中国移动 China Mobile

5G++创新实训基地

技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器

版权声明

中国移动 China Mobile | 5G+创新实训基地

技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器

Large-scale Anomaly Detection for Core Network Data Center

大规模异常检测在核心网数据中心的探索与实践

2020年11月

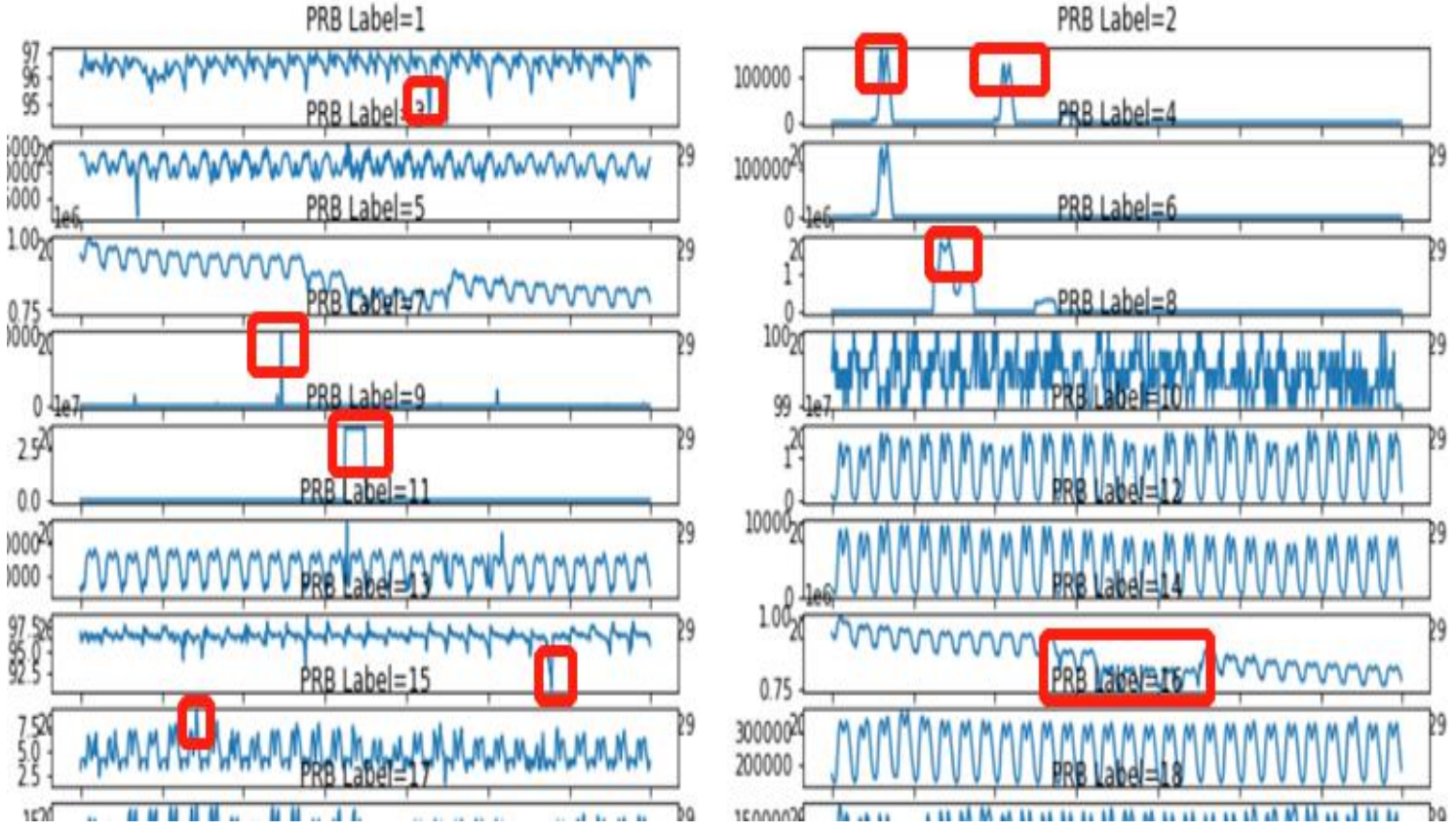
诺基亚-廖文哲

成果落地

目前，LKCED系统已经在浙江移动核心网运行部成功落地上线，对核心网的运维效率有了极大的提升，并获得2020年度浙江移动SRE优秀实践项目一等奖表彰，见公众号：https://mp.weixin.qq.com/s/rNfTsBwkObugx2lmubY6Zw

一、成果背景

- 异常检测需要监控的指标繁多（50万左右），覆盖了机器性能，业务用户数，率等众多指标检测。而**利用最少的人为参与同时及时准确发现这些指标数据的异常波动**，是业务稳定性的重要保证。
- 但是这些数据不但数量众多，而且不同业务的曲线也有截然不同的特征：



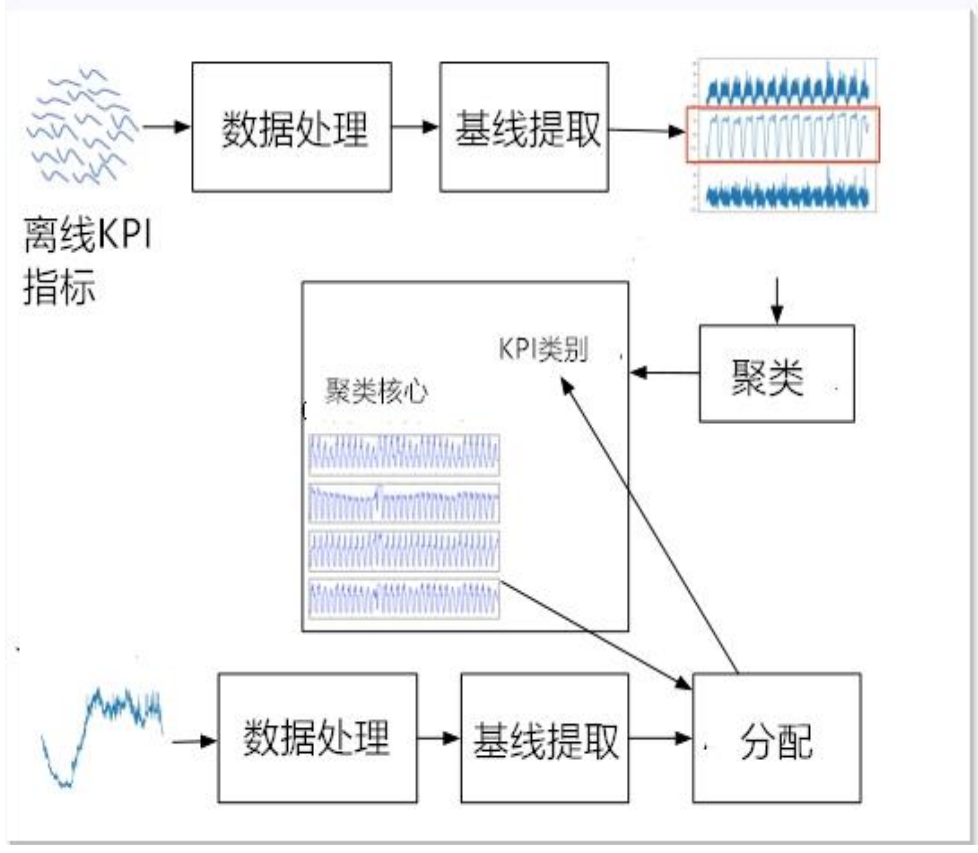
Contents

二、成果主要内容-大规模KPI分类

- 由于KPI数量众多，且形状各异，故先对海量KPI数据进行分类，包括离线和在线2个模块：

- 1.离线模块利用时间序列基线提取，数据标准化，聚类输出各种类别的时间序列模型。
- 2.在线模块利用离线模块进行实时类别分类，然后根据类别进行对应告警输出。

离线模块-线下训练



在线模块-线上运行

二、成果主要内容-KPI分类离线模块

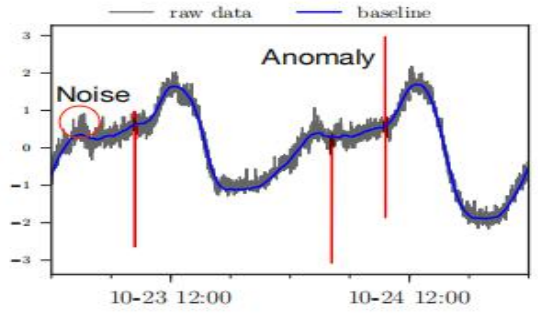
□ KPI分类离线模块技术方案：

1.预处理：包括数据填充，异常值剔除，数据标准化。

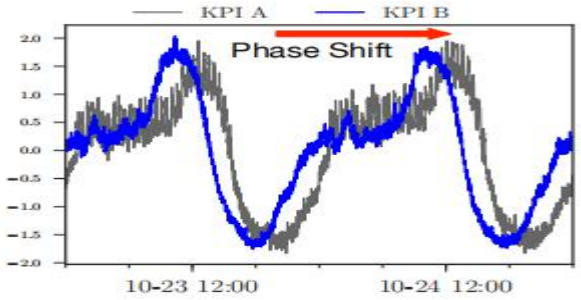
2.基线提取：消除噪音，提取一个粗略的基线来表示关键绩效指标的基本结构。

3.聚类：基于形状相似性对采样的关键绩效指标的基线执行基于密度的聚类。

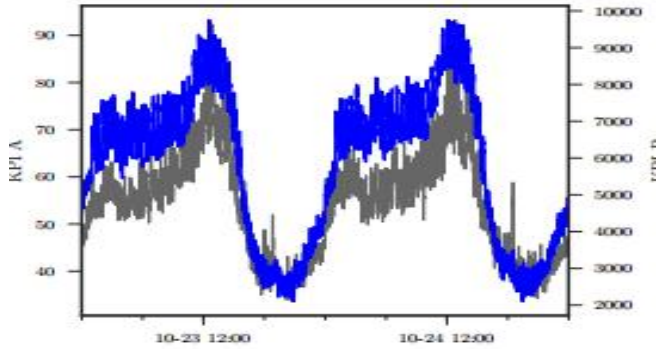
挑战：噪声，异常值，相位偏移，幅度变化.



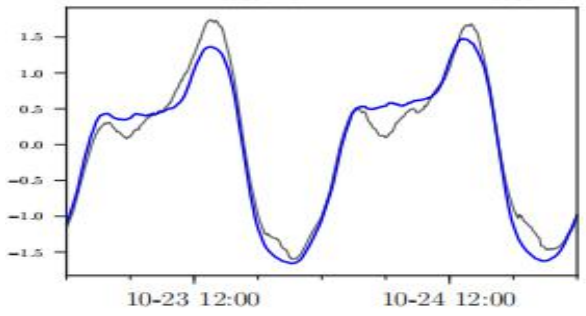
噪音和异常点



相位漂移



幅度变化1



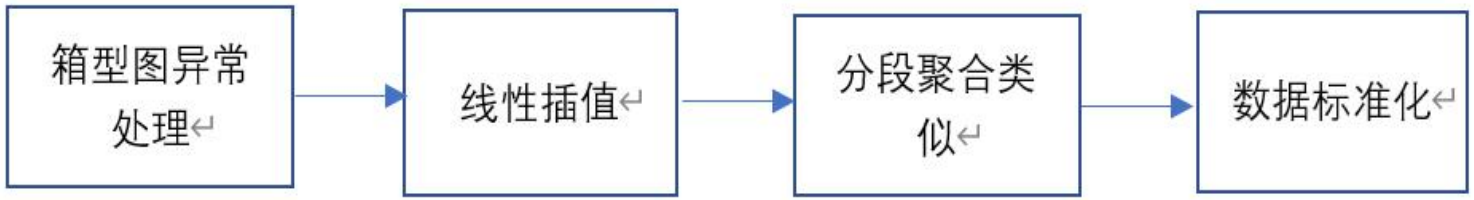
幅度变化2

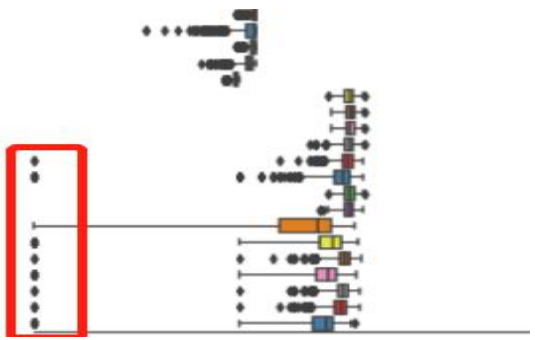
二、成果主要内容-KPI分类离线模块

1.数据预处理:

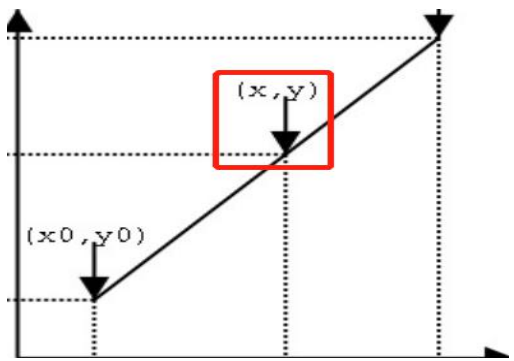
1.由于数据缺失并不严重，脏数据也不多，在利用箱型图IQR消除异常值后，对原先的缺失值，异常值剔除后的值用线性差值法插入。

2.对2个月指标分段聚合类似/平均压缩为1周数据，由于不同指标上下限不一致，将指标都转化为均值为0，方差为1的数据区间，**以消除幅度变化影响。**





箱型图异常处理



线性插值

$$q_i' = \frac{1}{k}\sum_{j=k(i-1)+1}^{k*i} q_j,\quad i = 1,2,\ldots,w$$

分段聚合类似

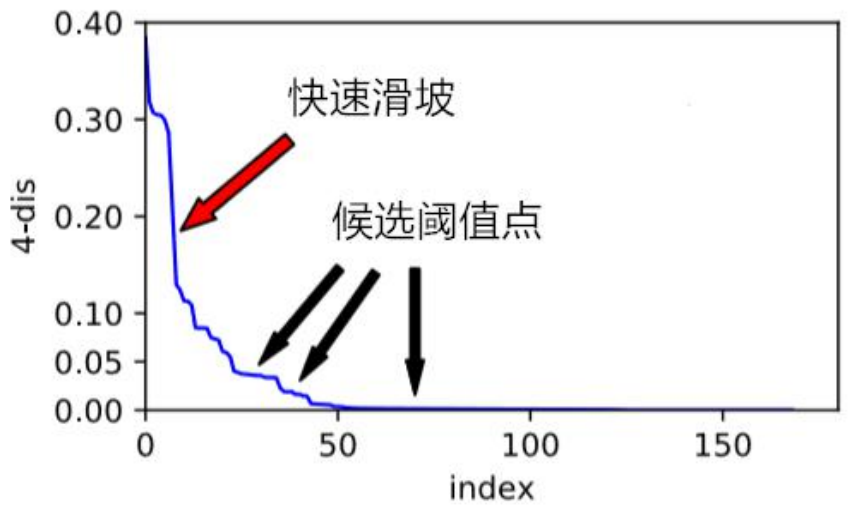
$$x^* = (x - \mu) / \sigma$$

数据标准化

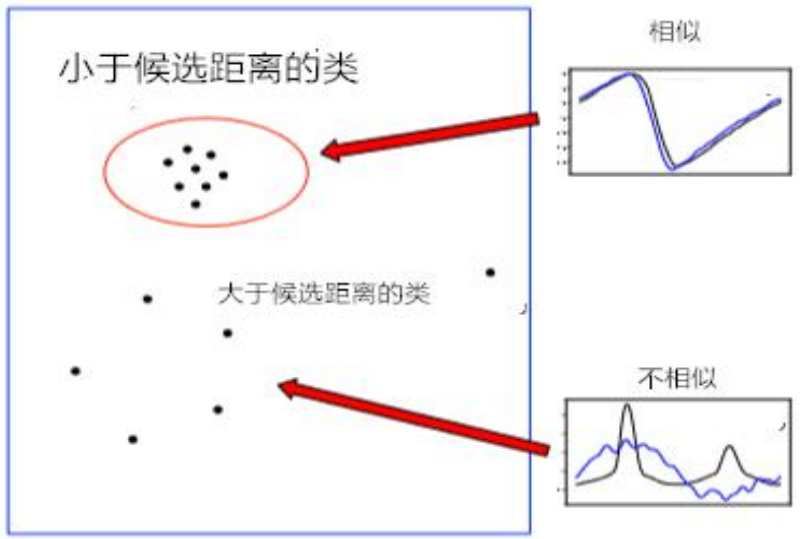
3.密度聚类：

1.曼哈顿距离/NCC-SBD/DTW距离度量：多维空间数学证明，曼哈顿距离比传统欧式距离更适合时间序列高维空间；而NCC-SBD距离度量非常适用于具有相位漂移的时间序列。

2.KNN+DBSCAN密度聚类：由于关键绩效指标是从各种应用程序和系统中收集的，因此很难预先确定集群的数量。基于密度的方法在密集区域形成簇，可以是任意形状和大小。其次我们可以利用形状相似性的传递性来扩展集群。例如，名为a、b、c的三个关键绩效指标衡量同一应用程序使用的机器的性能。a和b的形状相似，b和c的形状也相似。直觉上，a和c的形状也相似。因此，它们可以被分配到同一个集群中。



假设1000个样本，用KNN算出每个样本最近的5个点的距离，候选阈值点表示大部分点的最近5个点距离类似。



利用KNN算出的候选阈值DBSCAN聚类

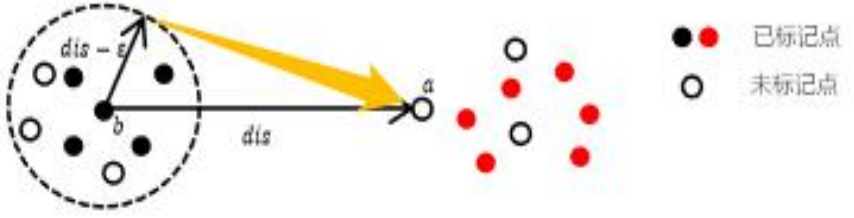
二、成果主要内容-KPI分类在线模块

3.未标记点在线分配：

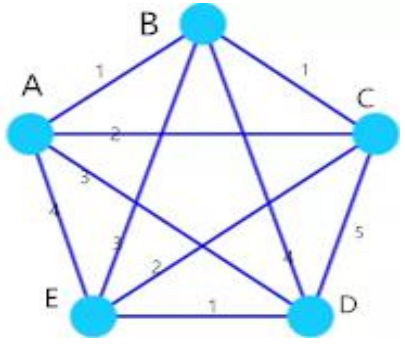
1.在离线模块训练完成后，已经将离线指标分为各个类别，对于未标记类别的指标，在经过数据预处理和基线提取后只要分别计算与各个类别核心点的曼哈顿距离即可判断是否属于该类别。

2.由于指标众多，逐个计算距离需要大量的计算资源和时间，故引入快速排序邻居图（SNG）数据结构和三角不等式定理/质心来替代类别里的所有点，以减少在线计算时间。

方式一：逐个计算距离，利用三角不等式和SNG减少计算量

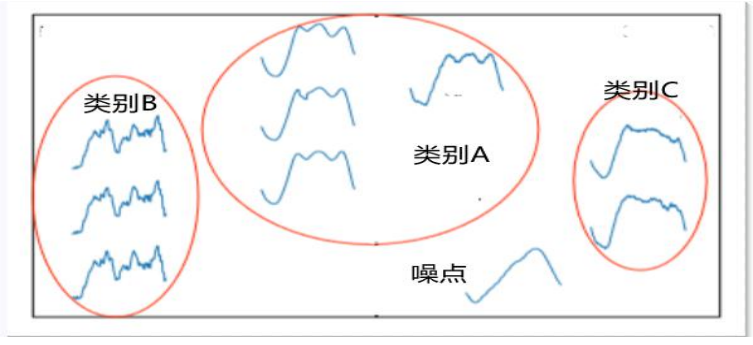


三角形两边之和大于第三边，对小于dis-密度距离阈值内的点可以不需要计算距离。

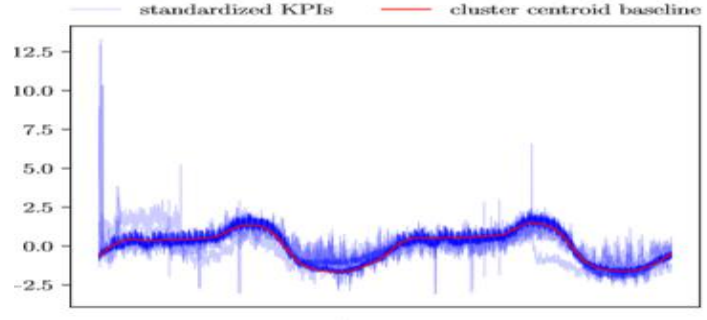


快速排序邻居图示例

方式二：只与质心计算距离



DBSCAN聚类结果



其中一个类别质心

Contents

二、成果主要内容-分类异常检测

利用KPI离线分类后的聚类类别：

1.对于满足时间周期性的指标，分为局部波动大和小类型：

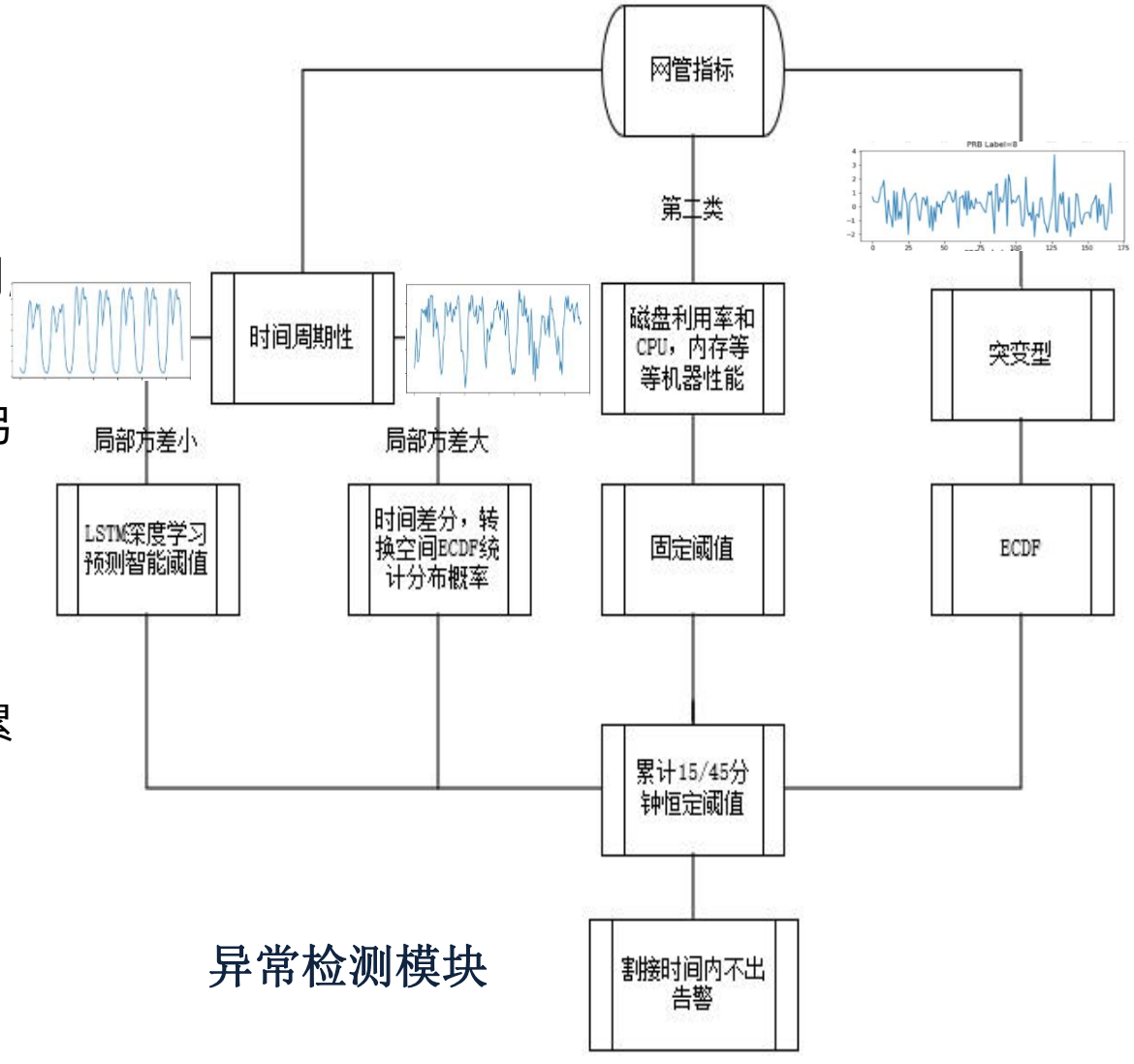
1.1对于局部波动比较小的指标，利用LSTM深度学习预测输出智能阈值。

1.2对于局部波动大的指标，利用时间差分将数据转换到另一个Z空间，再利用ECDF累计分布函数计算上下限。

2.对于磁盘，内存，CPU等指标，利用固定阈值的方式输出告警（例如[10,90]）.

3.对于时间规律不明显，突变类型的指标，直接利用ECDF累计分布函数计算上下限。

最后利用累计15/45/60分钟衡定均值输出告警短信



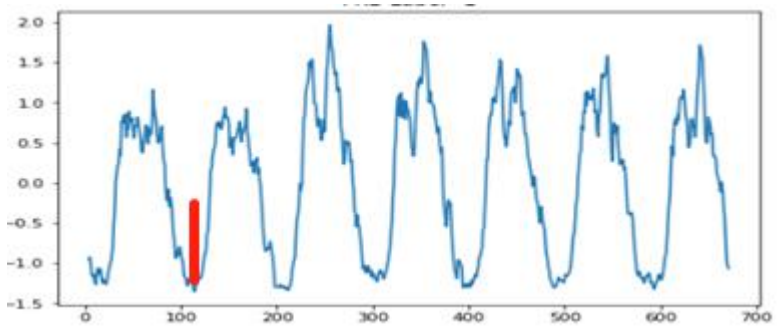
异常检测模块

二、成果主要内容-分类异常检测

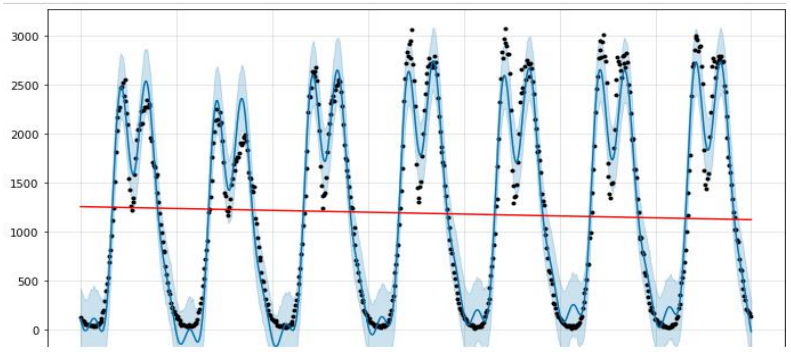
1.对于满足时间周期性的指标，分为局部波动大和小类型：

1.1对于局部波动比较小的指标，利用LSTM/prophet预测，输出智能阈值。

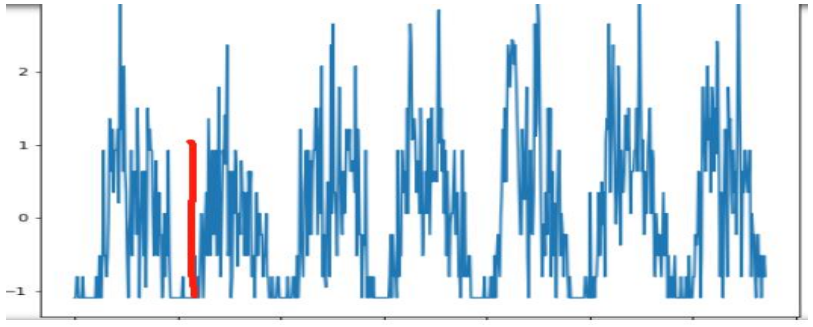
1.2对于局部波动大的指标，利用时间差分将数据转换到另一个Z空间，再利用ECDF累计分布函数计算上下限。



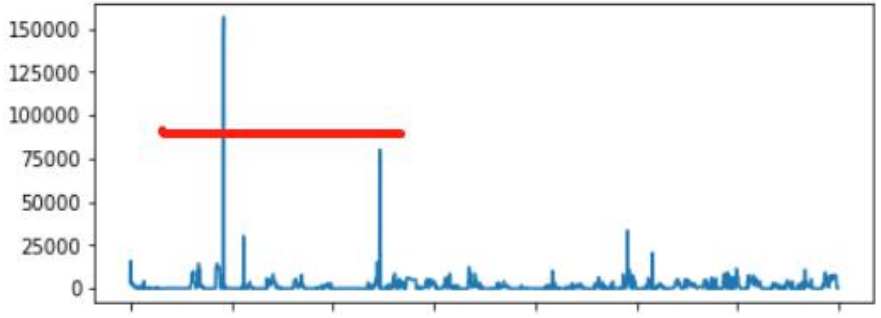
局部波动小的时间序列指标



LSTM深度学习/prophet智能阈值



局部波动大的时间序列指标

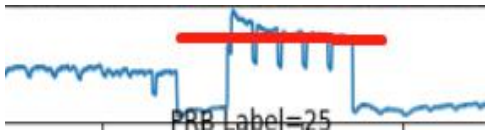


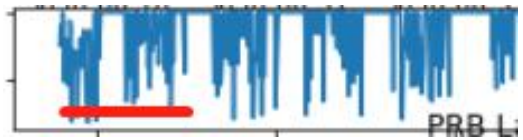
转换空间ECDF统计阈值

二、成果主要内容-分类异常检测

2.对于磁盘，内存，CPU等指标，利用固定阈值的方式输出告警（例如[10,90]）.

3.对于时间规律不明显，突变类型的指标，直接利用ECDF累计分布函数计算智能上下限。最后利用累计15/45/60分钟衡定均值输出告警短信





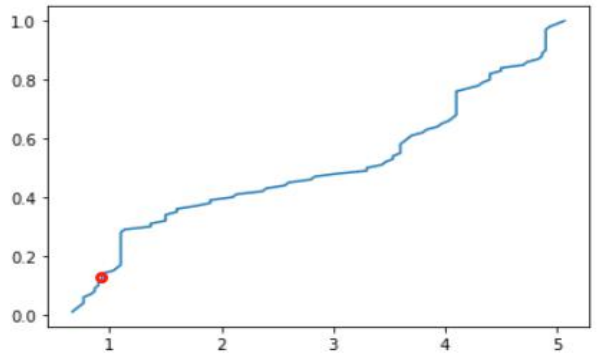
突变类型指标

$$F_X(x) = \mathrm{P}(X \leq x).$$

ECDF累计分布函数公式

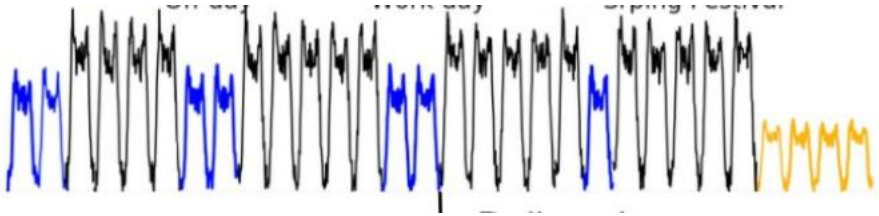
$$s(t) = \frac{x_t + x_{t-1} + \cdots + x_{t-w+1}}{w}$$

衡定均值输出告警短信

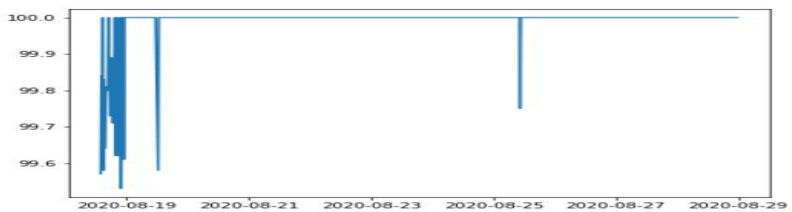


ECDF曲线，纵轴表示概率，横轴表示数值

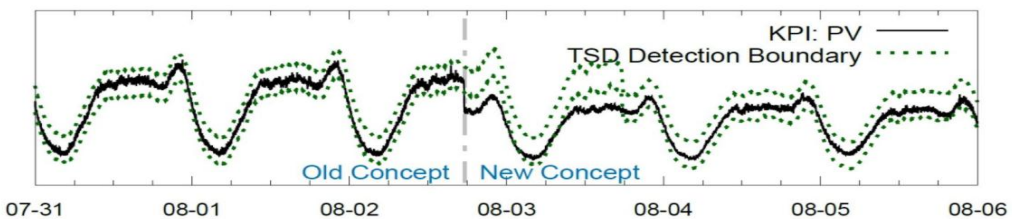
4.其他关注点



节假日与周末不一致



率相关指标用户基数少导致指标快速恶化

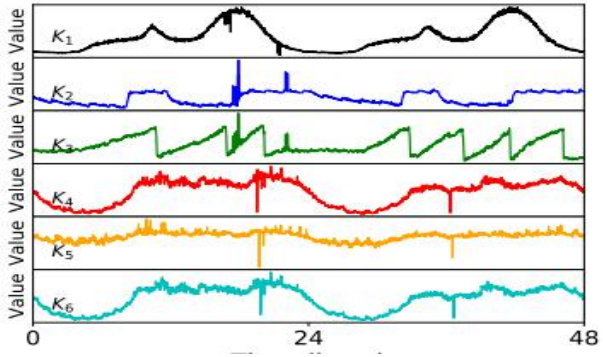


概念漂移

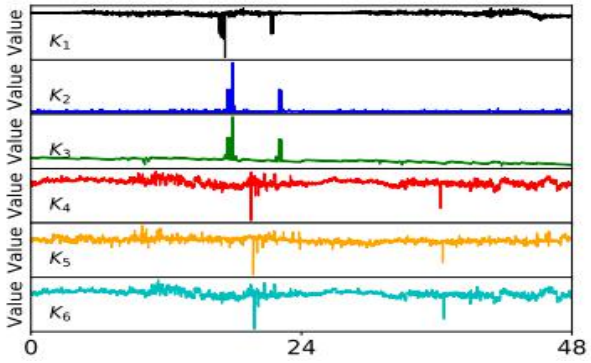
Contents

二、成果主要内容-关联分析与告警收敛

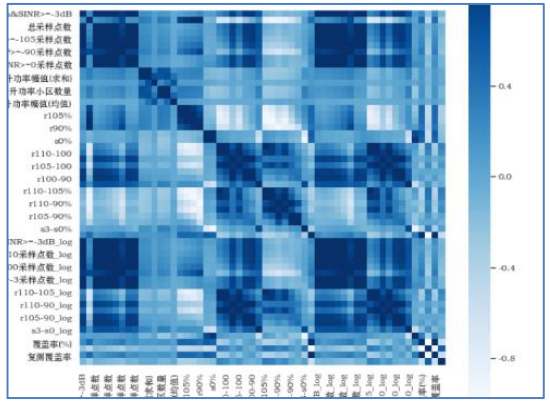
1.关联分析：由于指标众多，许多指标存在内在联系，某个指标的突变经常会造成其他指标的突变，为了找出这些指标，我们可以将指标进行1天或者1周的时间差分，以消除时间变化，然后进行Pearson/Spearman的关联分析，找出关联指标，如下所示，左图为原始指标，右图为时间差分后的指标，时间差分后的指标相关性非常明显。



原始指标



时间差分后的指标



Pearson/Pearman关联分析矩阵图

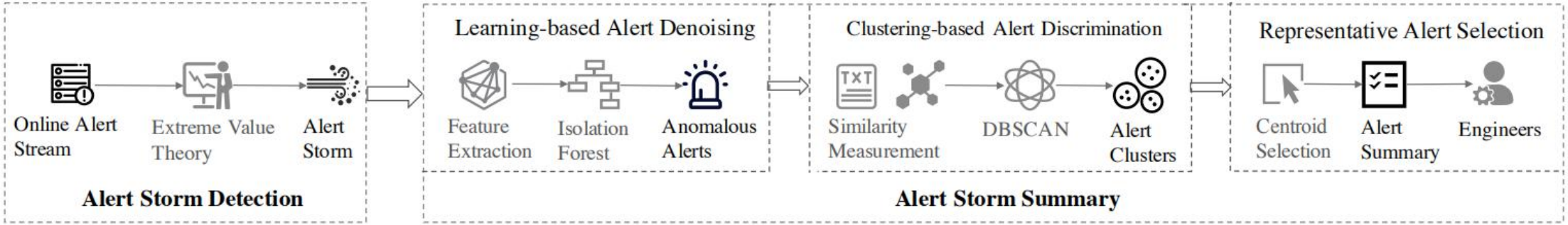
2.告警收敛：由于指标众多，且很多指标存在上下级和各种关系，某一个指标的迅速恶化可能会导致众多指标迅速恶化，并形成告警风暴，对运维人员排除故障会造成极大影响，我们可以利用极值理论/ECDF检测告警风暴，一旦发生告警风暴立即进行告警收敛并推送核心告警。

	网元名称	告警发现时间	告警标题	专业	设备厂家名称	设备类型	网管告警级别	定位信息	厂家告警号
0	ONENDS-HZHSS37BE02BNK	2019/11/11 13:44	COMMUNICATIONS FAULT	话音网	诺基亚	HSS	二级告警	EVENT_NUMBER\\: NDS-40403 FAULT_DETAILS\\: Lin...	34860
1	ONENDS-HZHSS37BE02BNK	2019/11/11 13:44	COMMUNICATIONS FAULT	话音网	诺基亚	HSS	二级告警	EVENT_NUMBER\\: NDS-40403 FAULT_DETAILS\\: Lin...	34860
2	ONENDS-HZHSS37BE02BNK	2019/11/11 13:44	COMMUNICATIONS FAULT	话音网	诺基亚	HSS	二级告警	EVENT_NUMBER\\: NDS-40403 FAULT_DETAILS\\: Lin...	34860
3	ONENDS-HZHSS37BE02BNK	2019/11/11 13:44	COMMUNICATIONS FAULT	话音网	诺基亚	HSS	二级告警	EVENT_NUMBER\\: NDS-40403 FAULT_DETAILS\\: Lin...	34860

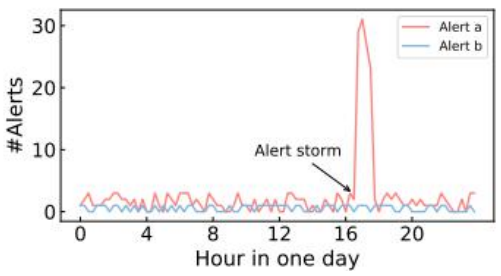
样例告警数据

二、成果主要内容-关联分析与告警收敛

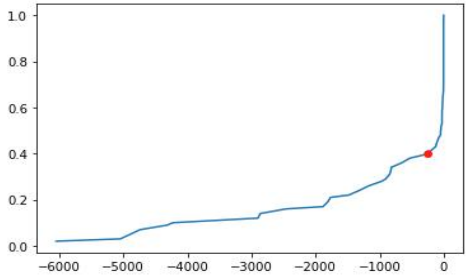
3.告警收敛具体流程：收集近一段时间（例如1分钟）在线告警数据，利用ECDF统计分布函数/极值理论进行告警风暴检测；一旦发现告警风暴，就对数据进行特征抽取，并利用孤立森林算法对告警数据进行异常值剔除；并输出杰卡德距离的距离度量矩阵给DBSCAN算法进行聚类，然后输出各个聚类类别的核心推送给相关工程师进行故障排除：



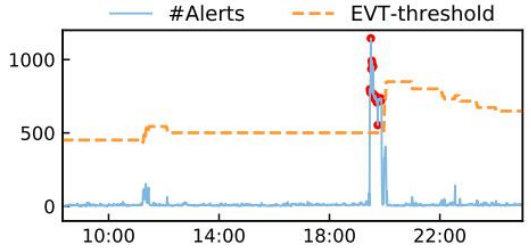
告警收敛具体流程



告警风暴



ECDF判断告警风暴



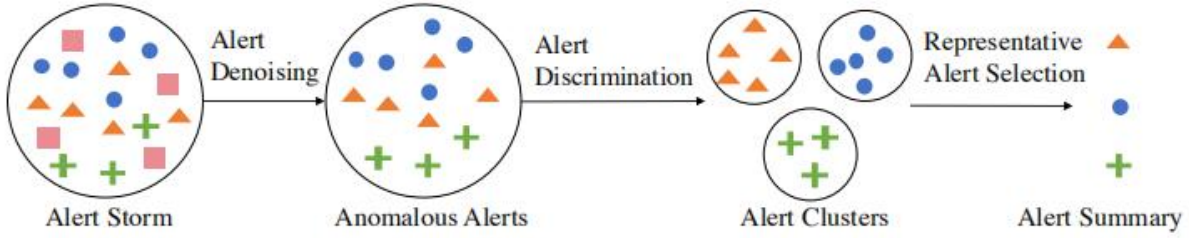
极值理论判断告警风暴

$$Jaccard(a,b) = 1 - \frac{|bow(a) \cap bow(b)|}{|bow(a) \cup bow(b)|}$$

杰卡德距离

$$centroid = \arg\min_{i \in cluster} \frac{1}{n} \sum_{j=1}^{n} similarity(i,j)$$

类别核心



告警分类图示

如左图所示，在对近1分钟告警数据进行收集时，告警数据分为很多个类别。在检测到告警风暴后，进行数据去噪，并进行告警分类，对各个类别的告警只输出类别核心告警，大幅度提高相关运维工程师的运维效率！

Contents

二、下钻根因分析

核心网络数据中心的监控指标是由多个指标汇聚而成，例如话务成功率指标是由多个设备组成的pool利用脚本计算宁波，杭州等全省各个地市的话务成功率汇聚而成，当话务成功率指标出现异常的时候，我们应当下钻到地市，分析哪个地市的指标异动导致的话务成功率指标出现异常。在调研了相关时间序列根因分析算法后，**基于微软提出的Adtributor做了适应核心网络数据中心的算法改进。**

```
1  Foreach m ∈ M // Compute surprise for all measures
2      Foreach E_ij // all elements, all dimensions
3          p = F_ij(m)/F(m) // Equation 5
4          q = A_ij(m)/A(m) // Equation 6
5          S_ij(m) = D_JS(p, q) // Equation 7
6  ExplanatorySet = {}
7  Foreach i ∈ D
8      SortedE = E_i.SortDescend(S_ij(m)) //Surprise
9      Candidate = {}, Explains = 0, Surprise = 0
10     Foreach E_ij ∈ SortedE
11         EP = (A_ij(m) − F_ij(m))/(A(m) − F(m))
12         if (EP > T_EEP) // Occam's razor
13             Candidate.Add += E_ij
14             Surprise += S_ij(m)
15             Explains += EP
16         if (Explains > T_EP) // explanatory power
17             Candidate.Surprise = Surprise
18             ExplanatorySet += Candidate
19             break
20 //Sort Explanatoryset by Candidate.Surprise
21 Final = ExplanatorySet.SortDescend(Surprise)
22 Return Final.Take(3) // Top 3 most surprising
```

对于异常KPI，计算所有元素的S值

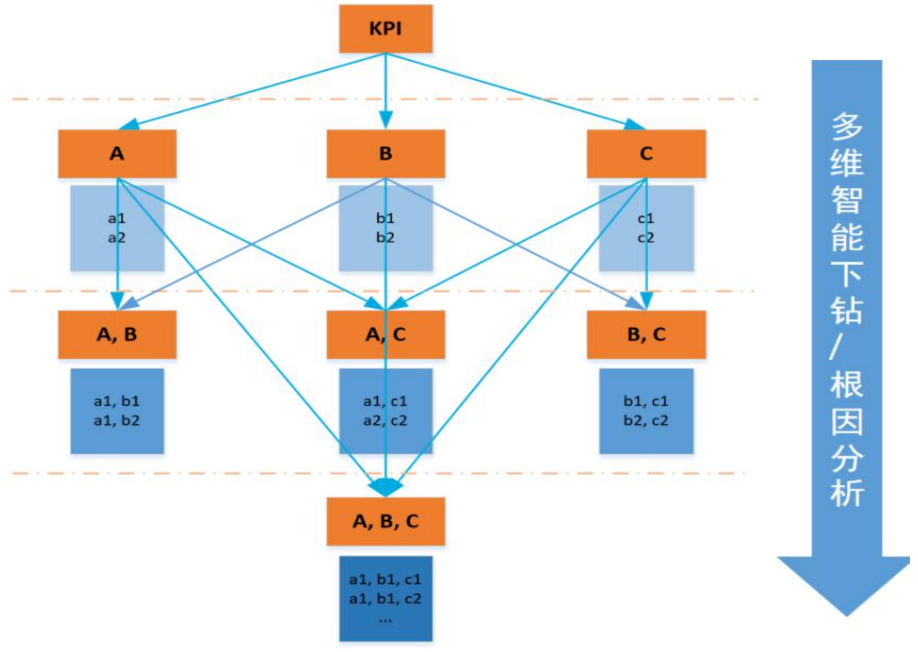
将每一维度下的元素按照S值降序排列

计算EP值，筛选可疑元素加入根因集合

计算维度EP值，忽略影响小的元素

根因集合按照S值降序排列，结果输出

Adtributor



Adtributor是微软2014年提出的多维智能下钻算法，但是在核心网络数据中心，时间序列KPI的异常波动主要由地市KPI的异常造成，故对Adtributor算法做了一些改进:

step1：对于异常KPI,利用MA算法预测每一个地市KPI值。

step2：利用每一个地市KPI的预测值和真实值，计算EP值。

step3：筛选大于EP阈值的所有地市，作为根因。

三、部分参考前沿论文

1. 《Unsupervised Anomaly Detection via Variational Auto-Encoder for Seasonal KPIs in Web Applications》
2. 《Clustering Intrusion Detection Alarms to Support Root Cause Analysis》
3. 《Time-Series Anomaly Detection Service at Microsof》
4. 《k-Shape: Efficient and Accurate Clustering of Time Series》
5. 《Probabilistic Alert Correlation》
6. 《FluxRank: A Widely-Deployable Framework to Automatically Localizing Root Cause Machines for Software Service Failure Mitigation》
7. 《Detecting Leaders from Correlated Time Series》
8. 全球运维大会：百度，携程，美团，必示，微众银行智能运维实践分享
9. 《HotSpot: Anomaly Localization for Additive KPIs With Multi-Dimensional Attributes》
10. 《Fast Time Sequence Indexing for Arbitrary Lp Norms》
11. 《A density-based algorithm for discovering clusters in large spatial databases with noise》
12. 《On the Surprising Behavior of Distance Metrics in High Dimensional Space》
13. 《Robust and Rapid Clustering of KPIs for Large-Scale Anomaly Detection》
14. 《YADING: Fast Clustering of Large-Scale Time Series Data》
15. 《Understanding and Handling Alert Storm for Online Service Systems》
16. 《CoFlux: Robustly Correlating KPIs by Fluctuations for Service Troubleshooting》

站在巨人的肩膀上，能让我们看的更远！

谢 谢!

中国移动 China Mobile

5G++创新实训基地

技术实操的练兵场 · 能力认证的人才站 · 5G应用的孵化器