



【A12】用户行为的深度追踪

用户行为分析平台

——【恒生电子】

江南大学-物设商联合队

尹达恒、姜海鹏、王宇歆、高超宏、刁寒钰

作品概要

- 用户画像
- 实时流量
- 顾客行为分析
- 埋点与无埋点
- 历史多时间维度关联
- 区域热点
- Spark+hadoop大数据分析
- 自适应显示

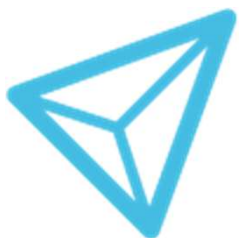


安全性

埋点成本

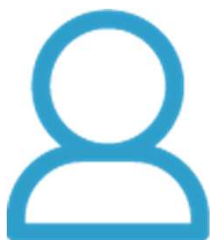
数据可视化

操作简单 数据精准



无埋点与埋点结合

- 更精准的采集数据，可应对软件升级造成的问题



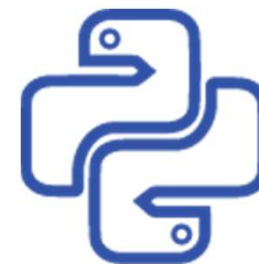
用户画像

- 无须请IT写代码，通过埋点数据分析形成客户画像。让企业及时掌握客户的最新动态



Pyspark大数据

- Spark是先进的大数据分布式编程和计算框架



Python快速开发

- 代码少易快速开发，可以快速为公司提供个性化开发

用户需求

统计网站流量、拥挤程度，为页面优化提供参考；
分析新增用户，确定网站发展态势，为未来的发展规划提供参考；
分析用户群的偏好与兴趣，为个性化宣传提供参考；
分析用户的集中分布地域，把握市场份额；
判断用户的忠诚度或是对网站的依赖/喜爱程度；
用户的注册完成度；
……等等等等

还有最重要的一点：

将这些数据把握在自己手里

功能概览

数据收集:

- 网页埋点
- 捕获用户行为
- 收集用户信息
- 监控网页访问情况

用户概况统计:

- 统计用户变化
- 统计用户活跃度
- 分析用户粘性
- 多角度统计用户组成

网站概况统计:

- 统计网站访问情况
- 统计热点页面
- 统计热门产品

WebSiteAnalysisKit

用户行为深度追踪:

- 用户群行为聚类
- 页面访问关联

统计结果呈现:

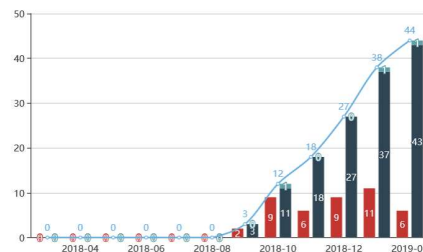
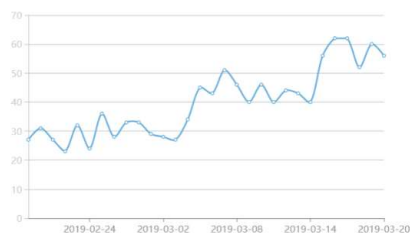
- 动态仪表盘
- 多维度数据图表联动
- 多时间跨度呈现

系统扩展:

- 低耦合的系统部件
- 尽可能完整的数据收集
- 容易扩展分析方法

系统层次

数据呈现层 (网站管理者终端)



数据分析层 (云端服务器)

MySQL分析结果数据库

流量统计

用户量统计

页面热度统计

产品热度统计

用户群聚类

页面关系统计

python

Spark

MySQL用户动作数据库

数据源 (网站用户终端)

用户动作数据采集

页面埋点

页面埋点

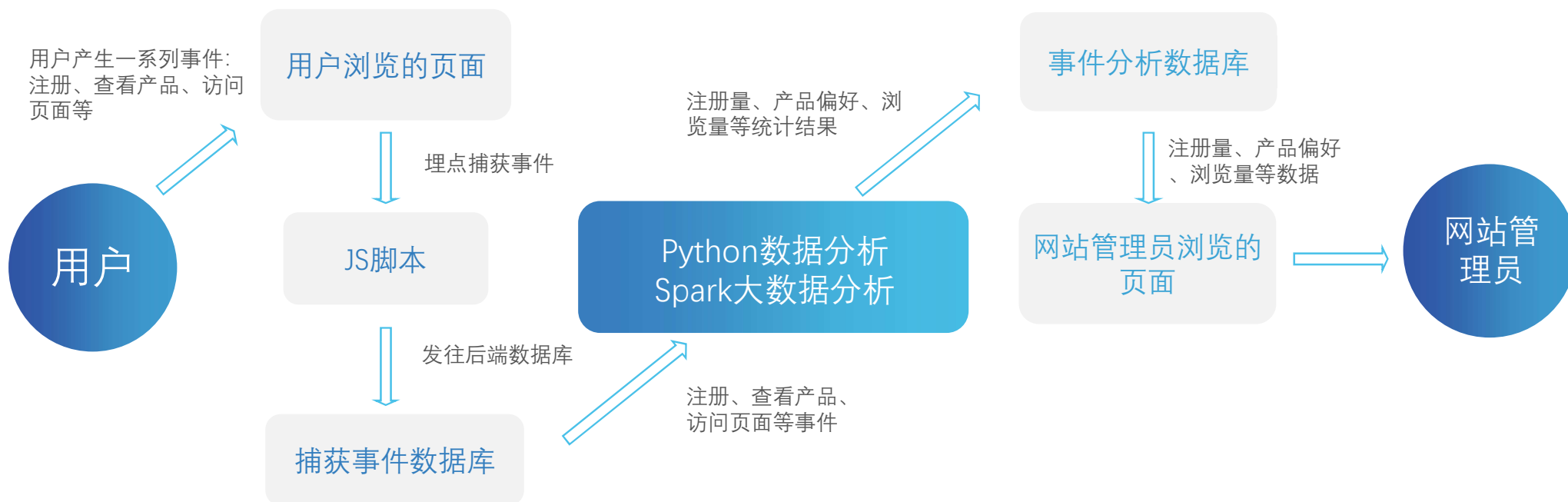
页面埋点

页面埋点

页面埋点

页面埋点

数据流



技术细节：网页埋点和数据收集



埋点数据收集原理：

WSK.js修改JavaScript原生prototype
监控页面上添加的所有click事件

定义函数WSK_RecordEvent()向后端
发送监控数据

基本用户信息+事件描述

后端接收到监控数据存入数据库

埋点方式：

在页面上引入埋点文件WSK.js

指定收集数据的发送位置

在要埋点页面元素的点击回调上添加
埋点函数WSK_RecordEvent()

数据收集：

后台Servlet接收前端WSK.js发来的
监控数据

收集数据通过JDBC写入事件数
据库

终端分析-python

基础数据：

前端捕获脚本获得的用户设备型号

后端通过IP地址判断的地理位置

分析方法：

统计各类设备占比

统计各地理位置的用户量

流量分析-python

基础数据：

存入数据库的“访问”事件

“访问”事件附带的页面地址

分析方法：

统计各页面的访问量

每天统计、每月统计

用户分析-python

基础数据：

存入数据库的“注册完成”事件

存入数据库的“点击产品”事件

“浏览产品”事件附带的用户信息

分析方法：

统计各用户对不同产品的浏览量

统计各产品的用户浏览量

技术细节：用户行为深度分析



用户消费能力分析-Spark

描述：让公司为顾客制定更精准的消费方案

基础数据：

前端捕获用户消费金额，浏览时间，访问页面次数等内容

分析方法：

将用户行为打包成特征集，通过Kmeans均值聚类将用户分类，分类标准通过改变模型训练的簇的个数变化来得到最优解。

呈现方式：饼图

页面访问分析-python

描述：发现页面之间的频繁项集和发现关联规则

基础数据：

前端捕获每个session访问的页面次序

分析方法：

通过apriori算法确定某两个页面中的频繁项

呈现方式：关系图

其他大数据分析

描述：数据（例如页面访问量）量大，通过大数据分析可减少分析时间

基础数据：

前端捕获的频繁数据集

分析方法：

通过pyspark将数据导入，处理，并保存至数据库

呈现方式：多种图例

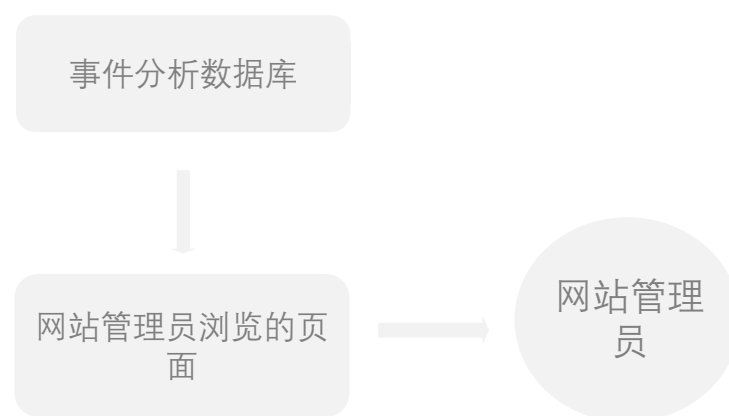
系统优势



- 部分数据自动埋点自动收集
- 手动埋点简单易行
- 收集到尽可能多的数据以备未来更加深入的数据分析
- 以最简单直接的方式发送和存储数据



- 与数据的收集和呈现部分互相独立，耦合度低
- 任何一种可以连接MySQL数据库的语言都能满足要求



- 使用Echart动态呈现数据
- 多图表联动使数据视图清晰简洁



谢谢

江南大学-物设商联合队

尹达恒、姜海鹏、王宇歆、高超宏、刁寒钰