

DTCC

数 / 造 / 未 / 来

第十二届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2021



2021 年 10 月 18 日 - 20 日 | 北京国际会议中心





58同城风控智能化体系的设计与实践

58同城信息安全部-张鹏



目录

风控业务的背景

风控智能化体系架构

风控智能化策略实践（感知策略 | 识别策略 | 关系计算）

总结



目录

风控业务的背景

风控智能化体系架构

风控智能化策略实践（感知策略 | 识别策略 | 关系计算）

总结



风控业务的背景 - 信息安全整体形态



58同城

招聘季月薪10000+ 搜索 免费发布信息

鱼缸 淘宝美工 夜班 日结 带车司机 招聘推广 修改/删除信息

首页 58APP 商家APP 58公众号 小程序 新车 金融 安居客 中华英才 58同镇 同城大学 我要推广

超职季 招聘 房产 二手车 二手车市场 宠物 本地商务 本地服务

全职 兼职 简历 租房 二手房 公寓 3万内 SUV 货车 手机 设备 回收 狗 猫 鱼 用品 工商 加盟 搬家 装修 修家电

北京房产 租房 整租 合租 品牌公寓 个人房源 二手房 安选房源 新房 商铺出租 商铺出售 生意转让 联合办公 写字楼出租 写字楼出售 办公楼盘 厂房 仓库 土地 车位

二手车市场 特惠九成新家具 摩托卖场 高性价比家电 数码通讯 二手手机 苹果/三星/小米 台式电脑 台式机/硬件 笔记本 苹果/联想 平板电脑 三星/iPad 数码产品 相机/游戏机 二手设备 机床/化工/箱房

北京招聘 北京求职简历 招聘会员 速聘直招 包吃包住 企业专区 周末双休 学习培训 热招职位 校园招聘 热门兼职 求职安全 高薪职位 求职神器 求职金卡 简历置顶 快速求职 求职机会翻倍 热招职位 销售 销售代表/电话销售/网络销售/销售经理

到家服务 招保洁/保姆上户快 家庭保洁 保姆月嫂 拉货搬东西 搬家 手机维修

本地服务大全 自主创业一站式服务 家政服务 搬家 居民/公司/小型/长途 空调移机/设备 保姆 月嫂/育儿嫂/钟点工/催乳师/保姆 保洁 家庭保洁/开荒/空气净化/除虫除蚁 家务 回收/开锁/疏通/鲜花/干洗改衣/白事 配送/桶装水/跑腿/粮油副食/煤气



北京 #一月辞职到现在, 2020年...

超职季 签到

全职招聘 兼职 租房 二手房 二手物品

二手车 本地服务 家政 商务服务 商铺写字楼

到家精选 借钱/理财 新房 游戏有奖 更多

[58部落]你新增了1个粉丝 '部落小助手'关注...

部落 最适合老实人的4个工作:“闷”却不费心,收...

教务管理 大山子 晒自拍 发放住房补贴 低价好房出租 分享美好生活

晒工资 美容美发圈 北京人 晒工资晒幸福 烫染洗剪吹 北京人一起聊

推荐 部落 视频 热议 家乡

来自部落频道 hvu7kribv 所在地保密

我在北京当保安已4年, 年龄一年一年大, 今年63岁, 本公司有把我辞退的打算, 我不想回老家, 还想

首页 发现 发布 消息 我的

本地生活服务入口
覆盖房产、招聘、二手车、本地服务等行业



数 / 造 / 未 / 来



风控业务的背景 - 风控业务案例

立减两个月租金 [模糊] 独卫 押一付

22小时前

 房屋租赁
合同范本

 经纪服务
合同范本



1900 元/月 押一付一 佣金: 无佣金

租赁方式: 合租 - 次卧 - 男女不限

房屋类型: 3室1厅1卫 15.00 平 精装修

朝向楼层: 南 高层 / 16层

所在小区: [模糊]

所属区域: [模糊]

详细地址: 西 [模糊] 路 [附近高薪工作](#) [查看地图](#)

电话联系TA

 微聊

你好，在吗

什么时候方便看房？

不常上，+我V A①②⑤⑤⑤⑤⑤③⑤



明天方便看房子吗，加一下我V详聊



数 / 造 / 未 / 来



风控业务的背景 - 风控业务案例



风控业务的背景 - 信息安全整体形态



内容安全层

涉政、涉恐、涉爆、色情、赌博、毒品、违禁品.....

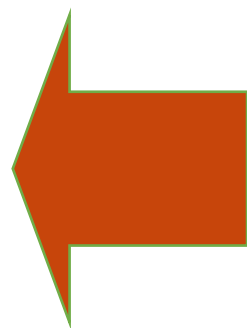


业务安全层

作弊、薅羊毛、信息虚假、灌水、吸粉、刷钻.....

账号安全层

恶意注册、拖库、洗库、撞库、盗号、养号、僵尸号.....



资源安全层

猫池、IMEI篡改、VPN、秒拨、接码、群控、云控.....

数据安全层

权限、脱敏、混淆、反爬、终端加固、接口加密.....

基础安全层

主机保护、病毒木马、漏洞管理、DDOS、渗透.....



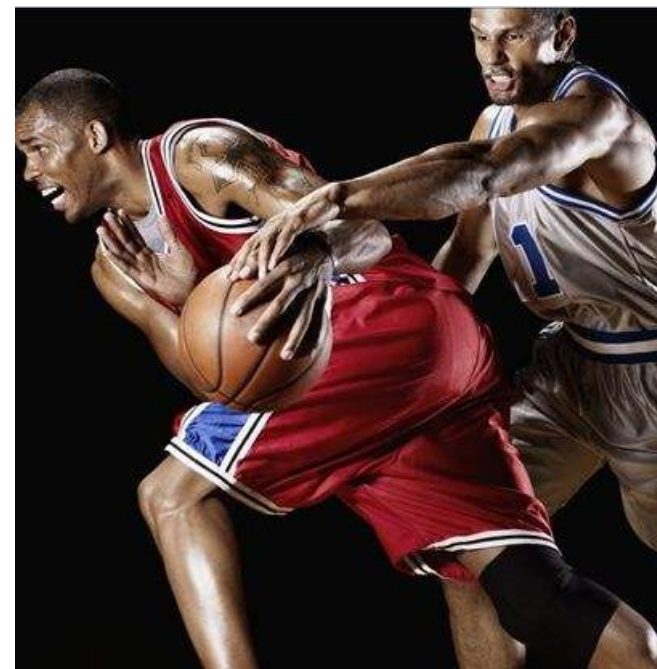
风控业务的背景 - 风控业务的特点



业务复杂



隐蔽性强



对抗性强



目录

风控业务的背景

风控智能化体系架构

风控智能化策略实践（感知策略 | 识别策略 | 关系计算）

总结





风控智能化体系架构 - 构建思路

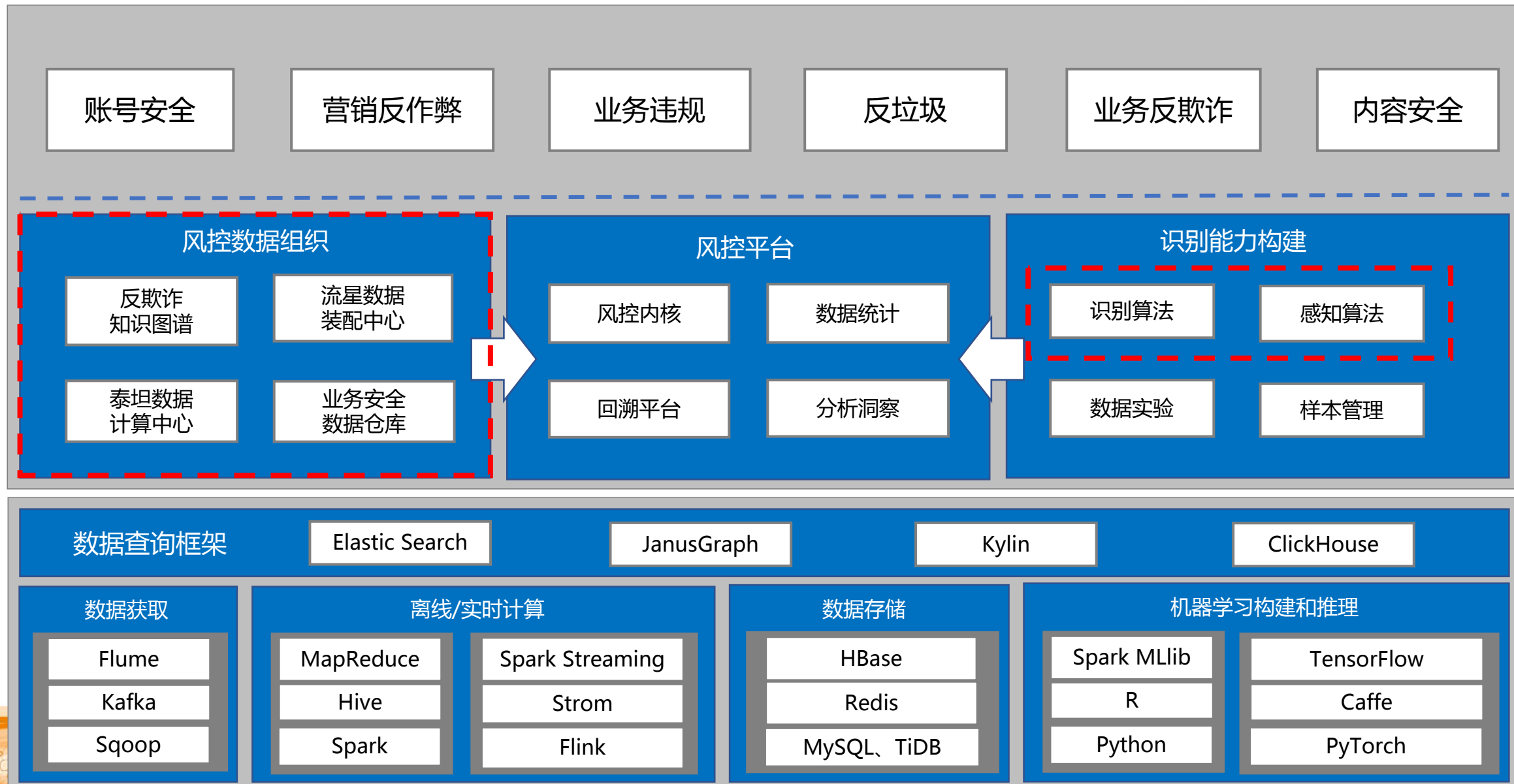


风控智能化体系架构 - 整体业务架构

风控应用

风控主体

技术支撑



风控智能化体系架构 - 数据架构

业务安全实时风控面临的挑战：

海量调用
千亿

特征值实时计算次数达千亿，高峰时段QPS近千万

低延迟
50ms

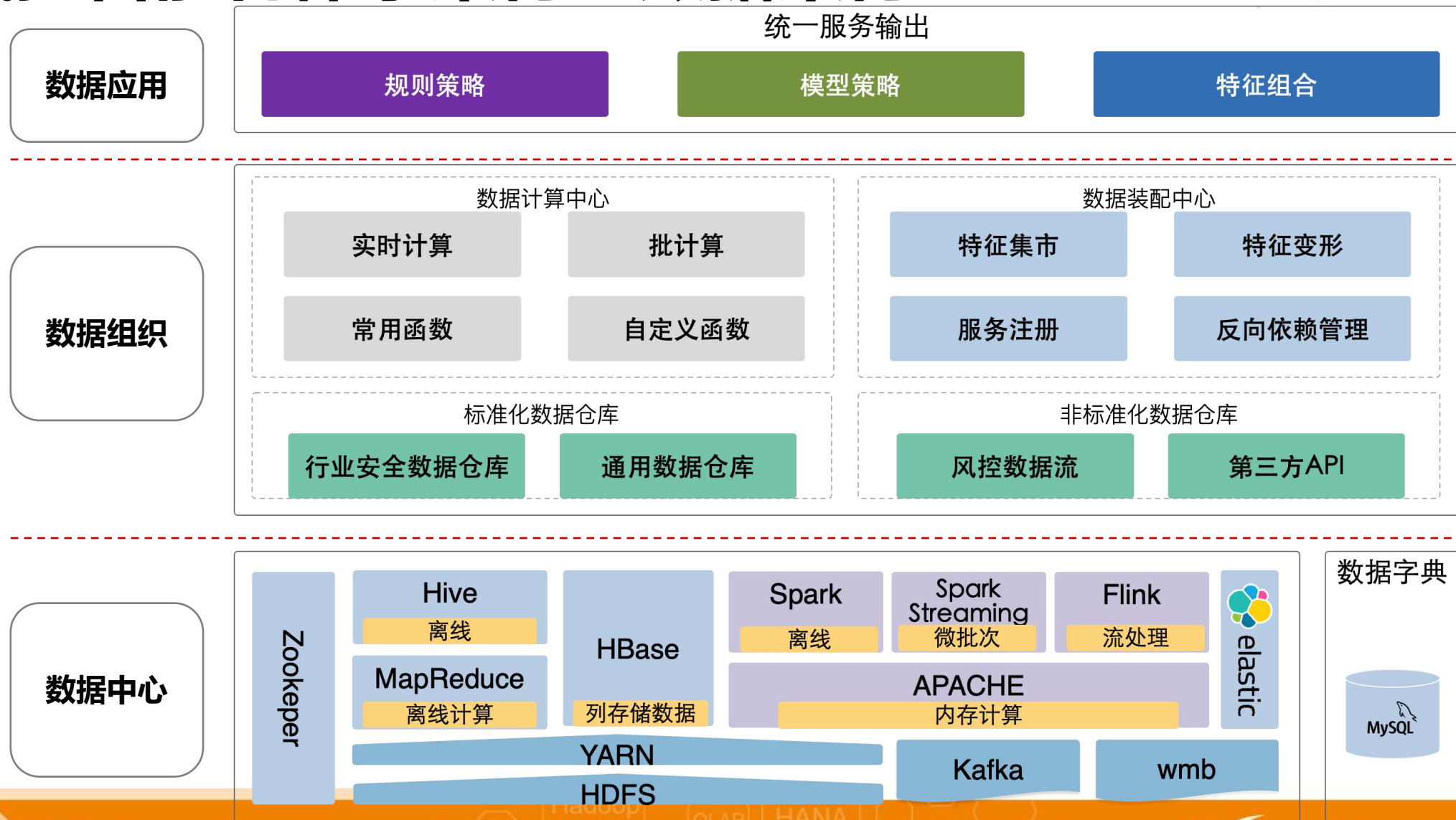
为保证线上体验，数据组织延时要控制在50ms

多路数据
40+

实时风控对应策略，需要的数据不在当前流中



风控智能化体系架构 - 数据架构



目录

风控业务的背景

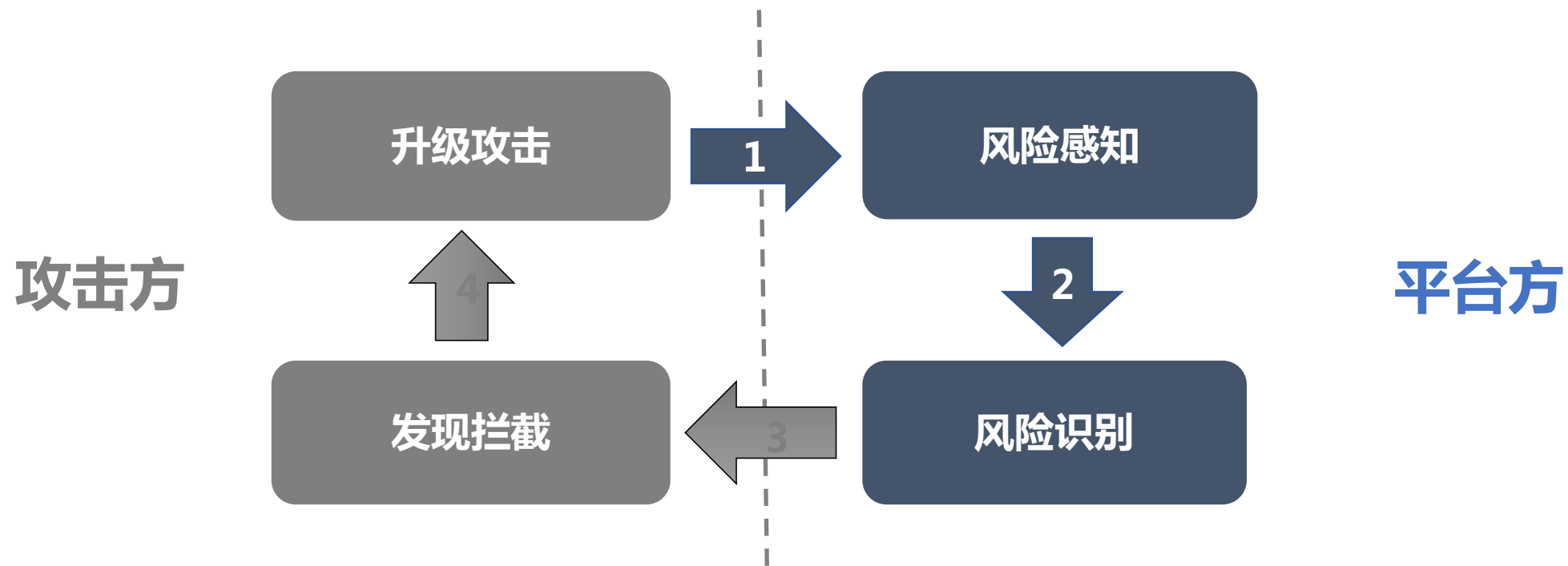
风控智能化体系架构

风控智能化策略实践 (感知策略 | 识别策略 | 关系计算)

总结



风控智能化策略实践 - 构建思路



风控智能化策略实践 - 构建思路

自动化处理

识别类策略要求：

- 建立在**高精度识别**基础上，提升召回量；
- 对问题进行**精准打击**，不做模糊问题的识别；
- 建立**自动化的迭代**的机制，降低维护成本；



监督学习 (风控融合决策引擎)

人机协同

感知类策略要求：

- 建立在**高召回量**的基础上，提升准确率；
- 对问题**不做定性**，交给人工进行分检和定义；
- 具备较强的**通用性和抗攻击能力**；

半监督&弱监督

专家处理

入侵检测&无监督

未知空间
(未召回&未定义)



风控智能化策略实践 - 构建思路

	感知类策略	识别类策略
建模方式	通用建模	问题建模
准确性	不要求	高
召回性	高	不要求
时效性	中	高
维护频率	低	高
处理方式	人工处理	自动处理

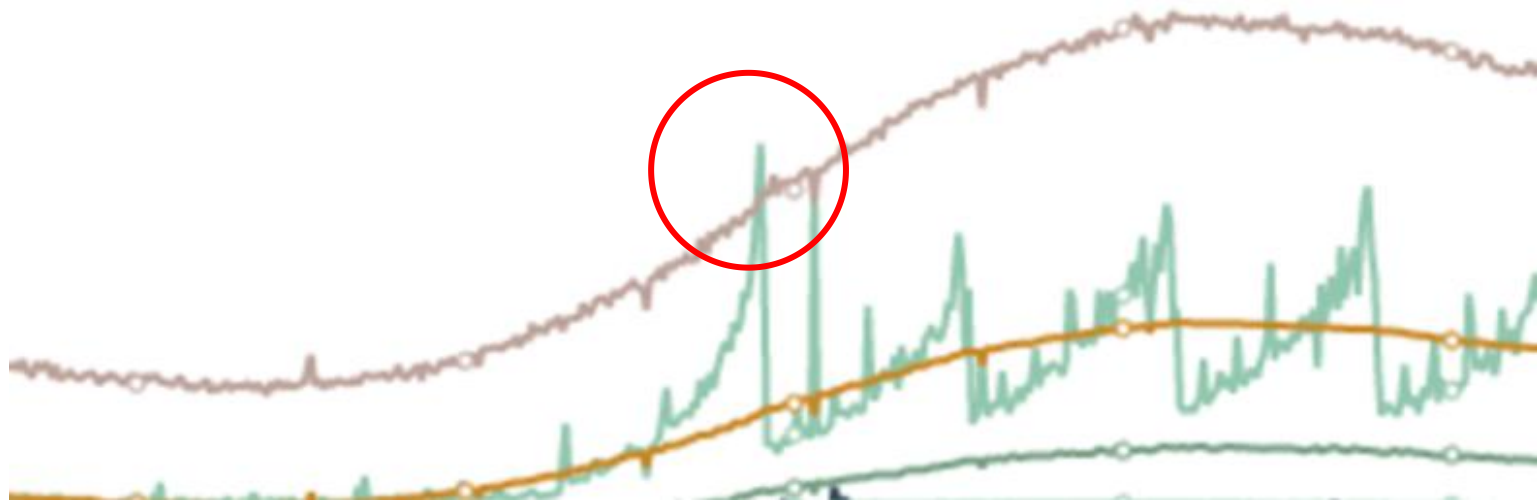




风控智能化策略实践 - 感知类策略

波动风险检测面临的挑战：

- 1、如何感知到波动。
- 2、如何定位到原因。



风控智能化策略实践 - 感知类策略

波动发现 - Prophet

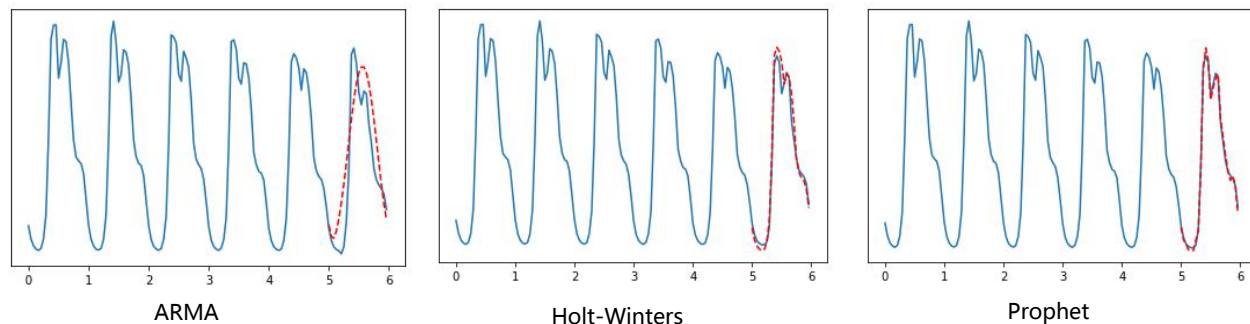
$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

趋势函数 $g(t)$ ：表示时间序列上的非周期变化趋势；

周期函数 $s(t)$ ：表示时间序列周期性变化，一般以周或年为单位；

节假日函数 $h(t)$ ：拟合节假日和特殊日期；

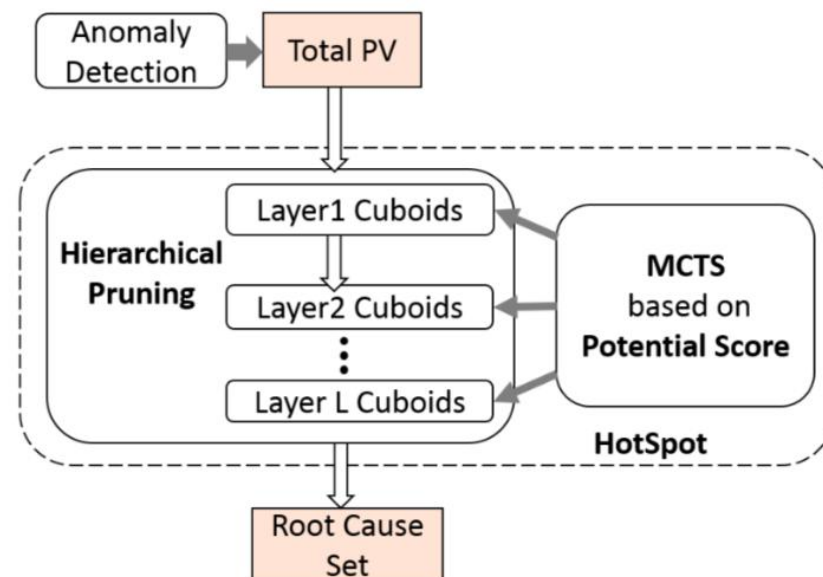
误差项 ϵ_t ：表示误差项或者称为剩余项。



根因分析 - HotSpot算法

$f(p, i) \rightarrow v(p, i)$		Province(p)			
		Beijing	Shanghai	Guangdong	*
ISP (i)	Mobile	20 → 8	15 → 15	10 → 10	45 → 33
	Unicom	10 → 4	25 → 25	20 → 20	55 → 49
	*	30 → 12	40 → 40	30 → 30	100 → 82

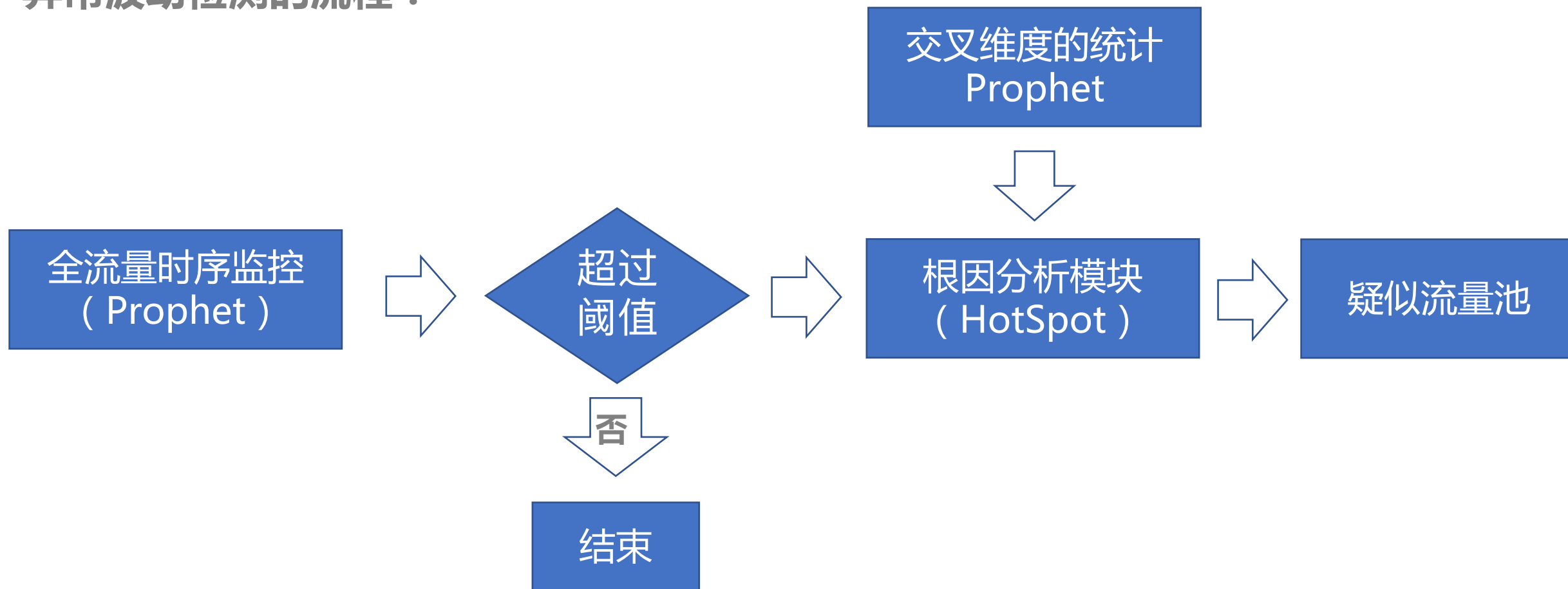
https://blog.csdn.net/watson_3583





风控智能化策略实践 - 感知类策略

异常波动检测的流程：



风控策略智能化 - 感知类策略

聚集行为检测 - Grid-density

离群行为检测 - isolation Forest



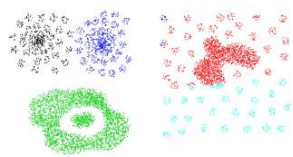
(f) MLLib K-Means|| on Spiral data



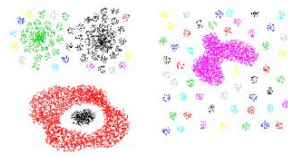
(g) Spark-DBSCAN on Spiral data



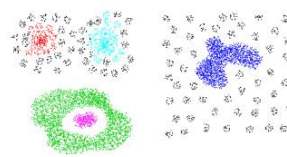
(h) PatchWork on Spiral data



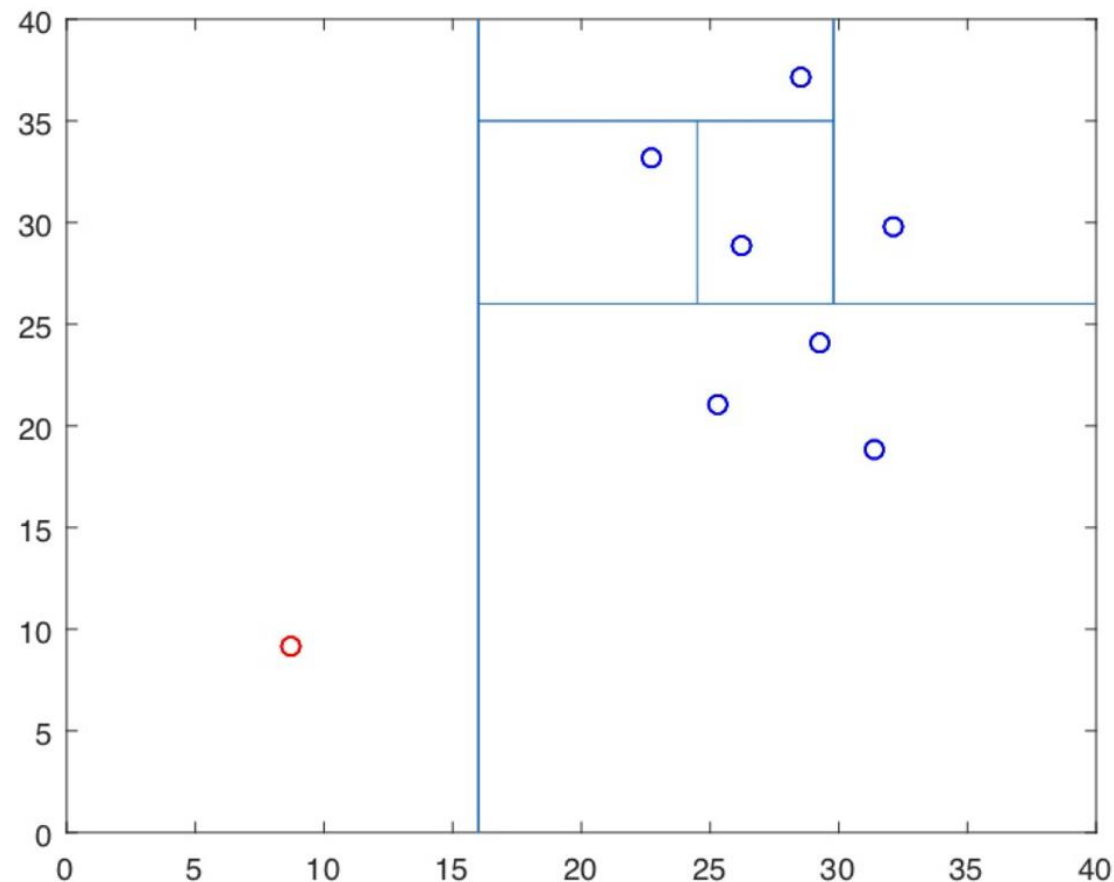
(n) MLLib K-Means|| on Compound data



(o) Spark-DBSCAN on Compound data



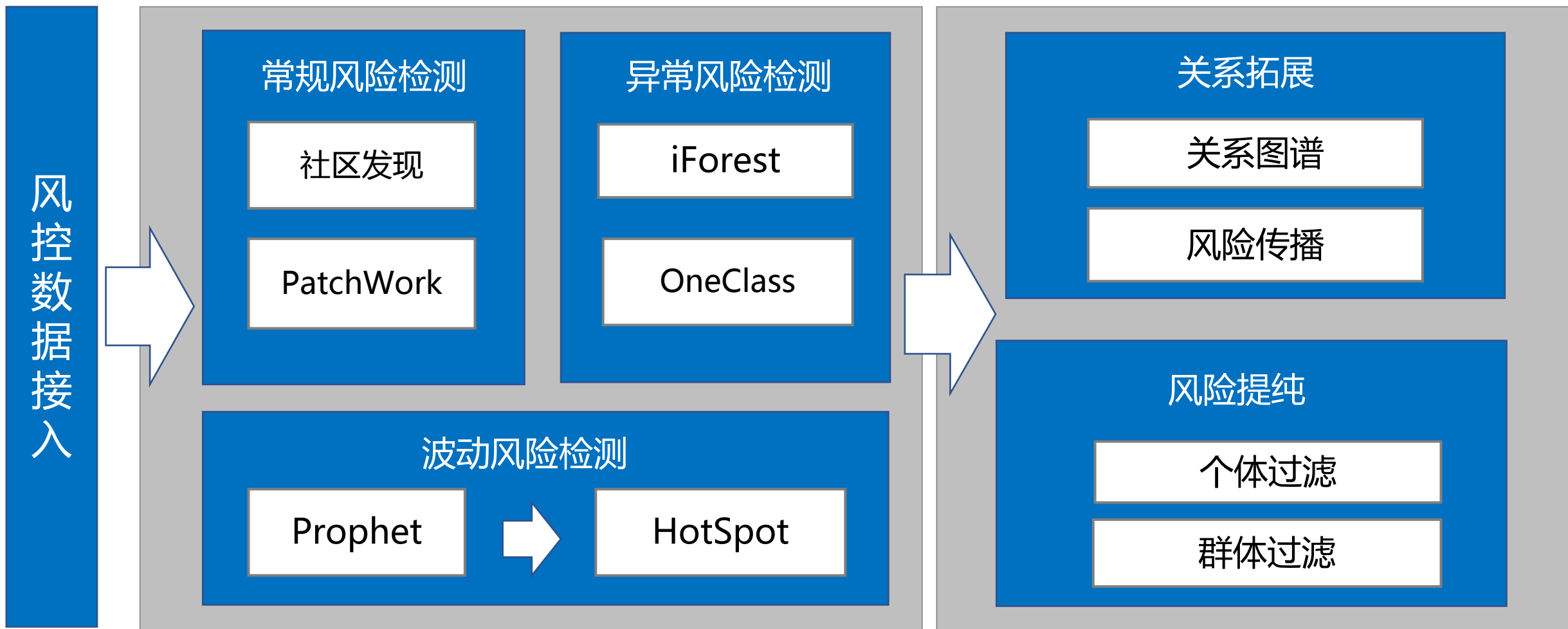
(p) PatchWork on Compound data



	10W	30W	50W	70W
DBSCAN(密度聚类)	5.2min	25min	>1h	-
PatchWork(网格密度聚类)	3.1min	3.7min	4.4min	5.3min



风控智能化策略实践 - 感知类策略



风控智能化策略实践 - 构建思路

	感知类策略	识别类策略
建模方式	通用建模	问题建模
准确性	不要求	高
召回性	高	不要求
时效性	中	高
维护频率	低	高
处理方式	人工处理	自动处理



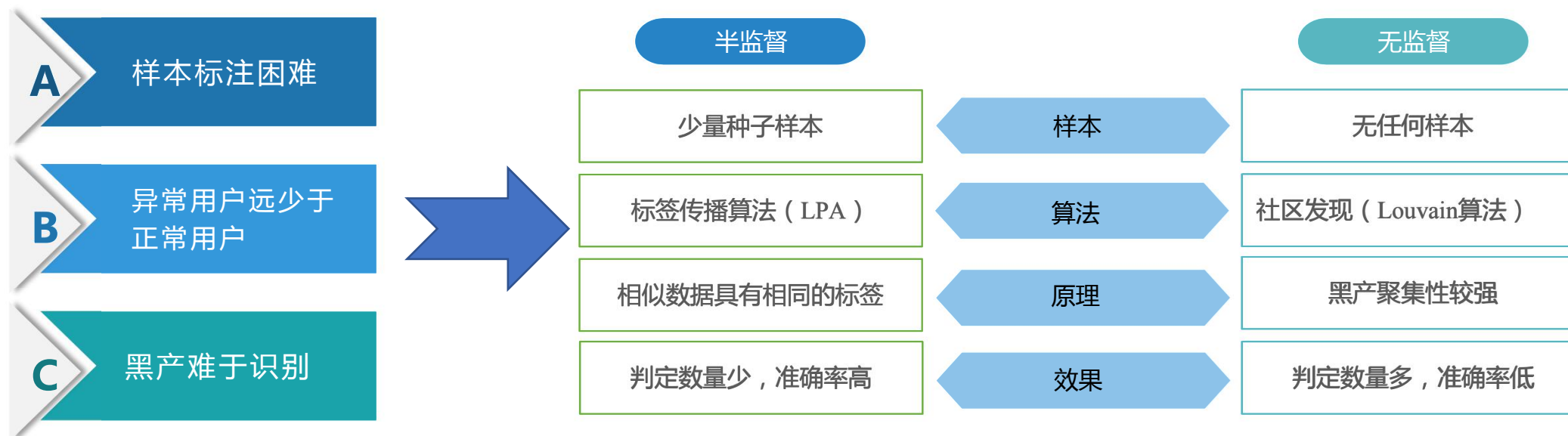
风控智能化策略实践 - 识别类策略

策略类型	构建成本	维护成本	对抗性	使用场景
名单类策略	低	低	差	快速止血
关键词策略	低	低	差	快速止血
规则类策略	低	低	差	迅速相应
事件标签策略	中	中	强	反欺诈精准打击
画像标签策略	中	中	强	业务违规长期治理
融合决策策略	高	高	强	长期治理



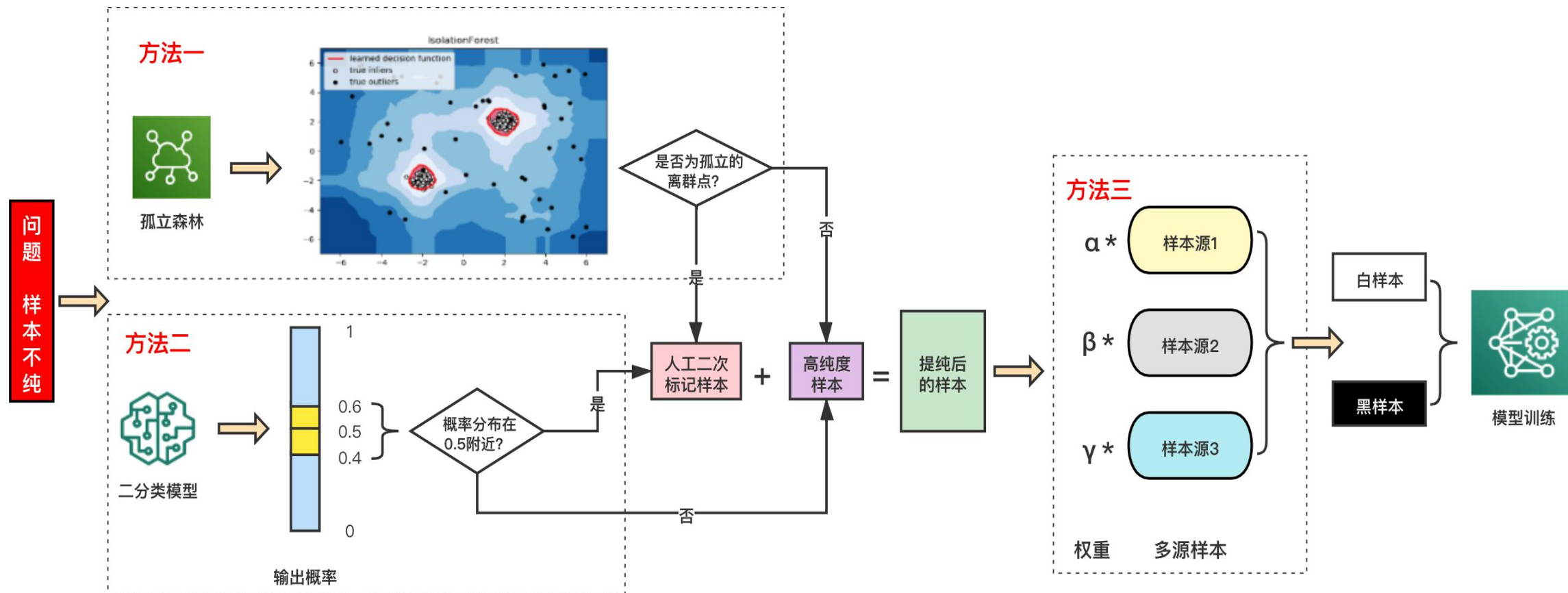
风控智能化策略实践 - 识别类策略

长尾问题的识别



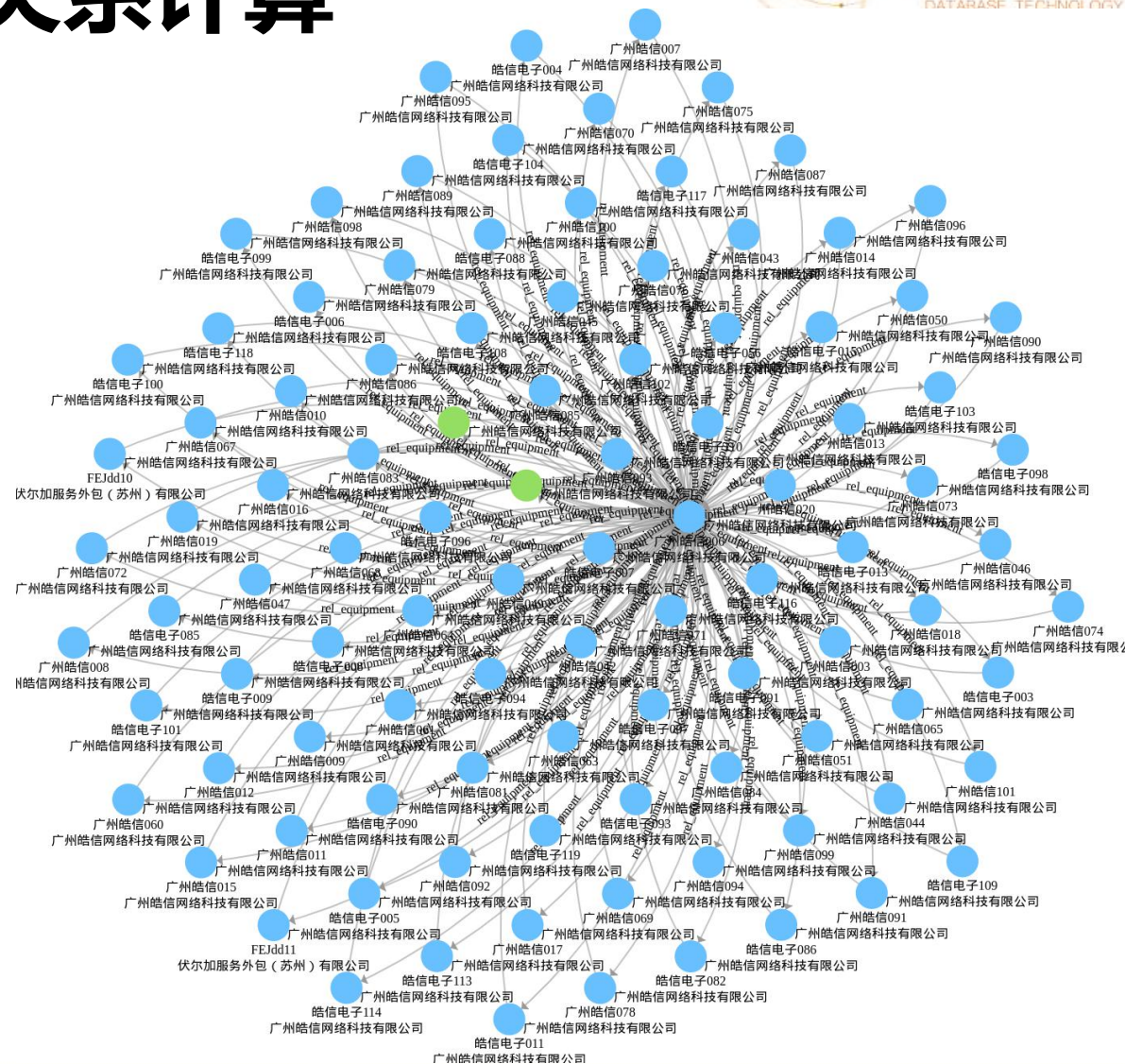
风控智能化策略实践 - 识别类策略

样本纯度的问题

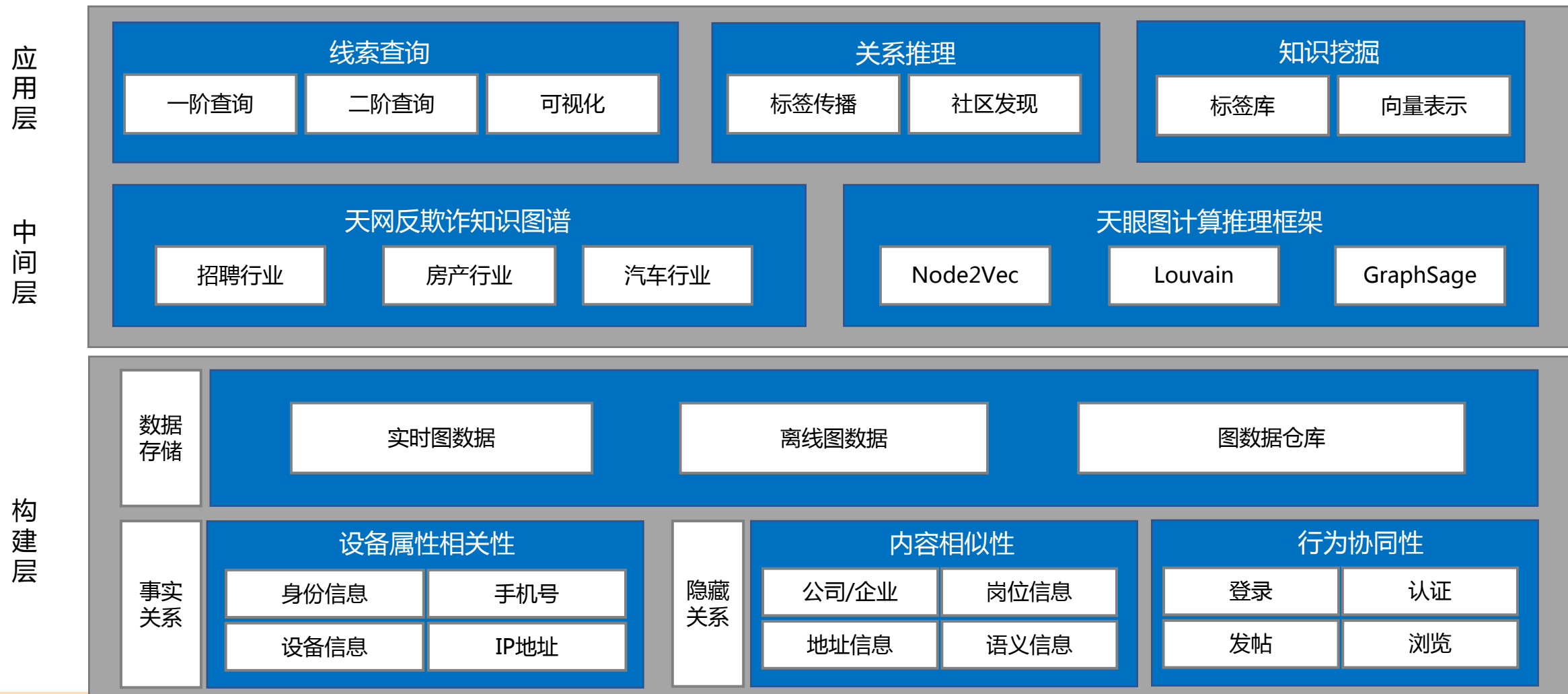




风控智能化策略实践 - 关系计算



风控智能化策略实践 - 关系计算



目录

风控业务的背景

风控智能化体系架构

风控智能化策略实践（感知策略 | 识别策略 | 关系计算）

总结



总结

分享回顾

- 1、风控业务的特性
- 2、风控智能化整体架构设计
- 3、数据组织架构设计
- 4、感知和识别策略能力的实践
- 5、关系计算的架构设计





THANKS