百度智能异常检测实践

百度 王博

个人介绍

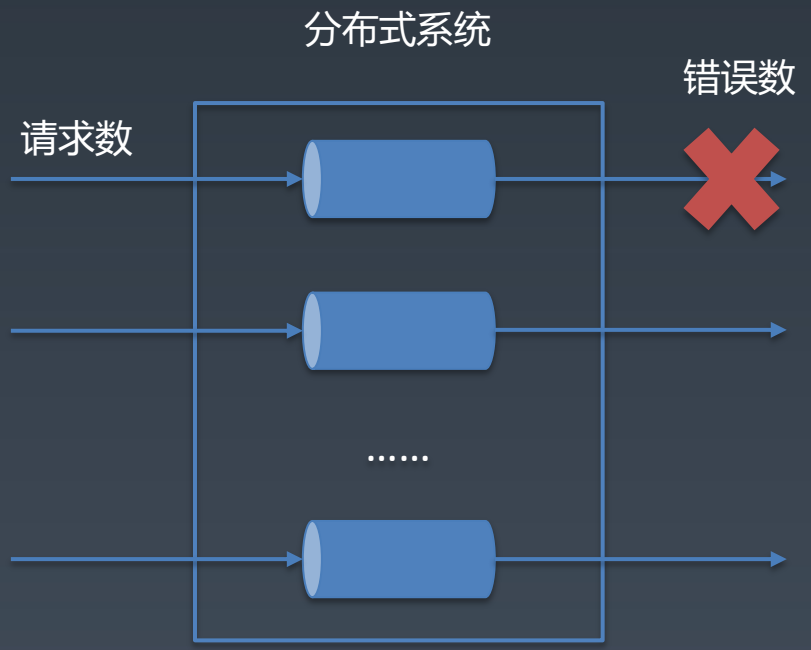
- 王博
- 2014年加入百度
- 智能运维平台Noah
- AIOps
  - 异常检测
  - 故障诊断
  - 实例排布

演讲大纲

- **黄金指标异常检测技术**
  - 异常检测问题及难点
  - 运维黄金指标监控方法
- 百度AIOps产品及效果
  - 产品形态
  - 产品使用效果
  - 百度内外部落地效果

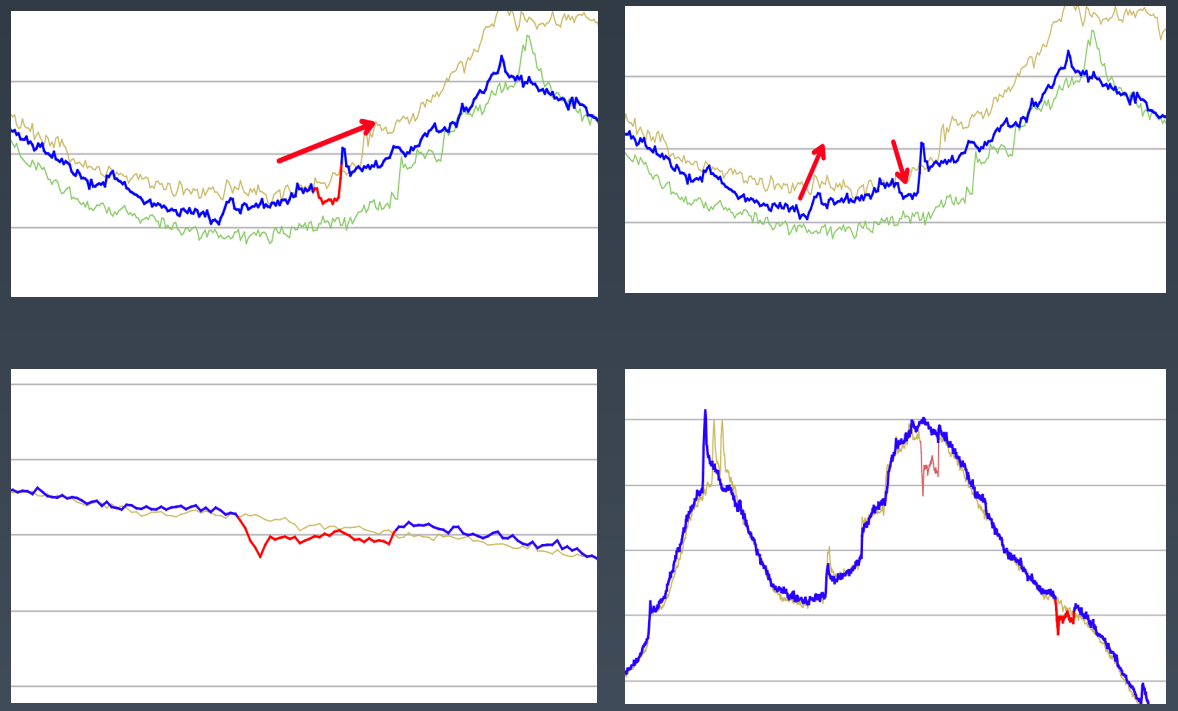
黄金指标异常检测

- 运维黄金指标
  - 请求数（流入状态）
  - 错误数（流出状态）
  - 响应时间（用户响应感受）
  - 系统容量（系统并发负载）
- 难点
  - 百万级指标
  - 配置成本高
  - 监控精度高



黄金指标异常检测思路

- 方案分析
  - 异常数据极其罕见（<1%）
  - 标注成本及错误率高
- 思路
  - 统计分析
  - 机器学习

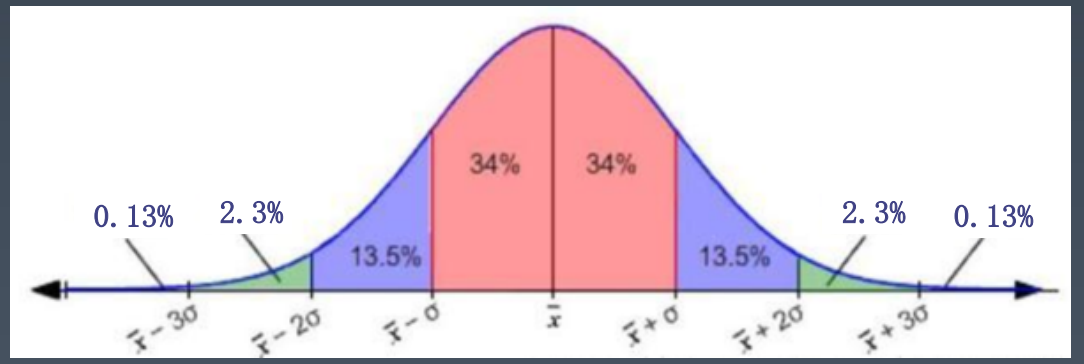


演讲大纲

- **黄金指标异常检测技术**
  - 异常检测问题及难点
  - 运维黄金指标监控方法
    - **响应时间**
    - 错误数
    - 请求数
- 百度AIOps产品及效果
  - 产品形态
  - 产品使用效果
  - 百度内外部落地效果

传统监控方法

- 判断响应时间是否过高的算法
  - 历史数据
    - $\{x_t | t = 1 \ldots n\}$
  - 计算样本均值和样本标准差
    - $\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}, s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$
  - 异常检测
    - $x_j > \bar{x} + 3s$
- 统计学视角
  - 假设响应时间服从正态分布
    - $x_t \sim \mathcal{N}(\mu, \sigma^2)$
  - 参数估计方法
    - $\mu = \bar{x}, \sigma = s$
  - 计算概率
    - $P\{x_j > \mu + 3\sigma\} \approx 0.13\% < p_0$

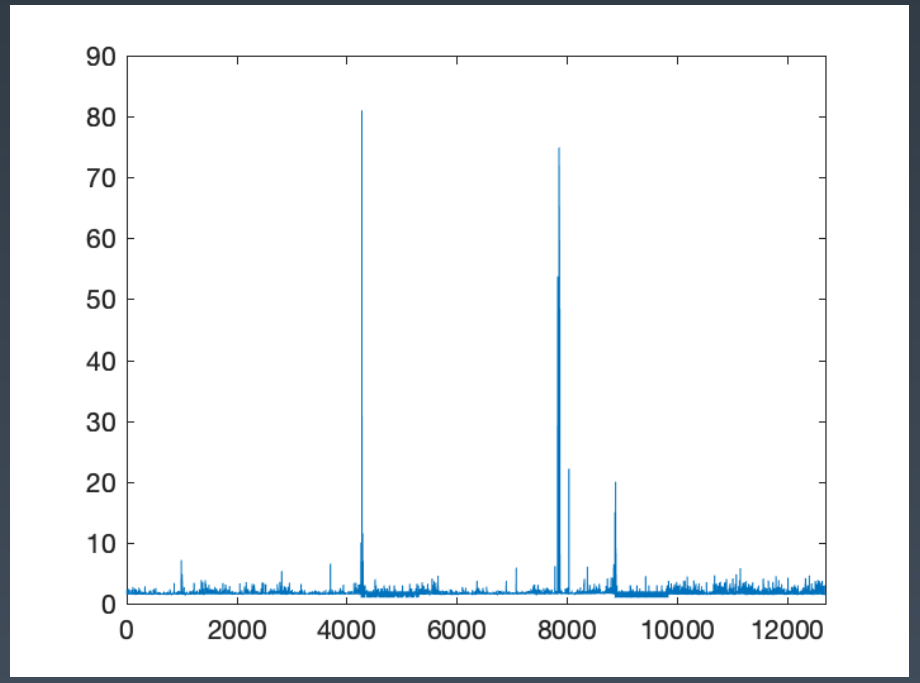


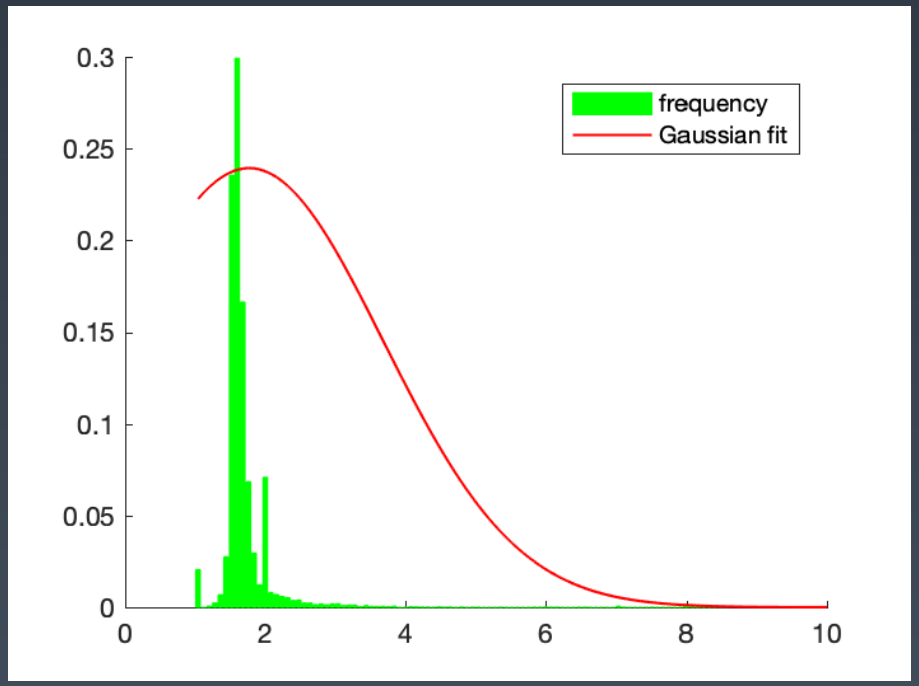
统计学建模三步法

- 统计学建模三步法
  - 概率分布假设
    - 假设数据x的取值服从某个概率分布
  - 估计分布模型
    - 使用历史数据$\{x_i\}$估计
  - 概率计算
    - 概率小于概率阈值$P\{x > t\} < p_0$时即异常

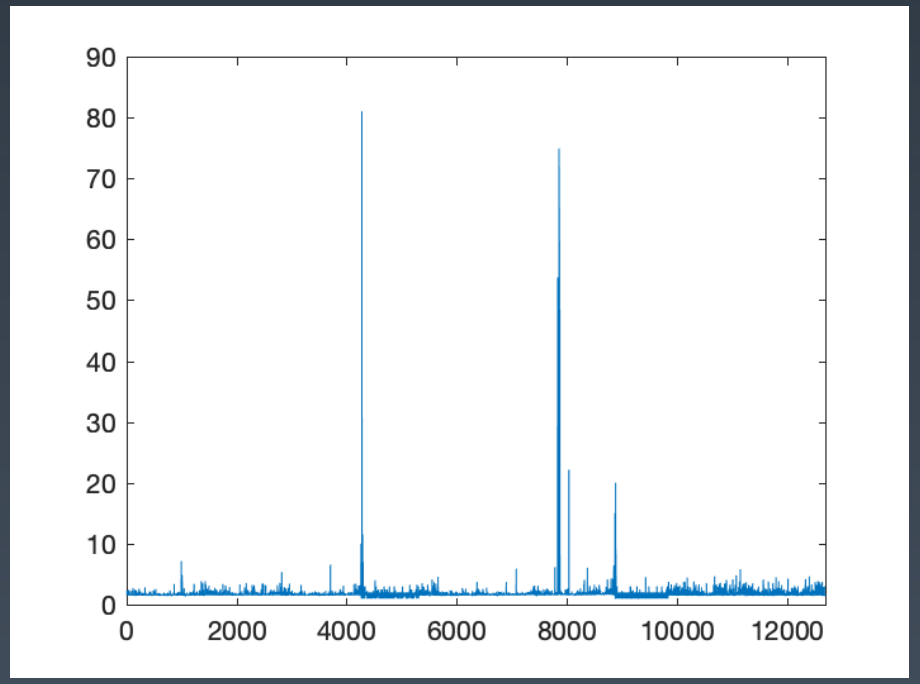
- 为什么在概率维度上设定阈值？
  - 与工程师的认知一致
  - 常数，与曲线性质、时间无关
  - 可以周期性重建分布模型，自动调节数据阈值t

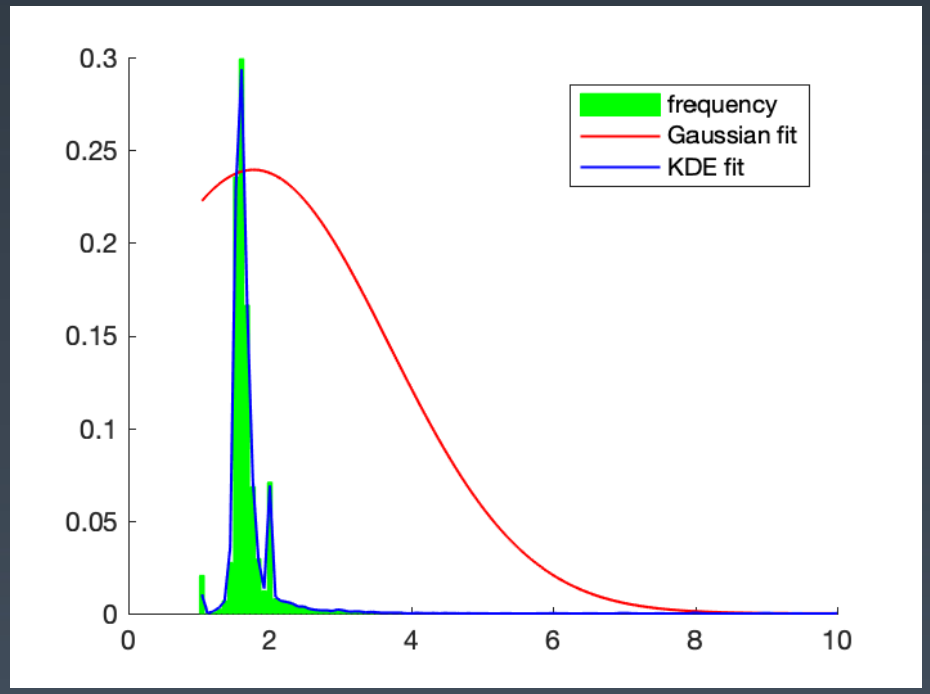
概率分布假设：响应时间服从正态分布吗？





估计分布模型：核密度估计





$$f(x) = \frac{1}{n}\sum_{i=1}^{n} K(x; x_i)$$

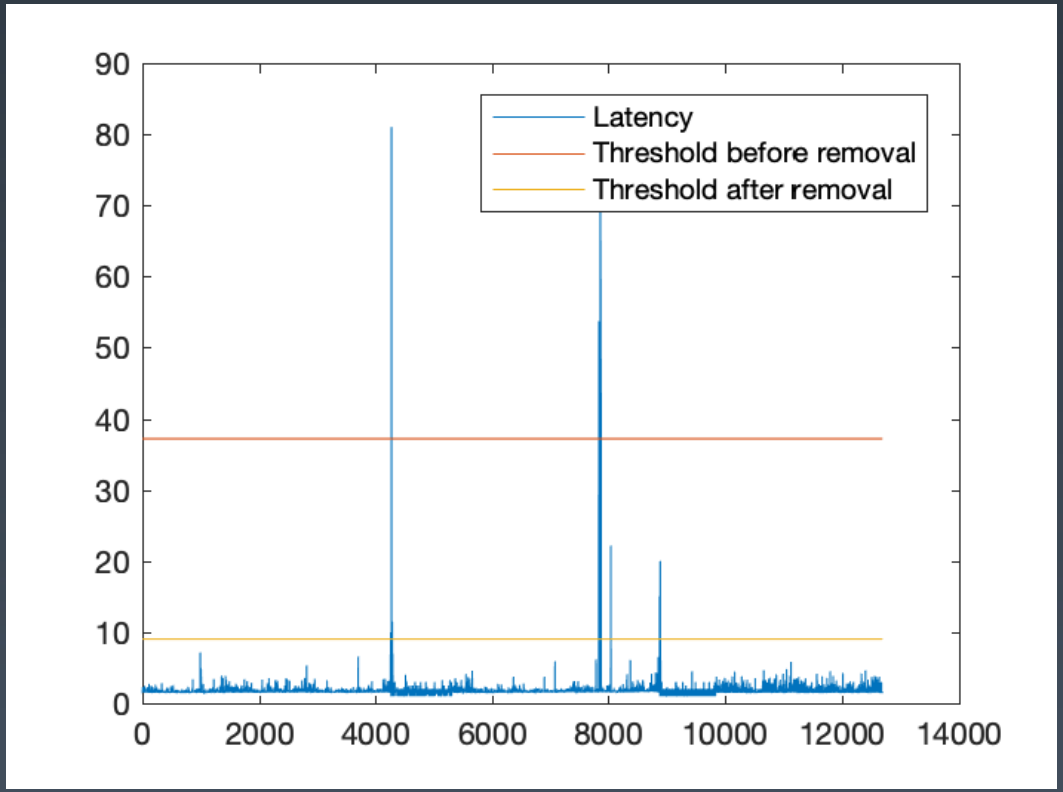
$$K(x; x_i) = \mathcal{N}(\mu = x_i, \sigma \approx 1.06sn^{-\frac{1}{5}})$$

概率计算

- 概率计算

$$P(x_j > t) = \int_{t}^{+\infty} f(x)dx$$

- 番外：历史异常数据的剔除算法
  - 数据聚类切分
  - 切分效果评估（LDA散度）

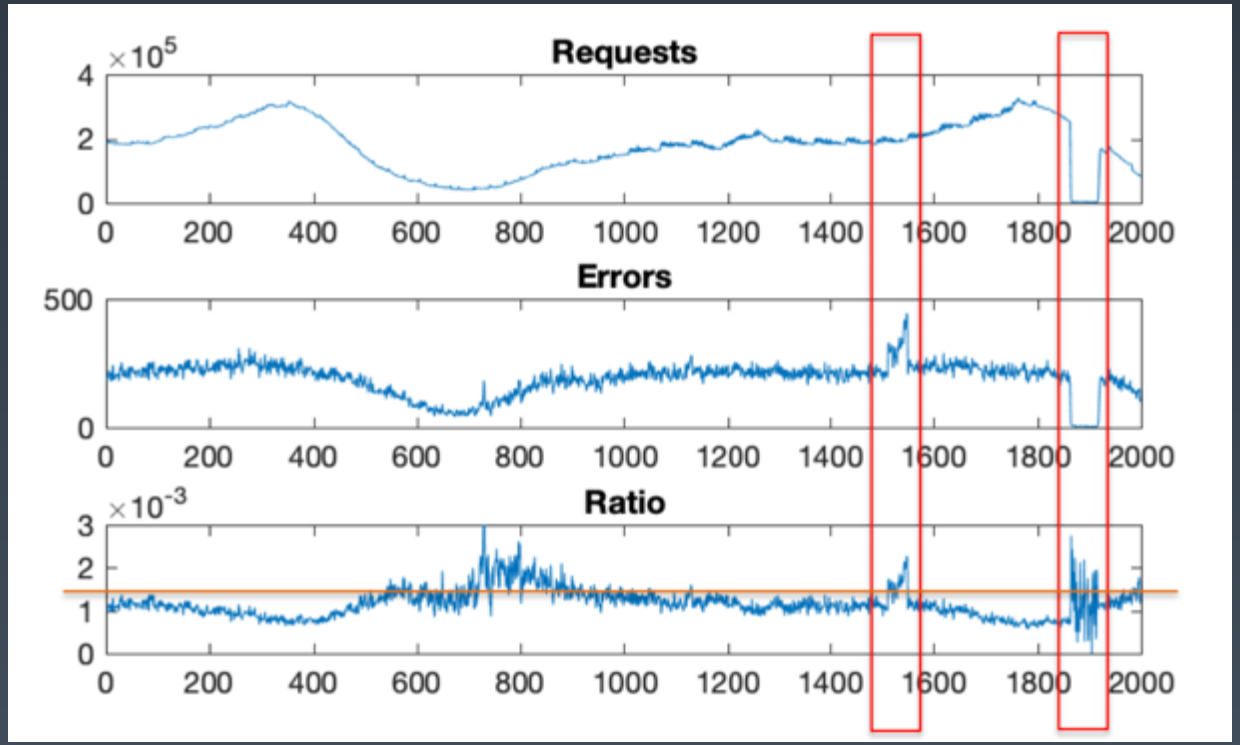


演讲大纲

- **黄金指标异常检测技术**
  - 异常检测问题及难点
  - 运维黄金指标监控方法
    - 响应时间
    - **错误数**
    - 请求数
- 百度AIOps产品及效果
  - 产品形态
  - 产品使用效果
  - 百度内外部落地效果

错误数监控

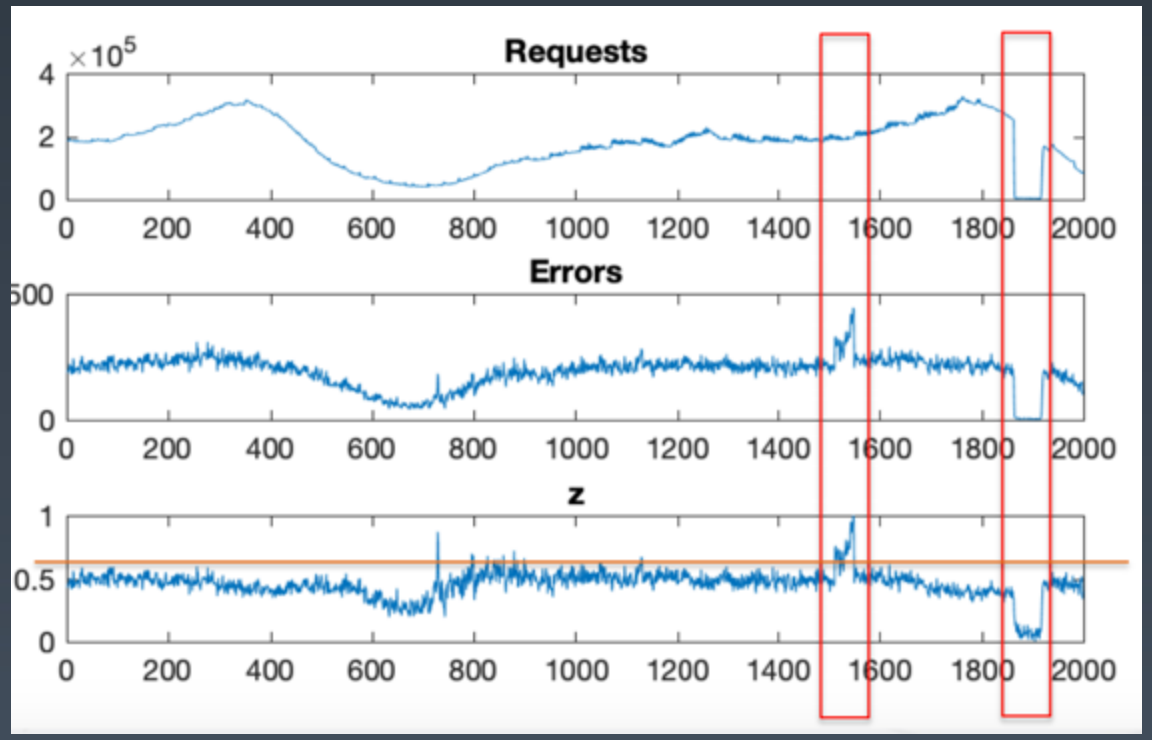
- 错误数监控
  - 请求数$\{n_i\}, i = 1,2, \dots, k$
  - 错误数$\{x_i\}, i = 1,2, \dots, k$
  - 判断当前的x_i是否过大
- 基于错误率的监控
  - 错误率$\left\{r_i = \frac{x_i}{n_i}\right\}, i = 1,2, \dots, k$
  - 请求数较小的时候需要调整阈值



错误数监控

- 统计学建模三步法
  - 概率分布假设
    - 二项分布是n个独立的是非试验中成功的次数的离散概率分布

      $P(X = x; n, r_0) = C_n^x r_0^x (1 - r_0)^{n-x}$
  - 估计分布模型：$r_0 = \frac{\sum x_i}{\sum n_i}$
  - 概率计算
    - 当n足够大时，近似计算
      - $\mu = nr_0, \sigma = \sqrt{nr_0(1 - r_0)}$
      - $z = \frac{x-\mu}{\sigma} = \frac{x-nr_0}{\sqrt{nr_0(1-r_0)}}$



演讲大纲

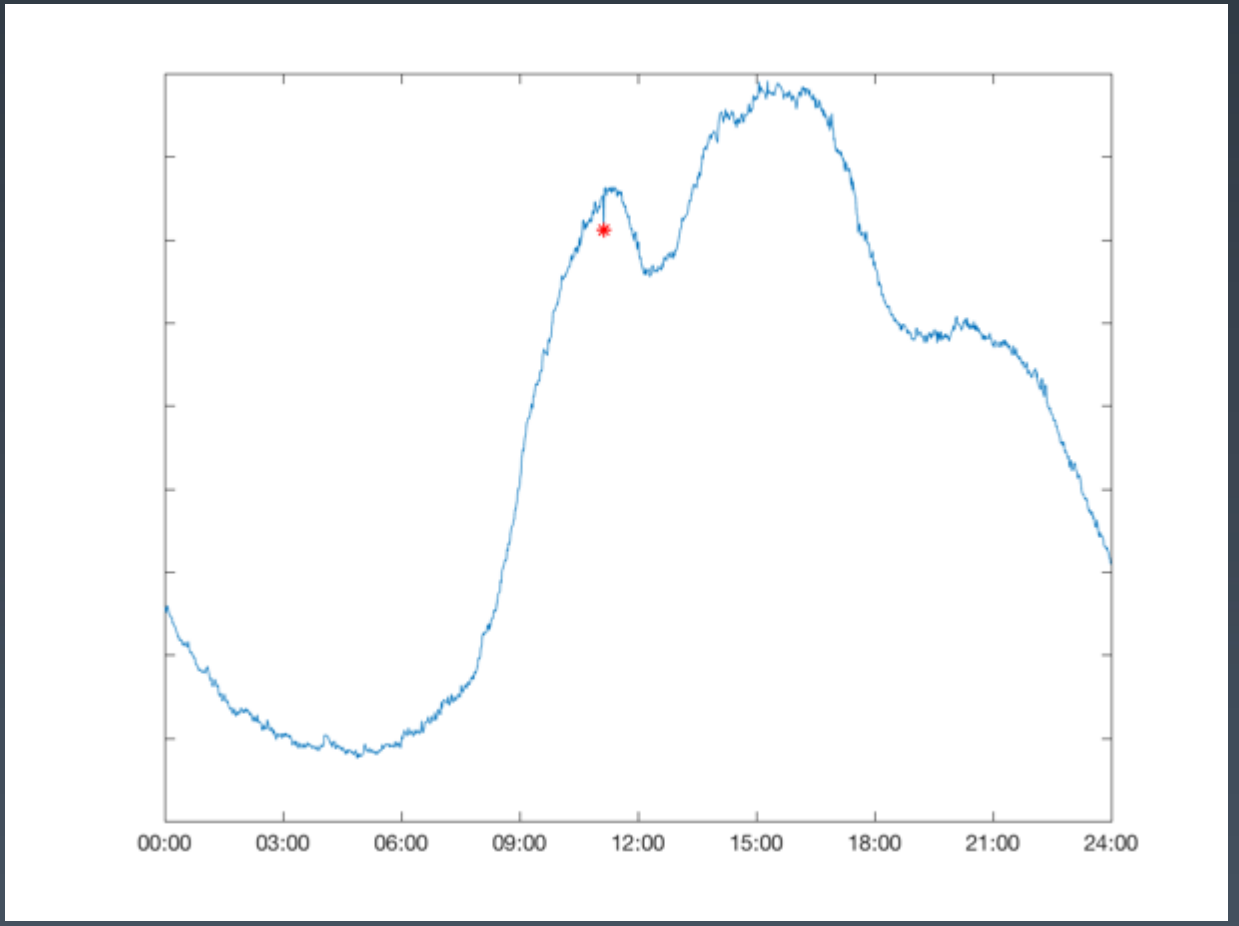
- **黄金指标异常检测技术**
  - 异常检测问题及难点
  - 运维黄金指标监控方法
    - 响应时间
    - 错误数
    - **请求数**
- 百度AIOps产品及效果
  - 产品形态
  - 产品使用效果
  - 百度内外部落地效果

请求数

- 统计学建模三步法
  - 概率分布假设
    - 泊松分布描述单位时间内随机事件发生的次数的概率分布

$$P(Y = k; \lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!}$$

  - 估计分布模型
    - 方案：基线预测算法
  - 概率计算

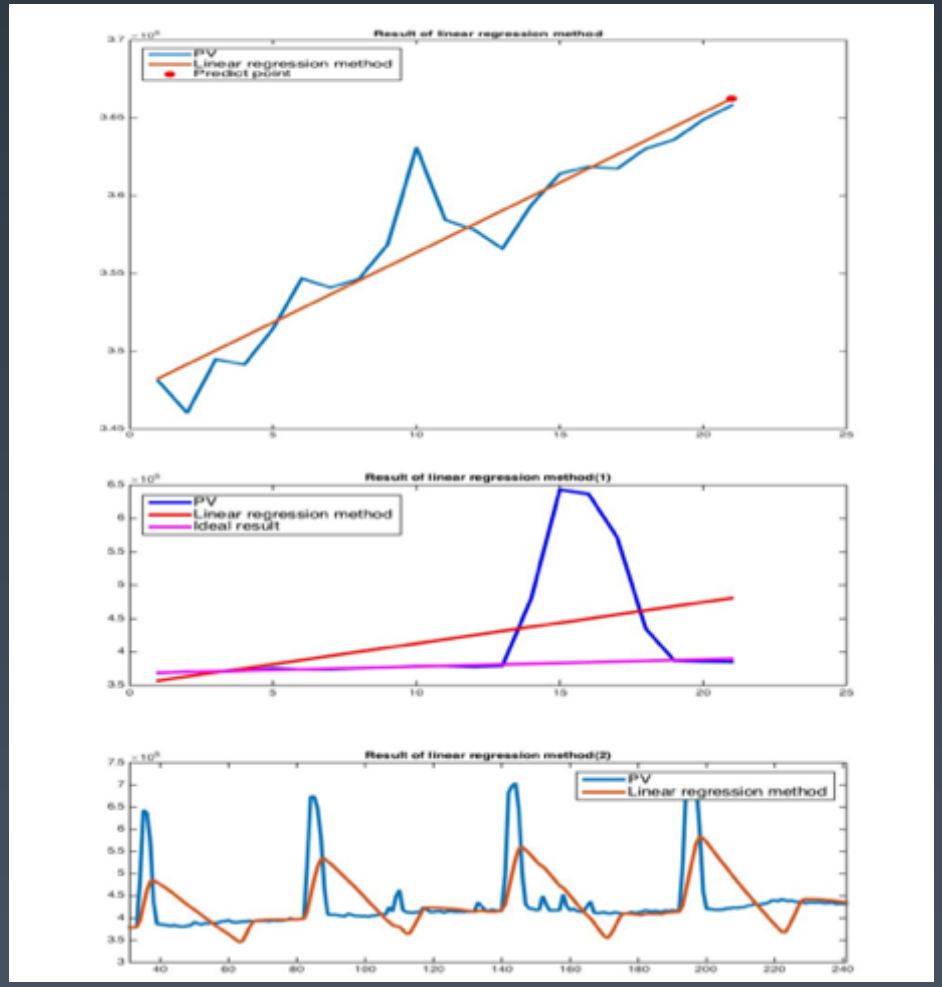


基线预测算法

- 移动平均
  - 窗口内权重相同
- 指数平滑
  - 近期数据权重更大
- 鲁棒回归
  - 假设较小窗口内符合线性趋势变化
- 变分自编码器（VAE）
  - 假设较大窗口内服从非线性趋势变化
- Holt-winters
  - 线性趋势+单一周期模式
- 周期数据多模式挖掘
  - 多种不同的周期模式

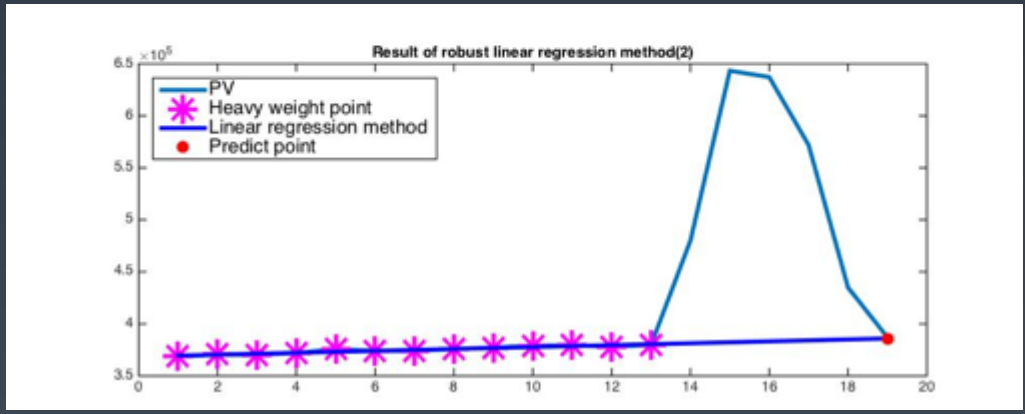
线性回归算法

- 线性回归算法原理
  - 局部符合线性
    - $y_t \approx kt + b$
  - 预测
    - $\hat{y}_t = kt + b$
  - 参数计算
    - $k, b = \underset{k,b}{argmin}\, L$
  - 最小二乘法（Least square）
    - $L_2 = \sum_{\tau=1}^{t}(y_\tau - \hat{y}_\tau)^2$
- 线性回归算法问题
  - 易受噪声点影响，产生误报

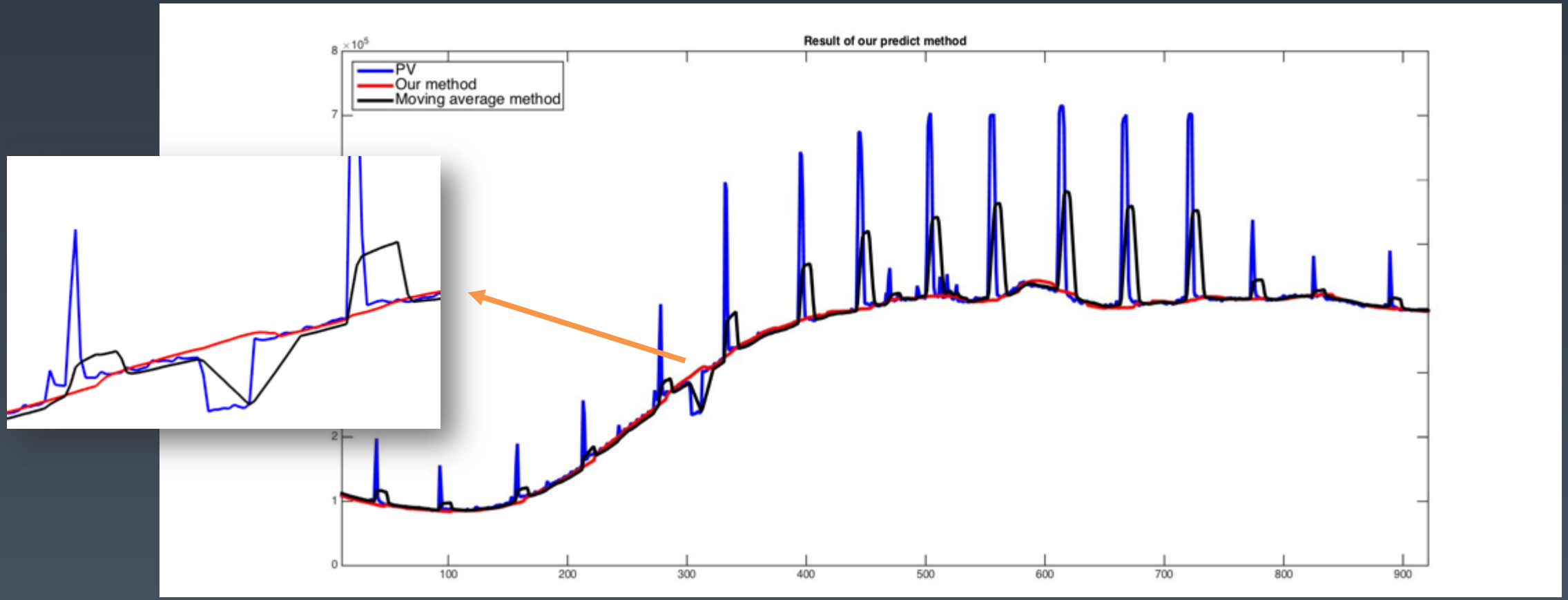


基于鲁棒回归的线性算法

- 鲁棒回归算法原理
  - 最小绝对误差（Least absolute deviations）
    - $L_1 = \sum_{\tau=1}^{t} |y_\tau - \hat{y}_\tau|$
  - 加权迭代最小二乘法（Iteratively re-weighted least squares）
    - $L' = \sum_{\tau=1}^{t} \omega_\tau (y_\tau - \hat{y}_\tau)^2$
    - $\omega_\tau = \frac{1}{|y_\tau - \hat{y}_\tau|}$

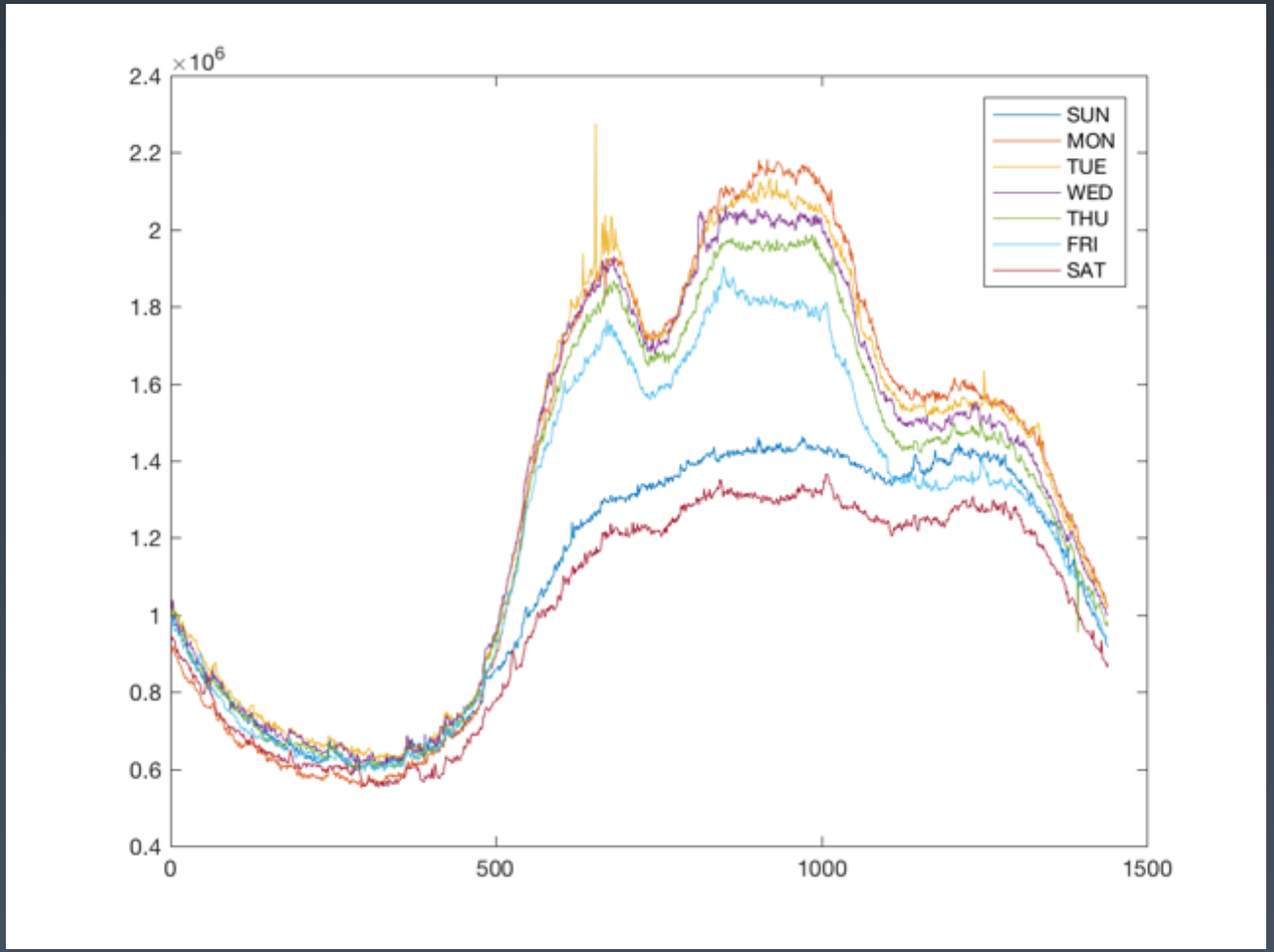


鲁棒回归的预测效果


Result of our predict method
×10^5
PV
Our method
Moving average method
0
1
2
7
8
100
200
300
400
500
600
700
800
900

周期数据多模式挖掘

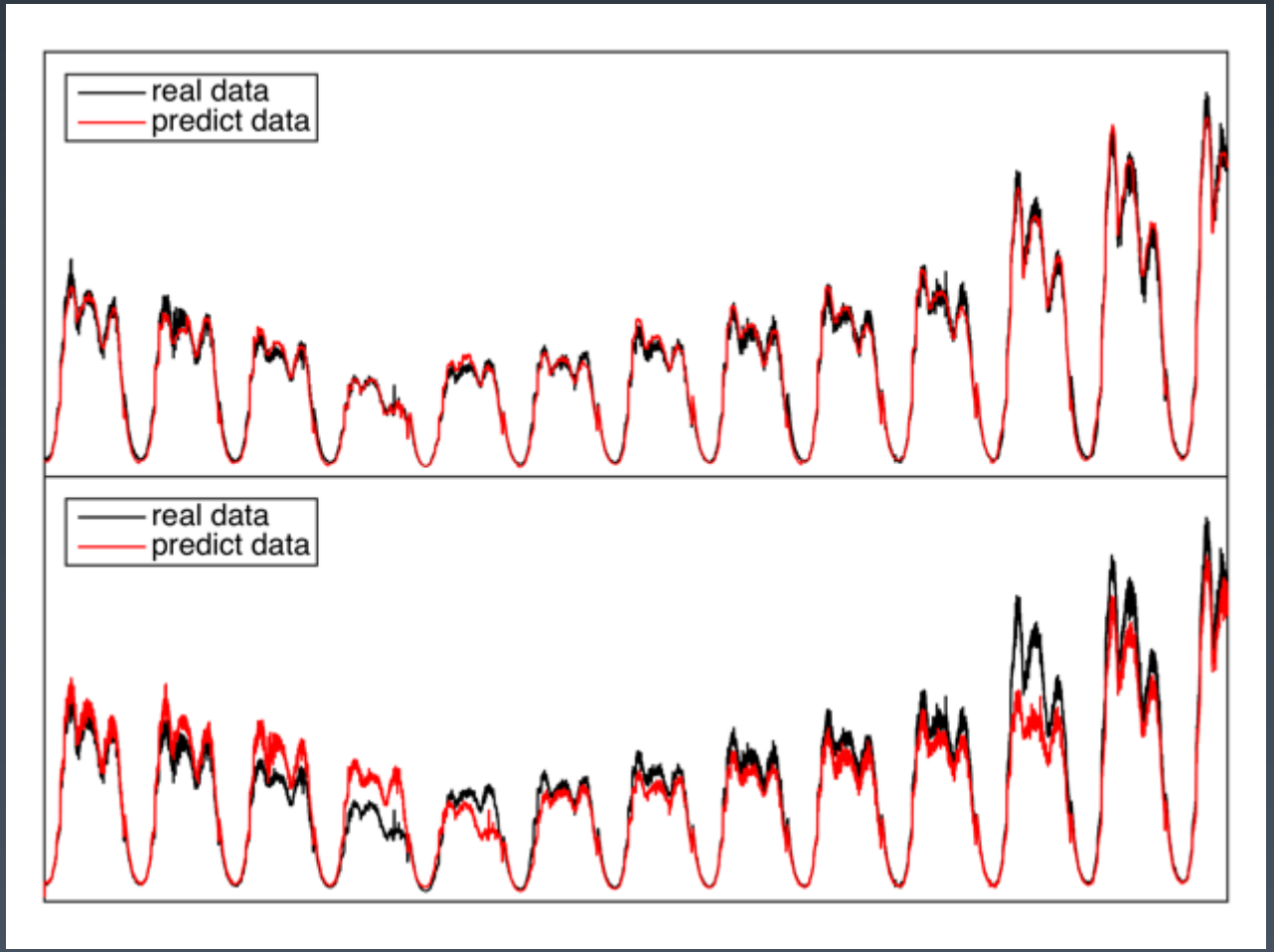
- 难点
  - 长时间缓慢下跌
  - 多种不同的周期
    - 工作日、休息日和假期
  - 水位漂移
- 想法
  - 形状模式提取
  - 自适应水位



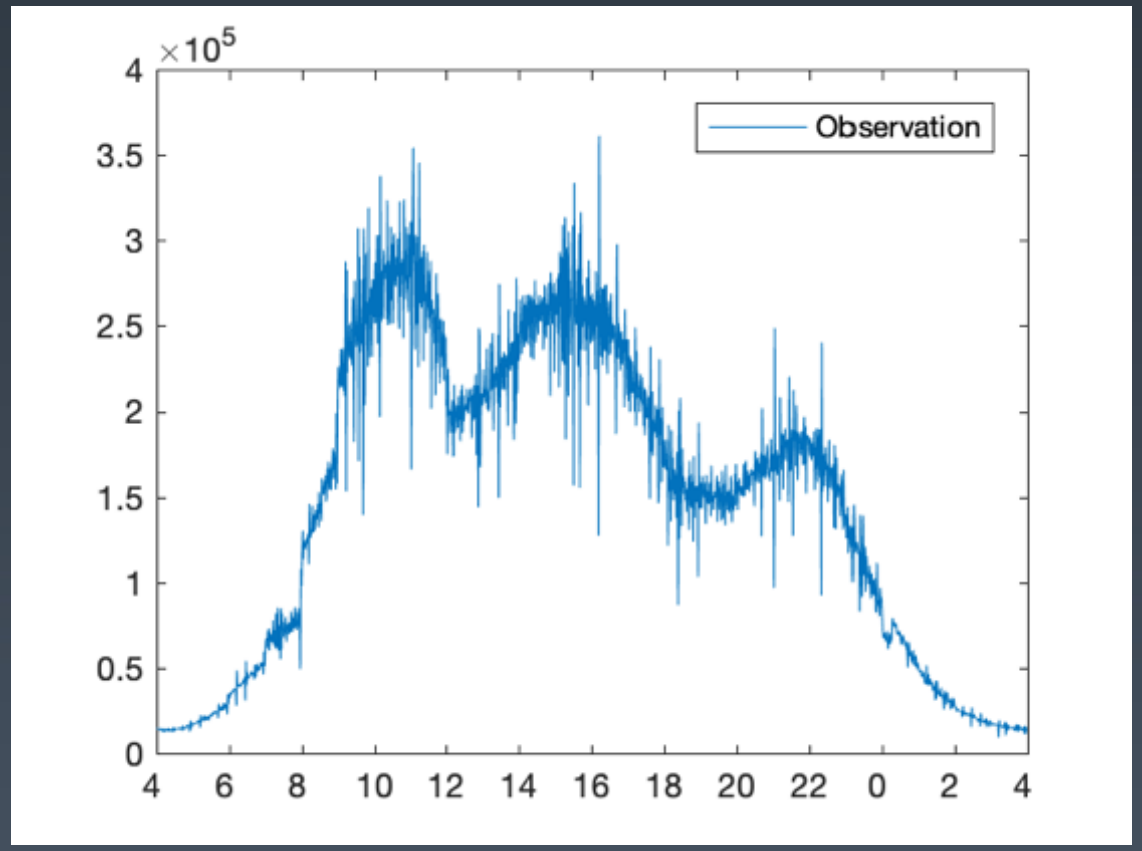
预测效果

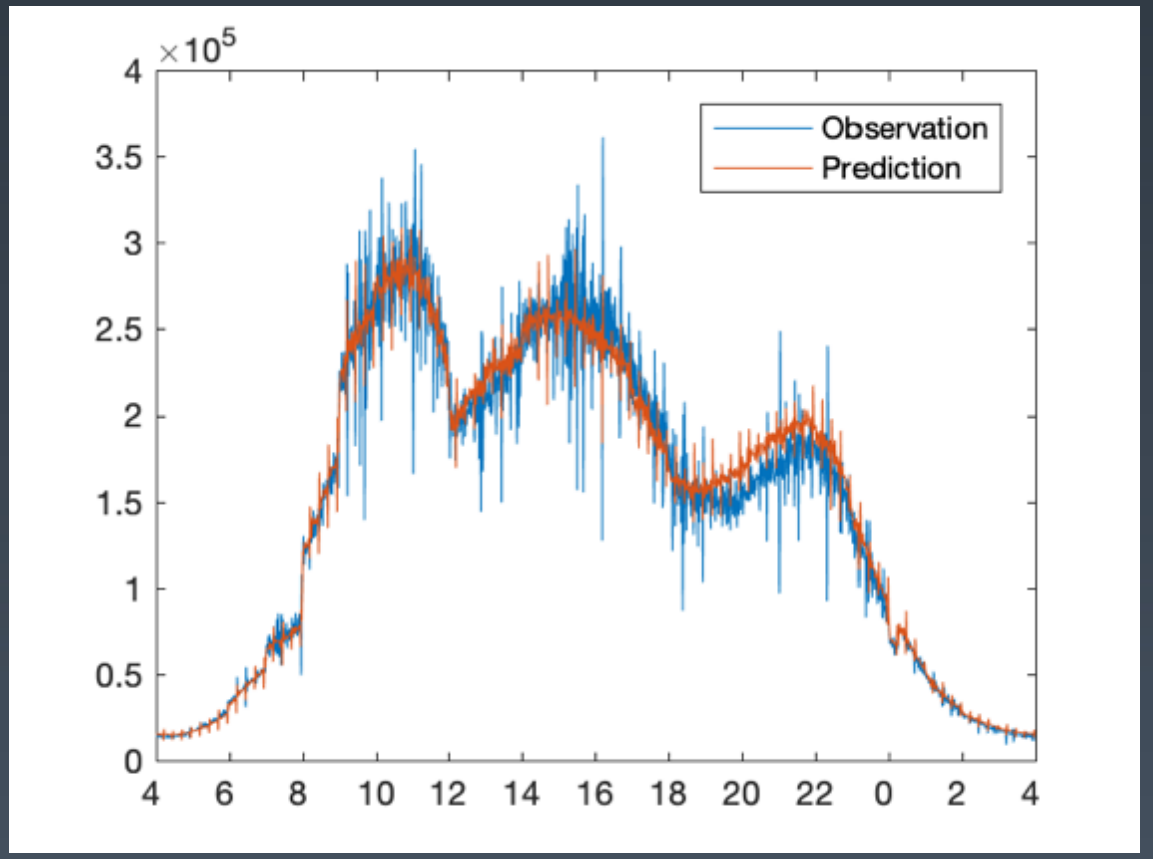
周期数据多模式挖掘的算法效果

简单同比的算法效果



预测效果


×10^5
4
3.5
3
2.5
2
1.5
1
0.5
0
4 6 8 10 12 14 16 18 20 22 0 2 4
Observation

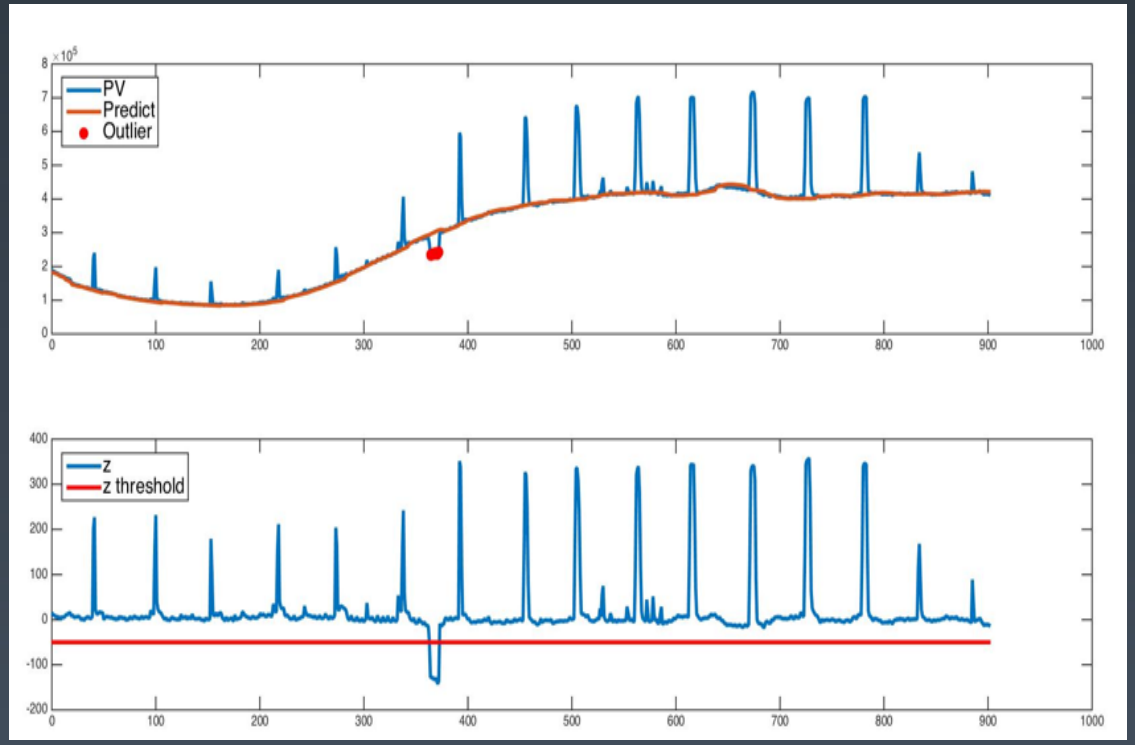

×10^5
4
3.5
3
2.5
2
1.5
1
0.5
0
4 6 8 10 12 14 16 18 20 22 0 2 4
Observation
Prediction

概率计算

- 概率计算
  - 泊松分布描述单位时间内随机事件发生的次数的概率分布

$$P(Y = k; \lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$

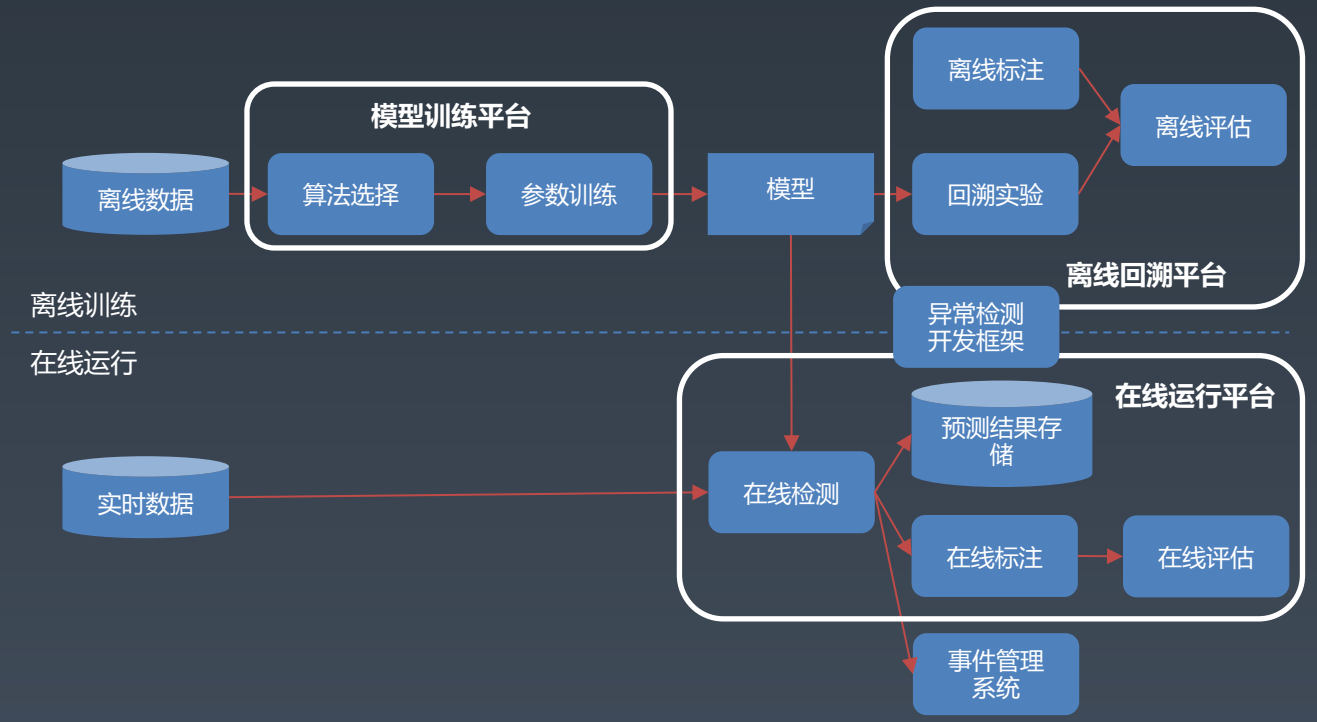
  - 当流量值较大时，泊松分布近似等于正态分布: $\mathcal{N}(y_t; \mu, \sigma^2)$
    - $\mu = \sigma^2 = \lambda = \hat{y}_t$
    - $y_t < C \Leftrightarrow y_t < \hat{y}_t - m\sigma$
    - $z = \frac{y_t - \hat{y}_t}{\sqrt{\hat{y}_t}} < -m$



演讲大纲

- 黄金指标异常检测技术
  - 异常检测问题及难点
  - 运维黄金指标监控方法
    - 响应时间
    - 错误数
    - 请求数
- **百度AIOps产品及效果**
  - 产品形态
  - 产品使用效果
  - 百度内外部落地效果

百度AIOps产品



- AIOps 平台产品（独立输出）
  - 模型训练平台
  - 标注平台
  - 在线运行平台

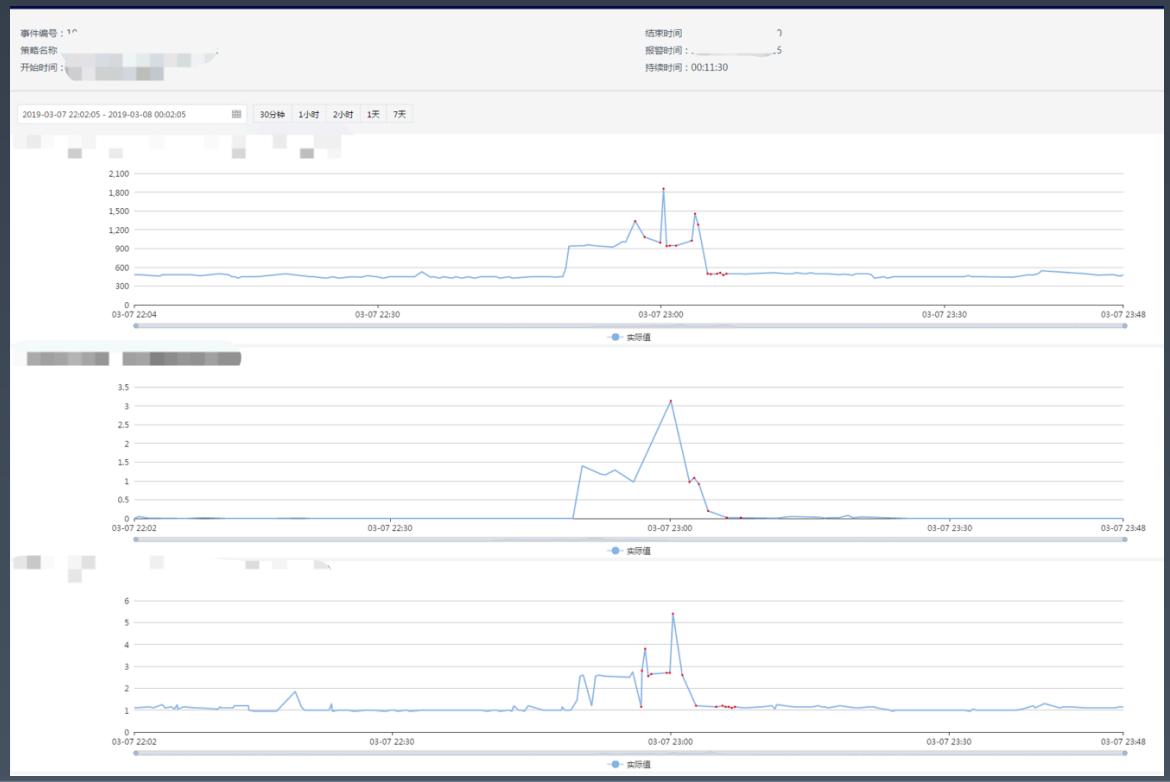
- AIOps SDK库产品（嵌入已有的监控平台）
  - 离线训练工具
  - 标注平台
  - 异常检测SDK
  - 指标排查SDK

智能异常检测配置

- 无需人工配置参数
- 自动排除历史故障数据干扰
- 自适应忙闲时等阈值变化
- 可查看历史回溯效果
- 防抖动策略

智能辅助业务故障诊断

- 实时异常检测，报警时效性<2s
- 自动指标排查

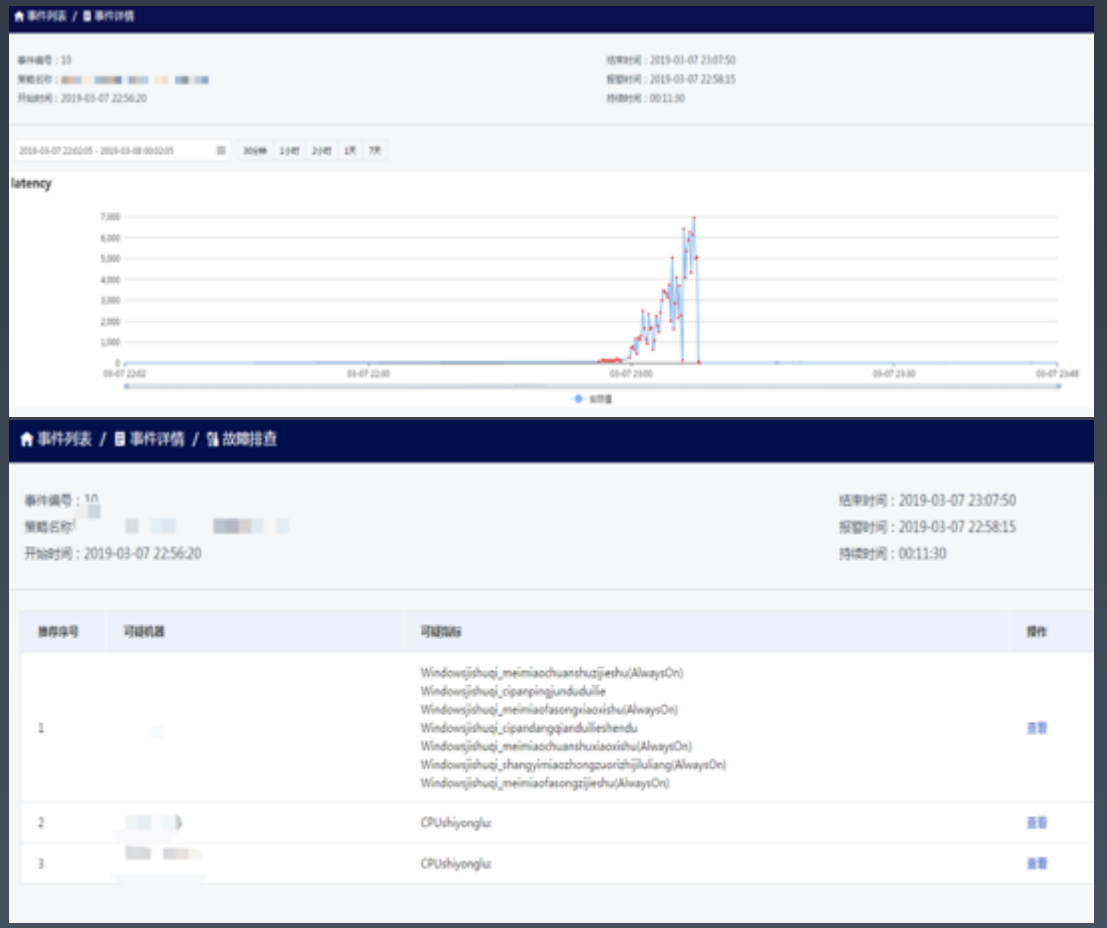


百度内部落地所有业务线

- 业务指标监控
  - 监控指标数量：百万级监控指标
  - 配置成本：SRE 0成本配置
  - 准确率>85%，召回率>99%
  - 报警时效性：2s
- 黄金指标排查
  - 核心业务涉及上百个接口，涉及指标近千个
  - 平均排查时效性：10s
  - TOP5召回率>90%
  - 平均定位时间从30+min下降到10min

ToB产品效果（以证券行业客户为例）

- 核心功能业务监控
  - 监控范围：几十个核心功能
  - 覆盖指标：交易量、交易延迟、交易响应率、交易成功率等
- 机器指标排查
  - 覆盖了千余个机器指标
  - 排查时效性15s
  - TOP5召回率>99%



Geekbang 极客邦科技 | InfoQ 极客传媒

InfoQ官网 全新改版上线

促进软件开发领域知识与创新的传播

关注InfoQ网站
第一时间浏览原创IT新闻资讯

免费下载迷你书
阅读一线开发者的技术干货

THANKS! | QCon 10th