

淘宝网定向精准广告 简介

李强

2011年11月

让用户不知不觉爱上你



数据，数据，还是数据

| 店铺浏览 | 宝贝浏览 | 店铺收藏 | 宝贝收藏 | 宝贝购买 | 店铺名称 | 店标 | 主营 | 店铺所在地 |
|------|------|------|------|------|-----------------------------------|---|------------------------------------|-------|
| 1 | 0 | 0 | 0 | 0 | 【雨的亭 型男装】百款夏装5折起包邮！迎61满61元减6.1元 |  | CCTV胡子大叔★100% 真人实物拍摄 让网购成为一件快乐的事儿！ | 北京 |
| 0 | 1 | 0 | 0 | 0 | 王道家居专营店 |  | 居家日用/收纳/礼品 | 杭州 |
| 1 | 0 | 0 | 0 | 0 | 匠工房 |  | 日本 潮牌 皇冠 正品 马克华菲 时尚 潮流 | 北京 |
| 1 | 4 | 0 | 0 | 0 | 时尚空间 男装旗舰店 正品 时尚男装 休闲男装 商务休闲 古老鲨鱼 |  | 精品男装 休闲男装 商务男装 专柜正品 古老鲨鱼 | 广州 |
| 0 | 1 | 0 | 0 | 0 | 尚蒂旗舰店 |  | 尚蒂/床上用品/四件套/被子/毛毯/枕头/床垫/靠垫/毛巾/ | 南通 |
| 0 | 1 | 0 | 0 | 1 | 领文图书专营店 |  | 外语/考试/小说/公务员/程序/正版/ | 南京 |
| 1 | 0 | 0 | 0 | 0 | 美库化妆品专营店 |  | 美容护肤/美体/精油 | 广州 |
| 3 | 2 | 0 | 0 | 0 | 尚蒂旗舰店 |  | 尚蒂/床上用品/四件套/被子/毛毯/枕头/床垫/靠垫/毛巾/ | 南通 |
| 0 | 1 | 0 | 0 | 0 | 恒丰数码专营 |  | 电脑硬件/显示器/电脑周边 | 北京 |
| 1 | 0 | 0 | 0 | 0 | LINDA自制欧美范★精致白领时尚家★ |  | 全球流行服饰 | 大连 |

我知道你昨晚在干什么！

Category:

| # | Category | Cat Name | Weight | Timestamp | Behavior [srch,brow,fav,list,buy] | Rankinfo |
|---|----------|----------|--------|------------------------------------|--------------------------------------|--|
| 1 | 1623 | 半身裙 | 10000 | Tue, 14 Jun 2011 15:45:02 +0800 | 0,1,0,0,0 | |
| 2 | 162105 | 小背心/小吊带 | 10000 | Tue, 14 Jun 2011 22:39:09 +0800 | 0,1,0,0,0 | |
| 3 | 50010850 | 连衣裙 | 23947 | Wed, 15 Jun 2011 09:40:43 +0800 | 0,6,0,0,0 | |
| 4 | 0 | | 25315 | Wed, 15 Jun 2011 14:10:31 +0800 | 13,1,0,0,0 | s,hmanh,1; s,臭美小野猪,2; s,aliang608,1; s,我是阿卡,1; s,hmanh,1; s,臭美小野猪,1; s,臭美小野猪,1; s,布丁可爱5555,1; s,zhouhongling1,1; s,aliang608,1 |

Shop ID:

购物兴趣类目

购买力

店铺

品牌

性别

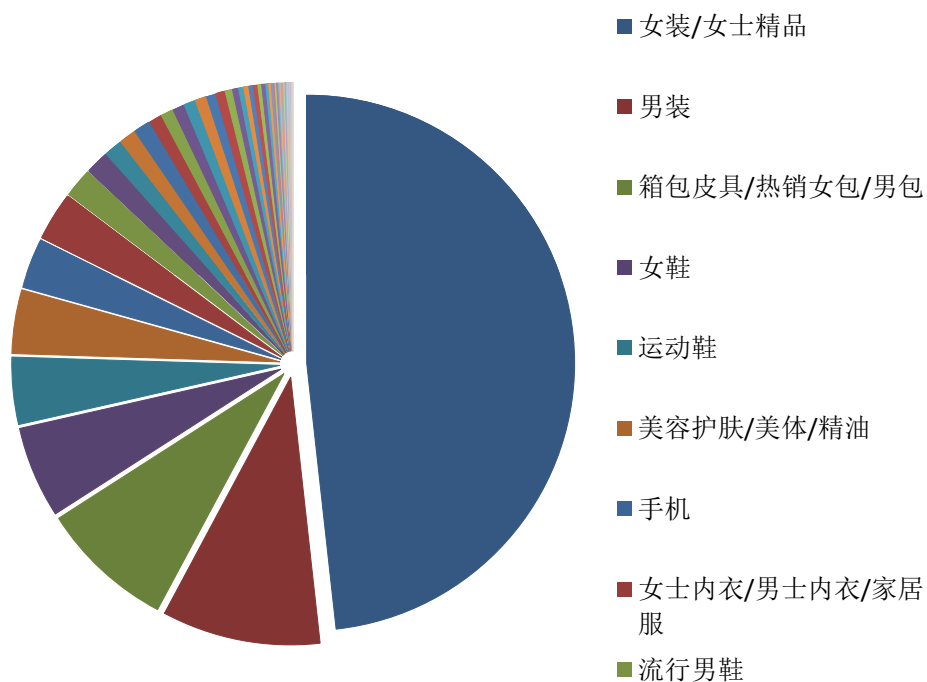
社会属性

年龄

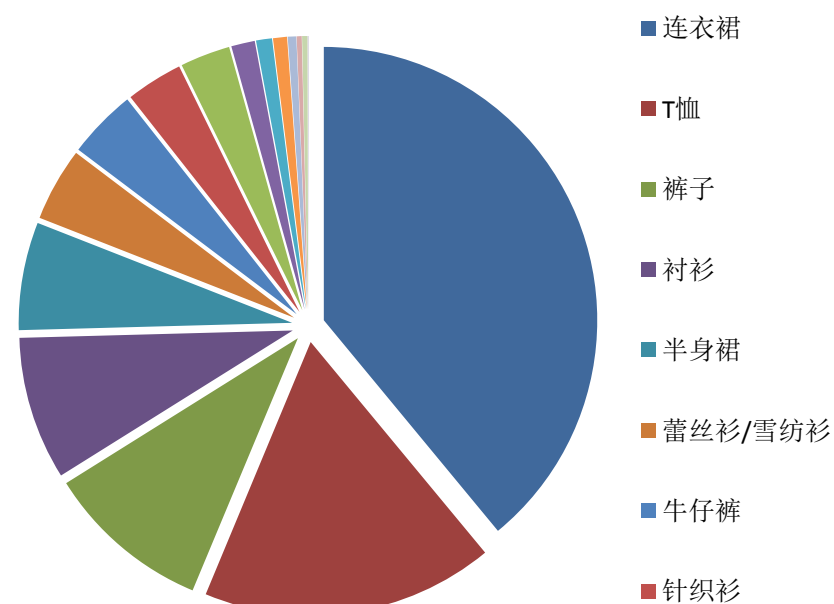


我知道你昨晚在干什么！ - 继续

淘宝一级类目



女装细分



25000000 per day 淘宝站内

竞价排名 (auction) 海量计算， 点击率预估等

| 二级类目 | | 直接成交 | 直接成交 | 间接成交 | 间接成交 | 成交金额 | 广告展现 | 消耗(元) | 点击 | ROI | CTR | PPC |
|------|------|------|--------|------|--------|--------|-----------|--------|-------|------|--------|------|
| 连衣裙 | 16 # | 294 | 43,126 | 37 | 3,268 | 46,394 | 2,715,248 | 10,373 | 21499 | 4.47 | 0.792% | 0.48 |
| 连衣裙 | 16 # | 12 | 2,218 | 85 | 13,962 | 16,180 | 2,208,176 | 4,535 | 6612 | 3.57 | 0.299% | 0.69 |
| 连衣裙 | 16 # | 29 | 8,001 | 98 | 11,642 | 19,643 | 2,128,505 | 6,398 | 9137 | 3.07 | 0.429% | 0.70 |
| 连衣裙 | 16 # | 5 | 438 | 56 | 3,626 | 4,064 | 1,861,398 | 4,605 | 4890 | 0.88 | 0.263% | 0.94 |
| 连衣裙 | 16 # | 12 | 3,581 | 42 | 8,681 | 12,262 | 1,751,266 | 4,204 | 5004 | 2.92 | 0.286% | 0.84 |
| 连衣裙 | 16 # | 67 | 3,148 | 112 | 5,735 | 8,883 | 1,544,450 | 2,355 | 2683 | 3.77 | 0.174% | 0.88 |
| 连衣裙 | 16 # | 49 | 13,560 | 19 | 2,191 | 15,751 | 1,460,929 | 3,806 | 3933 | 4.14 | 0.269% | 0.97 |
| 连衣裙 | 16 # | 5 | 1,246 | 23 | 3,560 | 4,806 | 1,365,573 | 2,799 | 3137 | 1.72 | 0.230% | 0.89 |
| 连衣裙 | 16 # | 16 | 1,156 | 52 | 3,472 | 4,628 | 1,256,801 | 2,540 | 3388 | 1.82 | 0.270% | 0.75 |
| 连衣裙 | 16 # | 23 | 3,019 | 20 | 1,929 | 4,948 | 1,183,086 | 1,972 | 2173 | 2.51 | 0.184% | 0.91 |
| 连衣裙 | 16 # | 51 | 4,097 | 57 | 3,977 | 8,074 | 1,131,643 | 1,877 | 3246 | 4.30 | 0.287% | 0.58 |
| 连衣裙 | 16 # | 44 | 3,385 | 74 | 4,366 | 7,751 | 891,786 | 2,523 | 4221 | 3.07 | 0.473% | 0.60 |
| 连衣裙 | 16 # | 18 | 2,085 | 8 | 519 | 2,604 | 785,910 | 1,099 | 1461 | 2.37 | 0.186% | 0.75 |
| 连衣裙 | 16 # | - | - | 3 | 1,679 | 1,679 | 663,984 | 1,305 | 1461 | 1.29 | 0.220% | 0.89 |
| 连衣裙 | 16 # | 2 | 364 | 5 | 442 | 806 | 660,085 | 1,746 | 2166 | 0.46 | 0.328% | 0.81 |
| 连衣裙 | 16 # | 2 | 287 | 3 | 336 | 623 | 541,008 | 1,266 | 1408 | 0.49 | 0.260% | 0.90 |
| 连衣裙 | 16 # | 4 | 1,079 | 7 | 1,626 | 2,705 | 486,792 | 1,426 | 1723 | 1.90 | 0.354% | 0.83 |
| 连衣裙 | 16 # | 11 | 1,096 | 3 | 183 | 1,279 | 475,776 | 1,052 | 1107 | 1.22 | 0.233% | 0.95 |
| 连衣裙 | 16 # | - | - | 7 | 2,264 | 2,264 | 471,949 | 1,806 | 2186 | 1.25 | 0.463% | 0.83 |
| 连衣裙 | 16 # | 6 | 1,525 | 8 | 1,743 | 3,268 | 460,855 | 1,647 | 1971 | 1.98 | 0.428% | 0.84 |

eCPM的定义

对于匹配步骤取出的 $N=1000$ 个广告，我们计算它们的 eCPM 值。eCPM 理论值定义如下：

$$eCPM = \frac{\text{成交额}}{\text{展现数}} * 1000$$

也就是每一千次展现带来的广告成交额。我们把上述公式分解为下面几块：

$$eCPM = \frac{\text{点击数}}{\text{展现数}} * \frac{\text{成交数}}{\text{点击数}} * \text{价格} * 1000$$

其中：

$$\text{点击率} = CTR = \frac{\text{点击数}}{\text{展现数}}$$

$$\text{转化率} = ACR = \frac{\text{成交转化数}}{\text{点击数}}$$

在绝大多数情况下，因为广告数目太多，没有足够的流量让我们直接计算 eCPM，点击率和转化率。所以我们在下面公式里采用相关性系数来模拟点击率以及转化率。

$$\text{伪 eCPM} = (\text{相关性系数})^{\alpha} * \text{bidPrice}$$

eCPM排序的case study

| rank | 伪 <u>eCPM score</u> | match quality score | bid | PPC |
|------|------------------------|---------------------------|-----|---------------------------|
| 1 | $90 \times 100 = 9000$ | 90 | 100 | $(80/90) \times 110 = 97$ |
| 2 | $80 \times 110 = 8800$ | 80 | 110 | $(95/80) \times 80 = 95$ |
| 3 | $95 \times 80 = 7600$ | 95 | 80 | $(80/95) \times 80 = 67$ |
| 4 | $80 \times 80 = 6400$ | 80 | 80 | -- |

Rank 是由 eCPM score 来确定的。最后如果点击发生的话,按相应的 PPC 计费。

eCPM排序减轻“赢者诅咒”

对第三步中抽取出来 M 个广告,进行 eCPM 竞价并且计算每位的 PPC。

$$eCPM = S^{\alpha} * B$$

其中 S 是相关性系数, B 是 bidPrice。

对于排名 K 的宝贝 PPC, 我们用排名下一位 (K+1 位) 的宝贝的 bidPrice, 也就是 B (K+1) 来计算。

$$PPC(K) = \frac{S^{\alpha}(K+1)}{S^{\alpha}(K)} B(K+1)$$

。

如果我们采取上面的计算公式, 因为我们的排序规则是使用 eCPM, 。

$$eCPM(K+1) < eCPM(K)$$

$$S^{\alpha}(K+1)B(K+1) < S^{\alpha}(K)B(K)$$

把上面公式带入 PPC 的计算公式, 我们就自然保证了。

$$PPC(K) < B(K)$$

什么是高质量的广告？

REMOVE TO IMPROVE YOUR CREATIVE IDEAS



都是美图，结果不同



¥ 255.00

镇店之宝 超气质 2011新款 气质雪纺连衣裙

最近成交3532笔



0.79% CTR

0.48元PPC

270W PV

成交46000 元



0.3% CTR

0.69元PPC

220W PV

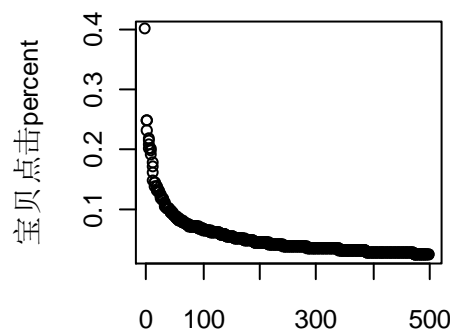
成交16000 元

Advertising 长尾规则

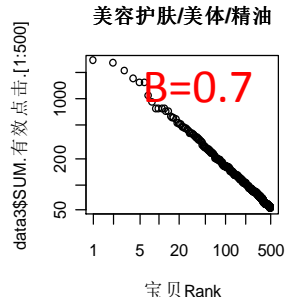
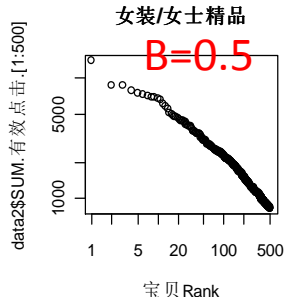
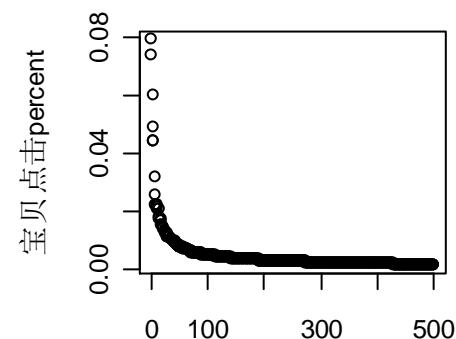
$$Z = \frac{A}{K^B}$$



女装/女士精品



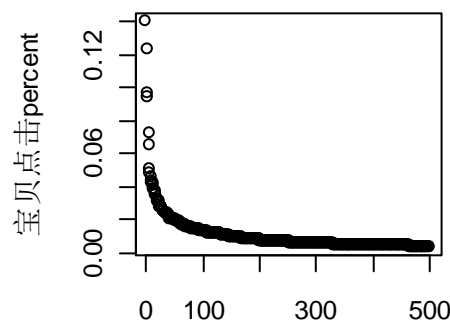
美容护肤/美体/精油



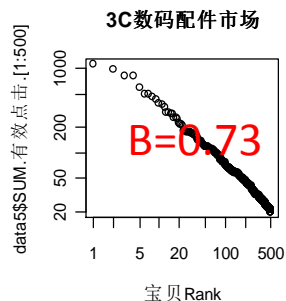
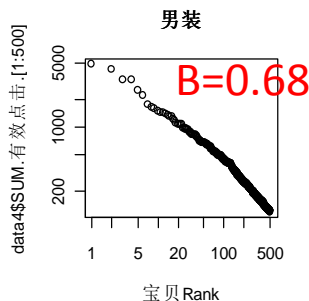
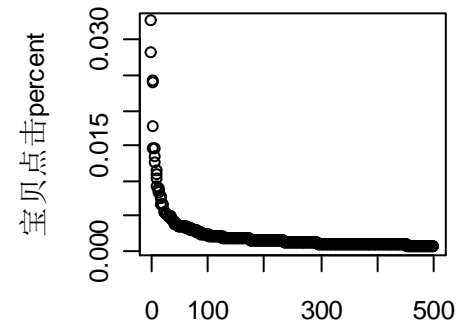
宝贝Rank

宝贝Rank

男装



3C数码配件市场



宝贝Rank

宝贝Rank



让市场更加公平：互联网让中小广告主获得流量：博弈论

- 互联网颠覆了商业规则中的
 - 传统规则中的80-20 rule
 - 假如我们有100件商品，最畅销的20个商品获得80% 成交
- 根据“长尾理论”的分析和我们从淘宝广告市场获得的数据
 - 女装：top20商品获得40%的成交
 - 男装：top20商品获得51%的成交
 - 美容：top20商品获得52%的成交
 - 数码：top20商品或得54%的成交
- 互联网降低了小人物进入市场的障碍
 - 淘宝女装市场的基尼系数：0.333
 - 男装：0.517
 - 美容：0.532
 - 数码：0.577
 - 按照经济学理论，基尼系数在0.6以上，则会引起市场的垄断，降低活力。过低的基尼系数则过于平均，市场也缺乏活力。
 - 比较：
 - 2010年中国居民收入的基尼系数已经超过0.5
 - 美国居民收入的基尼系数：0.4
 - 欧盟地区（英法德）：0.33左右

基于用户关系的推荐 Collaborative filtering



Like



2

Also Like

3

May Like



基于用户关系的推荐

- 思路：
 - 基于用户-店铺-用户 关系的挖掘
- 数据源：
 - 用户浏览收藏购买店铺
- 计算方法：
 - 关联推荐 & 协同过滤
- 效果：
 - CTR 是通投的6倍，人群定向的3倍

流量价值最大化 Optimization problem



流量价值最大化

- 思路：
 - N个item，要展示在K个位置上($K < N$)，使整体效果获得某种评价准则下的最大收益
- 数据源：
 - 页面所有item实时的点击数据
- 计算方法：
 - Exploration and Exploit (E&E)
- 效果：
 - CTR 是通投的6倍，人群定向的3倍

流量价值最大化-参考原理

Deterministic policy: UCB1.

Initialization: Play each machine once.

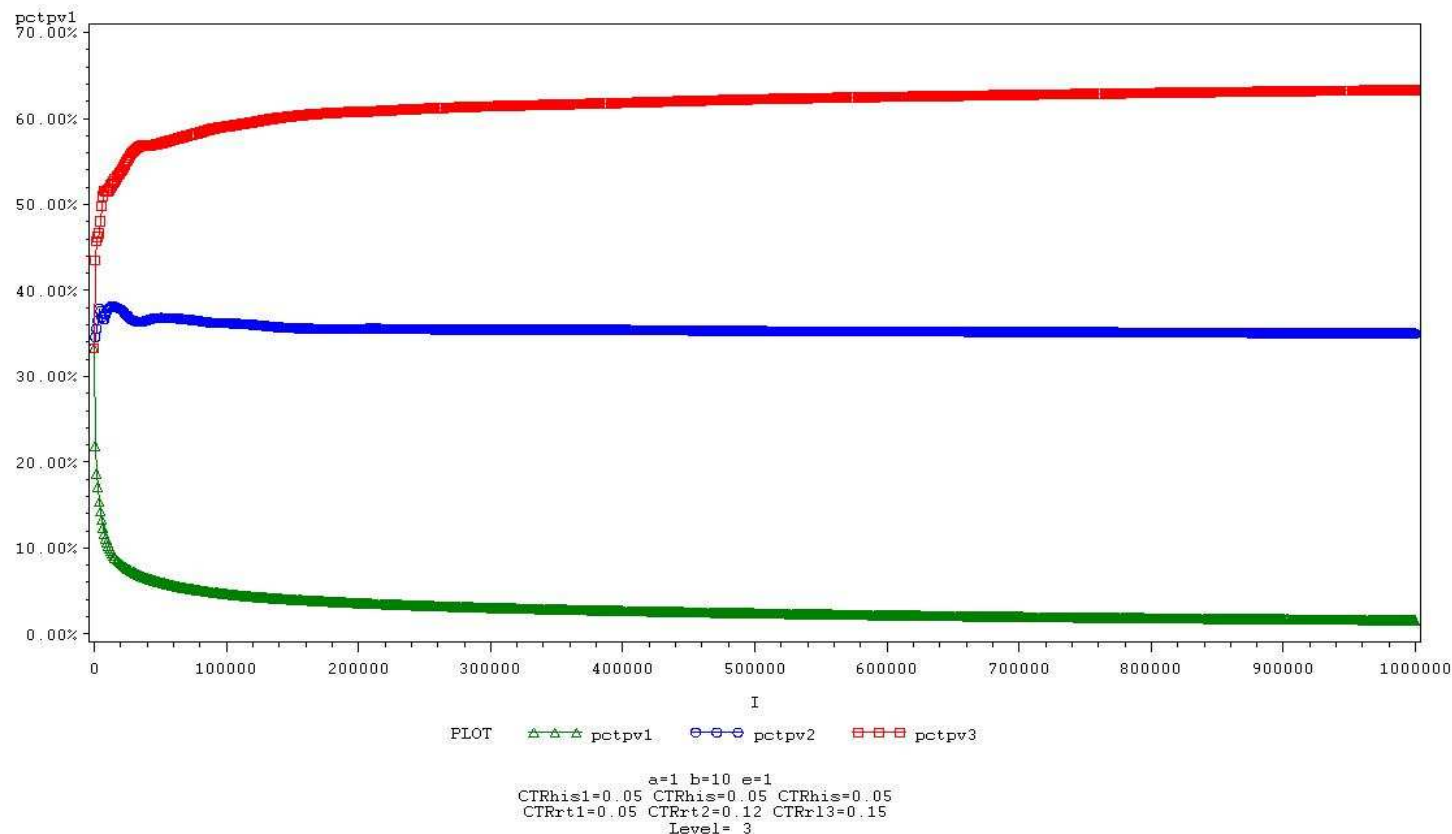
Loop:

- Play machine j that maximizes $\bar{x}_j + \sqrt{\frac{2 \ln n}{n_j}}$, where \bar{x}_j is the average reward obtained from machine j , n_j is the number of times machine j has been played so far, and n is the overall number of plays done so far.

参考文献

Auer P, Cesa-Blanchi N, Fischer P. Finite-time Analysis of the Multi-armed Bandit Problem[C]// Proceedings of the 19th International Conference on Machine Learning. Hingham, USA: [s. n.], 2002: 235-256

流量价值最大化-效果



2.87%

2.88%

2.91%

2.93%

2.97%

3.01%

3.06%

3.10%

3.10%

线上表现：
CTR持续上升

线下模拟：

开始时，商品的展现几率接近均等；

随着展现的增加，高CTR的商品会快速获得更高的展现概率