

# Geekbang>

极客邦科技

全球领先的技术人学习和交流平台

扫我，码上开启新世界



# Geekbang>

InfoQ | EGO NETWORKS | StuQ

## InfoQ

专注中高端技术人员  
的社区媒体

## EGO NETWORKS

EXTRA GEEKS' ORGANIZATION  
高端技术人员  
学习型社交网络

## StuQ

实践驱动的IT职业  
学习和服务平台

促进软件开发领域知识与创新的传播

InfoQ<sup>new</sup>

**QCon**  
全球软件开发大会

**[上海]** 2015年10月15-17日

**ArchSummit**  
全球架构师峰会

**[北京]** 2015年12月18日-19日



关注InfoQ官方微信  
及时获取ArchSummit演讲视频信息

# ArchSummit全球架构师峰会 深圳站2015

腾讯游戏大数据应用  
邓大付

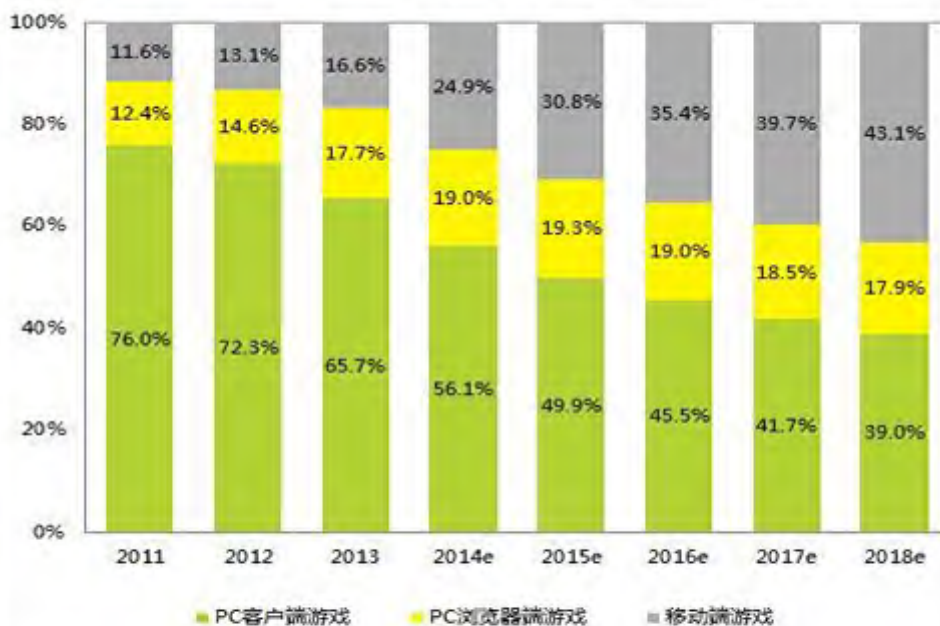
# 目录

- 1：腾讯游戏数据体系
- 2：腾讯游戏画像体系
- 3：道具推荐系统
- 4：总结

# 1: 腾讯游戏数据体系

# 网游市场现状

2011-2018年中国网络游戏市场规模结构



注释：1.中国网络游戏市场规模统计包括PC客户端游戏、PC浏览器端游戏、移动端游戏；2.网络游戏市场规模包含中国大陆地区网络游戏用户消费总金额，以及中国网络游戏企业在海外网络游戏市场获得的总营收；3.部分数据将在艾瑞2014年网络游戏相关报告中做出调整。

来源：综合企业财报及专家访谈，根据艾瑞统计模型核算。

©2015.1 iResearch Inc.

www.iresearch.com.cn

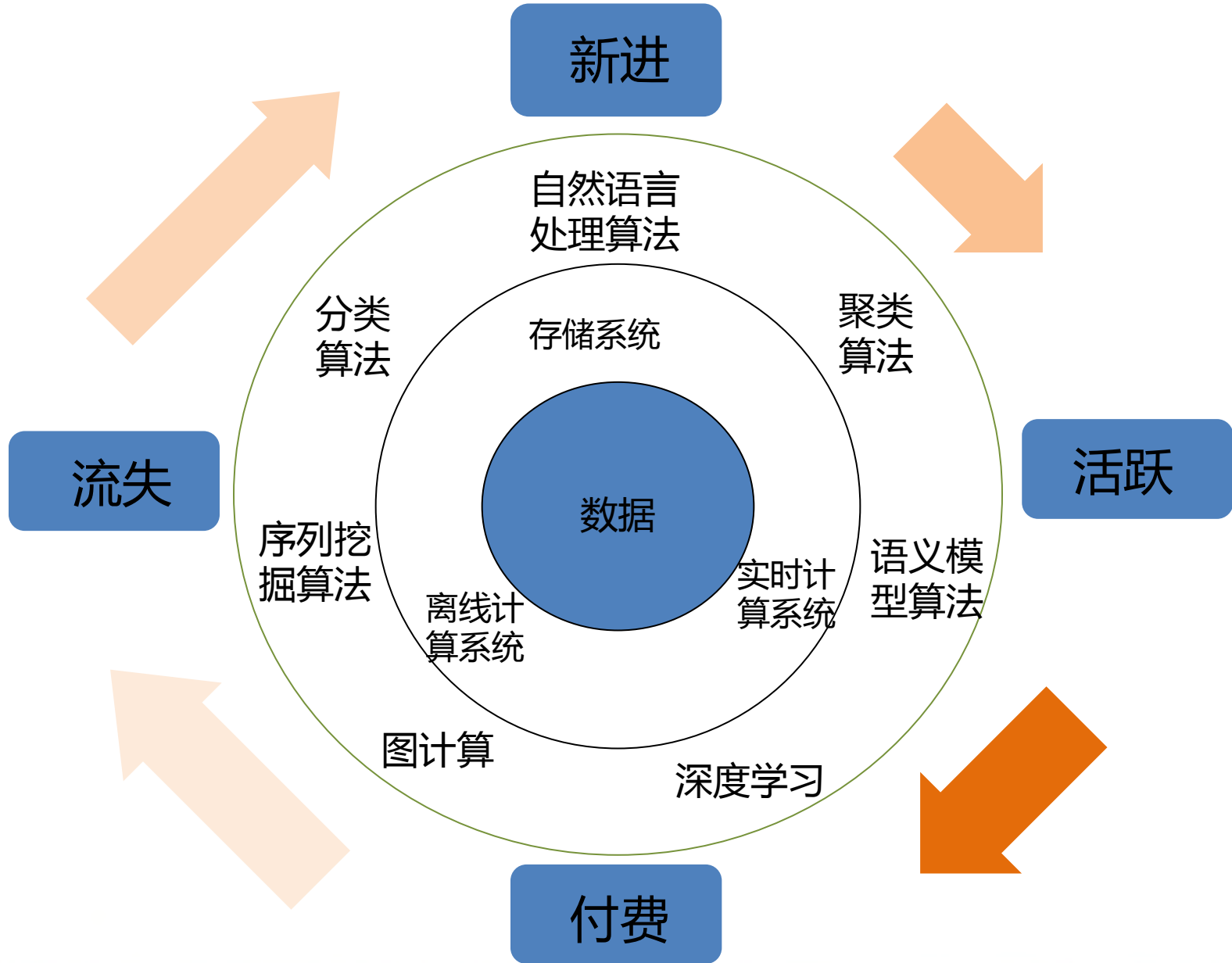
- 网游市场规模首破千亿
- 整体增幅持续放缓
- 手游占比增加，超过页游
- 如何能在逐渐饱和的市场中稳固领先地位，深挖用户价值？

# 游戏数据概况



每天入库近4000亿记录，共58T，库中总共60P+，占公司整体存储量26%

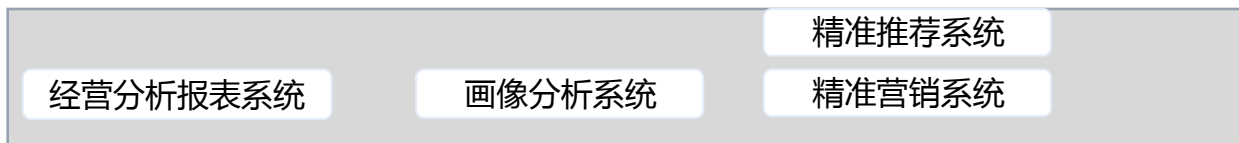
# 大数据落地应用=数据+系统+算法+应用场景



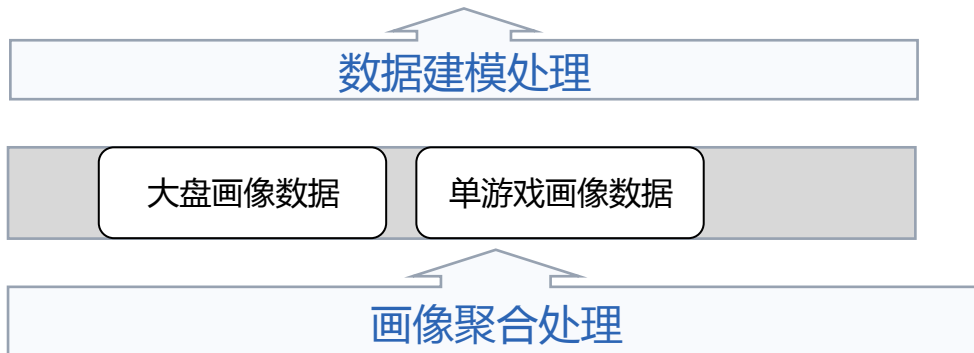


# 腾讯游戏用户数据分层体系

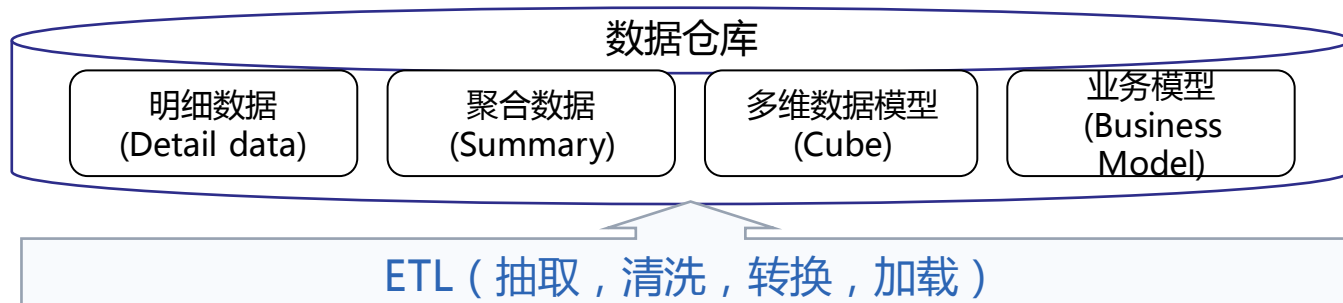
数据  
产品层



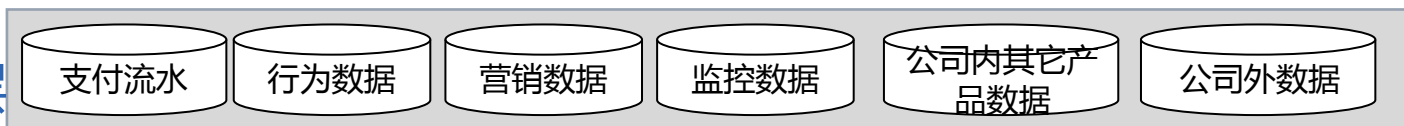
画像  
数据层



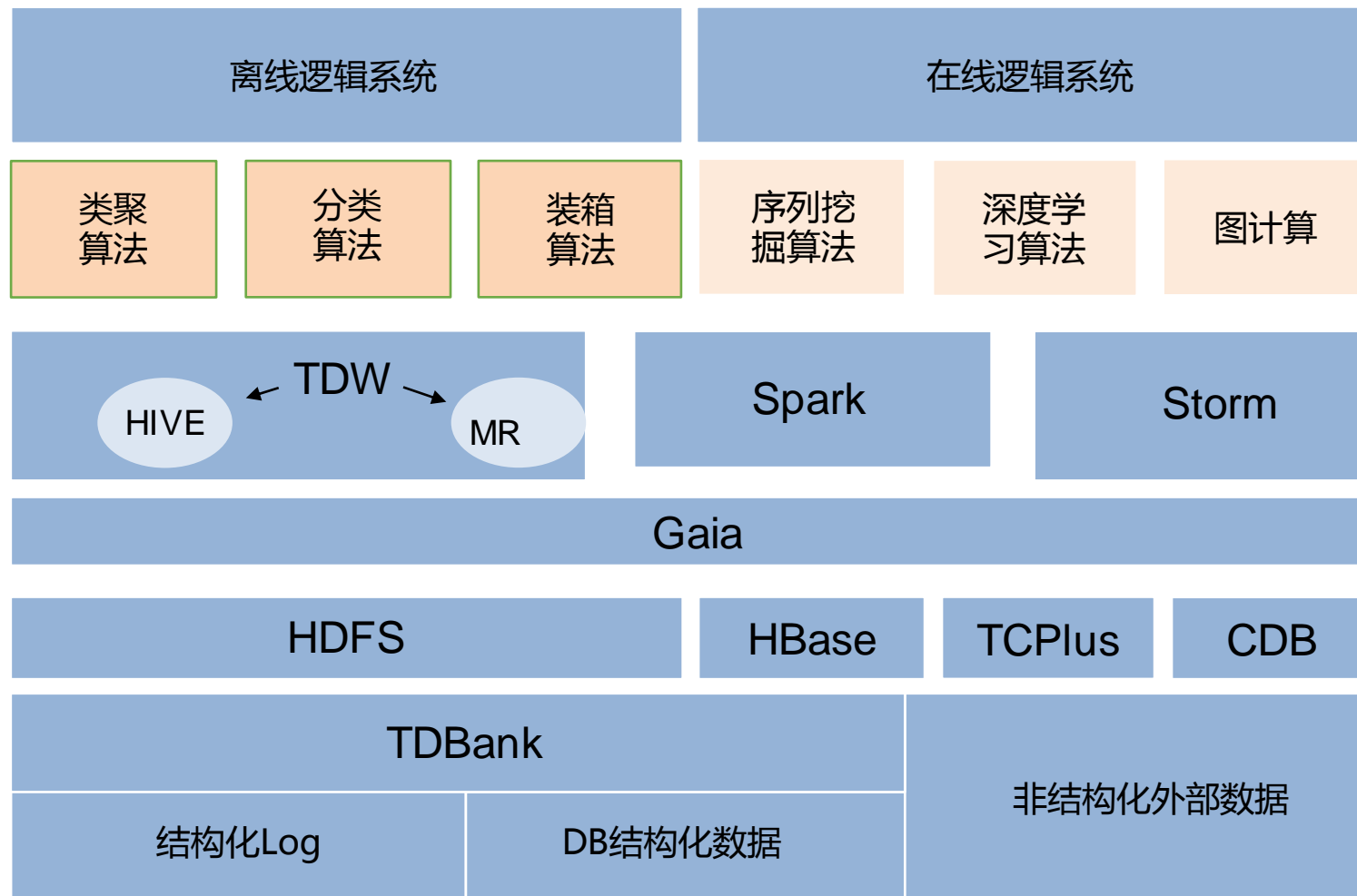
明细  
数据层



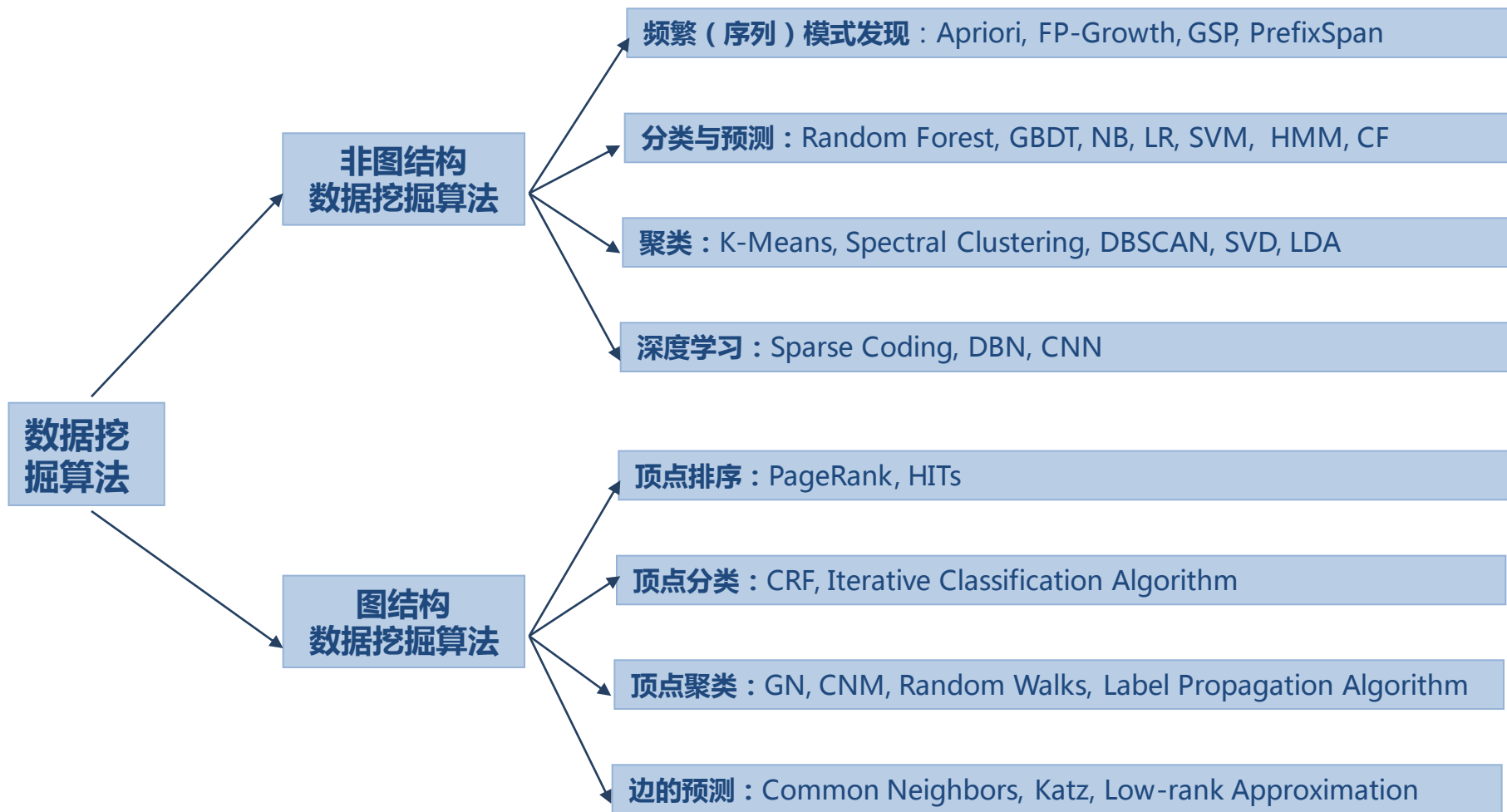
原始  
数据层



# 腾讯游戏数据处理系统架构

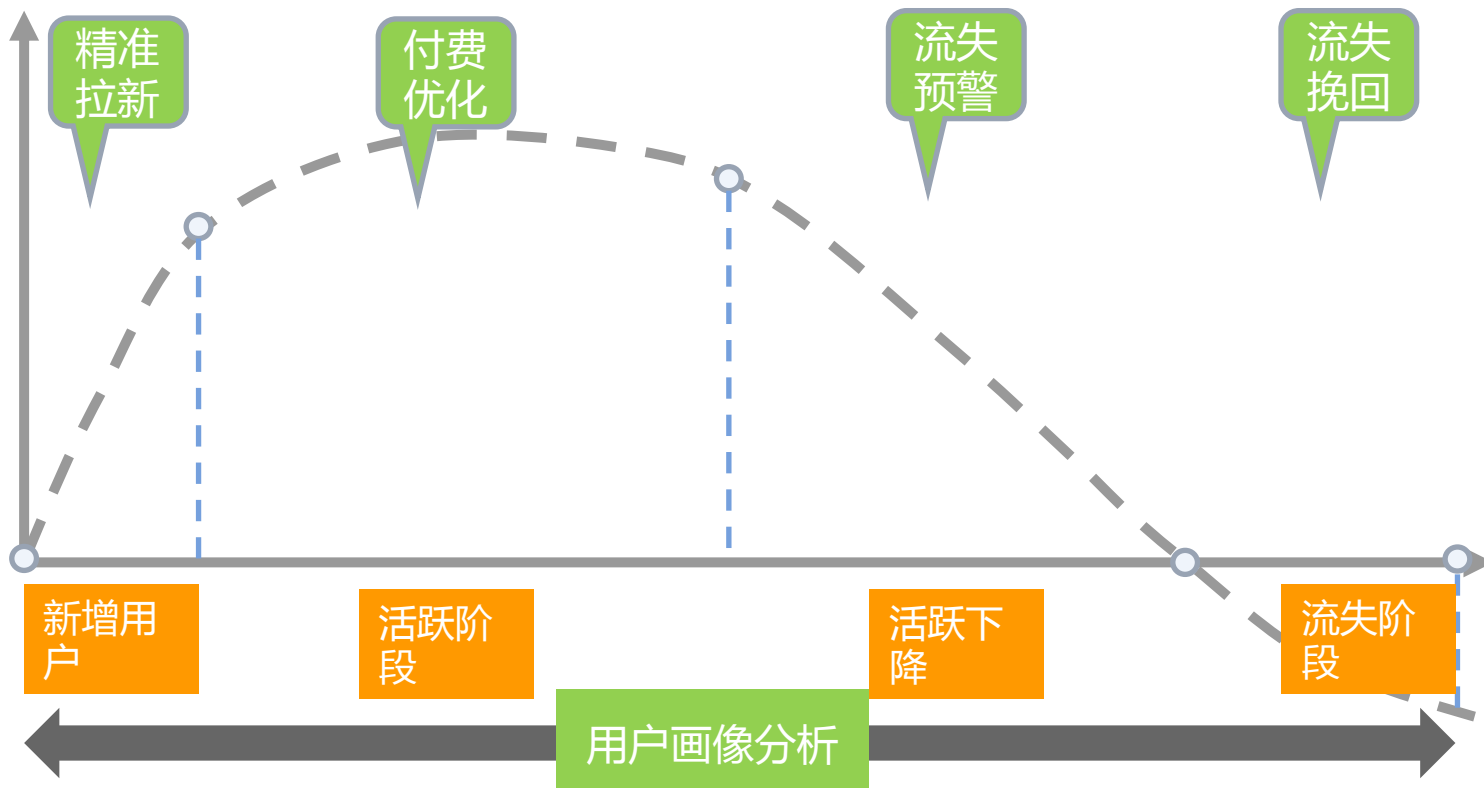


# 数据挖掘算法概览



# 游戏运营场景中数据挖掘算法的使用

结合用户生命周期模型对用户行为进行建模及预测



针对用户生命周期中的每个不同阶段，有针对性的进行建模，如对潜在新用户预测其对新游戏的兴趣高低，对活跃未付费用户预测其付费可能性等

## 2: 腾讯游戏画像体系

# 用户画像处理的问题

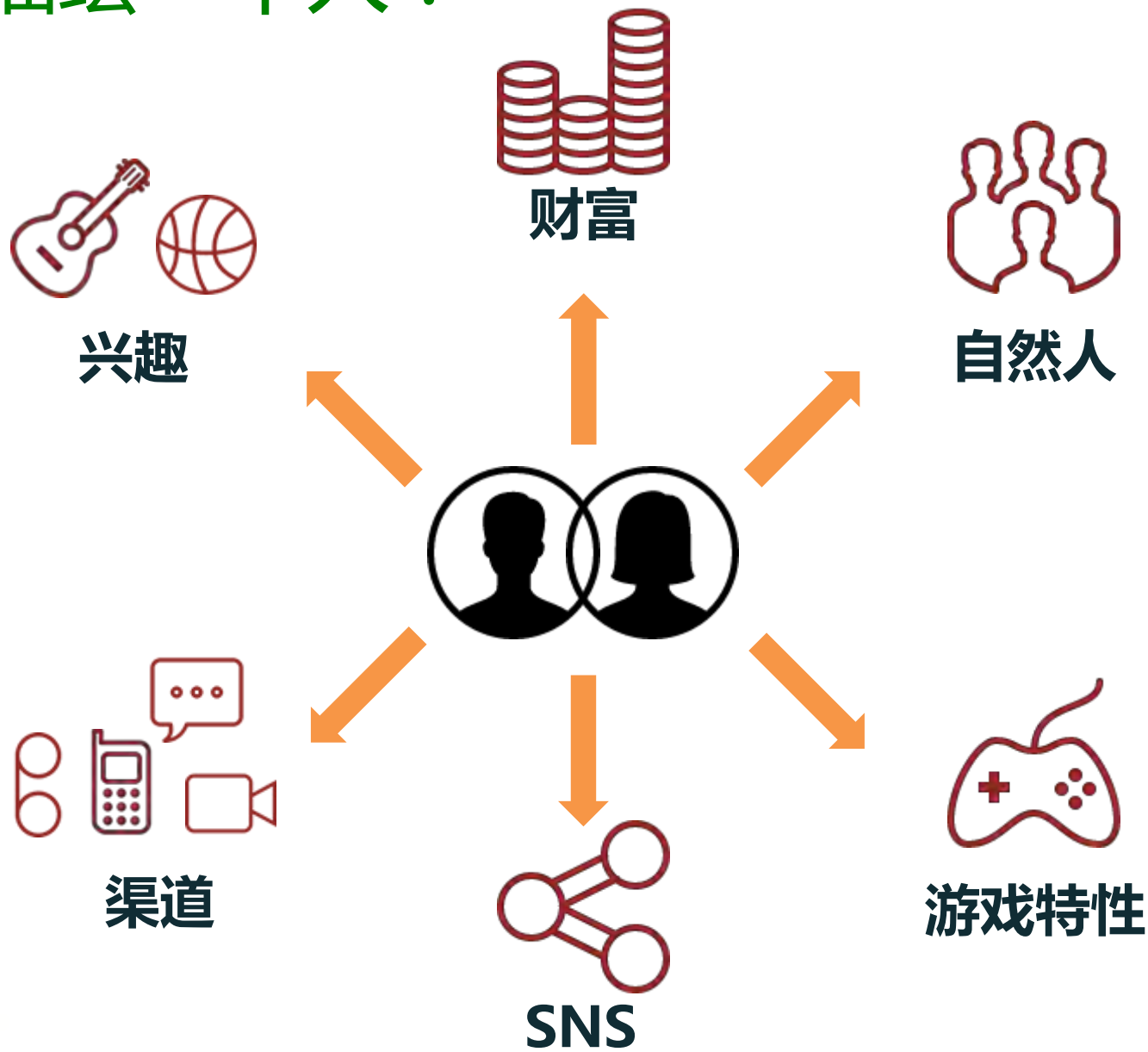
- 基础数据层

1. 整理数据标签，防止重复性工作。
2. 标签数据直接用于各种营销场景。

- 聚类分析层

- 1.
- 2.

# 如何描绘一个人？



# 画像建设的系统架构

## 应用层

画像可视化系统

权重检测系统

异常监控系统

标签系统

帐号转换系统

索引系统

## 组件层

ETL + 算法 + 数值

图计算

集合操作

图像特征

数据互导

分词

打横

日期类型

离散化

索引

帐号互转

## 底层框架

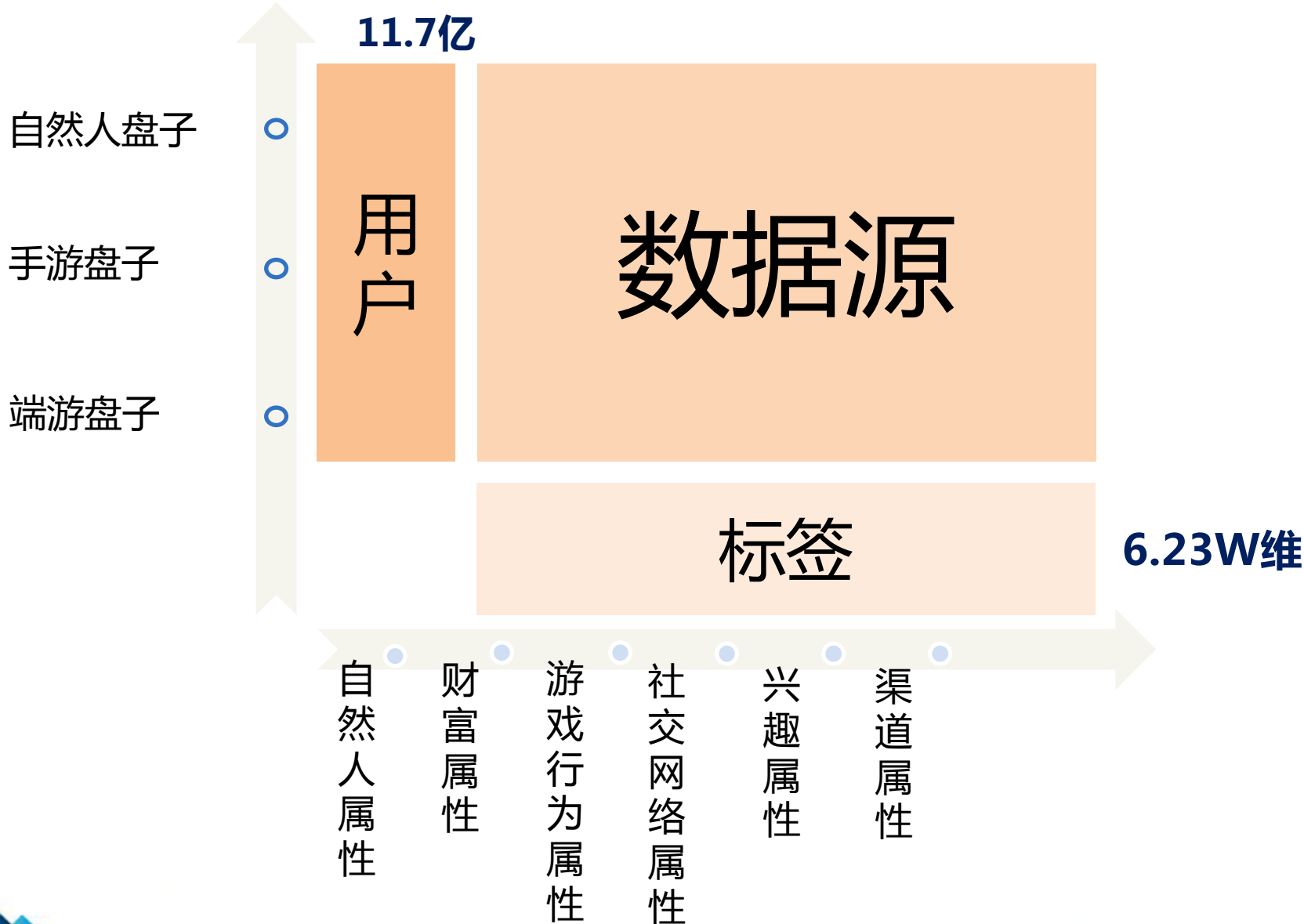
单元测试

单机调试

轻量作业流

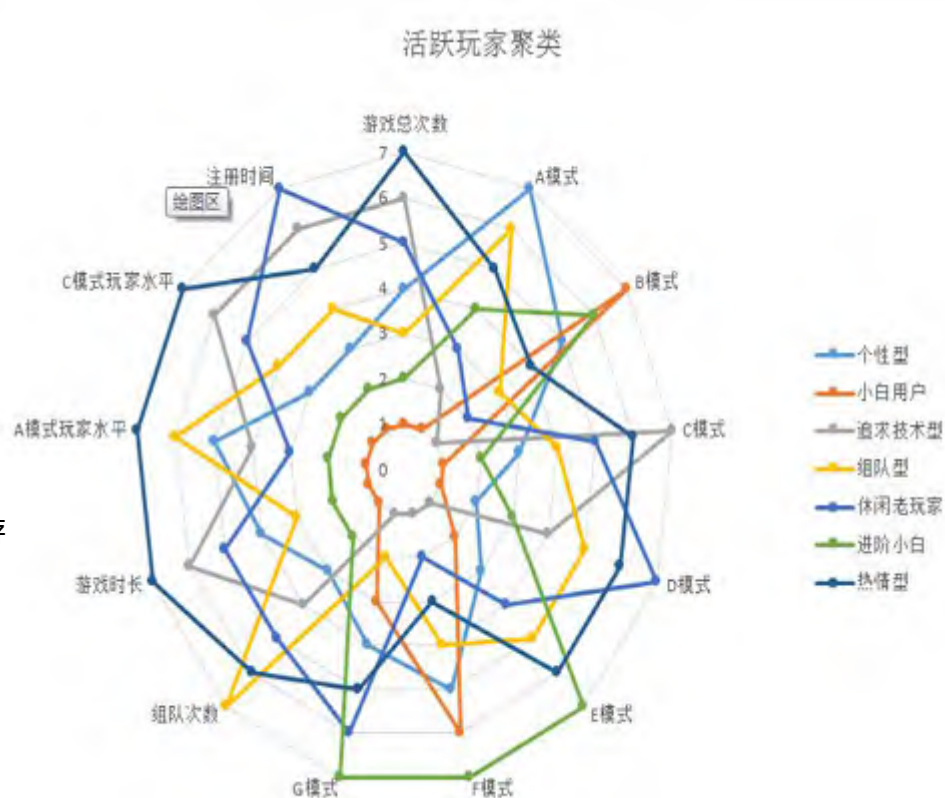


# 最终的用户画像标签数据



# 用户分群 - Kmeans的使用

- 小白用户 注册不久，游戏小白
- 进阶小白 对游戏有了一定了解，开始接触其它模式
- 个性型 孤胆英雄-A模式
- 社交型 独乐乐不如众乐乐-A模式
- 热情型 表现在游戏中投入度最高，水平也高，对游戏各模式都感兴趣，也爱组队
- 追求技术型 追求技术，追求名次
- 休闲老用户 注册时间最早的老用户，喜爱休闲玩法



# It is not easy

- 选择了不同的簇的数目，不同的数据预处理方法，不同的玩家特征得到较为理想的结果

*Not Too Easy.*

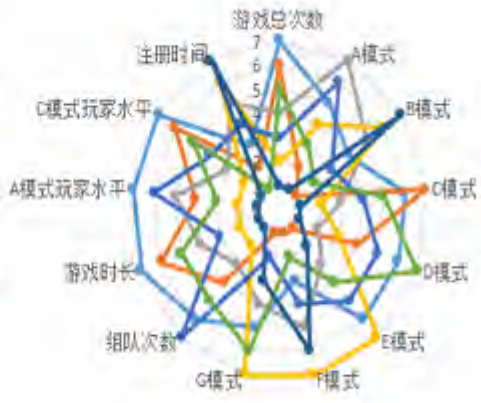
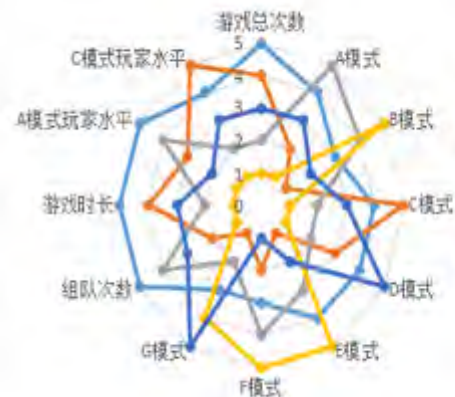
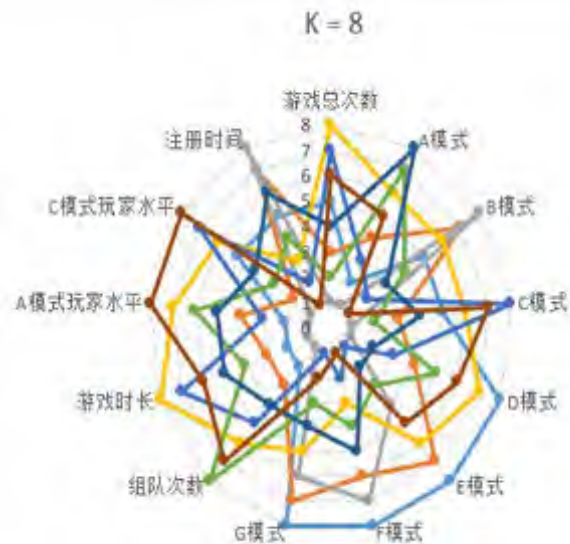
玩家 → 自然人

玩家拆分

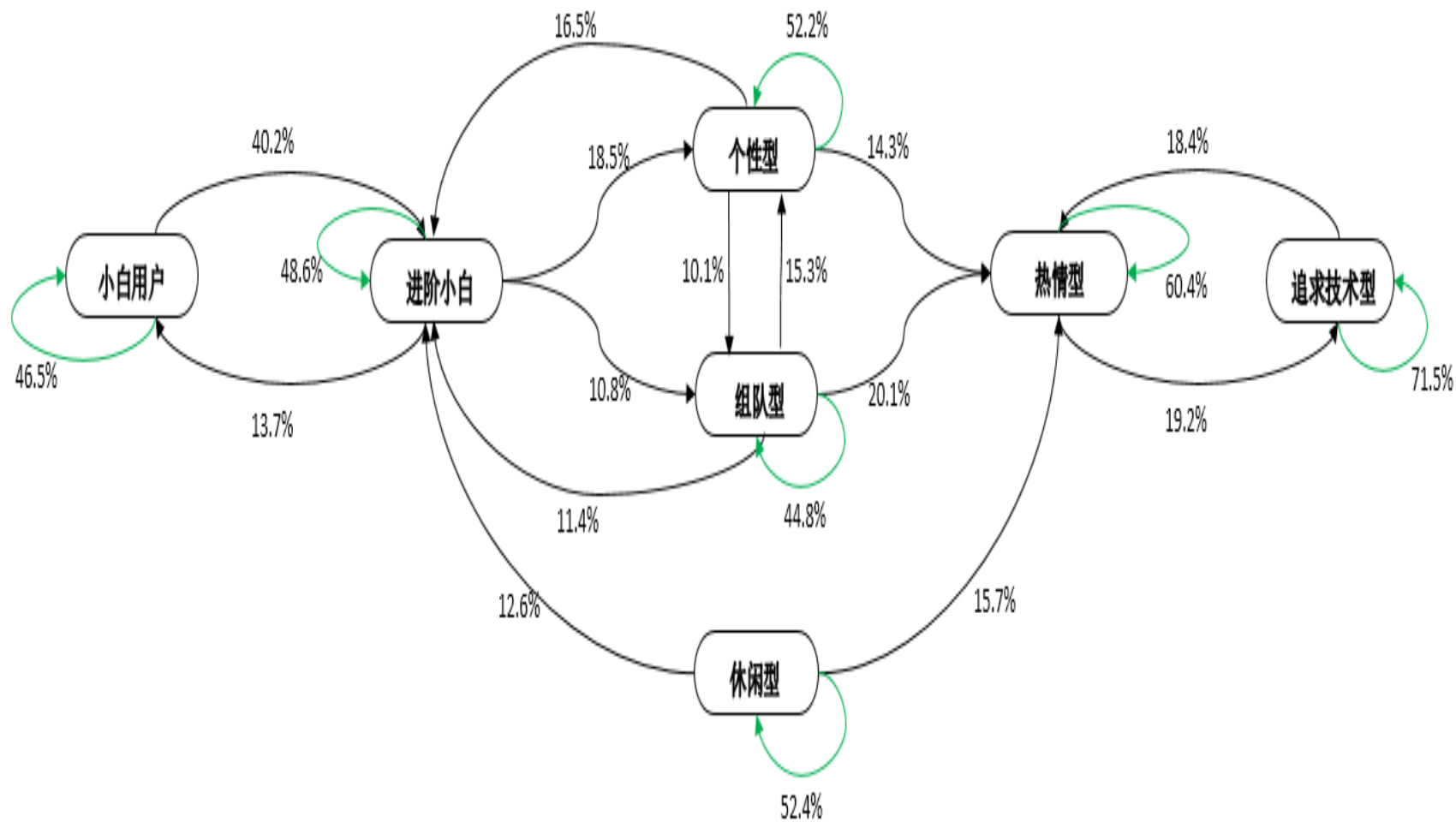
不同的数据预处理

不同玩家特征

不同的聚类个数



# 基于用户分群的状态变化分析



# 3: 道具推荐系统

# 解决的问题



# 如何做？











**Thanks!**

