

QQ大数据及其应用介绍

SNG 数据中心

QQ大数据

规模数据

QQ月活跃用户数超过	8.4亿
QQ最高同时在线超过	2.0亿
QQ空间月活跃用户数超过	6.5亿
QQ好友关系链对超过	900亿



活跃数据

一天QQ总发消息超过	100亿
一天发空间说说、个签超过	1亿



大数据应用一：好友/群推荐

数据：社交+标签+LBS

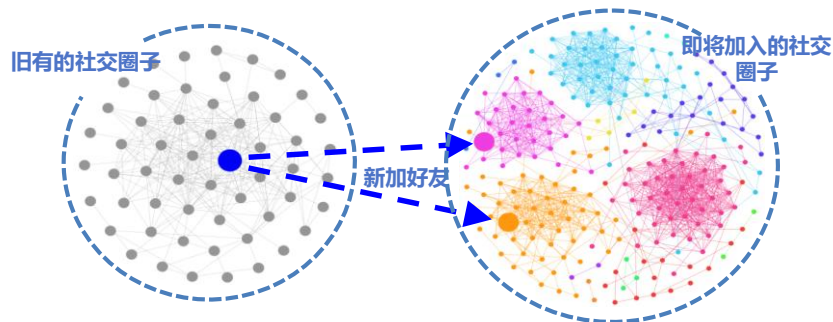
算法：识别并推荐适合的人与群

应用：在用户主动加好友/群时及时推荐

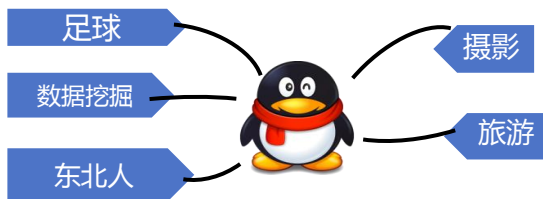
数据：SNS+标签+LBS

- 通过社交行为、用户标签以及活动地点感知用户状态以及社交需求。

基于社交行为



基于用户标签



基于活动地点



感知
用户

1

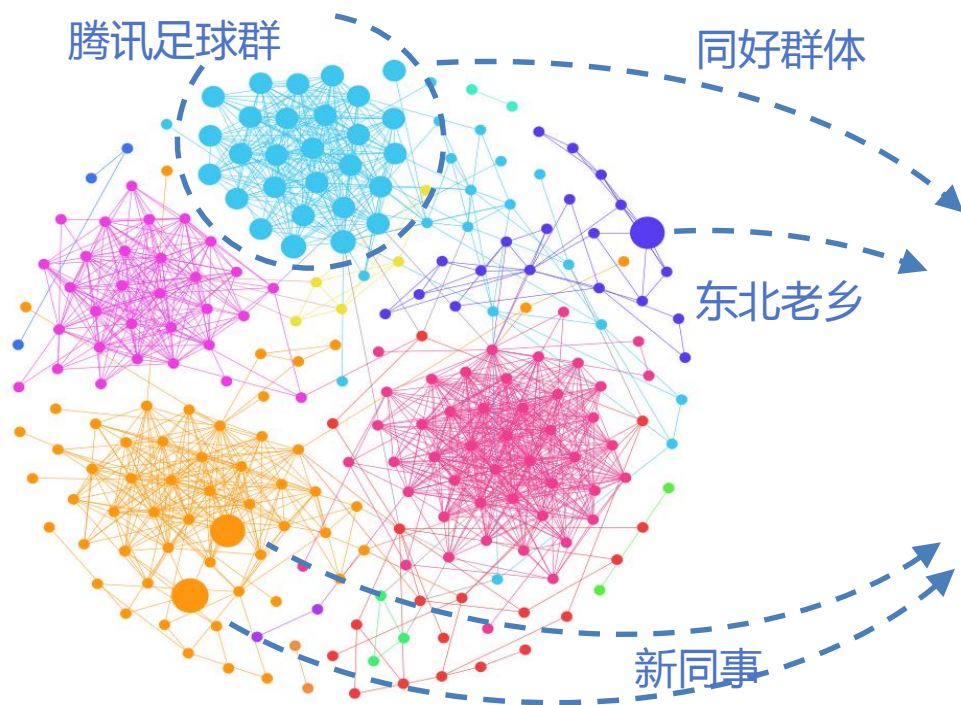
用户状态

2

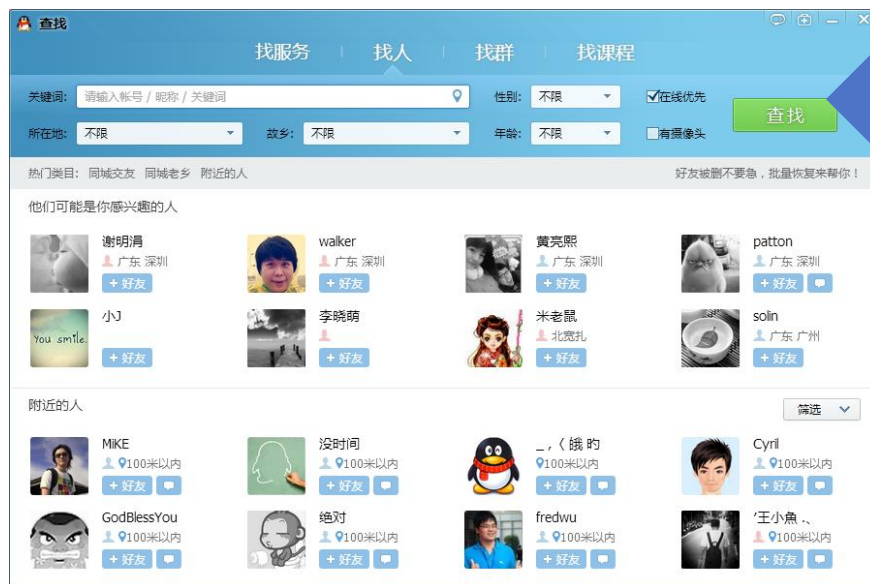
社交需求

算法：识别并推荐适合的人与群

- 从新的社交圈子中挑选与用户匹配的好友/群进行推荐，帮助其快速融入。



应用：用户主动加好友/群时及时推荐



好友推荐



群组推荐

大数据应用二：广点通

广点通助力长尾广告

广点通的大数据解决方案

案例应用：红米手机促销

广点通助力长尾广告

- 基于腾讯大社交平台独特基因，以海量用户为基础，以大数据洞察为核心，以智能定向推广为导向，广点通提供一站式网络推广营销平台。

精准聚焦

细粒度挖掘用户属性，精选出适合广告主推广的目标用户群



高效传播

智能广告竞价机制，高效提升广告效益。实现通过实际带来推广效果付费



优质资源

亿万活跃用户潜在投放量，提供多方位广告展现平台



1.上传广告
一键上传



2.精准定位
定位目标群



3.投放广告
社交传播



4.查看报表
效果实时查询

广点通的大数据解决方案

- 广点通应用大数据去了解每一位用户的喜好和需求，做到满足各个业务的个性化精准推荐需求。

数据：海量

基础属性数据

人口属性

- 年龄/性别
- 地域/学历
- 职业/收入
- 上网场景

兴趣属性(400余种类目)

- 女装
- 女包
-

行为属性数据

自定义标签(100万个)

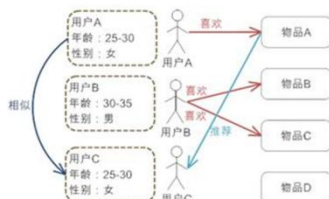
- 指环王
- 索尼
- 安踏
- 莫文蔚
-

生活状态

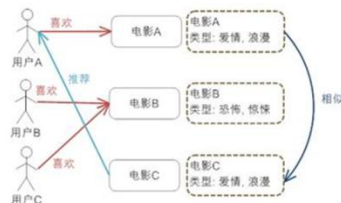
- 职业状态
- 婚恋状态
-

算法：精准

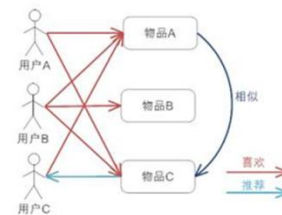
人口统计推荐



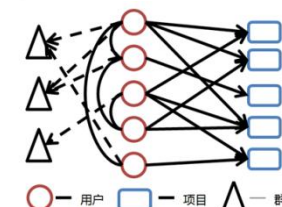
内容推荐



协同过滤



增强CF方法



应用：实时



案例：红米手机促销

推广需求

对QQ空间用户进行红米手机促销

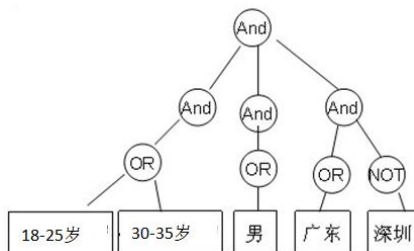
年轻上网用户，同时小米用户在空间中的使用占比较高

小米希望通过与QQ空间的合作促进红米手机首发销售，并扩大红米手机在目标客户群中的品牌影响力。

执行方案

广点通精准用户定位

- QQ空间活跃用户
- Android手机用户
- 18-35岁



推广效果

- **10万** 90秒红米全部售罄10W台；
- **15万** 首发第1秒超过15万用户抢购；
- **745万** 在空间预约创下史上最疯狂首发。

大数据应用三：用户行为研究

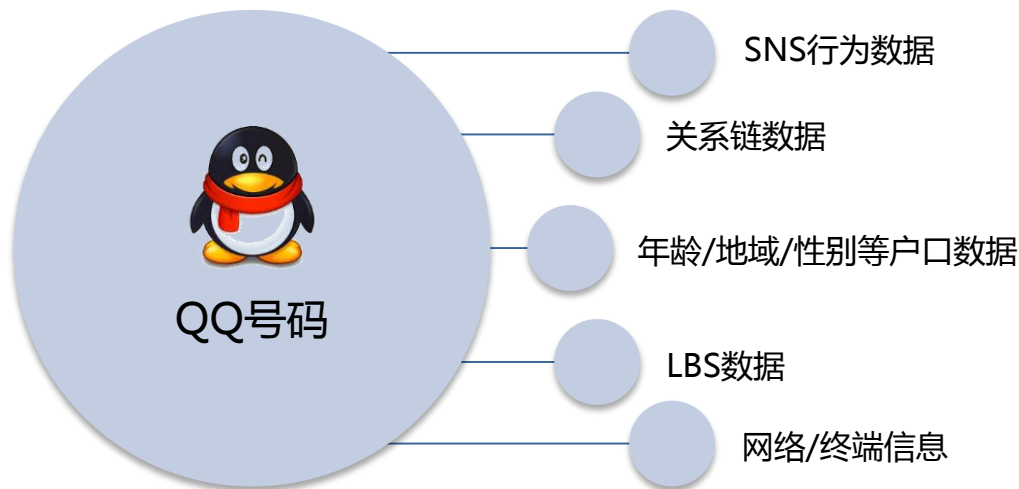
QQ用户行为的研究思路

案例：春节人口迁徙报告

QQ用户行为的研究思路

- 发挥QQ基数大、变量多的大数据优势，与社会热点话题结合，通过大数据计算技术与呈展技术开展全网用户行为研究，传播QQ大数据的社会价值。

数据记录与收集



行为分析

- 社会关系连接
- 用户社交行为
- 用户迁徙行为
- 用户上网习惯
- 影响力传播
- 产品易用性

案例：春节人口迁徙报告

- 基于QQ用户位置变化轨迹，建立模型测算迁徙规模，分析人口迁徙方向及其影响因素。该报告两次登录央视，并引来知乎日报、凤凰新闻等媒体主动扩散报导。

数据准备



- 全样本**：8亿以上QQ使用记录及LBS信息；
- 唯一ID**：QQ用户身份ID具有用户跟踪能力；
- 多样性**：支持多维分析，包括QQ聊天对象属性、位置属性、行为时间等交叉分析。

建模分析

迁徙规模测算模型



人口迁徙因素分析



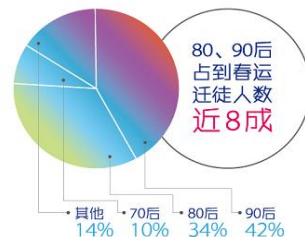
报告输出

手机QQ揭秘：春节都去哪儿了！

2014.1.16 - 2.16

16.52亿人次
加入春运的人潮人海中

人次统计方法：小明从A地到C地，如果在中间B地停留超过1天，则算迁徙2次。



大城市之所以大，之所以吸引人，是因为他们接纳了更多的外地人。



QQ数据大揭秘——之都市围城

逃离北上广，留守北上广，重回北上广……北上广深就像一座围城，城外的人想进去，城里的人想出来。腾讯QQ首次用大数据解读“逃离北上广”，带你一窥当代中国的“都市围城”。

有多少人离开北上广深？

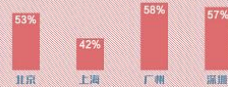


春节了，终于有假期可以逃出北上广了！

六观“人往何处”和“钱去哪儿了”（即将上演）

北上广深一半用户离开
和小A一样，春节（1月16日-2月1日）全国累计有4907万人离开北上广深，占当地总用户数的51%。

北上广深各自有多少人离开



大数据应用四：客户生命周期管理(CLM)

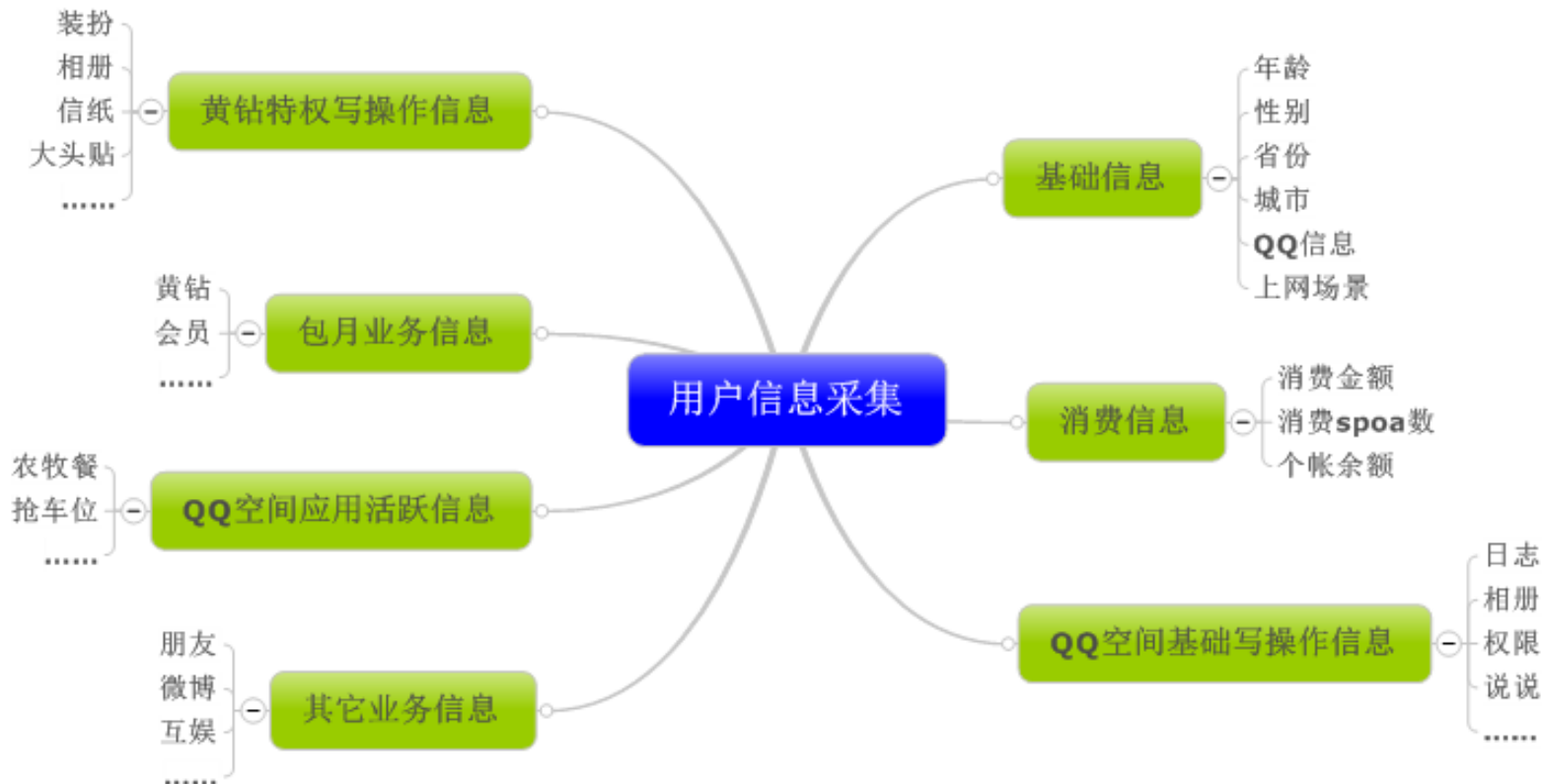
数据：多指标、多维度的海量数据采集

模型：在用户各生命周期进行预测与预警

应用：搭建CLM自动化系统








数据：多指标、多维度的海量数据采集

- 每周计算3.5亿的QQ用户，500多个指标，全面覆盖用户的腾讯平台数据



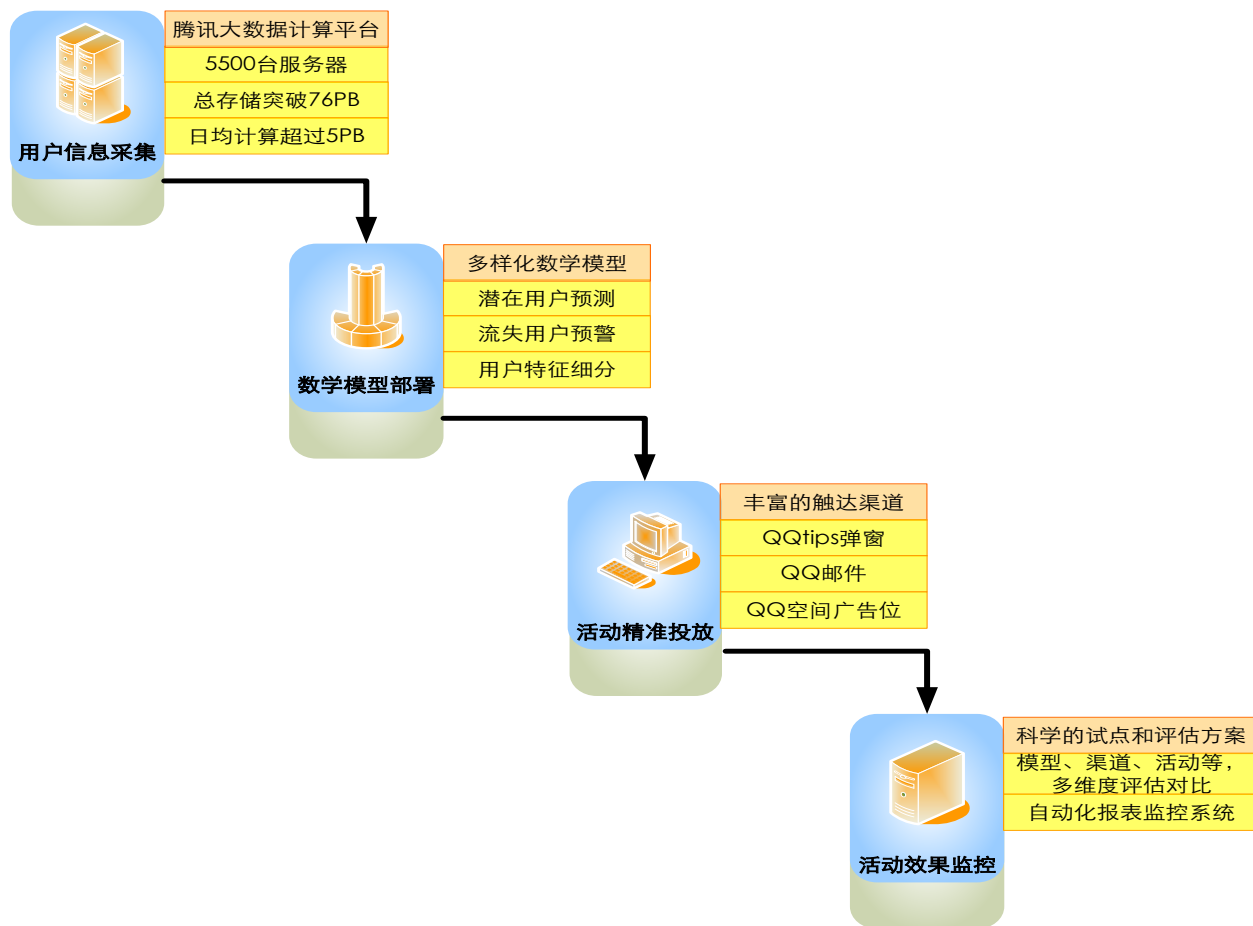
模型：在用户各生命周期进行预测与预警

- 建立预测、预警以及用户群特征分析模型，涵盖用户各个生命周期阶段

	装扮爱好者	钻族收集群	APP付费	作家	相册狂热者	QQ整体活跃	沉默用户
心理需求							
	喜欢漂亮的东西，尤其是装扮类的，例如花藤；兴趣广泛，比较博爱；活跃于多种活动；喜欢购买新颖和有吸引力的东西。年龄比较年轻，朋友数量和大多数用户相当	拥有3个钻以上；喜欢在第三方游戏或者应用上花钱，但对活动有特别偏好：仅在装扮和音乐盒等少数产品上活跃。该人群男性比较多	他们喜欢互动娱乐，愿意为此花钱；最近活跃次数是越来越高的，但是不大使用Qzone基础功能，对他们而言，有趣的游戏是关键。他们往往比一般用户年龄大，朋友较少，男性居多	喜欢在说说日志上表达观点，也喜欢在自己和朋友页面上分享、发表，和在别人的页面上评论；时刻都要跟朋友连在一起，手机功能活跃；不太关心相册。这个人群有很多朋友	非常喜爱相册，尤其是上传照片；对游戏/应用不感兴趣，对装扮也不感兴趣。他们有时候自己写写日志，他们喜欢作为客户看别人的Qzone。这群人一般Q龄较长，朋友较多，相对年龄较大	非常活跃的用户，而且变的越来越活跃。几乎活跃于所有基础功能；也比较喜欢买红钻，对绿钻，蓝钻和黑钻等不感兴趣。这个客户群比较年轻，有很多QQ好友	总体活跃度低。基础活跃最近8周呈下行趋势；QQ账户里还有余额。这部分人年龄较大。
细分规模百分比	15%	2%	27%	18%	7%	5%	26%
转化率是平均数的倍数 (转化率)	4X 0.4%	7X 0.7%	3X 0.3%	2X 0.2%	3X 0.3%	8X 0.8%	1X 0.1%
价值百分比	22%	5%	28%	15%	8%	12%	10%

应用：搭建CLM自动化系统

- 实现用户信息采集、数学模型部署、活动精准投放、活动效果评估的全面自动化，为用户提供及时准确的个性化服务，同时提高生产效率



大数据应用畅想

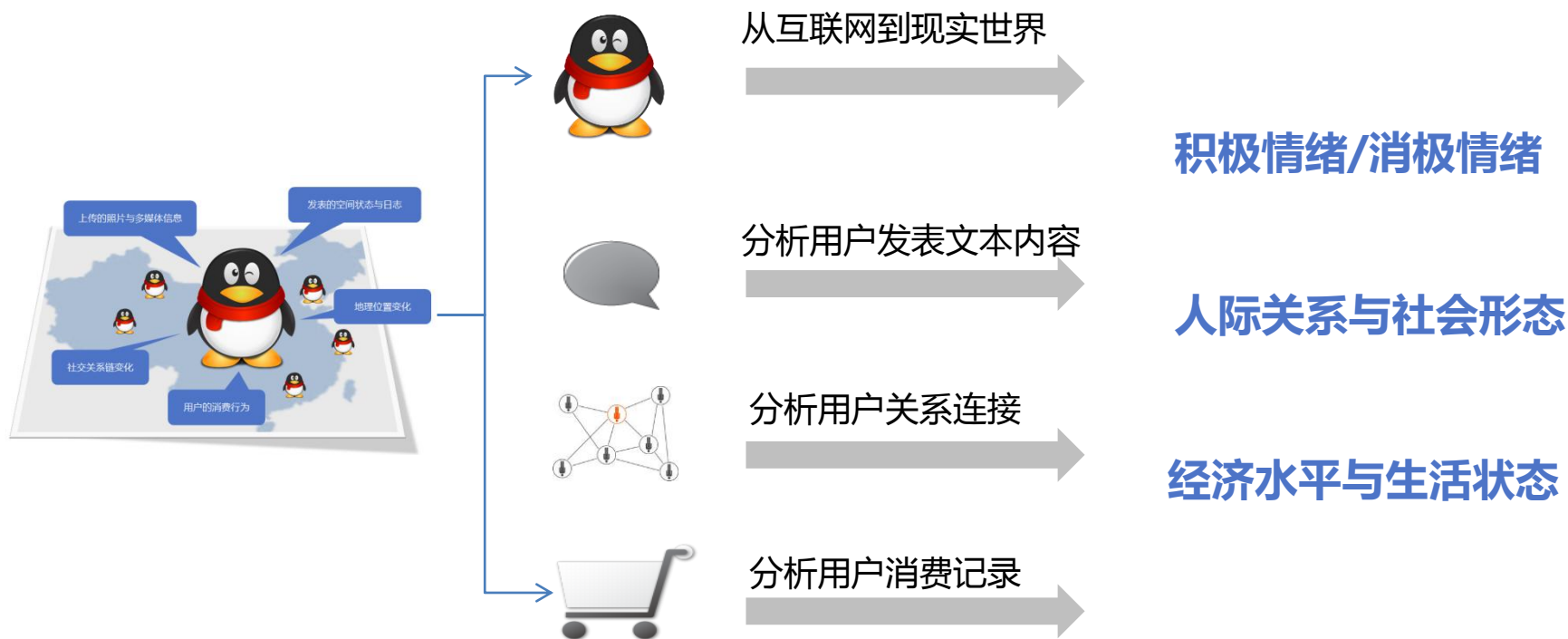
幸福指数与舆情监控

中国城镇化进程研究

个人征信

大数据应用畅想一：幸福指数

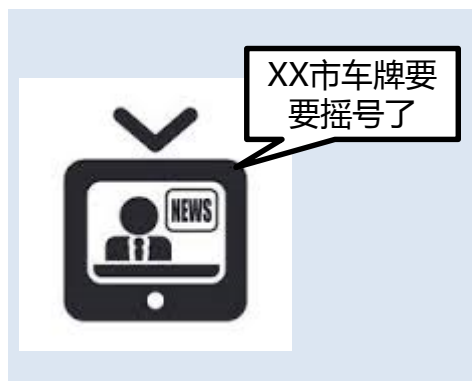
- 关注QQ用户幸福指数，做中国网民生活状态和社会运行状态的“晴雨表”。



大数据应用畅想二：舆情监控

- 通过对用户行为数据系统性的收集与分析，可以全面了解民众对民生问题的诉求，并能对突发事件进行及时监控与预警。

社会
事件



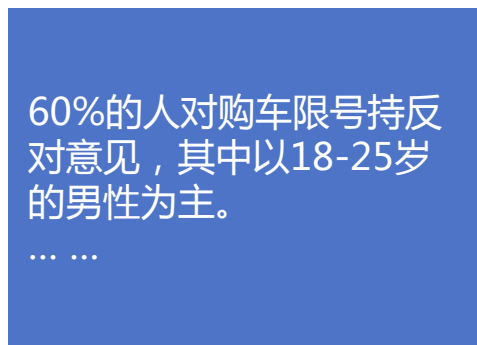
数据
采集



观点
汇总



得出
结论



大数据应用畅想三：中国城镇化进程研究

- 通过QQ用户LBS轨迹的变化跟踪计算，从人口迁移的角度看评估城镇化发展的动力——包括乡村的推力和城市的拉力。

1 通过LBS计算人口常住地



2 例行统计各地的常住人口



3 跟踪研究各时间切片变化



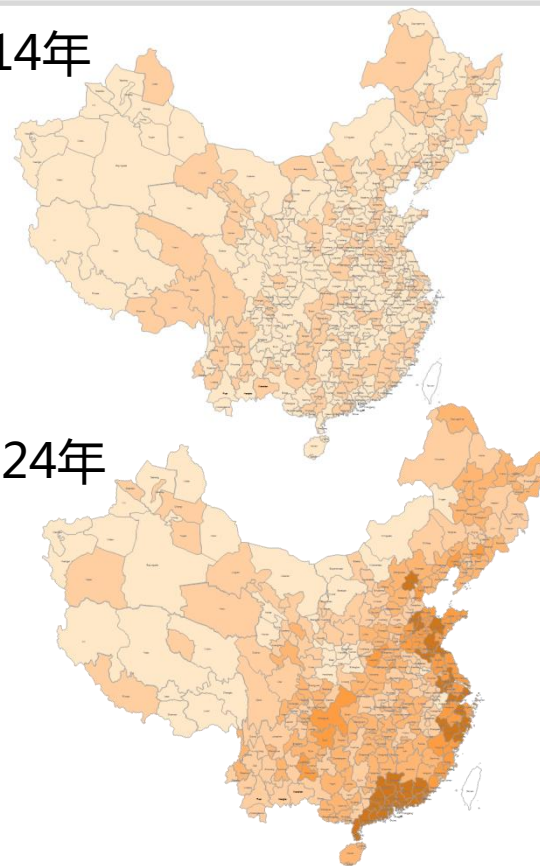
4 建立模型推算城镇化进程



2014年



2024年



大数据应用畅想四：个人征信

- 在传统个人征信的基础数据上，增加结合腾讯大数据，包含社交数据，消费数据，地理位置数据，社交评论数据等，形成个人征信综合评分系统。

传统型个人征信数据

- 个人基础信息：身份，婚姻，居住，职业
- 银行信息：各个银行的所有贷款数据、信用卡数据；
- 公用事业费用数据：社保、住房公积金、环保、养路费；
- 其他事业单位数据：工商局，交通局等。

腾讯新型个人征信数据

- 传统型个人征信基础数据
- 腾讯网上消费数据
- 腾讯社交网络数据

传统个人
征信数据



用户ID映射

个人传统信用

- 户口/社保信息
- 外部银行数据

腾讯消费



统计分析

个人网上消费力

- 腾讯网上消费力

社交网络



社团挖掘及节点影响力分析

个人社交信用及影响力

- 亲戚、同学、同事社团的信用
- 用户在社团的重要度

个人征信综合评分系统

谢谢聆听，敬请指导

By 腾讯社交网络事业群/数据中心



Thank
you