Robust and Rapid Adaption for Concept Drift in Software System Anomaly Detection

**Minghua Ma**, Shenglin Zhang, Dan Pei,
Xin Huang, Hongwei Dai

ISSRE 2018
Web-based Software System

Search Engine
- Google
- bing
- Sogou

Online Shopping
- amazon
- ebay
- Taobao.com

Social Network
- facebook
- twitter
- weibo
Web-based Software System Measurements
Web-based Software System Measurements

Servers

Processes

Key Performance Indicators

e.g. Page View, Search Response Time...
How to Build an Anomaly Detector

Value

Anomaly

Select / combine detectors

- Moving Average
- Holt-Winters
- Time Series Decomposition

Tune parameters & thresholds

[IMC15,KDD15]
How Anomaly Detector Works
Operators Often Conduct Software Changes
Operators Often Conduct Software Changes

Expected Concept Drift
Page view on Process#1 drops significantly

[ISSRE15,CoNEXT15]
Accuracy of Anomaly Detector Degrades

- Expected Concept Drift
- False Alarm! Annoying
Accuracy Degradation Reason #1: False Alarm

Anomaly detector uses one period of historical data (e.g. one day/week) to predict the upcoming "normal" data.

Trend of KPI changes

False Alarm!
Accuracy of Anomaly Detectors Degrade

Expected Concept Drift

Missing Alarm! Oops...

False Alarm! Annoying
Accuracy Degradation Reason #2: Missing Alarm

Anomaly detector cannot adapt to the amplitude of KPI without tune parameters and thresholds again

Amplitude of KPI changes

Expected Concept Drift

Missing Alarm!

Anomaly Detector

Time

Value
Accuracy of Anomaly Detectors Degrade

Expected Concept Drift

Trend changes
Amplitude changes

Missing Alarm!

Value

Anomaly Detector

False Alarm!

Mon. Tue. Wed. Thur. Fri. Time
Outline

What's concept drift?

Why it's challenging?

How to adapt to it?

Evaluation
Concept Drift in KPI Anomaly Detection

- Expected
- Unexpected
  - Jitters
  - Spikes
  - Dips...

Concept Drift → Anomaly

The distribution of KPI changes significantly in a short time
Unexpected Concept Drift: Roll Back

Expected

Unexpected

Jitters
Spikes
Dips ...

Concept Drift

Anomaly

Software change rolls back by FUNNEL

[CoNEXT15]

Unexpeceted concept drift
Expected Concept Drift: Our Focus

- **Expected**
- **Unexpected**
  - Jitters
  - Spikes
  - Dips

**Concept Drift**

**Anomaly**

Software change rolls back by FUNNEL [CoNext15]

- Need rapidly adapt to the concept drift

Unexpeceted concept drift

Expected concept drift

![Diagram showing concept drift and anomaly with timelines and value graphs]

16
Outline

What's concept drift? ➔ KPI distribution change

Why it's challenging?

How to adapt to it?

Evaluation
Why Is Adaption to Concept Drift Challenging

- **High Frequency**
  - 3000 Concept Drifts Per Day
  - 80% Are Expected Ones

- **Huge Amount of KPIs**
  - Tens of Thousands of Servers
  - Tens of Processes
  - Tens of Types of KPIs

- **Diverse Detectors**
  - Single Detectors or Combination

- **Rapid Adaptation**
  - False Alarm
  - Missing Alarm
Why Is Adaption to Concept Drift Challenging

3000 Concept Drifts Per Day
80% Are Expected Ones

High Frequency

➔ Need an **automatic** concept drift adaption approach
Why Is Adaption to Concept Drift Challenging

- Tens of Thousands of Servers
- Tens of Processes
- Tens of Types of KPIs

Huge Amount of KPIs

- Cannot tune parameters manually for each KPI stream

3000 Concept Drifts Per Day
80% Are Expected Ones

High Frequency

[CoNext15, IMC15]

False Alarm

Missing Alarm

Diverse Detectors

Rapid Adaption
Why Is Adaption to Concept Drift Challenging

- 3000 Concept Drifts Per Day
- 80% Are Expected Ones

High Frequency

- Tens of Thousands of Servers
- Tens of Processes
- Tens of Types of KPIs

Huge Amount of KPIs

Single Detectors or Combination

Diverse Detectors

[IMC15, SIGCOMM10, SIGCOMM13]

→ Need a generic adaption approach for various types of detectors
Why Is Adaption to Concept Drift Challenging

- 3000 Concept Drifts Per Day
- 80% Are Expected Ones
- High Frequency
- Huge Amount of KPIs
- Single Detectors or Combination
- Diverse Detectors
- False Alarm
- Missing Alarm
- Rapid Adaption

➤ Help anomaly detectors adapt concept drift in a **rapid** manner
Why Is Adaption to Concept Drift Challenging

High Frequency

3000 Concept Drifts Per Day
80% Are Expected Ones

Huge Amount of KPIs

Tens of Thousands of Servers
Tens of Processes
Tens of Types of KPIs

Diverse Detectors

Single Detectors or Combination

Rapid Adaption

False Alarm

Missing Alarm

StepWise
Detection, Classification and Adaption

Concept Drift

in Software System Anomaly Detection
StepWise Overview

- iSST-EVT
  - Improved SST
  - Extreme Value Theory
  - Score
  - Threshold

- Semi-Automatic
  - Difference in Differences
  - Software Change Roll Back
  - Concept Drift
  - Unexpected

- RLM-Adaption
  - Robust Linear Model
  - Old Concept, New Concept
  - Expected

- KPI Streams
- Anomaly Detectors
StepWise Overview

- **iSST-EVT**
  - Improved SST
  - Extreme Value Theory

- **Semi-Automatic**
  - Difference in Differences
  - Software Change Roll Back

- **RLM-Adaption**
  - Robust Linear Model
  - Old Concept, New Concept

- **KPI Streams**
  - Score
  - Threshold

- **Detection**

- **Anomaly Detectors**

- **Concept Drift**
  - Expected
  - Unexpected
StepWise Overview

KPI Streams

iSST-EVT
- Improved SST
  - Score
  - Threshold
- Extreme Value Theory

Semi-Automatic
- Difference in Differences
  - Software Change Roll Back
- Concept Drift
- Unexpected

RLM-Adaption
- Robust Linear Model
  - Old Concept, New Concept

Anomaly Detectors

Detection

Classification
StepWise Overview

KPI Streams

- Improved SST
  - Extreme Value Theory
  - Score
  - Threshold

Concept Drift

Difference in Differences

- Software Change Roll Back
- Unexpected

Expected

Robust Linear Model

Old Concept, New Concept

Anomaly Detectors

Detection
- Classification
- Adaption
iSST-EVT Detection

iSST-EVT

- Improved SST
- Extreme Value Theory

Semi-Automatic

- Difference in Differences
- Software Change Roll Back

RLM-Adaption

- Robust Linear Model
- Old Concept, New Concept

Anomaly Detectors

Detection

Classification

Adaption
iSST: Improved Singular Spectrum Transform

Detection time window: 6 min

Unbounded value, varies in different KPI streams
iSST: Improved Singular Spectrum Transform

KPI → iSST → Change Score → Threshold → Concept Drift

Tune threshold manually on a per-KPI basis is not affordable
Idea #1. iSST-EVT

KPI → iSST → Change Score → EVT → Concept Drift

Spike detection algorithm without making any assumption about the data distribution
EVT: Extreme Value Theory

Extreme Value Distribution $G_\gamma : x \mapsto \exp\left(-(1 + \gamma x)^{-\frac{1}{\gamma}}\right)$

Change score $X_1, X_2, \ldots, X_n \xrightarrow{\text{EVT}} z_q$

Compute $z_q$ such as $\mathbb{P}(X > z_q) < q$
iSST-EVT: no detection threshold to tune, no per-KPI model parameters to tune, and all model parameters can use model-wide empirical values
Semi-Automatic Classification

Reliable event feeds: eg, software change / holidays

KPI Streams

Improved SST
Score
Threshold
Extreme Value Theory

Concept Drift

Difference in Differences

Software Change Roll Back

Expected

RLM-Adaption

Robust Linear Model

Old Concept, New Concept

Anomaly Detectors

Detection

Classification
Semi-Automatic Classification

Reliable event feeds: eg, software change / holidays

KPI Streams

Improved SST

Score

Threshold

Extreme Value Theory

Concept Drift

Difference in Differences

Semi-Automatic

Software Change Roll Back

Expected

RLM-Adaption

Robust Linear Model

Operators make decisions

Old Concept, New Concept

Anomaly Detectors

Detection

Classification
RLM Adaption

KPI Streams

iSST-EVT
- Improved SST
  - Score
  - Threshold
- Extreme Value Theory

Semi-Automatic
- Difference in Differences
- Software Change Roll Back

Concept Drift

RLM-Adaption
- Robust Linear Model
  - Old Concept, New Concept

Anomaly Detectors

Detection
Classification
Adaption
Observation: Old and New Concept Can Be Linearly Fitted

Expected Concept Drift

Old Concept  New Concept
Trend changes
Amplitude changes

Value

Time

Mon. Tue. Wed. Thur. Fri.
Observation: Old and New Concept Can Be Linearly Fitted

Physical meaning: software on-demand scaling up/down
Challenge: Modeling Linear Relationship in Real-Time
Idea #3. RLM Adaption

Adaption Algorithm Using Robust Linear Model

Median Value for Every Time Bin
Idea #3. RLM Adaption

Adaption Algorithm Using Robust Linear Model

- Median Value for Every Time Bin
- Robust to anomalies than Least Squares Regression
- New Concept
- Old Concept
- Anomaly Detectors
StepWise Prototype Has Been Deployed in Sogou
Outline

What's concept drift? ➔ KPI distribution change

Why it's challenging? ➔ Four challenges

How to adapt to it? ➔ Detection, classification, adaption

Evaluation ➔ 288 labeled KPIs over six months
F-score of Anomaly Detection Algorithm

(EWMA, Holt-Winters, TSD etc.)

Substantial accuracy improvement by StepWise
Adaption Lag

Time when the concept drift starts

Delay of the Concept Drift Detection by iSST-EVT

Time when the concept drift is adapted

Without any adaption, it will take **days** for operators to tune the parameters and thresholds of anomaly detectors

**StepWise is rapid enough to adapt to concept drift**
Conclusion

• The accuracy of anomaly detection degrades because of the **concept drift**
• Challenges: scalability, robustness, and adaption delay

**StepWise** uses **domain-specific insights**
• iSST-EVT: no need to tune per-KPI parameter or threshold
• RLM: robust and rapid adaption to diverse types of detectors

Evaluation
• Real world KPI anomalies and concept drifts
• Improve F-score by 206% over a baseline without any concept drift adaption
StepWise
ISSRE 2018

Thank you

mmh16@mails.tsinghua.edu.cn