

文章编号:1000-5641(2013)03-0106-12

在线广告投放系统及技术的演变

宋乐怡¹, 宫学庆¹, 张蓉¹, 刘鹏²

(1. 华东师范大学 软件学院, 上海 200062;

2. 北京搜狐新媒体信息技术有限公司, 北京 100004)

摘要: 通过对在线广告市场的背景及相关技术的介绍,从广告的形式和投放位置这两个维度对在线广告进行划分,给出了涵盖在线广告领域现有及潜在广告形态的分类体系,分析了不同形态广告投放技术的差异.通过对在线广告投放平台演变和发展的了解,概述了以广告服务器、广告网络、广告交换平台为主的广告投放系统的架构及技术演变,为计算广告相关研究工作提供了关于在线广告投放系统的背景参考和技术概览.

关键词: 在线广告; 计算广告学; 投放系统

中图分类号: TP391 **文献标识码:** A **DOI:**10.3969/j.issn.1000-5641.2013.03.012

Online advertising systems and related technology evolution

SONG Le-yi¹, GONG Xue-qing¹, ZHANG Rong¹, LIU Peng²

(1. *Software Engineering Institute, East China Normal University, Shanghai 200062, China;*

2. *Beijing Sohu New Media Information Technology Co., Ltd., Beijing 100004, China*)

Abstract: This paper introduced the background and related techniques of online advertising market. While different types of online ads may imply different computational advertising techniques in the backend system, this paper provided a rational classification method of online advertising formats. The category of online advertising given in this paper covers existing and potential types in the domain. Further, the advertising platform have evolved in several major phases or generations, particularly ad server, ad network and ad exchange. We showed the system architectures of the advertising systems mentioned above, by discussing the main function modules and interfaces. Our work aims to give a comprehensive and detailed description of the online advertising systems from the view of computation. Besides, the surveys in our work can provide essential background knowledge for computational advertising related research.

Key words: online advertising; computational advertising; delivery system

0 引言

随着互联网的发展,广告营销模式被带入了数字媒体的时代,形成了一个巨大的以计算

收稿日期:2013-03

第一作者:宋乐怡,女,硕士研究生,研究方向为 Web 数据管理与数据挖掘. E-mail: songleyi@ecnu.cn.

通信作者:宫学庆,男,副教授. E-mail: xqgong@sei.ecnu.edu.cn.

和技术驱动的在线广告市场.美国互动广告局(IAB, Interactive Advertising Bureau)最新的报告显示,互联网广告收入已经达到了空前的高度,2012年上半年互联网广告收入高达170亿美元^[1].在线广告区别于传统的线下购买—投放广告模式,具有可交互性、可定制、可跟踪、可送达等优点^[2].技术使得在线广告可以进行精准的受众定向和广告决策,其中计算广告技术从互联网在线广告的实践中产生,以追求广告投放的综合利益最大化为目标,是支撑互联网广告产业链发展的基础.Andrei Broder在2008年指出,计算广告是综合了大规模搜索、文本分析、信息检索、统计模型、机器学习、分类、最优化、微观经济学、推荐系统等多个方向的交叉新兴学科,其核心挑战是解决用户、上下文与广告的“最佳匹配”问题^[3].

在线广告支撑了互联网中一个庞大的生态系统,计算广告技术为互联网市场和用户创造了大量的直接和间接价值.因此,本文将以在线广告生态体系为背景,首先为全文提供在线广告体系中概念和技术背景知识的介绍;其次从展示层分析在线广告的类型,以广告形式和投放位置为维度对在线广告进行划分,构建一个涵盖现有及潜在的广告形式和平台的分类体系;最后,介绍广告投放系统的演变,以广告服务器、广告网络、广告交换平台为例,对广告投放系统的功能模块和涉及角色进行调研和分析.

本文从上层的在线广告类型到底层的投放系统架构,对在线广告投放系统提供全面的介绍,为计算广告研究工作提供必要的背景知识和技术参考.全文组织如下:第1节主要介绍在线广告生态系统相关的概念和技术背景;第2节讨论计算广告的分类体系和技术差异;第3节详细分析计算广告生态系统发展中 Ad Server、Ad Network、Ad Exchange 等角色功能和体系结构演变;最后总结全文并展望计算广告生态系统的技术发展.

1 相关概念及背景

1.1 在线广告投放简介

在线广告系统是一个多角色博弈的平台,该系统中最初包含三个角色,即广告主、发布者(即互联网媒体)和互联网用户.计算驱动的广告投放平台以最大化用户对广告兴趣、最大化媒体收益和最大化广告主 ROI(投资回报率)为目标.2000年左右,互联网上涌现出大量的没有自身广告投放平台或广告位过剩的网络媒体,媒体和广告主通过接入广告网络(Ad Network)进行广告投放.以广告网络为中心的广告投放平台对媒体页面进行聚集和分类,提供支持上下文定向的流量售卖方式,一方面提高广告投放的精准度,另一方面减少媒体页面的剩余流量.2005年左右广告网络已颇具规模,在线广告投放体系进一步扩展,形成了广告网络、广告商、需求方/供应方平台(DSP/SSP)等多种角色通过广告交换平台(Ad Exchange)对广告位进行实时竞价和售卖的广告竞拍及精准投放的在线广告生态体系.^[4]

在线广告投放中的广告形态多样,常见的包括以搜索引擎为基础的搜索广告(Sponsored Search)和以网页内容为基础的上下文广告(Contextual Ads)或者内容匹配广告(Content Match Ads)^[5].广告主对广告流量的购买可以通过签订合约来规定媒体对广告展示次数的合约投放(GD, Guaranteed Delivery),或者仅受预算限制,不固定投放流量的非合约投放(NGD, Non-Guaranteed Delivery).广告售卖的依据可以是搜索关键词、网页类别或者具有相应特征的用户,根据不同的广告投放形态,相应采用以广告位页面内容为中心的上下文定向(Contextual Targeting),以浏览用户行为特征为中心的行为定向(Behavioral Targeting)等技术进行广告精准投放.根据不同的广告投放平台和广告形态,计价方式有基于

广告曝光次数的 CPM(Cost-Per-Mile, 每千次曝光)、基于点击次数的 CPC(Cost-Per-Click)、基于用户行动的 CPA(Cost-Per-Action)等计价法,其中 CPC/CPA 是基于广告投放效果的计价方式,另外也有结合 CPM 和基于效果的混合收费方式^[1].

1.2 相关技术

在线广告生态系统中为达到不同的利益目标,常采用一些预测或建模技术,本文将介绍如下三种:广告网络中实现精准营销的定向投放,DSP 等平台为提高收益率对点击率进行预测的技术,帮助广告主和广告投放平台识别恶意广告点击的检测方法.

1.2.1 定向投放

定向投放(Targeting)是互联网广告精准营销中的一项重要技术,被广泛用于广告辅助选择和提高点击率上.传统的定向方式有很多种,如基于人口学特征、基于地域、基于页面上下文、基于用户行为、重定向等.

基于上下文的定向(Contextual Targeting)和基于行为的定向(Behavioral Targeting)是目前两类重要的定向技术.前者多用于 Ad Server 等对于页面广告的投放中,适用于用户信息匮乏或者有隐私保护的场景下,使用广告与页面匹配的算法进行广告投放;定向的步骤包括对网页和广告数据的打标签、广告商的产品定向匹配等^[6,7].基于用户行为定向是最广泛应用的定向技术,也是较有效的定向技术.搜索引擎可以根据用户的搜索历史进行兴趣广告投放,电商网站可以根据用户购买和浏览历史投放产品推荐广告.基于用户行为的定向利用历史用户行为选择最相关的广告进行投放,其核心问题是学习一个从用户到广告的定向函数^[8],找到最优解,可以在用户历史点击、用户偏好等行为数据上,应用线性泊松回归模型(Linear Poisson Regression Model)等概率模型进行建模和预测^[9].相比于上下文定向,行为定向需要足够的用户数据并能够合理识别用户.定向技术都需要进行大量的后台数据加工和提取,能够在一定程度上提高广告的点击率.

1.2.2 点击率预测

点击率(Click Through Rate, CTR)预测,是通过历史点击数据或者用户行为,对广告投放后的效果进行预测的一类技术,通过预测点击率,决定广告竞标价和广告投放策略,从而使得广告投放利益最大化.

CTR 预测技术广泛用于 Ad Network、DSP 等平台中.如在 DSP 中对广告位进行竞标并选择合适的广告.首先对从上游系统接收到的广告位信息进行页面和用户特征抽取^[10],利用抽取的特征信息预测点击概率^[11-13],然后预测并优化竞标价格^[14],最后选择合适的广告进行投放.对点击率预测常采用机器学习、统计学等模型,如对关键词的聚类^[11]、采用逻辑回归的方法预测新广告点击率^[12]、贝叶斯方法计算 Term 的点击率^[13]等.

1.2.3 恶意点击检测

在线广告的投放中也面临诸多问题,其中恶意的广告点击是破坏在线广告生态系统信任度的源泉之一^[15].恶意点击或者点击欺诈被定义为带有恶意或者欺诈意图的点击,恶意点击属于无效流量的一部分.无效流量通常指不能给广告主带来收益,与真实用户兴趣和广告投放要求不符的广告点击和展示,其中包括故意的欺诈性流量、意外点击和其他记住机器产生的流量^[16].恶意点击一方面来源于广告主的竞争对手,竞争对手通过恶意点击增加广告主的费用或尽早耗尽预算;另一方面来源于媒体发布商,投放页面所属媒体通过恶意点击增加自己网站通过广告点击带来的收入.

现有的广告投放平台中均有一定的监督机制,如 Google 通过三种过滤器,分别基于异常检测、基于规则和基于分类器,以在线和离线的方式过滤无效点击^[17]. 离线检测恶意点击的方法以通过历史数据构建分类器为主,如 Random Forest 算法、Naïve Bayes 算法、SVM 等^[18]. 还可以在广告主投放广告页面,进行特定的信息收集,通过点击的统计数据,过滤非正常用户的点击,如投放反欺诈的广告(Bluff Ads)测试广告点击者的合法性^[19]. 在线的检测方法以数据流算法为主,在点击数据流上检测重复的用户点击^[20],对点击流量实时分析,发现具有一定特征的恶意点击攻击^[21],或者应用关联规则等数据挖掘算法检测日志流中的异常点击^[22]. 恶意点击检测在维护在线广告生态系统的稳定性中有重要意义,因此具有很大的研究价值.

根据该节所述在线广告投放系统的技术演变、交易方式和相关技术,接下来本文将对广告的形态和系统架构进行深入介绍.

2 在线广告分类

2.1 分类概览

根据在线广告的展现类型和投放方式有很多不同的分类方式,如 IAB 报告^[1]中对计算广告的格式就有搜索、显示条幅、分类、移动等数十种分类方式. 目前,有根据广告上下文将其分为以搜索引擎检索为上下文的搜索广告(Sponsored Search)、以网页为上下文的显示广告(Display Ads)、以便携设备应用为上下文的移动广告(Mobile Ads)等^[3,5]. 还有根据广告媒介平台,分为移动、社交网络、游戏、门户网站、视频、TV 等在线广告类别^[23,24].

从各种划分类别的方式来看,有些类别界限较为模糊,比如显示广告和移动广告,虽然前者基于网页后者基于移动终端,但前者强调广告投放的形式,后者强调投放给受众的广告位置. 从计算广告学研究角度讲,不同形式不同位置的广告投放,后台支撑的技术有所不同,如对于搜索广告通常会在搜索引擎算法基础上有其特有的选择排序^[6]、查询重写^[25]等相关技术. 对于在互联网页面上投放的显示广告,不同投放位置会有相应的系统设计,如投放在视频中^[26]和投放在网页条幅中^[27]的广告.

本节将详细讨论在线广告的分类划分及技术特点. 图 1 是以广告投放形式和投放位置作为两个划分的维度,对在线广告形态进行分类的概览图. 根据广告形式,可以分为纯文本广告(Text Ads)、显示广告(Display Ads)和互动式广告(Interactive Ads). 文本广告可以根据其投放位置不同,分为分类广告(Classified Ad)、基于关键词的搜索广告(Sponsored Search)和基于所在页面的内容广告(Content Match),显示广告和互动式广告根据其投放位置不同分别有电子邮件广告(Email Ad)、社交网络广告(Social Ad)、条幅广告(Banner Ad)、移动广告(Mobile Ad)、游戏广告(Game Ad)、视频广告(Video Ad)等不同类别. 总体而言,基于搜索引擎和基于所在页面或应用上下文的广告(Contextual Ads)是两类重要的广告形态.

2.2 按广告形式分类

在线广告的不同展现形式,即视觉上向用户表达广告内容的格式,可以分为纯文本格式和包含图片、动画、视频等多媒体元素的显示广告格式,以及互动式广告. 其中互动式广告是一类特殊的显示广告形式,用户可以通过拉模式获取广告,通过 HTML5 等技术获得更高的用户响应和点击率.

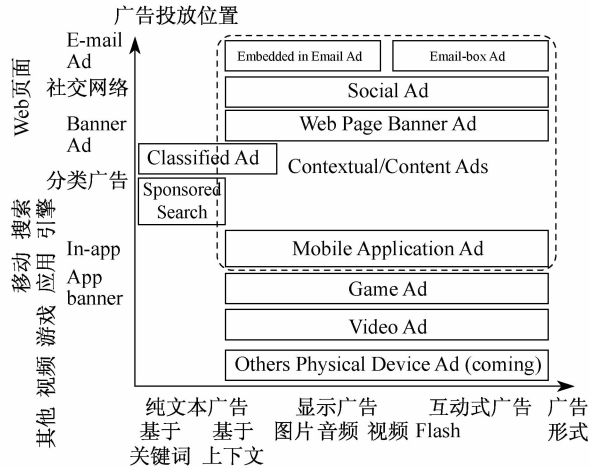


图 1 在线广告分类概览

Fig. 1 Landscape of online advertising category

2.2.1 纯文本广告

纯文本广告占据在线广告市场很大的份额,以短文本形式展现给用户,通常会包含类似于赞助链接(Sponsored Link)的广告页面链接.有两种重要的投放渠道:搜索广告(Sponsored Search)和页面上上下文广告(Contextual Advertising)^[28].

搜索引擎广告市场给文本广告带来了快速发展,广告商通过竞标关键词的方式来购买搜索引擎结果页面的广告展示,搜索引擎通过重写等方式将搜索关键词进行扩展,转换为进行广告位匹配和竞拍的关键词^[25],广告主竞拍关键词获取广告展示流量.

文本类型的上下文广告是网络媒体为提高页面带来的收益提供给广告主的文本广告位.该类型广告一般包括三个部分:标题、广告文本描述、超链接,其中超链接指向的页面称为着陆页(Landing Page),着陆页上包含更加详细的广告内容或产品介绍^[29].

2.2.2 显示广告

显示广告(Display Ads)是除文本以外,还包含 Logo、图片、位置地图等多媒体元素的一种广告展现形式^[30],提供在线广告展示的广告位是互联网网站收入的重要来源.当用户请求一个 Web 页面,一次显示广告的机会就产生了,广告请求通过网站的广告服务器,从广告池中选取投放的广告创意,广告将出现在用户浏览的页面指定的广告位上,网站发布商的计价系统会向广告所属的广告主收取费用^[31],这样就完成了一次典型的显示广告的投放.

2.2.3 交互式广告

交互式广告(Interactive Ads)又称为富媒体广告(Rich Media Ads),是一类由传统的显示广告发展起来的更加吸引用户的广告形式. Google 中富媒体广告被定义为包含图片、视频和其他具有良好用户交互性媒体的广告形式,用户可以根据需求对广告进行扩展、划动、弹出等操作^[32],常见的格式有展开式、漂浮式、撕页式、下推式等.由于交互式广告良好的用户体验性,对于品牌推广类型的广告来说,能够取得较之于其他广告更出色的效果.但交互式广告经常会打断用户正常的浏览行为,影响广告投放效果.因此,常将富媒体与定向投放结合起来,根据用户兴趣和上下文进行更加个性化的广告推荐,这样能够大大提高广告的投

放效果和用户的体验性^[33].

2.3 按广告投放位置分类

广告投放位置表示了在线广告投放的媒介和展示方式,隐含了对应的需求方和供应方,并且随着互联网媒体的发展,会不断出现新的广告投放位置和营销平台.以下将介绍目前互联网中重要的广告投放媒介及其展示方式.

2.3.1 Web 页面

投放在 Web 页面上的广告形式多样,有文字、图片、视频、富媒体等,主要目的是在网页中展示广告主的品牌或产品,是在线广告投放的重要方式,也是互联网页面收入的重要来源.根据所在页面的具体属性,又可以将页面广告分为静态页面的条幅广告、社交网络上的广告、分类广告和 Email 广告等.

条幅广告(Banner Ads)是嵌在或悬浮在页面中的固定广告位,通过广告活动(Advertising Campaign)投放广告主的广告,其收入已占在线广告市场的 20%左右^[1].社交网络可以利用用户生成数据来提高在线广告营销的用户认知度,尤其能够提升品牌广告的影响力^[34],也可以根据用户行为历史,预测用户的兴趣和点击行为^[35].分类广告(Classified Ads)是一种从传统的报刊杂志等发展到互联网的广告形式,通过分类导航的方式对不同广告进行组织^[36].Email 广告则是嵌在电子邮件内或者在收件箱边缝以显示广告的形式进行投放的广告.Web 页面广告根据媒体页面种类丰富多样,有上下文定向、行为定向、广告分配等多种技术作为支撑,在在线广告生态系统中占有重要地位.

2.3.2 搜索引擎

另外一种重要的广告投放途径,虽然同样为 Web 页面投放,但其源自搜索引擎技术,与页面的上下文广告有一定的区别,这类广告被称为搜索广告或者赞助商搜索.当用户通过搜索引擎查询时,有两种不同的查询,一种返回相关的页面信息,另外一种返回付费的相关广告信息,即搜索广告.首先,搜索引擎将检索关键词扩展为进行广告位匹配和竞拍的广告关键词^[25];然后,广告主通过需求方平台(DSP, Demand Side Platform)等接入搜索引擎广告位交换平台,预测广告点击率,提交竞拍价格;最后,当广告被点击后,转换为相应的点击或转化(如进行了对广告产品的购买)行为^[37],搜索引擎根据统计数据对广告主进行收费.2012 年上半年,搜索广告已占到接近 50%的在线广告市场^[1].

2.3.3 移动应用

移动手机、平板电脑等便携设备的渗透率逐年增加,通过移动设备进行网络搜索和浏览成为多数人的爱好,移动应用市场也逐步扩大.移动广告(Mobile Ads)即为投放在基于移动设备的应用中的广告,有些沿用了传统网页广告的技术,但由于移动环境下的可定位性,基于地理位置的移动推荐技术成为移动广告投放的关键技术,如根据用户所在位置对其进行旅行社的旅行路线推荐^[38].计算广告技术需要解决广告与用户之间的最佳匹配问题,这样既能满足广告被足够曝光的需求,又能满足对用户投放有选择性的少量广告.移动计算有更多投放广告方式的选择,能够选择更加合适的广告和投放途径,减少对用户的干扰.从统计数据来看,移动广告市场正在快速地逐年增长^[1],移动计算环境下的广告投放具有巨大的潜力.

2.3.4 游戏、视频等

早期的游戏广告(Game Ads)或者视频广告(Video Ads)直接植入游戏或视频播放器

中,可以采用与一般移动应用类似的广告投放方式. 游戏和视频广告正朝着更高交互性的方向发展,如微软的 Xbox 在游戏中投放交互式的广告^[39]. 视频流中针对合适的场景进行广告插入^[40]. 除此以外,在线广告可以投放在无处不在的数字媒体上,对于下一代的物理信息系统(Cyber-Physical System),未来在线广告将投放在更多的接入互联网的控制面板或其他平台上.

总之,计算广告作为互联网应用的一个技术体系,支撑了在线广告后台诸多重要的技术. 上文根据投放的形式和位置对在线广告体系进行分类,其中每个类别对应其相关的广告投放技术. 后文将深入剖析多种形态的在线广告后台复杂的投放平台及系统发展,尤其针对广告生态系统发展中三个重要阶段的广告投放平台的功能模块和体系结构进行详细分析和讨论.

3 在线广告投放系统发展

3.1 在线广告投放系统平台演变

从在线广告兴起到现在,投放系统的发展有三个重要阶段:① 网站发布商通过 Ad Server 系统直接进行广告投放;② 以 Ad Network 为中间平台,对网络媒体广告位库存进行购买,然后转售给广告主;③ Ad Exchange 形成一个开放的在线广告交易市场,通过实时竞价(RTB)的方式进行需求方和供应方的广告投放交易.

在线广告投放系统演变如图 2 所示:基于互联网广告的平台起始于 1995 年左右,产业链中仅包含广告主、发布者和用户,广告主与发布商签定广告投放的合约,发布商根据合约分配广告投放的流量,通过 Ad Server 进行广告选择和投放系统;2000 年左右出现的 Ad Network 是一个封闭的网络广告市场,广告网络作为中间平台首先向发布商采购广告为库存,然后再转售,广告网络对广告位进行打包、分类,提高广告位的价值,广告主可以进行更有受众针对性的广告投放,并提供广告网络管理和检测服务;2005 年左右出现的 Ad Exchange 有助于多个广告网络之间实现资源有效分配,降低广告位的闲置率,广告主能找到最合适的广告位,提高投放效果;2010 年左右在线广告市场出现了需求方平台(Demand Side Platform)、供应方平台(Sell Side Platform)、数据管理平台(Data Management Platform)等系统分别服务于不同的角色,如 DSP 服务于广告主,帮助广告主在指定的目标受众、投放区域、广告出价等要求下寻找优质的媒介和精准的用户,优化广告投放产出比^[4].

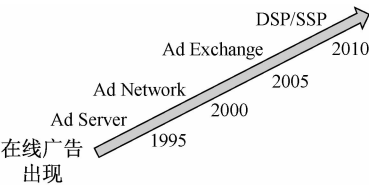


图 2 在线广告投放平台演变

Fig. 2 Evolution of online advertising platform

在线广告的市场演变与投放平台的进步是密不可分的,后文将针对不同阶段广告平台的主要功能分析其基本模块和流程框架.

3.2 Ad Server——广告服务器

狭义的 Ad Server 即供应端的广告服务器,用于管理广告主的广告投放合约(尤其是 GD 广告)及广告创意,进行网站页面广告位的广告选择和投放. 主要功能模块包括:离线的库存管理(Repository Manage)及点击预测(CTR Predict)模型建立,广告选择(Ad Retrieval),广告排序(Ad Ranking),投放优化(Optimization),广告投递(Ad Delivery)等. Ad Server 的示意系统架构图如图 3 所示:当新的广告创意和广告合约(Ad Contracts)加入广告服务器中,库存管理模块对其建立相关索引(Ad Index),并根据合约库建立目标受众预测模型,对广告投放进行预先调度;当用户浏览媒体页面,广告请求被发送到广告服务器,首先对页面信息和用户信息进行信息整合,通过广告服务(Ad Serving)引擎进行广告检索,广告排序,并进行相应的在线优化^[41],最后返回广告结果,发送回页面广告位.

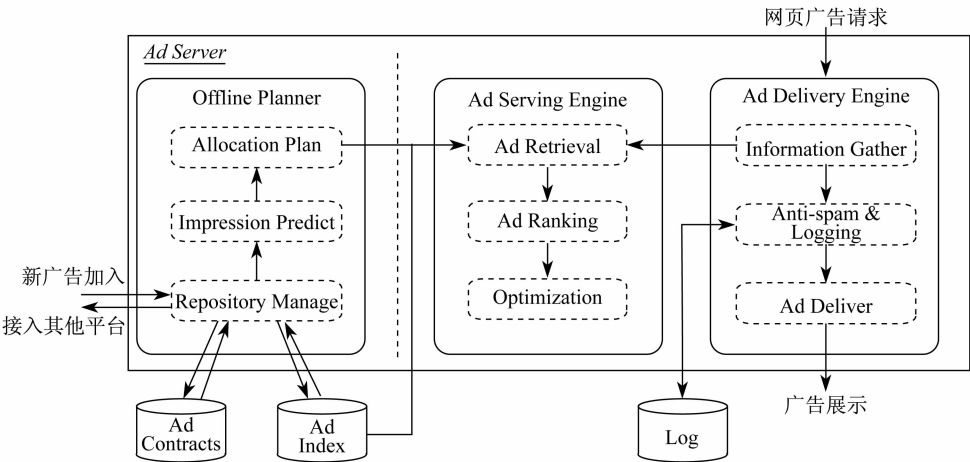


图 3 Ad Server 系统结构

Fig. 3 System architecture of ad server

3.3 Ad Network——广告网络

广告网络是通过集中采购各类网络媒体的媒介资源,通过技术手段实现分布投放的虚拟运营平台,帮助广告主实现多种媒体资源组合投放的广告服务网络^[42]. 广告网络负责对大量页面进行分析、聚集,完成广告投放转换的过程. 其优势体现在对网站媒介资源的整合能力、受众数据获取和挖掘的能力以及精准匹配广告主营销需求的能力. 从应用角度来看,广告网络的功能包括媒体采购、媒体管理、数据对接、精准定向^[43]等. 从系统角度,可能包括以下功能模块:广告位所属网页的内容分析、用户受众定向、广告匹配、广告选择与投放、展示跟踪与报告等.

广告网络的示意系统架构图如上图 4 所示. 当从发布者购入新的媒体广告位时,通过库存管理(Inventory Manage)、网络分析(Web Analysis)模块对广告上下文进行关键词抽取(Keywords Extract)/打标签(Tagging)/聚集(Aggregating)等,用于之后进行受众定向(Audience Targeting)和预测模型(Predict Model)建立. 当广告主从广告网络中购买指定受众设置的广告流量时,广告库存管理(Repository Manage)将其加入待投放的广告索引中,通过对广告进行特征抽取(Feature Extraction),用于进行投放时与广告位的匹配. 最后,当发布者页面发送广告投放请求到广告网络中,广告服务器(Ad Server)根据页面和用户的属性及预测模型选择最佳匹配并且收益最高的广告进行投放.

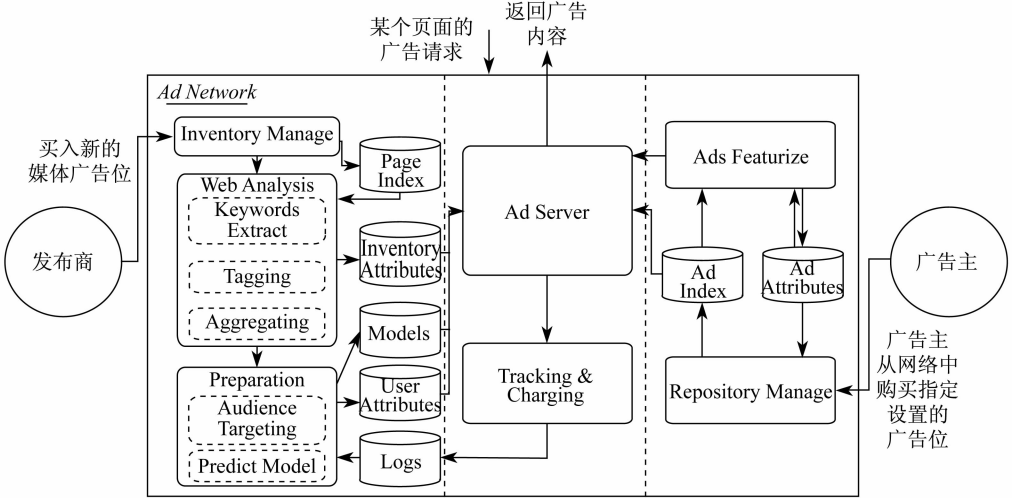


图 4 Ad Network 系统结构

Fig. 4 System architecture of ad network

3.4 Ad Exchange——广告交换平台

广告网络是一个相对封闭的系统,网络内部机制负责将网络内的媒体页面和广告进行匹配和投放,而广告交换平台提供了一个开放的交易场所,相比于广告网络具有更大的库存量.广告位的来源可以是多个广告网络,也可以是第三方的供应平台.广告交易针对每个展示机会进行实时竞价,并提供相关的用户行为属性、页面类别等.广告主可以以更加透明的方式进行竞拍.广告平台能够整合多个广告网络的信息,减少投放中多个广告网络之间人群重叠的问题,从而节省广告主的开支;也为网络媒体提供了一个公开的广告位拍卖平台,减少剩余流量和中间成本.

常见的广告交换平台有:Google 公司的 DoubleClick,OpenX 公司的 OpenX Market,雅虎公司的 Right Media,微软公司的 AdECN 等^[44]. Ad Exchange 作为一个大型的广告拍卖平台,内部有很多数据整合、优化、竞拍策略等技术,并且仍有很多值得研究的问题^[45].图 5 所示为广告交易平台的示意图.其核心为实时竞价(RTB, Real-Time Bidding)平台,通过整合数据管理平台、广告网络等多方面的信息,提供更有效的用户及广告位信息,帮助实现广告的精准投放.

综上所述,计算广告生态系统中逐渐出现了多种平台和角色,在线广告市场在技术的推动下,形成了一个具有一定规模以市场竞拍为基础的广告交易平台.虽然技术平台经历了 Ad Server、Ad Network、Ad Exchange 等的演化,但每一个阶段的技术并未被完全取代,而是加入新的系统中,以更高效的方式,协作完成在线广告展示的精准投放,维持整个计算广告生态系统的良好运转.

4 总结与展望

以计算驱动的在线广告投放市场支撑了互联网上一个庞大的生态系统,在线广告的投放形成了一个多角色博弈的平台,投放位置遍布数字媒体各种媒介形式,以不同的广告形态出现.本文介绍了计算广告技术中涉及的在线广告背景、相关术语和技术,给出了一个覆盖

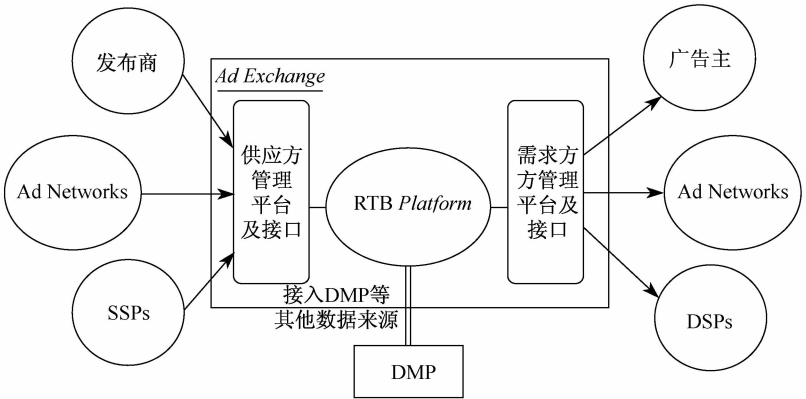


图 5 Ad Exchange 系统示意图

Fig. 5 System sketch of Ad Exchange

现有和潜在广告形态的分类体系,并在该体系基础上,根据在线广告投放系统的演变,给出了不同阶段的系统功能和架构. 在线广告投放平台是一个角色多样、功能复杂的生态体系,并随着计算技术和互联网媒介的发展,不断完善系统机制,有大量值得研究的问题,如移动广告投放、跨平台海量数据的受众定向、恶意点击检测等. 总之,本文通过对在线广告投放系统的全面介绍,为计算广告相关研究工作提供了必要的背景知识和参考.

[参 考 文 献]

[1] Internet Advertising Bureau. IAB internet advertising revenue report[R/OL]. 2012[2013 - 03 - 01]. <http://www.iab.net/AdRevenueReport>. .

[2] ZEFF R L, ARONSON B. Advertising on the Internet[M]. 2nd ed. New York: John Wiley & Sons, 1999.

[3] BRODER A. Computational Advertising[C]//Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA. San Francisco, California: SIAM, 2008: 992.

[4] 周傲英,周敏奇,宫学庆. 计算广告:以数据为核心的 Web 综合应用[J]. 计算机学报, 2011, 34(10): 1805-1819.

[5] BRODER A Z. Computational advertising and recommender systems[C]//Proceedings of the 2008 ACM conference on Recommender Systems. ACM, 2008: 1-2.

[6] HILLARD D, SCHROEDL S, MANAVOGLU E, et al. Improving ad relevance in sponsored search[C]//Proceedings of the 3rd ACM International Conference on WebSearch and Data Mining. New York: ACM, 2010: 361-370.

[7] TOUBIANA V, NARAYANAN A. Adnostic: Privacy Preserving Targeted Advertising[C]//17th Network and Distributed System Security Symposium. 2010: 1-23.

[8] PANDEY S, ALY M, BAGHERJEIRAN A, et al. Learning to Target: What Works for Behavioral Targeting [C]//Proceedings of the 20th ACM International Conference on Information and Knowledge Management. Glasgow, Scotland: 2011: 1805-1814.

[9] CHEN Y, PAVLOV D, CANNY J. Large-scale behavioral targeting[C]//Proceedings of the 15th ACM SIGKDD - International Conference on Knowledge Discovery and Data Mining. Paris: 2009: 209-217.

[10] CHENG H, CANTÜP E. Personalized click prediction in sponsored search[C]//Proceedings of the 3rd ACM international conference on Web search and data mining. New York, New York: ACM Press, 2010: 351-359.

[11] REGELSON M, FAIN D. Predicting click-through rate using keyword clusters[C]//Proceedings of the 2nd Workshop on Sponsored Search Auctions, EC06. 2006.

- [12] RICHARDSON M, DOMINOWSKA E, RAGNO R. Predicting clicks: estimating the click-through rate for new ads[C]//Proceedings of the 16th international conference on World Wide Web. Banff, Alberta, Canada: ACM, 2007: 521-529.
- [13] GRAEPEL T, CANDELA J Q, BORCHERT T, et al. Web-Scale Bayesian Click-Through Rate Prediction for Sponsored Search Advertising in Microsoft's Bing Search Engine[C]//Proceeding of the 27th International Conference on Machine Learning. Haifa, Israel:[s. n.], 2010.
- [14] CUI Y, ZHANG R, LI W, et al. Bid Landscape Forecasting in Online Ad Exchange Marketplace[C]//Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, California: ACM, 2011: 265-273.
- [15] DAVE V, GUHA S, ZHANG Y. Measuring and fingerprinting click-spam in ad networks[C]//Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication. Helsinki, Finland, 2012: 175-186.
- [16] Google. 广告流量质量资源中心[EB/OL]. [2013-03-01]. <http://www.google.com/ads/adtrafficquality/>.
- [17] TUZHILIN A. The Lane's Gifts v. Google Report[R/OL]. 2006[2013-03-01]. <http://googleblog.blogspot.com/pdf/Tuzhilin-Report.pdf>.
- [18] HÄGER M, LANDERGREN T. Implementing best practices for fraud detection on an online advertising platform [D]. Gothenburg: Chalmers University of Technology, 2010.
- [19] HADDADI H. Fighting Online Click-Fraud Using Bluff Ads[J]. Computer Communication Review, 2010, 40(2): 21-25.
- [20] ZHANG L, GUAN Y. Detecting Click Fraud in Pay-Per-Click Streams of Online Advertising Networks[C]//The 28th International Conference on Distributed Computing Systems. 2008: 77-84.
- [21] METWALLY A, AGRAWAL D, ABBADI A E. DETECTIVES: DETECTing Coalition hiT Inflation attacks in adVertisingnEtworks Streams[C]//Proceeding of the 16th international conference on World Wide Web. Banff, Alberta, Canada: 2007: 241-250.
- [22] METWALLY A, AGRAWAL D, ABBADI A E. Using Association Rules for Fraud Detection in Web Advertising Networks[C]//Proceedings of the International Conference on Very Large Data Bases. 2005: 169-180.
- [23] DAVE K. Computational Advertising: Leveraging User Interaction & Contextual Factors for Improved Ad Retrieval & Ranking[C]//Proceedings of the 20th International Conference on World Wide Web. Hyderabad, India: 2011: 349-353.
- [24] HA L. Crossing offline and online media: a comparison of online advertising on TV web sites and online portals [J]. Journal of Interactive Advertising, 2003, 3(2): 24-35.
- [25] BRODER A, CICCOLO P, GABRILOVICH E, et al. Online expansion of rare queries for sponsored search[C]//Proceedings of the 18th international conference on world wide web. Madrid, Spain: ACM, 2009: 511-520.
- [26] COLLINS-RECTOR M, SHAKLEY C. Targeting advertising using web pages with video: US, 006188398B1[P]. US Patent 6,188,398, 2001.
- [27] AUXIER R, KHERA V, SEIDMAN C. System and method for increasing click through rates of internet banner advertisements: US, 006379251B1[P]. US Patent 6,379,251, 2002.
- [28] BRODER A, FONTOURA M, JOSIFOVSKI V, et al. A semantic approach to contextual advertising[C]//Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Amsterdam: ACM, 2007: 559-566.
- [29] NIU X, MA J, ZHANG D. A Survey of Contextual Advertising[C]// Sixth International Conference on Fuzzy Systems and Knowledge Discovery. IEEE, 2009: 505-509.
- [30] Wikipedia. Display advertising[EB/OL]. http://en.wikipedia.org/wiki/Display_advertising.
- [31] LANG K, DELGADO J, JIANG D, et al. Efficient online ad serving in a display advertising exchange[C]//Proceedings of 4th ACM International Conference on Web Search and Data Mining. 2011: 307-316.
- [32] DoubleClick Rich Media Help. Doubleclick. What is Rich Media? [EB/OL].

- [33] MALHEIROS M, JENNETT C, PATEL S, et al. Too Close for Comfort: A Study of the Effectiveness and Acceptability of Rich-Media Personalized Advertising[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Austin, Texas: ACM, 2012: 579-588.
- [34] PROVOST F, DALESSANDRO B, HOOK R. Audience Selection for On-line Brand Advertising: Privacy-friendly Social Network Targeting[C]//Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris: ACM, 2009: 707-715.
- [35] BAGHERJEIRAN A, PAREKH R. Combining behavioral and social network data for online advertising[C]//2008 IEEE International Conference on Data Mining Workshops. 2008: 837-846.
- [36] Wikipedia. Classified advertising [EB/OL][2013-03-01]. http://en.wikipedia.org/wiki/Classified_advertising.
- [37] BECKER H, BRODER A, GABRILOVICH E, et al. What Happens after an Ad Click? Quantifying the Impact of Landing Pages in Web Advertising[C]//Proceedings of the 18th ACM conference on Information and Knowledge Management. Hong Kong: ACM, 2009: 57-66.
- [38] GE Y, LIU Q, XIONG H, et al. Cost-Aware Travel Tour Recommendation[C]//Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego, California: ACM Press, 2011: 983-991.
- [39] Microsoft. Xbox Advertising[EB/OL][2013-03-01]. <http://advertising.microsoft.com/international/xbox-live>.
- [40] SENGAMEDU S H, SAWANT N, WADHWA S. vAdeo: video advertising system[C]//Proceedings of the 15th international conference on Multimedia. ACM, 2007: 455-456.
- [41] CHEN P, MA W, MANDALAPU S, et al. Ad Serving Using a Compact Allocation Plan[C]//Proceedings of the 13th ACM Conference on Electronic Commerce. 2012, 1(212): 319-336.
- [42] DCCI 互联网数据中心. 中国广告网络蓝皮书 Bluebook of Ad Network Services[R/OL]. 2011[2013-04-15]. <http://www.docci.com.cn/report>.
- [43] WANG Y, BURGNER D, KUZMANOVIC A, et al. Understanding the Network and User-Targeting Properties of Web Advertising Networks[C]//International Conference on Distributed Computing Systems. IEEE, 2011: 613-622.
- [44] OpenX. Ad Networks vs. Ad Exchanges: How They Stack Up - An OpenX whitepaper[R/OL]. 2010[2013-04-15]. <http://www.openx.com>.
- [45] MUTHUKRISHNAN S. Ad Exchanges: Research Issues[C]//Proceedings of the 5th International Workshop on Internet and Network Economics. [S.l.]: Springer, 2009: 1-12.