

手机用户画像在大数据平台的实现方案

张 慷

(中国电信上海公司政企客户部,上海 200041)

摘要:大数据平台通过对 DPI 上网日志的加工处理和分析,形成用户上网行为的特征刻画,同时结合经分系统现有用户基础信息、套餐信息、终端信息、业务订购等信息,进行深度数据融合及交叉分析,形成手机用户画像,实现电信业务针对性营销。

关键词:Hadoop,DPI 上网日志;上网行为特征;用户画像;针对性营销

中图分类号:TP391.9

文献标识码:A

文章编号:1673-1131(2014)02-0266-02

1 系统概述

大数据平台通过对 DPI 上网日志的加工处理和分析,形成用户上网行为的特征刻画;同时结合经分系统现有用户基础信息、套餐信息、终端信息、业务订购等信息,进行深度数据融合及交叉分析,形成手机用户画像。然后根据用户画像分析结果,向用户有针对性地营销电信增值业务或合作方业务,提高电信市场竞争力。

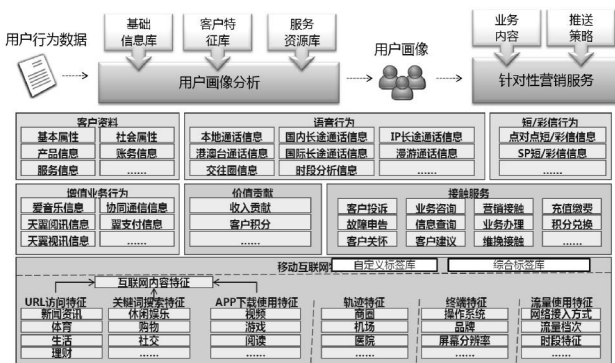


图 1 用户画像分析总体架构

2 解决方案

2.1 用户画像分析主要过程

用户画像分析主要过程如图 2 所示。

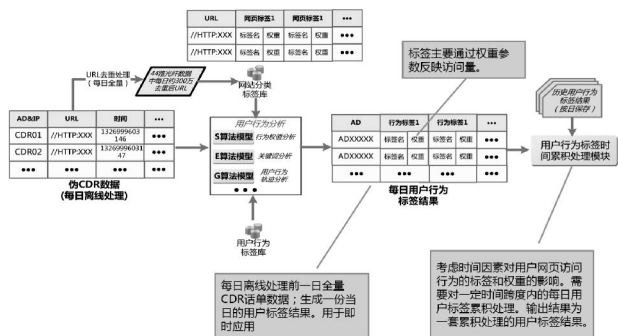


图 2 用户画像分析主要过程

网站分类标签可根据用户需求动态变化,支持多级标签,包括三级标签和四级标签。网站分类标签示例如图 3 所示。

3542	www.yymp3.com	音乐	YYMP3流行音乐网
3543	yalasol.com	音乐	亚拉索完全音乐空间
3544	yy.qqmusic.qq.com	音乐	QQ音乐
3545	zhangmen.baidu.com	音乐	百度MP3音乐掌门人
3546	eat.cncn.com	饮食	饮食资讯
3547	eat.online.sh.cn	饮食	上海热线吃在上海
3548	food.online.cq.cn	饮食	巴渝食府
3549	meal.dayoo.com	饮食	大洋网——食全食美

图 3 网站分类标签

用户行为标签示例如图 4 所示。

一级	二级
电信经营分类	3G手机潜在用户群、189邮箱用户群、社交爱好者、视频爱好者、宠物爱好者、时尚人群、婚恋男女、体育新闻关注者、高尔夫爱好者、赛车爱好者、大型网游玩家
娱乐消费群	家庭套餐潜在用户群、IPHONE潜在用户群、网购爱好者、文学爱好者、星座爱好者、娱乐八卦、追星群体、团购爱好者、户外运动爱好者、健身爱好者、羽毛球爱好者
运动爱好者	网球爱好者、网络游戏爱好者、对战平台玩家
游戏爱好者	

图 4 用户行为标签

用户的某个行为标签,需要包含访问次数和权重两个参数,其中权重需要通过纵向均衡和横向均衡两次修正后,才能得到具体的结果。例如:某个用户有三种访问行为,分别是新闻(150次)、小说(30次)、音乐(20次)。

(1) 聚类功能纵向分为 10 个等级,每个访问次数先确定等级,然后再根据等级范围,确定等级内的详细数字。新闻(150次)纵向均衡后的数值是 5.79,小说(30次)纵向均衡后的数值是 3.42,音乐(20次)纵向均衡后的数值是 4.58。

(2) 横向均衡就是将不同标签的数值,按照进行百分比折算即可。

(3) 最终输出的权重数值。

表 1 标签权重

新闻	小说	音乐
150 次/42.0%	30 次/24.8%	20 次/33.2%

2.2 数据挖掘

数据挖掘可以根据用户的要求建立多个手机上网的分析平面,通过数据聚类进行分析挖掘。例如:①第一平面为上网次数/流量;②第二平面为上网次数/时长;③第三平面为上网测试时段。

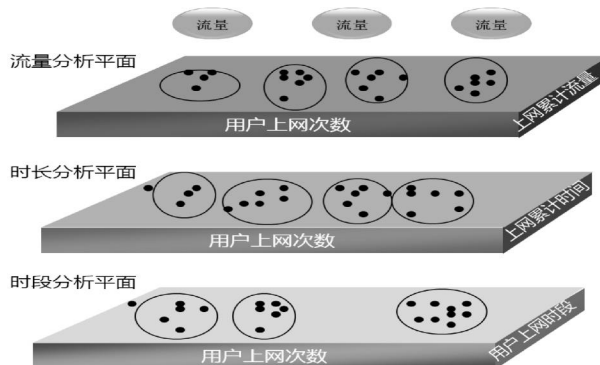


图 5 多平面分析数据模型

数据聚类方法如下:先将数据 Vector 向量化/归一化,接着设置聚类输出数目或中心点,最后选择聚类算法以及收敛参数。数据聚类算法包括:

- K-Means Clustering
- Canopy Clustering

LED 显示屏亮度及色度调整技术应用

王鹏宇, 宁 妍

(辽宁省葫芦岛市 92493 部队 3 分队 辽宁 葫芦岛 125000)

摘要 LED 显示屏作为一项新型的技术应用在娱乐市场当中,以其高亮度、宽色域以及节能环保等优点逐渐成为室外展示市场的主角。但基于 LED 具有一定的离散性,导致显示屏在显示的过程中不能保证亮度和色度的一致性和均匀性,需要对显示屏进行色度及亮度的调整。文章将介绍 LED 显示屏的色度亮度调整技术的应用。

关键词 LED ;显示屏 ;亮度 ;色度 ;调整 ;逐点

中图分类号 :TN9

文献标识码 :A

文章编号 :1673-1131(2014)02-0267-02

LED 显示屏是由很多个 LED 组成,由于离散性的问题导致各个 LED 在组成后会出现色度和亮度不均的情况。如果 LED 组成的批次不同,则离散性问题就会更大,色度和亮度的均匀性、一致性也会更差,影响观赏的效果。对此,必须对 LED 显示屏进行色度及亮度的调整,以保证色度和亮度的均匀性和一致性。

1 LED 显示屏的均匀性和一致性

评定 LED 显示屏的画质质量好坏的标准是亮度和色度是否具有均匀性和一致性。导致 LED 显示屏均匀性的问题的原因有很多,而 LED 的离散性问题是最主要的问题之一,其次还有导致 LED 显示屏一致性问题的时间衰减的问题等。离散性问题是 LED 普遍存在的问题,一般会在生产过程中根据一定的标准进行 LED 的分档和筛选,这也是 LED 生产工艺当中的最后一个流程,就是通过分光分色机将不合乎规格的 LED 进

行筛选和剔除,并将合乎规格的 LED 进行分档。

不同的厂家分档的标准是不一样的,有的是按电气指标进行分档,有的是按光学指标进行分档。当然,分档的情况可以通过设置进行调节,若想提高 LED 显示屏的整体画面质量的均匀性和一致性的话,可以通过两个指标共同实现分档。这样的分档方法虽然能够在一定程度上保证 LED 显示屏的画面质量,但却降低了 LED 显示屏的整体合格率,从而使 LED 显示屏的价格升高。目前常用的主波长分档属于光学指标分档,这种分档方式是十分不科学。这种分档方式只是单纯地根据光强和主波长来对 LED 的性能进行判断,而单纯的光强和主波长是无法完全的、准确地表示出光学的特性,所以这种判断方法的误差性很高。

2 LED 显示屏亮度及色度调整的原理

LED 显示屏以像素为基本单位,而每一个像素又是由

-Mean Shift Clustering
-Hierarchical Clustering
-Dimensional Reduction
-Expectation Maximization
-Logistic Regression
-Neural Network
-Hidden Markov Models
-Random Forests
-Parallel Frequent Pattern Mining

.....

用户上网行为分析结果示例如图 6 所示。

目标用户	关联次数	P2P 下载	P2P 流媒体	流媒体	浏览	彩信	下载	即时通信	游戏	娱乐	电子商务	VOIP	网管数据库	病毒类
133****0708	434			4										
133****7480	427		2										4	
189****3095	427												1	
133****0268	424	2		0	2				3	2	3	2	1	3
189****4280	421												4	
189****1119	413													
189****5971	413							0	2	0	4			
133****9983	412			0					1	4				
133****0326	410	0			0				0	4	0			
189****6571	409	0							2	3	0			4
133****6303	407													2
189****6830	398								1					4
189****6770	395	2							1	4		1		
189****6268	393			2										
189****3600	393	0	3		2									
189****7977	392													
133****0251	391	1								4		0		0
189****0021	390				4									
189****0030	385				4				0					3
189****0233	381													
133****1477	379				0				2	4	0		0	4

图 6 用户上网行为分析结果

2.3 客户画像示例

手机客户画像:

(1) 拥有的手机品牌和型号 HTC;

(2) 手机套餐标准 88RMB/月;

(3) 手机上网习惯:

——上网时间 8:00-9:00/11:30-14:30/18:30-21:00/.....

——喜欢浏览的网站 新浪娱乐、腾讯;

——喜欢使用的软件 视频(电视剧/电影/娱乐新闻)团购/微博/蘑菇街...

——是否购买产品 频次 是 2 次/周;

——产品的类型 美食团购/时尚衣裳。

3 结语

大数据平台按设定的行为标签、数据模型等规则对 DPI 上网日志的加工处理和数据挖掘,形成的用户行为特征数据,包括用户互联网关注点分析、相同兴趣点、用户群信息等。同时结合经分系统现有用户基础信息、套餐信息、终端信息、业务订购等信息,进行深度数据融合及交叉分析,形成手机用户画像。

电信根据大数据平台用户画像分析结果,可以向用户有针对性地营销增值业务或合作方业务,提高市场竞争力。

参考文献:

- [1] 周正,陈枫.“大数据时代”来了——专访国防信息学院研究所所长孟宝宏.解放军报,2013(1):17
- [2] 顾芳,刘旭峰,左超.大数据背景下运营商移动互联网发展策略研究[J].邮电设计技术,2012(8)
- [3] 陈如明.大数据时代的挑战、价值与应对策略[J].移动通信,2012(17)
- [4] 胡弘毅.大数据分析有助电信企业“有的放矢”[J].赛迪网,2013,6(20)