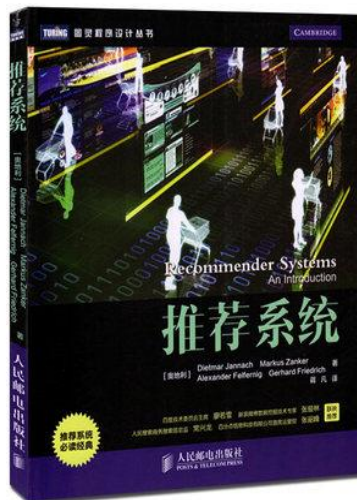


外卖物流配送的 大数据创新实践

SPEAKER

蒋 凡

- 2006年加入百度搜索团队
- 2012年加入百度知识图谱团队
- 2014年加入百度外卖团队，负责物流调度、个性化推荐、画像等技术方向
- 译著

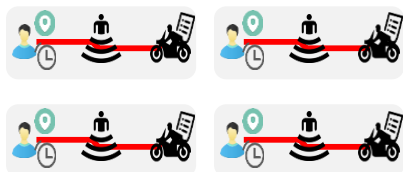


- 物流调度的云端计算模式
- 深度学习智能预测算法
- 大数据可视化平台

智能物流发展历程

调度系统 1.0

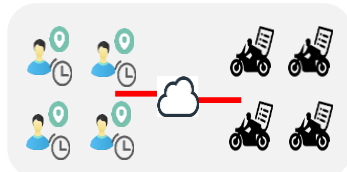
人工派单模式



- 调度员根据订单地址和骑手位置来进行订单分配
- 人力调度派单峰值有限

调度系统 2.0

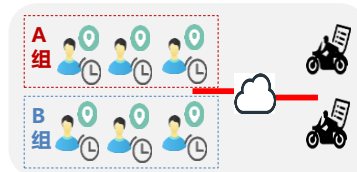
系统派单模式



- 系统综合考虑配送距离、骑士运力、期望送达时间等因素来自动派单

调度系统 3.0

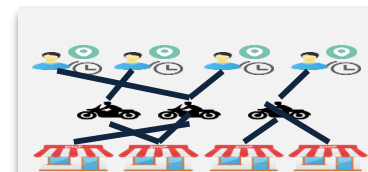
云端分组派单模式



- 系统综合考虑各因素进行订单分组，然后再指派给合适的骑手

调度系统 4.0

深度学习智能模式



- 出餐时间估算更准，缩短骑手到店等待时间，节省运力，提升用户等餐体验

10:00
登录

10:05
下单

10:15
压单

10:30
改派

10:50
送达

预估调度

骑士A

骑士B

骑士C

即时调度

骑士A

骑士B

骑士C

延时调度

订单1

订单2

订单3

订单X

订单Y

订单Z

骑士A

骑士B

骑士C

改派调度

骑士A

骑士B

骑士C

订单1

订单2

订单3

订单X

订单Y

订单Z

场景优化

商圈

城市

时段

天气

运力

场景调参

最优配送骑士及路径

动态规划

空间开销函数

时间开销函数

惩罚函数

函数建模

多项函数分值

约束条件

商户

用户

骑士

订单（预期时间、起止地点...）

数据采集

实时场景下的约束变量

云端计算模式

1. 代价打分函数

时间 T 有订单 O 要分配给候选骑士 R_1, R_2, \dots, R_n ,
 分别估算骑士 R_x 配送这笔订单的实际开销

$$f(x) = k_1 * f_1(x) + k_2 * f_2(x) + \dots + k_n * f_n(x)$$

0. 多目标优化

$\min(f_1(x), f_2(x), \dots, f_n(x)), s.t. g(x) \leq 0$

2. 场景模拟调参

解析骑士在岗率、平均压单数、订单出单位置密度等角度的基础物理分布函数

3. 多次最优匹配

二分图最大匹配算法

4. 云端虚拟队列

压单时间窗口



- **效果评估**

并单的情况下，系统为了完成订单配送实际耗费的时间成本

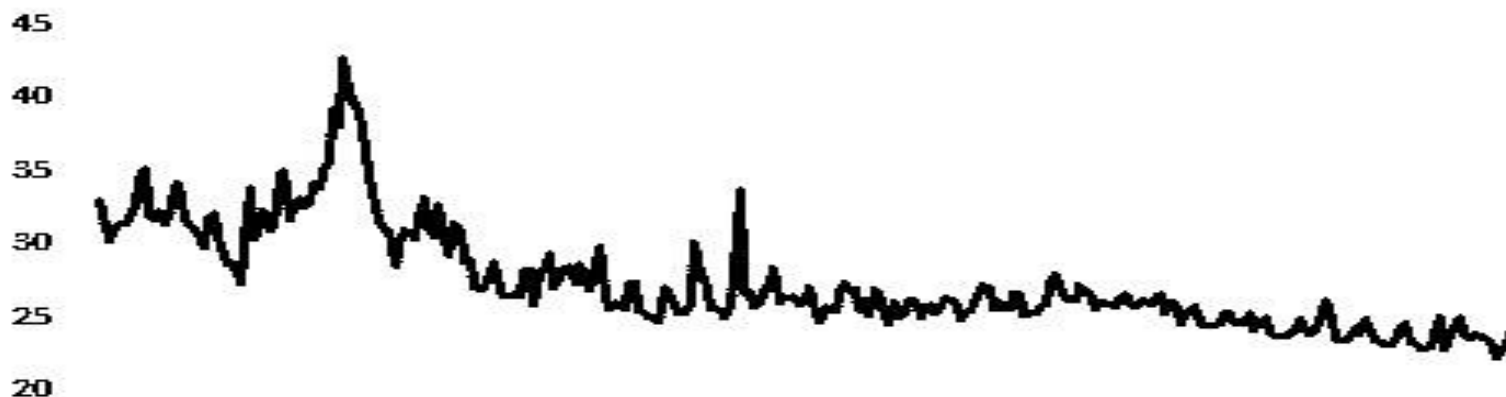
- **充分并单**

40分钟同时送到，单均耗时40分钟，单均成本是 $40/2=20$ 分钟

- **不充分并单**

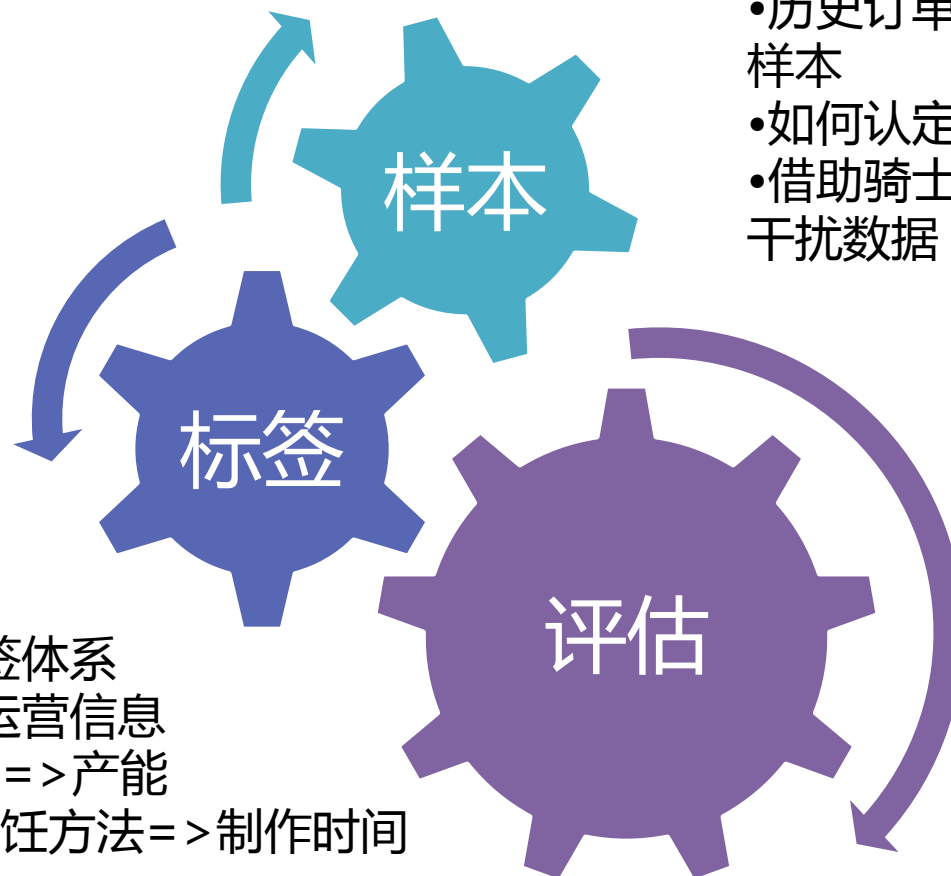
第1单30分钟送到，第2单50分钟送到，单均成本是 $50/2=25$ 分钟

单均成本（分钟）



上线前 上线后

深度学习智能预测算法

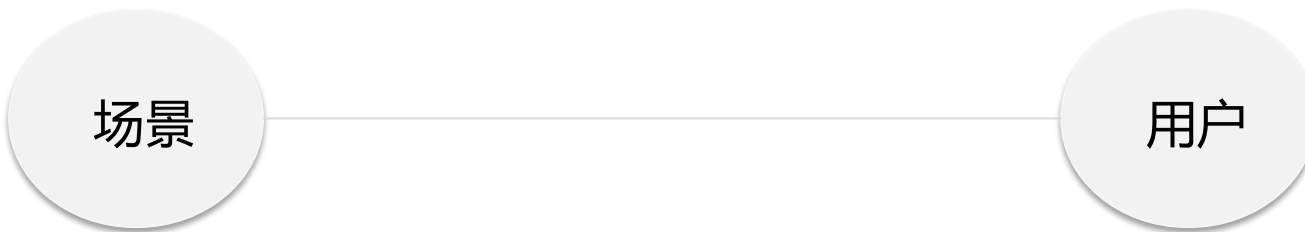


- 商户和菜品标签体系
- 竞品和本品的运营信息
- 品牌、营业面积=>产能
- 材料、价格、烹饪方法=>制作时间

- 历史订单数据中抽取出千万级训练样本
- 如何认定这些订单的实际出餐时间
- 借助骑士轨迹、停驻坐标，清洗出干扰数据

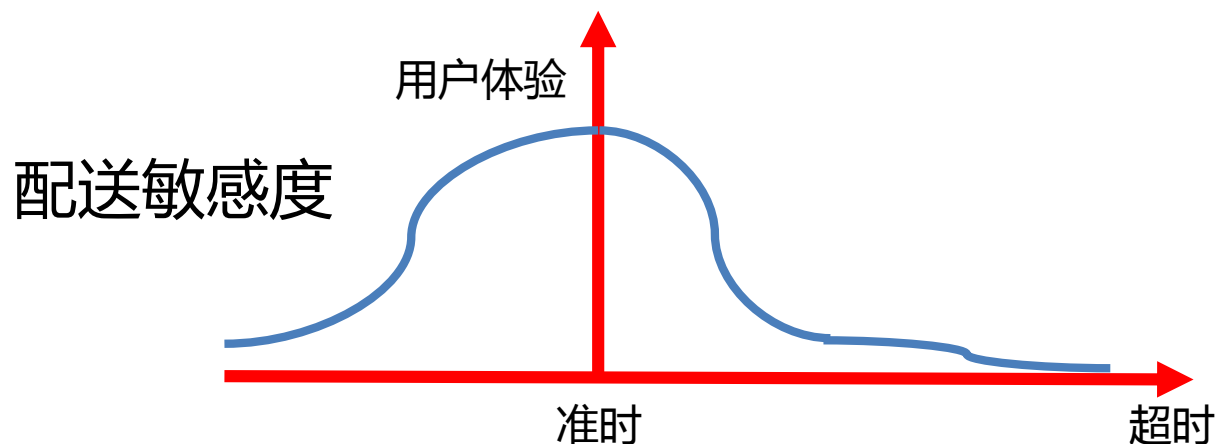
- 不同出餐时长的菜品预估时间要求误差是不同
- 时长区间样本敏感的深度学习评估框架

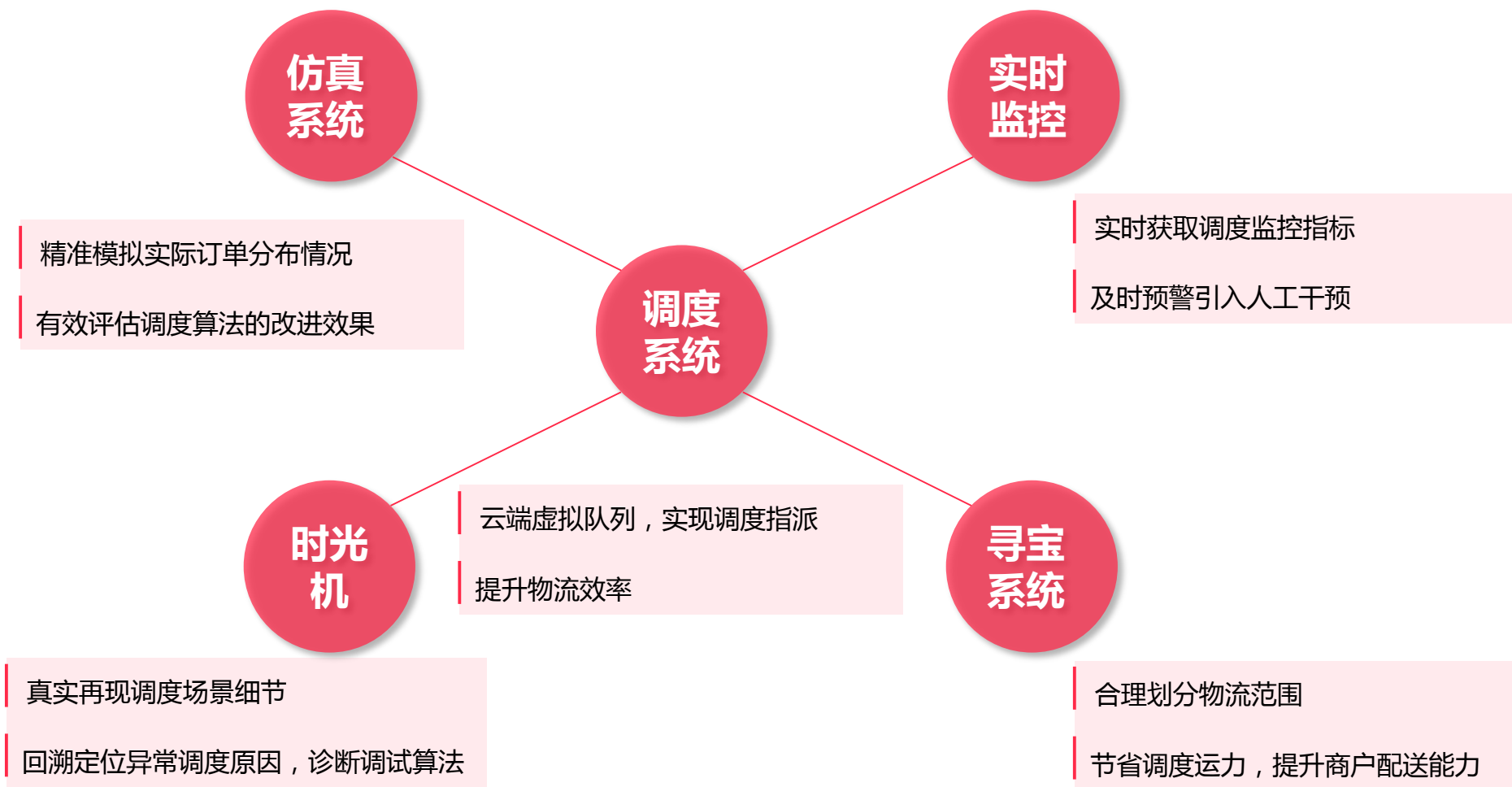




- **时间分段**：上午/中午/下午/晚上
- **日期分段**：工作日/节假日
- **运力情况**：是否爆单
- **天气情况**：雨雪雾霾

- **配送偏好**：配送敏感度
- **购买能力**：小区房价、商圈位置
- **用户价值**：消费频次、消费金额
- **用户身份**：校园、住宅、白领





实时监控平台

实时监控

实时获取调度监控指标

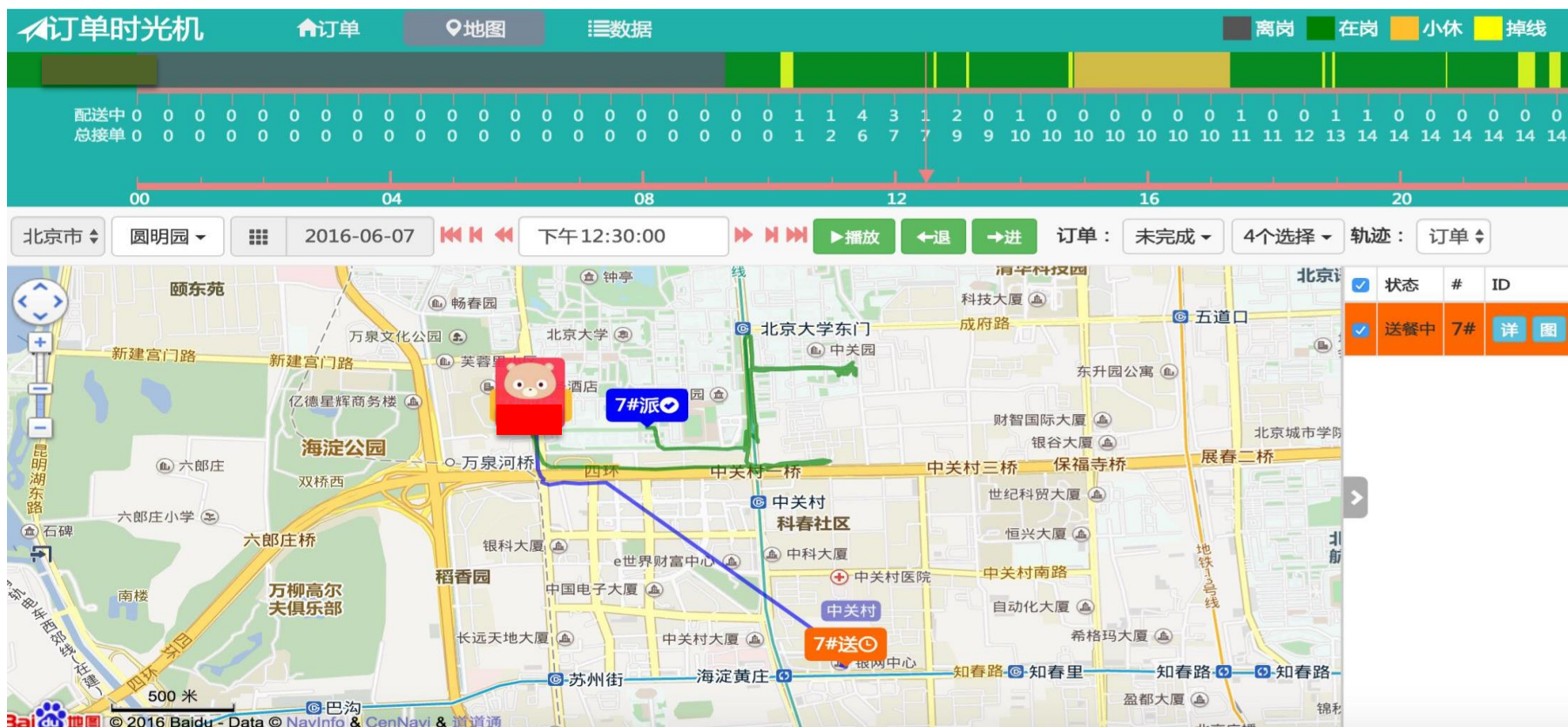
及时预警引入人工干预



时光
机

真实再现调度场景细节

回溯定位异常调度原因，诊断调试算法

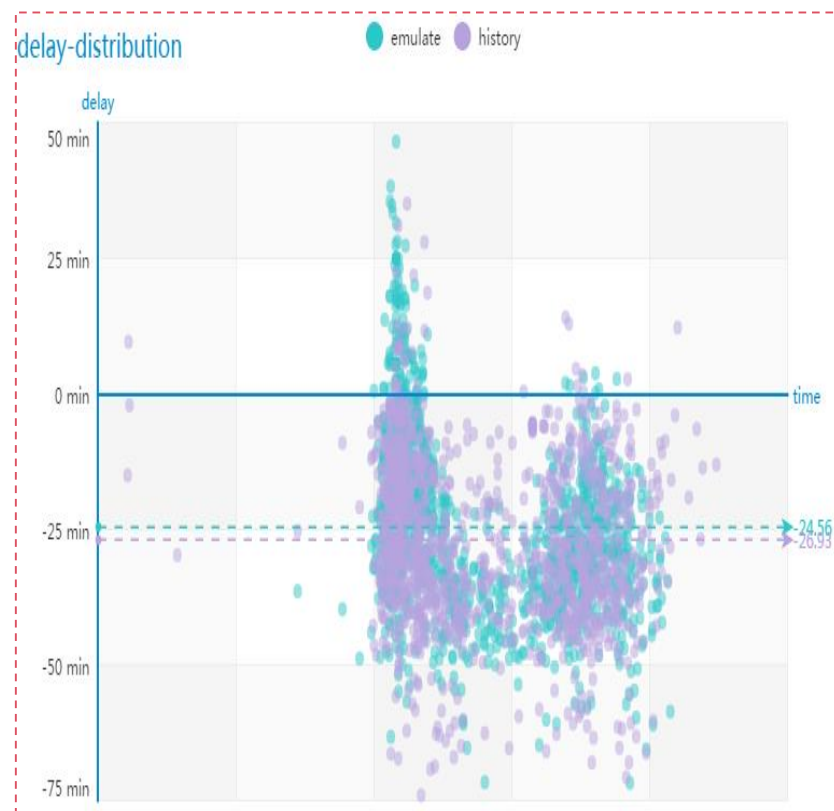
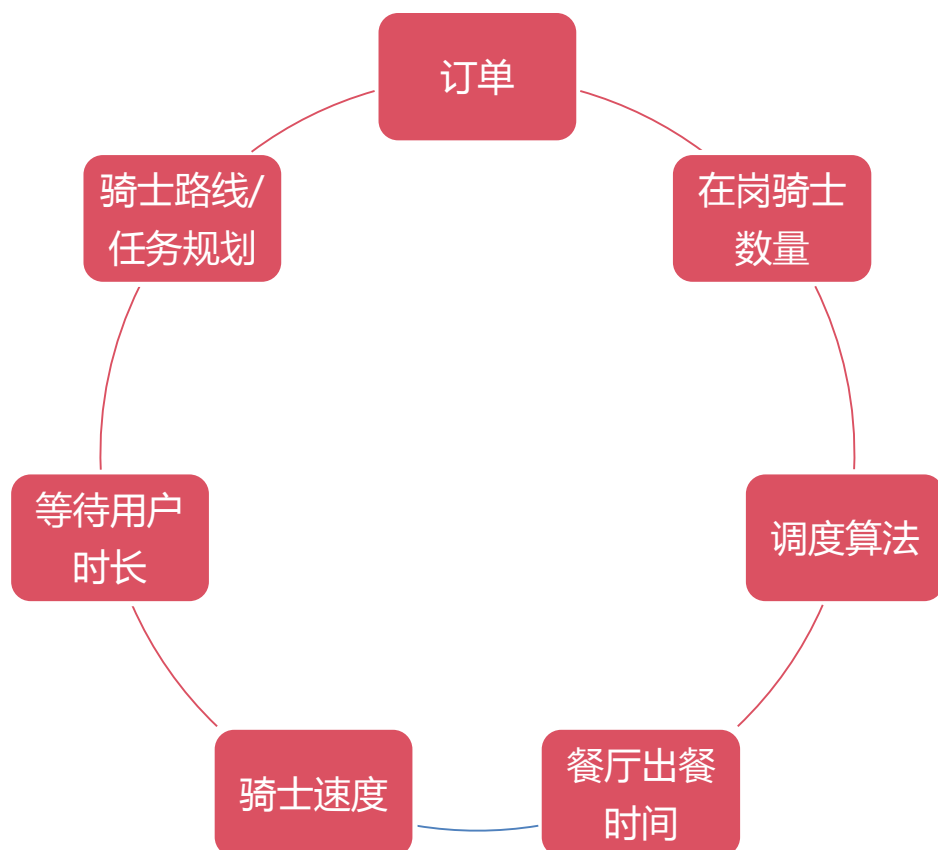


仿真平台

仿真
平台

精准模拟实际订单分布情况

有效评估调度算法的改进效果



寻宝
平台

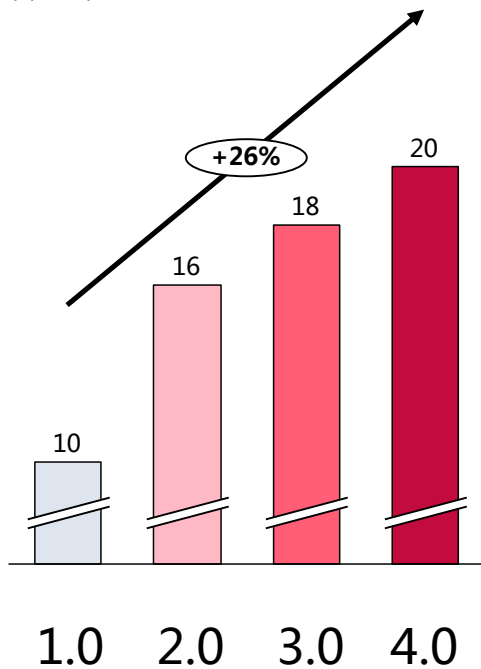
合理划分物流范围

节省调度运力，提升商户配送能力



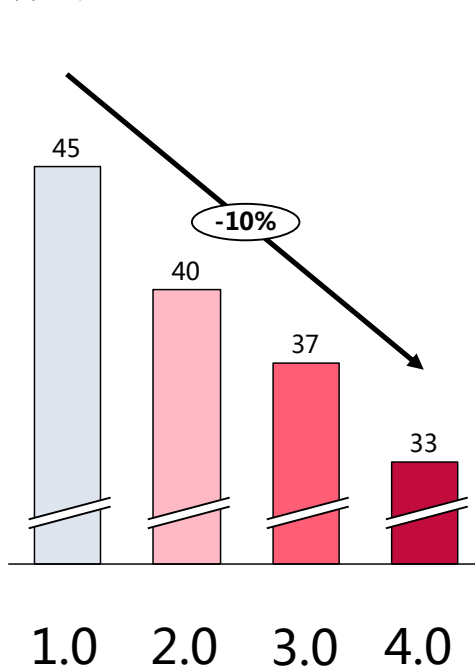
人均配送单数不断提升

单位：单/人/天



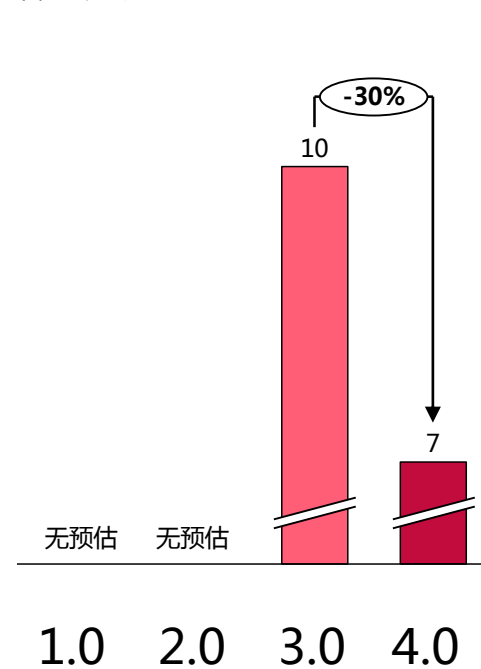
平均配送时长逐步降低

单位：分钟

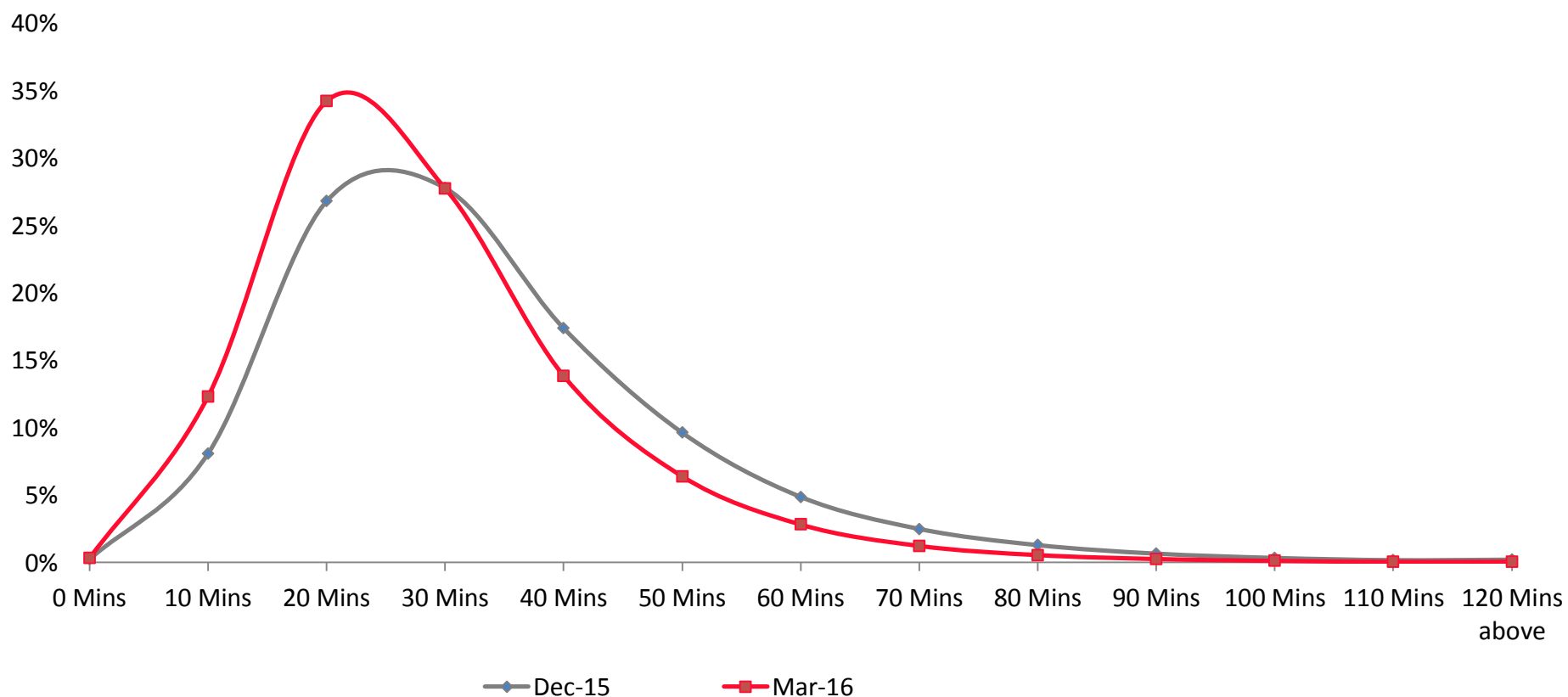


商户出餐时间预估误差缩减

单位：分钟

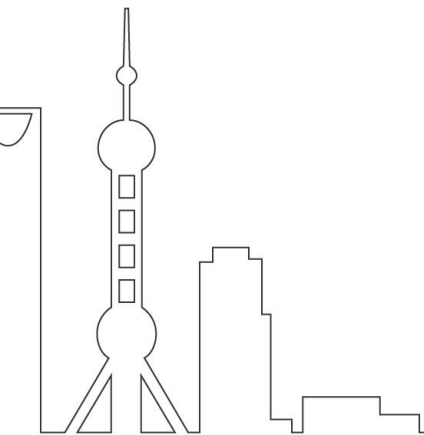


用户等待时长分布 - Dec 2015 vs. Mar 2016





- 利用**深度学习模型**对大量基础特征的组合训练能力
- 自动构造打分算法所依赖调度因素的组合或约束结构
- 形成多层反馈神经网络，找出最优分配方案
- 从**历史调度数据**找出影响整体指标和特定场景体验实例
- 作为训练学习的负面样本，结合线下仿真系统的模拟演练
- 自动修正机器学习模型的优化方向，取代人工参与方式



Thanks!

International Software Development Conference