Intelligent Operations through Machine Learning

Dan Pei

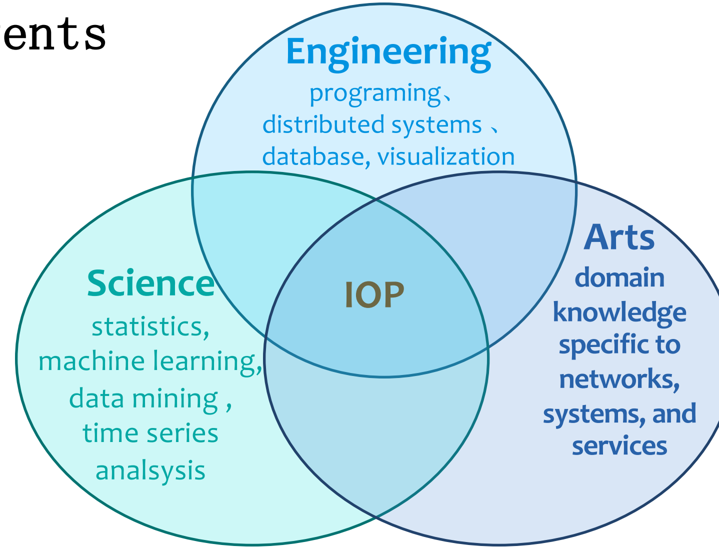
Tsinghua University

Introduction to Intelligent Operations (IOP)

In order to handle a large volume of events (QoE, performance, reliability, and security) in the distributed systems in the Internet, IOP does the following:

- Reconstruct and diagnose past events accurately
- Diagnose, detect and mitigate ongoing events
- Predict important future events.

IOP is at the intersection of engineering, science, and domain knowledge



IOP is critical to the Internet

"AI is the most important tool for managing the networks"

- Huawei Chairman Zhenfei Ren, Internal Speech in August 2016.

一、巨大的存量网络是人工智能最好的舞台

为什么要聚焦GTS、把人工智能的能力在服务领域先做好呢？对于越来越庞大、越来越复杂的网络，人工智能是我们建设和管理网络的最重要的工具，人工智能也要聚焦在服务主航道上，这样发展人工智能就是发展主航道业务，我们要放到这个高度来看。如果人工智能支持GTS把服务做好，五年以后我们自已的问题解决了，我们的人工智能又是世界一流。

首先，是解决我们在全球巨大的网络存量的网络维护、故障诊断与处理的能力的提升。我们在全球网络存量有一万亿美元，而且每年上千亿的增加。容量越来越大，流量越来越快，技术越来越复杂，维护人员的水平要求越来越高，经验要求越来越丰富，越来越没有这样多的人才，人工智能，大有前途。

IOP is critical to the Internet

"AI is the most important tool for managing the networks"
- Huawei Chairman Zhenfei Ren, Internal Speech in August 2016.

一、巨大的存量网络是人工智能最好的舞台

为什么要聚焦GTS、把人工智能的能力在服务领域先做好呢？对于越来越庞大、越来越复杂的网络，人工智能是我们建设和管理网络的最重要的工具，人工智能也要聚焦在服务主航道上，这样发展人工智能就是发展主航道业务，我们要放到这个高度来看。如果人工智能支持GTS把服务做好，五年以后我们自已的问题解决了，我们的人工智能又是世界一流。

首先，是解决我们在全球巨大的网络存量的网络维护、故障诊断与处理的能力的提升。我们在全球网络存量有一万亿美元，而且每年上千亿的增加。容量越来越大，流量越来越快，技术越来越复杂，维护人员的水平要求越来越高，经验要求越来越丰富，越来越没有这样多的人才，人工智能，大有前途。

"In addition to control plane and data plane, Internet needs an AI-based knowledge plane"
- Dave Clark, the Architect of the Internet, in his SIGCOMM 2003 paper.

A Knowledge Plane for the Internet

David D. Clark*, Craig Partridge♦, J. Christopher Ramming† and John T.

*M.I.T Lab for Computer Science
200 Technology Square
Cambridge, MA 02139
{ddc,jtw}@lcs.mit.edu

♦BBN Technologies
10 Moulton St
Cambridge, MA 02138
craig@bbn.com

†SRI
333 Rav
Menlo Pa
chrisramm

ABSTRACT
We propose a new objective for network research: to build a fundamentally different sort of network that can assemble itself given high level instructions, reassemble itself as requirements change, automatically discover when something goes wrong, and automatically fix a detected problem or explain why it cannot do so.

We further argue that to achieve this goal, it is not sufficient to improve incrementally on the techniques and algorithms we know today. Instead, we propose a new construct, the Knowledge Plane, a pervasive system within the network that builds and maintains high-level models of what the network is supposed to do, in order to provide services and advice to other elements of the network. The knowledge plane is novel in its reliance on the tools of AI and cognitive systems. We argue that cognitive techniques, rather than traditional algorithmic approaches, are best suited to meeting the uncertainties and complexity of our objective.

transparent network with rich end-sy
deeply embedded assumption of
administrative structure are critical stre
users when something fails, and high
much manual configuration, diagnosis a

Both user and operator frustrations aris
design principle of the Internet—the
with intelligence at the edges [1,2].
without knowing what that data is, or
combination of events is keeping dat
edge may recognize that there is a prob
that something is wrong, because the c
be happening. The edge understands
expected behavior is; the core only de
network operator interacts with the core
as per-router configuration of routes a
for the operator to express, or the netw

IOP problems have been a hot research topic for decades

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 Karp, University College London, UK

A top conference that is almost dedicated to OP problems.

Sponsored by ACM SIGCOMM and ACM SIGMETRICS

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

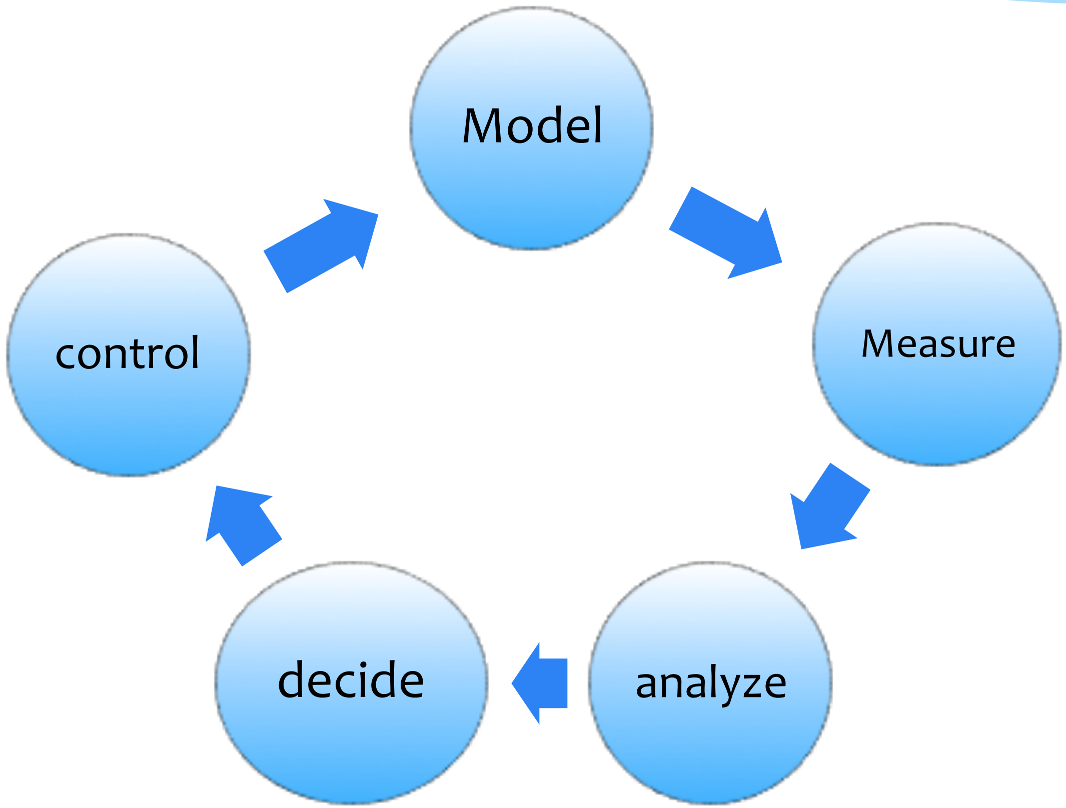
The Internet Measurement Conference (IMC) is a highly selective venue for the presentation of measurement-based research in data communications. The focus of IMC 2015 will be on papers that either (1) improve the practice of measurement or (2) illuminate some facet of an operational network. IMC takes 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)

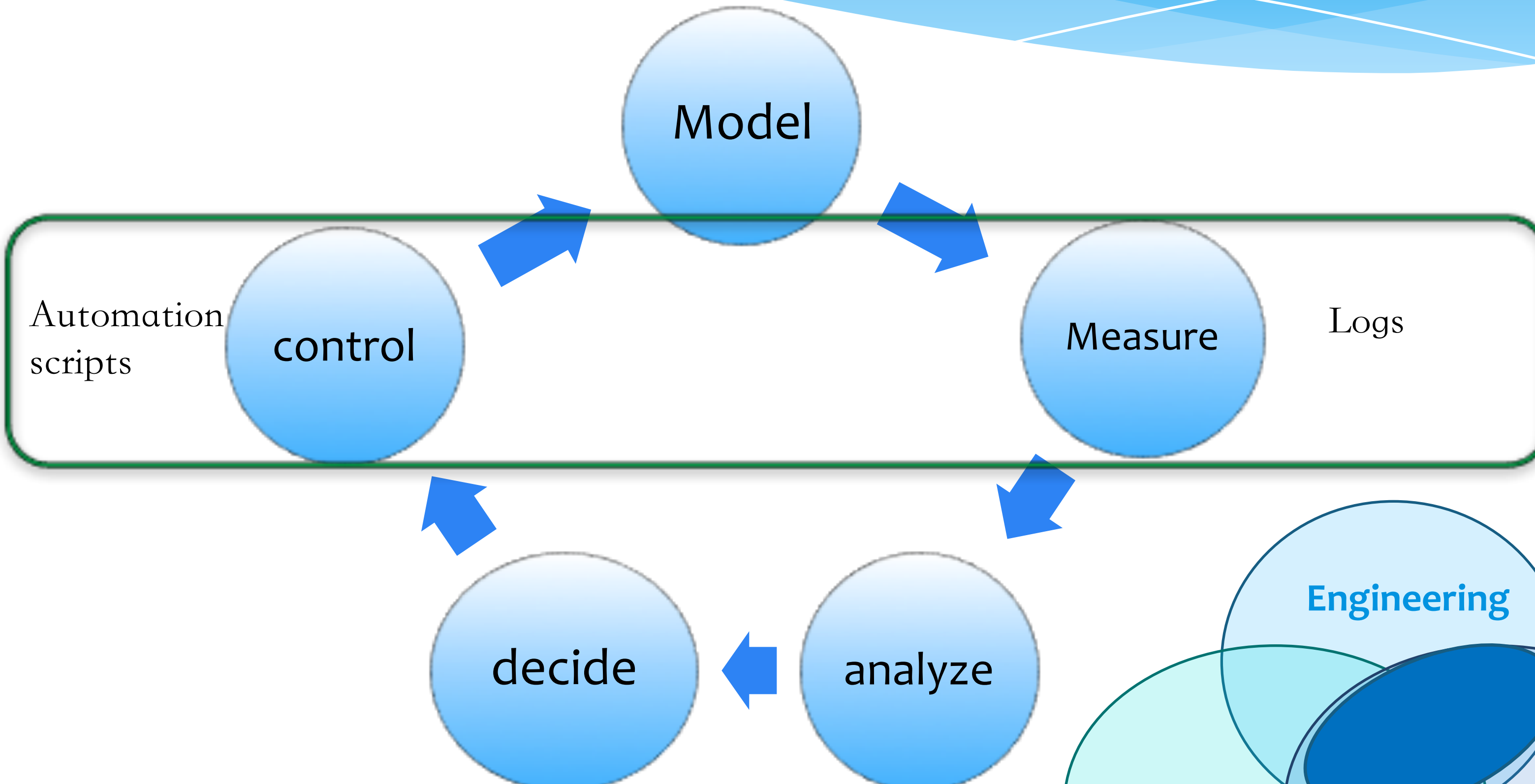
Types of contributions that the program committee would enjoy receiving submissions regarding include (but are not limited to):

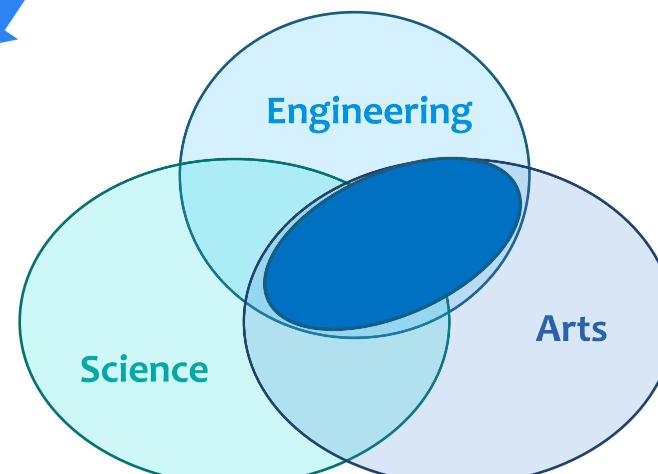
- collection and analysis of data that yield new insights about network structure and behavior
- methods and tools to monitor and visualize network-based phenomena
- systems and algorithmic techniques that leverage measurement-based findings in novel ways
- advances in data collection and handling (e.g.,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

IOP architecture

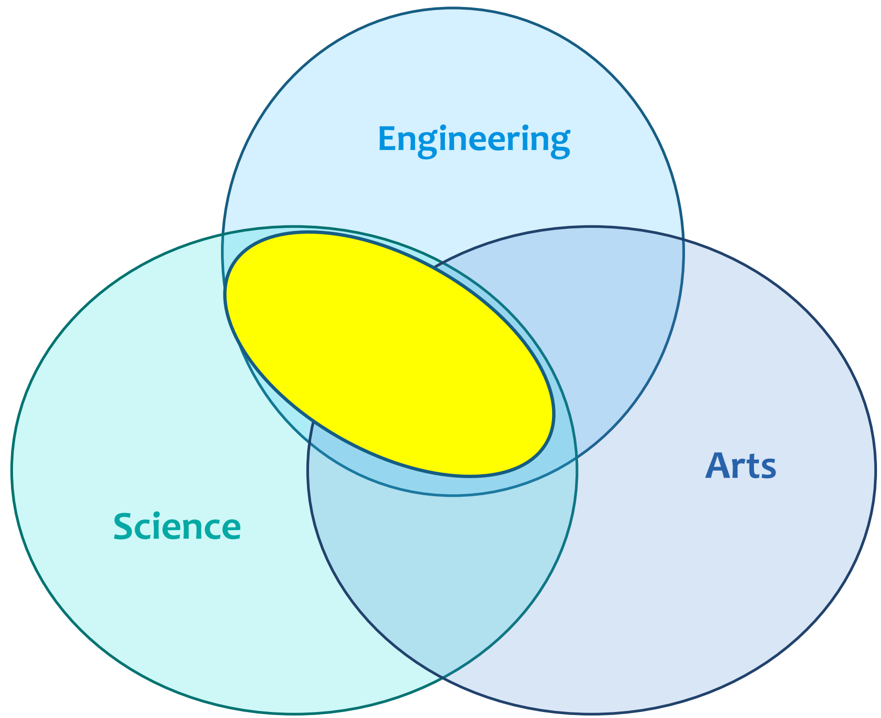


IOP architecture: rule-based

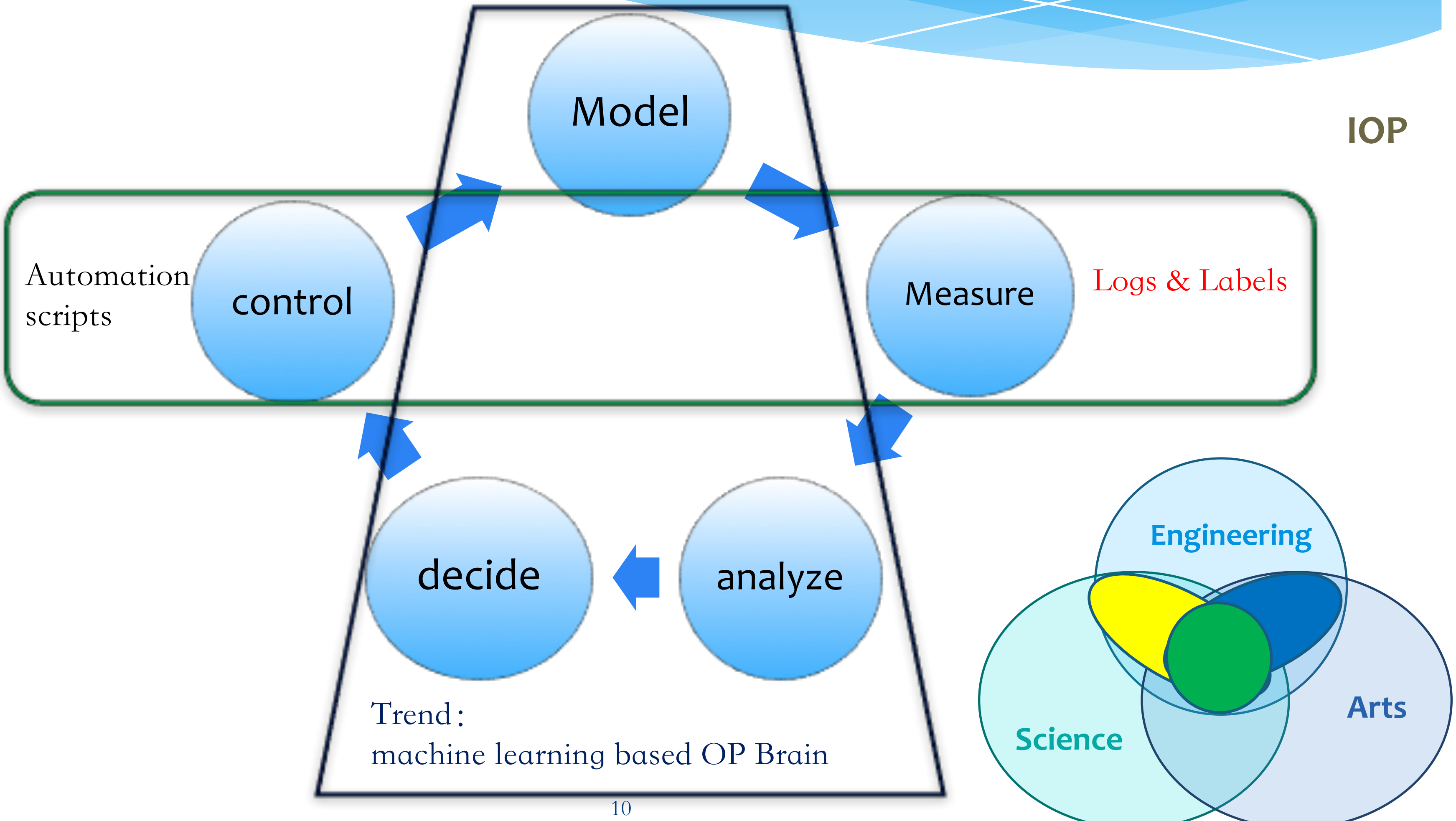




Machine learning tools (algorithms and systems)



IOP architecture: machine learning based



My life as an operator

My Official Resume

2000-2005 UCLA Ph.D., Best Ph.D. Thesis, working on BGP, OSPF etc.

Summer 2003, Intern at AT&T Research

2005-2011 Senior/Principal Researcher at AT&T Research

ACM, IEEE Senior Member

2012-now Associated Professor at Computer Science Department at Tsinghua University. “Expert of China Government's Global Talent Recruitment (Youth Program)” in 2012.

My Operator Resume

For five years, chased ISP OPs for data, experiences, and insights.

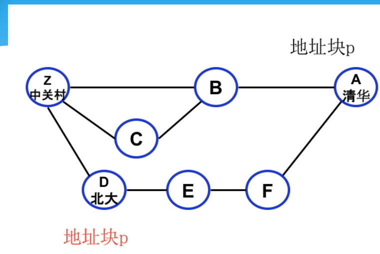
Felt in love with real OP data

Essentially a tier-5 OP

Worked on Performance, Reliability and Security.

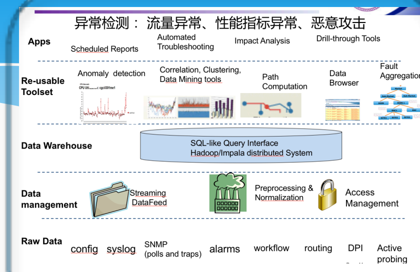
Teaching “Advanced Network Management. Almost all the projects are joint work with OPs at Baidu, Microsoft Azure, Petro China, Tsinghua Campus Network.

Our past and ongoing projects



IOP for BGP

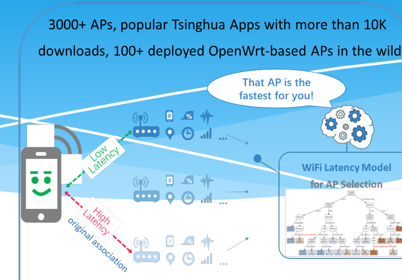
Global Routing



IOP for ISP

ISP

Access

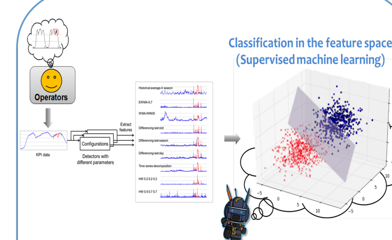


IOP for enterprise WiFi

Endpoint

IOP for mobile Apps

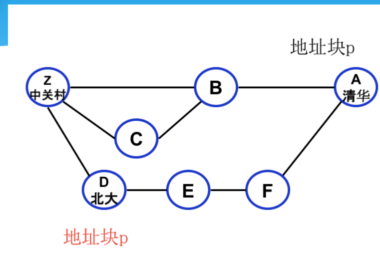
Cloud



OperRobot: an Operations Robot for managing networks and Internet applications

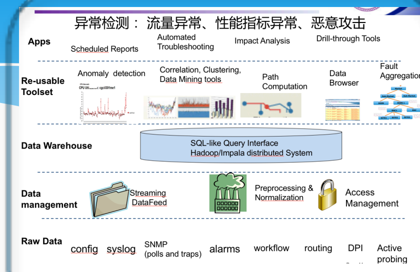
IOP for service providers

Our past and ongoing projects



IOP for BGP

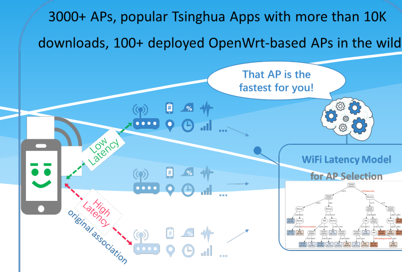
Global Routing



IOP for ISP

ISP

Access



IOP for enterprise WiFi

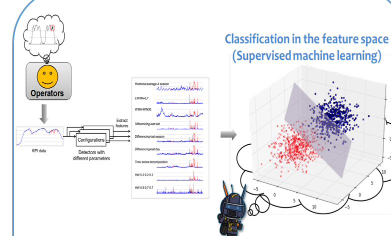
Endpoint

IOP for mobile Apps

AppMind engine

AppMind: Brain for Intelligent Operations

Cloud



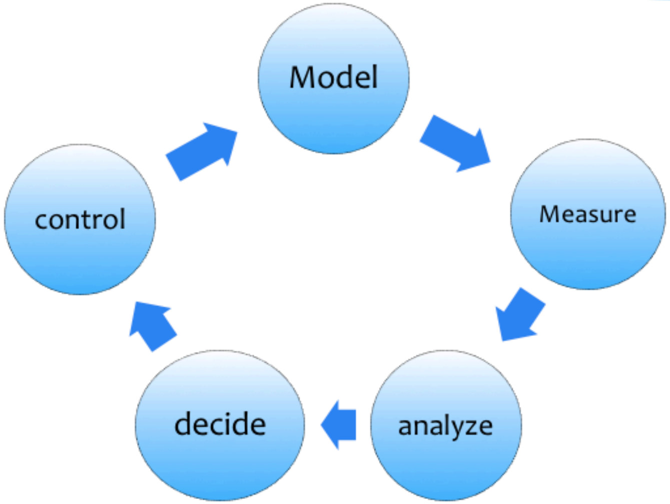
OperRobot: an Operations Robot for managing networks and Internet applications

IOP for service providers

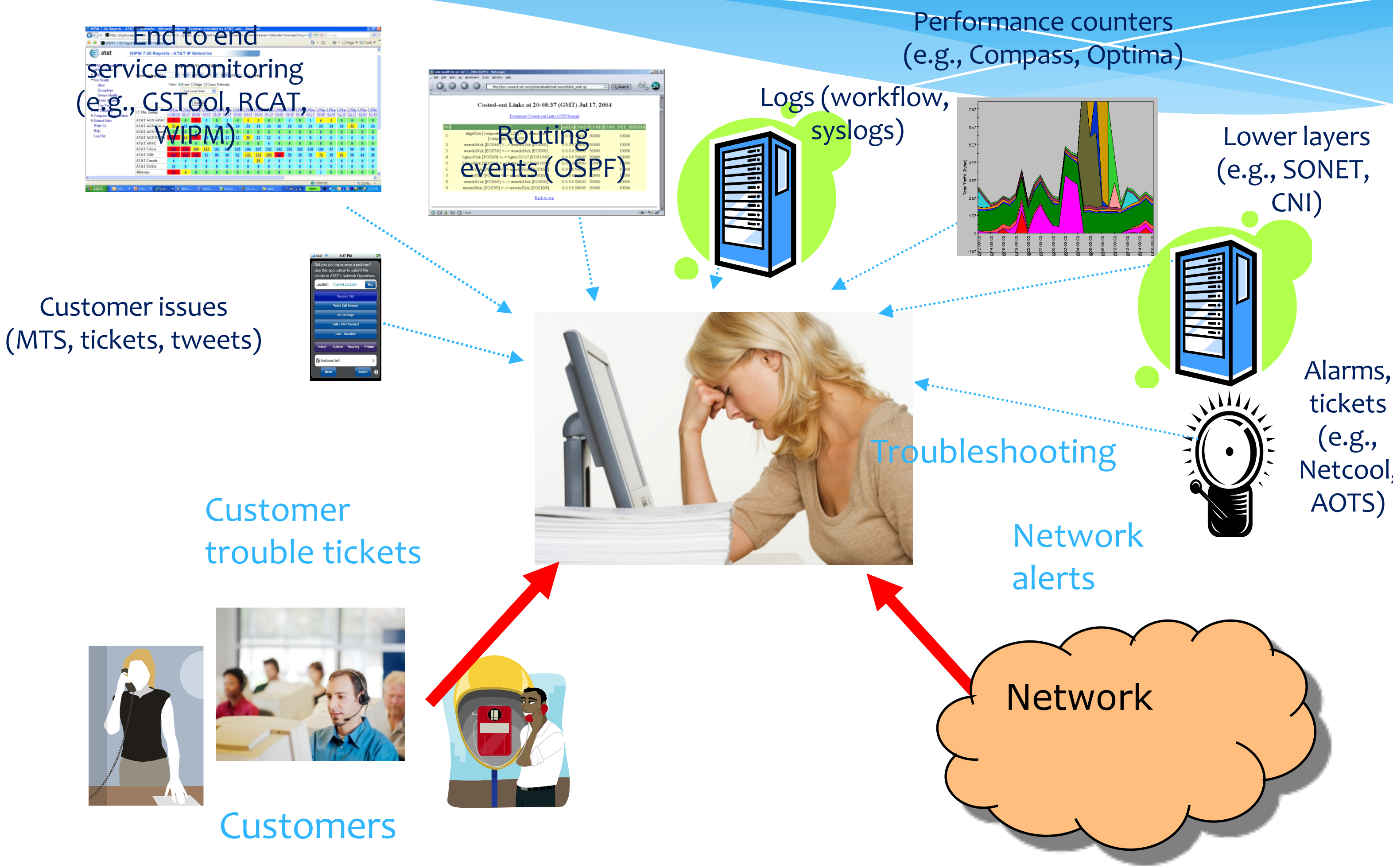
ML-based IOP will see rapid progress in the next few years

Have the necessities required for successful ML applications

* Machine learning tools (algorithms and systems)
* *Applications that show the value*
* *Large amount of data*
* *Labels and the experts who can label*



Data are abundant in OP, To make things better: new data can be generated by OP



OPs themselves are the experts who do the labels

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		

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 continously 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 continously 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.

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

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

Great process can be made in IOP in the next few years

Inherent Advantages of Intelligent Operations:

1. there are sheer amount of logs upon which features can be extracted for learning

2. the operators' daily actions can naturally serve as labels;

3. the learned model can be relatively easily integrated into production operations system.

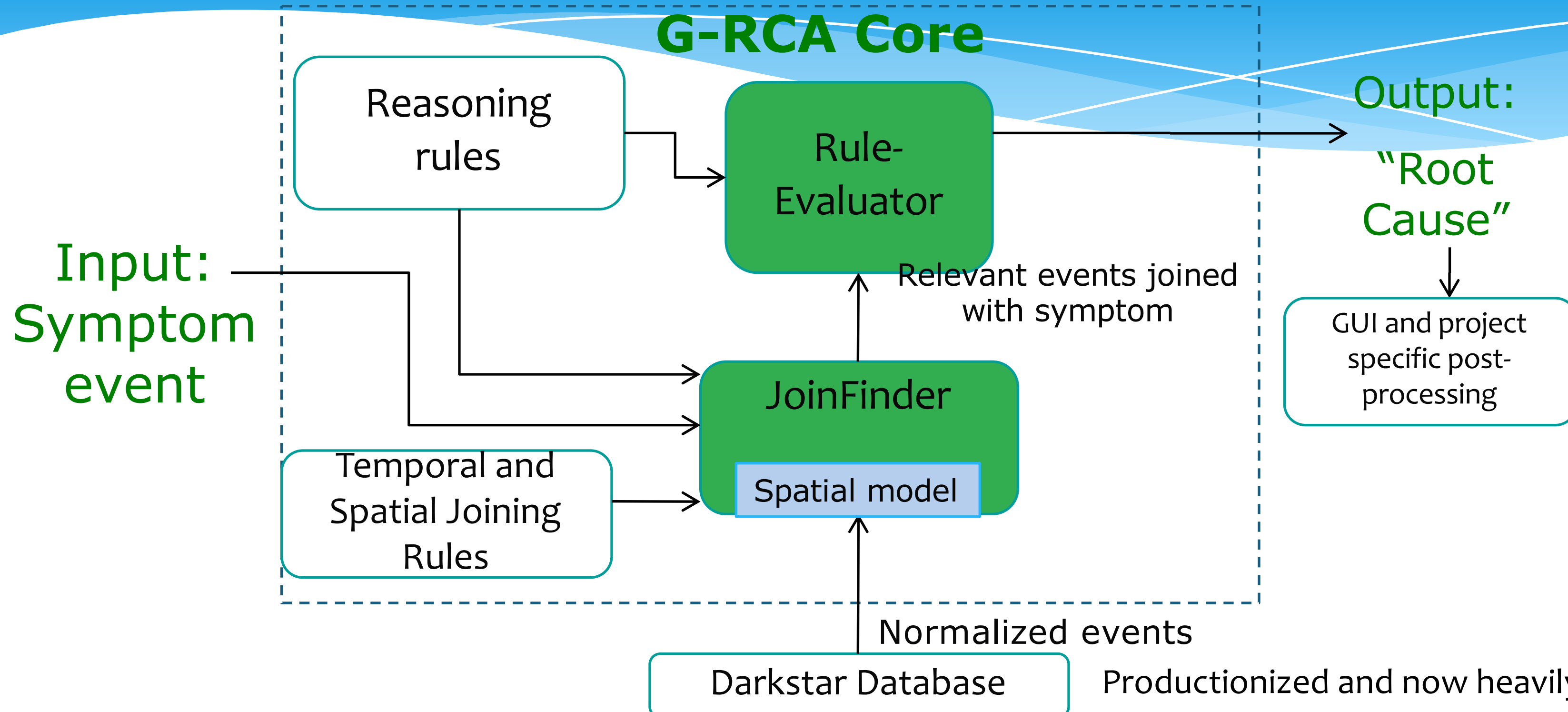
Outline

* Intelligent Operations: from “rule based” to “machine learning based”
* *Case Studies*
* Challenges and My thoughts

Case 1: Root Cause Analysis (Xiaohui Nie et al., IPCCC 2016)

Here is my personal (painful) journey from rule-based to learning-based OP intelligence….

G-RCA (Generic Root Cause Analysis) framework

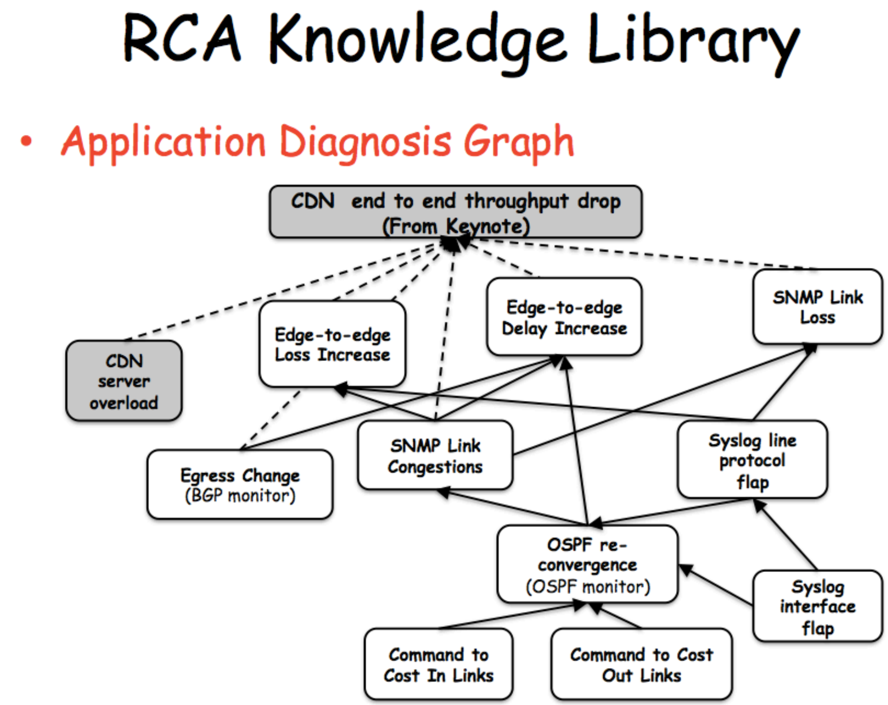


- Aiming at reusability: implement generic components only once:
  - A generic "language" to specify "rules" for joining and dependency
  - *JoinFinder* to find temporally and spatially joined events
  - *RuleEvaluator* to figure out the most likely root cause based on joined events

Productionized and now heavily used by AT&T operators.
Published in ACM **CoNEXT 2010**, and IEEE/ACM Transactions on Networking 2012.
Issued patents US #8,761,029; #8,411,577.
One CoNEXT 2010 reviewer commented: this tool **"revolutionizes troubleshooting Industry"**.

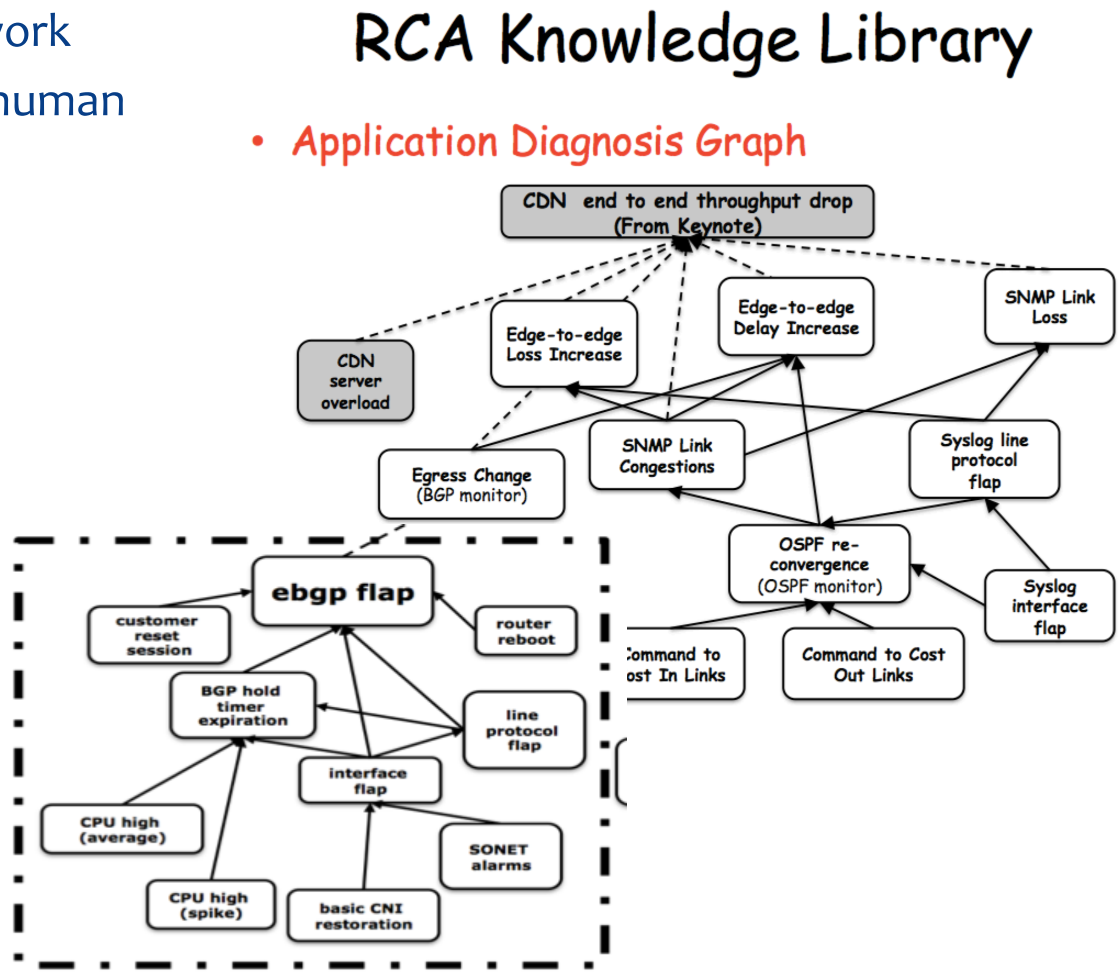
Rules in RCA

* Rule-based RCA framework
* Rule primarily given by human

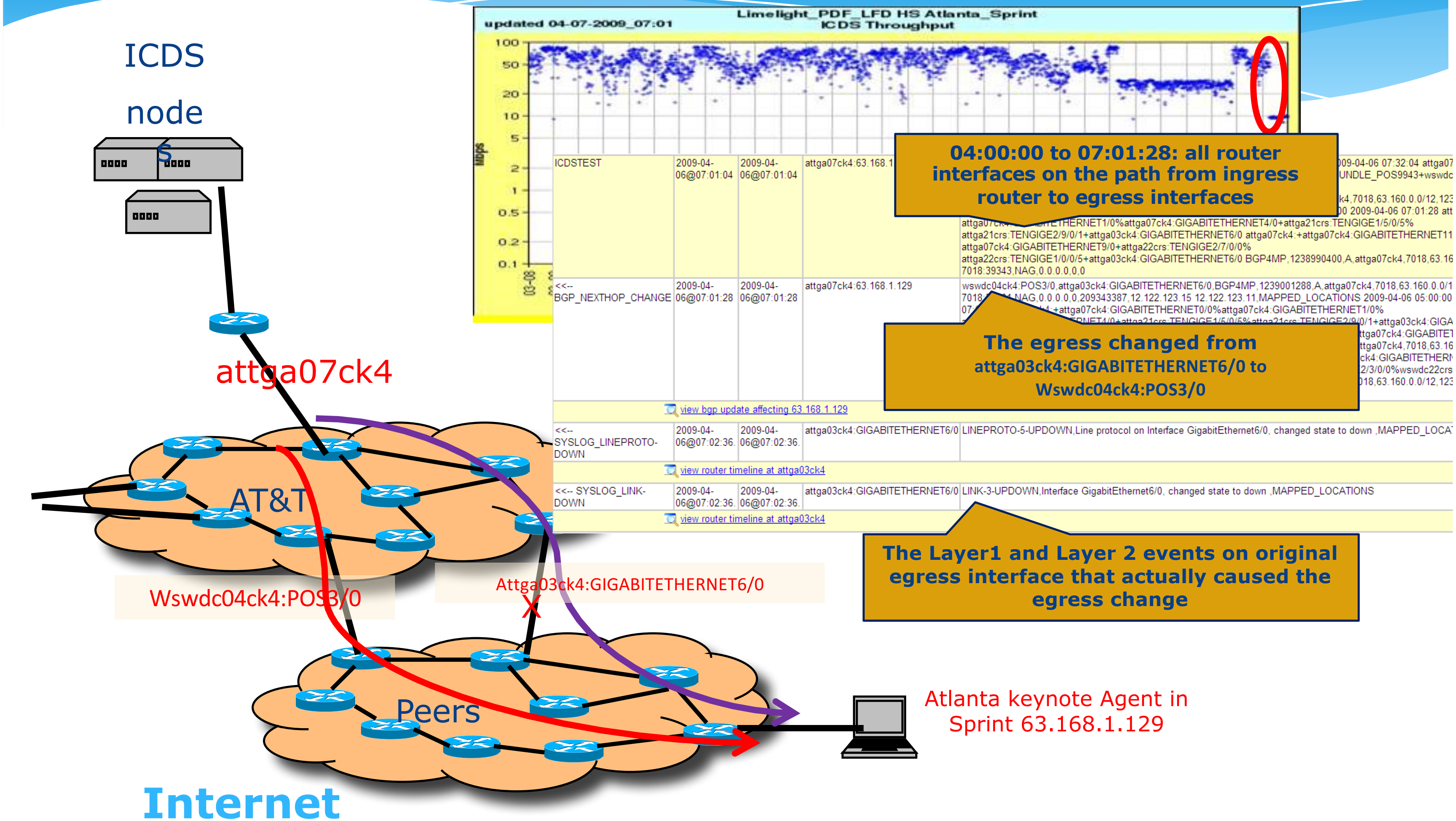


Rules in RCA

* Rule-based RCA framework
* Rule primarily given by human



Throughput drop between a router and dest ip



Reality Check at Baidu: infeasible to manually provide rules

- Scale
  - 100+ Internet-based services
  - 10k+ modules
  - 500+ thousand servers
  - Millions of KPIs are monitored
- Frequency of Changes:
  - 10k+ software changes per day
  - Developers come and go

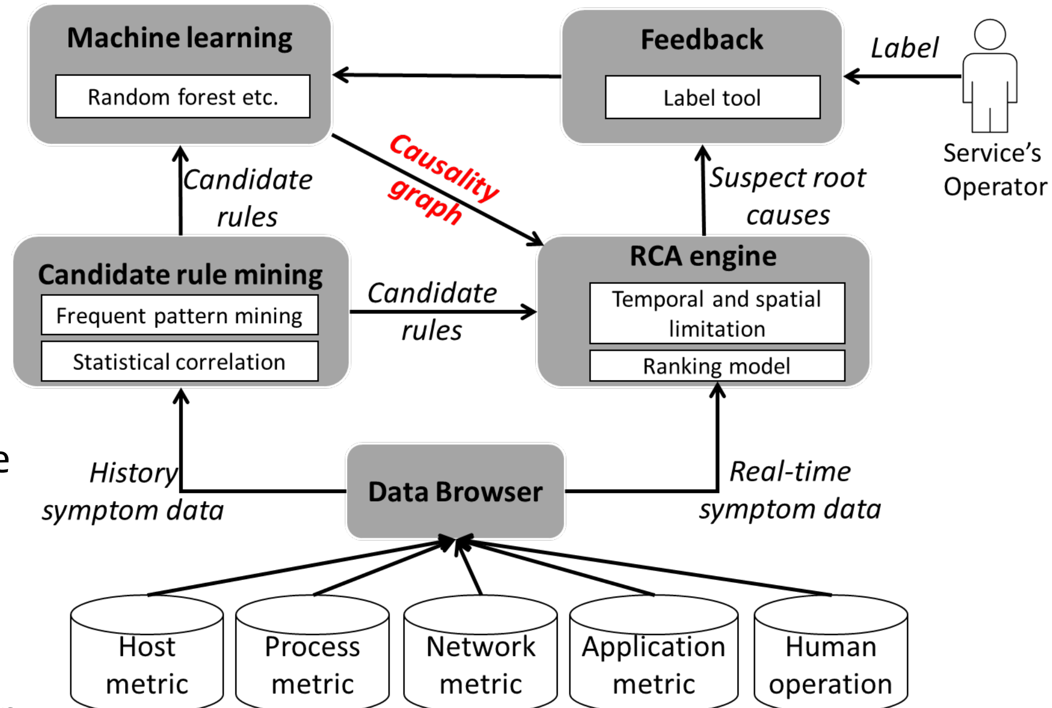
Machine Learning to the rescue: Automatically mining the dependency relationship between software modules

1. The key to the RCA problem is to build the dependency graph.

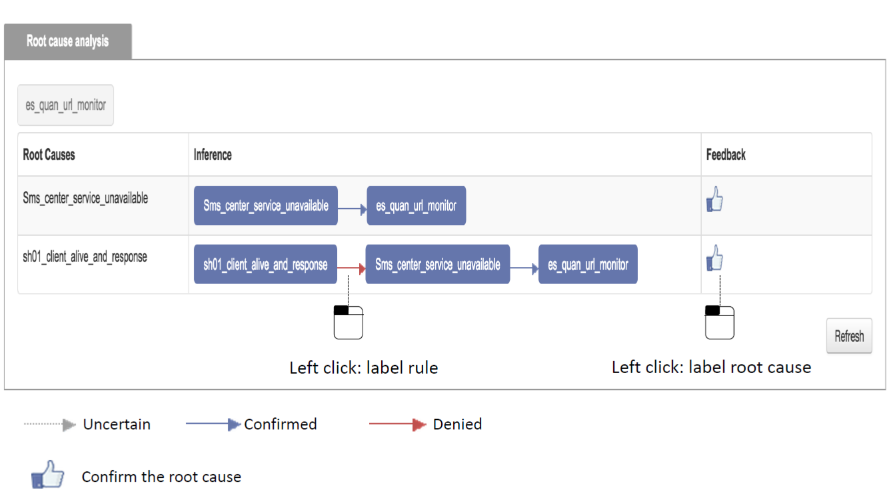
2. Challenge: dependency edges are distributed in the minds of many domain experts. How to make them collaborate to form a complete graph?

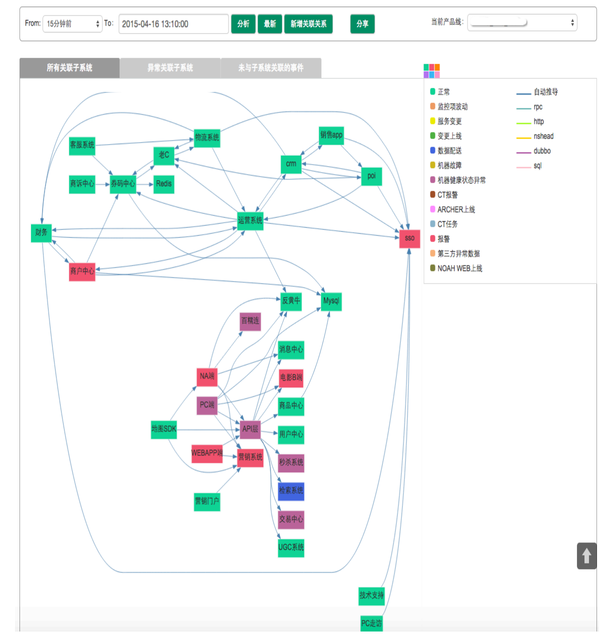
3. Our idea: use association mining to mine the rules (with mining parameters) , use the rules to provide a short list candidate root causes, operators label the candidate while browsing them.

4. Labels are used to train the algorithms which tune the parameters of association mining → supervised learning



Labeling is natural consequence when Operators use the RCA tool

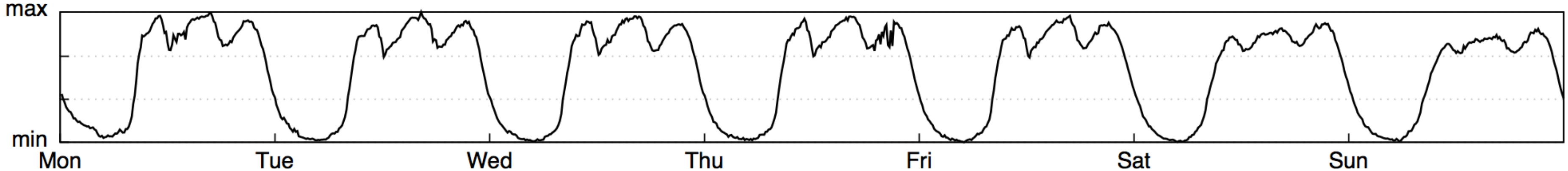




Localize the root cause in top 3 with 100% accuracy after a few rounds of learning.

Case 2： KPI Anomaly Detection (Dapeng Liu et al., IMC 2015)

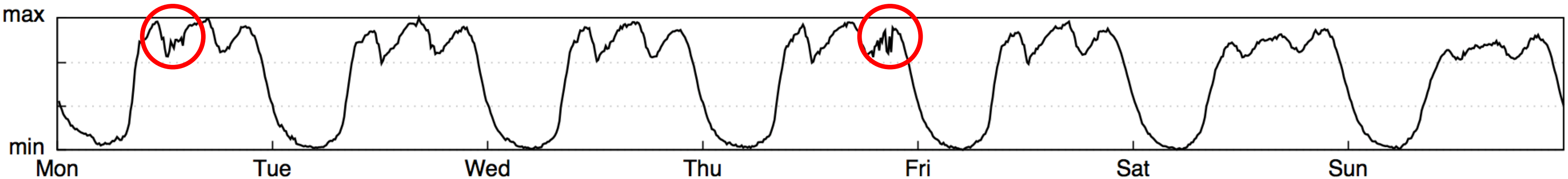
KPIs and Anomaly Detection



Page views (PV) of Baidu

KPIs (Key Performance Indicators): A set of performance measures that evaluate the service quality

KPIs and Anomaly Detection

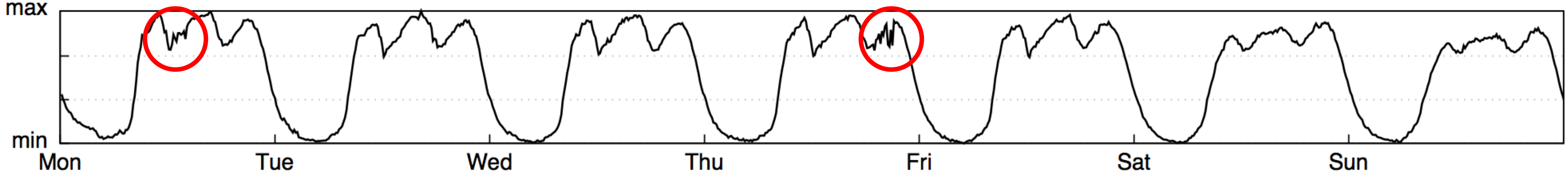


Page views (PV) of Baidu

KPIs (Key Performance Indicators): A set of performance measures that evaluate the service quality

KPI anomalous (unexpected) behaviors → Potential failures, bugs, attacks…

KPIs and Anomaly Detection



Page views (PV) of Baidu

KPIs (Key Performance Indicators): A set of performance measures that evaluate the service quality

KPI anomalous (unexpected) behaviors → Potential failures, bugs, attacks…

Anomaly detection matters: Find anomalous behaviors of the KPI curve

→ Diagnose and fix it

→ Avoid further influences and revenue losses

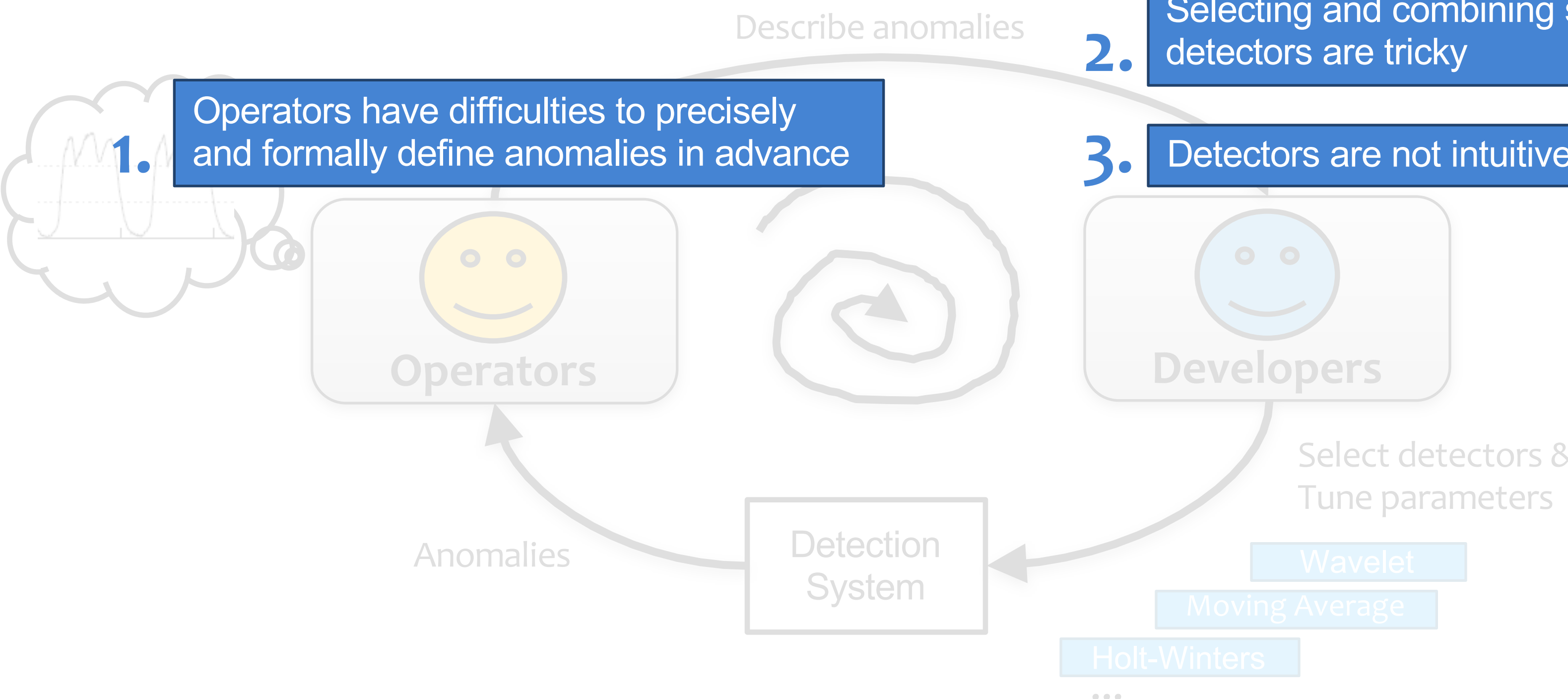
How to Build an Anomaly Detection System

Challenges

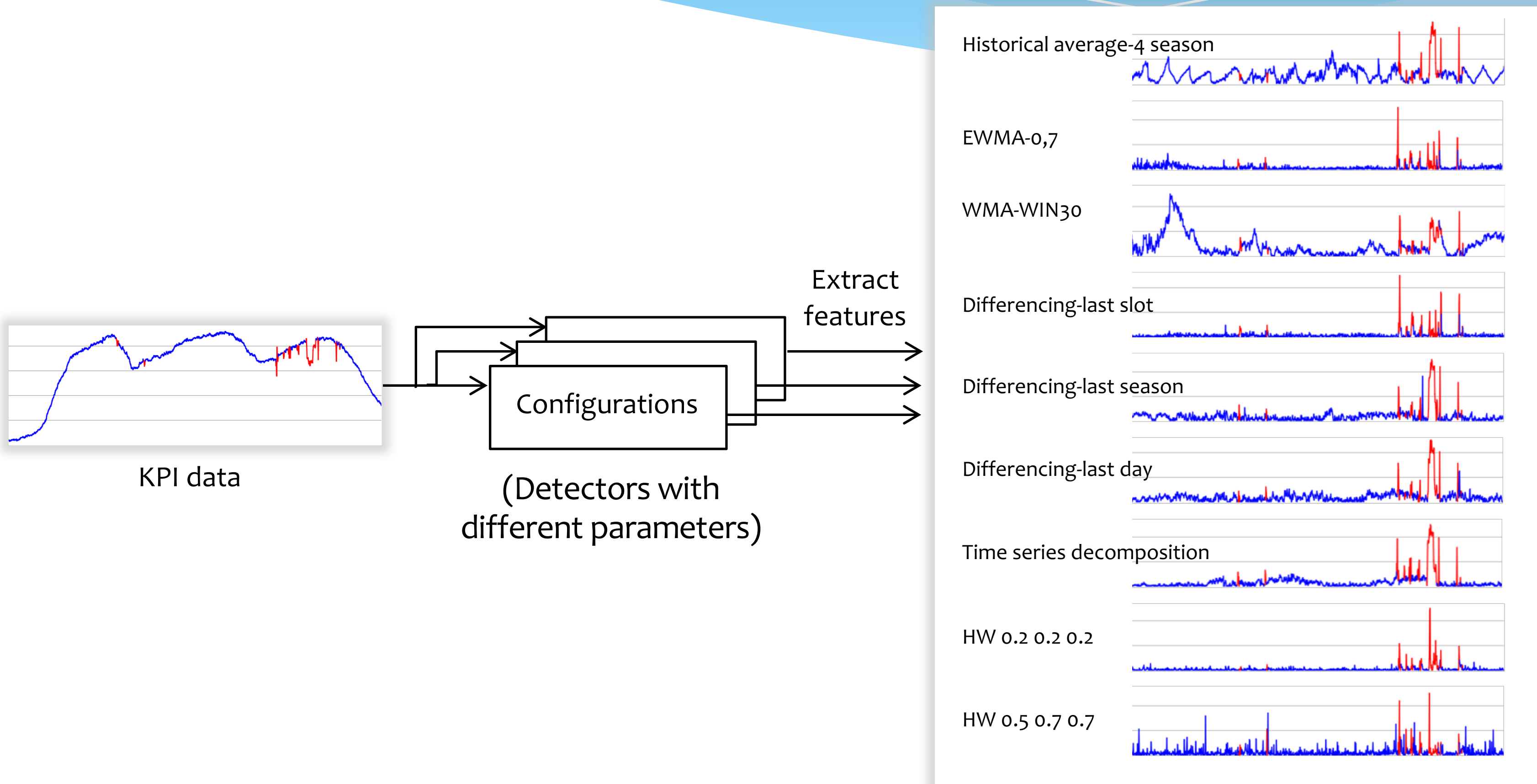
2. Selecting and combining suitable detectors are tricky

1. Operators have difficulties to precisely and formally define anomalies in advance

3. Detectors are not intuitive to tune

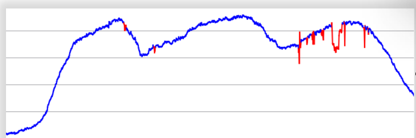


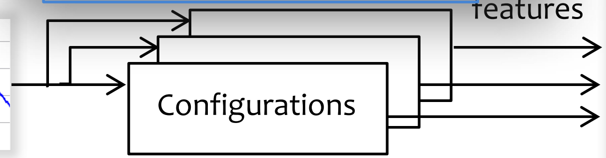
Key Ideas


KPI data
Configurations
(Detectors with different parameters)
Extract features
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

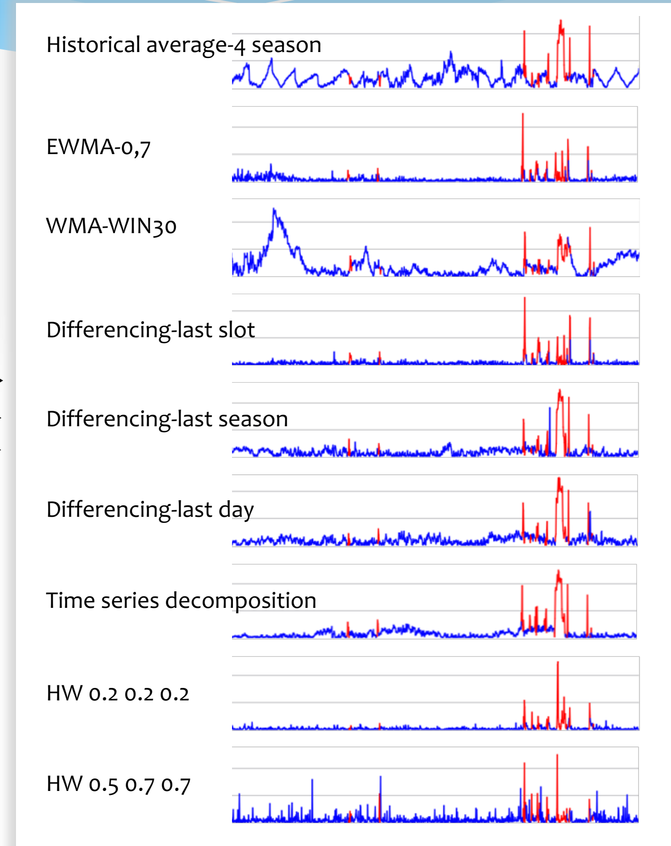
Key Ideas

Detector / # of configurations	Sampled parameters
Simple threshold [24] / 1	none
Diff / 3	last-slot, last-day, last-week
Simple MA [4] / 5	win = 10, 20, 30, 40, 50 points
Weighted MA [11] / 5	
MA of diff / 5	
EWMA [11] / 5	α = 0.1, 0.3, 0.5, 0.7, 0.9
TSD [1] / 5	win = 1, 2, 3, 4, 5 week(s)
TSD MAD / 5	
Historical average [5] / 5	
Historical MAD / 5	
Holt-Winters [6] / $4^3 = 64$	α, β, γ = 0.2, 0.4, 0.6, 0.8
SVD [7] / $5 \times 3 = 15$	row =10, 20, 30, 40, 50 points, column =3, 5, 7
Wavelet [12] / $3 \times 3 = 9$	win = 3, 5 ,7 days, freq = low, mid, high
ARIMA [10] / 1	Estimation from data
In total: 14 basic detectors / 133 configurations	


KPI data

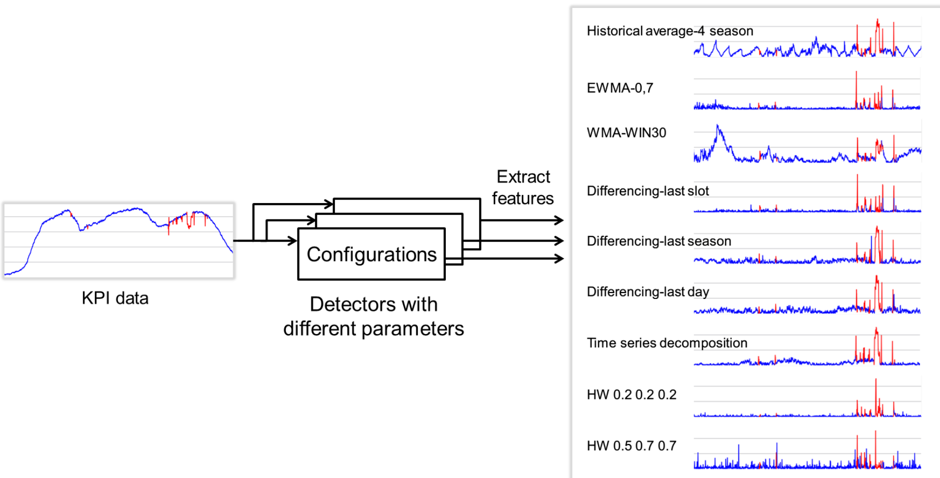


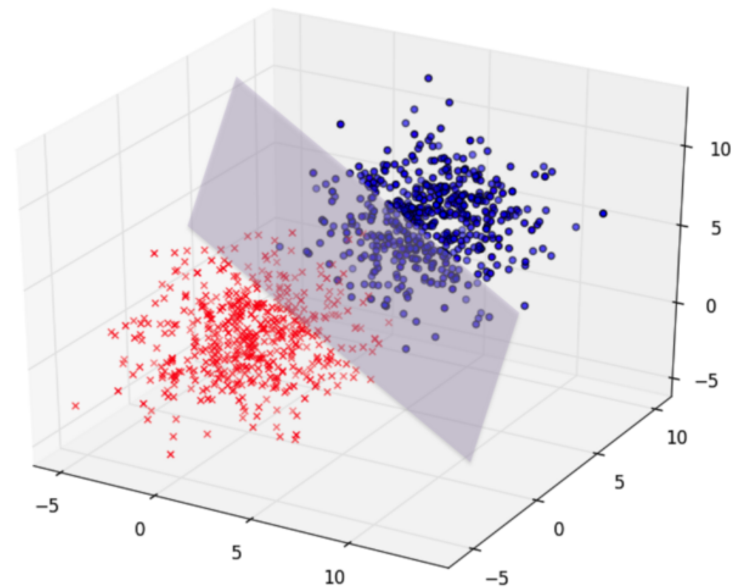
(Detectors with different parameters)



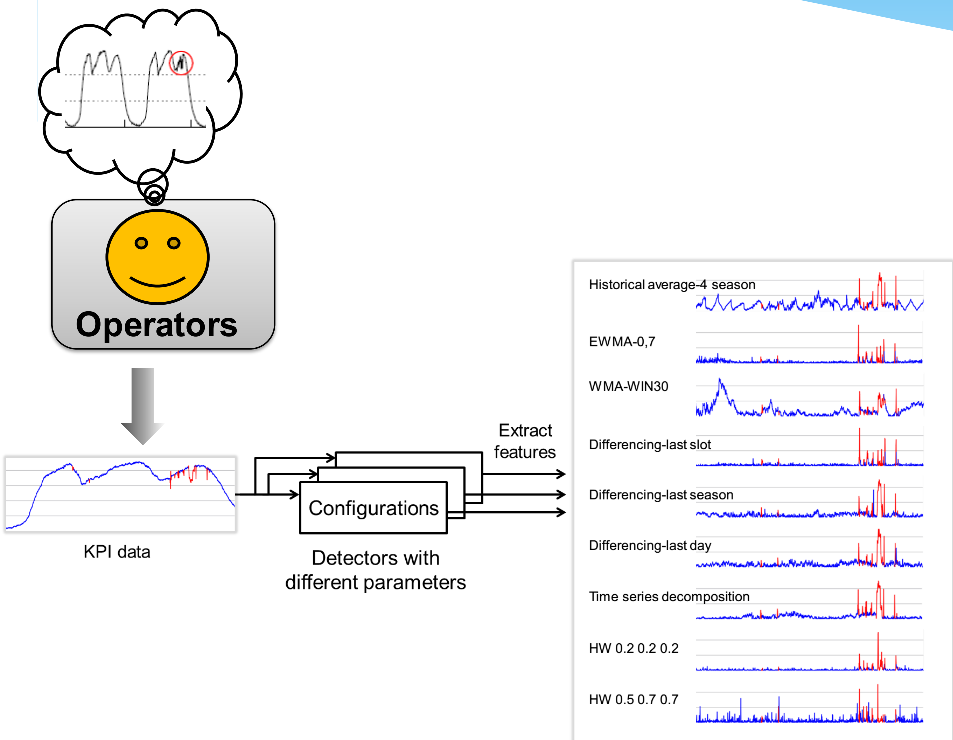
Key Ideas

Classification in the feature space (Supervised machine learning)

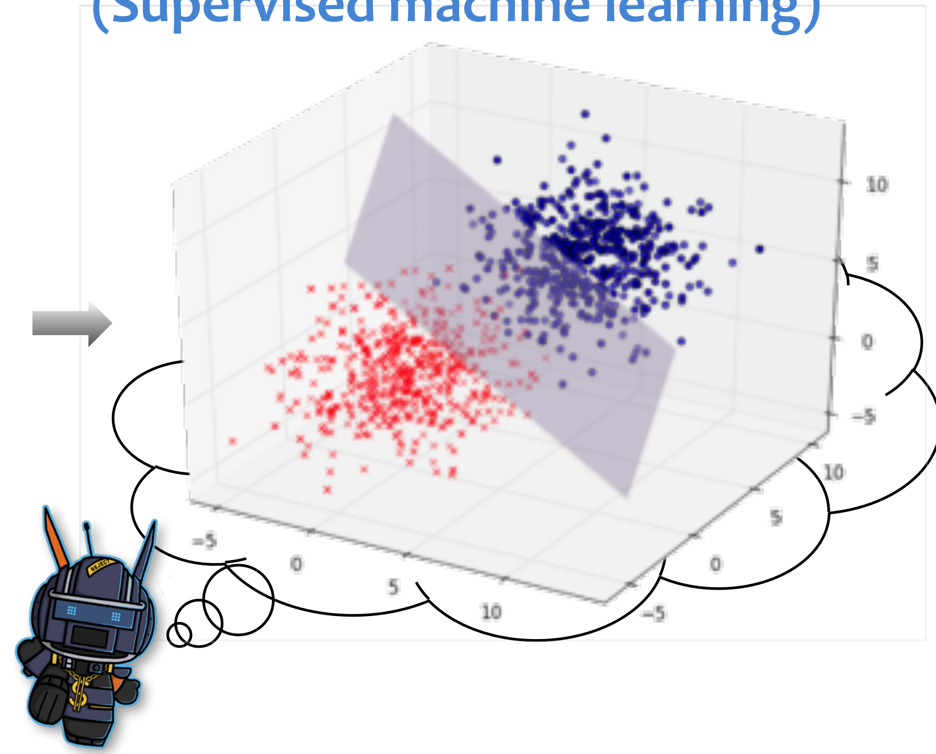




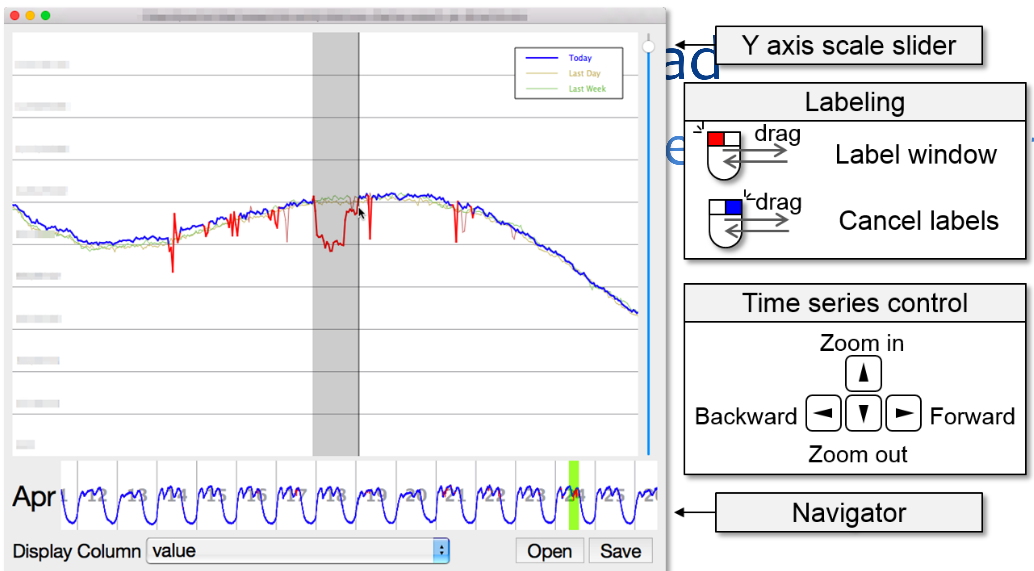
Key Ideas

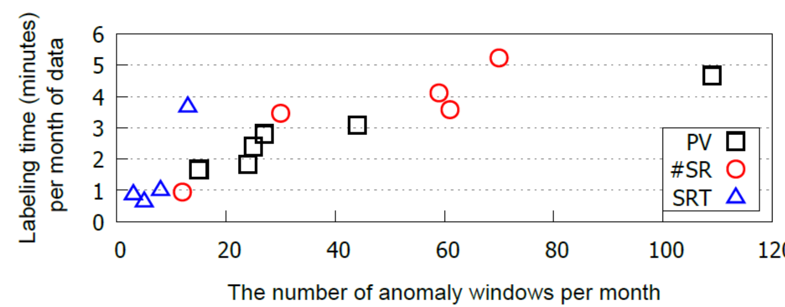


Classification in the feature space (Supervised machine learning)



Address Challenges of Designing Opprentice



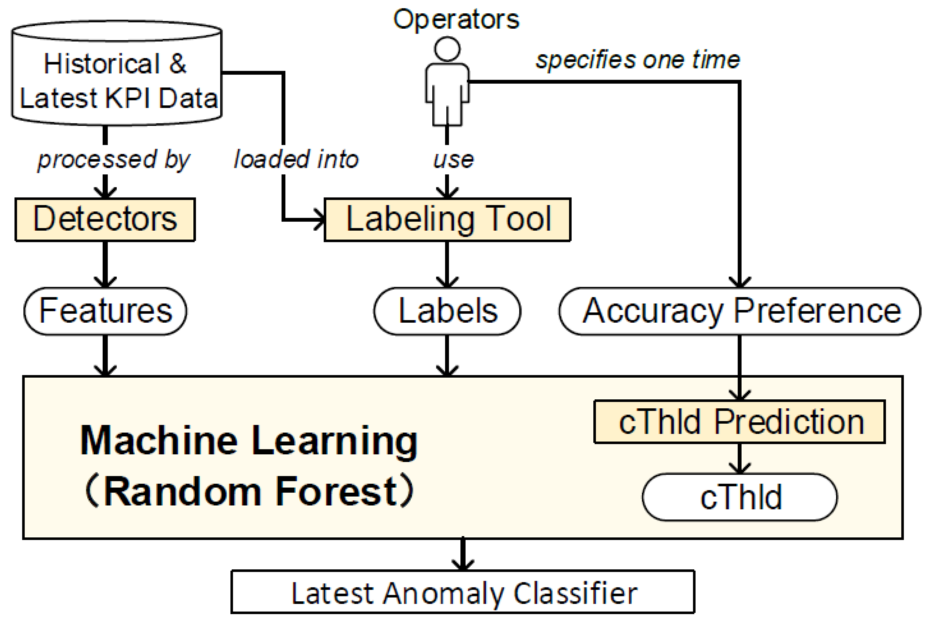


Address Challenges of Designing Opprentice

* Labeling overhead
  * Solution: an effective labeling tool
* Incomplete anomaly types in the historical data
  * Solution: incremental re-training with new data
* Class imbalance problem
  * Solution: adjusting classification threshold (cThld) based on the preference
* Irrelevant and redundant features
  * Solution: random forests

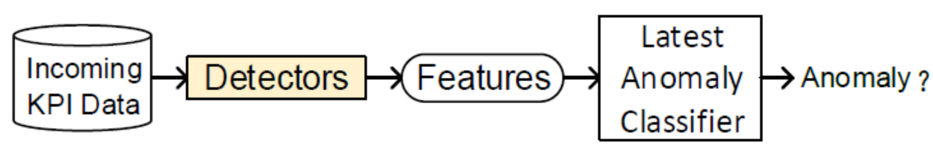
Design Overview

Training a classifier

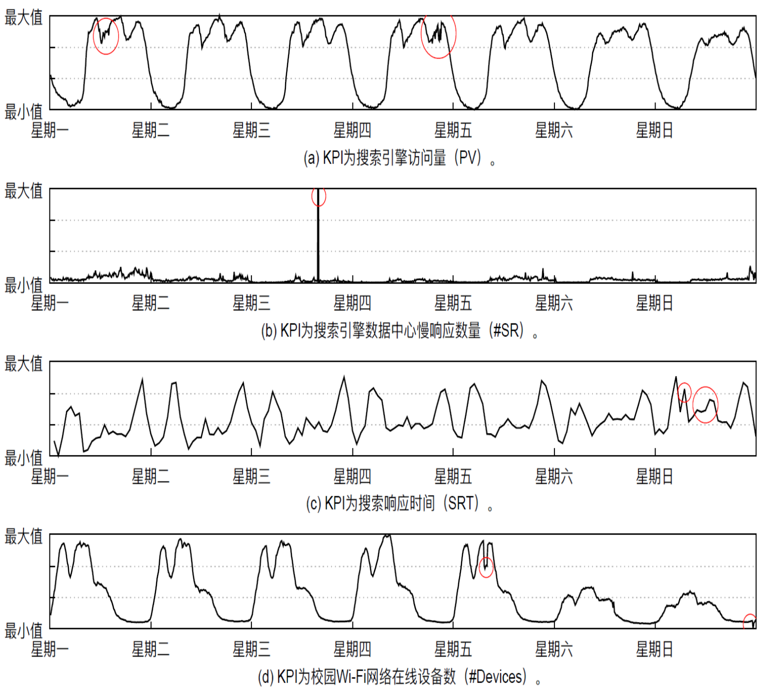


See the paper for full details

Detecting anomalies



Evaluation



(a) KPI为搜索引擎访问量（PV）。

(b) KPI为搜索引擎数据中心慢响应数量（#SR）。

(c) KPI为搜索响应时间（SRT）。

(d) KPI为校园Wi-Fi网络在线设备数（#Devices）。

Search PV (25 weeks) ）

#slow queries（19 weeks）

Baidu

Search Response Time（16 weeks）

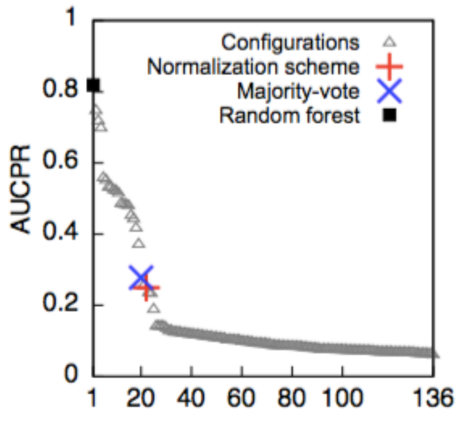
#online devices (15 weeks）

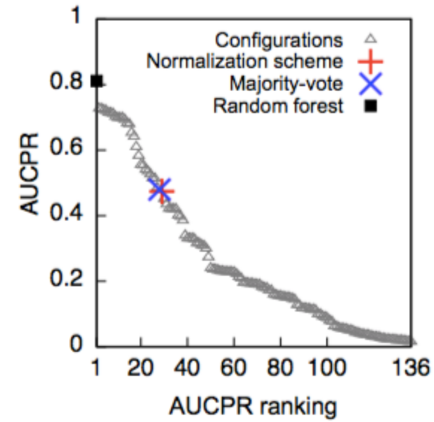
Tsinghua
Enterprise WiFi

Evaluation

* Compared with all existing detectors （Four KPIs）

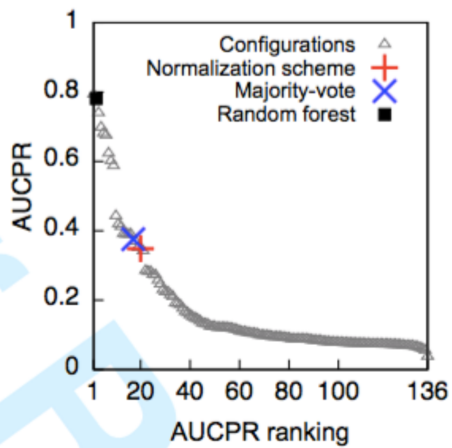
first

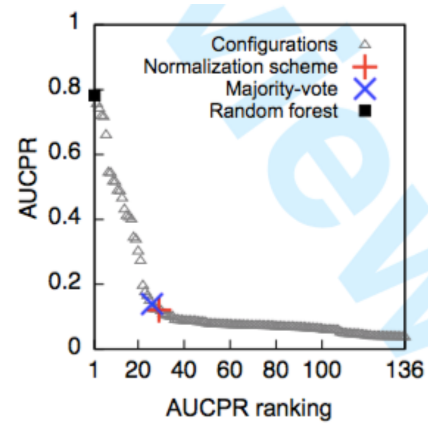




first

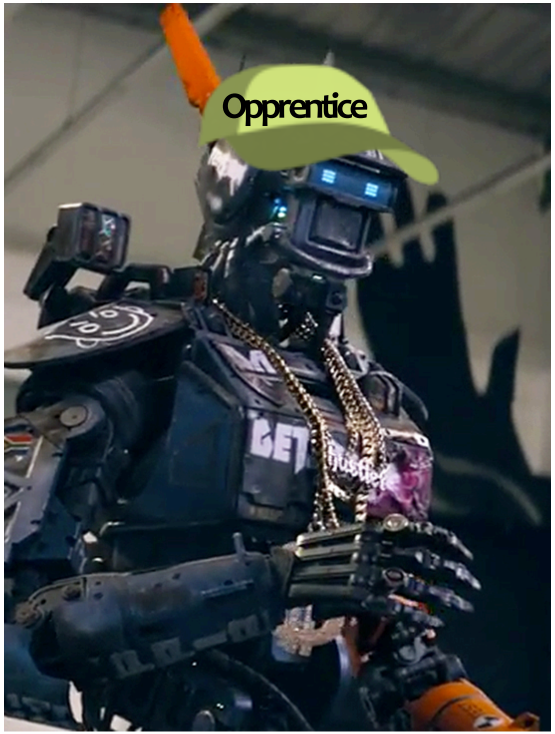
second





first

Case 2 summary



* Opprentice is an **automatic** and **accurate** machine learning framework for KPI anomaly detection

Defining anomalies | Selecting detectors | Tuning detectors

* Opprentice **bridges the gap** in applying complex detectors in practice

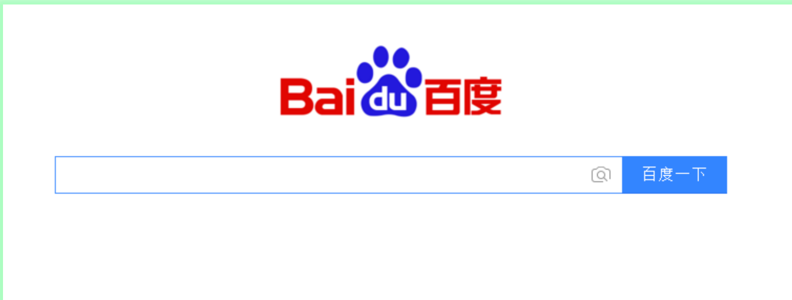
* The idea of Opprentice

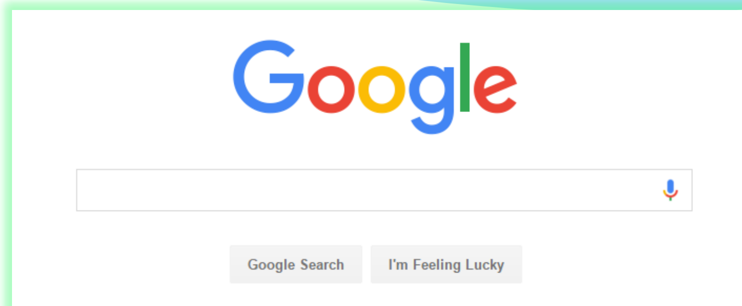
  i.e., **using machine learning to model the domain knowledge**

could be a very promising way to automate other service managements

Case 3： Bottleneck Identification for Search Response Time (Dapeng Liu et al., INFOCOM 2016)

Web Search Engines



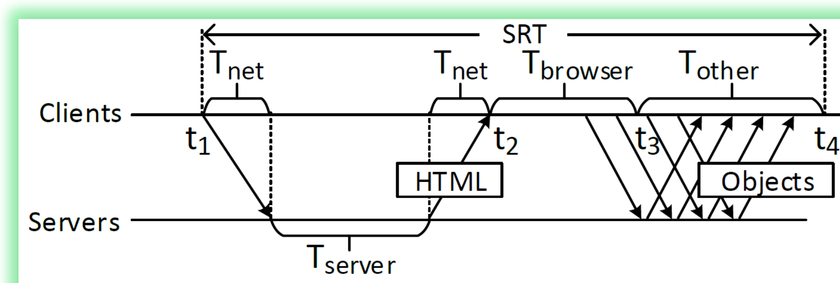


Search Response Time (SRT)

t_1 A search query is submitted

t_4 The result page Is rendered

$$SRT = t_4 - t_1$$



Search Response Time Matters

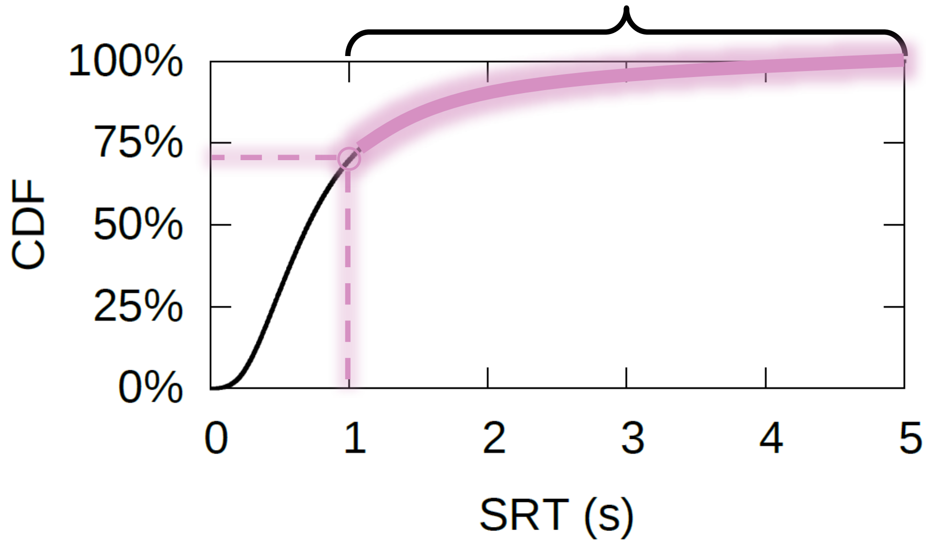
+500ms revenue 1.2%
[Eric Schurman, Bing]

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

Given two content-wise identical search result pages,
users are more likely to perform clicks on the fast page
[SIGIR 2014]

Search Response Time in the Wild

User's flow of thought is interrupted
if pages take **longer than 1s** to load



Why?

https://www.nngroup.com/articles/response-times-3-important-limits/

Monitoring SRT: Search Logs

Measurable attributes that can potentially impact SRT

SRT	User's ISP	Browser engine	# of Images	Ads	Server Load	…
800ms (Low SRT)	China Unicom	WebKit	10	Yes	1000 queries/s	…
1200ms (High SRT)	China Telecom	Trident 5.0	5	No	500 queries/s	…
……						

Goal of FOUCS

Measurable attributes that can potentially impact SRT

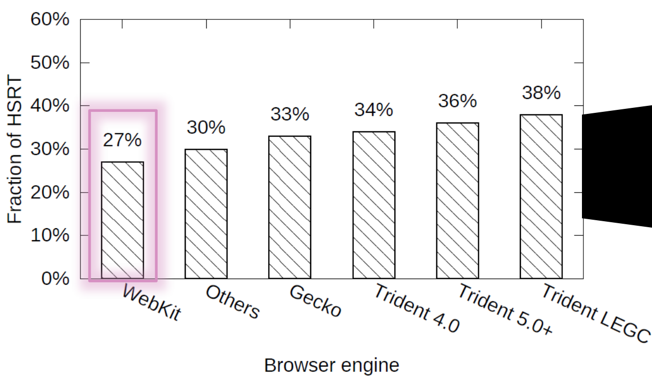
SRT	Client ISP	Browser engine	# of Images	Ads	Server Load	…
800ms (Low SRT)	China Unicom	WebKit	10	Yes	1000 queries/s	…
1200ms (High SRT)	China Telecom	Trident 5.0	5	No	500 queries/s	…
……						

We propose **FOCUS**, a search log analysis system to answer the following questions:

- Under what conditions **HSRT** （**High SRT**） is more likely to happen?
- Which HSRT conditions are similar (HSRT condition types)?
- How does each attribute affect SRT in HSRT condition types?

Challenges

Limited visibility of naïve single-dimension analysis



What we can see

WebKit **only 27%**

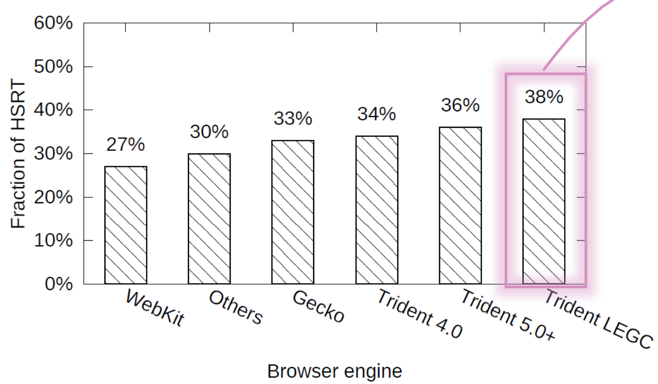
What we **cannot** see

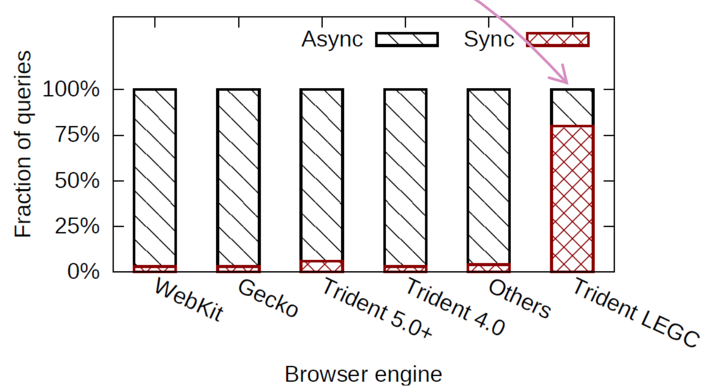
HSRT is **more than 38%**
when **"WebKit + #Images >30"**

Challenges

Limited visibility of naïve single-dimension analysis

Interdependencies between attributes





Which one should be blamed? Legacy Trident or sync page loading?

Challenges

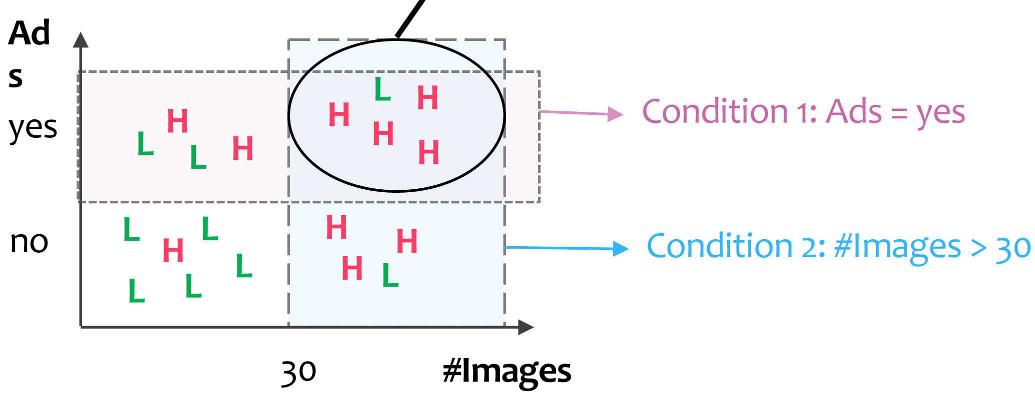
Limited visibility of naïve single-dimension analysis

Interdependencies between attributes

Overlapped HSRT conditions

For example: H(igh SRT) L(ow SRT)

HSRT in the overlapped part will be explained by more than one condition, but which one is better?



Challenges

Limited visibility of naïve single-dimension analysis

Interdependencies between attributes

Overlapped HSRT conditions

Key Idea of FOCUS

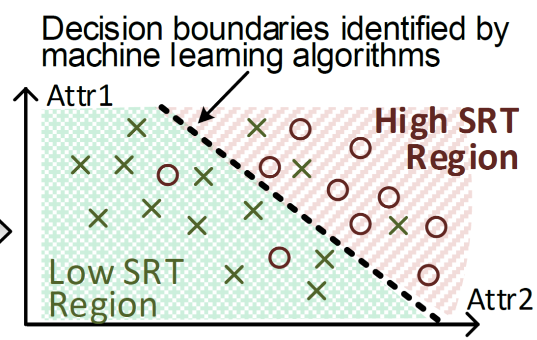
Limited visibility of naïve single-dimension analysis ← Multi-dimension analysis

Interdependencies between attributes ← Work with interdependencies

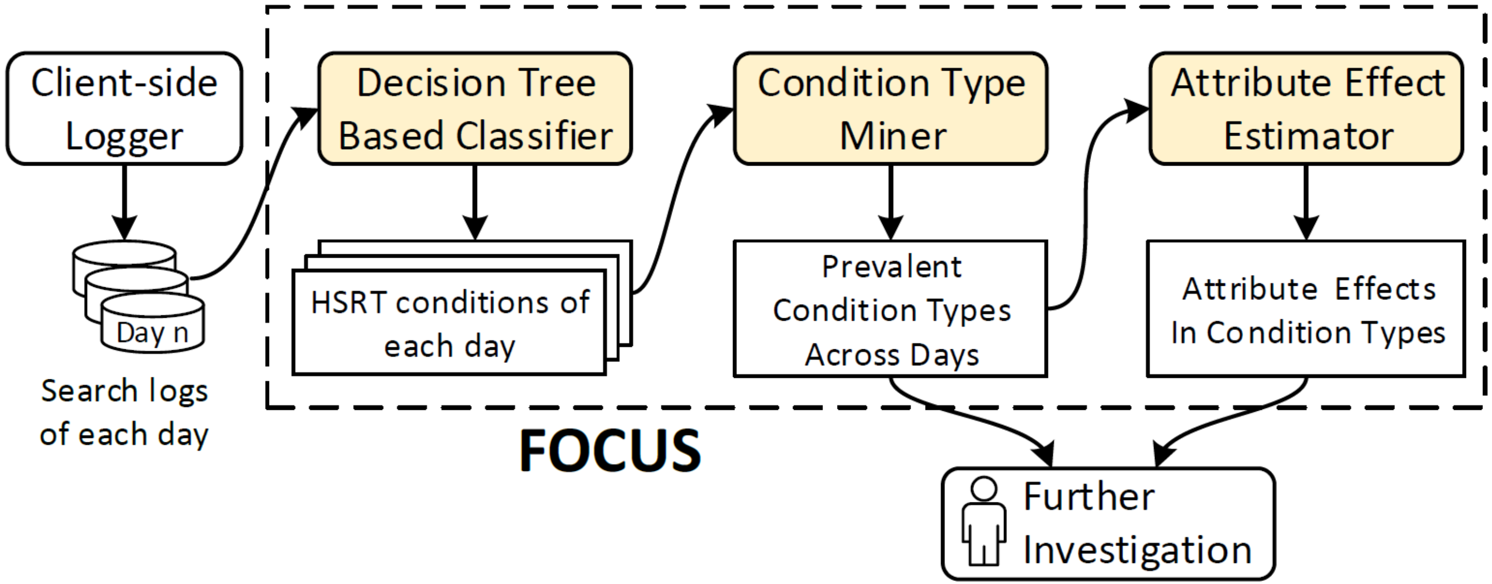
Overlapped HSRT conditions ← Classification is non-overlap

- Model it as a classification problem
- Solve it using decision trees

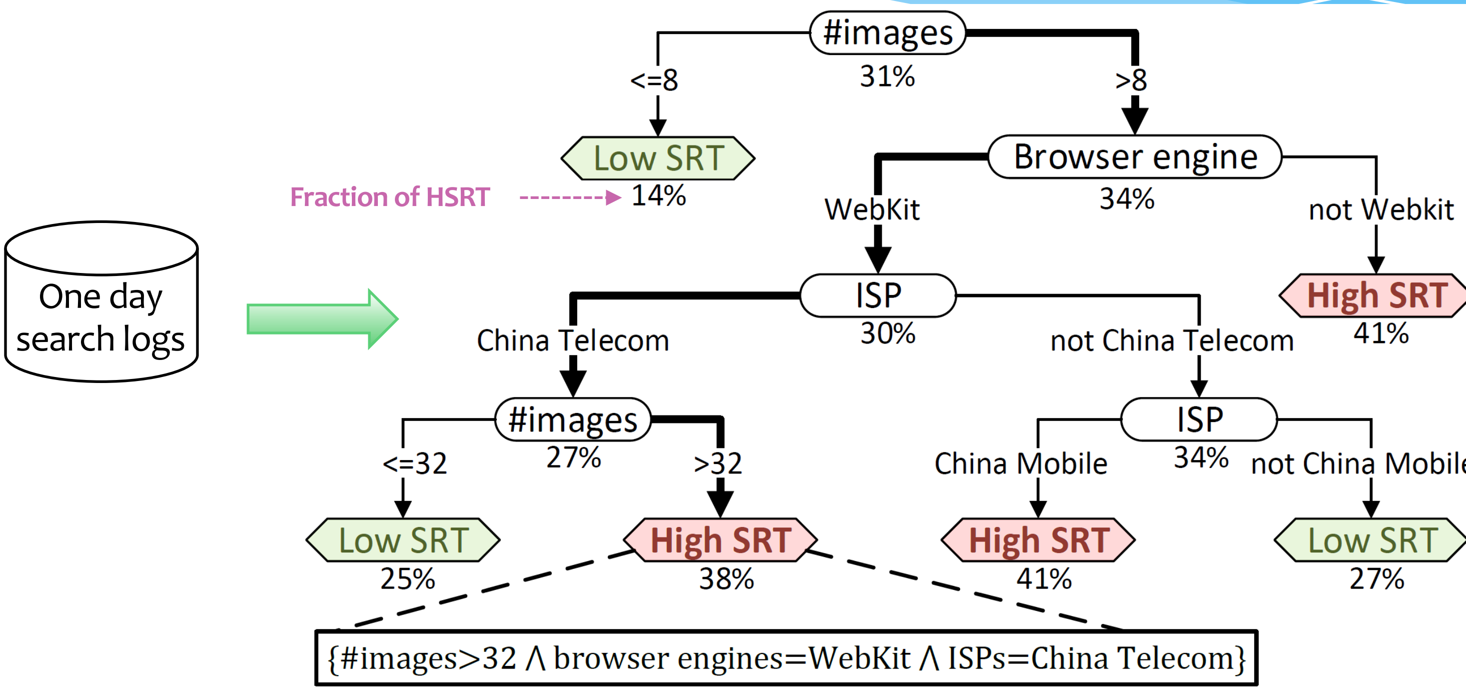
Attr1	Attr2	Label
…	…	High SRT ○
…	…	Low SRT ×
…	…	Low SRT ×
…	…	…



FOCUS Overview



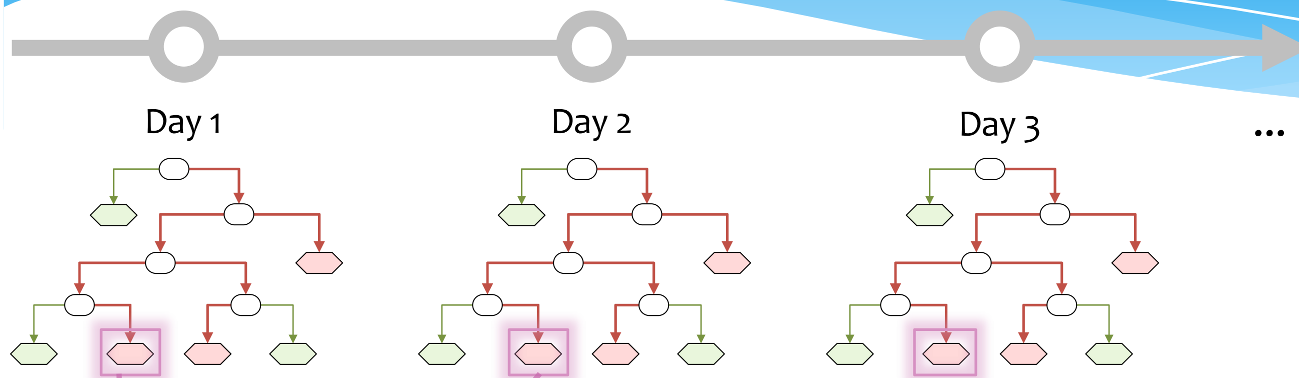
Identify HSRT Conditions Based on a Decision Tree



A HSRT condition

To build a reasonable tree, we **tailor** the mechanisms of decision trees
(Details are in the paper)

Find Similar HSRT Conditions (HSRT Condition Types)



ID	HSRT Conditions		
	#Images	Browser engine	Ads
1	> 9	Not WebKit	no
2	> 10	Not WebKit	no

HSRT Condition Type		
#Images	Browser engine	Ads
$> i, i \in \{9,10\}$	Not WebKit	no

- Same combination of attributes
- Same value for each categorical attribute
- **Similar** value for each numeric attribute

Hierarchical clustering

Estimate the Impact of Each Attribute

Inspired by controlled experiment

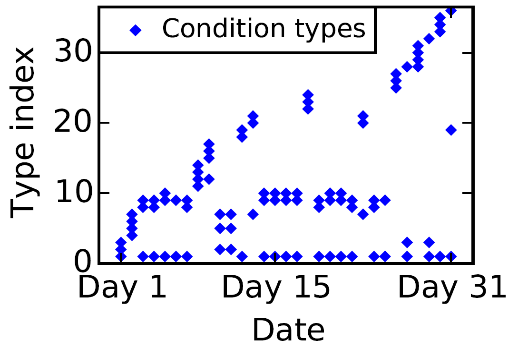
- **Control group:** the original HSRT contrition types
- **Experimental group:** changing one attribute at a time

Historical search logs

Compare performance in historical logs

ID	HSRT Condition Type		
	#Images	Browser engine	Ads
C	$> i, i \in \{9,10\}$	Not WebKit	no
C_1	$\le i, i \in \{9,10\}$	Not WebKit	no
C_2	$> i, i \in \{9,10\}$	WebKit	no
C_3	$> i, i \in \{9,10\}$	Not WebKit	yes

Results of FOCUS： Prevalent HSRT Condition Types



- Find 36 HSRT condition types in one month of search logs
- **Four** of them (11%) appear in more than five days

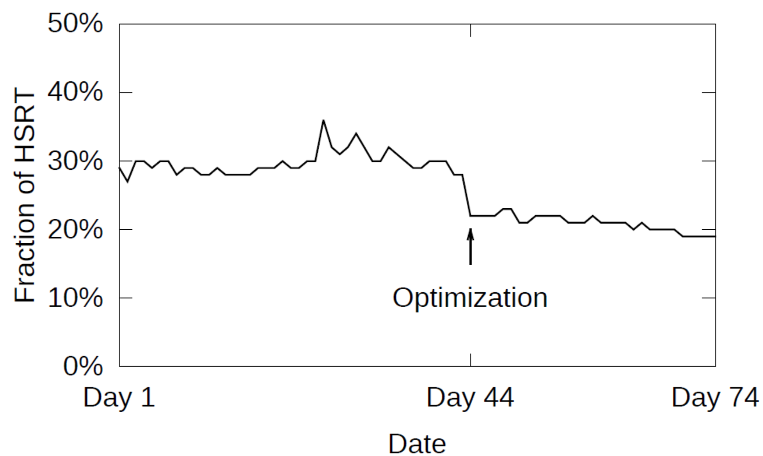
Images are the main bottleneck

(**Attributes** in bold have a bad effect on SRT)

Condition type ID	Prevalent condition type	Prevalence (days)
1	$\mathbf{\#images} > \boldsymbol{i}, \boldsymbol{i} \in \{\mathbf{5,6,7,8,9}\} \wedge$ **browser engine = not WebKit**	21
2	$\mathbf{\#images} > \boldsymbol{i}, \boldsymbol{i} \in \{\mathbf{5,6,7,8,9}\} \wedge$ **ISP = not China Telecom** $\wedge$ browser engine = WebKit	15
3	$\mathbf{\#images} > \boldsymbol{i}, \boldsymbol{i} \in \{\mathbf{25,26,27}\} \wedge$ ISP = China Telecom $\wedge$ browser engine = WebKit	7
4	$\mathbf{\#images} > \boldsymbol{i}, \boldsymbol{i} \in \{\mathbf{5,6,8}\} \wedge$ ISP = China Telecom $\wedge$ browser engine = WebKit $\wedge$ **ads = yes**	6

Real-world Optimization

- 1^{st} month results of FOCUS → images are the main bottleneck of SRT
- Deploy "image base64 encoding" to improve the transmission time of images



(a) Fraction of HSRT each day

HSRT percentage is reduced by 30%

SRT 80th-tile is reduced by 253 ms (20%)

The fraction of HSRT is reduced by 30%

Case 3 Summary

* FOCUS can
  * Narrow down the debugging space of High SRT in search logs
  * Analyze the effects of each attribute (potential improvements)

* With the output of FOCUS
  * We make several interesting observations
  * Deploy a solution in practice and greatly improve SRT

* FOCUS is a general method for analyzing multi-attribute logs
  * Web applications other than search engines
  * Performance of mobile apps
  * …

Outline

* Intelligent Operations: from "rule based" to "machine learning based"

* Case Studies

* *Challenges and My Thoughts*

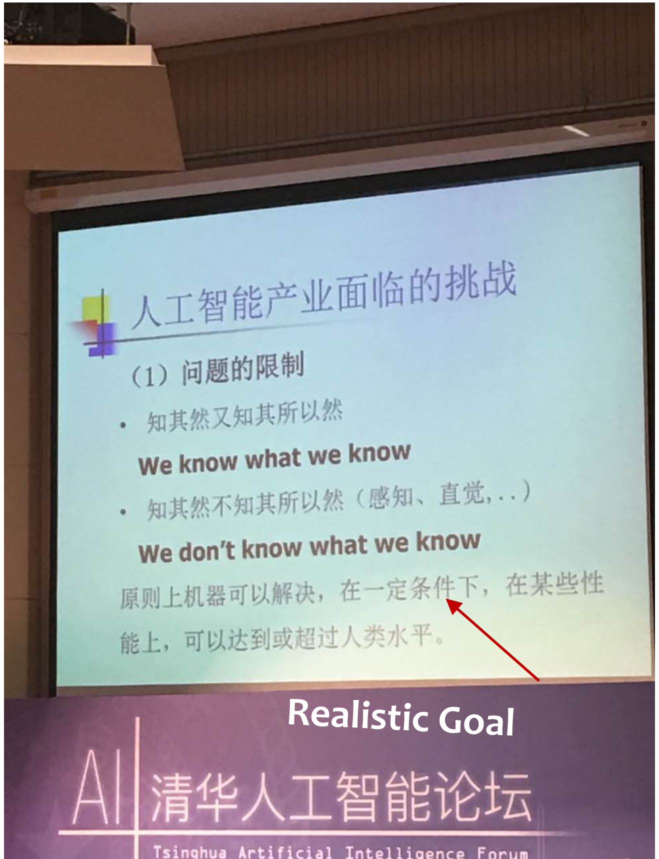
Challenge 1： T2 or R2-D2 for IOP?

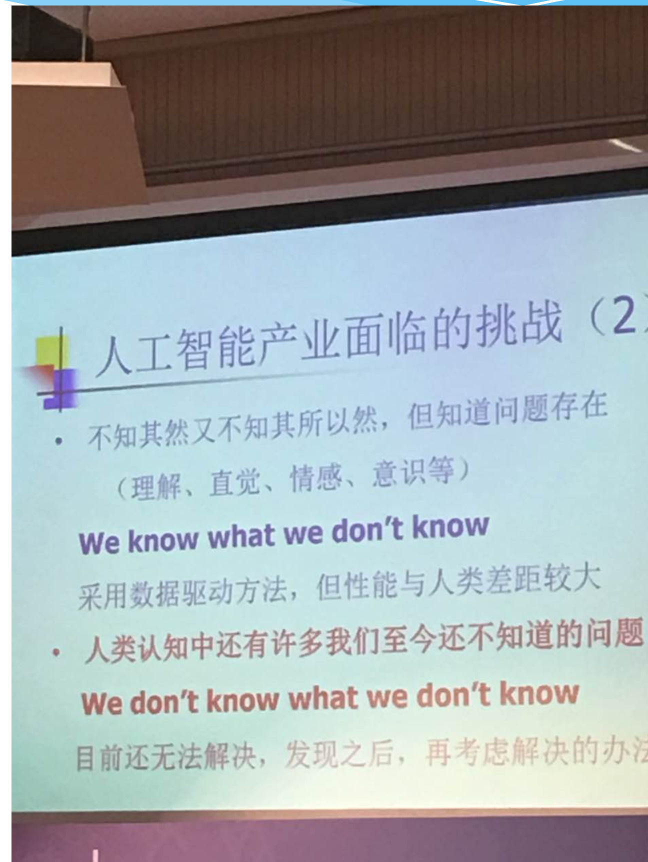
T2: completely automated, in charge of everything in OP ?

R2-D2: A reliable sidekick for Operator ?

Pictures are from the Internet

Thought: Automate on those OP tasks for which "we don't know what we know"

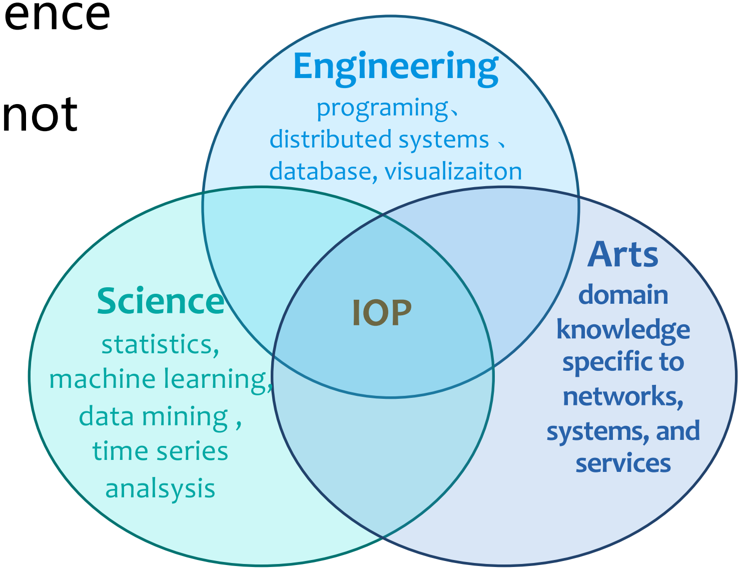




From Professor Bo Zhang, AI expert from Tsinghua University

Thought: Automate on those OP tasks for which "we don' t know what we know"

- Technology can gradually solve science and engineering problems, but cannot do much in arts



- Computer/Engineering cannot replace artists, but could provide better tools for artists.

The ultimate goal of Intelligent OP: Automate as much as possible such that

1. Routine tasks are automatically done
2. Operators can independently conduct data analysis

Challenge 2： How to retrieve useful labels from tickets?

NSDI 2013

Strawman Approach To Analyze Free-form Text

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 **continously 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 **continously 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

9

Thought： Design ticketing system as part of IOP architecture

- Ticket format and ticketing system should be carefully design as an important component of the IOP system.
- Tickets need contain enough information for machine learning.
- Ticketing system should be designed as a user friendly "product"
- Operators should be self-disciplined in filling the tickets.
- NLP-based tool to analyze the free-form text in the tickets

Challenge 3： Software upgrade on vendor devices cannot be very frequent

* Thought：

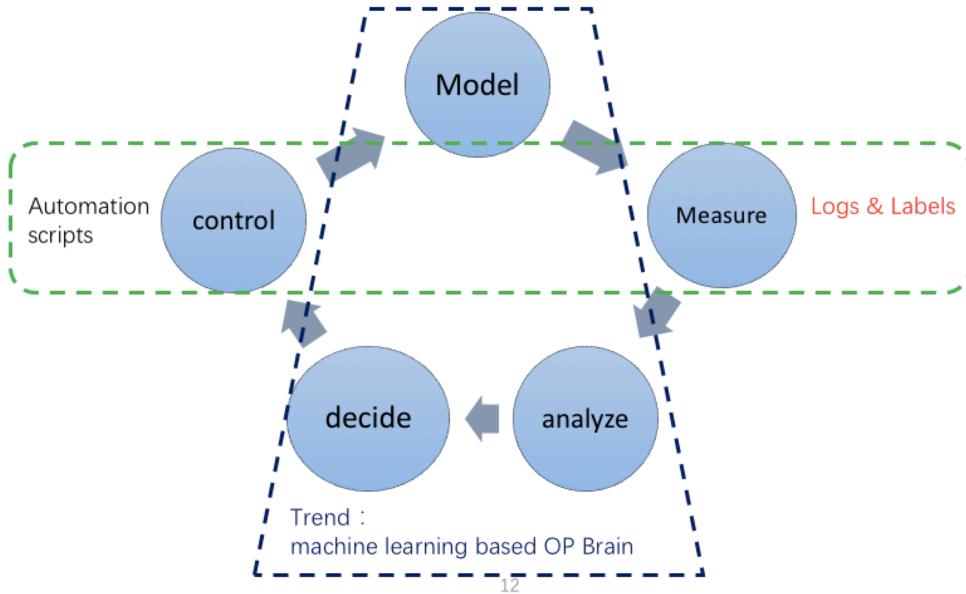
- IOP should be a key component of the device software architecture

- logging and control model should be programmable and evolvable

- The No. 1 goal for software UI/UX design is to collect labels from operators

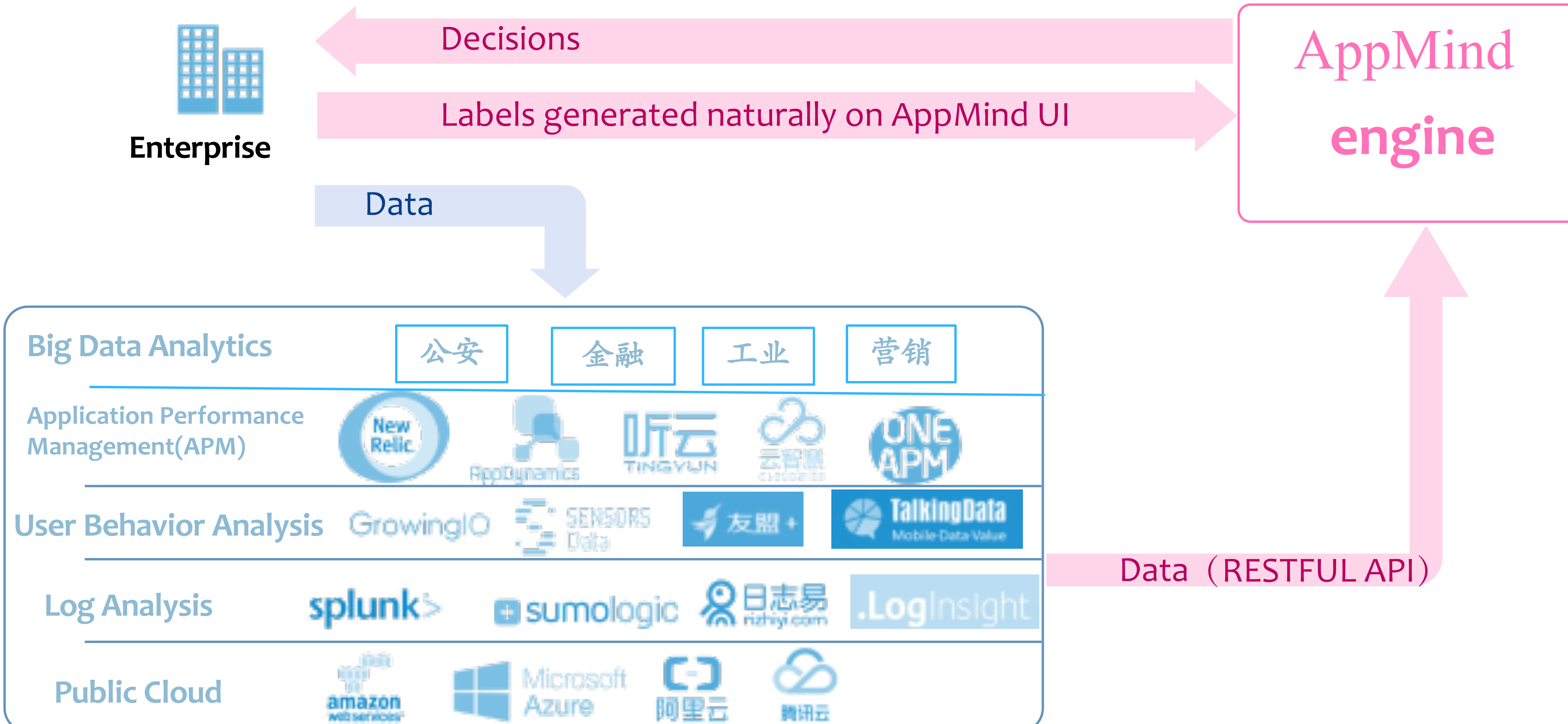
- Close collaboration between operators + scientist + engineer

- Look for real users and trial field.

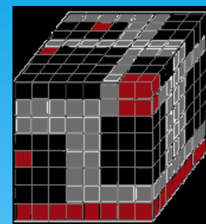


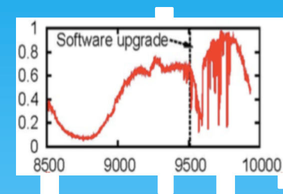
AppMind

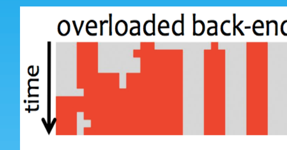
APPMind： a micro service module in the cloud, for some keyIOP tasks

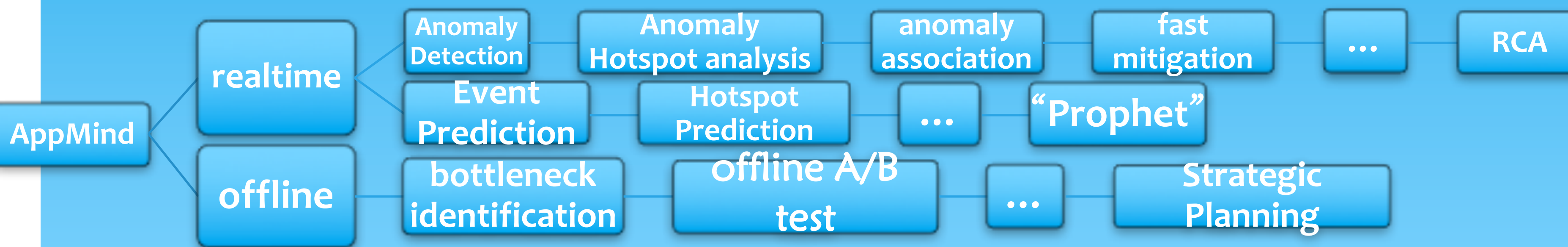


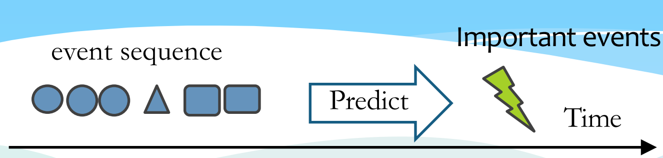
AppMind




Software upgrade
1
0.8
0.6
0.4
0.2
0
8500
9000
9500
10000


overloaded back-end
time


AppMind
realtime
offline
Anomaly Detection
Anomaly Hotspot analysis
anomaly association
fast mitigation
…
RCA
Event Prediction
Hotspot Prediction
…
"Prophet"
bottleneck identification
offline A/B test
…
Strategic Planning


event sequence
Predict
Important events
Time

Summary

* Intelligent Operations through machine learning will see rapid progress in the next few years.
  – abundant data, available labels, and ready applications

* Let's embrace it with
  * More principled application of cutting-edge machine learning techniques
  * Consciously produce more labels and more data to fuel machine learning
  * Aiming at a robot sidekick for OP first.
  * Vendor software' s measurement and control modules need to be programmable and evolvable.

THANK YOU

Email: peidan@tsinghua.edu.cn
Wechat： peidanwechat
http://netman.cs.tsinghua.edu.cn

《Advanced Network Management》：
http://netman.cs.tsinghua.edu.cn/courses/advanced-network-management-spring2016/

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