

基于眼动技术的互联网广告效果研究*

胡晓红^{1a,1b,2}, 王红^{1a,1b,2†}, 任衍具^{1c}

(1. 山东师范大学 a. 信息科学与工程学院; b. 生命科学研究院; c. 心理学院, 济南 250358; 2. 山东省分布式计算机软件新技术重点实验室, 济南 250014)

摘要: 网络广告效果研究中评定广告质量的指标不够准确, 广告配置科学性有待提高。提出观测度的概念, 将它作为衡量广告质量的新标准, 并结合多种信息源对广告记忆度、兴趣度进行建模。实验结果表明, 相比普遍使用显示反馈和广告自身特征, 融合眼动特征来预测广告的观测程度准确性有所提高。此外, 对搜索引擎结果页上位置不同、相关性不同的广告效果进行了定量分析。实验结果表明, 用户的注意力受位置影响显著, 广告的相关性在主体上方对注视次数影响显著, 而下方和右侧表现不显著, 所以可以考虑在主体下方、右侧进行与搜索内容无关商品的推广。

关键词: 观测度; 隐式反馈; 有序 Logistic 回归; 眼动追踪; 网络广告

中图分类号: TP181

文献标志码: A

文章编号: 1001-3695(2018)05-1345-05

doi:10.3969/j.issn.1001-3695.2018.05.014

Research on network advertisement effect based on eye tracking technology

Hu Xiaohong^{1a,1b,2}, Wang Hong^{1a,1b,2†}, Ren Yanju^{1c}

(1. a. School of Information Science & Engineering, b. Institute of Life Sciences, c. School of Psychology, Shandong Normal University, Jinan 250358, China; 2. Shandong Provincial New Technology Key Laboratory for Computer Software Novel Technology, Jinan 250014, China)

Abstract: At present, in the study of the effect of network advertisement, the measure of advertising quality is not accurate enough, and the scientific nature of the ad configuration needs to be improved. So, this paper firstly proposed the new concept, degree of observation, as a new criteria of the advertisement quality, and firstly combined the various information sources to model the advertising memory and interest. The experimental results show that after combining eye movement features, the accuracy to predict the effectiveness of advertising has improved. In addition, it makes a quantitative analysis on the effect of the advertising in different locations and with different relevance to the main content. The experimental results also show that the user's attention is significantly affected by the location of the advertising. And the advertisement relevant to the main content has different influence on the fixation count according to its position. So, it might consider placing ads that are not related to the topic of the content to the bottom or right side of the subject.

Key words: degree of observation; implicit feedback; ordinal Logistic regression; eye tracker; network advertisement

互联网广告作为网络营销的重要组成成分之一, 凭借其覆盖面广、交互性强、成本低廉等诸多特点, 成为主要的网络营销方法之一, 在网络营销方法体系中具有举足轻重的地位。个性化推荐和合理分配广告, 可以完善用户体验, 提高客户满意度, 提高广告系统的性能。

目前, 测量方法有评价指标及分析用户行为日志。评价指标和行为日志的大部分内容是点击日志, 如点击序列、驻留时间等。然而鼠标行为滞后, 产生数据源单一, 所以简单地应用鼠标信息来评估广告效果并不是一种准确、科学的方法。

近年来眼动追踪技术不断成熟。它可以将获取到的视觉信息进行可视化处理, 提供了一种观察用户无意识行为信息的方法。目前该方法已经应用到商业、工业设计、软件评估等众多领域。将眼动追踪技术应用到网络广告效果研究中, 可有效弥补用户信息少、信息不够准确等缺点。深入了解用户心理行

为, 真实地反映用户的兴趣度, 提高广告质量, 使研究信息更加完善, 更加直观。

1 相关工作

点击率是互联网广告最传统和最广泛使用的衡量标准。通过点击模型, 可以根据用户的历史点击数据来描述用户的兴趣和行为, 并预测用户未来的点击行为。现在主要有两种类型的点击模型, 即位置模型和级联模型。2008年, Craswell 等人^[1]提出级联模型。相比位置模型, 级联模型可以更好地解释更高位置的 URL(uniform resource locator)上为何有更多的点击, 但是此时的级联模型仅适用于单点击会话。2009年, Guo 等人^[2,3]在级联模型的基础上进行拓展, 提出适用于多点击会话的 DCM(dependency click model)以及基于贝叶斯结构的 CCM(click chain model)。雅虎实验室的 Chapelle 等人^[4]提

收稿日期: 2017-01-03; 修回日期: 2017-03-08 基金项目: 国家自然科学基金资助项目(61672329, 61373149, 61472233, 61572300, 81273704); 山东省科技计划资助项目(2014GGX101026); 山东省教育科学规划资助项目(ZK1437B010); 山东省泰山学者基金资助项目(TSHW201502038, 20110819); 山东省精品课程项目(2012BK294, 2013BK399, 2013BK402)

作者简介: 胡晓红(1993-), 女, 山东枣庄人, 硕士, 主要研究方向为眼动追踪、机器学习、数据挖掘; 王红(1966-), 女(通信作者), 天津人, 教授, 博士, 主要研究方向为移动社会软件、复杂网络、工作流(wanghong106@163.com); 任衍具(1977-), 男, 山东济宁人, 副教授, 博士, 主要研究方向为视知觉、视觉注意、工作记忆及其应用。

出的 DBN (dynamic Bayesian network click model) 成为目前最广泛应用的点击模型。然而这种方式却忽略了浏览广告与点击广告具有同样的重要性。美国广告研究公司 Adknowledge 的高级副总裁 Zinman 先生说过:“营销人员更应该关注那些占用用户总数的 99% 的无点击广告”。

通过眼动追踪技术,精确收集用户眼动信息,可以确定用户的关注点和浏览轨迹,进而准确判断出用户的兴趣区域和用户行为。Wedel 等人^[5]将广告分为文案、图片和品牌三部分,利用注视次数和面积的比率,发现品牌要素在广告中的意义比重大。Drèze 等人^[6]在研究经验知识对个人注意力的影响时发现,用户对广告浏览路径因个人经验而异;同时指出,目前使用点击率来评估横幅广告的效果是有缺陷的。因为用户可能会注意到了广告但没有点,或者只专注于广告处理的程度。这与 Zinman 先生的结论不谋而合。

虽然目前部分眼动研究已注意到广告记忆度这一概念,但并没有一个完整的模型能够准确描述用户行为信息。并且如今眼动技术研究工作与点击模型工作弊端类似,仅仅是通过眼动参数这一类信息源对用户的行为进行研究,并且研究过程中将不同类型的眼动参数分开考虑,这种方法信息源单一、综合性不高。所以将眼动追踪信息与鼠标行为信息结合成为一种更加理想的方式。

本文旨在以下方面针对网络广告进行定量研究:首次提出广告观测度的概念,并首次结合多种信息源对广告记忆度建模。与以往研究的广告记忆度不同,不再单一统计用户记忆时间等因素,而是结合广告自身特性、用户的显示反馈信息和隐式反馈信息三个方面,构建广告观测度模型,将观测度作为衡量广告质量的重要标准。此外,本文运用数理统计的方法有针对性地分析位置不同、相关性不同的广告对于用户注意力、吸引力的影响。结合观测度模型以及统计检验,分析可知广告效果与各特征之间的关系,从而可以提高互联网广告的质量,优化广告分配效果,使得互联网用户获得更好的交感体验,对互联网广告策略具有一定的意义。

2 观测度

广告的成功不仅取决于阅读广告后是否购买该商品,而更多地应该表现为用户是否注意到并记住该商品,形成品牌效应。鼠标行为具有滞后性,用户反馈具有主观性,因此无法获知用户的真实信息。所以本文提出观测度作为衡量广告质量的重要指标,结合三种信息源建立观测模型,更科学、有效地衡量广告质量。

观测度指的是用户浏览完网页之后对该网页的记忆程度,由于品牌效应追求的是用户长期的了解,所以希望能够反映出用户对广告的注意力。本文将广告的观测度分为四个等级,程度递减。为方便表达,对四个观测等级进行量化:

$$y = \begin{cases} 0 & \text{负向肯定} \\ 1 & \text{负向模糊} \\ 2 & \text{正向模糊} \\ 3 & \text{正向肯定} \end{cases}$$

影响量化等级的特征即为自变量,表示为 $x = (x_1, x_2, \dots, x_k)$ 。

2.1 特征提取

衡量广告质量的标准有很多,但基本上都是基于显示反馈和广告自身信息。根据研究工作,本文发现眼动追踪、鼠标行为以及广告自身参数与广告效果之间关系密切。鼠标特征中

很多特征都与眼动追踪特征具有较强的相关性,但眼动更真实,所以这里只选取具有较强代表性的点击特征。表 1 展示了在本次实验初期预测可能会对用户观测度造成影响的八个特征。

表 1 观测模型特征

特征		特征描述
广告自身特征	广告位置	广告在 SERPs 位置排序,取值为 1、2 或 3,分别代表上方、下方和右方
	广告相关性	广告与搜索内容的相关性,取值为 0 或 1,分别代表不相关、相关
	广告覆盖率	广告面积占 SERPs 面积的百分比
	回访	用户在广告区域回访次数
眼动追踪特征	注视次数	用户在广告区域的注视总数
	注视时效	用户在广告区域的注视时间与总时间的百分比
	平均瞳孔直径	用户在广告区域内每次注视时平均瞳孔直径累加和与注视次数的百分比
点击特征	点击	广告区域是否点击,取值为 0 或 1

为了消除因用户个人阅读风格、网页类型不同带来的差异,将对部分特征进行归一化。例如,将眼睛在广告区域内的注视时间 (fix_time) 修改为

$$\text{注视时效} = \frac{\text{注视时间 (fix_time)}}{\text{结束时间 (end_time)}} \quad (1)$$

其中:结束时间 (end_time) 为用户浏览完整网页所需要的时间。

眼动追踪参数中的瞳孔直径重定义为

$$\text{平均瞳孔直径} = \sum_{i=1}^p \text{注视点 } i \text{ 的平均瞳孔直径} / p \quad (2)$$

其中: p 为该广告区域中注视点总数。

眼动追踪装置得到的眼动参数较多,均方误差、 C_p 统计量和 AIC 准则等特征选择方法计算复杂度非常大,因此本文选取逐步回归方法对连续型变量进行特征选择。

逐步回归方法结合前进法和后退法,在第一、二步,按前进法进行特征引入。例如,第 q 步时将剩余特征集 $\{x_{i_n}^q, x_{i_{n+1}}^q, \dots, x_{i_m}^q\}$ 中的每个元素分别与已选取的特征子集 $\{x_{i_1}^q, x_{i_2}^q, \dots, x_{i_n}^q\}$ 同 y 作多元线性回归,计算 F 统计量 $\{F_{i_n}^q, F_{i_{n+1}}^q, \dots, F_{i_m}^q\}$,选取 F 统计量最大,且大于预先给定的显著性水平 α^{in} 的特征引入。为了防止新引入的特征与已选择特征存在相关关系,从第三步开始进行筛选剔除。同样地,将选择好的全部特征与被选特征结合,同 y 进行多元线性回归,若 F 统计量小于预先给定的显著性水平 α^{out} 的特征,踢出特征子集。依次进行迭代计算,直至特征不再被选择、剔除为止。为了避免死循环,令 $\alpha^{\text{in}} < \alpha^{\text{out}}$ 。

2.2 有序 Logistic 回归

由于观测度的概念首次提出,所以无法与其他研究进行对比。但通过 SVM、决策树、EM 等多种先进的分类算法比较,有序 Logistic 回归预测观测度等级的效果表现最好。

将传统的 Logistic 回归模型写成关于 x 的线性形式:

$$\hat{y} = \text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \gamma + \varphi x \quad (3)$$

其中: γ 和 φ 是待估计的回归系数; $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_k)$; p 是预测 y 取某个值的概率,即样本根据所选取的特征预测观测度属于某一等级的概率。

由于观测等级是程度递增的,所以本文选择 Logistic 回归模型中的累积概率模型。计算达到某种程度的累积概率 \hat{p} 的公式为

$$\begin{aligned}\hat{p}_0 &= p(y=0|x_1, x_2, \dots, x_k) = \frac{1}{1 + \exp^{-(\gamma_1 + \varphi x)}} \\ \hat{p}_1 &= p(y=1|x_1, x_2, \dots, x_k) = \frac{1}{1 + \exp^{-(\gamma_2 + \varphi x)}} \\ \hat{p}_2 &= p(y=2|x_1, x_2, \dots, x_k) = \frac{1}{1 + \exp^{-(\gamma_3 + \varphi x)}}\end{aligned}\quad (4)$$

由此可得 y 取某个值时的概率变量 p 为

$$\begin{aligned}p(y=0|x_1, x_2, \dots, x_k) &= \hat{p}_0 \\ p(y=1|x_1, x_2, \dots, x_k) &= \hat{p}_1 - \hat{p}_0 \\ p(y=2|x_1, x_2, \dots, x_k) &= \hat{p}_2 - \hat{p}_1 \\ p(y=3|x_1, x_2, \dots, x_k) &= 1 - \hat{p}_2\end{aligned}\quad (5)$$

选取概率变量最大时的 y 值作为预测结果。

2.3 模型算法

Logistic 回归的目标函数为

$$J(\gamma, \varphi) = \frac{1}{2}(\hat{y}^{(i)} - y^{(i)})^2 \quad (6)$$

因此,目的就是求能够使目标函数 $J(\gamma, \varphi)$ 的值最小的回归系数 γ 和 φ 。为了表达方便,令 $\theta = (\gamma, \varphi) = (\gamma, \varphi_1, \varphi_2, \dots, \varphi_k)$ 。目标函数作相应变化:

$$J(\theta) = \frac{1}{2}(\hat{y}^{(i)} - y^{(i)})^2 \quad (7)$$

显然,目标函数是一个关于 θ 的二次函数,因此极小值即为最小值。随机梯度下降算法每个样本迭代更新一次参数,更快捷、有效。所以选取随机梯度下降算法求解 θ 同样可以取得较好效果。因此,观测模型算法如算法 1 所示。

算法 1 观测模型算法

输入:训练数据集 $\{(x^1, y^1), (x^2, y^2), \dots, (x^m, y^m)\}$, 迭代次数 T 。

输出:系数矩阵及观测度。

1 初始化回归系数矩阵、学习率

2 for $t=1$ to T do

3 for $i=1$ to m do

4 for $j=1$ to k do

5 利用式(3)得出观测度预测值 $\hat{y}^{(i)}$ 。由 $(y^{(i)} - \hat{y}^{(i)})$

$x_j^{(i)}$ 计算下降梯度,计算学习率 α 。更新回归系数 θ_j : $\theta_j = \theta_j + \alpha(y^{(i)} - \hat{y}^{(i)})x_j^{(i)}$

6 end for

7 end for

8 end for

9 将得到的回归系数矩阵代入式(4)(5),计算最大概率变量,相应的 y 值为观测度预测值

3 实验

3.1 实验准备

3.1.1 实验被试

实验招募 28 名被试人员,男女比例 1:1,年龄在 22~25 岁,平均年龄 24 岁,所有的参与者均为在校研究生。由于希望得到的实验结果更符合大众真实情况,所以被试选取于不同专业的学生。此外,被试人员的裸眼视力或矫正视力均在 1.0 以上,无色盲、色弱等影响视觉效应因素。

3.1.2 实验设备

本实验所采用的是德国普升科技有限公司研发的 SMI RED 眼动仪(version 2.4),其采样频率为 60 Hz。记录和分析过程使用其自带的 IViewX、Experiment Center 和 BeGaze,实验数据的处理和分析使用 IBM SPSS Statistics 22。

3.1.3 实验材料

为保证实验效果更加真实,让被试尽可能像往常一样自由

浏览,如点击和滚动屏幕;并且较好地控制实验变量,保证被试看到的某一类型的搜索引擎结果页是一样的,将从搜索引擎上爬下来搜索引擎结果页,只保留所要研究位置的商业推广。实验过程中,通过眼动仪记录被试的眼动信息,通过网页中嵌入的 JavaScript 代码获取用户的点击信息。

随机选取多种不同类别的物品作为搜索任务,分别为手机、汽车等,并在不同位置分别配置相关、不相关的网页广告。采用 15(搜索任务)×3(广告位置)×2(广告相关、不相关)的设计。

3.2 实验程序

每个实验开始前将记录被试者个人信息,并为被试者展示关于本次实验的简短介绍,包括实验任务及设备硬件。但为了得到最真实的眼动数据和鼠标行为数据,实验所研究的具体内容将对被试者隐藏。随后进行校准工作,被试者要求坐在距离屏幕 65 cm 处,头部固定在桌前的 U 型支架上,调整到舒适的坐姿后开始校准,当被试者追踪水平达到要求后开始正式实验。实验完成后,进行再测试,考察被试者对广告的观测情况。对广告的观测程度分成四个等级,分别为 A 肯定见过、B 好像见过、C 好像没见过、D 肯定没见过,分别为 3~0 分。

4 实验结果与分析

4.1 数据集

从 28 位被试者实验过程中采集 3.1 节中预先提出的特征,其中三个由于机器故障或被试个人原因数据缺失。实验数据如表 2 所示。

表 2 实验数据汇总

		N	百分比/%
问卷统计	A	31	7.4
	B	49	11.8
	C	116	27.9
	D	221	52.9
位置	上方	147	35.3
	下方	141	33.8
	右侧	129	30.9
相关性	相关	208	50
	不相关	209	50
点击	否	337	80.9
	是	80	19.1
合计		417	100

4.2 眼动行为分析

原始材料缩略图如图 1 所示。BeGaze 眼动分析软件可以自动生成可视化的数据图,如 scan path(扫视路径图,图 2)及 heat map(热区图,图 3),可以通过这些数据图进行初步定性分析。

通过对比用户在 SERPs 上的数据图,可以发现用户对搜索结果主体上方的广告注视较多,下部区域次之,右侧区域关注很少,扫视路径更倾向于自上而下的流线型。

4.2.1 不同位置的广告关于眼动参数的比较

统计位于主体不同方位的网络广告区域内注视时效和注视次数的均值,并将其与网络广告区域内总体注视时间、注视次数的均值相比较,如表 3 所示。



图1 原始材料缩略图

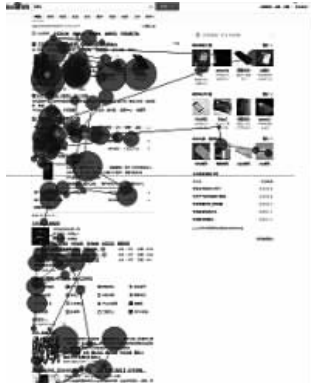


图2 扫视路径图



图3 热区图

表 3 不同位置眼动参数比较

广告位置	注视次数		注视时效		平均瞳孔直径	
	M	SD	M	SD	M	SD
上	20.62	9.514	13.958	7.571	3.715 0	0.422 3
下	15.65	8.768	10.065	5.496 2	3.785 4	0.379 7
右	8.86	7.378	5.795	6.042 3	3.702 6	0.347 8
总计	15.31	8.884	10.121	6.385 4	3.735 0	0.380 8

从表 3 可知,位于 SERPs 上方的网络广告的注视次数、注视时效均呈现出依次大于下方、右侧的网络广告的效果(注视次数为 $20.62 > 15.65 > 8.86$, 注视时效为 $13.958 > 10.065 > 5.795$),而对于平均瞳孔直径参数也表现出上方、下方网络广告的瞳孔直径均值要明显高于右侧平均瞳孔直径。

LSD 检验是在方差分析之后所作的事后比较分析。对所有的多重比较进行简单的多次 T 检验,目的是对每个因素的均值逐对进行比较,以判断具体是哪些水平间存在显著差异。通过不同位置三个眼动参数的方差齐型检验,发现注视次数、注视时效比的 sig 值都小于 0.05,方差齐。因此,为了了解广告不同位置对注视次数、注视时效的影响效果,可以进一步使用 LSD 检验。不同位置眼动的多重均数比较如表 4 所示。

表 4 不同位置眼动的多重均数比较

(I) 位置	(J) 位置	注视次数			注视时效		
		MD(I-J)	SE	sig	MD(I-J)	SE	sig
上	下	4.97	2.891	0.090	3.893	2.278	0.092
	右	11.77 *	2.961	0.000	8.163 *	2.333	0.001
下	上	-4.97	2.891	0.090	-3.893	2.278	0.092
	右	6.80 *	2.991	0.026	4.270	2.357	0.075

注: * 表示 $p < 0.05$ 。

分析表 4 可知,在注视次数方面,上方和右侧的均值差为 11.77, sig = 0.000,即在显著性水平 $\alpha = 0.05$ 的条件下,上方和

右侧的网络广告对于注视次数的影响存在显著差异。同样,下方和右方的均值差为 6.80, sig = 0.026,即在显著性水平 $\alpha = 0.05$ 的条件下,上方和右侧的网络广告对于注视次数的影响存在显著差异,而上方和下方对于注视次数的影响差异不显著。在注视时效比方面,上方和右侧的均值差为 8.163, sig = $0.001 < 0.05$,因此可以认为上方和右侧的网络广告对于注视时效的影响存在显著差异。由此可以得出初步结论,上方和下方的网络广告更加吸引用户,印证了 BeGaze 呈现出的可视化数据结果图中得到的定性分析,用户在搜索引擎结果页面上对上部注视较多,下部区域次之,右侧区域关注很少,扫视路径基本呈现自上而下的流线型。广告位置对用户注意力具有显著影响,是衡量广告质量的重要特征之一。

4.2.2 相关性不同的网络广告对于眼动参数的比较

为了解决一些产品因其特殊性而长期不被大多数人所了解的问题,本文尝试将与主题内容无关的广告加入到搜索引擎结果页面上,探究用户对于无关广告的注意力情况。进一步分析,当网络广告处于不同位置时,与搜索内容相关、不相关的广告对于眼动参数的影响,进而探究广告相关性对于广告质量的影响。相关和不相关的 T 检验如表 5 所示。

表 5 相关和不相关的 T 检验

比较项	平均值	标准差	T	sig
注视次数	15.31	8.884	2.753 *	0.009
注视时效	10.121	6.385 4	-0.959	0.345
平均瞳孔直径	3.735 0	0.380 8	-1.77	0.861

注: * 表示 $p < 0.05$ 。

由表 5 可以看出,网络广告与搜索内容是否相关对注视次数的影响显著 $T = 2.753$, sig = 0.009,说明对于与搜索内容相关的广告,用户注视次数相对较多,吸引力上要强于不相关的网络广告,但是广告相关性的注视时效和平均瞳孔直径差异不显著。不同相关性下眼动参数的描述统计如表 6 所示。

表 6 不同相关性下眼动参数的描述统计

比较项		上方		下方		右方	
		M	SD	M	SD	M	SD
注视次数	无关	17.25	9.64	12.82	5.00	7.36	4.18
	相关	23.17	8.88	16.83	5.44	10.50	6.38
注视时效	无关	12.68	8.95	9.982	5.36	4.518	2.68
	相关	14.22	9.70	10.142	5.86	7.200	8.30
平均瞳孔直径	无关	3.648	0.40	3.847	0.38	3.695	0.33
	相关	3.782	0.45	3.729	0.38	3.711	0.38

表 6 说明总体上在标准差相似的情况下,不同位置,无论是注视次数、注视时效比还是平均瞳孔直径,与搜索内容相关的广告都要高于无关的广告。也就表明网络广告与搜索内容的相关性对用户是有影响的,同样是衡量广告质量的重要特征。

通过对比发现,广告的相关性在主体上方位置差异尤为显著,也就是说,当本文在上方加入无关广告时,会对用户的浏览造成影响,但是在搜索结果主体的下方和右侧加入无关广告,对用户影响较小,这为较特殊物品的推广提供了机会。

4.3 观测度模型建立

4.3.1 特征确定

在本次实验收集到的数据集上对预先提出的特征使用逐步回归方法进行特征选择并建立 Logistic 回归方程。这里令引

人和剔除的显著性水平分别为 $\alpha^{\text{in}} = 0.1$, $\alpha^{\text{out}} = 0.15$ 。通过分析得出,第一步引入的特征为注视点数目,第二步引入的特征为广告平均覆盖率,其余三个连续型特征依次引入,并无被剔除的特征。

广告位置、广告相关性以及点击为分类特征,在 4.2 节中已经验证其重要性,因此不再进行筛选。

4.3.2 模型验证

使用 SPSS Statistics 22 统计软件进行数据分析,通过对比引入特征后卡方值的大小来判断模型是否有效。模型拟合信息如表 7 所示。

表 7 模型拟合信息

模型	2 * 对数似然值	卡方	df	显著性
仅常数项	154.586			
最终	104.619	49.966	9	0.000

从表 7 可知,引入特征后,在显著性水平 $p < 0.001$ 的情况下,卡方值为 49.966,说明在该显著性水平下,至少有一个自变量的偏回归系数不为 0,即本模型是有效的。

为了证明本文提出加入隐式反馈信息即眼动追踪信息这种方式的有效性,将使用不同的信息源进行有序多分类 Logistic 回归,对比包含不同信息源模型的准确率。由于数据集较小,所以这里采用五折交叉验证。

图 4 展示了不同特征预测用户观测度的准确率。“自身”代表仅含有广告自身信息特征,“自身+鼠标”代表既有广告自身特征又含有点击特征,“自身+鼠标+眼动”代表含有三种信息源的特征。

加入眼动追踪信息的综合模型,即使在低迭代次数的情况下仍明显高于前两种模型,当迭代次数达到 50 次时,三种信息源模型比仅广告自身信息预测类别的准确率高 7%,比广告自身信息结合鼠标行为信息预测类别的准确率高 3.5%,说明在广告自身信息以及显示反馈的基础上,增加隐式反馈特征的方法来预测用户观测度的方法表现出了不错的效果。

由于观测度的概念为首次提出,首次结合多种信息源对广告记忆度、兴趣度进行建模,并且相关广告研究工作中的行为数据不可获取,所以本文在数据集上进行多种算法测试。这里给出了宏平均表现最好的三种方法,详见图 5。可以看出有序 Logistic 回归表现最好。分析原因是观测度具有等级递增的性质,而有序 Logistic 回归使用的是累积概率,这与观测度的性质是一致的,所以分类效果更准确、科学。

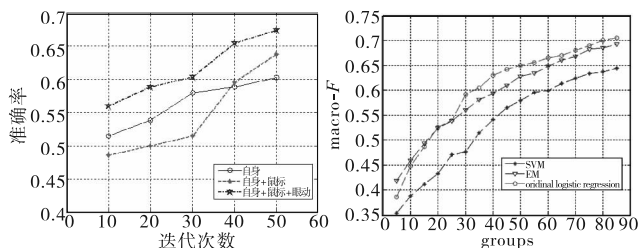


图4 准确率对比

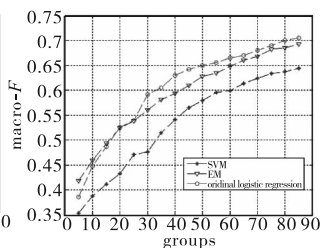


图5 Macro-F对比

选用准确率 P 、召回率 R 以及 F_1 值来评测不同分类算法在本实验数据集上的分类效果。为了方便显示,这里只展示宏平均 F 。

$$\text{macro_}P = \frac{1}{n} \sum_{i=1}^n P_i \quad (8)$$

$$\text{macro_}R = \frac{1}{n} \sum_{i=1}^n R_i \quad (9)$$

$$\text{macro_}F = \frac{2 \times \text{macro_}P \times \text{macro_}R}{\text{macro_}P + \text{macro_}R} \quad (10)$$

5 结束语

本文提出利用观测度作为广告质量的衡量指标,解决了之前广告记忆度、点击率等目前网络广告研究中存在的数据源单一、信息不可靠等问题,通过结合广告自身、眼动以及鼠标行为数据三种信息源,将自身、隐式反馈信息与显示反馈信息结合起来,预测用户对广告的观测程度,替代传统的评估原则,更加科学、可信。实验结果证明,对比仅广告自身信息与结合鼠标行为信息,加入隐式反馈信息,模型准确率有所提升。

通过对位置不同、相关性不同的网络广告进行眼动行为分析,发现广告位置和相关性是影响用户注意力的显著变量,同时印证了本文针对观测度提出的多个眼动追踪特征是较好的分类特征。对比不同位置广告的眼动追踪参数发现,用户对位于主体上方的广告注意力多于下方和右侧。广告的相关性在上方表现出的差异显著高于下方和右侧,对用户影响较大。因此得出结论,与搜索内容相关性的广告应尽量放在主体上方,而与搜索内容无关的广告可以考虑放在主体下方或右侧,为较特殊物品的推广提供了机会。

结合观测度模型和数理统计,可以提高互联网广告的质量,优化广告分配效果以及用户的兴趣度,使得互联网用户获得更好的交感体验,对于互联网广告策略具有一定的意义。

参考文献:

- [1] Craswell N, Zoeter O, Taylor M, et al. An experimental comparison of click position-bias models[C]//Proc of International Conference on Web Search and Data Mining. New York: ACM Press, 2008: 87-94.
- [2] Guo Fan, Liu Chao, Kannan A, et al. Click chain model in Web search[C]//Proc of International Conference on World Wide Web. New York: ACM Press, 2009: 11-20.
- [3] Guo Fan, Liu Chao, Wang Yimin. Efficient multiple-click models in Web search[C]//Proc of the 2nd ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2009: 124-131.
- [4] Chapelle O, Zhang Ya. A dynamic Bayesian network click model for Web search ranking[C]//Proc of the 18th International Conference on World Wide Web. New York: ACM Press, 2009: 1-10.
- [5] Wedel M, Pieters R. Eye fixations on advertisements and memory for brands: a model and findings[J]. Marketing Science, 2000, 19(4): 297-312.
- [6] Drèze X, Hussherr F X. Internet advertising: is anybody watching? [J]. Journal of Interactive Marketing, 2003, 17(4): 8-23.
- [7] Joachims T, Granka L, Pan Bing, et al. Accurately interpreting click through data as implicit feedback[C]//Proc of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2005: 154-161.
- [8] Mao Jiaxin, Liu Yiqun, Zhang Min, et al. Estimating credibility of user clicks with mouse movement and eye-tracking information[J]. Communications in Computer & Information Science, 2014, 496: 263-274.
- [9] 丁锦红,王军,张钦. 平面广告中图形与文本加工差异的眼动研究[J]. 心理学探新, 2004, 24(4): 30-34.
- [10] 瞿珍. 网络广告视觉搜索的眼动研究[D]. 杭州: 浙江师范大学, 2009.

需要改进。首先,CG和DTG的构建不能完全自动化。这是因为CPF的语法非常复杂,例如可以包含多层嵌套的条件表达式或概率分布函数,有些情况下需要人工解析或者使用额外的数据结构来标注。然而这并非本文方法独有的缺点,因为在Fast Downward中,这两类图的创建也是手工完成的。但是如何从RDDL描述中自动构建CG和DTG,是一个非常值得研究的子课题。其次,在CGH_{RDDL}算法中,固定轮数 h_{fix} 是一个可调参数,其值目前本文设置为最大轮数 H 的1/2或者1/3。 h_{fix} 的意义在于,由于PPP没有固定的目标状态,所以状态的演化无论在求解模拟中还是在估值中必须设置固定的步数。与其他的启发步长一样, h_{fix} 的设置显然会影响启发值计算的效果,太大则耗费时间,太小则不够精准。在未来工作中,可研究如何设置一个可变的 h_{fix} 值,使得其随着估值效果自适应地调整。最后,CGH_{RDDL}算法目前只能用于RDDL描述的MDP版本,需进一步将其扩展到RDDL描述的POMDP版本。这会涉及到对可观测流文字(observ-fluent)的特殊处理。

参考文献:

- [1] 饶东宁,李建华,蒋志华,等. 并行概率规划综述[J]. 计算机应用研究,2016,33(6):1607-1611.
 - [2] Ghoshchi N, Namazi M, Newton M, *et al.* Transition constraints for parallel planning[C]//Proc of the 29th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2015:3268-3274.
 - [3] Vianna L, Barros L, Sanner S. Real-time symbolic dynamic programming[C]//Proc of 29th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2015:3402-3408.
 - [4] Sanner S. Relational dynamic influence diagram language (RDDL): language description [EB/OL]. (2011). <http://users.cecs.anu.edu.au/~ssanner/IPPC2011/RDDL.pdf>.
 - [5] Keller T, Helmert M. Trial-based heuristic tree search for finite horizon MDPs[C]//Proc of the 23rd International Conference on Automated Planning and Scheduling. Palo Alto, CA: AAAI Press, 2013:135-143.
 - [6] Kolobov A, Dai peng, Mausam D S, *et al.* Reverse iterative deepening for finite-horizon MDPs with large branching factors[C]//Proc of the 22nd International Conference on Automated Planning and Scheduling. Palo Alto, CA: AAAI Press, 2012:146-154.
 - [7] Srinivasan S, Talvitie E, Bowling M. Improving exploration in UCT using local manifolds[C]//Proc of the 29th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2015:3386-3392.
 - [8] Bonet B, Geffner H. Labeled RTDP: improving the convergence of real-time dynamic programming[C]//Proc of the 13th International Conference on Automated Planning and Scheduling. Palo Alto, CA: AAAI Press, 2003:12-21.
 - [9] Ghallab M, Nau D, Traverso P. 自动规划——理论和实践[M]. 姜云飞,杨强,凌应标,译. 北京:清华大学出版社,2008.
 - [10] Geffner H, Bonet B. A concise introduction to models and methods for automated planning [M]//Synthesis Lectures on Artificial Intelligence and Machine Learning. Williston, VT: Morgan & Claypool, 2013:1-141.
 - [11] Hoffmann J, Nebel B. The FF planning system: fast plan generation through heuristic search [J]. *Journal of Artificial Intelligence Research*, 2001, 14(1):253-302.
 - [12] Bonet B, Geffner H. Planning as heuristic search[J]. *Artificial Intelligence*, 2001, 129(1-2):5-33.
 - [13] Helmert M. The fast downward planning system[J]. *Journal of Artificial Intelligence Research*, 2006, 26(1):191-246.
 - [14] Helmert M, Geffner H. Unifying the causal graph and additive heuristics[C]//Proc of the 18th International Conference on Automated Planning and Scheduling. 2008:140-147.
 - [15] Rao Dongning, Jiang Zhihua. Learning planning domain descriptions in RDDL [J]. *International Journal on Artificial Intelligence Tools*, 2015, 24(3):1550002.
 - [16] Piacentini C, Alimisis V, Fox M, *et al.* An extension of metric temporal planning with application to AC voltage control [J]. *Artificial Intelligence*, 2015, 229(11):210-245.
 - [17] Núñez S, Borrajo D, López C L. Automatic construction of optimal static sequential portfolios for AI planning and beyond [J]. *Artificial Intelligence*, 2015, 226(9):75-101.
 - [18] Gerevini A, Long D. Plan constraints and preferences in PDDL3: the language of the fifth international planning competition, R. T. 2005-08-47 [R]. Italy: University of Brescia, 2005.
 - [19] Younes H L S, Littman M L. PPDDL 1. 0: an extension to PDDL for expressing planning domains with probabilistic effects, CMU-CS-04-167 [R]. Pittsburgh, PA: Carnegie Mellon University, 2004.
 - [20] Manfredotti C V, Messina E V. Relational dynamic Bayesian networks to improve multi-target tracking[C]//Proc of the 11th International Conference on Advanced Concepts for Intelligent Vision Systems. Berlin: Springer-Verlag, 2009:528-539.
 - [21] 饶东宁,蒋志华,姜云飞,等. 在部分观测环境下学习规划领域的派生谓词规则[J]. 计算机学报, 2015, 38(7):1372-1385.
-
- (上接第1349页)
- [11] 周象贤,金志成. 国外对平面广告受众注意心理的眼动研究[J]. 心理科学进展, 2006, 14(2):287-293.
 - [12] Eickhoff C, Dungs S, Tran V. An eye-tracking study of query reformulation[C]//Proc of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2015:13-22.
 - [13] Liu Zeyang, Liu Yiqun, Zhou Ke, *et al.* Influence of vertical result in Web search examination[C]//Proc of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2015:193-202.
 - [14] Chen Ye, Liu Yiqun, Zhou Ke, *et al.* Does vertical bring more satisfaction? Predicting search satisfaction in a heterogeneous environment [C]//Proc of the 24th ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2015:1581-1590.
 - [15] Liu Yiqun, Chen Ye, Sun Jiashen, *et al.* Different users, different opinions: predicting search satisfaction with mouse movement information[C]//Proc of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM Press, 2015:493-502.
 - [16] 曹卫真,车笑琼,祁祿,等. 画面主体位置布局的眼动实验及对网络视频资源建设的启示[J]. 远程教育杂志, 2013(5):97-106.
 - [17] 裴道方,钟悦. 考虑行为和眼动跟踪的用户兴趣模型[J]. 河南科技大学学报:自然科学版, 2014, 35(1):49-52.
 - [18] 施卓敏,郑婉怡. 探秘不同认知风格的个体关注广告的差异——广告位置和认知风格对广告效果影响的眼动研究[J]. 营销科学学报, 2014, 10(3):128-145.
 - [19] 柴林麟. 大数据时代下互联网广告及计算广告学的应用研究[J]. 信息与电脑:理论版, 2015(16):29-31, 56.
 - [20] 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.