

基于机器学习的智能运维

清华大学 裴丹

- 背景介绍
- 智能运维：从基于规则到基于学习
- 百度案例
- 挑战与思路

我的运维之路

我的官方简历

2005年：UCLA计算机系最佳博士论文, 研究BGP

2003年夏：AT&T研究院实习

2005-2011：AT&T研究院资深研究员和主任研究员 ACM 和IEEE Senior Member，与20+美国教授合作，23项美国专利

2012至今清华大学计算机系副教授、博导、入选千人计划（青年项目）、ACM/IEEE Senior Member, 80多篇学术论文

我的运维简历

与ISP运维人员密切打交道五年

喜欢上分析实际运维数据

第五级运维，基于大数据技术管理网络和应用（BGP, OSPF, DSL, IPTV, CDN, Cellular, Web, App, Video Streaming）的性能、可靠性和安全

开设网络/应用管理课程

所有科研项目都是运维相关：与百度、微软Azure云计算、清华校园网、中石油数据中心的运维部门合作



ACM SIGCOMM 2015 Call for Papers

London, UK; August 17-21, 2015

<http://conferences.sigcomm.org/sigcomm/2015>



The ACM SIGCOMM 2015 conference seeks papers describing significant research contributions to the field of computer and data communication networks. We invite submissions on a wide range of networking research, including, but not limited to:

- Design, implementation, and analysis of network architectures and algorithms
- Enterprise, datacenter, and storage area networks
- SDNs and network programming
- Experimental results from operational networks or network applications
- Economic aspects of the Internet
- Energy-aware communication
- Insights into network and traffic characteristics
- Network management and traffic engineering
- Network security and privacy
- Network, transport, and application-layer protocols
- Networking issues for emerging applications
- Fault-tolerance, reliability, and troubleshooting
- Operating system and host support for networking
- P2P, overlay, and content distribution networks
- Resource management, QoS, and signaling
- Routing, switching, and addressing
- Techniques for network measurement and simulation
- Wireless, mobile, and sensor networks

the SIGCOMM 2015 PC includes experts in the core EE areas of optical and wireless communications. They will contribute reviews for these submissions.

Authors must as part of the submission process attest that their work complies with all applicable ethical standards of their home institution(s), including, but not limited to privacy policies and policies on experiments involving humans. The PC takes a broad view of what constitutes an ethical concern, and authors agree to be available at any time during the review process to rapidly respond to queries from the PC chairs regarding ethical standards.

Important Dates

Paper registration: January 23, 2015 (7:59 PM GMT)

Paper submission: January 30, 2015 (7:59 PM GMT)

Decision notification: April 24, 2015

Organizing Committee

General Chairs

Steve Uhlig, Queen Mary Univ. of London, UK

Olaf Maennel, Tallinn University of Technology, Estonia

Program Committee Chairs

Brad Katz, University College London, UK

SIGCOMM 2015 评委会中的AT&T运维人员

Chairs

Brad Karp, University College London, UK
Jirendra Padhye, Microsoft Research, USA

PC Members

Aditya Akella, Univ. of Wisconsin, Madison, USA
Mohammad Alizadeh, MIT and Cisco, USA
Katerina Argyraki, EPFL, Switzerland

Aruna Balasubramanian, Stony Brook University, USA

Hitesh Ballani, Microsoft Research, UK

Sujata Banerjee, HP Labs, USA

Keren Bergman, Columbia University, USA

John Byers, Boston University, USA

Jeff Chase, Duke University, USA

Mung Chiang, Princeton University, USA

Jon Crowcroft, University of Cambridge, UK

Bruce Davie, VMware, USA

Nandita Dukkkipati, Google, USA

Anja Feldmann, Technische Univ. Berlin, Germany

Bryan Ford, Yale University, USA

Nate Foster, Cornell University, USA

Lixin Gao, Univ. of Massachusetts Amherst, USA

Brighten Godfrey, UIUC, USA

Sharon Goldberg, Boston University, USA

Kyle Jamieson, University College London, UK

Colin Jones, University of Cambridge, UK

Ethan Katz-Bassett, Univ. of Southern California, USA

Teemu Koponen, VMware, USA

John C. S. Lui, Chinese Univ. of Hong Kong, Hong Kong

Z. Morley Mao, Univ. of Michigan, USA

Dave Oran, Cisco, USA

George Papen, UC San Diego ECE, USA

Kyoungsoo Park, KAIST, Korea

George Portet, UC San Diego CSE, USA

Luigi Rizzo, Università di Pisa, Italy

Ashutosh Sabharwal, Rice University, USA

Stefan Savage, UC San Diego CSE, USA

Michael Schapira, Hebrew University, Israel

Vyas Sekar, Carnegie Mellon University, USA

Scott Shenker, ICSI and UC Berkeley, USA

Kun Tan, Microsoft Research, China

Pramod Viswanath, UIUC, USA

Geoff Voelker, UC San Diego CSE, USA

Michael Walfish, NYU, USA

Jia Wang, AT&T Research, USA

Philip Watts, University College London, UK

David Wertherall, Google and Univ. of Washington, USA

Walter Willinger, Nixun, USA

Keith Winstein, Stanford University, USA

Xiaowei Yang, Duke University, USA

Mulan Yu, Univ. of Southern California, USA

Yi Yang, University of California, Berkeley, USA

实习运维

运维员工



Sponsored by ACM SIGCOMM and ACM SIGMETRICS

Call for Papers (full CFP at <http://conferences2.sigcomm.org/imc/2015/cfp.html>)

The Internet Measurement Conference (IMC) is a highly selective venue for the presentation of measurement-based research in data communications. The focus of IMC 2015 will be on papers that either (1) improve the practice of measurement or (2) illuminate some facet of an operational network. IMC takes a broad view of what constitutes an operational network. This view includes (but is not limited to):

- the Internet backbone and edge networks (e.g., home networks, cellular networks, WLANs)
- data centers and cloud computing infrastructure
- peer-to-peer and content distribution networks
- infrastructure for online social networks
- experimental networks affiliated with the Internet (e.g., overlay networks, future internets or other prototype networks)

Types of contributions that the program committee would enjoy receiving submissions regarding include (but are not limited to):

- collection and analysis of data that yield new insights about network structure and behavior
- methods and tools to monitor and visualize network-based phenomena
- systems and algorithmic techniques that leverage measurement-based findings in novel ways
- advances in data collection and handling (e.g., anonymization, querying, storage, facilitating sharing)
- modeling of network structure and behavior (e.g., workload, scalability, assessment of performance bottlenecks)
- reappraisal of previous empirical findings

清华大学NetMan实验室科研项目简介

<http://netman.cs.tsinghua.edu.cn>

国家自然科学基金支持



青年千人项目支持



国家自然科学基金支持



- **工业界与学术界应该在运维领域密切合作**
 - 工业界获得算法层面的深度支持
 - 学术界获得现实世界的前沿问题及数据，有利发表论文和申请国家项目

值得工业界运维同仁关注的顶级学术会议

最相关的single-track顶会
(Google, Facebook, Microsoft, LinkedIn 在这些会议中发标过运维相关论文)

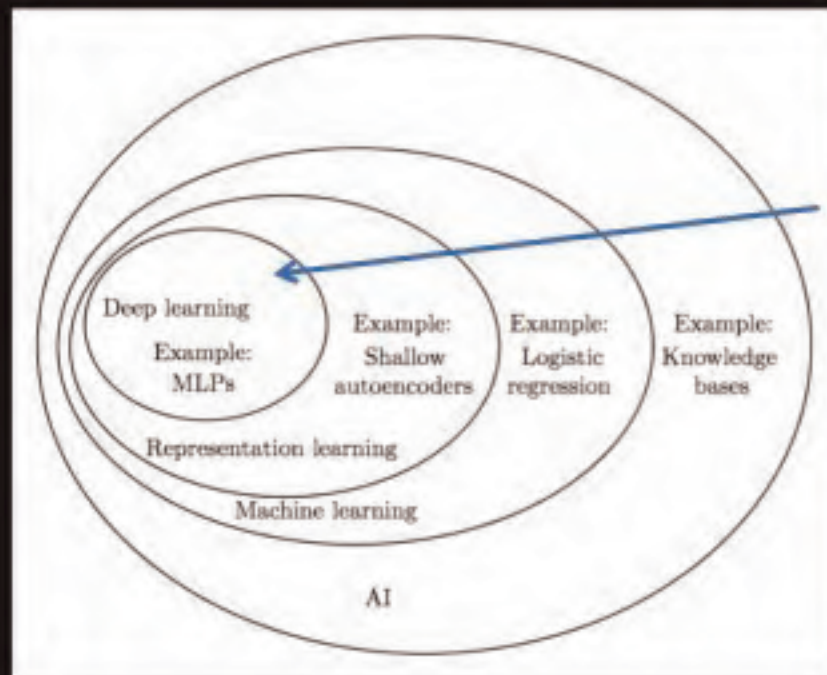
ACM SIGCOMM
ACM IMC
ACM/USENIX NSDI
ACM MobiSys
ACM CoNEXT
ACM MobiCom
ACM SIGMETRICS

相关Multi-track 顶会或偏安全方面的顶会

IEEE INFOCOM
ACM KDD
USENIX Security
IEEE Security & Privacy
ACM CCS
NDSS

- 背景介绍
- 智能运维：从基于规则到基于学习
- 百度案例
- 挑战与思路

人工智能 VS. 机器学习 VS. 深度学习



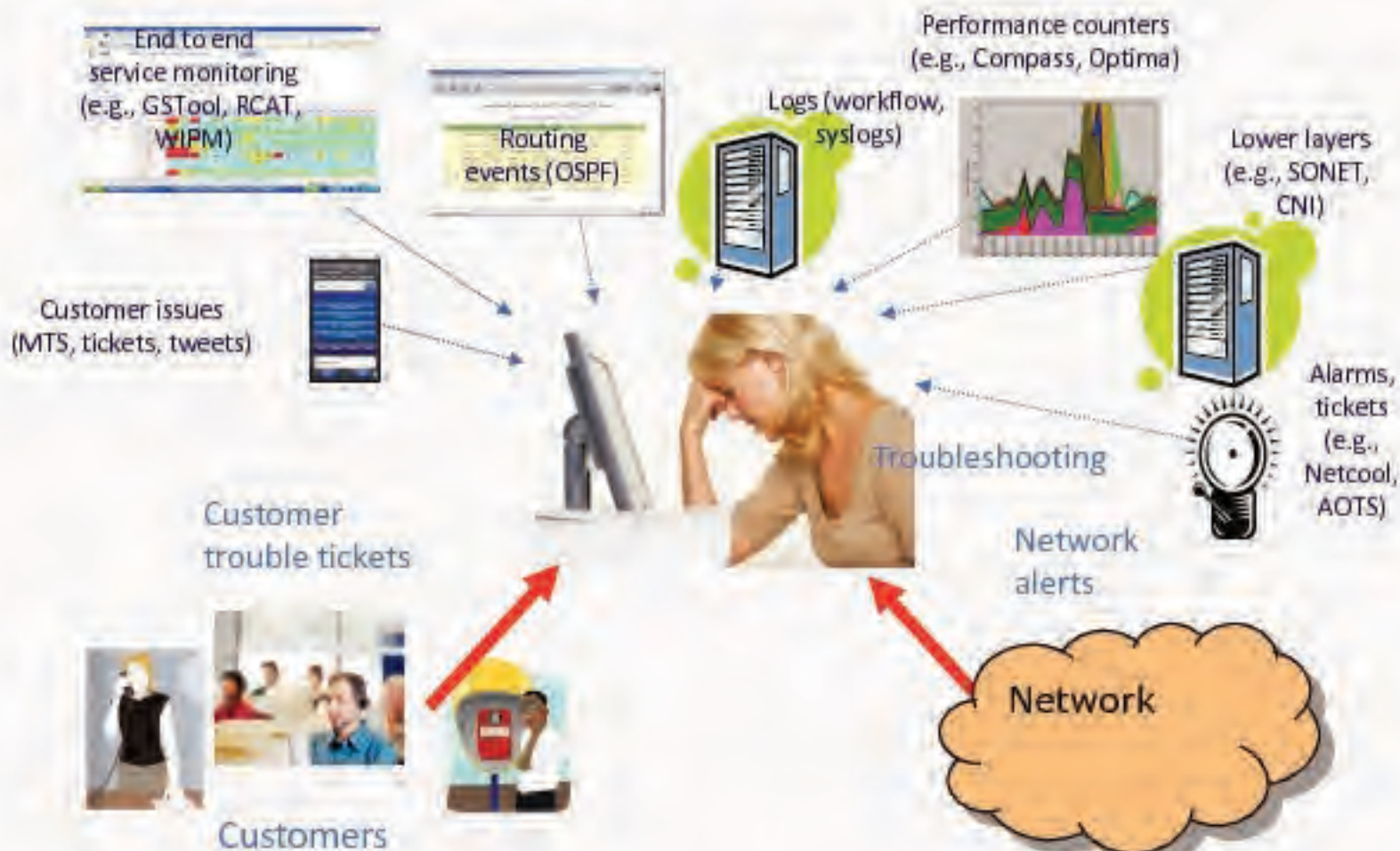
数据

标注

工具（算法和系统）

应用

互联网应用天然有海量日志作为特征数据；还可以按需自主生成新的日志数据



NSDI 2013

What Does a Ticket Contain?

STRUCTURED

Ticket Title	Ticket #xxxxxx-NetDevice: LoadBalancer Down 100%. Summary: Indicates that the root cause is a failed system		
Problem Type	Problem SubType	Priority	Created
Severity - 2	2: Medium		



STRUCTURED FIELDS
E.g., ticket title, problem type, priority etc.

UNSTRUCTURED (Diary)

Operator 1: I replaced the memory chips on this device and both power supplies have been reseated
Operator 2: The device has been powered back up. It should be back online shortly.
Operator 1: Ok. Let me check..
Operator 1: Yes. It is functional. Thanks!

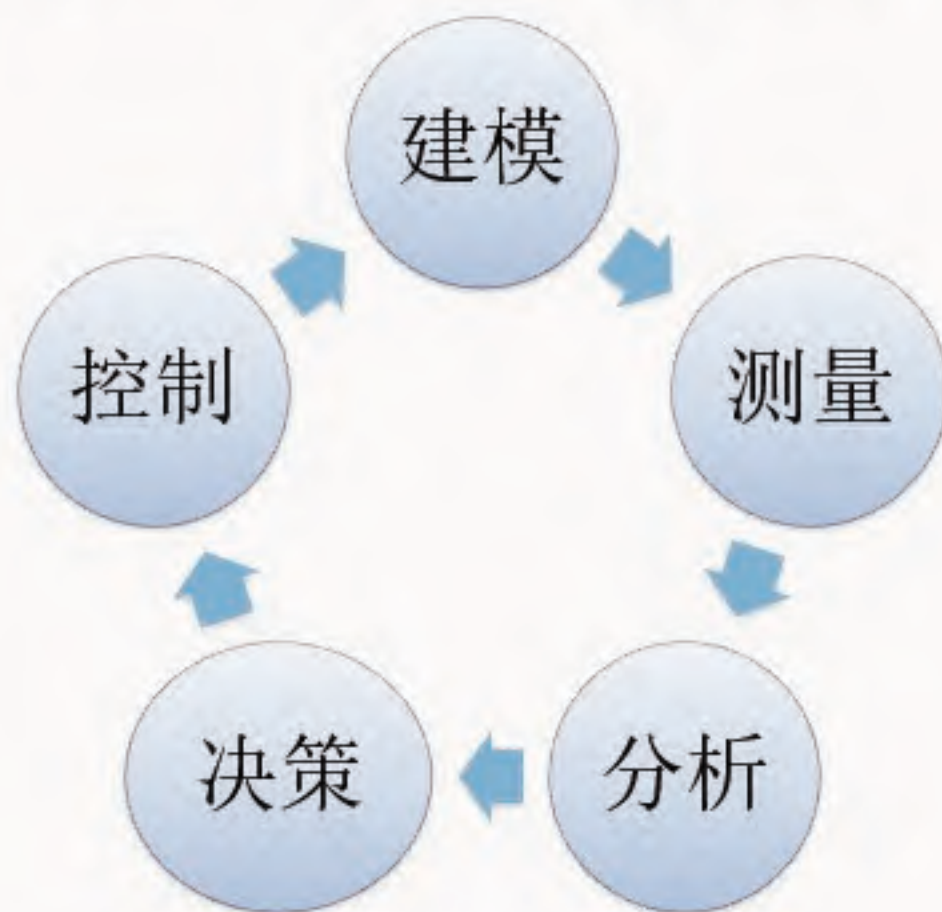
--- Original Message ---
From: Vendor Support
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continuously rebooting
As discussed, the device has bad memory chips as such we replace it. Please completely fill the RMA form below and return it.

--- Appended Message ---
From: Operations
Subject: Regarding Case Number #yyyyyy
Title: Device xxx-xxx-xxx-130b v9.4.5 continuously rebooting
We have cleaned the cable connecting the load balancer to the access router. Please invoke device diagnostics and send the logs to the vendor for further troubleshooting.



FREE-FORM TEXT
E.g., operator notes, emails, device debug logs, etc.

应用：运维人员就可以设计、部署、使用、并受益于智能运维系统，形成有效闭环



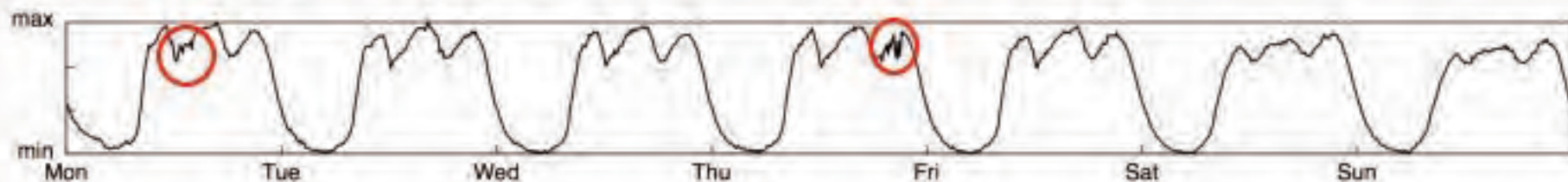
小结：智能运维在今后若干年会飞速发展

- “基于机器学习的智能运维” 具有得天独厚的基础
 - 互联网应用天然有海量日志作为特征数据
 - 运维日常工作日志产生标注数据
 - 大量成熟的机器学习算法和开源系统
 - 直接用于改善互联网应用

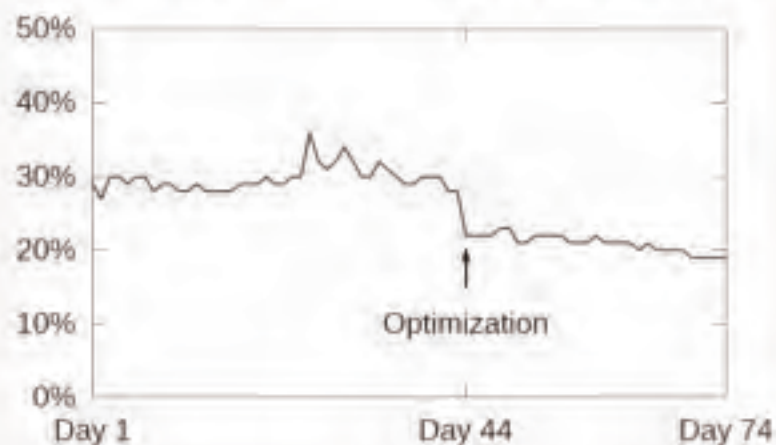
- 背景介绍
- 智能运维：从基于规则到基于学习
- *百度案例*
- 挑战与思路

智能运维的三个案例： 基于与百度运维、搜索部门的合作

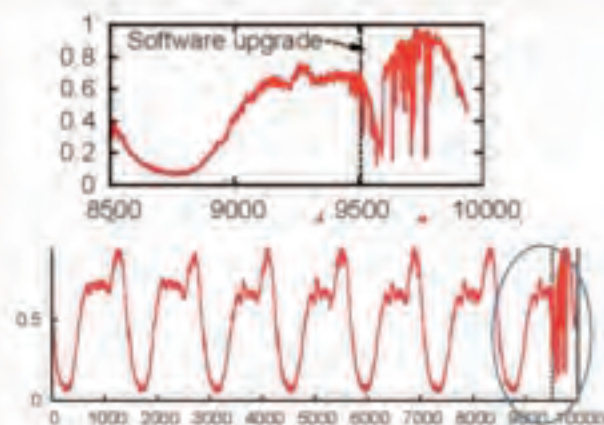
1. 自动检测PV异常



2. 自动分析性能瓶颈并提出优化建议



3. 自动关联KPI异常与版本上线



案例1：基于机器学习的KPI自动化异常检测

(Dapeng Liu et al. IMC 2015)

KPI异常检测



KPIs (Key Performance Indicators) : 用来衡量服务性能的关键指标

KPI异常行为 → 潜在的风险、故障、bugs、攻击.....

KPI异常检测 : 在KPI时序曲线上识别异常行为

→ 诊断和修复

→ 阻止进一步损失或潜在风险

构建KPI异常检测系统



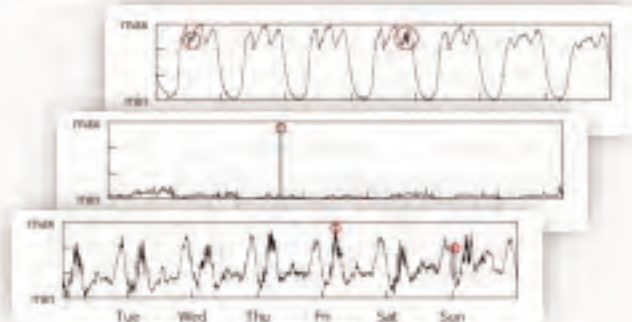
领域专家(运维人员)

- 对KPI负责
- 熟悉KPI的行为



算法开发人员

- 负责构建KPI异常检测系统
- 熟悉一些异常检测器(算法)



Simple threshold

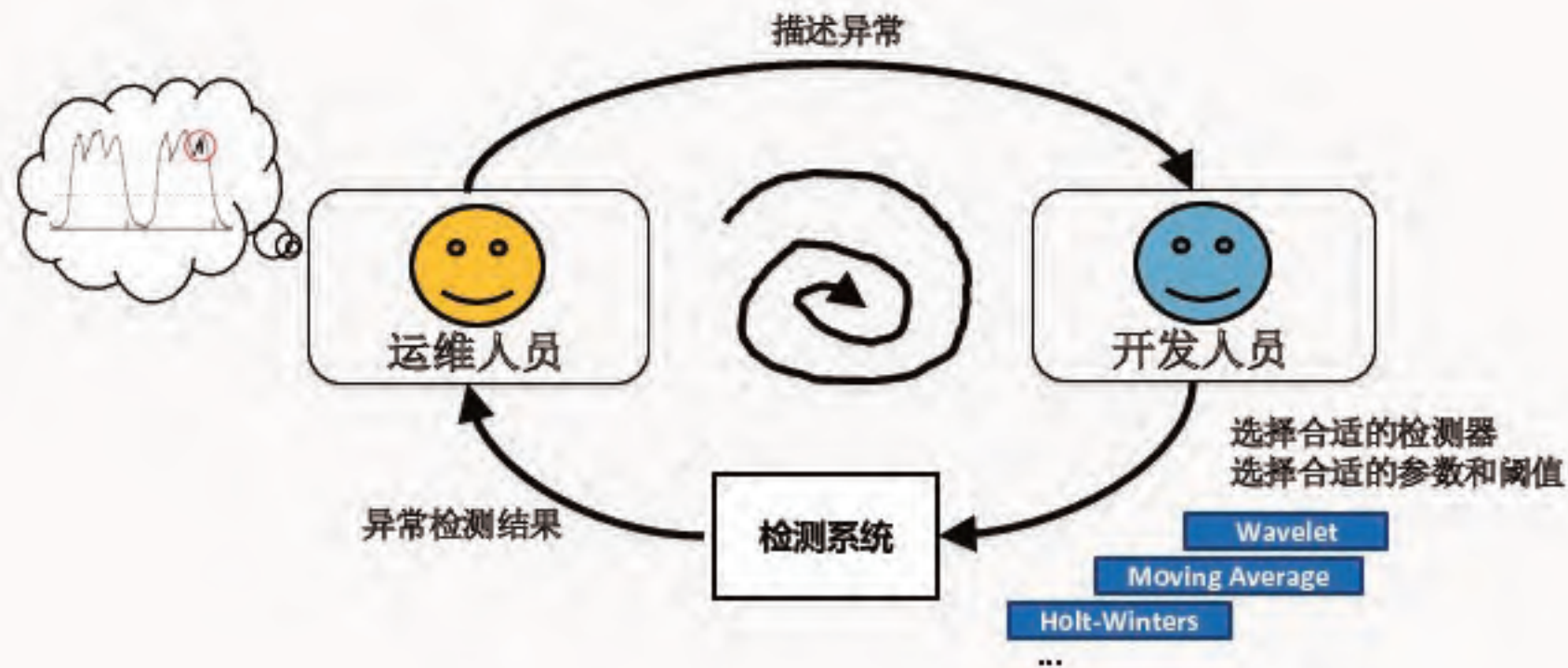
Historical Average

Wavelet

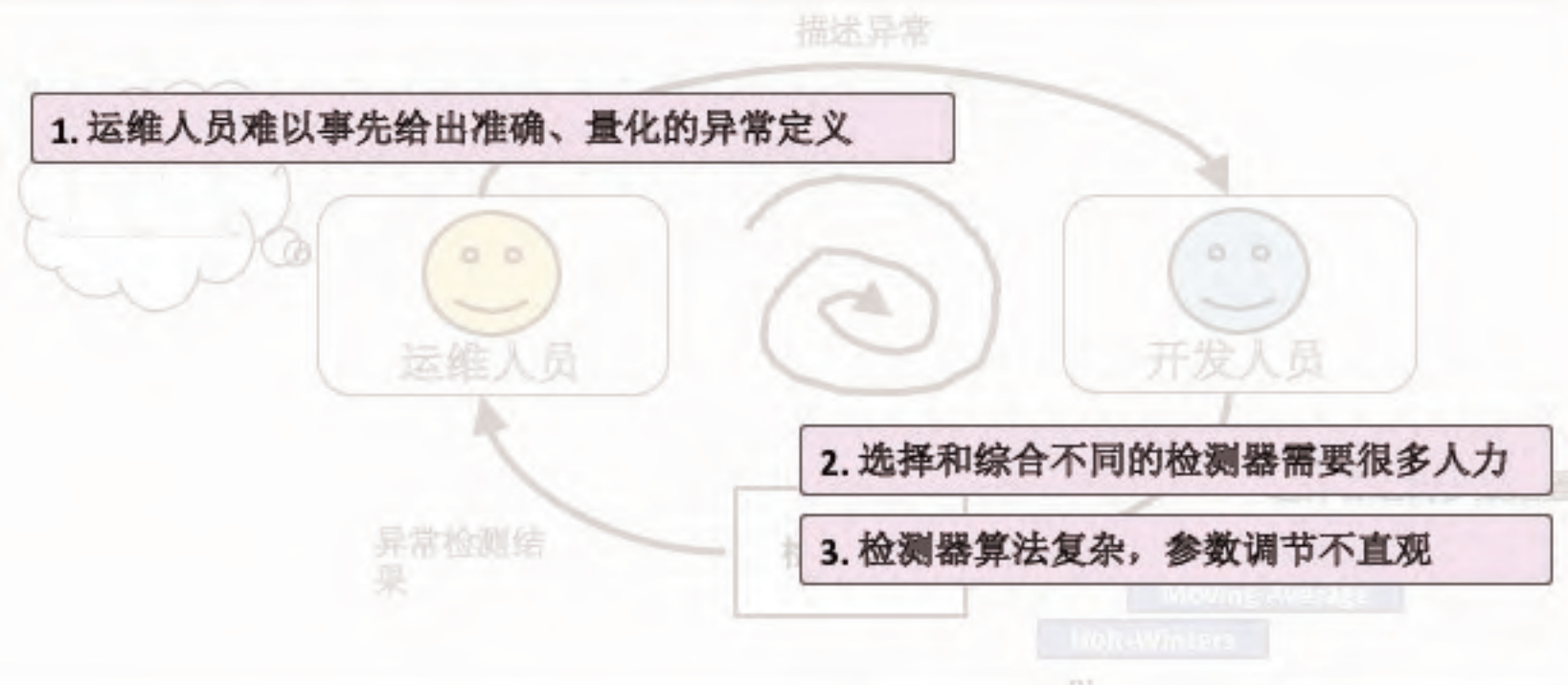
Holt-Winters

...

实践与挑战

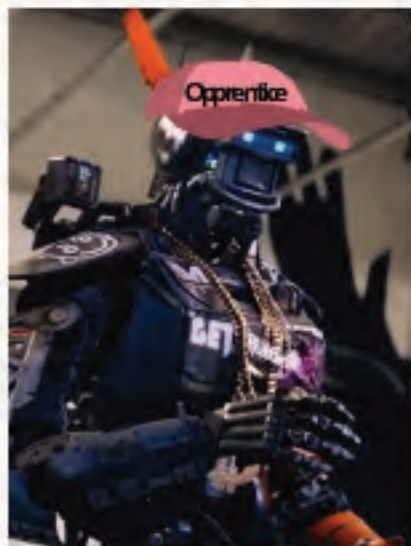


实践与挑战

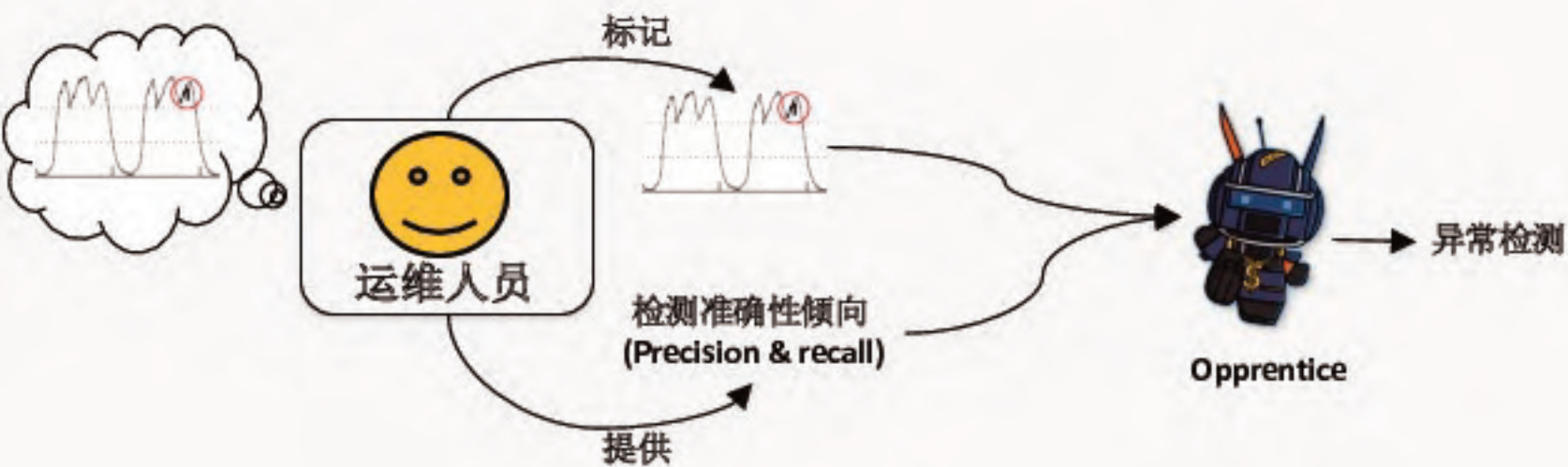


主要思想

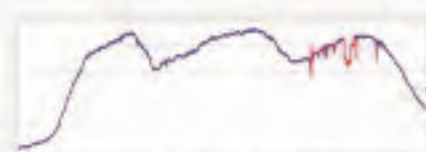
Opprentice (Op_{er}ator's App_{rentice}): 跟着运维人员从历史异常中学习



主要思想



主要思想



KPI 曲线

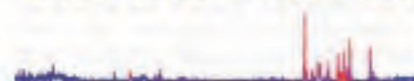


Simple threshold [23] / 1	none
Diff / 3	last-slot, last-day, last-week
Simple MA [4] / 5	
Weighted MA [10] / 5	win = 10, 20, 30, 40, 50 points
MA of diff / 5	
EWMA [10] / 5	$\alpha = 0.1, 0.3, 0.5, 0.7, 0.9$
TSD [1] / 5	
TSD MAD / 5	
Historical average [5] / 5	win = 1, 2, 3, 4, 5 week(s)
Historical MAD / 5	
Holt-Winters [6] / $4^3 = 64$	$\alpha, \beta, \gamma = 0.2, 0.4, 0.6, 0.8$
SVD [3] / $5 \times 3 = 15$	row = 10, 20, 30, 40, 50 points, column = 3, 5, 7
Wavelet [11] / $3 \times 3 = 9$	win = 3, 5, 7 days; freq = low, mid, high
ARIMA [9] / 1	Estimation from data
In total: 14 basic detectors / 133 configurations	

Historical average-4 season



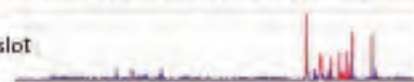
EWMA-0,7



WMA-WIN30



Differencing-last slot



Differencing-last season



Differencing-last day



Time series decomposition



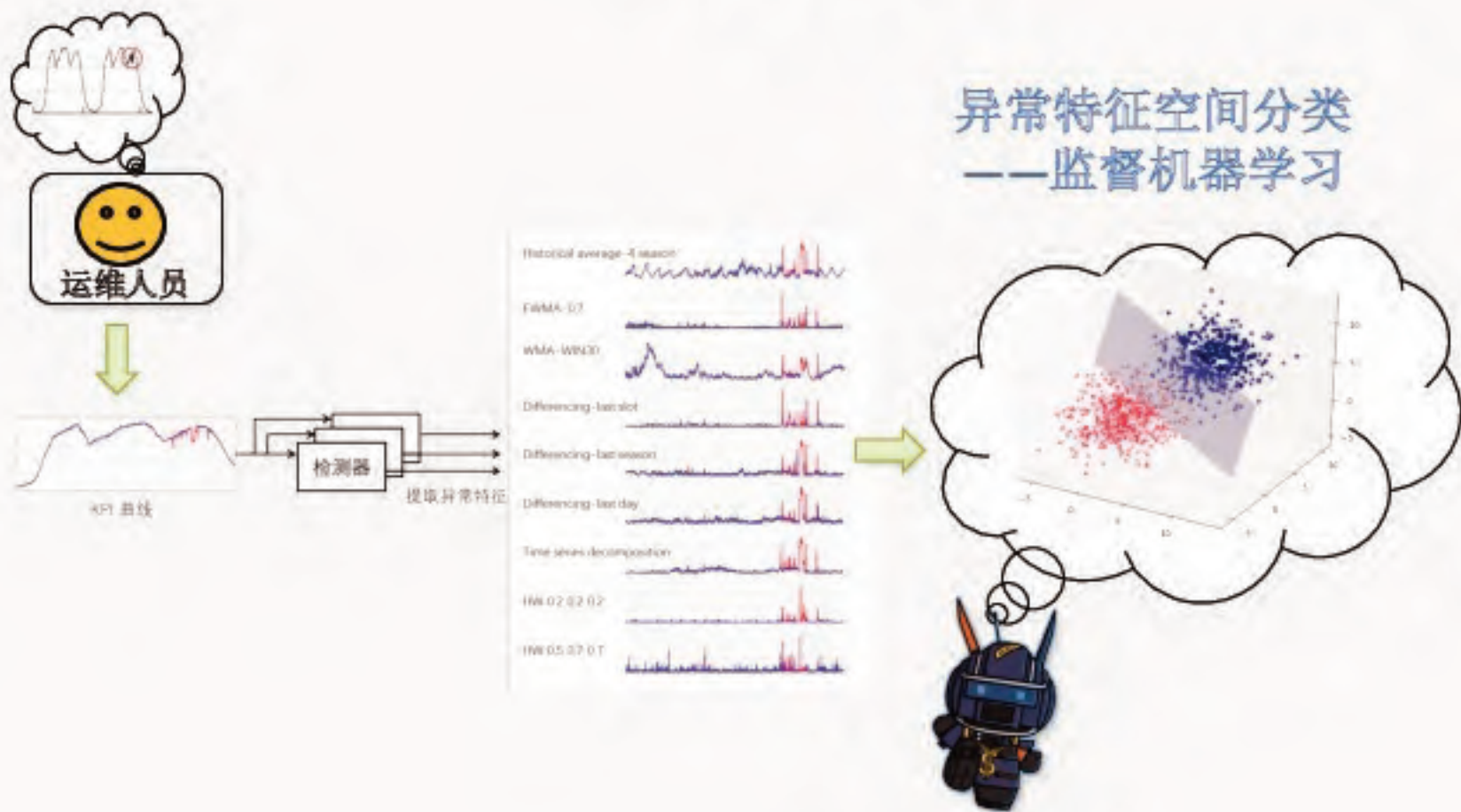
HW 0.2 0.2 0.2



HW 0.5 0.7 0.7



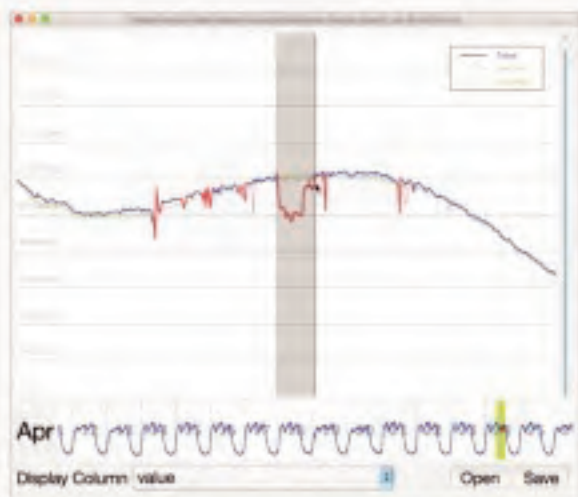
主要思想



挑战与解决方案

挑战1: 标记历史数据的开销

方案: 高效的标记工具



Y轴最大值调节

标记操作

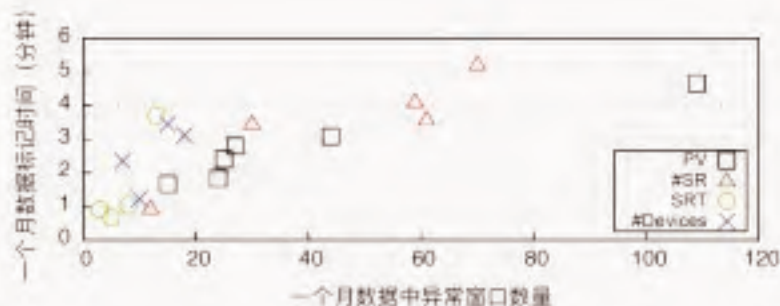
拖拽 标记异常窗口
拖拽 取消标记

时序数据导航

缩小时间粒度

向后移动 向前移动
放大时间粒度

导航器



挑战与解决方案

挑战1: 标记历史数据的开销

方案: 高效的标记工具

挑战2: 历史数据中异常种类少

方案: 用最新的数据增量学习

挑战3: 类别不均衡问题

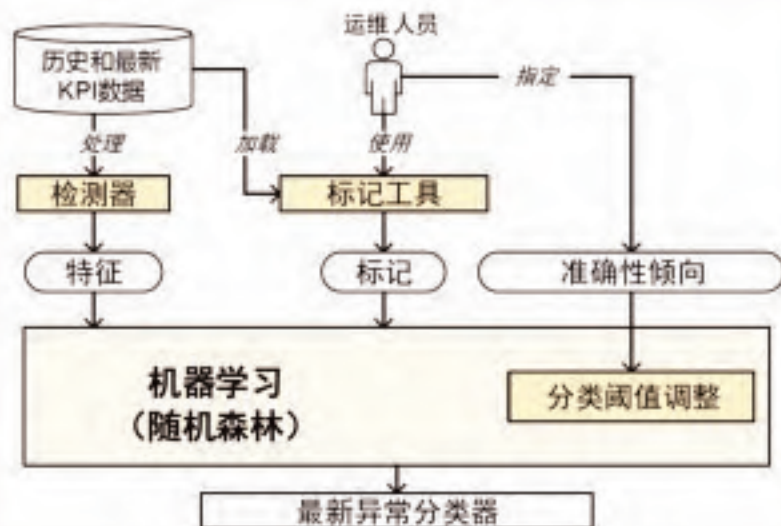
方案: 根据检测准确性倾向调整分类阈值

挑战4: 冗余和无关特征

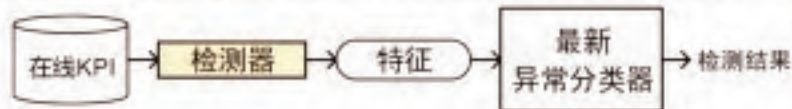
方案: 随机森林

Opprentice设计

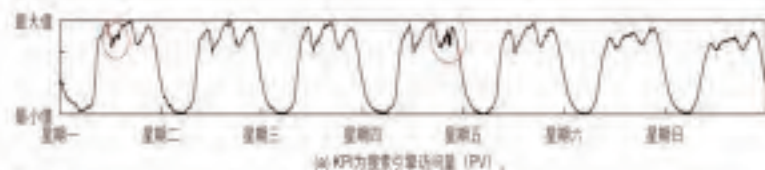
离线训练分类器



在线检测

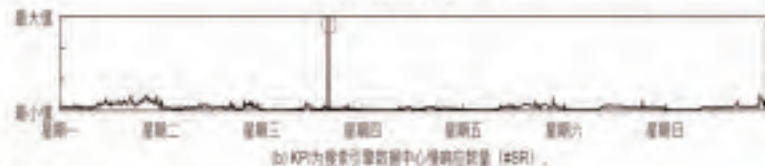


四种真实KPI数据



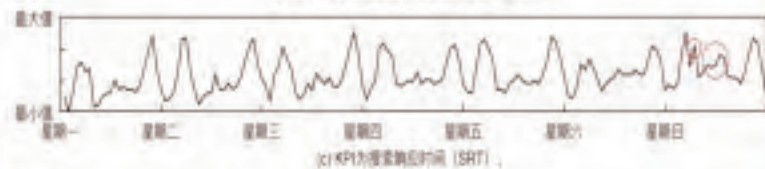
(a) KPI为搜索引擎访问量 (PV)。

搜索访问量 (25周)



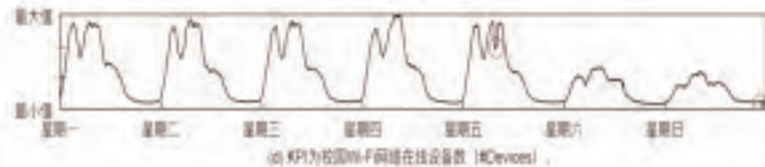
(b) KPI为搜索引擎数据中心慢响应数 (#SR)。

数据中心慢响应数 (19周)



(c) KPI为搜索响应时间 (SRT)。

搜索响应时间 (16周)



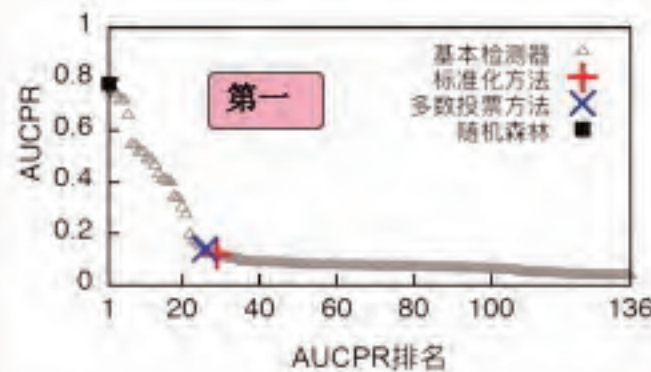
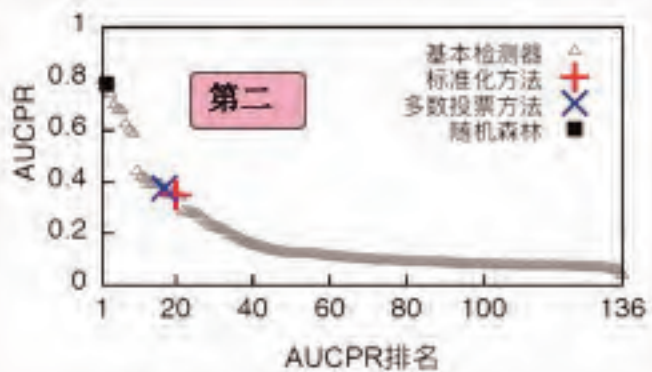
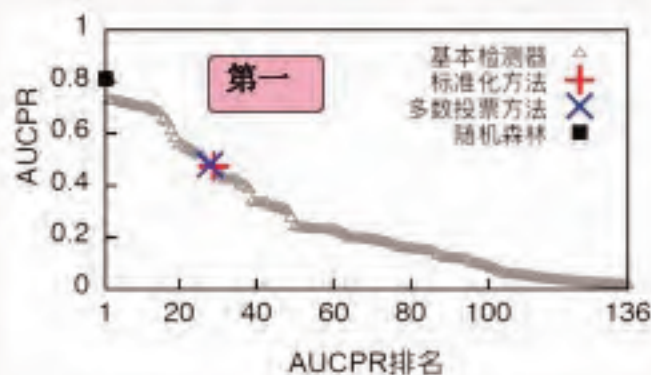
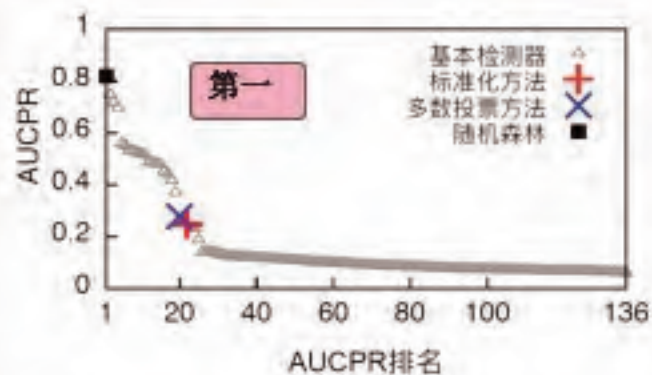
(d) KPI为校园Wi-Fi网络在线设备数 (#Devices)。

在线设备数 (15周)

百度

清华校园
无线网

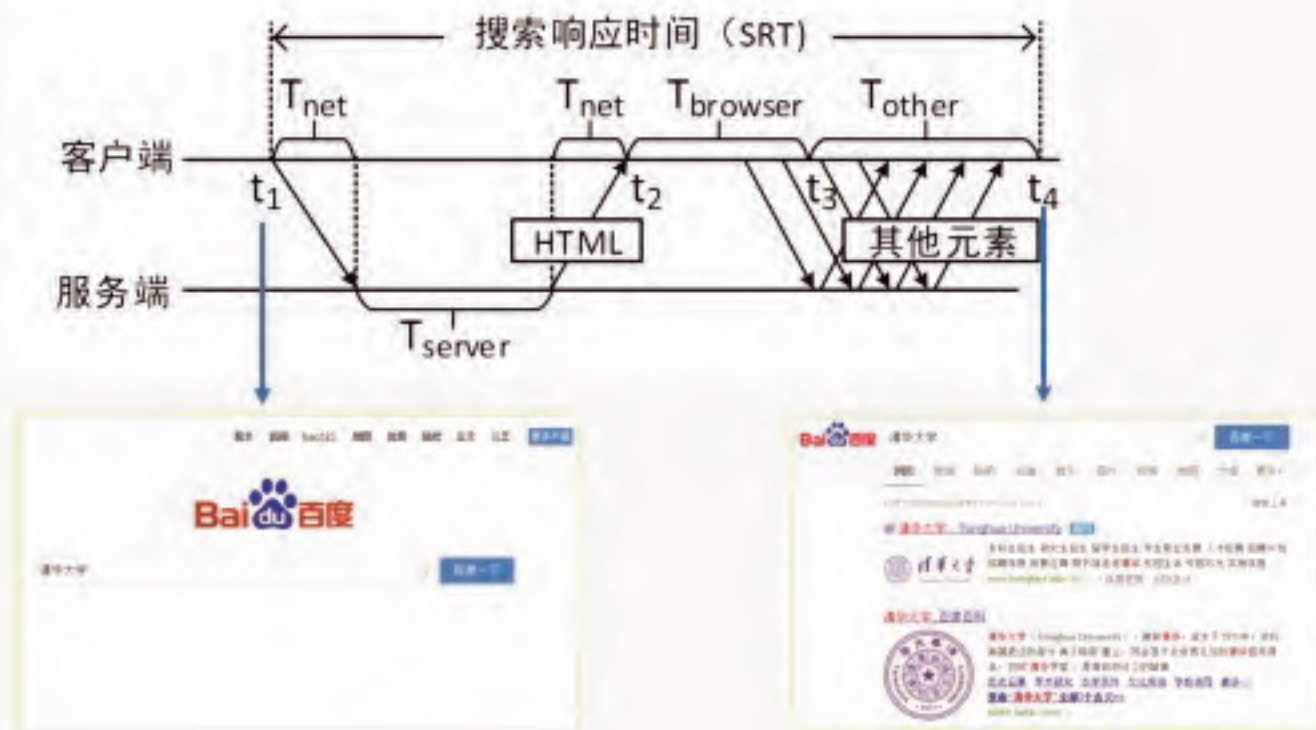
与已有检测器方法比较（四种KPI）



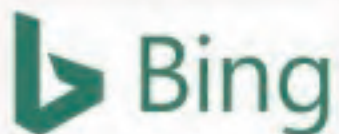
- 通过学习历史异常数据自动构建异常检测系统
 - 无需人工选择繁杂的检测器和调参
 - 为复杂检测器的实际应用提供自动化框架
- 采用来自百度、清华校园网的数个月的真实数据验证

案例2：多属性日志中的搜索响应时瓶颈分析 *(Dapeng Liu et al., INFOCOM 2016)*

用户在搜索中实际等待的时间



Web响应时间的重要性



+500ms 利润 ↓ 1.2%
[Eric Schurman, Bing]



+100ms~400ms 搜索 ↓ 0.2%~0.6%
[Jake Brutlag, Google]

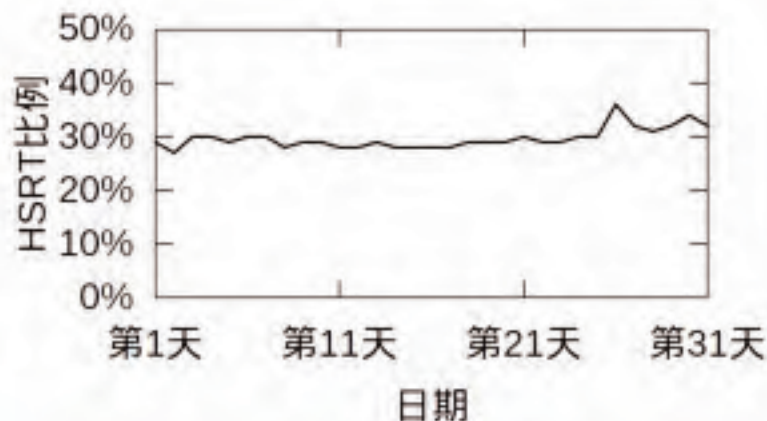
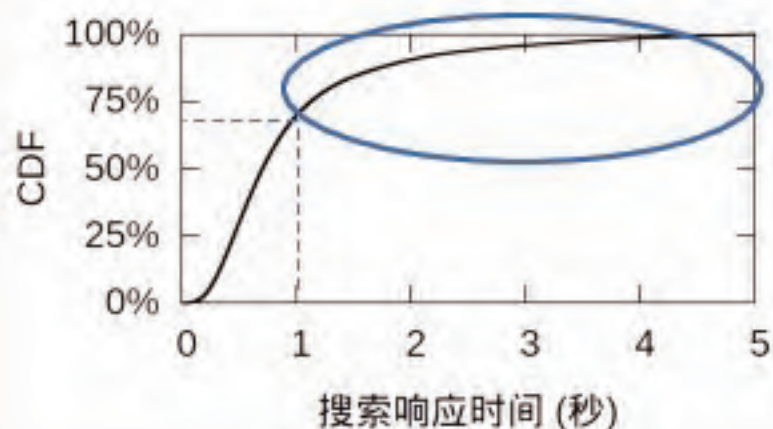


+100ms 销量 ↓ 1%
[Greg Linden, Amazon]



+1000ms 访问量 ↓ 11%
[Simic Bojan, Aberdeen]

实际中的搜索响应时间



问题：大于1秒的搜索（HSRT）是为什么？

High SRT

搜索引擎通过**搜索日志**来监测搜索响应时间SRT

潜在可能影响SRT的可测量属性

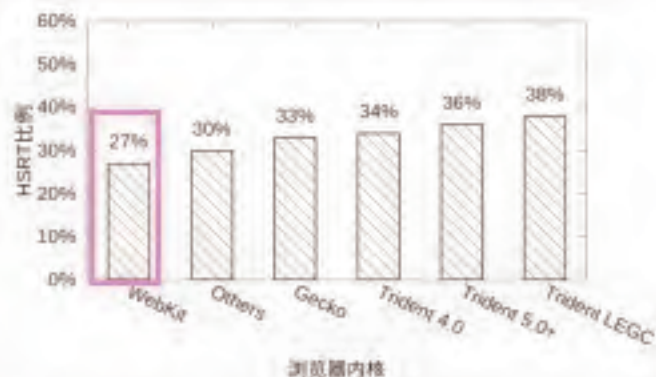
SRT	Client ISP	浏览器内核	图片数量	有无广告	后台负载
800ms (Low SRT)	China Unicom	WebKit	10	Yes	1000 PV/s
1200ms (High SRT)	China Telecom	Trident 5.0	5	No	500 PV/s
.....						

本项目提出搜索日志分析框架**FOCUS**来回答下面三个问题:

- HSRT容易发生的条件是什么?
- 哪些HSRT条件是相近的(HSRT条件类型), 并且比较流行?
- 流行的HSRT条件类型中的各个属性和值对SRT有怎样的影响?

分析多维属性搜索日志的挑战

单维度属性分析方法无法揭示属性组合的影响



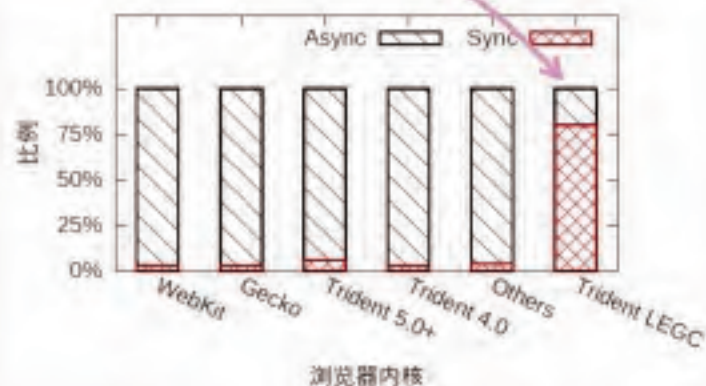
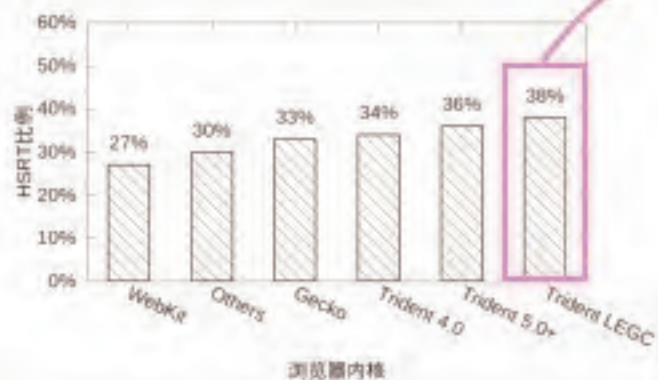
看到的：条件“**WebKit**”下的HSRT比例只有**27%**

未看到的：条件“**WebKit+图片数量多于30**”下的HSRT比例可以多于**38%**

分析多维属性搜索日志的挑战

单维度属性分析方法无法揭示属性组合的影响

属性间的潜在依赖关系 → 单维度分析的结论可能是片面的



Trident LEGC内核浏览器 还是同步加载?

分析多维属性搜索日志的挑战

单维度属性分析方法无法揭示属性组合的影响

属性间的潜在依赖关系→单维度分析的结论可能是片面的

得到的HSRT条件可重叠，每次HSRT被计算多次，不易理解

比如得到下面三个条件：

- “图片数量>30” → 贡献50%的HSRT
- “有广告” → 贡献40%的HSRT
- “图片数量>20，有广告” → 贡献30%的HSRT

重叠部分还是
非重叠部分？

总计120%？

主要思想

单维度属性分析方法无法揭示属性组合的影响

属性间的潜在依赖关系 → 单维度分析的结论可能是片面的

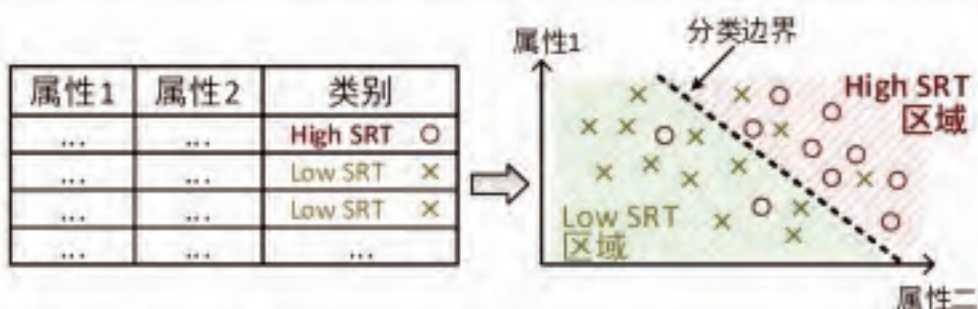
得到的HSRT条件可重叠，每次HSRT被计算多次，不易理解

将其建模为分类问题，利用监督机器学习算法——决策树得到直观分类模型

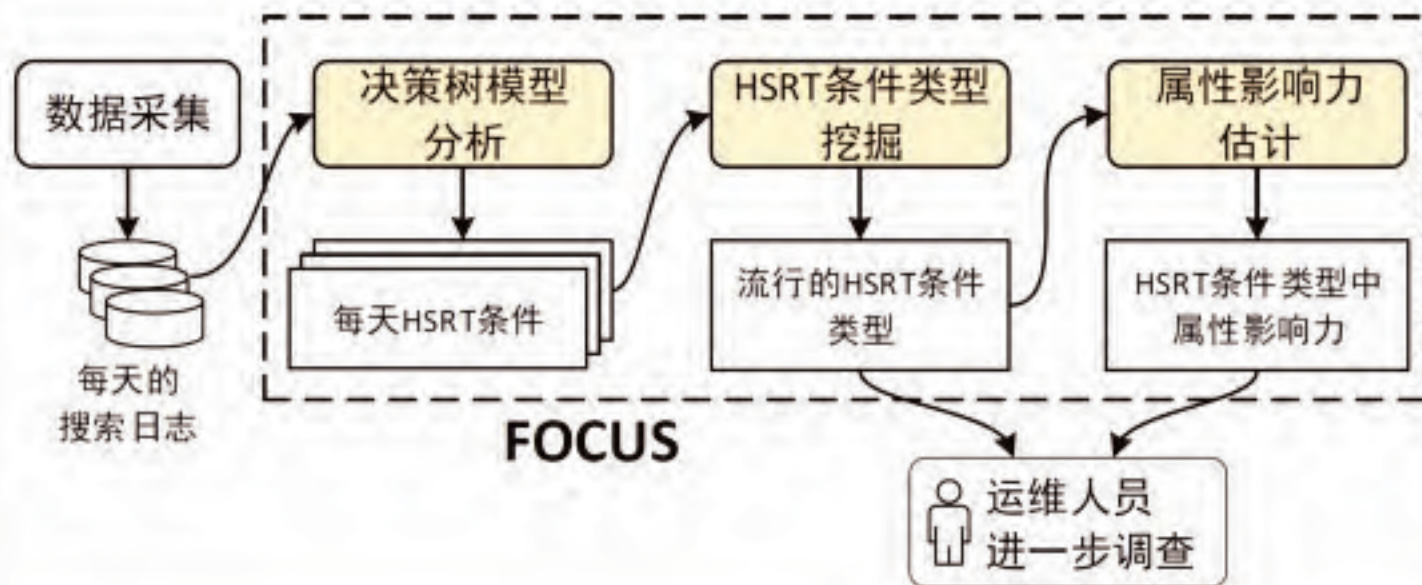
多维度分析

可以解决属性依赖关系

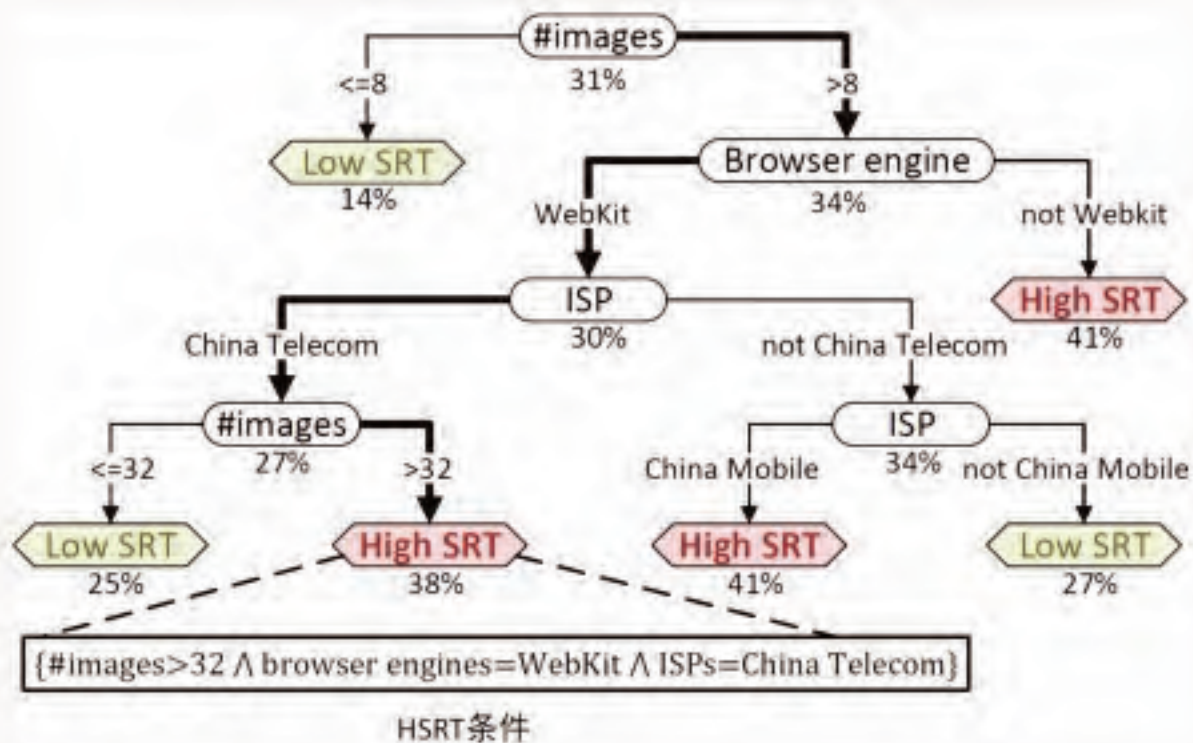
分类划分没有重叠



FOCUS概览



基于决策树的HSRT条件识别



挖掘相似HSRT条件 (HSRT条件类型)



HSRT 条件 编号	#images	示例属性取值 browser engine	ads
1	> 9	not WebKit	no
2	> 10	not WebKit	no
3	> 22	WebKit	yes

相似HSRT条件:

- 属性种类相同
- 类别型属性值相同
- 值型属性相似 → 层次化聚类

属性影响力估计

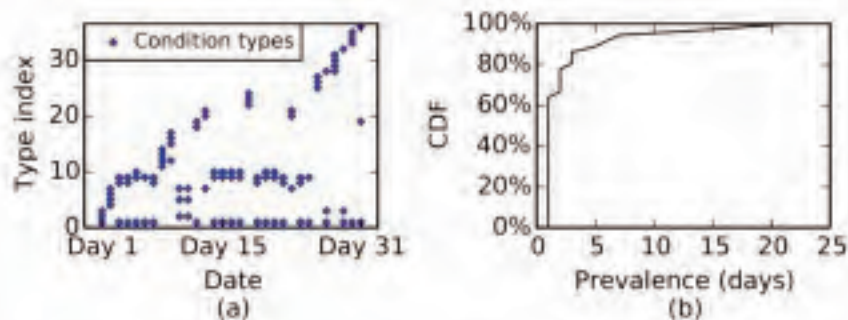
受控制实验启发:

- 每次变化一个属性条件（取补集）产生实验组
- 在历史数据中对比实验组和对照组条件下的SRT差别

HSRT 条件类型	属性条件 c_1 #images	属性条件 c_2 browser engine	属性条件 c_3 ads
原始对照组 C	$> i, i \in \{9, 10\}$	not WebKit	no
实验组 C'_1	$\leq i, i \in \{9, 10\}$	not WebKit	no
实验组 C'_2	$> i, i \in \{9, 10\}$	WebKit	no
实验组 C'_3	$> i, i \in \{9, 10\}$	not WebKit	yes

一个月的真实搜索日志中发现36种HSRT条件类型

其中4个出现超过5天



HSRT 条件 类型编号	HSRT 条件类型	出现天数	覆盖 HSRT
1	$\#images > i, i \in \{5, 6, 7, 8, 9\} \wedge browser\ engine = not\ WebKit$	21	43%
2	$\#images > i, i \in \{5, 6, 7, 8, 9\} \wedge ISP = not\ China\ Telecom \wedge browser\ engine = WebKit$	15	25%
3	$\#images > i, i \in \{25, 26, 27\} \wedge ISP = China\ Telecom \wedge browser\ engine = WebKit$	7	9%
4	$\#images > i, i \in \{5, 6, 8\} \wedge ISP = China\ Telecom \wedge browser\ engine = WebKit \wedge ads = yes$	6	9%

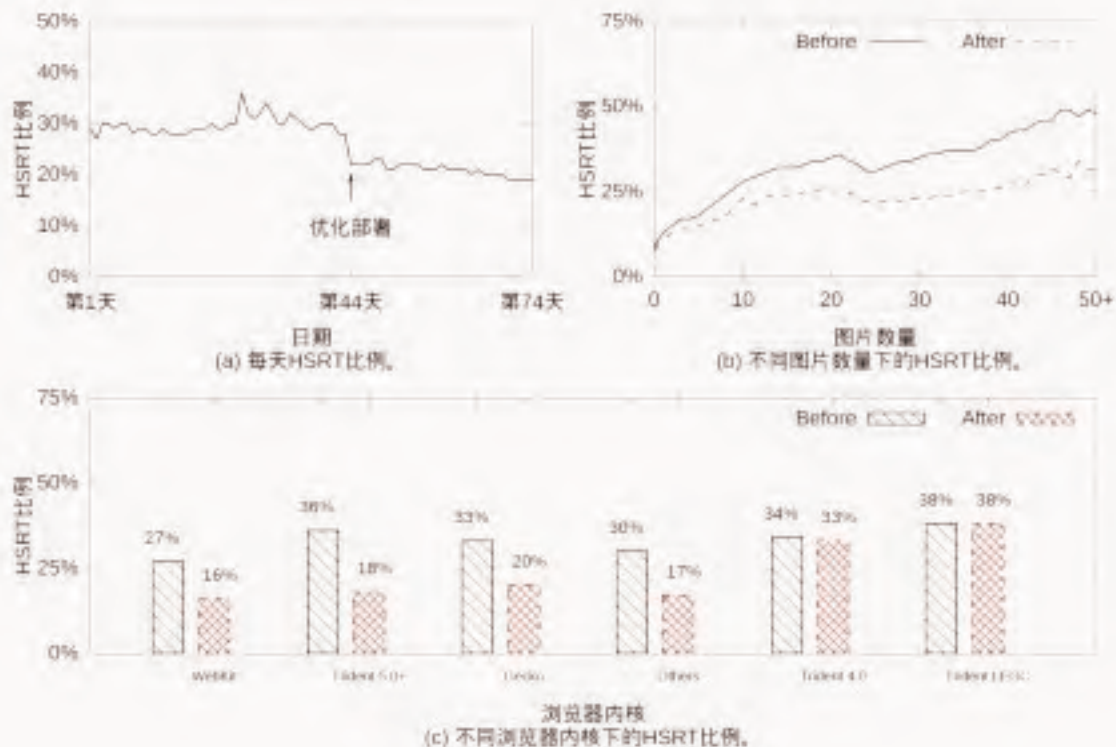
图片数量是主要的瓶颈

FOCUS的分析结果显示优化图片有最大提升潜力

部署base64 encoding提高“数量多、体积小”的图片传输速度

HSRT比例
减少30%

SRT 80分位数
下降253 ms (20%)



- 搜索日志中响应时间瓶颈分析系统
 - 为高搜索时间调查提供更具体方向
 - 根据历史数据估计不同属性的影响力
 - 通用性很高
- 部署于搜索引擎并分析2个月真实数据
- 根据分析结果实际部署优化方案，效果显著

案例3：软件更新对应用的影响
(Shenglin Zhang et al. CoNEXT 2015)

2014.1, Dropbox

Outage post-mortem

On Friday, we had a planned maintenance scheduled to upgrade the OS on some of our machines. During this process, the upgrade script checks to make sure there is no active data on the machine before installing the new OS.

A subtle bug in the script caused the command to reinstall a small number of active machines. Unfortunately, some master-replica pairs were impacted which resulted in the site going down.

- 部分服务器上规划中的操作系统升级
- Dropbox 服务下线3小时

2014.6, Facebook

Facebook outage caused by software system update

20 June 2014 By Helen Luibin

Facebook

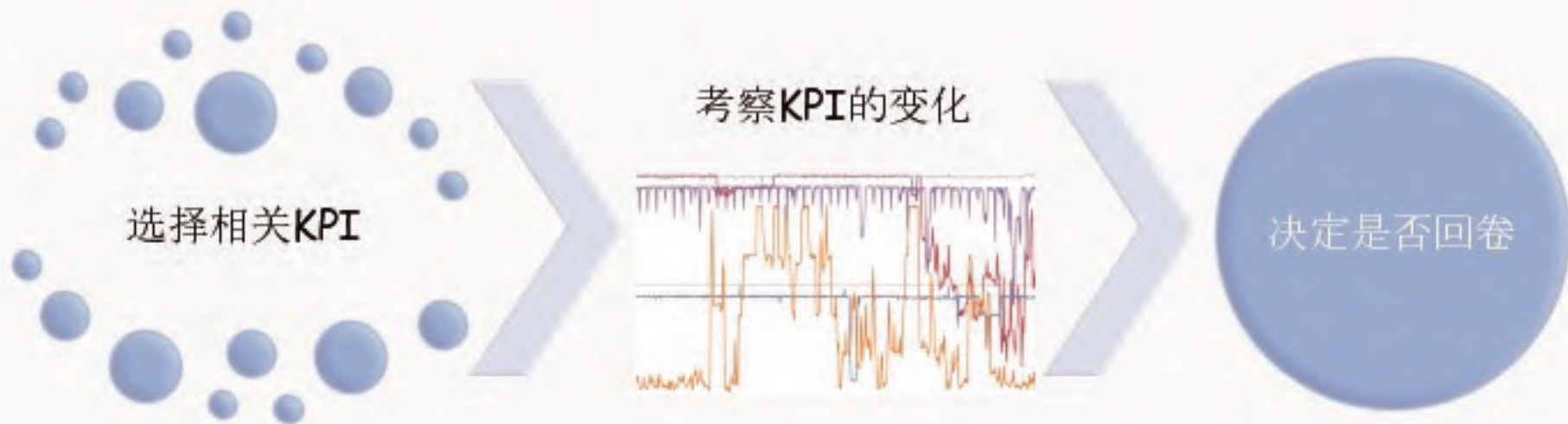
Shares Tweet Like

Social networking site Facebook suffered a worldwide outage yesterday after an issue while updating the configuration of one of its software systems.

The worldwide outage lasted for about 31 minutes.

Facebook

- 软件配置更新错误
- Facebook 下线31分钟



- 自动
- 可扩展
- 鲁棒性强

挑战1：检测延迟短 vs. 监测鲁棒性强

用户体验下降

营收损失

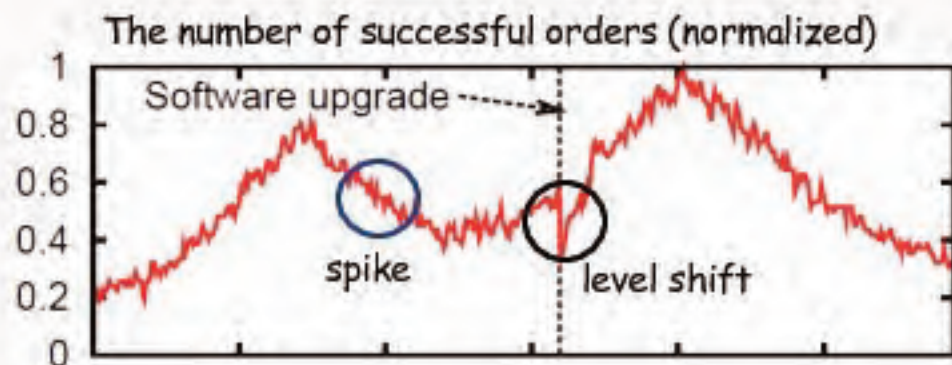


真实事件

挑战1：检测延迟短 vs. 监测鲁棒性强

用户体验下降

营收损失



真实事件

挑战1：检测延迟短 vs. 监测鲁棒性强

用户体验下降

营收损失



KPI变化检测要又快又准



真实事件

挑战2：大量KPI



挑战2：大量KPI & 大量软件更新



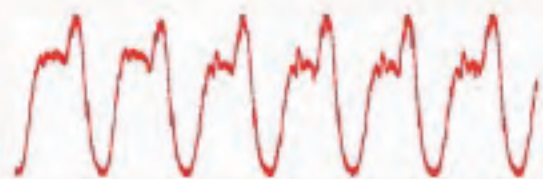
- 规模大
 - 100多个产品线
 - 上万个模块
 - 几十万台服务器
 - 百万级KPI监控
- 变化快
 - 每天上万个软件更新



计算开销要小

不同类型的KPI数据

季节性



PV

多变



网卡吞吐率

静态

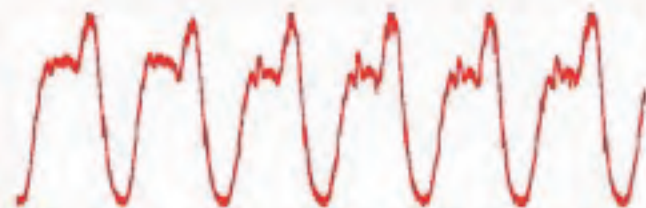


内存利用率

需要适应多样性的KPI数据

挑战4：KPI变化可能是其它因素导致的

季节变化



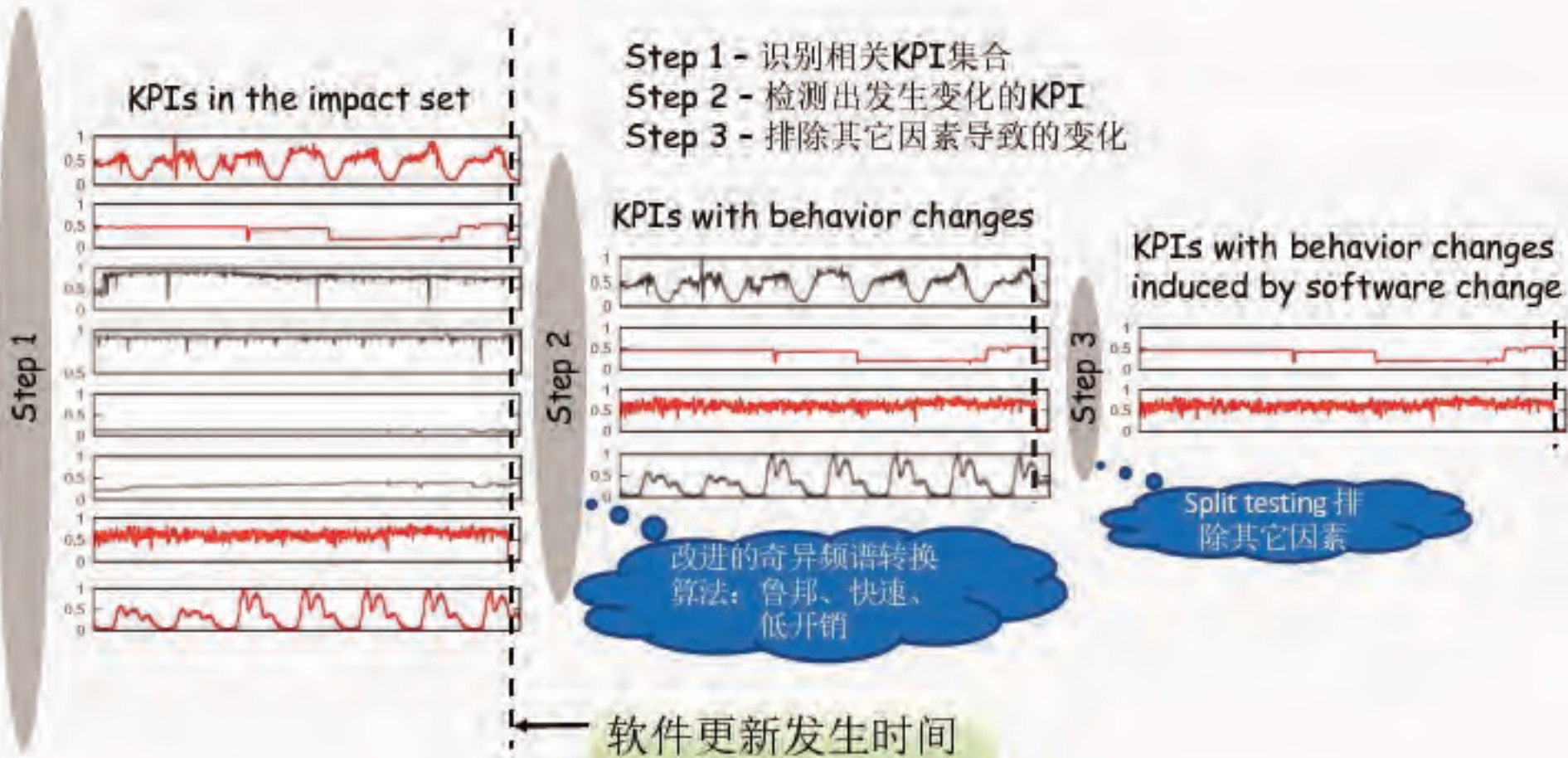
网络故障



恶意攻击



需要消除其它因素的影响



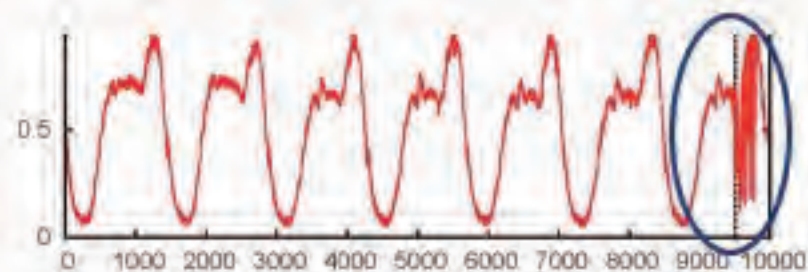
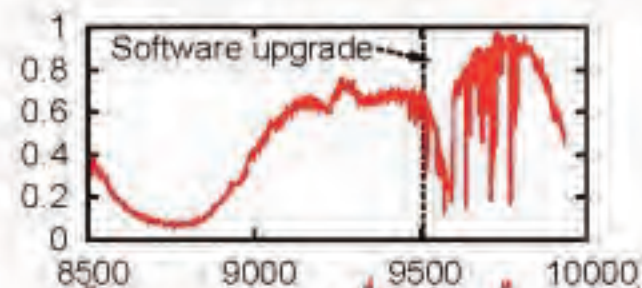
案例举例: 一个带bug的版本上线把来自iPhone的搜索流量都屏蔽了

FUNNEL :

- 能十分钟准确检测出问题并定位到该版本上线

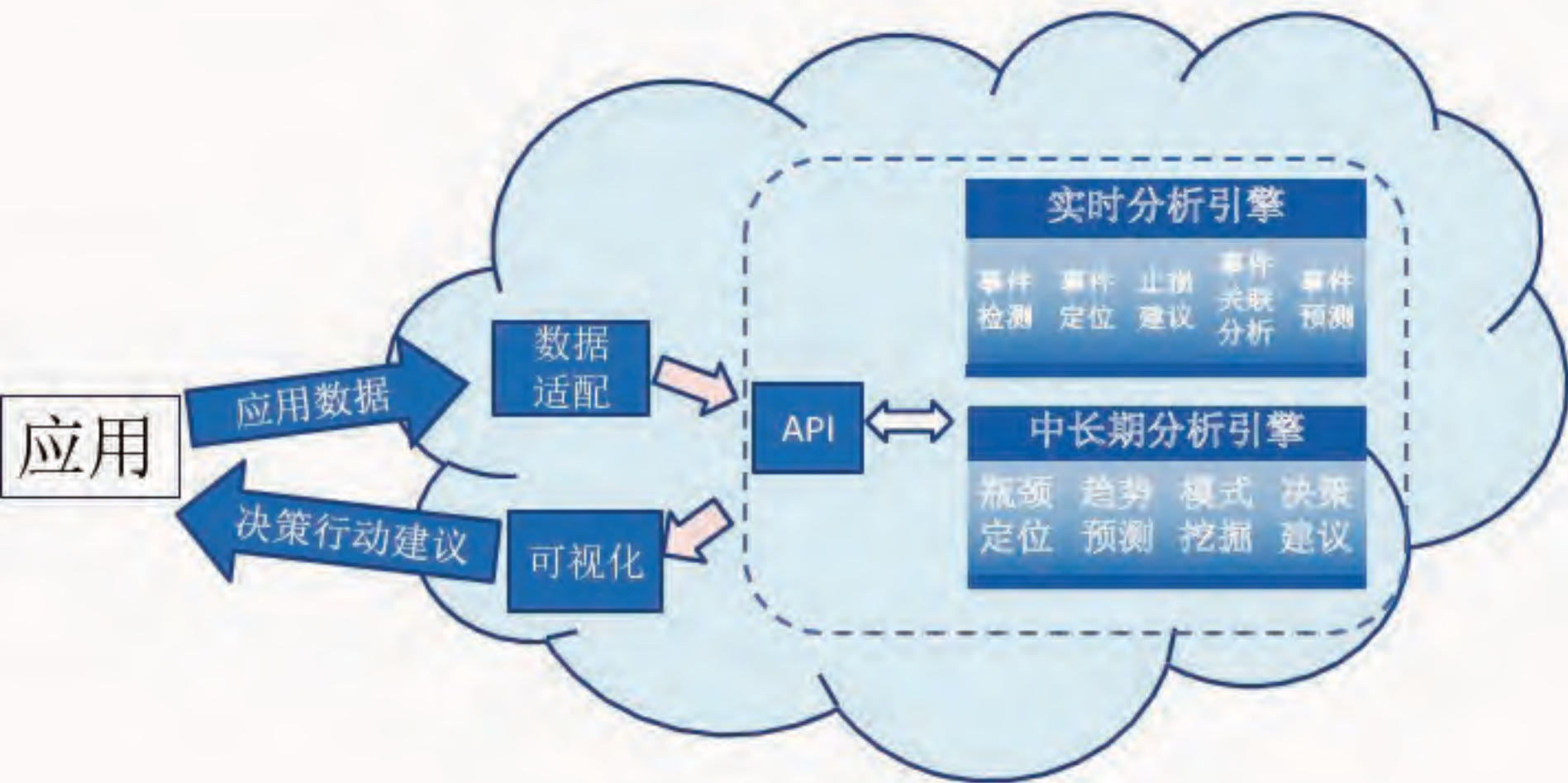
管理员人工定位 :

- 1.5 小时检测完成问题定位
- 客户申述->检查KPI->定位问题



- 异常检测之后的故障定位
- 故障止损建议
- 故障根因分析
- 数据中心交换机故障预测
- 海量Syslog 日志压缩成少量有意义的事件
- 基于机器学习的系统优化（如TCP运行参数）

AppMind 智能运维算法云： 把数据转化为决策和行动



标准API：支持任意时序数据

时间戳	关键指标	属性1	属性2	...	属性n
-----	------	-----	-----	-----	-----

销售额、利润、订单数、PV、转化率、用户数、用户增速、留存率、首屏时间、闪退率、投诉率...

- 背景介绍
- 智能运维：从基于规则到基于学习
- 百度案例
- *挑战与思路*

挑战1：智能运维的可行目标是什么？

T2: 代替运维人员，接管所有工作？

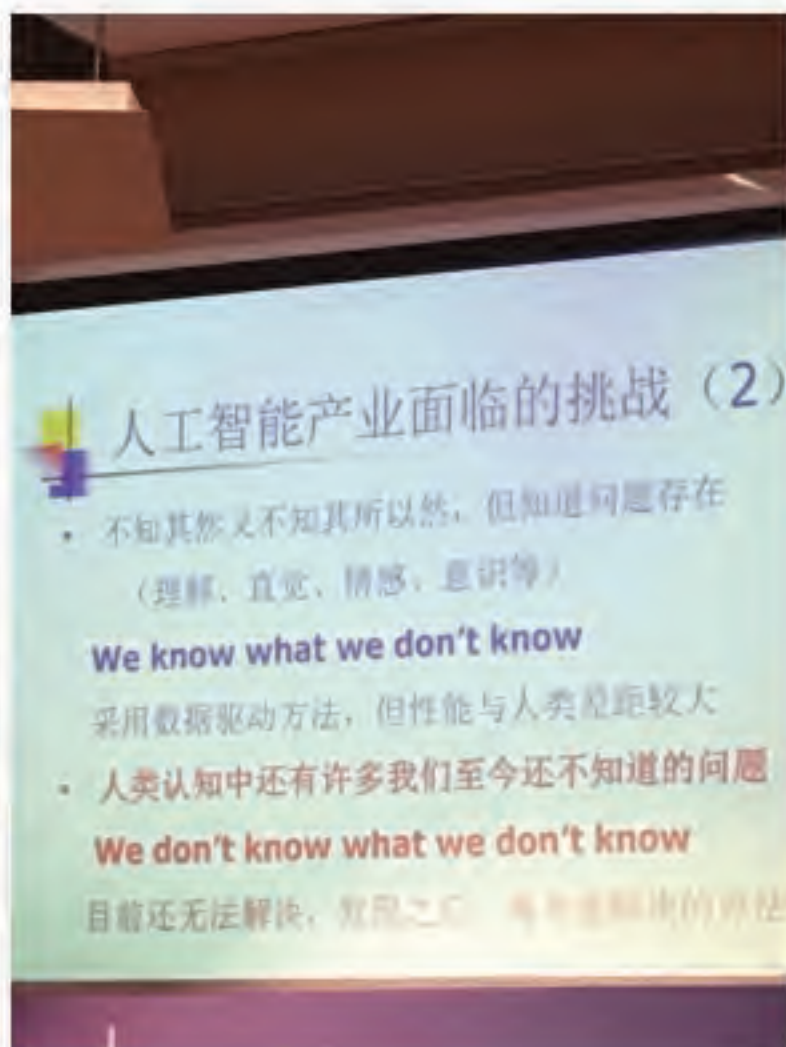
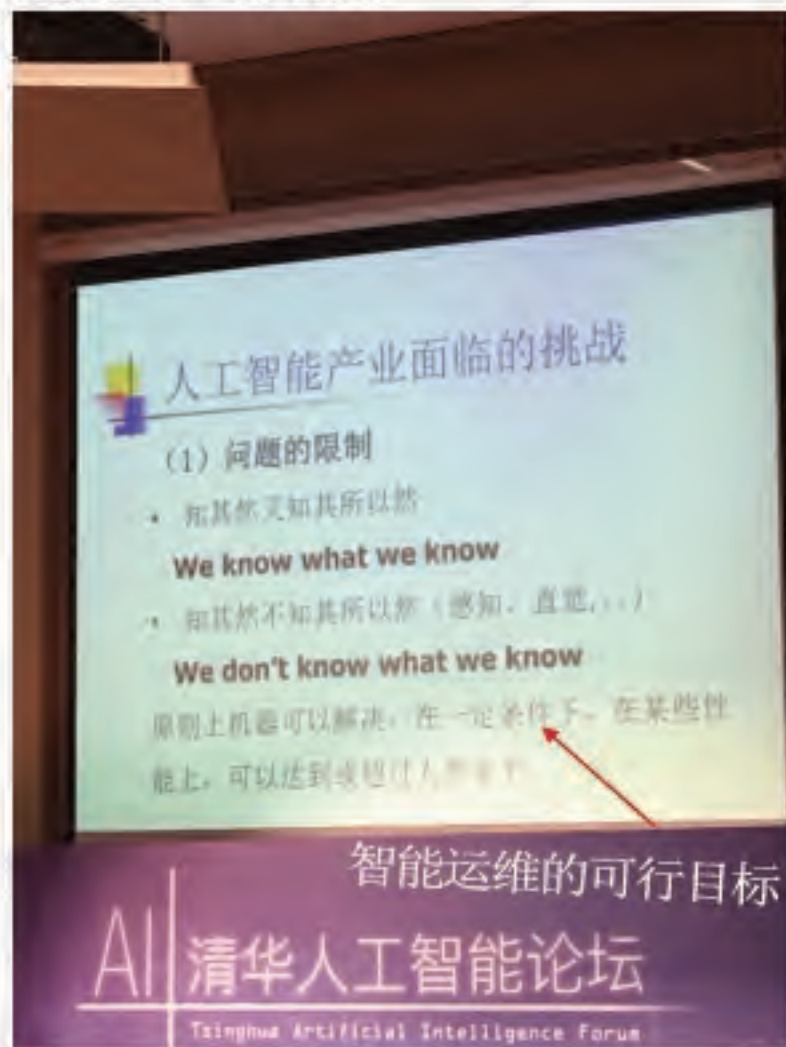


R2-D2: 运维人员的高效可靠助手？



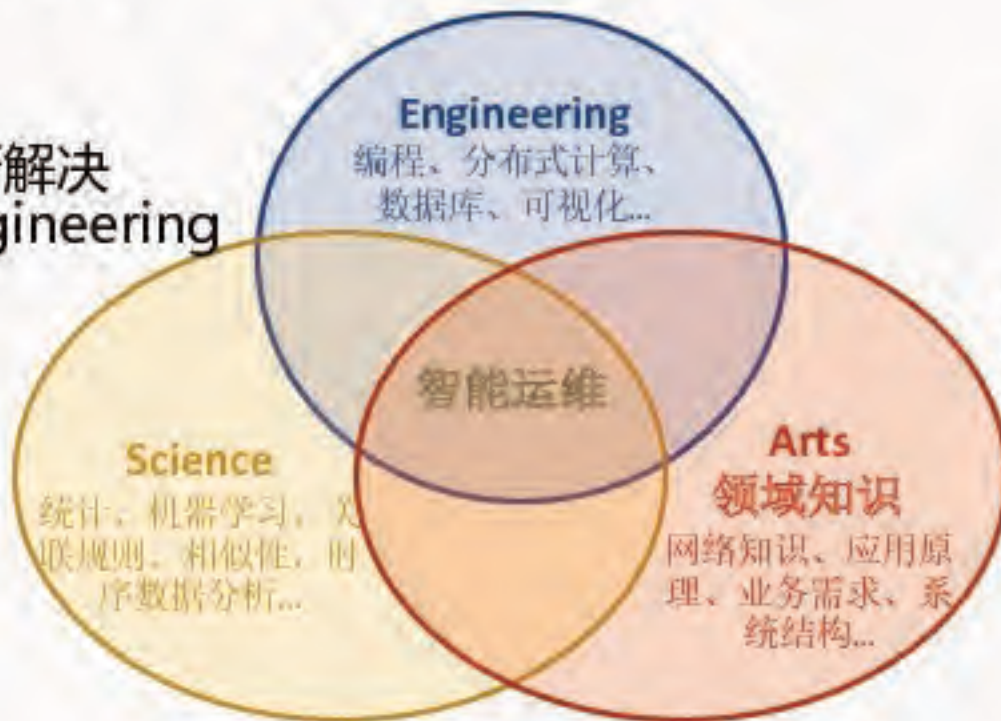
图片来自互联网

思路：自动化那些“知其然而不知其所以然”的运维技能



思路：自动化那些“知其然而不知其所以然”的运维任务

技术正在逐渐解决
Science+Engineering
的问题



技术可能永远也无法代替领域专家（艺术家），但是可以为领域专家提供更好的工具

智能运维的终极可行目标

1. 日常工作都能自动完成
2. 运维人员能够独立进行数据分析

挑战2：如何更系统的应用机器学习技术？

- 特征选取：
 - 全部数据+容忍度高的算法（如随机森林）
 - 特征工程
 - 自动选取（深度学习）
- 不同机器学习算法适用不同的问题
 - *Tree-based*: 决策树, 回归树, 随机森林等
 - *Convolutional Neural Networks*
 - *Recurrent Neural Networks*
 - *Deep Belief Network*
 - *Monte Carlo Tree Search*
- 有效策略：工业界和学术界针对具体问题
进行密切合作

挑战3：如何从现有ticket 数据中提取有价值信息

Strawman Approach To Analyze Free-form Text

UNSTRUCTURED (Diary)

Operator 1: **replaced** the **memory chips** on this **device** and both **power supplies** have been **reseated**

Operator 2: The **device** has been **powered back up**. It should be back online shortly.

Operator 2: Ok. Let me check.

Operator 1: Yes. It is functional. Thank!

— Original Message —

From: Vendor Support

Subject: Regarding Case Number #yyyyyy

Title: **Device** xxx-xxx-xxx-1300 v9.4 is **continuously rebooting**

As discussed, the device has **bad memory chips** as such we **replace** it. Please

completely fill the **RMA** form below and return it.

— Appended Message —

From: Operations

Subject: Regarding Case Number #yyyyyy

Title: **Device** xxx-xxx-xxx-1300 v9.4 is **continuously rebooting**

We have **cleaned** the **cable** connecting the **load balancer** to the **access router** so don't **replace** the cable. We are currently checking for on-going

maintenance. Please invoke **device diagnostics** and send the logs to the **vendor** for further **troubleshooting**.

Strawman #1: Use NLP techniques

Limitation: Work only on well-written text such as news-articles.

Strawman #2: Keyword selection

Limitations: Ignores contextual semantics

Strawman #3: Clustering tickets based on manual keyword selection

Limitations: 1. Significant time and effort to build the keyword list
2. Limited coverage or risks becoming outdated as the network evolves

思路：把ticketing系统作为智能运维的一部分 APMCon
来设计

- Ticket 格式、系统的设计都应该是智能运维的工作的一部分
- ticket需要包含足够的信息以供机器学习使用
- ticket系统要向互联网产品一样简洁易用
- 运维人员要自律并认真填写ticket
- 开发工具自动分析ticket自由文本部分

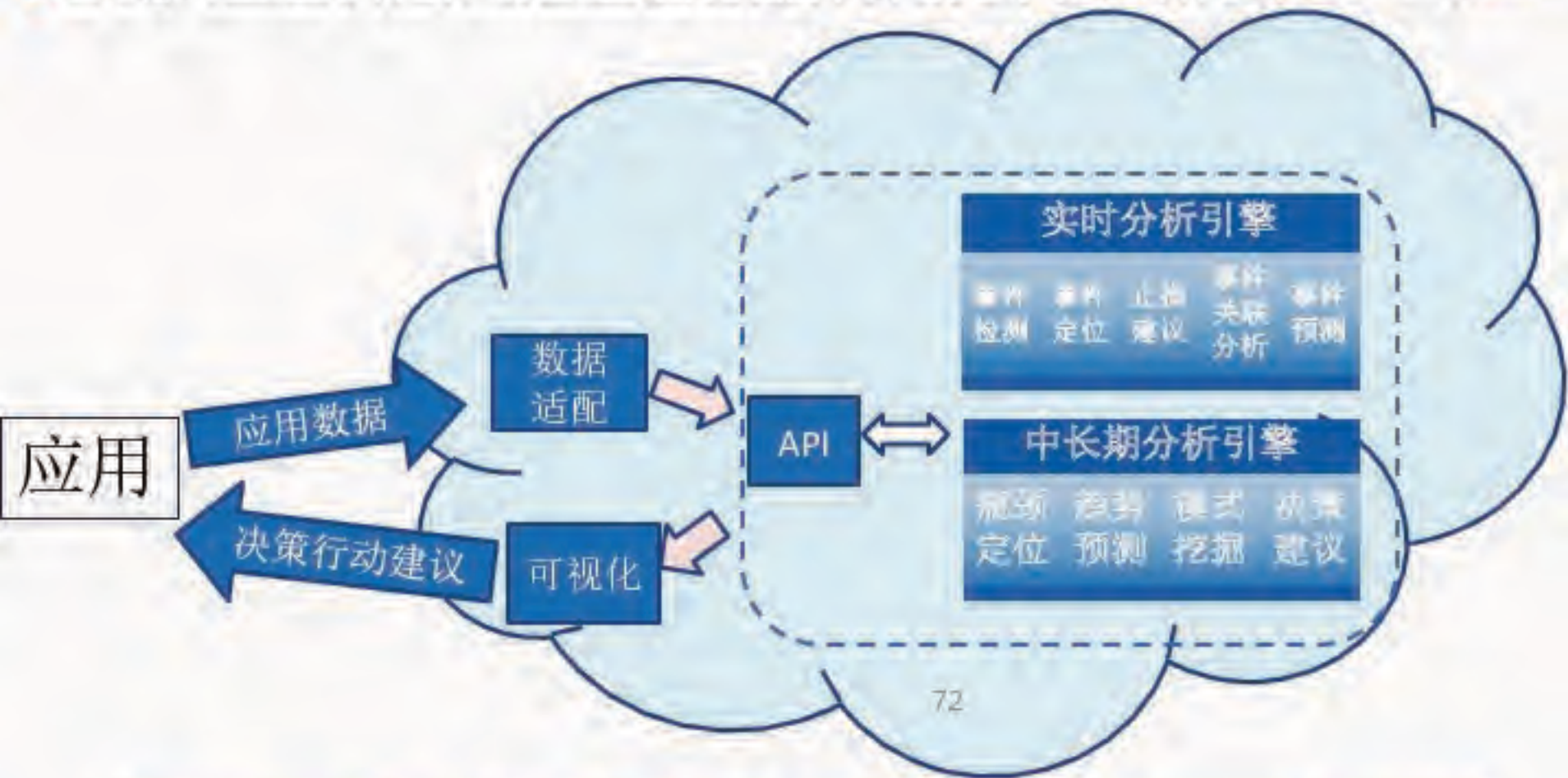
挑战4：如何把智能**运维**延伸到智能**运营**？

企业痛点：

运营数据与精准决策/行动之间的巨大鸿沟



思路: 通过算法云把运营数据转化为决策和行动 APMCon



时间戳	关键指标	属性1	属性2	...	属性n
-----	------	-----	-----	-----	-----

销售额、利润、订单数、PV、转化率、用户数、用户增速、留存率、首屏时间、闪退率、投诉率...

- 基于机器学习的智能运维在今后若干年会飞速发展
 - 得天独厚的数据、标注和应用
- 智能运维的终极可行目标: 运维人员高效可靠的助手
 - 日常工作都能自动完成
 - 运维人员能够独立进行数据分析
- 智能运维应更系统应用机器学习技术
 - 工业界与学术界应在具体问题上密切合作
- 更系统的数据采集和标注会帮助智能运维更快发展
- 从智能运维延伸到智能运营

THANK YOU

Email: peidan@tsinghua.edu.cn

微信: peidanwechat

<http://netman.cs.tsinghua.edu.cn>

《高等网络管理》课件:

<http://netman.cs.tsinghua.edu.cn/courses/advanced-network-management-spring2016/>

Many thanks to Baidu Search & OP team, and the entire Tsinghua NetMan team