Heterogeneous Anomaly Detection for Software Systems via Semi-supervised Cross-modal Attention

Cheryl Lee[*], Tianyi Yang[*], Zhuangbin Chen[*], Yuxin Su[†], Yongqiang Yang[‡], and Michael R. Lyu[*]

*The Chinese University of Hong Kong, †Sun Yat-sen University, ‡Huawei Cloud

May, 2023

Table of Contents

01 INTRODUCTION

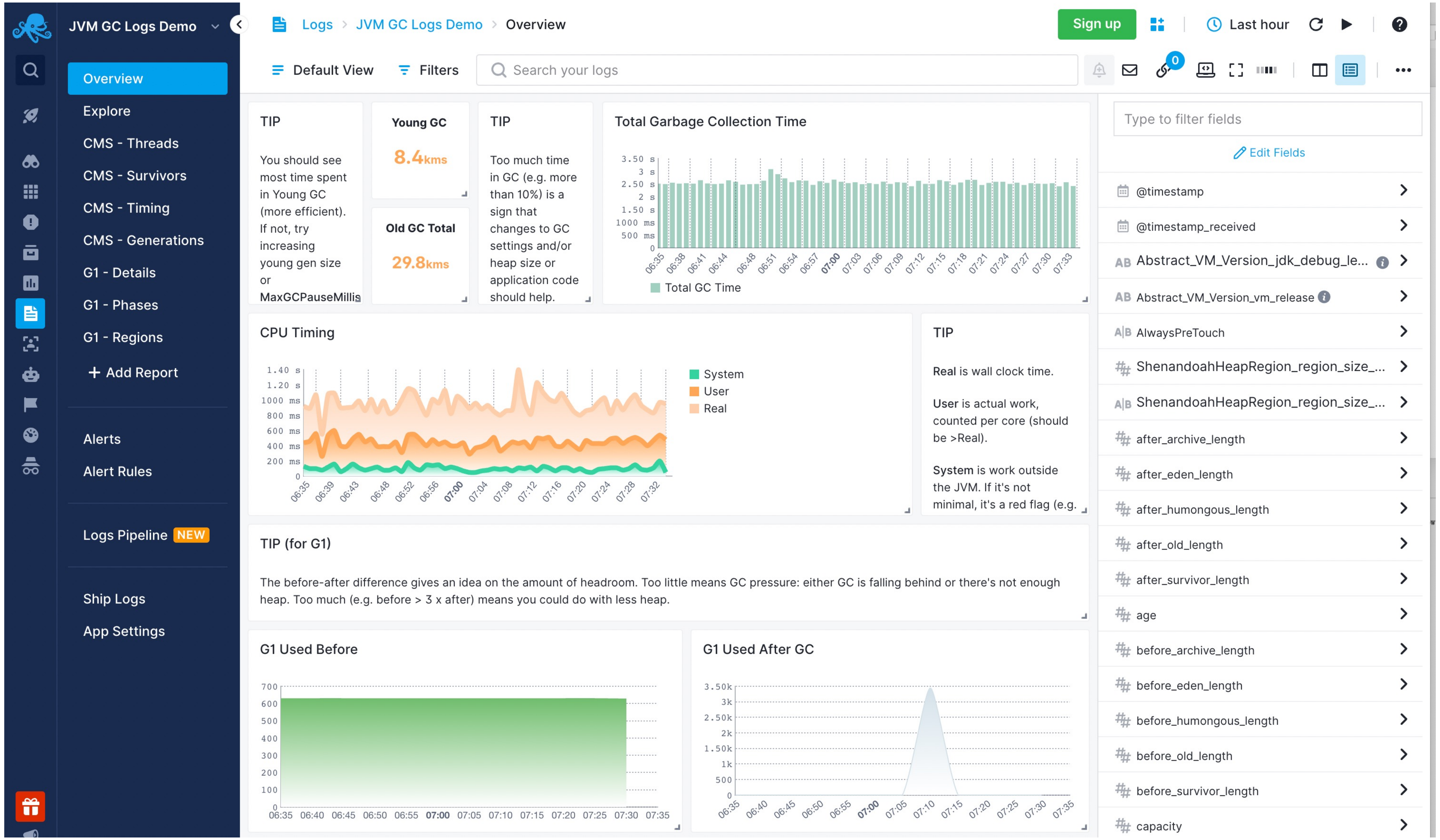
Background, Preliminary...

Background

Anomaly detection is essential

Background

Single-source data may be insufficient



Background

Anomaly detection is essential

Single-source data may be insufficient

Combining multi-source data may be effective

PRELIMINARY

Logs

```
17/06/09 20:10:48 INFO executor.Executor: Finished task 0.0 in stage 0.0 (TID 0). 2703 bytes result sent to driver
17/06/09 20:10:52 INFO executor.CoarseGrainedExecutorBackend: Got assigned task 42
17/06/09 20:10:52 INFO executor.Executor: Running task 0.0 in stage 1.0 (TID 42)
17/06/09 20:10:52 INFO executor.CoarseGrainedExecutorBackend: Got assigned task 56
17/06/09 20:10:52 INFO executor.Executor: Running task 1.0 in stage 1.0 (TID 56)
```

Log Sequence

Log Message

Parsing

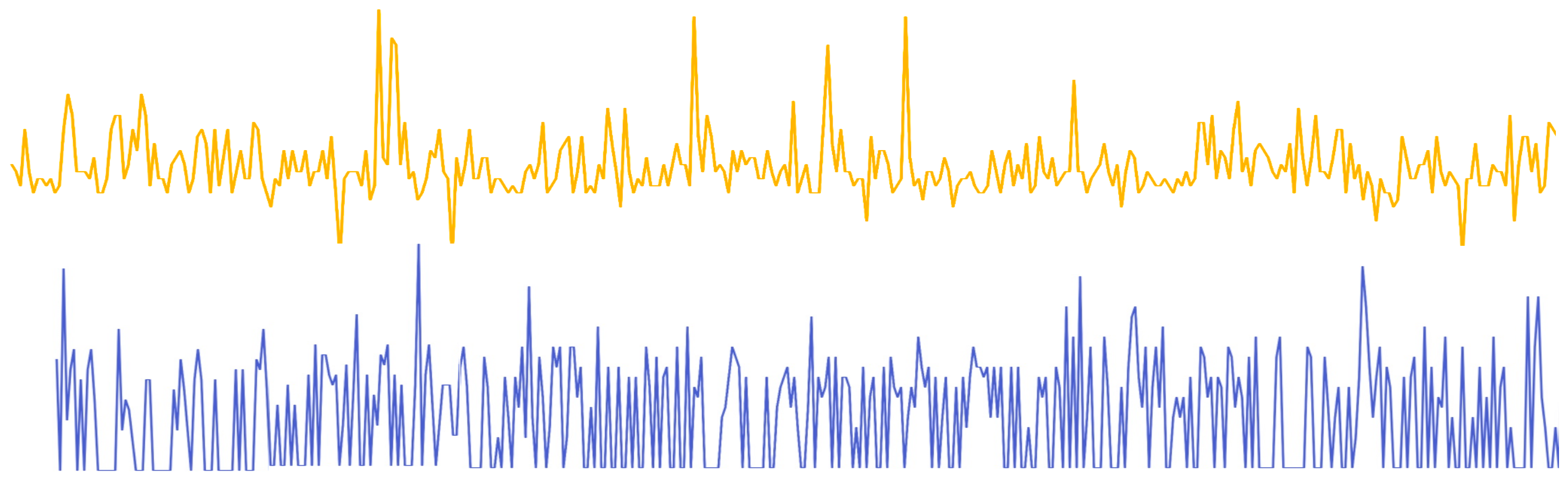
Timestamp	Level	Component	Log Event
17/06/09 20:10:48	INFO	executor.Executor	Finished task * in stage * (TID *). * bytes result sent to driver.
17/06/09 20:10:52	INFO	executor.CoarseGrainedExecutorBackend	Got assigned task *
17/06/09 20:10:53	INFO	executor.Executor	Running task * in stage * (TID *)
17/06/09 20:10:54	INFO	executor.CoarseGrainedExecutorBackend	Got assigned task *
17/06/09 20:10:55	INFO	executor.Executor	Running task * in stage * (TID *)

PRELIMINARY

Log events

INFO util.SignalUtils: Registered signal
WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18

Metrics

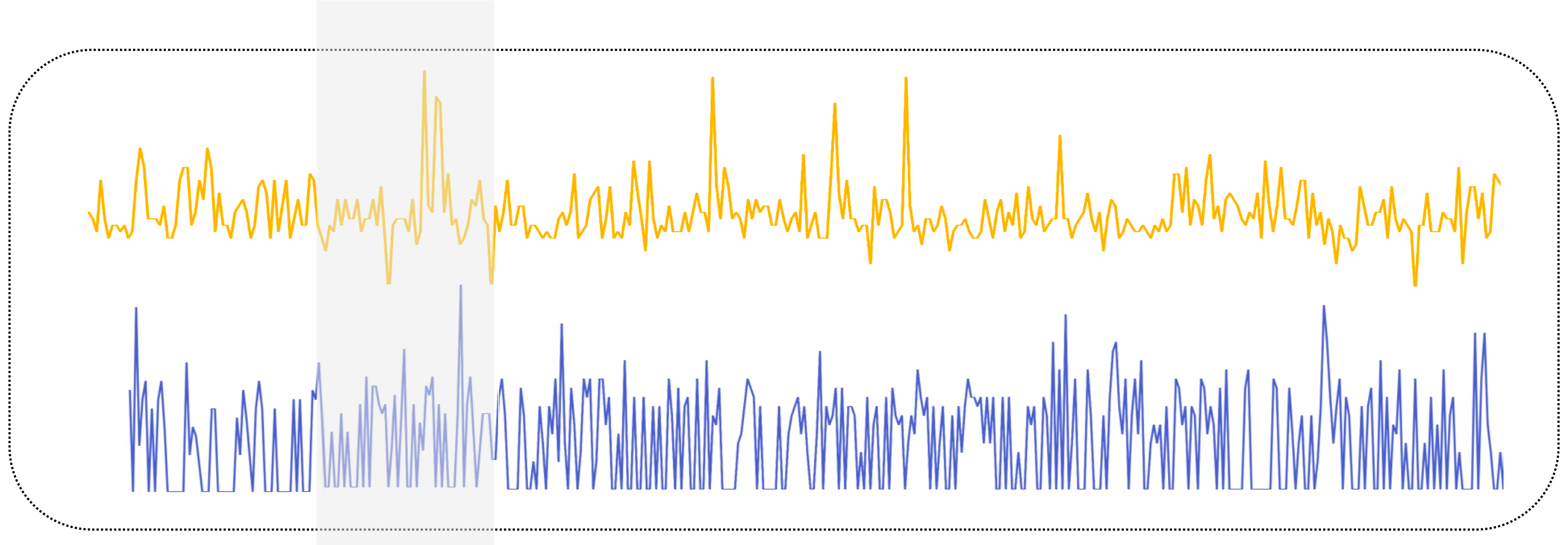


PRELIMINARY

Log events

INFO util.SignalUtils: Registered signal
WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18

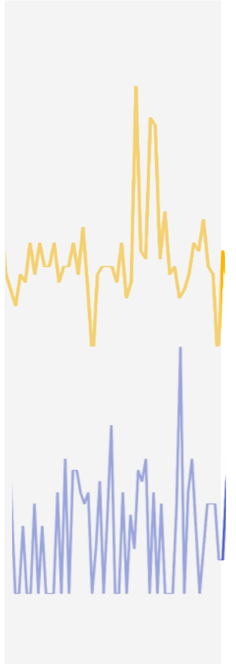
Metrics



PRELIMINARY

A chunk

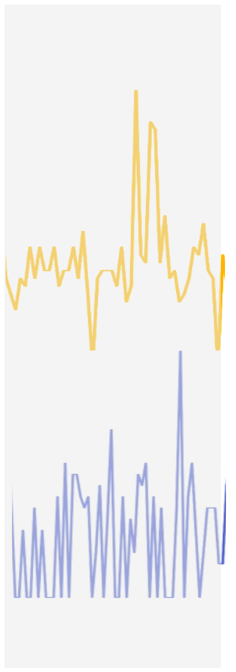
WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18



PRELIMINARY

A chunk

WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18



 ?

 ?

02

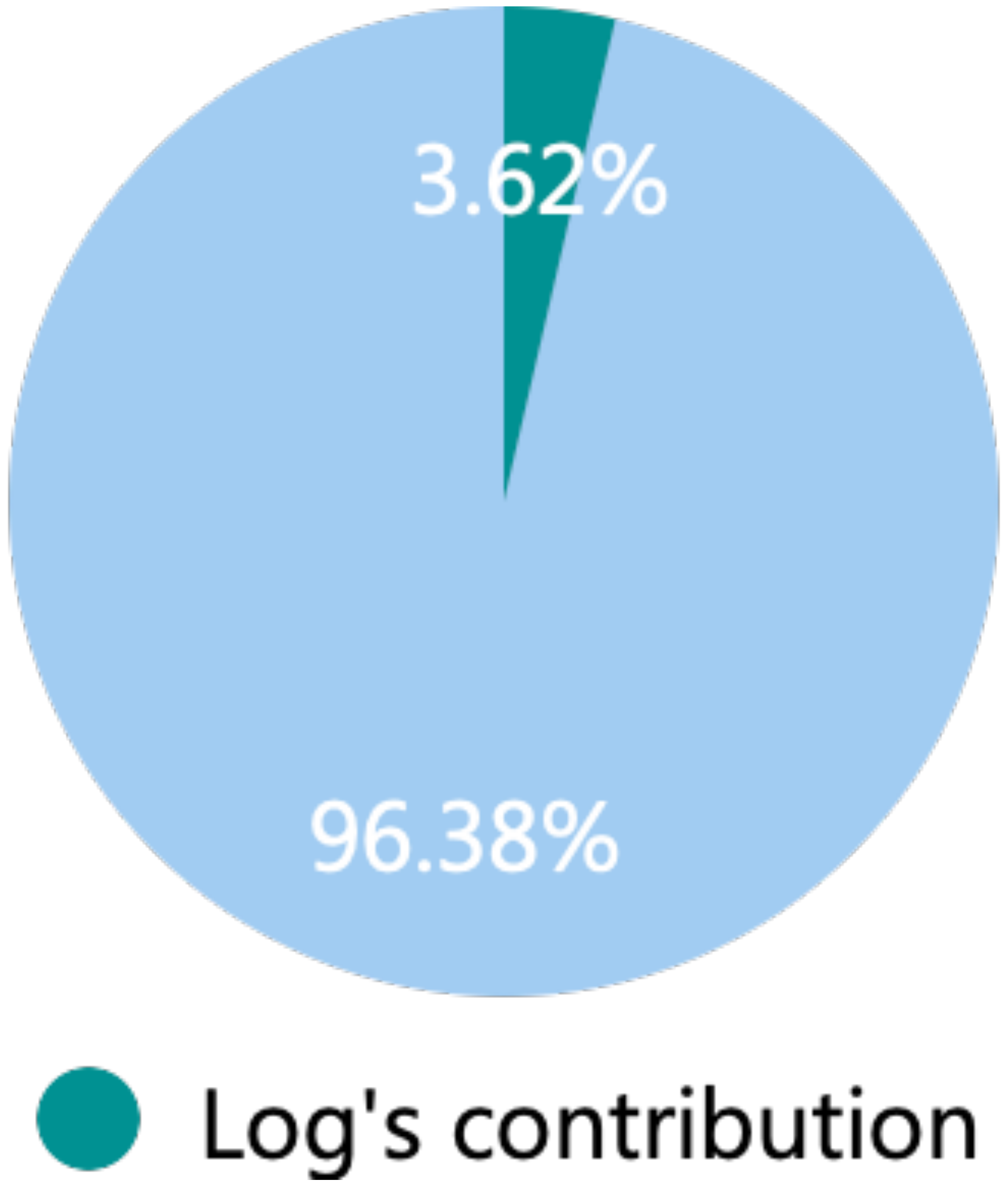
MOTIVATION

Anomaly Characteristics, Case Studies

How do logs manifest system anomalies?

Finding 1

Logs sometimes cannot record fine-grained information and therefore, are not susceptible enough to manifest all system anomalies.

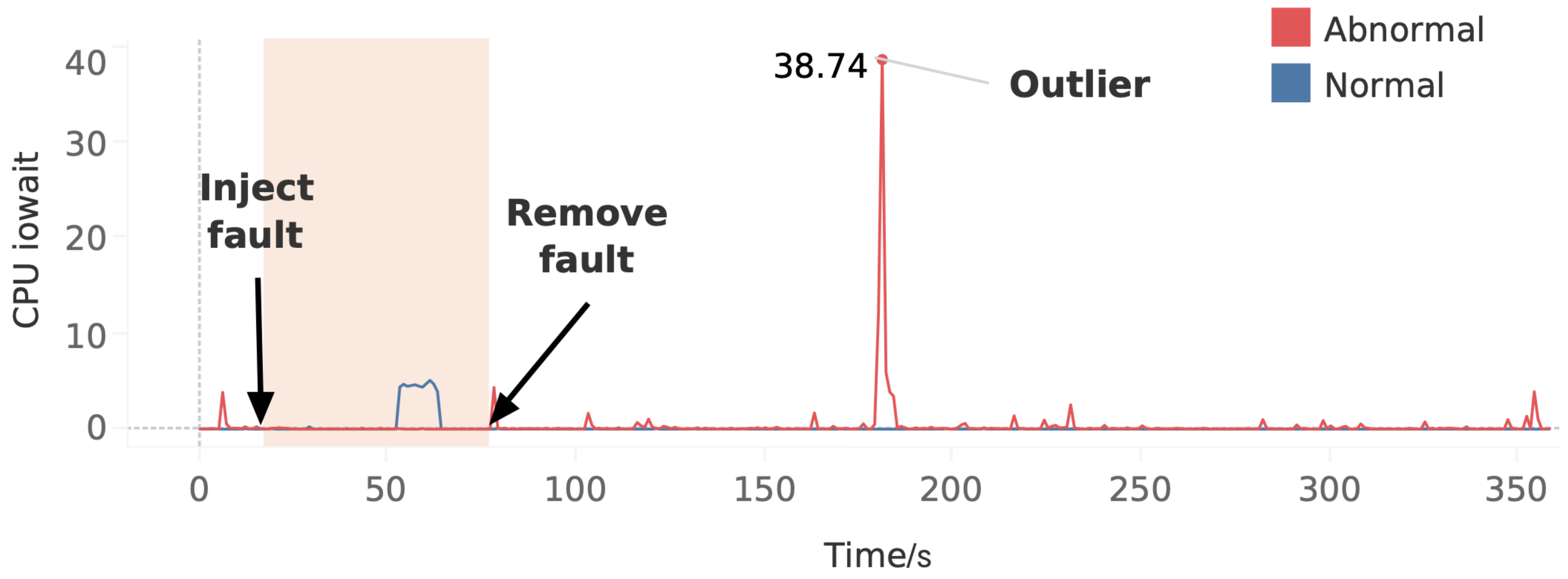


Only 3.62% of positively labeled chunks are anomalous from the log's perspective.

How do metrics manifest system anomalies?

Finding 2

Metrics are insufficient sometimes. Their over-sensitivity may cause false alarms on uncommon yet acceptable fluctuations.



``CPU iowait'' generates a rare heartbeat spike even in the fault-free period.

How do logs & metrics manifest anomalies?

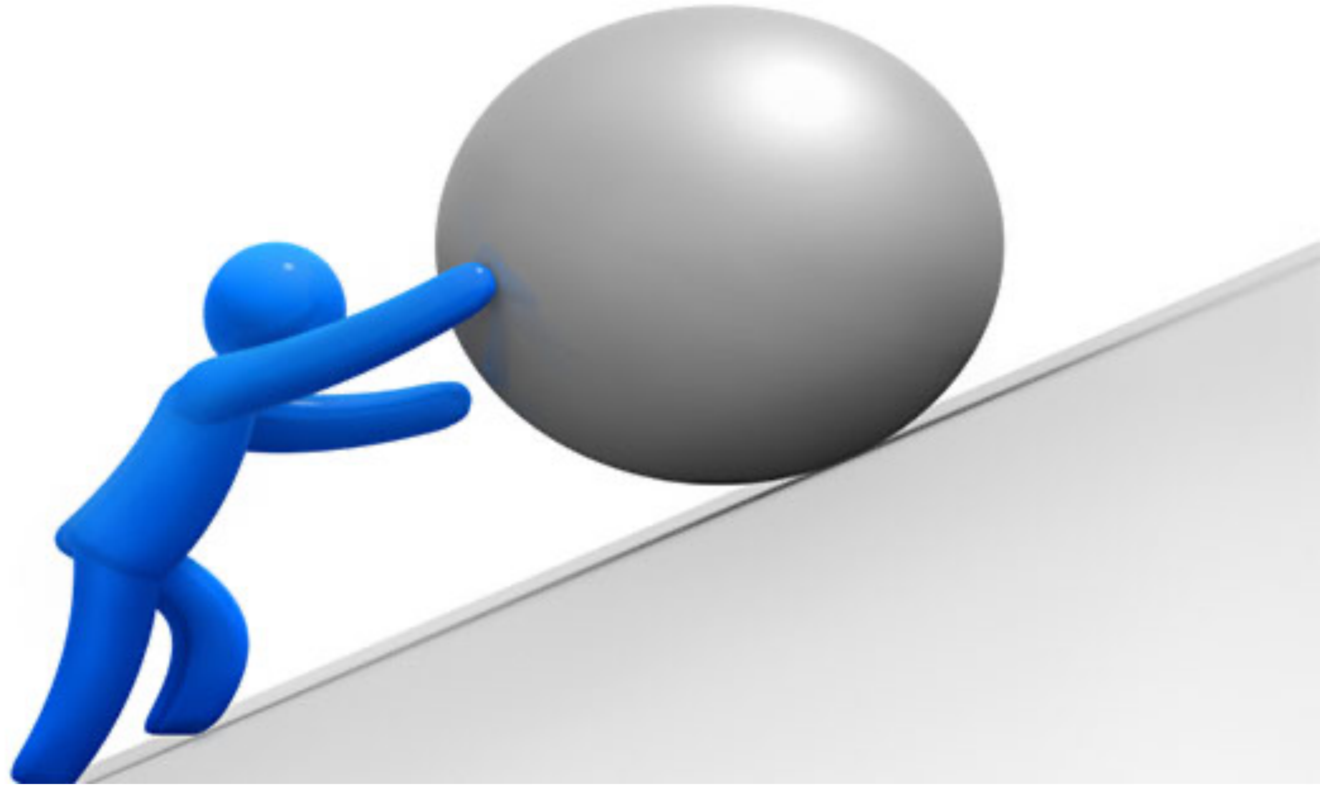
Finding 3

Metrics and logs can both respond to anomalies, but neither is sufficient. They have collaborative and complementary relationships in reflecting anomalies.

TABLE I: Typical faults and the corresponding anomalous manifestations of logs and metrics

Faults	**Anomalies in logs**	**Anomalies in metrics**
Memory hog	Warnings (reaches the memory limit)	Memory-related metrics rise steeply
Virtual memory hog	Errors (reporter thread fails)	CPU and memory-related metrics jitter
I/O hog	Warnings (slow ReadProcessor)	I/O-related metrics rise steeply
Network delay	Warnings (executor heartbeat timeout)	Network-related metrics suddenly drop
Connection flash	Nothing **(silent)**	Network-related metrics suddenly drop and quickly restore
Datanode killed	Errors (excluding datanode)	Related metrics plummet to zero **(silent)**
Secondary namenode killed	Errors (failed to connect to <IP>)	Related metrics plummet to zero **(silent)**

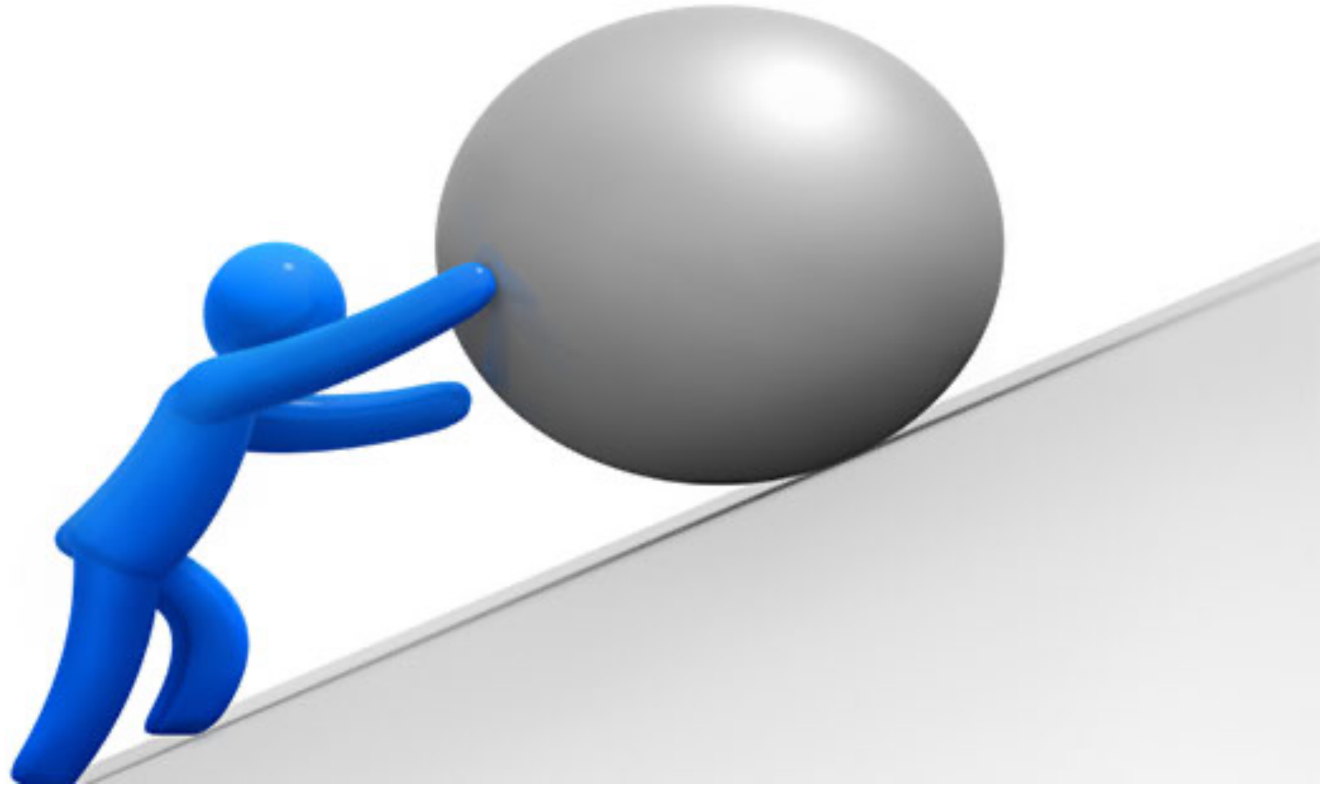
Challenges



Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics’ diverse aspects.

Challenges



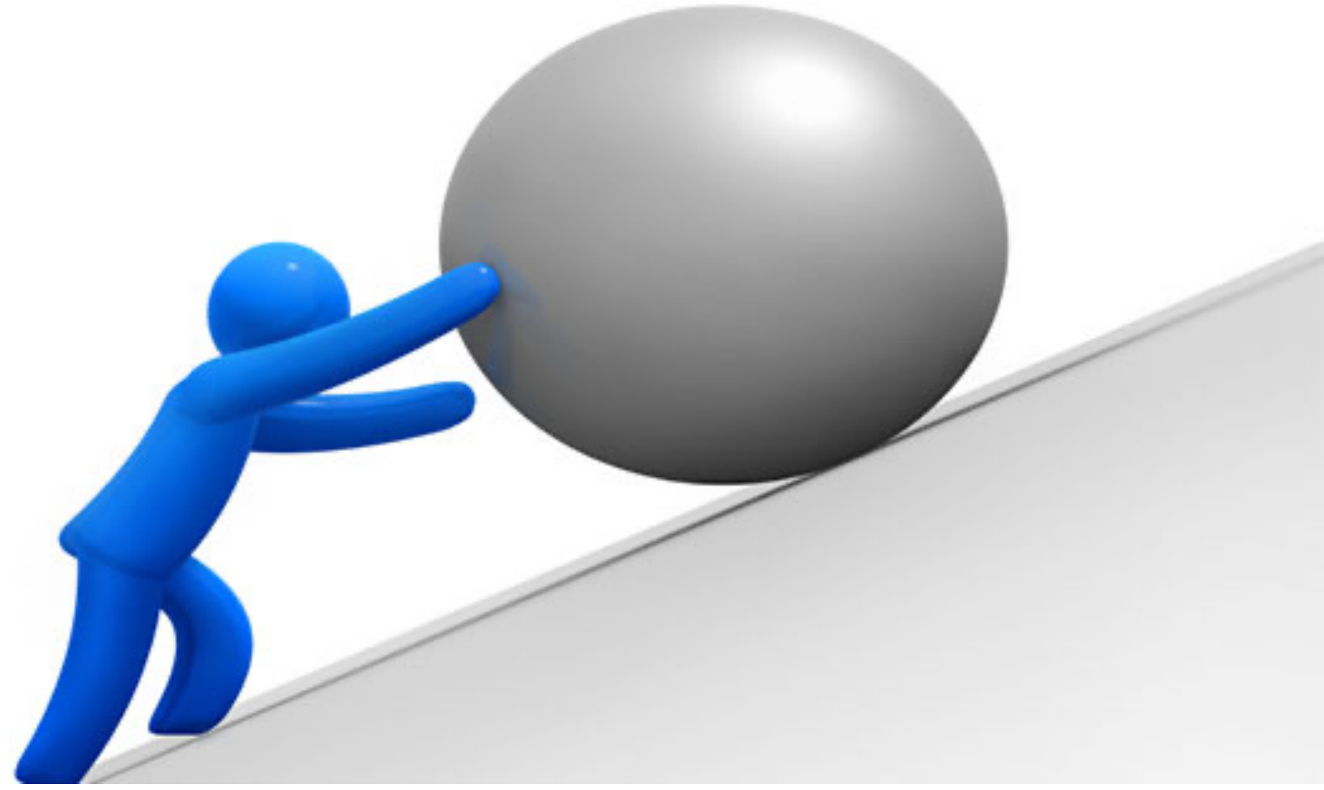
Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics’ diverse aspects.

Significant inter-modal gap:

- Logs and metrics are in different forms.
- Different degrees of anomaly affectedness.

Challenges



Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics’ diverse aspects.

Significant inter-modal gap:

- Logs and metrics are in different forms.
- Different degrees of anomaly affectedness.

Trade-off between cost and accuracy:

- Supervised learning is accurate but costly.
- Unsupervised learning ignores human oversight.

Our Solution

Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics' diverse aspects.

Properly modeling each modality

Significant inter-modal gap:

- Logs and metrics are in different forms.
- Different degrees of anomaly affectedness.

Trade-off between cost and accuracy:

- Supervised learning is accurate but costly.
- Unsupervised learning ignores human oversight.

Our Solution

Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics' diverse aspects.

→ Properly modeling each modality

Significant inter-modal gap:

- Logs and metrics are in different forms.
- Different degrees of anomaly affectedness.

→ Cross-modal attention

Trade-off between cost and accuracy:

- Supervised learning is accurate but costly.
- Unsupervised learning ignores human oversight.

Our Solution

Complex intra-modal information:

- Log semantics and sequential dependencies.
- Metrics' diverse aspects.

→ Properly modeling each modality

Significant inter-modal gap:

- Logs and metrics are in different forms.
- Different degrees of anomaly affectedness.

→ Cross-modal attention

Trade-off between cost and accuracy:

- Supervised learning is accurate but costly.
- Unsupervised learning ignores human oversight.

→ Semi-supervised

03 METHODOLOGY

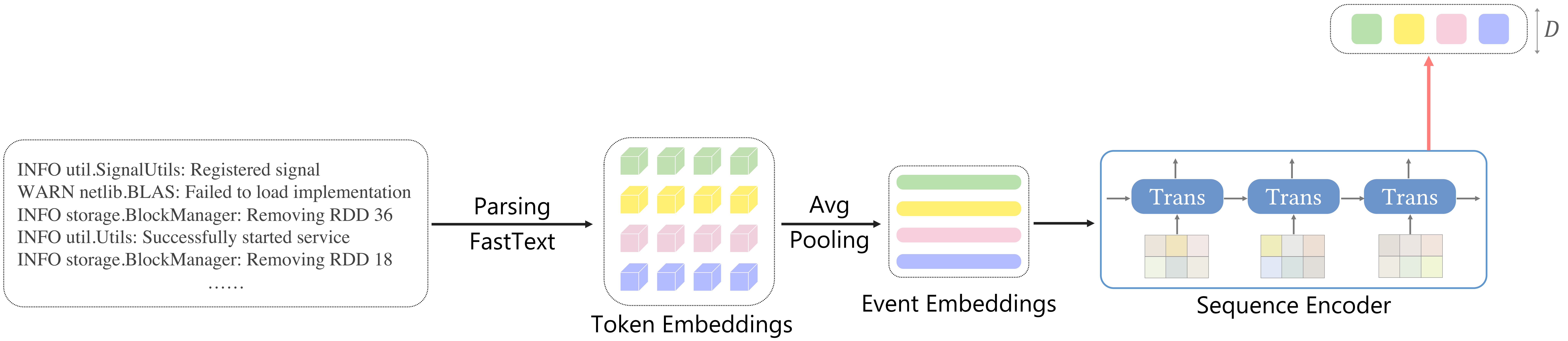
Modal-wise Modeling, Cross-modal Attention

Overview

Metric Modeling
Multivariate Metrics
Grouping
Intra-aspect Encoder
Conv
Conv
Conv
Max
Pooling
Inter-aspect Encoder
Conv
Conv
Conv
Detection
Details
Yes
No
Abnormal?
FC
Heterogeneous Fusion
Global R^g
D
Concat
Fused R^α
Fused R^β
Attnα
Attnβ
$K=V=R^m$
$Q=R^l$
$Q=R^m$
$K=V=R^l$
D
D
Log Modeling
INFO util.SignalUtils: Registered signal
WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18
……
Parsing
FastText
Token Embeddings
Avg
Pooling
Event Embeddings
Trans
Trans
Trans
Sequence Encoder

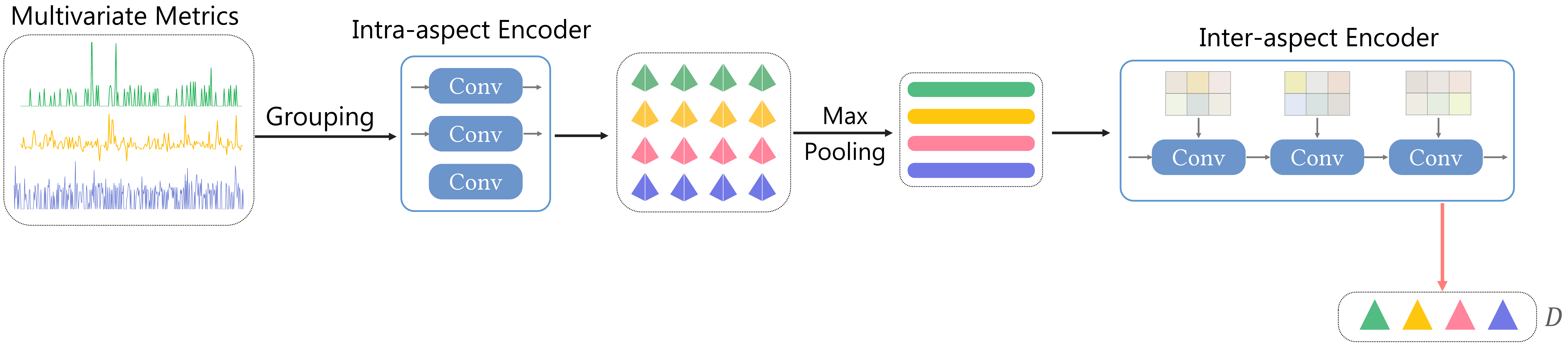
Log Modeling

Log parsing ⇨ Log vectorization ⇨ Log representation learning

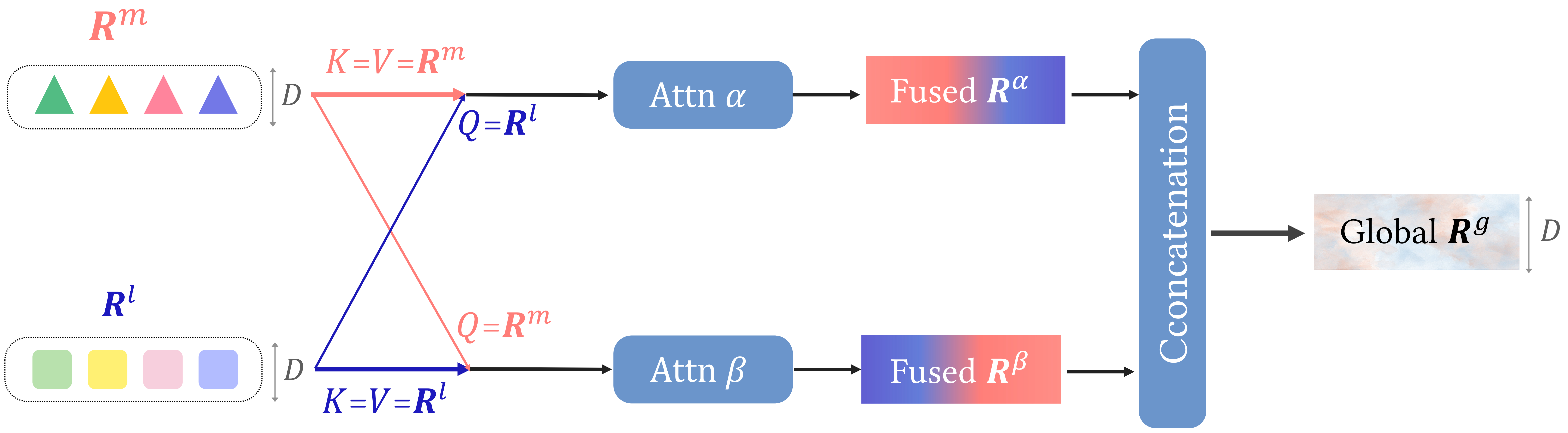


Aspect-aware Metric Modeling

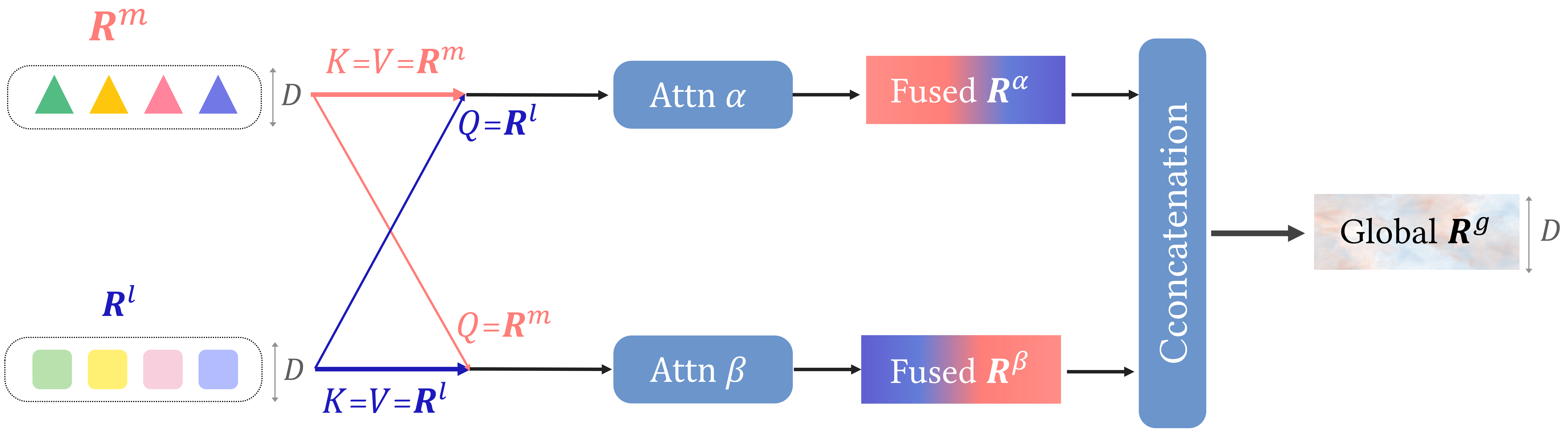
Pre-processing ⇨ Intra-Aspect Encoder ⇨ Inter-Aspect Encoder



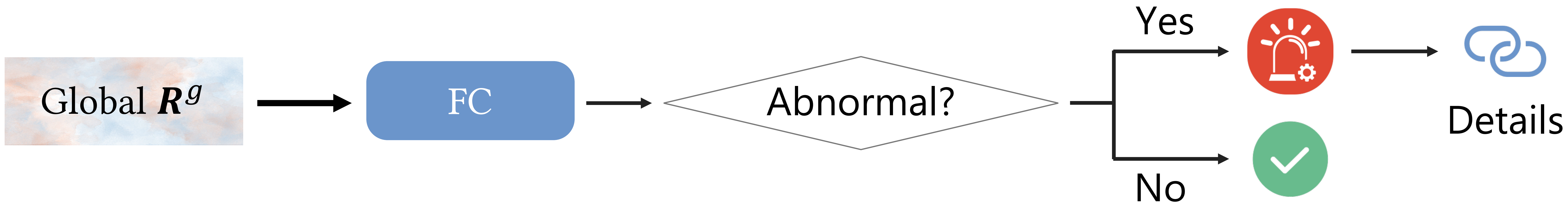
Heterogeneous Representation Fusion



Detection

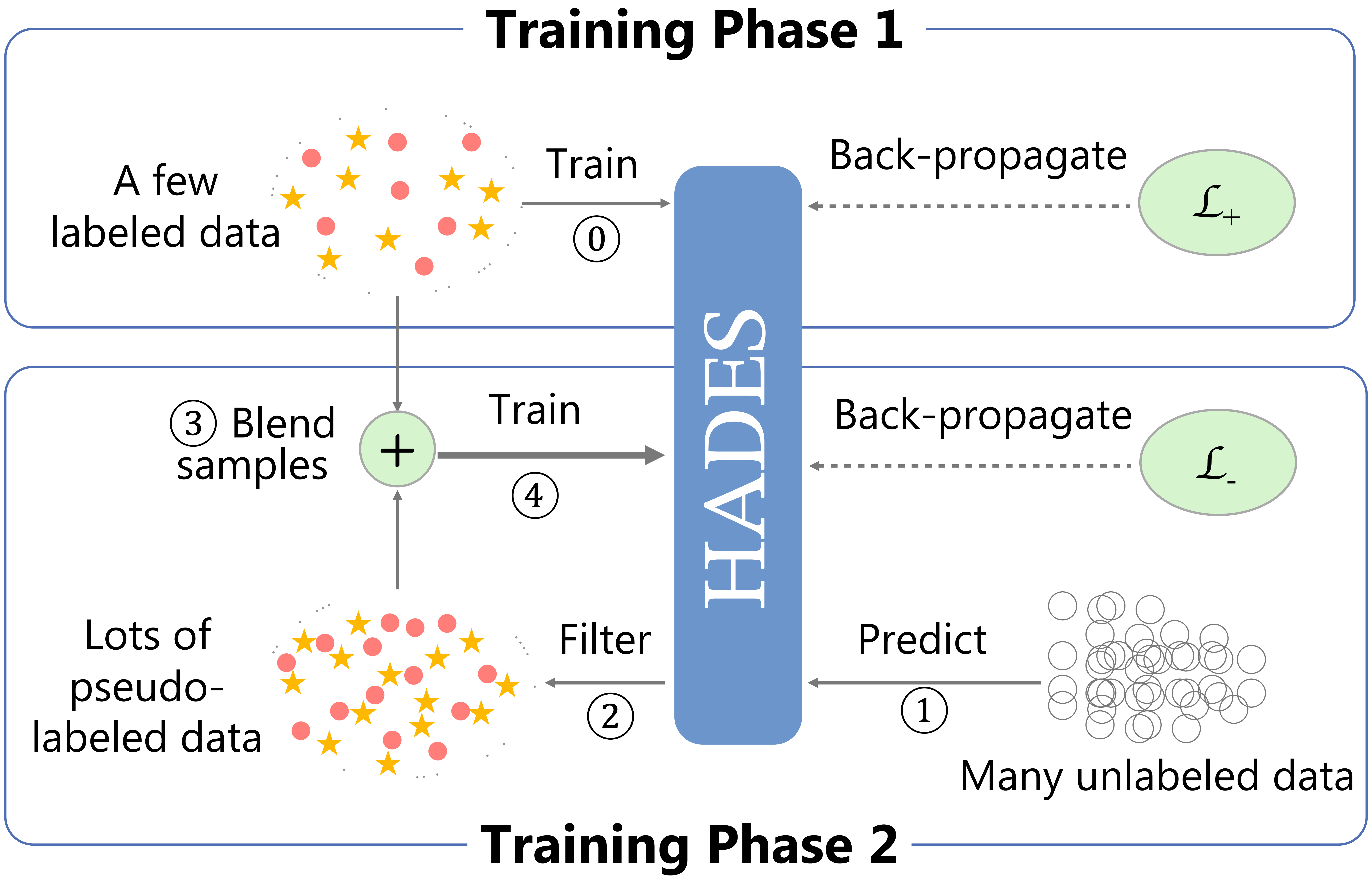


R^m
D
$K=V=R^m$
$Q=R^l$
Attn α
Fused R^α
Cconcatenation
Global R^g
D
R^l
D
$Q=R^m$
$K=V=R^l$
Attn β
Fused R^β



Global R^g
FC
Abnormal?
Yes
No
Details

Semi-supervised Learning




ades
Overview
Details
Settings
Auto Fresh
10
Workload
ID
c0d17d481f47bdd9
Status
Running
Start at
22/03/01T07:00:00
Chunk Info
Time
22/03/01T09:28:00 ~ 22/03/01T09:38:00
Status
Abnormal
Source
Log, Metric
Log File
Path
http://127.0.0.1/
root/workspace/...
Download
Key Metrics
Aspect
Name
Metrics
CPU
%user
I/O
tx/b
I/O
rx/b
Log Preview
INFO storage.BlockManager: Found block rdd_2_3 locally
INFO storage.BlockManager: Found block rdd_2_4 locally
INFO util.SignalUtils: Registered signal
WARN netlib.BLAS: Failed to load implementation
INFO storage.BlockManager: Removing RDD 36
INFO util.Utils: Successfully started service
INFO storage.BlockManager: Removing RDD 18
INFO python.PythonRunner: Times: total = 42, boot = -4131,
init = 4172, finish = 1

04

Evaluation

Effectiveness Comparison, Ablation Study…

How effective is Hades in anomaly detection?

The F1-score of Hades is **9.12%~174.41%** higher than competitors on average.

Table 1: Overall Performance Comparison.

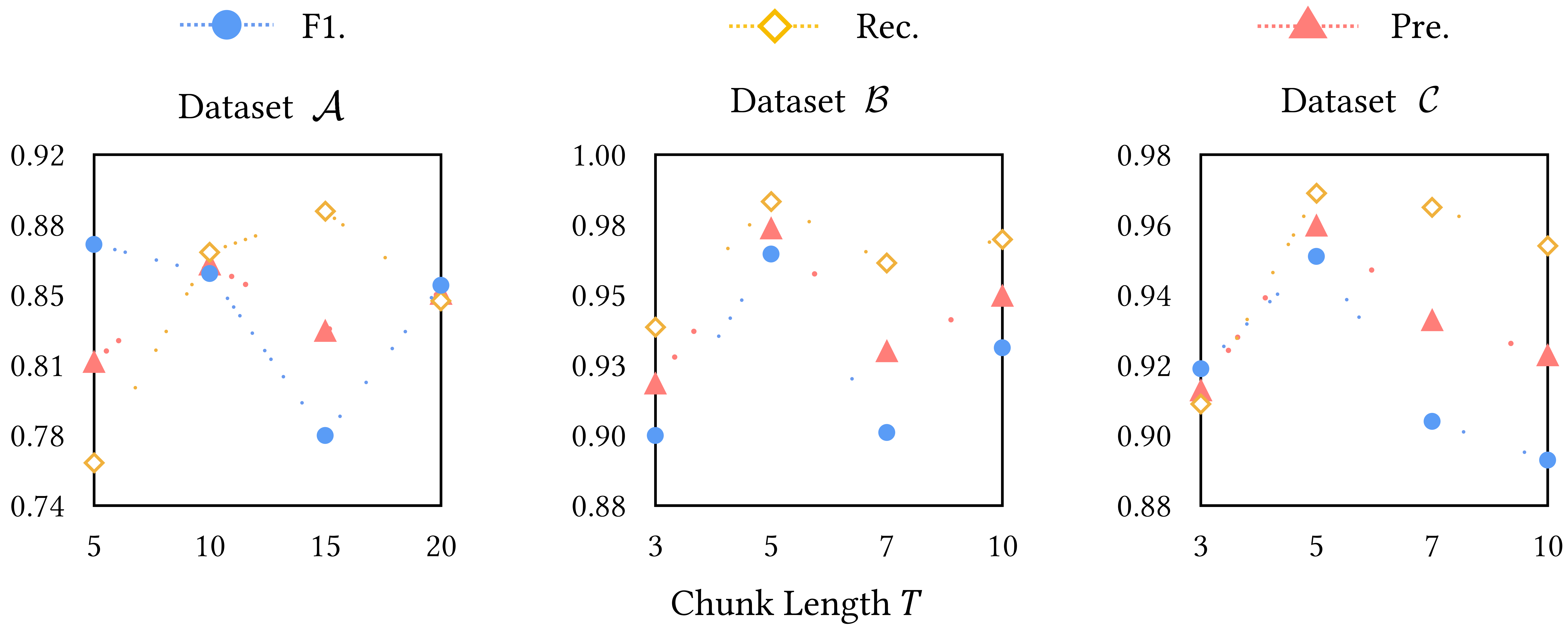
Models	Source	Manner	Dataset $\mathcal{A}$			Dataset $\mathcal{B}$			Dataset $\mathcal{C}$		
			F1	*Rec*	*Pre*	*F1*	*Rec*	*Pre*	*F1*	*Rec*	*Pre*
SVM-$\mathcal{L}$	Log	Supervised	0.289	0.707	0.181	0.541	0.756	0.421	0.481	0.742	0.356
DeepLog	Log	Unsupervised	0.259	0.386	0.194	0.386	0.526	0.305	0.375	0.524	0.292
PLELog	Log	Semi-supervised	0.314	0.602	0.213	0.463	0.618	0.371	0.434	0.597	0.341
LogRobust	Log	Supervised	0.404	0.684	0.287	0.524	0.718	0.413	0.495	0.698	0.383
SVM-$\mathcal{M}$	Metric	Supervised	0.536	0.833	0.395	0.608	0.839	0.477	0.556	0.801	0.426
Adsketch	Metric	Semi-supervised	0.404	0.476	0.351	0.543	0.644	0.470	0.538	0.649	0.459
OmniAnomaly	Metric	Unsupervised	0.681	0.788	0.601	0.827	0.863	0.794	0.812	0.896	0.743
SRCNN	Metric	Unsupervised	0.342	0.614	0.237	0.467	0.701	0.350	0.472	0.586	0.394
SRCNN-s	Metric	Supervised	0.784	0.826	0.745	0.898	0.938	0.861	0.883	0.926	0.844
SCWarn	Log & Metric	Unsupervised	0.321	0.389	0.273	0.497	0.643	0.405	0.491	0.585	0.423
Hades	Log & Metric	Semi-supervised	**0.864**	**0.870**	**0.859**	**0.975**	**0.984**	**0.966**	**0.960**	**0.969**	**0.951**

What is the contribution of each design of Hades?

Table 2: Experimental Results of the Ablation Study.

Models	Dataset $\mathcal{A}$			Dataset $\mathcal{B}$			Dataset $\mathcal{C}$		
	F1	*Rec*	*Pre*	*F1*	*Rec*	*Pre*	*F1*	*Rec*	*Pre*
Hades-supervised	**0.866**	0.878	0.855	**0.979**	0.972	**0.986**	**0.961**	0.953	**0.970**
HADES	0.864	0.870	**0.859**	0.975	**0.984**	0.966	0.960	0.969	0.951

How sensitive is Hades to the length of a chunk?



ARISE
Automated Reliable Intelligent Software Engineering

香港中文大學
CUHK

HUAWEI

Huawei Cloud

THANK YOU

Presenter: Cheryl LEE

Full Paper

Home Page