介绍,特别分析了强化学习(Reinforcement learning)在 RTB 出价中的应用,然后对比分析了典型出价策略的特点,最后介绍预测市场价格分布的竞价愿景(bid landscape)问题.

3.2.1 问题描述

在出价策略的研究中,通常将出价描述为一个约束条件下的优化问题,如式(5)所示,然后利用最优化理论求出最优解或近似最优解. 这里 KPI (Key Performance Indicator)通常从广告商的角度定义为收益(revenue)或利润(profit),其中 v_{ij} 表示广告活动 j 将广告创意投放到广告展示机会 i 后获得的价值, c_{ij} 表示广告活动 j 赢得展示后付出的费用,在第二高价拍卖中, c_{ij} 通常是第二高价或者保留价, x_{ij} 表示广告活动 j 是否赢得展示机会 i ,如果出价是整个拍卖中的最高出价且大于等于保留价,则 $x_{ij} = 1$,否则 $x_{ij} = 0$, I 表示满足广告活动 j 的广告展示机会集合. 约束条件表示在广告活动 j 的整个投放周期的花费不能超过其预算. 除了考虑广告商的收益(利润)最大化,目前的最新研究也有从需求方平台或整体系统的角度,将平台收益作为优化目标,例如文献[49]和[50].

$$\begin{aligned} & \max_{\text{bidding strategy}} KPI \quad (5) \\ & \text{s. t. } cost \leq budget \end{aligned}$$

其中 $KPI = \sum_{i \in I} v_{ij} x_{ij}$ 或 $KPI = \sum_{i \in I} (v_{ij} - c_{ij}) x_{ij}$.

3.2.2 线性出价策略

文献[3]已经证明在第二价格密封拍卖中,每个竞标者按照拍卖物品对自己的真实估值出价,将获得最优收益. 在 RTB 中,假设对于同一个广告活动点击产生的价值是相同的,那么出价应该正比于该广告展示机会的期望点击率,这就是线性出价策略的基

本思路. 文献[51]基于这一思路,提出一个简单的基于点击率(转化率)的分段线性出价策略,如式(6)所示,这里 $b(j, i)$ 表示广告活动 j 对广告展示机会 i 的出价, $b(j)$ 表示广告活动 j 的基础价格,由系统管理者设定, $\Phi^*(j, i)$ 表示一个比例因子, $p(c|u, i, j)$ 表示对于广告展示机会 i 的定向用户 u , 广告活动 j 投放后发生转化的概率, $E_a[p(c|u, a, j)]$ 表示广告活动 j 投放到该用户 u 对应的其他广告展示机会 a 的期望转化率.

$$b(j, i) = \begin{cases} 0, & \Phi^*(j, i) < 0.8 \\ b(j) \times \Phi^*(j, i), & 0.8 \leq \Phi^*(j, i) \leq 1.2 \\ 2 \times b(j), & \Phi^*(j, i) > 1.2 \end{cases} \quad (6)$$

其中 $\Phi^*(j, i) = \frac{p(c|u, i, j)}{E_a[p(c|u, a, j)]}$.

文献[52]首先基于历史数据建立点击率预测模型和赢标价格预测模型,然后引入一个出价效率分界值(bid efficiency cutoff value)指标 $\rho(x_i)$,如式(7)所示,该指标与广告展示机会 i 的预测点击率 $pCTR(x_i)$ 与预测赢标价格 $pWP(x_i)$ 有关. 对于每个到达的广告展示机会,首先计算该广告展示机会的预测 CTR 和预测赢标价格,然后计算 $\rho(x_i)$,只有当 $\rho(x_i)$ 大于给定阈值,广告活动才按照预测赢标价格 $pWP(x_i)$ 出价,否则不进行出价.

$$\rho(x_i) = pCTR(x_i) / pWP(x_i) \quad (7)$$

3.2.3 非线性出价策略

由于预算和投放周期库存的约束,以及竞价市场的高度动态性,完全按照广告展示机会对给定广告活动的真实价值出价很难使广告活动的收益最大化,为此有研究者在线性出价的基础上,通过建立最优化问题,利用控制理论或统计学方法对线性出价策略进行了改进,例如文献[14]和[15]. 在文献[14]中,出价如式(8)所示, $eCPI(j, i)$ 表示将广告活动 j 的广告投放到展示机会 i 期望收益,出价 $b(j, i)$ 为 $eCPI(j, i)$ 减去广告活动 j 的价格调整因子 $\alpha(j)$,通过实时调整 $\alpha(j)$ 达到广告活动收益(利润)最大化. 为此作者首先按照收益(利润)最大化将出价问题建模为一个线性规划问题,然后证明可以通过将其转化为对偶问题,离线计算出最优解,以初始化 $\alpha(j)$;在此基础上分别从控制原理和统计方法的角度提出两种不同的 $\alpha(j)$ 实时更新方案. 为了进行在线更新,广告活动投放周期(通常为一天)被划分为若干时段,例如 $T=24$,然后在每个时段对 $\alpha(j)$ 进行更新.

$$b(j, i) = eCPI(j, i) - \alpha(j) \quad (8)$$

文献[15]与[14]的思路相似,在线性出价公式的基础上增加与时段相关的调整因子 $\phi(t)$,如式(9)所示,这里 b_0 是设置的符合广告活动定向规则的展示机会的基准价格, $b(t)$ 则表示广告活动在时段 t 的出价,它正比于广告活动在时段 t 的预测 CTR 的均值 θ_t 与其整个投放周期的 CTR 均值 θ_0 的比值. 作者将最大化给定预算下的点击次数作为优化目标,首先利用拉格朗日法求出基准价格 b_0 , 以及 $b(t)$, 再通过控制理论方法,例如 PID 控制器和水位控制器,对调整因子 $\phi(t)$ 进行更新,最后得出实时的出价 $b_a(t)$.

$$b_a(t) = b(t) \exp\{\phi(t)\}, b(t) = b_0 \cdot \frac{\theta_t}{\theta_0} \quad (9)$$

文献[16]提出一种新的概率统计的方法来建立最优公式,如式(10)所示,优化目标是最大化点击次数,这里 N_T 表示在预算约束下能够进行出价竞标的次数, B 为广告活动的预算, θ 是预测点击率, $w(b(\theta))$ 是出价 $b(\theta)$ 后的赢标率, $p_x(\theta)$ 为 CTR 的概率密度函数. 论文利用拉格朗日法求解,得到出价函数 $b(\theta)$ 与赢标函数 $w(b(\theta))$ 的关系;在此基础上,作者通过历史数据集观察赢标函数 $w(b(\theta))$ 的形式,用 $w(b(\theta)) = b(\theta)/(c + b(\theta))$ 拟合赢标函数,最后可得到 $b(\theta)$ 的解. 很明显,这里的出价与预测点击率是非线性关系,但其不足主要在于,拟合的赢标函数形式是否符合新的广告投放周期,并且该出价策略也是不能实时调整的.

$$b(\cdot)_{\text{ORTB}} = \arg \max_{b(\cdot)} N_T \int_{\mathbf{x}} \theta(\mathbf{x}) w(b(\theta(\mathbf{x}), \mathbf{x}), \mathbf{x}) p_x(\mathbf{x}) d\mathbf{x} \quad (10)$$

$$\text{s. t. } N_T \int_{\mathbf{x}} b(\theta(\mathbf{x}), \mathbf{x}) w(b(\theta(\mathbf{x}), \mathbf{x}), \mathbf{x}) p_x(\mathbf{x}) d\mathbf{x} \leq B$$

不同于上述将出价策略作为独立优化问题求解的思路,文献[50]指出,RTB 中出价主要面临三大挑战:(1) CTR 预测;(2) 广告展示机会的市场价格(成本)预测;(3) 计算最优竞标价格. 大多数的出价策略将三个挑战分别作为独立问题求解;然而文献[50]的作者认为三个问题是密切相关的,每个问题独立求解将使得广告活动的整体收益不能达到最优. 为此,作者提出一个全局的学习竞标框架,将三个挑战作为一个整体来考虑,通过三个组件的联合优化来实现广告活动收益的最大化.

3.2.4 基于强化学习的出价策略

随着强化学习(Reinforcement Learning, RL)在游戏、决策领域取得巨大进展^[53-54], RL 的研究吸引了大量研究者的关注,目前已经广泛应用于游戏博弈、机器人控制、优化决策等领域. 2016 年,文献

[55]首先尝试将强化学习引入 RTB 的出价策略中,随后京东、淘宝的研究人员纷纷将强化学习引入到商业化 RTB 平台中. 由于具有动态、自学习等特性,基于强化学习的出价策略已经成为出价策略研究的趋势. 本文将基于强化学习的出价策略作为研究的一个重点,在介绍主要出价策略之前,首先简要介绍强化学习的基本概念.

RL 的基本思想是通过最大化智能体(agent)从环境中获得的累积奖励,以学习能够达到目标的最优策略^[56]. RL 问题通常采用马尔可夫决策过程(Markov Decision Process, MDP)进行建模. 将 MDP 定义为一个四元组 $\langle S, A, \mathbf{P}, R \rangle$, 这里 S 表示从环境中观测到的状态集合, $s_t \in S$ 表示 agent 在 t 时刻所处的状态; A 为 agent 可执行动作的集合, $a_t \in A$ 表示 agent 在 t 时刻所采取的动作; \mathbf{P} 表示状态转移概率矩阵, $p(s_{t+1} | s_t, a_t)$ 表示 agent 在状态 s_t 执行动作 a_t 转移到新状态 s_{t+1} 的概率; R 表示从环境中获得的奖励, $R_t(s_t, a_t) = E[R_{t+1} | s_t, a_t]$ 表示 agent 在状态 s_t 执行动作 a_t 后的奖励;策略 π 是状态空间到动作空间的一个映射,表示 agent 在状态 s_t 选择动作 a_t 的映射关系,执行该动作并以概率 $p(s_{t+1} | s_t, a_t)$ 转移到下一状态,同时接受来自环境的奖励 $R_t(s_t, a_t)$, 这里如果策略是随机的,则根据 $\pi(a_t | s_t)$ (agent 在状态 s_t 选择动作 a_t 的概率)选择动作,如果策略是确定的,则根据状态选择动作 $a_t = \pi(s_t)$. 在 RL 中引入两个价值函数:状态值函数 $V(\cdot)$ 和状态-动作值函数 $Q(\cdot)$, 分别表示指定“状态”上以及指定“状态-动作”上的累积奖励. 关于 RL 的更详细的介绍可参考文献[57].

文献[55]是最早将出价决策过程建立为 RL 问题进行求解的方案,模型如图 9 所示. 假设每个广告活动的出价是一个已知模型的马尔可夫决策过程 $(\mathbf{S}, \{\mathbf{A}_s\}, \{P_{ss'}^a\}, \{R_{ss'}^a\})$, 其中状态 \mathbf{S} 是指对拍卖环境的部分观察,包括时间、初始投标参数(例如设置的超参数、预算约束等)、预算消费、获得报酬、拍卖环境竞争水平和活动信息;动作 $\{\mathbf{A}_s\}$ 指设置竞标价格;奖励 $\{P_{ss'}^a\}$ 是指在状态 s_t 执行动作 a_t 转移到新状态 s_{t+1} 的奖励,例如获得点击或转化;环境 $\{R_{ss'}^a\}$ 指 RTB 系统.

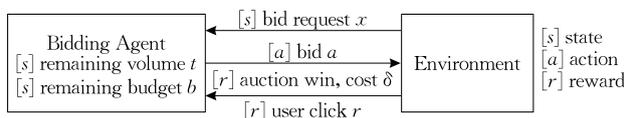


图 9 基于强化学习的出价决策过程

这里将每个广告展示机会的拍卖竞价请求的到来过程视为一个偶发过程(episodic episode): 假设对于一个广告主, episode 包括 T 个竞价请求, 每个竞价请求的特征向量记为 \mathbf{x}_t , 广告商的预算为 B , 因此在强化出价问题中, 状态用 (t, b, \mathbf{x}_t) 表示, 这里 $t \in \{0, \dots, T\}$ 表示剩余的竞价请求数量, $b \in \{0, \dots, B\}$ 表示当前可用预算; agent 根据当前状态决定出价 $a_t \in \{0, \dots, b\}$; $P_{ss'}$ 表示从状态 (t, b, \mathbf{x}_t) 转移到下一个状态的概率, 按照是否赢标可分为两种情况, 计算如式(11)所示。

赢得拍卖:

$$\begin{aligned} \mu(a_t, (t, b, \mathbf{x}_t), (t-1, b-\delta, \mathbf{x}_{t-1})) &= p_x(\mathbf{x}_{t-1})m(\delta, \mathbf{x}_t) \\ r(a_t, (t, b, \mathbf{x}_t), (t-1, b-\delta, \mathbf{x}_{t-1})) &= \theta(\mathbf{x}_t) \end{aligned}$$

没有赢得拍卖:

$$\begin{aligned} \mu(a_t, (t, b, \mathbf{x}_t), (t-1, b, \mathbf{x}_{t-1})) &= p_x(\mathbf{x}_{t-1}) \sum_{\delta=a_t+1}^{\infty} m(\delta, \mathbf{x}_t) \\ r(a_t, (t, b, \mathbf{x}_t), (t-1, b, \mathbf{x}_{t-1})) &= 0 \end{aligned} \quad (11)$$

广告活动赢得拍卖, 获得的奖励为 $\theta(\mathbf{x}_t)$, 并按照第二高价 δ 付费, 这里通常称为市场价格, $p_x(\mathbf{x}_{t-1})$ 表示广告竞价请求中特征向量 \mathbf{x}_{t-1} 的概率密度函数, $m(\delta, \mathbf{x}_t)$ 表示在给定特征向量 \mathbf{x}_t 的条件下市场价格 δ 的概率密度函数, $\theta(\mathbf{x}_t)$ 表示竞价请求的特征向量为 \mathbf{x}_t 时的点击率; 如果没有赢得拍卖, 则获得的奖励为 0, 不付费, 没有赢得拍卖意味着市场价格大于 agent 的出价 a_t . 其中 $\mu(\cdot)$ 表示从当前状态 (t, b, \mathbf{x}_t) 执行动作 a_t 后转移到新状态的概率, $r(\cdot)$ 表示从当前状态执行动作后获得的奖励。

根据文献[55]的描述, 在 RTB 的出价中, 策略 π 是确定的, 即从每个状态 $s \in S$ 到动作 $a \in A_s$ 的映射是确定的, $a = \pi(s)$. 根据策略 π , 可以计算状态值函数 $V^\pi(s)$, 即从状态 s 开始按照策略 π 执行动作得到的累积奖励, 如式(12)所示, 因为在 RTB 中是以总点击数为优化目标, 而点击可能发生在广告展示的几分钟后, 也有可能不发生点击, 因此是一个偶发过程, 所以这里衰减因子 $\gamma = 1$. 最优状态值函数定义为 $V^*(s) = \max_{\pi} V^\pi(s)$, 因此有最优策略 $\pi^*(s)$ 如式(13)所示. 由于是已知模型的, 论文提出利用动态规划方法求解 $\pi^*(s)$.

$$V^\pi(s) = \sum_{s' \in S} \mu(\pi(s), s, s') (r(\pi(s), s, s') + V^\pi(s')) \quad (12)$$

$$\begin{aligned} \pi^*(s) = & \\ \arg \max_{a \in A_s} \left\{ \sum_{s' \in S} \mu(a, s, s') (r(a, s, s') + V^*(s')) \right\} & \quad (13) \end{aligned}$$

文献[58]在强化学习框架下, 将出价过程建模

为一个预算受限的马尔可夫决策过程. 不同于文献[55], 它利用六元组建模 $(S, \{A_s\}, P, R, C, V)$, 这里状态是指针对广告展示机会竞价请求的点击率 $\theta(\mathbf{x})$, $S = \Theta$, 动作仍然是出价, 状态转移概率 P 是预测点击率的概率密度函数, 独立于当前状态和执行动作; 奖励(reward)仍然使用获胜后的预测点击率; $C(s, a)$ 表示 agent 在状态 s 出价 a 的期望成本(市场价格); V 表示每个广告展示机会竞价请求的预算限制. 论文利用批量强化学习(Batch-RL)^[59] 从先验知识中推导最优策略, 确定每个状态的竞标价格。

上述两个出价策略是基于有模型的 RL 方案, 然而 RTB 中高度的动态性和随机性, 导致很难建立准确的模型来描述状态转移过程, 因此更多的出价策略是基于无模型的强化学习算法来求解. 文献[60]的建模与文献[55]相似, 不同之处在于状态定义为第 t 次广告展示机会的相关信息, 动作仍然是出价, 奖励为平台的利润, 即对第 t 次展示机会的收益减去成本. 然后作者提出一个基于 DQN(Deep Q Network)的异步随机算法来求解, 这种算法引入了一个完全异步的学习机制. 文献[61]针对赞助搜索的实时竞价场景, 提出一个深度强化学习方法来进行出价. 不同于文献[55], 它的最优策略不是直接生成竞标价格, 而是得到一个调节参数 α , 出价采用了一个类似线性出价的形式, 通过调节 α , 得到最大累积奖励. 目标是通过迭代更新最大化状态-动作值函数 $Q(s_t, a_t)$, 从而找到最优策略. 论文采用 DQN 算法得到近似最优解。

文献[62]提出 RTB 的场景由于 agent 的状态、状态转移概率、奖励函数等往往是未知的, 因此很难准确地利用动态规划方法来求解, 因此提出一个无模型(model-free)的强化学习出价策略. 首先作者分析了 RTB 中出价与广告展示机会真实价值的关系, 从而将出价问题转化为一个对参数的调节问题. 仍然假设整个 episode 包括 T 次广告展示机会的竞价请求, 这是一个衰减因子 $\gamma = 1$ 的有限马尔可夫决策过程. 状态是对拍卖环境的部分观察, 包括时间、初始投标参数、预算消费、获得报酬、拍卖环境竞争水平和活动信息; 动作是调整参数 λ ; 奖励是执行动作后赢得的广告展示机会的价值. 论文使用经典的 DQN 算法求解, 使用 ϵ -greedy 策略根据状态-动作值函数 $Q(s_t, a_t)$ 返回的累积奖励来选择动作。

3.2.5 典型出价策略的特点对比

目前出价策略主要分为线性出价、非线性出价、

基于强化学习的出价三类. 早期的线性出价策略从广义第二高价拍卖的最优出价策略出发, 按照广告展示机会对广告活动的真实价值出价, 即出价正比于广告展示机会的预测点击率(预测转化率). 线性出价的优点是简单易于部署; 然而对于存在条件约束的 RTB 场景, 按照真实价值出价通常不能达到最优目标. 为此研究者提出基于历史数据, 通过最优化理论直接求解最优出价或对真实价值进行调整, 称为非线性方法(即出价与广告展示机会的预测点击率是非线性关系). 考虑到 RTB 是一个高度动态、随机的场景, 在新的广告投放周期, 参与竞价的广告活动、广告展示机会的质量和数量、甚至竞争者的出价策略都可能发生变化, 为此研究者提出应该针对出价函数进行以时段(time slot)为粒度的在线实时调整, 例如文献[14]和[15].

遗憾的是, 无论是线性还是非线性出价策略, 在针对每个广告展示机会出价时, 都只针对这次广告展示能够带来的回报, 而没有考虑对后续广告展示机会竞价产生的影响, 即出价决策只考虑已经发生的事实. 一种更为理想的出价策略是将每次出价决策与整个投放周期的竞标过程关联起来(包括未发生的后续竞拍), 将整个投放周期的广告展示机会的拍卖作为一个连续动态的决策过程, 不仅关注眼前回报, 也需要关注未来的回报, 动态地调整出价. 而以马尔可夫决策过程为载体的深度强化学习正好符合这个目标, 将出价作为一个连续的决策, 为每一个广告展示机会生成最优的出价(曝光粒度的优化), 来优化动态 RTB 环境下的广告效果, 从而获得最大化的长期回报. 因此, 基于强化学习的出价策略被认为是获得最优出价的理想解决方案. 表 2 分别从出价策略类别、是否进行在线调整、优化目标三个方面对典型出价策略的特点进行分析.

最后, 对比不同类型典型出价策略在 iPinYou 数据集上的性能. 这里仅采用了 1458 广告活动在

表 2 典型出价策略的特点分析

	类别	在线调整	优化目标	优化算法
[51]	线性	不支持	收益	最优出价理论
[52]	线性	不支持	收益	优化理论
[14]	非线性	时段粒度	收益/利润	优化理论+控制理论
[15]	非线性	时段粒度	收益/利润	优化理论+控制理论
[16]	非线性	不支持	点击次数	优化理论
[50]	非线性	不支持	点击次数	使用机器学习方法联合优化
[55]	有模型强化	曝光粒度	收益/利润	动态规划
[58]	有模型强化	曝光粒度	收益/利润	Batch-RL
[61]	无模型强化	曝光粒度	收益/利润	DQN
[62]	无模型强化	曝光粒度	收益/利润	DQN

2013 年 6 月 6 日的竞价和点击日志作为训练集, 6 月 7 日的竞价和点击日志作为测试集. 由于数据集是由 DSP 方提供的, 对于一个广告展示机会的市场价格, 只有当广告活动获胜时才能获知, 因此在实验中, 仅采用获胜的广告展示机会作为达到的广告展示机会(大幅低于实际达到数量), 计算竞标价格, 出价大于等于市场价格时为获胜, 统计获得的点击数, 以及相关性能指标, 如表 3 所示, 数据集中所有费用的单位为人民币(10^{-3} 分). 由于测试集本身是按照每个展示机会固定出价 0.3 分/次得到的, 因此固定出价 0.3 分所捕获的点击数就是测试集中所有的真实点击数 307, 固定出价 0.2 分/次和 0.1 分/次分别捕获的点击数为 277 和 216, 线性和非线性出价捕获的点击数为 296 和 299, 强化出价捕获的点击数为 301. 但是从平均点击花费(CPC)指标分析, 性能由高到低依次为: 固定出价(0.1 分)、线性出价、非线性出价、固定出价(0.2 分)、强化出价、固定出价(0.3 分). 这里强化出价获得了最多的点击数(固定出价 0.3 分除外), 但是也获得了最高的 CPC, 这是因为在实验中, 使用的是固定出价 0.3 分策略的总花费作为预算, 因此强化出价策略在自主学习时会分析出预算非常充足, 而使用点击数作为最大化累计奖励的指标, 会使学得的出价策略偏向于出高价, 即在预算充足的条件下, 出价越高, 收益越大.

表 3 典型出价策略的性能对比

出价策略	曝光数	展示数	点击数	花费/ 10^{-3} 分	CTR/%	CPC/分	CPM/ 10^{-3} 分
固定出价(0.3 分/次)	478 109	478 109	307	30 228 554	0.0642	98.50	63.22
固定出价(0.2 分/次)	478 109	457 628	277	25 385 986	0.0605	91.65	55.47
固定出价(0.1 分/次)	478 109	405 138	216	18 018 885	0.0533	83.42	44.48
线性出价 ^[16]	478 109	433 630	296	24 772 734	0.0683	83.69	57.13
非线性出价 ^[16]	478 109	457 606	299	26 742 433	0.0653	89.44	58.44
强化出价 ^[55]	478 109	467 067	301	28 432 348	0.0644	94.46	60.87

3.2.6 竞价愿景

如前所述, 在出价策略中对于赢标价格的分布

进行预测是非常重要的一个问题. 在广义第二高价拍卖中, 赢标价格是指的出价中的第二高价, 也称为

市场价格,不仅体现了广告展示机会的质量,更体现了拍卖市场中的竞争程度.对市场价格分布预测的工作被称为竞价愿景(bidding landscape).竞价愿景建模的一种可行方法是构建一个启发式函数,在历史数据集上拟合市场价格分布.因此给定市场价格分布的概率密度函数 $p_z(z)$ 和出价 b ,可利用式(14)计算赢标率 $w(b)$.这种方法^[63]的缺点是现实中市场价格的分布可能与假设的函数形式有很大的偏差.

$$w(b) = \int_0^b p_z(z) dz \quad (14)$$

市场价格分布预测的一个难点是数据删失问题.由于 RTB 中会接入多个 DSP,每个 DSP 通常只有在自己代理的广告活动赢得拍卖时,才会获知赢标价格(即支付费用),对于失败的竞标,DSP 只知道自己的出价,即市场价格的下界,这种偏差被称为右删失(right-censored).此外,如果提供方平台设置了保留价,且保留价介于第一高价和第二高价之间,则获胜的 DSP 是按照保留价付费,而不是市场价付费,因此 DSP 只能观察到市场价格的上界.这种偏差被称为左删失(left-censored).删失问题的存在会导致拟合的市场价格分布与真实分布之间存在较大偏差.为此,文献[64]提出市场价格预测应同时考虑赢标数据和删失数据,如式(15),这里赢标数据是指 DSP 赢得广告展示机会的数据,DSP 有真实的市场价格,而删失数据是指 DSP 没有赢得的广告展示机会,DSP 只有观察到的市场价格(即自己的出价,市场价格大于等于出价).

$$v_i = [P(v_i < b_i) \beta_{lm} + (1 - P(v_i < b_i)) \beta_{clm}]^T \mathbf{x}_i = \beta_{\max}^T \mathbf{x}_i \quad (15)$$

在式(15)中, v_i 表示广告展示机会 i 的预测市场价格, \mathbf{x}_i 表示广告展示机会 i 的特征向量, $P(v_i < b_i)$ 表示该拍卖获胜(市场价格小于出价)的预测概率, $(1 - P(v_i < b_i))$ 表示该拍卖失败的预测概率, $P(v_i < b_i)$ 采用逻辑回归模型进行预测; β_{lm} 表示基于真实市场价格建立的市场价格预测模型的参数向量, β_{clm} 表示基于删失数据建立的市场价格预测模型的参数向量.

目前通常的市场价格预测都是假设市场价格分布服从相同的分布.文献[65]的作者通过真实数据分析,发现市场的赢标率分布在不同的特征或特征组合上是不同的,因此用一个函数(例如文献[64]中的标准正态分布)来拟合市场价格分布是不恰当的.为此文献[65]提出通过构建具有 K -means 聚类的

决策树,在无任何函数假设和约束的情况下,自动学习给定广告展示机会到市场价格分布的函数映射.作者的主要贡献是通过建立一棵二叉树,将样本按照特征相似性进行聚类,每个叶子节点将包含一组特征相似的样本;针对每组特征相似的样本子集,引入生存模型来建立市场价格预测模型.为了构建二叉树,作者提出在每个节点使用划分后 KL 散度最大的属性进行划分,直到叶子节点中的样本数据量小于预设值,最后使用减小误差的剪枝方法来对二叉树进行剪枝.

文献[66]指出在实际的 RTB 系统中,广告活动的定向规则可能会经常变化,如果为每个广告活动建立市场价格预测模型,当广告活动的定向规则发生变化或面对一个新的广告活动,将无法准确预测.为此,文献[66]提出不应该针对单个广告活动建立市场价格预测模型,而是针对广告活动的定向特征的组合向量进行预测.每种定向特征值的组合向量作为样例,例如一个广告活动有 3 个定向特征,发布媒体($P1, P2$),用户性别(男,女),用户年龄($18 \sim 23, 23 \sim 30$),那么 $\langle P1, 男, 18 \sim 23 \rangle$ 就构成一个样例.样例是由所有特征的“叉积”产生的,因此数量巨大且稀疏,为此作者提出用出价星形树(bid star tree)来解决稀疏性问题,同时降低存储开销.文献[66]的整体思路如下:首先对每个广告活动的定向特征进行分解,得到多个特征值组合向量(样例);然后在样例粒度上对市场价格进行预测;最后将样例粒度的预测聚合到广告活动粒度上.

综上,竞价愿景的研究通常是在历史数据的基础上对赢标率函数进行拟合或建立市场价格的预测模型,这类方法的主要问题在于拍卖市场中广告主出价策略的高度动态性,新广告投放周期的市场价格分布可能与历史数据有较大差异.

3.3 预算与步进管理

预算与步进管理机制主要解决的是实时竞价过程中预算花费的速度问题.在 RTB 中,如果不对预算花费速度进行管理,当出价策略相对新的广告投放周期过于激进(出价偏高),将会使广告主的预算被提前花光,导致无法购买后续时段可能带来点击(转化)的高质量曝光机会;当出价策略过于保守,将会使广告主因为出价偏低丢掉大量可能带来点击(转化)的高质量曝光机会.为了使预算的花费速度适应新的广告投放周期,研究者提出需要对预算的花费或者投放步进进行管理.

目前预算与步进管理的基本解决思路是将整个

广告投放周期分为若干时段(通常为一天,划分为24个时段),为每个时段预分配一定的预算;在整个广告投放周期中,每个时段的花费不能超过它的预算;如果已经结束的时段有剩余预算,可将剩余预算在后续时段重新分配^[67].最简单的预算分配方案是在整个广告投放周期均匀分配.然而,均匀分配不能有效区分不同时段广告展示机会数量和质量的差异,可能导致预算浪费在低质量的广告展示机会上.在文献[68]中,作者假设所有曝光机会对于广告主的价值是相同的,因此每个时段的预算应该与该时段的曝光机会数目成正比.在文献[69]中,作者仍然按照时段粒度进行预算分配的思路,首先基于每个时段的预测曝光机会数目分配预算,然后根据当前时段的可用预算计算该时段的步进率(pacing rate),最后根据步进率和曝光机会的 pCTR 计算竞标价格.因此步进率越高,竞标价格越高,广告主在该时段成功竞标的曝光机会就越多.文献[70]的作者沿用了文献[69]中步进率的思路,首先将曝光机会根据 pCTR 划分为若干组,为每个组确定步进率,这里同一个组的曝光机会具有相同的步进率,具有更高 pCTR 的组的曝光机会具有更高的步进率.因此在文献[70]中,具有更高 pCTR 的曝光机会的竞标价格将会更高,即更多地将预算花费在高 pCTR 的曝光机会上.

综合现有研究表明,通过为每个时段分配预算并根据剩余预算进行自适应调整,是实现预算与步进管理的有效方法.当预算不充足时,通过为每个时段计算广告投放的步进率,可以有效避免预算花费到那些低质量的曝光机会上,从而提升预算的使用效率.

4 提供方平台的研究进展

提供方平台的主要功能是登记发布媒体的广告库存,然后向接入的 ADX 发起拍卖请求,因此 SSP 是为媒体服务的,目标是最大化发布媒体的收益.目前 SSP 的研究主要集中在保留价优化上,即通过为每个广告展示机会设置拍卖保留价,来提升发布媒体的收益.此外,发布媒体和 SSP 通常同时采用离线和在线两种渠道售卖广告库存(广告展示机会对于提供方通常称为库存),一般策略是优质库存采用离线合约渠道销售,剩余库存采用在线 RTB 渠道销售.然而随着 RTB 发展成为主要的广告库存销售渠道,SSP 需要解决的一个问题是如何在两个(或多

个)渠道间动态分配广告库存,以使得发布媒体和 SSP 的收益最大化,这就是库存动态分配问题.本节将分别介绍这两个关键问题的研究进展.

4.1 保留价优化

首先回顾发布媒体的收益是怎么产生的.当一个用户浏览发布媒体的一个嵌入广告位的网页时,网页上嵌入的脚本代码会向 SSP 发起一次广告展示机会的请求,SSP 将向交易平台(ADX)发起对该广告展示机会的拍卖请求,ADX 根据各接入的 DSP 返回的出价,决定获胜方,由获胜方按照市场第二高价或者广告展示机会的保留价付费给媒体,因此对于发布者来说,其收益主要来自于拍卖市场中各竞标者的出价,这通常由该次广告展示机会对每个竞标者(参与出价的广告活动)的价值决定,而从 SSP 的角度,提升发布媒体收益则可以通过优化每个广告展示机会的保留价(reserve price)设置来实现.这是因为在 RTB 中,若假设针对广告展示机会 t 有 K 个竞标者,他们的出价按由大到小降序排列为 $b_1(t), b_2(t), \dots, b_K(t)$,那么获胜者为广告展示实际支付的费用 $r(t)$ 为拍卖中第二高出价 $b_2(t)$ 和广告展示机会保留价 $\alpha(t)$ 的较大值,如式(16)所示.

$$r(t) = \begin{cases} \alpha(t), & b_1(t) \geq \alpha(t) > b_2(t) \\ b_2(t), & b_2(t) \geq \alpha(t) \\ 0, & \alpha(t) > b_1(t) \end{cases} \quad (16)$$

式(16)说明了三种情况:如果拍卖中的最高出价小于保留价时,则该次广告展示机会流拍,通常将降低保留价继续发起第二次拍卖;如果最高出价大于等于保留价且保留价大于第二高出价,则获胜者将按照保留价付费给媒体;如果第二高出价大于等于保留价,则获胜者将按照第二高出价付费给媒体.因此对于发布媒体来说,保留价设置在第一高价和第二高价之间,且尽可能逼近第一高价,将可以最大化发布媒体的收益;针对一个广告展示机会,如何找到最优保留价,便是保留价优化问题^[3].

理论上,如果假设 SSP 或发布媒体已知竞标者对展示机会的私有估值(private value)的分布,保留价优化问题可以利用最优拍卖原理来求解^[3].首先假设 K 个竞标者是风险中性且同质的,具有相同的私有价值分布,每个竞标者能够计算出对于一个展示机会的私有价值.定义 $F_k(x)$ 表示竞价者 k 预估的展示机会的私有价值不高于 x 的概率,通常假设 $F_k(x)$ 服从对数正态分布或者均匀分布,由于每个竞标者的出价是独立同分布的,则 K 个竞标者出价的联合概率分布如式(17)所示.根据文献[71],最优

保留价 r^* 可通过求解式(17)得到, 这里 $F(r^*)$ 表示出价不高于保留价的概率, $F'(r^*)$ 是概率密度函数, 是 $F(r^*)$ 的一阶导数.

$$F(\cdot) = F_1(\cdot) \times F_2(\cdot) \times \dots \times F_K(\cdot) \quad (17)$$

$$r^* = \frac{1 - F(r^*)}{F'(r^*)} \quad (18)$$

如何设置保留价在拍卖理论和搜索广告的竞价机制中已经有较为广泛的研究^[72]. 遗憾的是, RTB 场景与最优拍卖理论甚至搜索广告中的场景都有较大的差异. 首先, 最优拍卖理论要求在计算最优保留价之前, SSP 需要知道每个竞标者对广告展示机会的私有估值分布, 在 RTB 场景中 SSP 或者发布媒体很难获得每个竞标者的私有估值, 因为在 RTB 中每个竞标者对广告展示机会的价值估计是一个复杂的过程, 依赖于对该次广告展示机会的点击率预测以及竞标者自己的出价策略等, 而这些都不会对外公开. 其次, 竞标者很可能自己都无法准确估计广告展示机会的私有价值, 如前所述, 出价策略中私有价值通常被定义为点击率(转化率)与点击价值(转化价值)的乘积, 对点击率的预测存在偏差, 或出价策略发生变化, 都将直接影响到竞标者对展示机会的价值估计.

目前 RTB 的保留价优化研究才刚起步. 工业界常采用的保留价计算方法, 包括: (1) 给每个展示机会设置一个固定的保留价, 即 $r = a$, 当 $a = 0$ 时即不设置保留价, 这里保留价通常反映的是发布者对展示机会的价值评估, 可以由发布者签订的担保合同中 CPM 的价格来得到; (2) 一种更为复杂的计算方法是利用历史收益的均值来计算保留价, 如式(19)所示, $r(t)$ 表示第 t 次展示机会的保留价, 利用第 t 次展示机会之前的 M 次展示机会的收益的平均值作为最新一个展示机会的保留价. 该方法的一种简单扩展是给不同的展示机会的收益赋予不同的权重, 越接近当前的展示机会的收益, 权重越高, 如式(20)所示. 这类计算方法的优点是直观简单, 在实际生产平台中易于部署且效果较好.

$$r(t) = \frac{1}{M} \sum_{i=t-M}^{t-1} rev(i) \quad (19)$$

$$r(t) = \frac{1}{M} \sum_{i=t-M}^{t-1} w(i, t) rev(i) \quad (20)$$

另一类计算保留价的方法仍是基于竞标者对广告展示机会(曝光机会)的私有价值分布来计算最优(或近似最优)保留价. 这类方法的前提是假设所有竞标者的私有价值(即出价)是独立同分布的, 且 SSP(publishers)知道这个分布. 文献[73]的作者在

最优保留价的基础上, 针对 LinkedIn 特有的社交广告定向投放场景, 基于最优保留价计算公式提出了针对用户等级和针对广告活动等级的保留价计算方法. 首先利用回归模型预测每个广告活动(即竞标者)对符合其定向规则的用户们的出价分布; 然后利用式(18)计算每个广告活动对定向用户的保留价; 在此基础上, 结合 LinkedIn 社交广告的具体需求, 将符合一个广告活动定向规则的用户划分为一个用户片段, 然后计算每个广告活动对其定向用户片段的保留价, 将该片段内的用户保留价由大到小降序排列, 然后按照分位数规则选择某个用户的保留价作为该广告活动的保留价. 例如片段内有 10 000 个用户, 按照二分位规则确定保留价, 即选择排名第 5000 的用户的保留价作为该广告活动的保留价. 因此在该方案中, 广告活动的出价必须同时大于等于用户和它自身的保留价, 在扣费时按照用户保留价和广告活动保留价中的较大者扣费. 该方案按照定向用户对于广告活动的价值来设置保留价, 不同的广告活动保留价也不同, 这对于广告商是不公平的, 却可以最大化售卖方的收益.

文献[74]仍然沿用最优保留价的思路, 主要针对雅虎的 BrightRoll 视频广告交易平台的保留价进行研究, 特别是存在如图 10 所示的下游交易平台的场景. 作者研究发现 BrightRoll 上有一部分曝光机会是来自于其他交易平台(即下游平台), 即在 BrightRoll 中获胜的竞标者还需要参与下游交易平台的竞价, 只有在下游交易平台获胜, 才能赢得最终的展示机会, 这样的竞价场景会影响上游交易平台对最优保留价的设置, 因此作者提出一个基于博弈论的结构化模型来推导保留价修正公式.

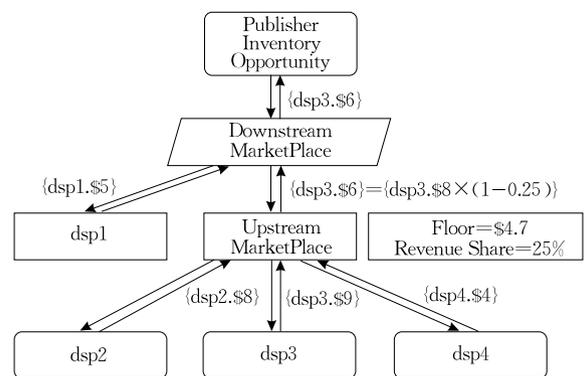


图 10 有下游交易平台的 RTB 系统

文献[75]首先分析了文献[74]中的结构化模型对于雅虎的展示广告实时交易平台 BRXD 不太适合, 在展示广告交易中曝光机会的量更大, 参与竞标

的广告活动更多,竞标价格的变化范围更大,且很难满足竞标者出价独立同分布的要求.为此,文献[75]的作者提出了一个新的方法来实现对高价值库存保留价的动态调整.该方法首先构建若干分类器来预测一个展示机会的最高出价,若预测出价高于某个阈值,则将该展示机会视为高价值库存;在此基础上,构建另一组分类器来预测展示机会的最高出价和次高出价之间的价格差,从而计算推荐的保留价.注意,为了降低噪声影响,该方法仅对高价值曝光机会计算相应的推荐保留价,并根据两个最高出价之间的价格差,灵活采用不同的策略.特别需要说明的是,文献[75]的方法无需对竞标者的出价分布进行假设和预测,也不要求所有竞标者的出价是独立同分布的,而仅仅假设每个库存的两个最高出价是相对稳定的.

综上所述,由于 RTB 在竞价机制、影响因素等方面的复杂性,导致目前 RTB 中关于保留价优化的研究尚处于起步阶段,而一个良好的保留价计算方法既可以满足广告商的投放需求,又可以使库存供应方(SSP 或 publisher)获得更大的收益,进而促进 RTB 市场的良性发展.

4.2 库存分配策略

库存分配是随着 RTB 不断发展出现的新问题.对于广告展示机会的提供方来说,有多种渠道销售广告库存(广告展示机会),既可以通过担保合约模式销售,也可以通过 RTB 模式销售,或者采用担保合约与 RTB 的混合模式销售,这就产生了库存分配问题.当采用混合模式时,不同销售模式的成本与收益是不同的,如何合理地将广告库存实时分配到各种销售模式以实现最大化收益或利润,是工业界非常关注的问题^[76-77],示意图如图 11 所示.

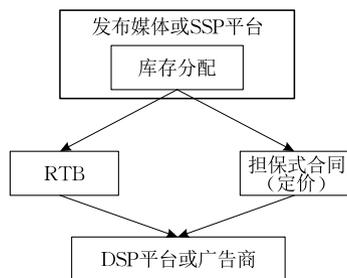


图 11 库存分配问题示意图

文献[76]是较早研究库存动态分配的文献,其基本思想是对两种渠道销售广告库存的收益以及担保合同违约产生的罚金进行建模,提出一种最优动态销售方案,以实现三者之间的精准权衡.此外,作

者研究了存在多个担保合约渠道的情况下库存分配的调度问题,并提出一个简单的解决方案:将可用库存分配到即将最早到期的担保合约中,以尽量避免过早的罚金.文献[77]将库存分配问题建模为一个随机控制问题,推导出一种基于广告位质量和 ADX 出价的联合分布环境下的有效分配策略,并证明了该策略的渐进最优性.文献[78]针对两种销售渠道提出一个数学模型来分配广告库存,并为通过担保合约渠道销售的库存给出一个动态的最优定价方案.此外,该模型还讨论了市场竞争程度对库存分配和定价策略的影响:当市场竞争不激烈时,广告主可能更愿意通过 RTB 购买广告库存,此时应该降低担保合约渠道的库存定价,以激励广告主购买;当市场竞争激烈时,广告主更愿意通过有保障的合约渠道购买,此时发布方或 SSP 应该提高担保合约渠道的库存定价,以获得更高的收益.

文献[79]针对两种销售渠道,讨论了三种不同的库存分配策略:全部库存通过 RTB 渠道销售;优先满足担保合约的需求,剩余库存通过 RTB 销售;只将小部分库存通过担保合约销售,更多库存通过 RTB 销售.作者首先将库存分配问题描述为 SSP 收益优化问题.SSP 收益是来自两个渠道的收益之和,RTB 渠道的收益为通过 RTB 销售的所有库存的收益(可能存在库存流拍而无收益的情况),合约渠道的收益为广告实际展示次数乘以固定展示价格并减去可能的违约罚金.论文通过分析比较三种库存分配策略的期望收益,可以确定最佳的库存分配方式.同时文献[79]中的实验表明,不存在一种在任何场景下都能达到收益最优的库存分配策略,不同的分配策略在不同场景下都表现出了各自的优越性.因此,作者提出一种基于竞标价格分布预测和合约价格预测的实时分配策略,即如果一个广告库存的预测合约价大于所有的预测竞标价,则将其投放到合约渠道销售,如果小于最高的预测竞标价,则将其投放到 RTB 渠道销售.

综上所述,可以发现只要通过 RTB 拍卖库存,始终存在流拍的风险(无广告活动出价或最高出价低于保留价),因此为了避免库存浪费,最大化收益,SSP 或发布媒体通常会对流拍的库存再次拍卖.图 12 展示了目前工业界普遍采用的一种广告库存分配方式——瀑布(waterfalling)模式^①.首先当用户浏

① What is Waterfalling and How Does it Work. <https://clearcode.cc/blog/what-is-waterfalling>, 2016

览的网页发起一次广告展示请求时, SSP 会判断该次广告展示机会是否满足发布媒体通过担保合约渠道直接销售的库存的投放要求, 如果满足则直接展示合约广告; 如果不满足投放要求, 则将该广告展示机会的竞价请求(携带保留价)发到与之相连的第一个 ADX 进行拍卖, 如果在 ADX(1) 中流拍, 则继续请求相连的 ADX(2) 进行拍卖, 以此类推, 直到广告展示机会被成功拍卖或遍历所有交易平台。瀑布模式的优点在于可以提升库存的销售比例, 避免库存浪费; 问题在于管理复杂高, 需要对相连的各个 ADX 进行优先级排序, 而且重复拍卖次数越多, 消耗的时间也越多, 而每个广告库存是有时限的, 可能导致无法在有效时间内销售出去。

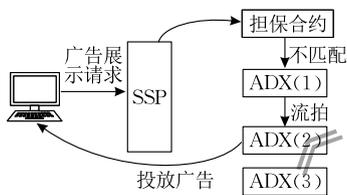


图 12 广告库存在 SSP(发布媒体)的销售过程

5 交易中心的研究进展

RTB 的核心是采用一种在线的实时拍卖机制对广告展示机会进行售卖, 使广告商能够合理利用预算将广告精准投放给自己的目标用户。首先简要介绍常见的拍卖形式: (1) 英式拍卖(English auction), 拍卖由拍卖师主持, 开始时拍卖师报出一个较低的价格, 只要还存在至少两个竞拍者, 就以一个较小的增量逐渐提高价格, 直到只剩下一个竞拍者时才停止, 最后剩下的竞拍者将赢得拍卖物品, 并按照倒数第二个竞拍者放弃竞拍时的价格来支付费用^[80]; (2) 荷式拍卖(Dutch auction), 与英式拍卖相反, 它是一个公开降价的过程, 开始时拍卖师报出一个足够高的价格, 然后逐步降低价格, 直到某一个竞拍者在当前价格下表现出了兴趣, 那么该竞拍者将赢得拍卖物品^[81]; (3) 第一价格密封拍卖(Sealed-bid first price auction), 所有竞拍者以密封形式提交出价, 出价最高者将赢得拍卖, 并按其出价支付费用^[82]; (4) 第二价格密封拍卖(Sealed-bid second price auction), 所有竞拍者以密封形式提交出价, 出价最高者将赢得拍卖, 并按第二高出价支付费用^[83]。由于 RTB 的实时性要求, 因此 RTB 中一般采用第一价格密封拍卖和第二价格密封拍卖, 而第二价格拍卖由于能够鼓励竞拍者(广告主)按照广告展示机会

对自己的私有真实价值出价, 因此在 RTB 中得到了广泛的应用。本节将对 RTB 中的拍卖机制展开研究, 首先介绍目前广泛采用的第二价格密封拍卖机制及其相关推广, 讨论其中的关联以及存在的问题。

5.1 第二价格密封拍卖

5.1.1 Vickrey 拍卖

第二价格密封拍卖也叫做 Vickrey 拍卖, 即所有竞拍者以密封形式独立出价, 出价最高的竞拍者赢得拍卖, 并按照第二高价支付费用^[83]。在 Vickrey 拍卖中, 已经证明了在竞拍者风险中性的假设下, 竞拍者的占优策略(Dominant Strategy)是按照自己对拍卖品的真实估值出价, 即 Vickrey 拍卖机制是真实的(truthful)。因此对于竞拍者来说, 只需将注意力集中在对拍卖品的私有估值上, 从而可以节省计算成本, 提升利润。这里的关键和难点是如何理解和证明 Vickrey 拍卖的真实性, 它是 RTB 出价策略的基础, 也是在 RTB 交易机制中选择第二价格密封拍卖的基础。在拍卖理论中, 如果一个拍卖机制可以激励竞拍者按照他们对物品的真实估值来出价, 即竞拍者的出价等于其对拍卖品的真实估值是占优策略, 那么该机制是真实的。一个真实的拍卖机制可以使竞价的博弈过程更快地达到均衡点, 使各方的收益比较稳定。

文献^[83]已经给出了 Vickrey 拍卖机制真实性的严格证明。本节仅给出真实性的直观解释, 并通过一个例子加以说明。在 Vickrey 拍卖中, 由于竞拍者获胜后实际支付的费用是独立于其出价的, 因此如果竞拍者出价低于其对拍卖品的真实估值, 其赢得拍卖的概率也会降低; 如果出价高于真实估值, 其赢得拍卖的概率虽然会增加, 但是获得的利润却不一定增加, 这是因为竞拍者支付的费用有可能大于其真实估值, 因此对于竞拍者来说, 其最优出价应该等于其对拍卖品的真实私有估值。具体例子如下: 假设某竞拍者对竞拍品的私有估值为 1000, 如果他出价 800(低于真实估值), 其他竞拍者出价 900, 那么他输掉拍卖, 利润为 0; 如果他提升出价为 950, 则他赢得拍卖, 实际付费 900, 获得利润 100; 但是当竞拍者出价高于真实估值时, 例如出价 1200, 其他竞拍者出价 1100, 虽然竞拍者赢得拍卖, 但是他需要按照 1100 支付费用, 实际利润只有一 100, 因此对于竞拍者来说, 出价越接近真实估值, 获得利润的概率越高。

5.1.2 广义第二价格拍卖

由于 Vickrey 拍卖是针对单物品拍卖的, 而在赞助搜索广告中, 通常需要同时对多个广告位进

行拍卖,即多物品拍卖,为此 Google 的研究者在 2012 年针对赞助搜索广告的场景,提出了广义第二价格(Generalized Second Price, GSP)拍卖^[13]. GSP 在短时间内取得了巨大的商业成功,并被运用到展示广告中,Google 2005 年的总收入是 61.4 亿美元,超过 98% 的收入来自于 GSP 拍卖,目前 GSP 已经成功应用于全球大多数的 ADX 中^[13]. GSP 是 Vickrey 拍卖的推广,在赞助搜索广告的 GSP 拍卖中一般有由上至下的多个广告位,因而会产生多个获胜者,获胜者的产生机制与 Vickrey 拍卖相似,按照分数(出价与点击率的乘积)由高到低排序,分数高者得到位置更靠前的广告位,获胜者支付的费用是紧随其后的下一名竞拍者的出价,以此类推.

本文通过一个简化的例子对 GSP 机制进行说明. 假设竞拍品为某网页的 2 个广告位 1 个小时的展示,3 个广告活动参与竞标,按照点击次数(CPC)付费,广告位 1 平均每小时可获得 200 次点击,广告位 2 平均每小时可获得 100 次点击,广告活动 1、2、3 对于每次点击的估值分别是 10 元,4 元,2 元. 假设每个广告活动(竞拍者)都按照真实估值出价,点击率只与广告位有关,因此,在 GSP 中,广告活动 1 赢得广告位 1,为每次点击支付的费用为 4 元,支付费用为 $200 \times 4 = 800$ 元;广告活动 2 赢得广告位 2,为每次点击支付的费用为 2 元,支付的费用为 $100 \times 2 = 200$ 元,因此拍卖方的总收益为 $800 + 200 = 1000$ 元. 需要说明的是,GSP 虽然在 RTB 中广泛应用,是事实上的标准. 但是很多理论研究者甚至包括 google 都认为 GSP 是一种“错误”的拍卖方式^[84],因为它并不是 Vickrey 拍卖的合适推广,理论上 VCG 拍卖是一种比它更合适的选择.

5.1.3 VCG 拍卖

VCG(Vickrey-Clarke-Groves)拍卖是由 William Vickrey,Edward H. Clarke,Theodore Groves 三人共同提出的,作为 Vickrey 拍卖在多物品拍卖中的推广. 已经证明 VCG 拍卖是真实的^[85],且对于按照真实估值出价的竞拍者,在竞拍结束后不会嫉妒其他人的竞拍结果(envy free). VCG 拍卖仍然按照竞拍者的分数(出价乘以点击率)高低来决定获胜者,不同的是,每个竞拍品的获胜者需要支付的费用为他对其他竞拍者造成的损失. 仍然采用 GSP 的例子来说明 VCG 拍卖机制. 在 VCG 中,每个广告活动仍然按照真实估值出价,因此,广告活动 1 赢得广告位 1,广告活动 2 赢得广告位 2. 不同的是,广告活动 1 为广告位 1 支付的费用为: $(200 - 100) \times 4 + 100 \times 2 = 600$ 元,这里 $(200 - 100) \times 4$ 表示如果

广告活动 1 不参加拍卖,广告活动 2 赢得广告位 1 而增加的收益, 100×2 表示广告活动 3 赢得广告位 2 而增加的收益;广告活动 2 为广告位 2 支付的费用为: $100 \times 2 = 200$,表示广告获得 2 如果不参加拍卖,广告活动 3 赢得广告位 2 而增加的价值;因此,拍卖方的总收益为 $600 + 200 = 800$ 元.

5.1.4 三种拍卖的特点分析

纵观三种第二价格拍卖机制,当只有一个拍卖品时,三种拍卖机制是完全等价的;当有多个拍卖品时,GSP 拍卖机制对于广告位提供方或者交易平台是有益的,可以获得高于 VCG 的收益,例如本文的例子,但是对于竞拍者来说却非最优的,相反 VCG 可以使全体竞拍者的收益获得最大化. 在目前的实际 ADX 中,大多数采用的是 GSP 机制,例如腾讯、百度、Google,少数采用的是 VCG 机制,例如 Facebook. 分析原因主要有以下几个方面:(1) GSP 机制能提升交易平台和广告位提供方的收益;(2) GSP 机制更简单,易于向广告商解释付费机制;(3) VCG 机制要求计算获胜者对其他竞拍者造成的损失,这在实际场景中很难准确估计;(4) GSP 已经在 ADX 中广泛部署,重新部署 VCG 代价太大.

5.2 信任危机与第一价格密封拍卖

尽管第二价格密封拍卖受到 Google 等交易平台的大力推广,且已经在工业界获得广泛部署,但是近几年却遭到了来自发布媒体和广告主的巨大质疑^[86-87]. 首先在目前的 RTB 系统中,交易平台不会将每个 DSP 的出价、获胜广告主的支付价格、以及支付给发布媒体(SSP)的价格对外公开,这种不透明操作在第二价格拍卖中为交易平台作弊提供了充足的条件^[86],因为各种费用支付都是以第二高价为基础,而只有 ADX 知道真实的第二高价. 举例来说,ADX 可能并不如实地将自己与获胜广告主的成交价格告知发布媒体,而是通知发布媒体一个仅仅略高于保留价的价格,从而将其中的差价作为自己的“手续费”. 另一种作弊情况是,ADX 对于获胜广告主,也可以不按照真实第二高价收费,而是按照一个仅略低于最高出价的价格收费. 此外,ADX 还可以违规地选择非最高出价的广告主作为获胜者,从而赚取“额外费用”. 这种作为拍卖者的不诚实行为导致 RTB 各参与方存在严重的信任危机^①.

① Nearly Half of Programmatic Impressions Are Sold through First-Price Auctions. <https://www.emarketer.com/content/nearly-half-of-programmatic-impressions-are-sold-through-first-price-auctions>

文献[86]针对拍卖方作弊,提出一种高效、保密的整数比较技术,并设计了一种可验证的 ADX 协议,来对拍卖过程中的各种信息进行审查,以检查作弊行为,并在此基础上实现并部署了一个名为 VEX 的实时交易系统.实验结果显示该系统在可接受的额外网络带宽和时间消耗下,保证了拍卖过程的完整性,但它仍然存在很多局限,例如只能检测已发生的作弊行为,并不能防止作弊,并且不能检测到某些特殊形式的作弊.在此背景下,文献[87]进一步从理论上严格证明了第一价格拍卖是唯一可信任的拍卖机制.产业界更早地反映出对第二价格拍卖机制不透明问题的担忧,从 2017 年开始就不断有交易平台公司(例如 OpenX^①)推动采用第一价格密封拍卖机制,且获得了成功,根据文献[4]的报道,2018 年 3 月采用第一价格拍卖成交的广告库存已经提升到整个采用 RTB 销售的广告库存的 43.3%.因为按照广告主的真实出价收费,确实增加了拍卖过程的透明度,避免了 ADX 对获胜广告主作弊.然而第一价格拍卖机制并不是真实的,有研究^②指出第一价格拍卖易使广告主陷入价格战,而不再仅仅关注广告库存对自己的真实价值,从而使广告主为曝光支付更多的费用,因此设计更“干净的”第二价格拍卖机制才是未来的发展趋势.

5.3 Header 拍卖

为了防止 ADX 作弊,发布媒体更希望由自己来掌握最终的广告库存拍卖权.针对这一需求,Header 拍卖^③被提出.作为最新的拍卖方式,Header 拍卖通过位于网页头部的一条 JavaScript 向各种渠道发送对广告曝光机会的出价请求,并根据收到的报价,确定最终的投放渠道,如图 13 所示.在 Header 拍卖中,发布媒体理论上可以同时请求所有销售渠道报价,因此能够最大程度降低库存流拍的可能性,提升资源配置效率;此外,发布媒体直接获知各种渠道广告主对广告库存的真实估值,能够避免交易平台作弊,增加了市场透明度和公平性.在 Header 拍卖中,也会存在多级拍卖,拍卖机制既可以采用第一价

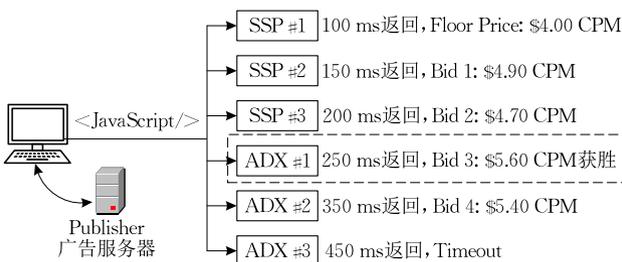


图 13 Header 拍卖示意图

格拍卖,也可以采用第二价格拍卖,但是为了保证市场的透明度,OpenX 建议^[5]在所有的非最终的拍卖中都采用第一价格拍卖,在最终的 Header 拍卖中,采用第二价格拍卖,如图 14 所示.这种同时包含第一价格和第二价格拍卖的机制,能够充分利用两种拍卖的优点,是一种非常具有前景的拍卖方式.

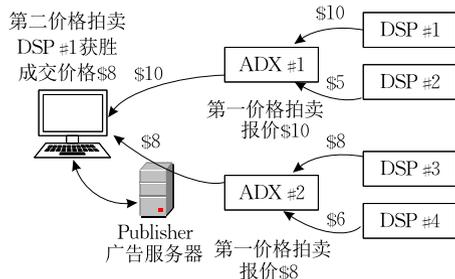


图 14 第一价格和第二价格结合的 Header 拍卖

5.4 拍卖机制的研究和讨论

目前学术界关于拍卖机制的研究主要集中在赞助搜索广告的实时拍卖,专门针对 RTB 拍卖机制的研究还非常少,主要集中在经济学领域,原因是目前 RTB 的拍卖机制是由市场驱动的,公开资料和数据非常少,而早期第二价格拍卖的研究已经相对成熟^[13,84].本节对已有的研究工作做一个简单的整理.文献[11]讨论了在包含多级拍卖的第二价格拍卖中可能存在第二价格非真实的情况并给出了简单的解决方案;文献[88]和[89]讨论了目前广泛使用的 GSP 机制;文献[90]提出了一种新的动态拍卖机制,其基本思想是每次拍卖不是独立进行的,拍卖结果不仅取决于本次竞拍者的出价,而且依赖历史竞拍的结果,从而优化广告库存提供方的收益;文献[86]和[87]则是针对第二价格拍卖导致的交易平台作弊问题展开研究,特别是文献[87]严格证明了第一价格拍卖是唯一可信的静态拍卖机制,为第一价格拍卖在 RTB 市场的推广,奠定了理论基础.相信随着第一价格拍卖在 RTB 市场的进一步推广,将会激发更多对该领域的研究.

6 广告欺诈检测

随着在线广告产业规模的不断扩大,在线广告

- ① Understanding how OpenX determines when to run first-price auctions. <https://docs.openx.com/Content/demand-partners/first-price-auctions.html#understanding-first-price>
- ② Hearts & Science Calls to Replace First-Price Auction with Reformed Second-Price Model. <https://adexchanger.com/online-advertising/hearts-science-calls-to-replace-first-price-auction-with-reformed-second-price-model/>
- ③ What is Header Bidding and How Does it Work. <https://clearcode.cc/blog/what-is-header-bidding/>

已成为目前众多互联网公司的主要收入来源,然而在线广告生态系统的机制和组织都相当复杂,在丰厚利益的驱动下,在线广告的各参与方都有动机进行欺诈活动,以获得额外的收益.实际上,欺诈活动已经成为在线广告产业中的一个大问题,正在破坏这项数百亿美元的产业.根据互联网广告局的报告,在 2015 年广告欺诈活动使得美国媒体产业损失 82 亿美元左右^[91].谷歌在 2016 年对于“广告欺诈”的定义^①是“那些包含展示、点击和转化的无效流量,这些都不是真正的用户兴趣的结果”.直观地说,广告欺诈活动就是广告生态系统中的一个或多个成员为了试图增加他们的收益,而以损害其他成员的利益为代价所采取的欺骗行为,欺诈者可能是 RTB 中的提供方,也可能是需求方.本节对已有的广告欺诈展开研究,首先介绍广告欺诈的主要方式,然后介绍已有的欺诈检测与防范方法.

6.1 广告欺诈方式

目前广告欺诈的主要方式可分为曝光机会欺诈、点击欺诈和转化欺诈三类.其中曝光机会欺诈是指广告展示机会的提供方通过欺骗广告主购买虚假曝光机会来赚取收益,通常存在于以 CPM 或 dCPM 为结算方式的展示广告投放中;点击欺诈是指欺诈者通过雇佣人力或点击机器人等手段对投放到曝光机会中的广告产生虚假点击,以骗取广告主预算,这是目前最主要的广告欺诈方式,在展示广告和赞助搜索广告中都广泛存在;转化欺诈是指欺诈者通过雇佣人力或机器人对特定广告展示,执行一系列有意义的用户业务行为,如将商品加入购物车、填写注册表、下载并安装 app 等,以产生虚假转化.下面分别介绍这三类广告欺诈方式的研究进展.

6.1.1 曝光机会欺诈

曝光机会欺诈(impression spam)的实质是发布媒体(网站)作为欺诈者通过提供大量虚假广告展示机会来骗取广告主的预算.目前产生虚假广告展示机会的方式主要有两种.第一种方式是通过对网页进行处理,例如在网页上堆叠大量广告位^[92],使用户浏览网页时产生大量曝光机会,或者在网页上添加大量“高价值”但是与网页内容无关的关键词或内容标签^[93],以提升广告主对这些曝光机会的估值,并吸引更多的广告主参与竞拍.在这种方式中,大量广告和无关关键词展示会损害用户体验,因此通常这些虚假广告位、虚假关键词被设计为对用户不可见.另一方面,曝光机会的产生需要用户浏览行为来触发,因此第二种产生曝光机会欺诈的方式是

通过开发欺诈机器人来模拟用户对网页的浏览行为,以触发更多的曝光机会.通过上述两种方式产生的曝光机会并不会将广告展示给真正感兴趣的用户,因此对于广告主来说是无效展示机会.由于曝光机会欺诈能为发布媒体带来巨额收益,目前互联网上已经出现了大量以从 RTB 赚取广告费为目的的虚假网站(Fake Site)^[94].这些网站通常具有合法域名,其内容通常是从其他网站复制的或者无意义的,网页上堆叠了大量广告位及高价值关键词,通过加入各种 RTB 交易平台,销售自己的虚假曝光机会.

6.1.2 点击欺诈

点击欺诈(Click Fraud)是赞助搜索广告中最主要的广告欺诈方式,在赞助搜索领域已经有大量的研究^[93].在展示广告中点击欺诈的方式与之相似,根据点击欺诈行为的意图不同,主要分为两类.第一类是由发布媒体(网站)发起的,目的是人为扩充目标网站曝光机会的点击量(click inflation);第二类是由广告主发起的竞争对手点击(competitor click)^[93],目的是广告主通过对其竞争对手的广告进行虚假点击,以尽快耗尽竞争对手的广告预算.点击欺诈在以 CPC 为主要计费模式的赞助搜索中被大量采用,在以 dCPM 为主要计费模式的 RTB 系统中,点击欺诈也能帮助发布媒体提升 DSP 对其曝光机会的估计价值(预测 CTR),从而提升广告主对曝光机会的出价.

点击欺诈的主要实现方式分为点击农场(Click Farm)和点击机器人(Click Bots)两类.点击农场是指欺诈者以低成本雇佣大量人力来点击广告,虽然这些点击是由真实人类产生的,但他们并不是可能产生转化的正常用户,带有恶意的欺诈意图.点击机器人^[93]是指由欺诈者开发部署的恶意软件,该软件可以根据配置自动化地实施不同类型的欺诈行为,例如浏览指定的网站网页,点击网页上的指定广告等,通常这些点击机器人被部署在一个或多个静态主机上,还可以改变 IP 地址以躲避检测.部署在单台或多台主机上的点击机器人进一步演变为部署在僵尸网络上(Botnet Click Bots)的机器人.作为目前互联网点击欺诈的主要来源,僵尸网络^[95]有其独特之处,攻击者首先入侵正常用户的计算机,安装恶意软件,形成欺诈的基础——僵尸网络;恶意软件在受感染的用户机器上自动、隐秘地运行,时刻监听并执

① Google Ads traffic quality resource center. <http://www.google.com/ads/adtrafficquality/index.html>

行欺诈者所发出的命令,例如在后台运行去访问特定网站并点击特定的广告^[96].这种欺诈行为类似于分布式拒绝服务(DDoS)攻击^[97],它使用分布式的机器产生大量流量,同时使得IP地址和地理位置多样化,虚假流量隐藏于正常流量中,更加不易被发现.即使被发现,由于分布式攻击的特点,想要完全消除其影响也非常具有挑战性^[98].

在点击欺诈中,除了主动的虚假点击,还有一种强制(Coercion)点击行为,例如网站要求用户在获得期望的内容之前必须点击网页上的广告,或者故意将网站实际内容与广告内容混淆,诱骗用户点击广告,而这些点击行为也并不代表用户真实意图.

6.1.3 转化欺诈

在转化欺诈(Conversion Fraud)^[93]中,欺诈者通常需要完成一些特定的动作,例如将商品加入购物车、填写注册表、下载并安装app等,以产生虚假转化.与点击欺诈相同,转化欺诈的发起者既可以是发布媒体,也可以是广告主.其实施方式也与点击欺诈相似,可以雇佣低成本人力或开发机器人来实现.通常转化欺诈只在转化行为免费的情况下才能实施.

6.1.4 其他欺诈方式

除了上述三类欺诈方式,近期有研究者提出一种针对重定向广告的欺诈方式,称为重定向欺诈(Re-targeting Fraud).所谓重定向广告是指广告系统利用重定向技术分析用户在网站的历史行为(例如浏览和购买记录),为广告主的潜在客户建立用户画像,从而帮助广告主将广告投放到那些最有可能产生收益的曝光机会上.在文献^[99]中研究者提出了一种利用重定向技术的针对广告主的广告预算流失攻击.首先,攻击者使用基于黑盒测试的技术逆向推断目标广告主的定向策略,并基于定向策略构建用户画像,然后通过机器人模拟定向客户的行为,以吸引广告主购买由机器人触发的曝光机会.攻击者通过在多台机器上大量部署机器人程序,反复诱骗广告主购买虚假曝光机会,从而导致广告主的预算被很快耗尽,如图15所示.文中的模拟攻击实验表明,对于选出的291个公开广告主,这种攻击方式成功地揭示了其中254名广告主的定向策略,并在一小时内骗取到了4万多个广告,流失的广告预算达上百美元.

此外,广告注入(Ad Injection)是另一种由第三方发起的具有侵略性的广告欺诈行为,它通过篡改HTTP网页内容来实现^[100].例如,互联网服务提

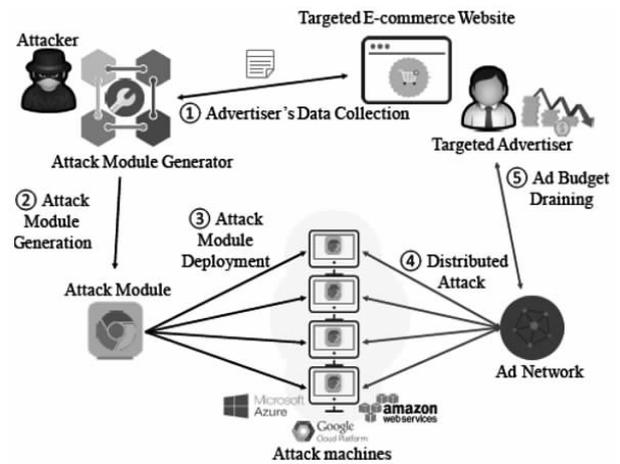


图15 针对广告主的广告预算流失攻击

供商和WiFi提供商都有可能篡改网络传输中的HTTP内容来注入广告^①.运行在客户端的恶意广告软件或浏览器插件也可以通过修改网页代码来达到广告注入的目的.

6.2 广告欺诈的检测与防范方法

针对各种广告欺诈方式,本文对已有的广告欺诈检测方法进行了整理,总结如下.

6.2.1 基本方法

首先介绍三种通用的欺诈检测方法,这三种基本方法既可以用于检测曝光机会欺诈,也可以用于检测点击欺诈和转化欺诈^[96].

基于规则的检测方法(Rule-based Detection).该方法^[100]使用静态规则匹配来确定哪些广告流量是有效的,哪些广告流量是恶意的欺诈流量.这些规则是从已确定的欺诈流量中提取的,例如一个通用规则是在连续双击广告的行为中,可以认为第二次点击是无效的^[94].这种检测有助于通过已知的欺诈模式来发现已知的欺诈流量,但对于模式未知或者不遵循静态规则的欺诈,则不奏效,并且具有滞后性.

基于异常的检测方法(Anomaly-based Detection).该方法^[101]使用有关发布媒体的历史信息来发现广告流量模式的异常变化.例如当发布媒体的曝光机会数量或曝光机会质量(CTR)突然升高时,这种异常变化意味着发布媒体可能正在进行欺诈行为.异常检测对于发现发布媒体的欺诈行为或者在发布媒体更改他们正在实施的欺诈类型时非常有用^[94,101].

基于分类器的检测方法(Classifier-based Detec-

① Comcast Wi-Fi serving self-promotional ads via JavaScript injection. <https://arstechnica.com/tech-policy/2014/09/why-comcasts-javascript-ad-injections-threaten-security-net-neutrality/>

tion). 该方法根据已知的发布媒体的历史流量信息,即知道哪些流量是有效的,哪些流量是无效的,构建出带标签的数据集,然后利用机器学习模型建立流量有效/无效的二分类器. 该方法主要被用于点击欺诈检测,也可以扩展到曝光机会欺诈和转化欺诈的检测.

6.2.2 曝光机会欺诈的检测方法

曝光机会欺诈主要出现在展示广告中,针对曝光机会欺诈的研究主要集中在对发布媒体方的欺诈检测,例如是否存在广告位堆叠和关键词堆砌,网站是否为虚假网站等. 文献[102]针对广告位堆叠问题,利用其通常对用户不可见的特性,提出两种检测方法,第一种方法通过将广告和从网页代码中提取的图形图像与此网页的快照进行对比,通过图像分析技术,任何在网页快照中没有出现的展示广告都被标记为对用户不可见,即可能被其他广告遮挡或广告位尺寸被设置为 0×0 ;第二种方法是基于几何分析,为了检查广告是否可见,利用嵌入到网页中的代码片段计算广告位的位置、浏览器的可见区域和打开的浏览器窗口的大小等数据,以判断一个展示机会是否是用户可见的. 文献[103]直接针对虚假网站展开研究,基于支持向量机模型 SVM,提出一种检测虚假网站的分类器. 该分类器基于大量特征,包括网站正文、HTML 设计、图像、链接等. 给定一个网站,该网站中的每个页面可以用两个向量(即最大相似度向量和平均相似度向量)来表示,这两个向量是在网站网页和目标页面之间计算出来的,然后将两个向量的内积作为 SVM 分类器的输入,进行训练与分类,最后虚假网站可以通过检测页面的数量以及虚假页面的数量和所占百分比来确定.

此外,文献[96]提出还可以基于信誉度来检测一个网站是否存在曝光机会欺诈,即将网站产生的曝光机会数与已知可信排名网站(例如 Alexa)上该网站的信息进行对比,如果实际检测到的曝光机会数量远远超过网站排名所显示的流量,则表明该网站可能存在曝光机会欺诈.

6.2.3 点击欺诈的检测方法

点击欺诈检测方法的研究一直是广告欺诈检测研究的热点,吸引了大量研究者的关注. 早期点击欺诈的检测方法通常是基于数据挖掘算法. 例如 Metwally 等人在文献[104]中提出可以通过挖掘点击流中的频繁模式来建立欺诈检测的关联规则;在文献[105]中提出可以利用布隆过滤器(Bloom Filters)来检测点击流中的重复点击. 后期的研究主

要是利用机器学习模型建立二分类器来检测欺诈点击. 例如在文献[106]中,首先通过主成分分析、公共空间模式等方法来进行特征选择;然后采用过采样算法 SMOTE 来解决数据集中的类不平衡问题;在此基础上,利用决策树建立多个基分类器;最后使用集成学习(ensemble learning)理论将多个基分类器的结果进行结合,得到最终的分类结果.

此外,也有一些从不同角度对欺诈点击检测的研究. 文献[107]从交易平台角度,提出一种基于蜜罐的假验广告(Bluff Ads)方法. 交易平台通过将 Bluff Ads 提供给发布媒体展示,以探测发布媒体是否存在欺诈活动. 具体来说,当发布媒体向 ADX 发起广告展示请求的拍卖时,ADX 按照一定比例将 Bluff Ads 作为获胜广告返回给发布媒体展示,而这些 Bluff Ads 被故意设计为不那么吸引人(例如只有一幅图片,没有文字来吸引用户关注并点击). 如果这些假验广告的点击率或转化率没有比正常展示广告低很多,则说明在发布媒体方可能执行了欺诈点击. 文献[108]从发布媒体角度,提出三种检测是否存在欺诈点击的方法. 第一种是要求用户执行一段 JavaScript 代码和触发鼠标事件来证明其是真实用户;第二种是要求用户执行浏览器特征检测,即网站服务器发送包含真实特征和虚假特征的集合,要求用户向网站服务器回复集合中有多少真实的浏览器特征,通常点击机器人无法通过检测,因为它无法测试出集合中包含的虚假特征;对于通过前两种检测的高级点击机器人或者人类欺诈者,作者进一步提出第三种检测方法,即检测用户的浏览行为是否符合正常用户的行为模式,通常欺诈者更倾向于快速浏览广告网站后进行下一次点击任务.

文献[109]则是从用户角度,检测用户主机是否成为僵尸网络的节点并在后台执行自动点击任务,它包括检查和分析两个模块. 检查模块从 HTTP 报文中收集数据,从每个进程中捕获输入事件信息,并用广告屏蔽过滤器来筛选出与广告相关的进程;分析模块为每个进程创建 HTTP 请求树,在 HTTP 请求树中,把对 HTTP 页面的请求作为根节点,对静态资源的请求作为子节点,然后使用机器学习来对 HTTP 请求树中的广告请求进行分类,以识别出广告请求节点,并将此节点下的子节点标记为广告点击,如果发现进程在后台运行时使用软件模拟点击或者使用 HTTP 请求来完成广告点击,而非人为的点击行为,则将对对应进程标记为欺诈,通知用户并阻止该进程继续访问网络.

随着检测技术的发展,点击机器人技术也在不断发展,最新的点击机器人可以模拟正常用户的点击行为,以避免被检测.文献[110]针对这种情况展开研究,发现正常用户产生的点击流是完全随机的,而模拟点击机器人产生的点击流会与历史点击流在分布上存在更高的相关性,据此作者提出了两种检测方法:模仿(mimicry)和诱饵攻击(bait-click).在模仿检测方法中,首先使用压缩函数将点击时间序列划分为若干基本模式,然后通过聚类算法将点击分布相似的基本模式进行聚类,如果存在与历史点击分布高度相关的基本模式,则认为该点击序列中存在点击欺诈.诱饵攻击是一种主动检测方法,它通过浏览器执行的客户端脚本向用户设备注入有标记的点击模式,如果点击机器人模拟这种点击模式来执行欺诈点击,将使得这类欺诈点击行为很容易被检测出.

6.2.4 转化欺诈的检测方法

针对转化欺诈的预防方法,广告主可以使用基于回报的结算方式(Performance-based Pricing)^[111]来有效避免.使用这种按回报定价的付费方案,不再以展示次数或点击次数为基础,而是以广告商从提供方获得的投资回报比(ROI)为基础,广告主只为那些真正产生了回报收益的展示机会进行付费,这也符合广告主投放广告的最终目的.例如,某广告主将转化行为定义为“将商品加入购物车”,欺诈者可以无支出成本地进行转化行为,若转化行为被定义为“将商品加入购物车并付款购买”,则可以有效保证自身实际收益,避免欺诈者.但在实际中,目前许多网站为了自身利益并不接受这种模式.

6.2.5 广告注入的检测方法

针对广告注入欺诈,由于广告注入主要是通过修改网页代码或 HTTP 传输内容来实现的,Thomas 等人利用谷歌网站开发了一种基于网页文档对象模型(Document Object Model, DOM)的广告注入脚本检测方法^[100].DOM 是描述网页文档结构的编程接口,可以通过它来访问或修改网页的结构、样式和内容.作者首先扫描 DOM 以识别其中的可疑元素和被修改的区域,然后通过排除浏览器扩展插件和反病毒引擎等常规程序,过滤不是由广告注入引起的脚本代码,最后根据脚本内容手动检查发现广告注入欺诈.

综上所述,本节对一些已知的欺诈行为和检测防范方法进行了介绍,虽然已有不少研究者正对该问题进行研究,但实际中在线广告中的欺诈现象仍

然广泛存在.“反欺诈”仍有很长的路要走,单纯依靠某一种手段来完全阻止欺诈活动是很困难且不现实的,需要事前预防、事后追溯、实时监控、智能算法挖掘新型欺诈模式等多管齐下,才能与欺诈行为相抗衡.更重要的是,在线广告中的各参与方要一起努力,打造一个透明、健康、互利的生态环境,才能促进产业的持续良性发展.

7 未来的研究趋势及方向

RTB 自诞生以来,无论在工业界还是在学术界都吸引了大量的关注,目前已经发展成为展示广告的一种主要的投放渠道,一些传统采用合约担保模式的展示广告,例如视频和横幅广告都在逐渐转向 RTB 模式.可预期地,随着人们生活方式进一步地“互联网化”,展示广告流量也将获得大幅提升,而 RTB 作为近年来最成功的广告投放模式也必然会继续受到工业界和学术界的关注.本节从科学技术和业界发展的角度对其中关键问题的研究趋势和发展方向展开评述.表 4 展示了关键问题所涉及的主要研究领域.

表 4 关键问题涉及的研究领域

关键问题	研究领域
用户响应预测(hot)	机器学习
出价策略(hot)	拍卖理论、博弈论、最优化理论
预算与步进管理	最优化理论
保留价优化	拍卖理论、博弈论、最优化理论
库存分配	最优化理论
拍卖机制(hot)	博弈论、拍卖理论
欺诈检测	网络安全、机器学习

7.1 用户响应预测

用户响应预测是目前 RTB 中最活跃的研究方向.作为在线广告、推荐系统的关键部件,用户响应预测已经在相关产业产生了巨大的经济价值,为了促进学术界对该方向的研究,各著名在线广告公司及广告投放平台(例如 Criteo、iPinYou、腾讯)纷纷开放公开数据集或发起相关的点击率(转化率)预测竞赛,为该方向的研究提供了数据基础.另一方面,用户响应预测主要涉及机器学习理论,而近年来深度神经网络的巨大进展使得利用深度学习模型进行预测成为可能.已有的研究结果显示基于深度神经网络的用户响应预测模型能够取得比传统浅层结构模型更优的性能.然而简单的前馈神经网络仅能探索高阶特征的交互信息,而样本的低阶特征对于提升预测性能也非常重要,因此该方向研究的一

一个重要趋势是设计能够同时兼顾低阶和高阶特征信息的融合模型,并利用自动特征工程来代替手工特征,如图 16 所示,其中选择哪些模型进行融合以及如何融合将是研究的重点,可以借鉴集成学习中“好而不同”的研究思路。

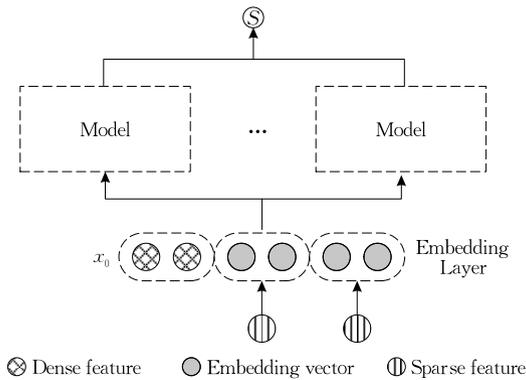


图 16 CTR 预测的融合模型架构

用户响应预测研究的另一个值得关注的方向是解决样本正负比例严重不平衡问题。这是因为在实际应用中,相比于广告展示次数,用户发生的点击/转化行为是非常稀少的,在目前已经公开的展示点击(转化)日志中,正负样本比例非常低^[10](不足 0.1%),使用类不平衡的数据集进行模型学习,很容易对数量多的大类样本产生过拟合,无法对数量少的小类样本(正样本)进行充分地学习。目前的解决方案通常是采用负采样,使得类比例相对平衡;或者使用代价敏感学习,对于分类错误的小类样本给予较大的惩罚权重;同时使用 AUC、ROC 曲线作为模型性能的评估标准,而不是准确率,可以取得更好的泛化性能。然而,利用负采样以达到正负样本比例平衡,可能会破坏真实数据中的样本特征分布,导致训练数据与真实数据的分布出现偏差,从而降低模型的预测性能。因此,未来的研究将更多集中在提升预测模型的泛化能力,以及根据偏差对预测值进行校正方面。在最新的研究中,也有研究者提出了新的解决思路。例如在文献[112]中,作者提出利用伪装对抗网络(Disguise Adversarial Networks)在小类样本上进行对抗学习训练,进而有效区分小类样本,解决数据集中类别不平衡的问题。

综上,用户响应预测的性能仍然有非常大的提升空间,随着越来越多不同场景的点击率(转化率)数据集公开,以及新的基于深度学习的融合模型提出,该方向的研究将会吸引更多的来自学术界和工业界的关注。同时,该方向的研究也为利用大数据预测小概率事件提供了理论模型和研究方法。

7.2 出价策略

出价策略研究也是目前 RTB 研究的热点之一,其目标主要从广告商的角度出发,希望最大化每个广告商的收益或利润。好的出价策略能够帮助广告商更有效的花费预算,从而吸引更多的广告营销活动,为整个在线广告产业带来更大的收益。早期出价策略的研究包括利用最优拍卖和博弈论等经济学理论对出价行为和策略进行分析,利用最优化理论以及控制论的方法对出价策略进行设计。随着强化学习的提出,将强化学习应用到出价策略设计上已经成为出价策略研究的必然趋势。其基本思路是将广告活动对每个广告展示机会的出价作为一个动态决策过程,每次出价不仅与广告展示机会的预测点击率有关,还与当前的可用预算和后续时段的广告展示机会有关,如文献[55,58,60]。已有的研究结果证明基于深度神经网络的强化出价策略通过奖励和惩罚机制,确实能够有效学习历史投标数据中的出价行为,使累积奖励最大化;但是对于一个新的广告投放周期,由于拍卖市场的高度动态性(整个广告投放周期的曝光机会数量和参与竞争的广告活动都可能发生较大的变化),导致基于历史数据学习的最优出价策略对于新的投放周期可能失效,因此如何将动态调整机制融入到强化出价策略的研究中将是该方向的研究趋势之一。最后,广告商的目的是希望目标用户通过广告产生转化行为而不仅仅是点击行为,因此有必要对用户的转化行为与广告展示或点击的关系进行归因分析(Attribution Analysis),对于那些归因概率高的广告展示机会应该提高竞标价格^[113]。受阻于缺乏相关数据集,目前关于转化行为归因分析的研究才刚起步,相信未来将会有更多这方面的研究出现。

综上,出价策略的研究趋势将集中在:(1)基于深度神经网络的强化出价策略设计;(2)结合动态自适应调整机制的强化出价策略设计;(3)结合归因分析模型的出价策略设计。

7.3 预算与步进管理

目前与预算与步进管理相关的研究工作较少,这是因为文献[68]提出的按时段进行预算分配和动态调整的策略已经能保证通过预算的平滑花费实现广告的平滑投放。与已有研究观点不同的是,本文作者认为预算管理与出价策略应该是密不可分的,在整个广告投放周期,无论时段的预算是否充足,采用相同出价策略的思路会导致非最优的结果。一个更合理的研究思路是,将预算分配与出价策略关联起

来,研究与当前可用预算相关的动态出价策略,这也是目前强化出价策略的思路。此外,目前每个时段的预算分配通常只考虑了广告展示机会数量和质量,忽略了该时段对这类广告展示机会的竞争程度,相信为那些具有更多的高性价比展示机会的时段分配更高的预算,将取得更为理想的结果。

7.4 保留价优化与库存分配

目前关于保留价优化的解决方案中,通常假设出价者对展示机会的估值独立地来自于同一个价格分布,并在此基础上计算得到一个合适的保留价。但在实际中由于市场的复杂性和动态性,该假设往往不能成立,并且出价价格分布也很难准确估计,在竞价愿景预测中也存在着相似的问题,因而目前的大部分方案存在较大的局限性。同时实际市场中可能存在多级拍卖的现象,使用固定的保留价往往不能在复杂市场中得到最优的收益。因此,脱离价格分布估计的、动态的保留价设置方法是更理想的。使用无模型的强化学习设置合适的保留价是一个值得探索的方向,通过将广告相关特征建模为环境,可以在每次拍卖后学习、调整保留价,从而最大化累积收益。

另一方面,对于发布者的库存分配问题,传统的做法是将优质的价值高的库存通过合约方式售出,未被售出的低质量库存再通过 RTB 方式售卖,但 RTB 的高效、准确和回报正在吸引着越来越多的库存。一次广告展示机会在产生时,面临着多个离线合约渠道和 RTB 渠道的选择(每个渠道有不同的期望收益的惩罚风险),并且在实际中,发布者的展示机会供应量、每个渠道的展示机会价格是不断变化的,因而这可以自然地构建为一个有约束的动态优化问题,可以使用变分法、动态规划、控制理论等方法来计算得到一个库存分配方案。

总体而言,目前关于保留价优化与库存分配的研究相对较少。部分原因是,目前缺少公开可用的 RTB 实际场景下的保留价数据集,使得研究分析无法展开。另一方面,市场实践中的拍卖机制出现了新变化,随着 header 拍卖和第一价格拍卖的兴起,在新的拍卖环境下保留价优化和库存分配问题的解决方案需要重新考虑,亟待更多的研究工作。

7.5 拍卖机制与 RTB 系统

拍卖机制对于整个 RTB 的稳定性、各方收益等都有较大影响。目前关于拍卖机制的学术研究更多来自于经济学领域,主要从博弈论和拍卖理论的角度对市场采用的拍卖机制的性能进行理论分析。但在市场实践中,主要还是由工业界来决定所采用的

拍卖机制,例如 Google 提出并推广的 GSP 机制,即使后来在理论上证明该机制存在缺陷,但由于技术、成本等原因,并未广泛采用理论上更优的 VCG 机制。相比之下,能够拿到第一手数据的工业界对于拍卖机制的实证研究一直非常活跃,并引领市场实践中拍卖机制的演化。

目前的 RTB 市场交易机制还远非是最优的,仍然有较大的创新空间。一方面,目前被广泛采用的 GSP 拍卖机制被证明是非真实的,可能导致收益是不稳定的。另一方面,RTB 市场存在严重的信息不对称问题,拍卖过程对于发布媒体(网址)和广告主是不透明的,容易产生作弊行为,并且由于缺少相关信息导致广告主不能准确估计曝光机会的价值,导致高质量的曝光机会不能获得更高的销售价格,进而产生逆向选择,迫使优质发布媒体离开 RTB 市场。从 2017 年开始,就不断有公司提出交易平台存在不透明操作的问题,并着力推动第一价格密封拍卖机制的部署。由于第一价格拍卖机制可以提高拍卖过程的透明度,它得到了各著名 DSP 公司和交易平台公司的广泛支持,已经成为目前 RTB 系统中重要的交易机制。同样出于拍卖过程不透明的原因,发布媒体为了保证自己的收益,越来越多地使用 header 拍卖,它可以使发布媒体接触到更多潜在合作伙伴。但第一价格拍卖和 header 拍卖也并非完美的,在第一价格拍卖机制中不存在均衡状态,因此出价者的出价以及 ADX 的收益可能会出现剧烈波动;而 header 拍卖由于在网页头部嵌入过多代码,可能影响页面的加载速度和用户体验,并且 header 拍卖为需求方收集用户数据提供了途径,这会对用户隐私安全产生损害。因此,本文认为设计更加透明、高效、公开、真实的拍卖机制将是 RTB 拍卖机制研究的必然趋势。

另一方面,从市场结构的角度分析,一些优质的发布媒体(如新浪、网易等门户网站)为了保证自己的广告位以高价值顺利售出,建立私有广告交易市场(Private Marketplace, PMP)渐渐成为趋势。私有广告市场中只存在特定的发布媒体和广告主,可以使用 header 拍卖进行优先交易,优质的广告主可以提早锁定优质的广告资源。对于广告主而言,与开放广告交易市场相比,使用 PMP 可以避开低质量流量、虚假流量,并且展示机会价格一般稳定,市场竞争程度有限;虽然竞价价格稍高,但同时也能获得有保证的投放回报。总体而言,RTB 市场中广告主的需求是多样的,虽然有越来越多的私有广告交

易市场出现,但目前私有市场与公开市场是共存的,两者相互影响,相互促进;也许未来公开市场可能引入流量质量分级的制度来增加市场透明度,而私有市场也会扩大市场规模以纳入更多的发布媒体和广告主。

7.6 新理论在 RTB 研究中应用

迁移学习和强化学习是近期在机器学习领域兴起的两个重要理论,本节讨论这两种新理论在 RTB 研究中的一些有益的探索。

迁移学习(Transfer Learning)的目标是把从一个或多个源领域的任务中学习到的知识、经验,应用中另一个目标领域,帮助目标领域的任务取得更好的效果。根据文献[114],迁移学习可以大致分为四类:基于实例的迁移学习、基于特征表示的迁移学习、基于参数的迁移学习、基于关系知识迁移学习。目前在 RTB 中迁移学习并没有得到广泛的关注,但针对其中新广告的冷启动问题,迁移学习是一个值得尝试的方向,通过把与新广告属性相似的旧广告上学习到的知识,运用到新广告的预测中,可以有效提高预测准确率^[115],缓解冷启动问题。另一方面,一些深度神经网络模型虽然学习能力强,但是其参数量巨大且学习耗时,针对这一问题,可以将已训练好的神经网络模型的部分参数,迁移到新任务的神经网络模型中,这样可以共享某些通用的任务知识,帮助新模型更快地收敛。

强化学习(Reinforcement Learning)是近来 RTB 出价策略研究中新出现的趋势。文献[55]的作者尝试把强化学习应用到 RTB 的出价算法中,不同于大多数出价策略使用一个静态的出价函数,他们使用马尔可夫决策过程来对 RTB 进行建模,取得了不错的效果。基于同样的思路,可以尝试利用强化学习来优化预算分配(预算步进算法),或者利用强化学习来优化 CTR 预测,以获取最大化点击次数作为奖励来进行反馈、调整,达到最优的 CTR/CTR 预测效果。强化学习与有监督学习相比,它可以使用无标签数据进行训练,并且动态地与环境进行交互,去优化算法性能。

8 结束语

近十年,在线广告在产业界取得的巨大成功吸引了来自学术界和工业界大量的研究人员关注。RTB 作为目前在线广告最主要的投放模式,相比传统的担保模式具有更高的资源配置效率。本论文对 RTB

的最新研究进行了调研,并分别从需求方平台、提供方平台、交易平台三个主要参与平台的角度,梳理其中的关键问题,以及目前已有的研究方法、理论和模型,特别是对目前热度最高的用户响应预测及出价策略的最新研究进行了详细地介绍和分析。希望通过本文的介绍,对计算广告领域感兴趣的研究者和工程师能够更多地了解 RTB 的相关知识和研究进展,促进这一研究方向及其相关研究的发展。

参 考 文 献

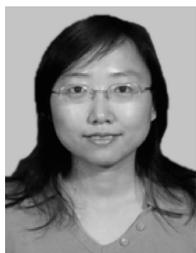
- [1] Liu Peng, Wang Chao. Computational Advertising: Market and Technology of Internet Business Realization. Beijing: The People's Posts and Telecommunications Press, 2015 (in Chinese)
(刘鹏,王超. 计算广告: 互联网商业变现的市场与技术. 北京: 人民邮电出版社, 2015)
- [2] Fain D, Pedersen J. Sponsored search: A brief history. Bulletin of the American Society for Information Science & Technology, 2006, 32(2): 12-13
- [3] Wang J, Zhang W, Yuan S. Display advertising with real-time bidding (RTB) and behavioral targeting. Foundations & Trends in Information Retrieval, 2017, 11: 4-5
- [4] Ghose A, Yang S. An empirical analysis of search engine advertising: Sponsored search in electronic markets. Management Science, 2009, 55(10): 1605-1622
- [5] Ye S, Aydin G, Hu S. Sponsored search marketing: Dynamic pricing and advertising for an online retailer. Management Science, 2015, 61(6): 1255-1274
- [6] Animesh A, Ramachandran V, Viswanathan S. An empirical investigation of the performance of online sponsored search markets//Proceedings of the International Conference on Electronic Commerce. New York, USA, 2007: 153-160
- [7] Yuan S, Wang J, Zhao X. Real-time bidding for online advertising: Measurement and analysis//Proceedings of the 7th International Workshop on Data Mining for Online Advertising. New York, USA, 2013: 19-26
- [8] Zhang H, Zhang W, Rong Y, et al. Managing risk of bidding in display advertising//Proceedings of the 10th ACM International Conference on Web Search and Data Mining. Cambridge, UK, 2017: 581-590
- [9] Shen J, Orten B, Geyik S C, et al. A. From 0.5 million to 2.5 million: Efficiently scaling up real-time bidding//Proceedings of the IEEE International Conference on Data Mining. Atlantic, USA, 2015: 973-978
- [10] Zhang W, Yuan S, Wang J. Real-Time Bidding Benchmarking with iPinYou Dataset. London: University College London, UCL Technical Report; 1, 2014
- [11] Muthukrishnan S. Ad exchanges: Research issues//Proceedings of the 5th International Workshop on Internet and Network Economics. Rome, Italy, 2009: 1-12

- [12] Yuan Y, Wang F, Li J, et al. A survey on real time bidding advertising//Proceedings of the IEEE International Conference on Service Operations & Logistics, and Informatics. Qingdao, China, 2014: 418-423
- [13] Benjamin E, Michael O, Michael S. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 2007, 97(1): 242-259
- [14] Chen Y, Berkhin P, Anderson B, et al. Real-time bidding algorithms for performance-based display ad allocation//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. San Diego, USA, 2011: 1307-1315
- [15] Zhang W, Rong Y, Wang J, et al. Feedback control of real-time display advertising//Proceedings of the 9th ACM International Conference on Web Search and Data Mining. San Francisco, USA, 2016: 407-416
- [16] Zhang W, Yuan S, Wang J. Optimal real-time bidding for display advertising//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 1077-1086
- [17] Matthew R, Dominowska E, Ragno R. Predicting clicks: Estimating the click-through rate for new ads//Proceedings of the International Conference of World Wide Web. Banff Alberta, Canada, 2007: 8-12
- [18] Steffen R. Factorization machines with libFM. *ACM Transactions on Intelligent Systems and Technology*, 2012, 3(3): 1-22
- [19] Juan Y, Zhuang Y, Chin W S, et al. Field-aware factorization machines for CTR prediction//Proceedings of the ACM Conference on Recommender Systems. Boston, USA, 2016: 43-50
- [20] Zhang W, Du T, Wang J. Deep learning over multi-field categorical data: A case study on user response prediction//Proceedings of the European Conference on Information Retrieval. Padua, Italy, 2016: 45-57
- [21] Qu Y, Cai H, Ren K, et al. Product-based neural networks for user response prediction//Proceedings of the IEEE International Conference on Data Mining. Barcelona, Spain, 2016: 1-6
- [22] Cheng H, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems//Proceedings of the Workshop on Deep Learning for Recommender Systems. Boston, USA, 2016: 1-4
- [23] Chapalle O, Manavoglu E, Rosales R. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology*, 2014, 5(4): 1-34
- [24] Yang H, Lu Q, Qiu A, Han C. Large scale CVR prediction through dynamic transfer learning of global and local features. *Proceedings of Machine Learning Research*, 2016, 53: 103-119
- [25] Lee K C, Orten B, Dasdan A, et al. Estimating conversion rate in display advertising from past performance data//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Beijing, China, 2012: 768-776
- [26] Ketkar N. *Deep Learning with Python*. Berkeley, USA: Apress, 2017
- [27] Liu Meng-Juan, Zeng Gui-Chuan, Yue Wei, et al. Review on click-through rate prediction models for display advertising. *Computer Science*, 2019, 46(7): 38-49(in Chinese)
(刘梦娟, 曾贵川, 岳威等. 面向展示广告的点击率预测模型综述. *计算机科学*, 2019, 46(7): 38-49)
- [28] Dalessandro B, Chen D, Raeder T, et al. Scalable hands-free transfer learning for online advertising//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 1573-1582
- [29] Shioji E, Arai M. Neural feature embedding for user response prediction in real-time bidding (RTB)//Proceedings of the Workshop on Social Media for Personalization and Search. Aberdeen, UK, 2017: 8-13
- [30] Yang H. Bayesian heteroscedastic matrix factorization for conversion rate prediction//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. Singapore, 2017: 2407-2410
- [31] Pan J, Xu J, Ruiz A L, et al. Field-weighted factorization machines for click-through rate prediction in display advertising//Proceedings of the 2018 World Wide Web Conference. Lyon, France, 2018: 1349-1357
- [32] Xiao J, Ye H, He X. Attentional factorization machines: Learning the weight of feature interactions via attention networks//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 3119-3125
- [33] He X, Pan J, Jin O, et al. Practical lessons from predicting clicks on ads at facebook//Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 1-9
- [34] Gai K, Zhu X, Li H, et al. Learning piece-wise linear models from large scale data for ad click prediction. *arXiv preprint arXiv: 1704.05194*, 2017
- [35] Agarwal A, Chapelle O, Dudik M, et al. A reliable effective terascale linear learning system. *Journal of Machine Learning Research*, 2014, 15(1): 1111-1133
- [36] Norris C M, Ghali W A, Saunders L D, et al. Ordinal regression model and the linear regression model were superior to the logistic regression models. *Journal of Clinical Epidemiology*, 2006, 59(5): 448-456
- [37] Ma Xiao-Shuan. *Research and Application of Matrix Factorization in Recommender Systems* [M.S. dissertation]. University of Electronic Science and Technology of China, Chengdu, 2017(in Chinese)

- (马小栓. 矩阵分解在推荐系统中的研究与应用[硕士学位论文]. 电子科技大学, 成都, 2017)
- [38] Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks// *Proceedings of the Fourth ACM Conference on Recommender Systems*. Barcelona, Spain, 2010: 135-142
- [39] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. *Computer*, 2009, 42(8): 30-37
- [40] Shan Li-Li. Response Prediction on Real Time Bidding Via Tensor Factorization [Ph. D. dissertation]. Harbin Institute of Technology, Harbin, 2016(in Chinese)
(单丽莉. 基于张量分解的实时竞价广告响应预测方法[博士学位论文]. 哈尔滨工业大学, 哈尔滨, 2016)
- [41] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks//*Proceedings of the International Conference on Neural Information Processing Systems*. Lake Tahoe, USA, 2012: 1097-1105
- [42] Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Vancouver, Canada, 2013: 6645-6649
- [43] Shen Y, He X, Gao J, et al. A latent semantic model with convolutional-pooling structure for information retrieval// *Proceedings of the ACM International Conference on Conference on Information and Knowledge Management*. Shanghai, China, 2014: 101-110
- [44] Guo H, Tang R, Ye Y, et al. DeepFM: A factorization-machine based neural network for CTR prediction//*Proceedings of the International Joint Conference on Artificial Intelligence*. Melbourne, Australia, 2017: 1-7
- [45] Zhou G, Zhu X, Song C, et al. Deep interest network for click-through rate prediction//*Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, USA, 2018: 1059-1068
- [46] Wang R, Fu B, Fu G, et al. Deep & cross network for ad click predictions//*Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Halifax, Canada, 2017: Article No. 12
- [47] Lian J, Zhou X, Zhang F, et al. XDeepFM: Combining explicit and implicit feature interactions for recommender systems//*Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, USA, 2018: 1754-1763
- [48] Liu Meng-Juan, Zeng Gui-Chuan, Yue Wei, et al. A hybrid network based CTR prediction model for online advertising. *Chinese Journal of Computers*, 2019, 42(7): 1570-1587(in Chinese)
(刘梦娟, 曾贵川, 岳威等. 基于融合结构的在线广告点击率预测模型. *计算机学报*, 2019, 42(7): 1570-1587)
- [49] Grigas P, Lobos A, Wen Z, et al. Profit maximization for online advertising demand-side platforms//*Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Halifax, Canada, 2017: Article No.11
- [50] Ren K, Zhang W, Chang K, et al. Bidding machine: Learning to bid for directly optimizing profits in display advertising. *IEEE Transactions on Knowledge and Data Engineering*, 2018, 30(4): 645-659
- [51] Perlich C, Dalessandro B, Hook R, et al. Bid optimizing and inventory scoring in targeted online advertising//*Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Beijing, China, 2012: 804-812
- [52] Lin C C, Chuang K T, Wu C H, et al. Combining powers of two predictors in optimizing real-time bidding strategy under constrained budget//*Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. Indianapolis, USA, 2016: 2143-2148
- [53] Mnih V, Kavukcuoglu K, Silver D, et al. Playing Atari with deep reinforcement learning//*Proceedings of the Deep Learning Workshop on Neural Information Processing Systems (NIPS 2013)*. Lake Tahoe, USA, 2013: 1-9
- [54] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. *Nature*, 2015, 518(7540): 529-533
- [55] Cai H, Ren K, Zhang W, et al. Real-time bidding by reinforcement learning in display advertising//*Proceedings of the 10th ACM International Conference on Web Search and Data Mining*. Cambridge, UK, 2017: 661-670
- [56] Kaelbling L P, Littman M L, Moore A P. Reinforcement learning: A survey. *Artificial Intelligence Research*, 1996, 4(1): 237-285
- [57] Sutton R, Barto A. Reinforcement Learning: An Introduction. Cambridge, USA: The MIT Press, 1998
- [58] Du M, Sassioui R, Varistean G, et al. Improving real-time bidding using a constrained Markov decision process// *Proceedings of the International Conference on Advanced Data Mining and Applications*, Singapore, 2017: 711-726
- [59] Lange S, Gabel T, Riedmiller M. Batch Reinforcement Learning//Wiering M, Otterlo M eds. *Reinforcement Learning*. Berlin, Germany: Springer, 2012: 45-73
- [60] Wang Y, Liu J, Liu Y, et al. LADDER: A human-level bidding agent for large-scale real-time online auctions. *arXiv preprint arXiv: 1708.05565v2*, 2017
- [61] Zhao J, Qiu G, Guan Z, et al. Deep reinforcement learning for sponsored search real-time bidding//*Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. London, UK, 2018: 1021-1030
- [62] Wu D, Chen X, Yang X, et al. Budget constrained bidding by model-free reinforcement learning in display advertising// *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. Torino, Italy, 2018: 1443-1451
- [63] Zhang W, Zhou T, Wang J, Xu J. Bid-aware gradient descent for unbiased learning with censored data in display advertising//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, USA, 2016: 665-674

- [64] Wu C H, Yeh M Y, Chen M S. Predicting winning price in real time bidding with censored data//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Sydney, Australia, 2015: 1305-1314
- [65] Wang Y, Ren K, Zhang W, et al. Functional bid landscape forecasting for display advertising//Proceedings of the European Conference on Machine Learning & Knowledge Discovery in Databases. Riva del Garda, Italy, 2016: 115-131
- [66] Cui Y, Zhang R, Li W, et al. Bid landscape forecasting in online ad exchange marketplace//Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, USA, 2011: 265-273
- [67] Fernandez-Tapia J. Optimal budget-pacing for real-time bidding. SSRN Electronic Journal. [ssrn.2576212](https://ssrn.com/abstract=2576212), 2015
- [68] Agarwal D, Ghosh S, et al. Budget pacing for targeted online advertisements at LinkedIn//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA, 2014: 1613-1619
- [69] Lee K, Jalali A, Dasdan A. Real time bid optimization with smooth budget delivery in online advertising//Proceedings of the 7th International Workshop on Data Mining for Online Advertising. Chicago, USA, 2013: Article No.1
- [70] Xu J, Lee K C, Li W, et al. Smart pacing for effective online ad campaign optimization//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, Australia, 2015: 2217-2226
- [71] Yuan S, Wang J, Chen B, et al. An empirical study of reserve price optimization in real-time bidding//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, USA, 2014: 1897-1906
- [72] Yang W, Feng Y, Xiao B. Optimal reserve price in dynamic sponsored search auction//Proceedings of the International Conference on Service Systems & Service Management. Hong Kong, China, 2013: 647-652
- [73] Cui T, Peng L, Pardoe D, et al. Data-driven reserve prices for social advertising auctions at LinkedIn//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Halifax, Canada, 2017: Article No. 5
- [74] Lisbona M A A, Chammas S, Lee K C. Optimal reserve prices in upstream auctions: Empirical application on online video advertising//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. San Francisco, USA, 2016: 1395-1404
- [75] Xie Z, Lee K C, Wang L. Optimal reserve price for online ads trading based on inventory identification//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Halifax, Canada, 2017: Article No. 6
- [76] Chen Y J. Optimal dynamic auctions for display advertising. *Operations Research*, 2017, 65: 897-913
- [77] Balseiro S R, Feldman J, Mirrokni V, et al. Yield optimization of display advertising with ad exchange. *Management Science*, 2014, 60(12): 2886-2907
- [78] Chen B, Yuan S, Wang J. A dynamic pricing model for unifying programmatic guarantee and real-time bidding in display advertising//Proceedings of the 8th International Workshop on Data Mining for Online Advertising (ADKDD'14). New York, USA, 2014: 1-9
- [79] Li J, Ni X, Yuan Y, et al. Optimal allocation of ad inventory in real-time bidding advertising markets//Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics. Budapest, Hungary, 2016: 3021-3026
- [80] Avery C. Strategic jump bidding in English auctions. *Review of Economic Studies*, 2010, 65(2): 185-210
- [81] Robinson M A, Robinson R. Dutch-auction IPOs: Institutional development and underpricing performance. *Journal of Economics & Finance*, 2012, 36(3): 521-554
- [82] Guerre E, Perrigne I, Vuong Q. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 2010, 68(3): 525-574
- [83] Krishna V. *Auction Theory*. 2nd Edition. Amsterdam: Academic Press, 2009
- [84] Wilkens C A, Cavallo R, Niazadeh R. GSP: The Cinderella of mechanism design//Proceedings of the International Conference on World Wide Web. Perth, Australia, 2017: 25-32
- [85] Varian H R, Harris C. The VCG auction in theory and practice. *American Economic Review*, 2014, 104(5): 442-445
- [86] Angel S, Walfish M. Verifiable auctions for online ad exchanges //Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China, 2013: 195-206
- [87] Akbarpour M, Li S. Credible mechanisms//Proceedings of the 2018 ACM Conference on Economics and Computation. Ithaca, USA, 2018: 371
- [88] Caragiannis I, Kaklamanis C, et al. On the efficiency of equilibria in generalized second price auctions//Proceedings of the 12th ACM Conference on Electronic Commerce. San Jose, USA, 2011: 81-90
- [89] Caragiannis I, Kaklamanis C, Kanellopoulos P, Kyropoulou M. Revenue guarantees in the generalized second price auction. *ACM Transactions on Internet Technology*, 2014, 4(2): Article No.17
- [90] Mirrokni V, Leme P R, Tang P, Zuo S. Dynamic auctions with bank accounts//Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16). New York, USA, 2016: 387-393
- [91] Fulgoni G M. Fraud in digital advertising: A multibillion-dollar black hole. *Journal of Advertising Research*, 2016, 56(2): 122-125
- [92] Edelman B. Securing online advertising: Rustlers and sheriffs in the new wild west//Oram A, Viega J eds. *Beautiful Security: Leading Security Experts Explain How They Think*. USA: O'Reilly Media, 2009: 89-105

- [93] Jakobsson M, Ramzan Z. *Crimeware: Understanding New Attacks and Defenses*, Sebastopol; Addison-Wesley Professional, 2008
- [94] Stone-Gross B, Stevens R, Zarras A, et al. Understanding fraudulent activities in online ad exchanges//Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference. Berlin, Germany, 2011: 279-294
- [95] Feily M, Shahrestani A, Ramadass S. A survey of Botnet and Botnet detection//Proceedings of the 3rd International Conference on Emerging Security Information, Systems and Technologies. Athens, Greece, 2009; 268-273
- [96] Zhu X, Tao H, Wu Z, et al. *Fraud Prevention in Online Digital*. Cham, Switzerland; Springer, 2017
- [97] Jaafar G A, Abdullah S M, Ismail S. Review of recent detection methods for HTTP DDoS attack. *Journal of Computer Networks and Communications*, 2019, 4: 1-10
- [98] Vratonjic N, Manshaei M H, Raya M, Hubaux J P. ISPs and ad networks against Botnet ad fraud//Proceedings of the International Conference on Decision and Game Theory for Security. Berlin, Germany, 2010: 149-167
- [99] Kim I L, Wang W, Kwon Y, Zheng Y, et al. AdBudgetKiller: Online advertising budget draining attack//Proceedings of the 2018 World Wide Web Conference. Lyon, France, 2018: 297-307
- [100] Thomas K, Bursztein E, Grier C, et al. Ad injection at scale: Assessing deceptive advertisement modifications//Proceedings of the 2015 IEEE Symposium on Security and Privacy. San Jose, USA, 2015: 151-167
- [101] Kshetri N. The economics of click fraud. *IEEE Security & Privacy*, 2010, 8(3): 45-53
- [102] McLaughlin M, Rosenfeld R K, Abu L M, Simon L. System and method for identifying hidden content. USA, 2015. 10. 01
- [103] Abbasi A, Chen H. A comparison of tools for detecting fake websites. *Computer*, 2019, 42(10): 78-86
- [104] Metwally A, Agrawal D, Abbadi A. Using association rules for fraud detection in Web advertising networks//Proceedings of the 31st International Conference on Very Large Data Bases. Trondheim, Norway, 2005: 169-180
- [105] Metwally A, Agrawal D, El Abbadi A. Duplicate detection in click streams//Proceedings of the 14th International Conference on World Wide Web. Chiba, Japan, 2005: 12-21
- [106] Perera K S, Neupane B, Faisal M A, et al. A novel ensemble learning-based approach for click fraud detection in mobile advertising//Proceedings of the International Conference on Mining Intelligence and Knowledge Exploration (MIKE). Virudhunagar, India, 2013: 370-382
- [107] Haddadi H. Fighting online click-fraud using bluff ads. *ACM SIGCOMM Computer Communication Review*, 2010, 40(2): 21-25
- [108] Xu H, Liu D, Koehl A, et al. Click fraud detection on the advertiser side//Proceedings of the 19th European Symposium on Research in Computer Security. Wroclaw, Poland, 2014: 419-438
- [109] Iqbal M S, Zulkernine M, Jaafar F, et al. Protecting internet users from becoming victimized attackers of click-fraud. *Journal of Software Evolution & Process*, 2017, 30(2): 1871-1886
- [110] Nagaraja S, Shah R. Clicktok: Click fraud detection using traffic analysis//Proceedings of the Conference on Security and Privacy in Wireless and Mobile Networks. Miami, USA, 2019: 105-116
- [111] Hu Y. Performance-based pricing models in online advertising. *SSRN Electronic Journal*. ssrn. 501082, 2014
- [112] Deng Y, Shen Y, Jin H. Disguise adversarial networks for click-through rate prediction//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 1589-1595
- [113] Diemert E, Meynet J, Galland P, et al. Attribution modeling increases efficiency of bidding in display advertising//Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Halifax, Canada, 2017: Article No. 2
- [114] Pan S J, Qiang Y. A survey on transfer learning. *IEEE Transactions on Knowledge & Data Engineering*, 2010, 22(10): 1345-1359
- [115] Su Y, Jin Z, Chen Y, et al. Improving CTR prediction accuracy in online advertising by transfer learning//Proceedings of the International Conference on Web Intelligence. Leipzig, Germany, 2017: 1018-1025



LIU Meng-Juan, Ph. D., associate professor. Her research interests include machine learning, computational advertising.

YUE Wei, M. S. candidate. His research interests include computational advertising and machine learning.

QIU Li-Zhou, M. S. candidate. His research interests include computational advertising and machine learning.

LI Jia-Xing, M. S. candidate. His research interests include computational advertising and machine learning.

QIN Zhi-Guang, professor, Ph.D. supervisor, professor. His research interests include machine learning and network security.

Background

In recent years, online advertising has developed into a multi-billion dollar industry. As one of the most exciting advances in online advertising, advertising targeted delivery has received increasing attention, since it improves the efficiency and transparency in the online advertising ecosystem. This article explores RTB, a new delivery method for display advertising. At present, the international research on RTB has just started, and the research main focus on the click rate prediction and bidding strategy. Unfortunately, in addition to Jun Wang et al.'s 2016 paper "Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting", few papers have sorted out the entire RTB transaction process and the latest developments in key issues.

In order to help researchers and algorithm engineers in this field understand the current core issues and solutions of RTB, especially its applications in industry. This paper investigates the latest researches in RTB, and sorts out the key issues from the perspectives of the three main platforms: demand side platform, service side platform and AD eXchange. These issues include user response prediction, bidding strategy, bidding landscape, reserve price setting, auction theory,

advertising fraud detection, etc. At the same time, this paper analyzed the existing research methods, theories and models. In particular, the latest research on the most popular user response prediction and bidding strategy is introduced and analyzed in detail. In user response prediction, the paper introduces the current common shallow models, such as linear model, FM, FFM, etc., deep models such as FNN, PNN, DIN, etc., and mixed models such as Wide&Deep, DeepFM, etc. In the bidding strategy, the paper introduces linear bidding models, nonlinear bidding models and bidding models based on reinforcement learning. Finally, the paper looks forward to the future research trends and directions in this field and introduces some new technologies and methods that may be used in future research. This work is supported by the National Natural Science Foundation of China under Grants (Nos. 61202445, 61502087), and the Fundamental Research Funds for the Central Universities (No. ZYGX2016J096). The main job of the project is to conduct recommendations and targeted advertising based on user interest and behavioral predictions. In this direction, the research team published 5 related papers and obtained 2 authorized patents.