

doi:10.16055/j.issn.1672-058X.2015.0001.019

# 用户行为轨迹研究\*

林雪云

(福建师范大学福清分校 电子信息学院,福建 福清 350300)

**摘要:**针对目前使用最广泛的谷歌搜索、淘宝网、开心网用户行为轨迹的优缺点,提出“从用户寻找信息”的被动服务模式到“信息寻找用户”的主动服务模式,利用 Clickstream 研究建设相应的数据信息采集方法,构造用户轨迹研究建设思路——采集、分析、推荐、评价、反馈,抽象提出六大模型体系,分析了三大应用场景:应用内容推荐、关注应用推荐、辅助重点应用保障,并对用户行为轨迹改造进行案例分析应用。

**关键词:**应用场景;用户行为轨迹;被动服务模式;主动服务模式

**中图分类号:**TP131.2      **文献标志码:**A      **文章编号:**1672-058X(2015)01-0078-05

## 0 前言

第一种谷歌搜索——因人而异的搜索结果、提升搜索质量,google 根据用户行为分析结果,将搜索引擎做到因人而异,不同的用户,虽然输入相同的关键字,却产生不一样的搜索结果,例如将用户关注的内容放在用户搜索结果的前面,从而提升用户的搜索体验和搜索质量。第二种淘宝网——优化界面服务,增强用户购物体验。淘宝网通过对用户行为的微观研究,近 3 年来各种类型统计研究,更加了解用户网购行为,更清晰的看到一个完整的网购路径,进而能更好的指导淘宝对不同特征的网购用户进行针对性的服务优化。第三种开心网——细分用户群体,打造个性化应用服务。开心网分析新老用户对应用的使用偏好,发现新用户对社会化媒体有着更强的娱乐性偏好,而老用户更多注重社会化媒体的多元化内容和资讯,以及与好友间的互动与分享。通过用户偏好的分析结果,可以为不同的用户群体提供个性化的应用服务。

如何能够从“用户寻找信息”的被动服务模式转变为“信息寻找用户”的主动服务模式成为目前亟待解决的关键问题<sup>[1]</sup>。借助 google、淘宝和开心网等互联网标杆企业的思路和应用场景,构建符合中国电子商务的用户行为轨迹分析技术,实现由“用户寻找信息”的被动服务模式转变为“信息寻找用户”的主动服务模式,满足差异化的用户需求,提升企业系统的用户体验和人性化服务水平<sup>[1]</sup>。

## 1 用户轨迹研究建设思路

以“信息寻找用户”的主动服务模式为基本原则和目标,借鉴互联网业界标杆企业的分析思路和成功经验,方案采用“采集、分析、推荐、评价、反馈”闭环的 5 步法建设思路。从系统的各个模块,采集用户行为轨迹的相关数据,为用户的行为轨迹研究提供数据保障。建立用户行为轨迹模型,对用户行为轨迹进行分析,研究用户偏好,为“信息寻找用户”的主动服务模式提供技术支撑。根据用户行为轨迹的分析结果,获取用户偏好信息,为用户推荐符合自己喜好的内容,实现“信息寻找用户”的主动服务。根据内容推荐的评价机制,方便收集用户对主动服务的评价,方便后续行为轨迹分析模型的优化。对用户的评价内容进行分析,找

收稿日期:2014-06-01;修回日期:2014-07-09.

\* 基金项目:福建省教育厅 B 类项目(JB13197).

作者简介:林雪云(1976-),女,福建闽侯人,硕士,副教授,从事数据挖掘研究.

出问题的根源,对轨迹分析模型进行循环优化,并将优化结果反馈给用户,实现闭环可持续的精准主动服务<sup>[2]</sup>。表 1 介绍了用户行为轨迹功能结构。

表 1 用户行为轨迹研究功能架构

访问层	系统门户				
应用层	用户行为偏好推荐	用户行为分析统计			用户评价反馈
	用户偏好应用推荐	应用访问情况统计			用户推荐评价 8
	用户偏好内容推荐	重点应用使用情况统计			
	用户偏好界面推荐	用户贡献度分析统计			模型优化反馈
	...	...			
用户行为轨迹基础分析					
服务层	用户关注应用分析	用户关注内容分析	用户操作过程分析	系统优化分析	
	应用关联分析	用户活跃度分析	...		
数据层	用户行为分析服务	用户行为偏好推荐服务	用户行为分析统计服务	用户评价反馈服务	...
	用户行为轨迹模型				
获取层	用户关注应用模型	用户关注内容模型	用户操作过程模型	系统优化模型	
	应用关联模型	用户活跃度模型	...		
	用户行为采集数据	用户基本信息 鼠标滚动条信息	应用基本信息 搜索关键词	用户登陆信息 应用收藏 评价信息	应用访问信息 页面提交 响应信息
获取层	AIG		Clickstream		...

1.1 目前热门电子商务业界用户行为轨迹信息采集的参考

谷歌分析是由 Google 提供的网站流量统计服务,它记录了用户在浏览网站时的浏览器类型、访问途径、搜索关键字等信息,为网站优化提供详细报告支持(图 1)。淘宝用户人群行为特征分析是在产品设计时,通过对商品浏览、消费类型、消费偏好、使用偏好等行为进行分析,针对各个特征群体投放精准商品广告。

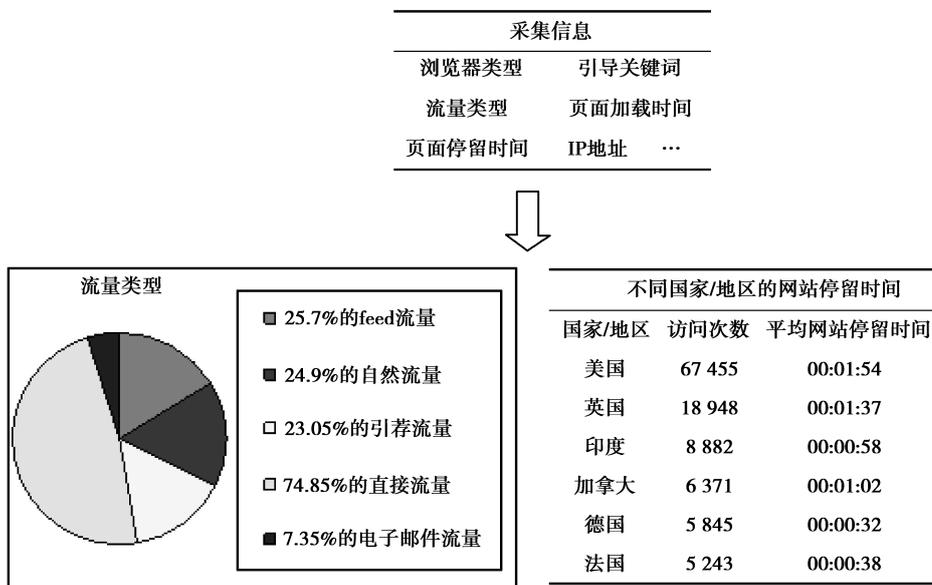


图 1 谷歌用户浏览状况

## 1.2 用户行为轨迹信息采集的实现方法

以电子工业出版社出版的,由美国数据仓库领域专家,斯威格特等作者编写的《点击流数据仓库》的用户行为轨迹理论为基础,借鉴 Google、淘宝等互联网业界成功经验,用户行为轨迹信息的采集主要以“用户行为”为主线,贯穿用户访问系统、使用系统、退出系统的整个过程。具体采集的数据和实现方法如表 2 所示。

信息的采集主要从页面与后台,基于开源的 Clickstream 等用户行为采集技术,配合 Javascript 脚本插件、日志截止等方法来获取。

表 2 采集的数据和实现方法

系统访问	基础数据			用户行为轨迹库
	IP 地址	分辨率	浏览器类型版本	用户关注应用模型
界面操作	使用网段	登录时间	用户编码	应用内容模型
	应用编码	应用路径	应用访问时间	用户操作过程模型
	应用类型	维度条件	维度条件值	系统优化模型
业务操作	鼠标位置	事件节点	鼠标事件时间	关联应用模型
	输出字段	响应时间	活动时间 ...	用户活跃度模型

## 1.3 用户行为轨迹信息采集关键技术 (Clickstream)

Clickstream 是一个基于 Java 的开源项目,主要被用于跟踪用户在 Web 服务器上的浏览历史,通过历史数据可以分析出用户访问网站的路径,瓶颈,热点连接等。用户行为轨迹信息的采集主要从页面与后台,基于开源的 Clickstream 用户行为采集技术,配合 Javascript 脚本插件、日志截取等方法来获取。图 2 介绍了其功能流程。

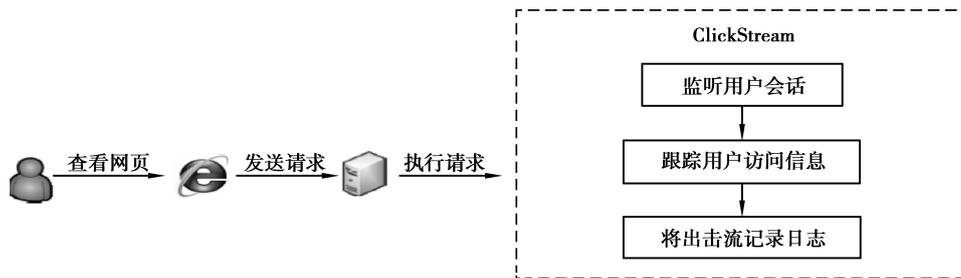


图 2 信息采集关键技术 (Clickstream)

## 2 用户行为轨迹的六大模型体系

(1) 用户关注应用模型:其目标是将用户对各个应用的使用次数以及有效使用时间综合评估,预测对该用户可能使用到的应用。应用目标在于常用应用推荐、重点应用推荐,提升用户应用关注度及应用访问效率。提升新上线应用推广能力。模型构建主要通过用户编码、访问时间、应用编码、应用路径、应用类型、活动时间等要素来创建相对应的用户关注应用模型,主要包含常用应用关注度、重点应用关注度、新上线应用匹配度。

(2) 应用内容模型:其目标是通过用户对用户常用维度条件、维度值、查询输出字段、指标查看等信息收集,统计分析最常用的内容。应用目标在于在页面初始化时自动应用分析后的数据,调整页面结构,提高用户

使用感知,自动剔除非常用输出字段,提高用户查看指标效率。模型构建主要通过操作时间、维度条件、输出字段、指标编码等等要素来创建相对应的用户关注应用模型,主要包含维度条件使用率、维度条件值使用率、输出字段使用率、指标关注度。

(3) 用户操作过程模型:其目标是通过收集用户在页面操作的顺序和响应时间,鼠标在控件上的操作效率以及内容操作的完整性,协助应用优化改造。应用目标在于提高用户应用效率,分析、判断可能存在的性能问题,改进应用设计,优化用户体验。模型构建主要通过操作顺序、响应时间、内容完整性操作次数,鼠标点击控件及时间等要素来创建相对应的应用内容模型,主要包括操作效率、响应效率、内容完整情况。

(4) 应用优化模型:其目标是通过用户的应用基本信息和应用运行信息等,构建系统优化相关的信息模型,包括用户界面行为特征、系统应用运行特征、系统集中访问行为特征等。

(5) 关联应用模型:其目标是通过用户以及同角色用户的相关应用访问信息,构建关联应用模型,主要包括用户管理应用行为特征等。应用目标主要在可用于同角色关注应用分析、应用关联度分析、个人关联访问应用分析等,提升用户的应用访问效率,方便用户开展日常工作。

(6) 用户活跃度模型:其目标是根据用户的系统登录信息,应用访问信息等系统使用信息,来构建用户活跃度模型,主要包括用户活跃度行为特征等内容。应用目标主要应用于分析用户活跃时间、应用活跃时间、活跃终端情况等,方便进行用户贡献度的统计等,提升用户使用系统的积极性。

### 3 应用场景

(1) 应用内容推荐。用户在应用过程中发现维度条件太多、经常使用的维度不是在最前面、条件值每次都要重新选择,操作不方便。通过分析,发现用户经常使用的只有归属地市、时间和品牌等几个维度。解决措施:对用户经常使用的维度进行推荐展现,不经常使用维度进行隐藏,推荐的操作顺序可以根据用户经常操作的维度顺序,自动进行维度展现的排版,自动推荐填充用户经常选择或者输入的条件值<sup>[3-6]</sup>。

(2) 关注应用推荐。用户在应用过程中发现系统太大,功能节点太多,每次都必需点击好多层才能找到自己需要使用的功能页,且关注的应用分散在不同地方,不方便使用。根据用户关注应用模型,分析用户日常工作经常访问的应用,发现功能节点太过繁多,系统左边的树节点过于庞大,造成用户使用很不方便,且用户关注的应用分散在不同地方,不方便使用,而且用户经常使用的仅有几个 KPI 和即席查询应用。解决措施:将用户经常使用的 KPI 和即席查询应用主动推荐到用户的个性化办公桌面,或者收藏夹,或者首页等用户可以直接使用应用的地方,方便用户使用这些应用。

(3) 辅助重点应用保障。用户在应用过程中发现某些重点应用,用户反馈经常重复查询了好多次,都查不到数据。根据用户关注应用模型和系统优化模型进行重点应用确认,并分析重点应用的集中访问时间。解决措施:后台 UTAP 根据用户的访问时间要求,进行任务的优先级调整,保障重点应用优先执行。

### 4 用户行为轨迹实施改造示例

用户行为轨迹信息的采集,是通过 JS 插码,Http 访问拦截,JAVA 代码动态植入等技术来实现的,因此用户行为轨迹的实施,需要原有系统进行一定工作量的配合改造。如经分系统需要进行系统管理、框架管理、即席查询、KPI、固定报表、多维分析和客户研究平台等相关功能模块的改造。

经过实践,用户行为轨迹信息流如下:用户登录→选择应用→操作应用→退出系统。

用户登录需进行系统管理的改造,采集用户登录时的客户端类型、版本、分辨率,登录的门户类型等信

息,为用户活跃度模型、用户操作过程模型等提供基础数据。选择应用需进行框架管理的改造,采集用户选择应用的操作路径、应用类型、应用选择时间、应用关闭时间等信息,为用户关注应用模型、用户操作过程模型、关联应用模型、系统优化模型等提供基础数据。操作应用需进行即席查询、KPI、固定报表、多维分析等功能模块的改造,采集用户应用操作的相关信息,包括鼠标点击、维度、指标选择、查询条件选择顺序、界面停留时长等,为用户关注内容模型、用户操作过程模型、用户事件点击模型,系统优化模型等提供基础数据。

#### 参考文献:

- [1] 陈康,黄晓宇,王爱宝,等.基于位置信息的用户行为轨迹分析与应用综述,电信科学,2013(4):118-123
- [2] 刘分,葛国栋,赵宇,等.移动通信网络位置群体关联用户挖掘[J].计算机应用,2013,33(8):2100-2103
- [3] 陈熙,用户浏览轨迹算法分析与应用[J].信息与电脑,2010(8):129-130
- [4] 苏孝强,倪宏,马承帅.移动环境下基于用户轨迹的社交亲近度研究[J].微计算机信息,2011,27(5):7-9
- [5] 郭镇虎.基于事件轨迹的局部行为模型的合并[J].计算机软件和原理,2010(4):129-130
- [6] 贾金营,张凤荔.位置隐私保护技术综述[J].计算机应用研究,2013(3):641-646

## Research on User Behavior Trajectory

**LIN Xue-yun**

(School of Electronics and Information, Fuqing Branch of Fujian Normal University,  
Fujian Fuqing 350300, China)

**Abstract:** According to the advantages and disadvantages of the user behavior trajectory in the most widely used networks such as Google search, Taobao and Happy Network, that the passive service model of “user seeking information” is changed into the active service model of “information seeking users” is presented, the corresponding data information collection method is built by clickstream research, the user trajectory research construction including six model systems such as collection, analysis, recommendation, evaluation, feedback and abstraction is built, three application scenarios such as application content recommendation, concerned application recommendation and aided important application security are analyzed, and the case analysis and application for user behavior trajectory reconstruction is conducted.

**Key words:** application scenario; user behavior trajectory; passive service model; active service model

责任编辑:田 静