实时分析驱动用户资产成长

QCon[上海站] 全球软件开发大会 2016

如何通过 APP+Wi-Fi 数据给企业做大数据精准用户画像

SPEAKER

郭炜





促进软件开发领域知识与创新的传播



关注InfoQ官方信息

及时获取QCon软件开发者 大会演讲视频信息



[北京站] 2016年12月2日-3日

咨询热线: 010-89880682



[北京站] 2017年4月16日-18日

咨询热线: 010-64738142

郭炜 易观 CTO

毕业于北京大学,在Teradata, IBM,中金负责大数据方向架构师及研发总监,后任万达电商数据部总经理,联想研究院大数据总监。

现任易观技术负责人,负责易观整体技术架构及分析产品线。在电商、移动互联网、商业地产、百货、移动通信、零售等多个业务领域大数据方面具有搭建团队、系统以及分领域的分析与算法经验。



易观是谁?



易观是中国互联网市场领先的大数据分析公司。自2000年成立以来,易观打造了以海量数字用户数据及专业大数据算法模型为核心的大数据与分析师服务生态体系,并致力于帮助所有拥有互联网产品及服务的企业,洞察自身的产品和用户,对标竞争和市场,并通过对数字用户资产的持续运营,实现增收,节支,提效和避险。



目录



- 1 用户画像的定义
- 2 PC时代用户画像的特点
- 3 移动互联网时代用户画像
- 4 移动互联网时代画像算法特点举例

为什么使用用户画像









生活数据化

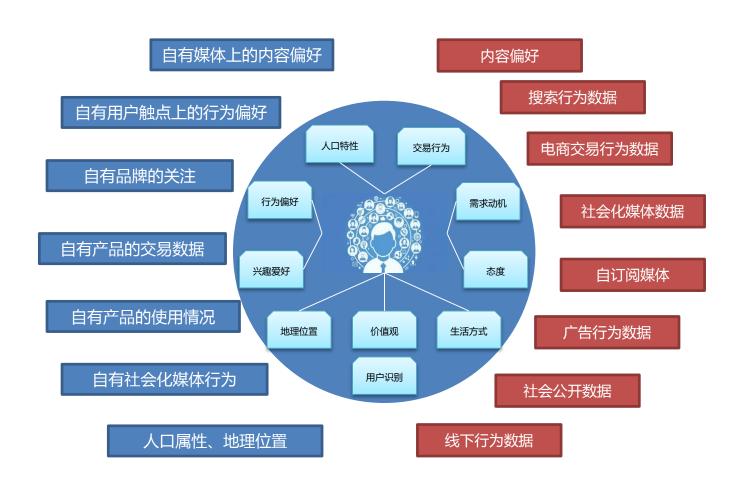
沟通多样化

信息采集智能化

通过对用户的准确画像,不仅可以为产品设计与完善提供支持,也服务于营销及用户管理

用户画像是用户属性数字化的描述





目录

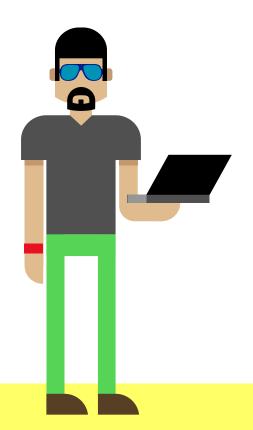


- 1 用户画像的定义
- 2 PC时代用户画像的特点
- 3 移动互联网时代用户画像
- 4 移动互联网时代画像算法特点举例

PC互联网用户画像



实时分析驱动用户资产成长



WHO?

描述性数据

用户标识:Cookie、IMEI、 DeviceID、强关联标识

人口属性:用户自描述

WHEN? WHAT?HOW?

行为数据

- 内容偏好
- 交易数据
- 行为偏好 品牌使用
- 品牌关注
- 社交行为

WHY?

意识形态标签

- 生活方式 •
- 观点态度

- 价值观
- 需求动机

PC用户画像的用途



实时分析驱动用户资产成长

用户行为数据

· 用户浏览行为 产品类别 产品喜好 产品驱动 使用习惯

用户心理数据

• 产品使用行为动机 信息/终端对决策的影响 对产品/服务/事件的评价 过程体验 满足感-愉悦度

用户标签数据

- 人口统计学标签
- 社会属性标签
- 品牌拥有标签
- 媒体接触标签
- 产品偏好标签
- 产品厌恶标签
- 信息偏好标签
- 生活习惯标签

用户画像数据

- 基本维度
- 兴趣维度
- 产品使用维度
- 生活习惯维度
- 品牌接触维度

用户分析模型

- 竞争分析模型
- 用户分类模型
- 用户价值分析模

型

用户行为预测模

꿱

- 用户评价模型
- 用户体验模型

PC用户画像的举例——自然语言



实时分析驱动用户资产成长

用户感兴趣的文字→用户特征

- 通过词频+语意刻画提炼用户特 征词
- 通过分类、聚类、机器学习将特 征词提炼为用户标签或用户基本 属性
- 常见于门户网站、论坛、搜索引



PC用户画像的举例——商品互动

商品→商品特征→用户特征

- 提炼商品特征,商品单价
- 主成分分析, 汇聚到客户
- 通过分类、聚类、机器学习将商品特征提炼为用户标签或用户基本属性
- 常见于交易类网站



〈 我的证	T单▼	
30,431.	1+.	
全部 待付款 待发	货 待收货 待评	价
七天退换		
婚庆用品婚房婚礼 真玫瑰绢布花瓣10 颜色分类:紫渐变 七天退换		.20 x1
婚庆用品婚房婚礼· 真玫瑰绢布花瓣10 颜色分类:玫红渐变 七天退换	00片床撒花	.20 x1
婚庆用品婚房婚礼· 真玫瑰绢布花瓣10 颜色分类:浅粉 七天退换		.20 x1
婚庆用品婚房婚礼· 真玫瑰绢布花瓣10 颜色分类:香槟 七天退换		.20 x1
凡享结婚用品 婚庆 色牙刷 新人必备情 颜色分类:红色牙刷 七天退换	信软毛牙刷	.00 x1

目录



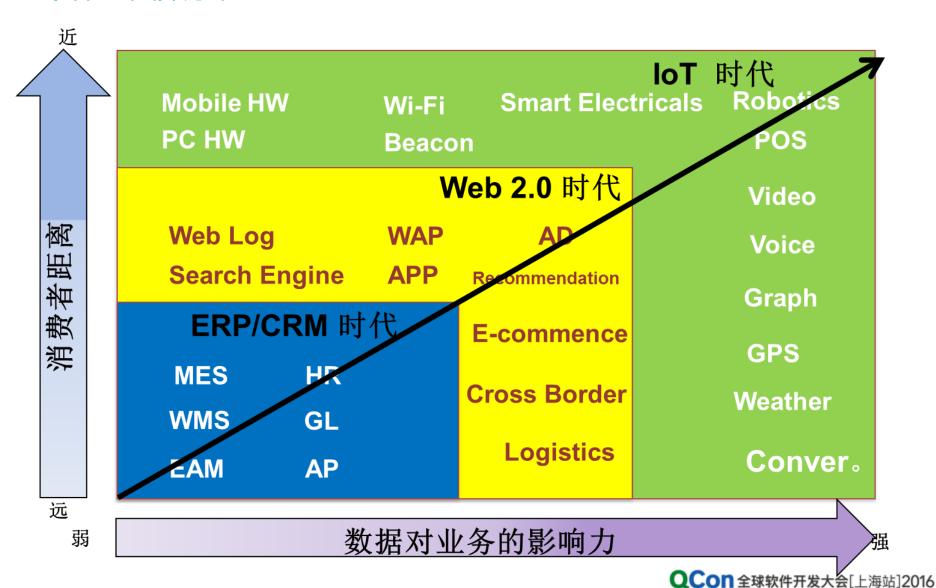
- 1 用户画像的定义
- 2 PC时代用户画像的特点
- 3 移动互联网时代用户画像
- 4 移动互联网时代画像算法特点举例



移动互联网究竟改变了什么?

| 设备的可移动性扩大了人类的感知器官,同时也将 | 人个体全面数字化





用户生活形态全数据描述



阅读偏好	娱乐偏好		
视频偏好	媒体接触 习惯		
餐饮习惯	出行习惯		
空间+时间			

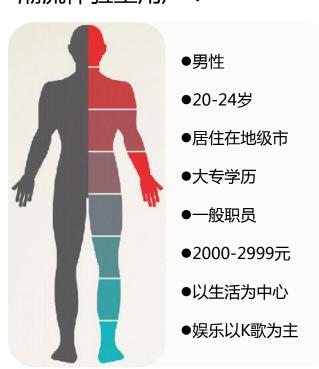
因子	因子描述	意识类型
F1	喜欢网购,喜欢追求时尚	谨慎消费意识
	喜欢外出吃饭,喜欢美食,对饮食有讲究	观点分享意识
	喜欢网购,对品牌有好感	网购意识
	有品牌意识,追求质量,但是消费能力不强,对生活充满热爱	追赶潮流意识
	购物达人,喜欢时尚穿着,喜欢分享	观点分享意识
	购物达人,年轻时尚,喜欢流行饰品与穿着	观点分享意识
	热爱旅游,喜欢自由行	自由意识
	年轻态,明星崇拜,非主流的亚文化	追赶潮流意识
F2	喜欢展现自己的唱功,以歌会友	自我表现意识
	文艺范,喜欢摄影	文艺时尚意识
	喜欢下载音乐	注重体验意识
F3	关注小生活,关注亲友	家庭生活意识
F3	喜欢下载音乐,了解好友在音乐方面的喜好	娱乐社交意识
F4	港剧爱好者,励志人群	励志意识
F5	英剧爱好者	品质生活意识
F6	动漫爱好者	娱乐意识
F7	快速获取身边信息	信息获取意识
	英文学习爱好者	自我学习意识
	社交圈以校友为主	社交分享意识
	文艺范,制作自己的MV	自我实现意识
F8	关心天气情况,喜欢收看适时街景	信息获取意识
F9	流行,时尚,喜欢修图	流行时尚意识
F10	在线支付达人,善于小额理财	理财意识
	移动办公达人	自我实现意识
	外地人士, 经常乘坐火车	网购意识
F11	重度移动人群	网络依赖意识
F12	自信,美丽,自拍爱好者	完美自信意识
F13	美剧爱好者	时尚生活意识
F14	对自己的花销非常谨慎,文艺范	经济消费意识
F15	习惯通过网址导航获取信息	品牌忠诚意识
F16	喜欢志同道合的人群,希望结识新朋友,并在线下相识	个性张扬意识
F17	喜欢旅游、汽车、财经、时尚、音乐等影音	时尚潮流意识



核心用户的生活形态分析,了解用户偏好



潮流体验型用户:

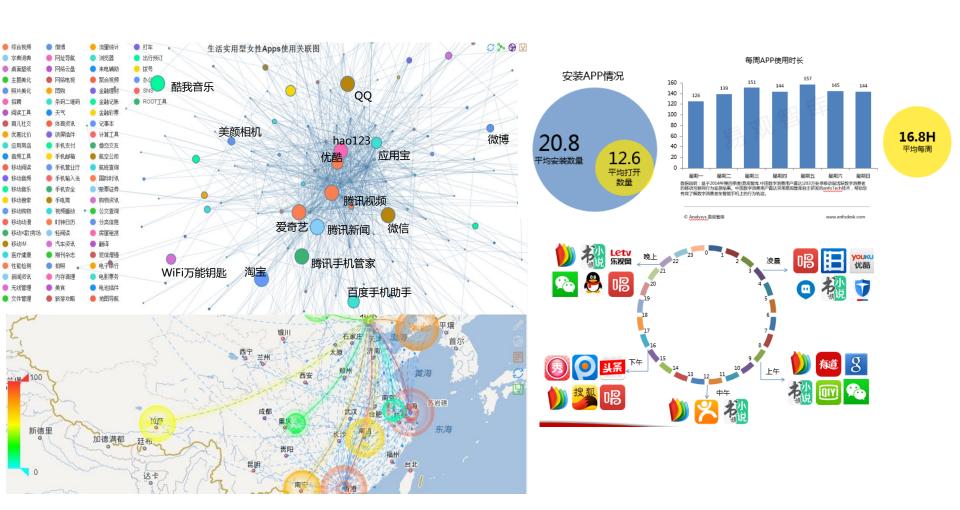


他是**文艺范**,**明星**崇拜,喜欢非主流的亚文化,喜欢歌唱,喜欢摄影。

热衷用相机记录生活中的点点滴滴。每个人会看到不尽相同的世界,而通过镜头,让摄影成为了最直观生动了解和分享世界的方式。同时,他还喜欢在移动互联的世界中展现自己的唱功,以歌会友。

细分核心用户的行为分析,制定运营与营销方案





目录



- 1 用户画像的定义
- 2 PC时代用户画像的特点
- 3 移动互联网时代用户画像
- 4 移动互联网时代画像算法特点举例

在移动化联网下用户画像成为虚拟的"你"







Wi-Fi基本原理

三点定位原理

- 指纹感知,多用于APP厂商
- 服务器端计算,多用于Wi-Fi厂商





0.35



0.055



|举例,室内大型商厦三点定位+单点感知模型

Onalysys 易观





Onalysys 易观



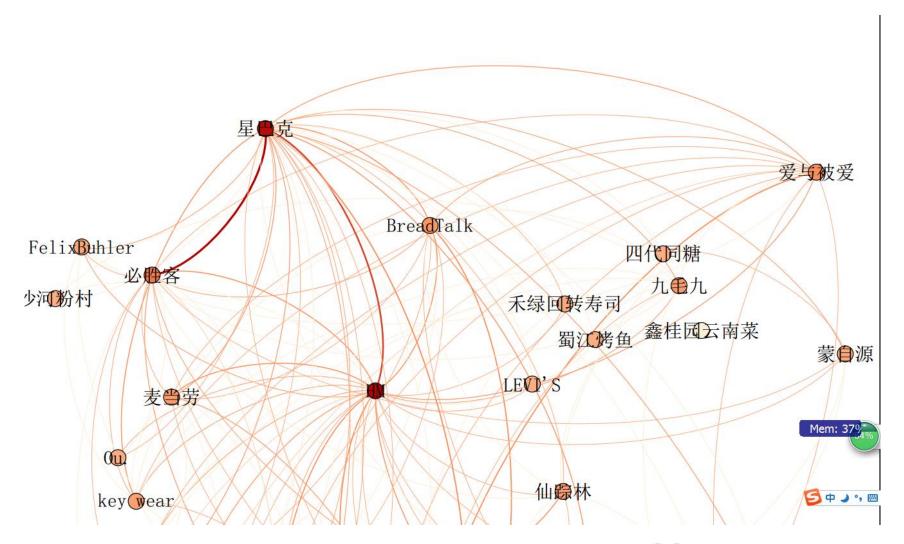


Onalysys 易观

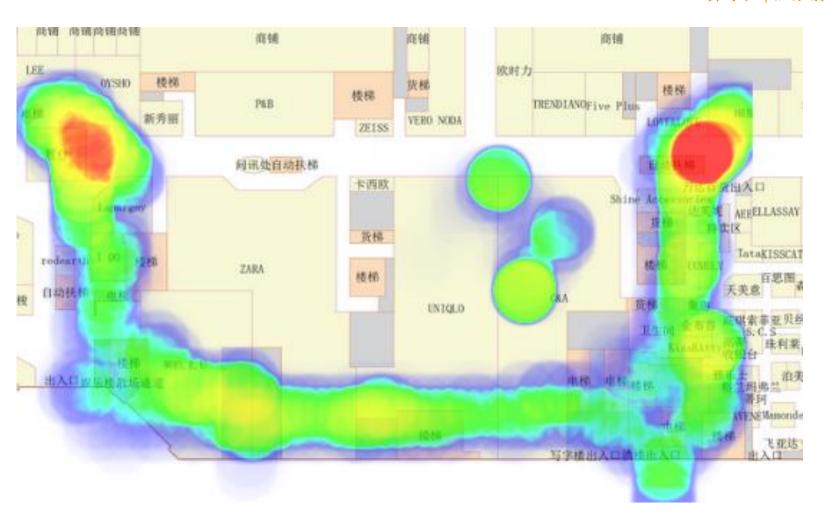








Onalysys 易观



延伸室内场景

Onalysys 易观

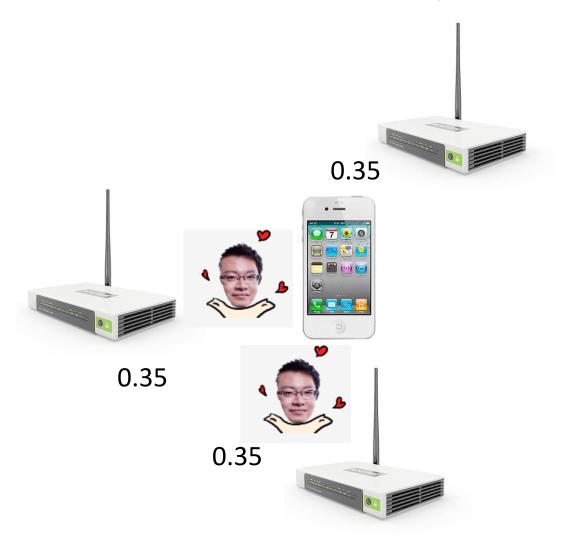
- 根据室内感知的Wi-Fi信 号判断所处位置
- 智能家居自适应
- APP智能判别场景



Wi-Fi特征算法的问题

- Onalysys 易观
 - 实时分析驱动用户资产成长

- 噪音与漂移
- 房间位置的变化



移动终端比PC多了什么?



排名	全天 (0:00~24:00)	凌晨 (0:00~06:00)	上午 (06:00~12:00)	中午 (12:00~14:00)	下午 (14:00~18:00)	晚上 (18:00~24:00)
1	❷ 微信	҈ 微信	@ QQ	❷ 微信	微信	微信
2	优酷视频	8 QQ	◇ 微信	@ QQ	🚇 QQ	② 美图秀秀
3	■ 爱奇艺视频	优酷视频	₩ 今日头条	/	(尤語视频	支 支付宝
4	@ QQ	美國秀秀	② 美图秀秀	● 美图秀秀	美图秀秀	🔀 大众点评
5	📸 手机百度	淘淘	₩ 手机百度	■ 爱奇艺视频	☆ 百度	₩ 今日头条
6	淘淘宝	♂ 微博	淘淘宝	淘淘宝	淘淘宝	📸 手机百度
7	6 微博	■ 爰奇艺视频	♂ 微博	♂ 微博	6 微博	优酷视频
8	👱 QQ空间	📸 手机百度	■ 爱奇艺视频	🚖 QQ空间	👱 QQ空间	🚖 QQ空间
9		Ż QQ空间	👱 QQ空间	原 喜马拉雅FM	↓ UC浏览器	
10	Letv 乐视视频	↓ UC浏览器	¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿	¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿ ¿	○ 酷狗音乐	🧹 高德地图

用户的两种基本计算方法

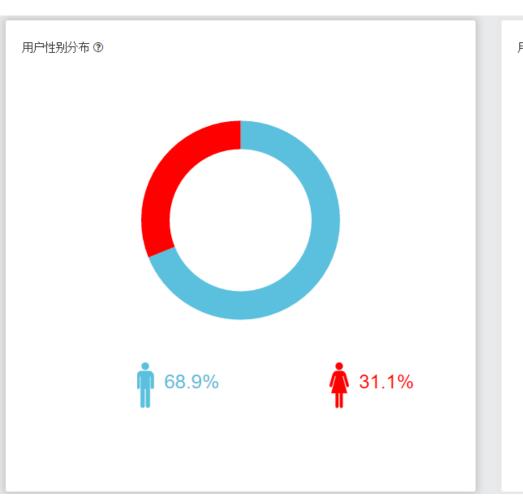


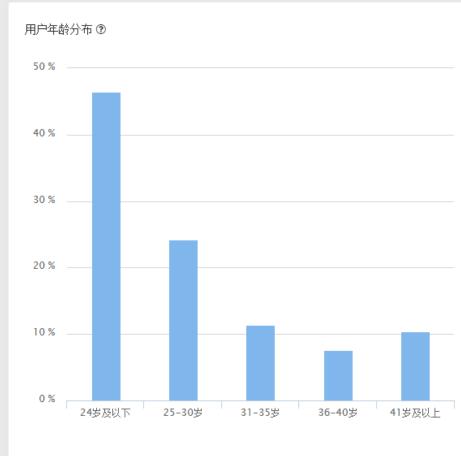
TGI:即Target Group Index(目标群体指数),可反映目标群体在特定研究范围(如地理区域、人口统计领域、媒体受众、产品消费者)内的强势或弱势。

RFM变种:最近一次消费(Recency)→最近一次发生、消费频率(Frequency)→发生频次、消费金额(Monetary)→发生的权重(时长、金额、量化)

利用分类算法刻画用户属性







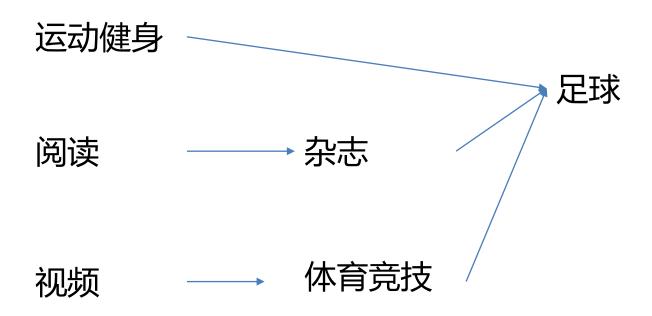
将APP标签赋予用户标签



标签如何设置

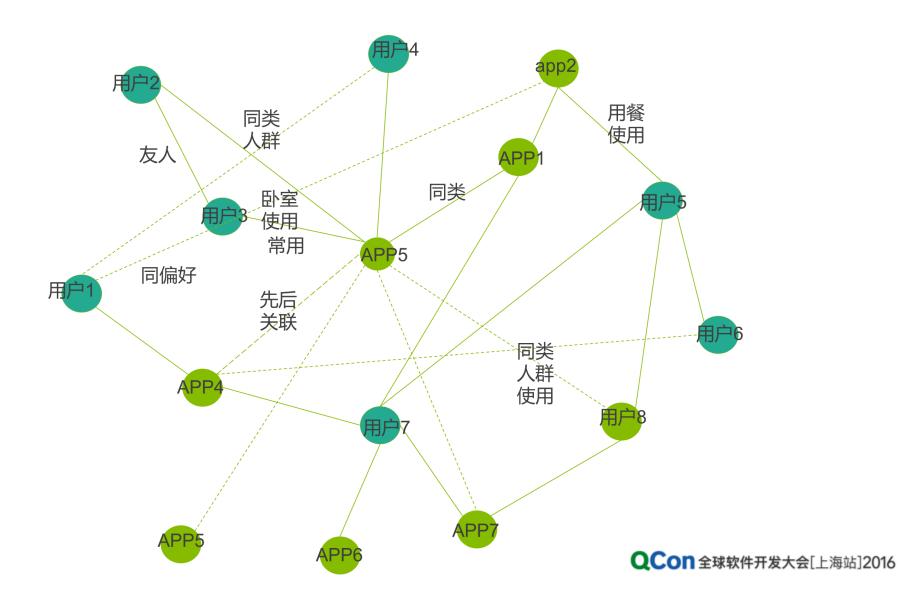


树形标签不适合用户群使用,正确的标签,应该是有向图 状标签,多种分类多种链接



应用与用户的图谱





泛地域标签



智能硬件时代来临,线下线上的标签融合,给用户画像带来更高的准确度与挑战

居住地与工作地:地域本身属性,时间属性,频次属性

线下活动属性:个性化偏好





International Software Development Conference