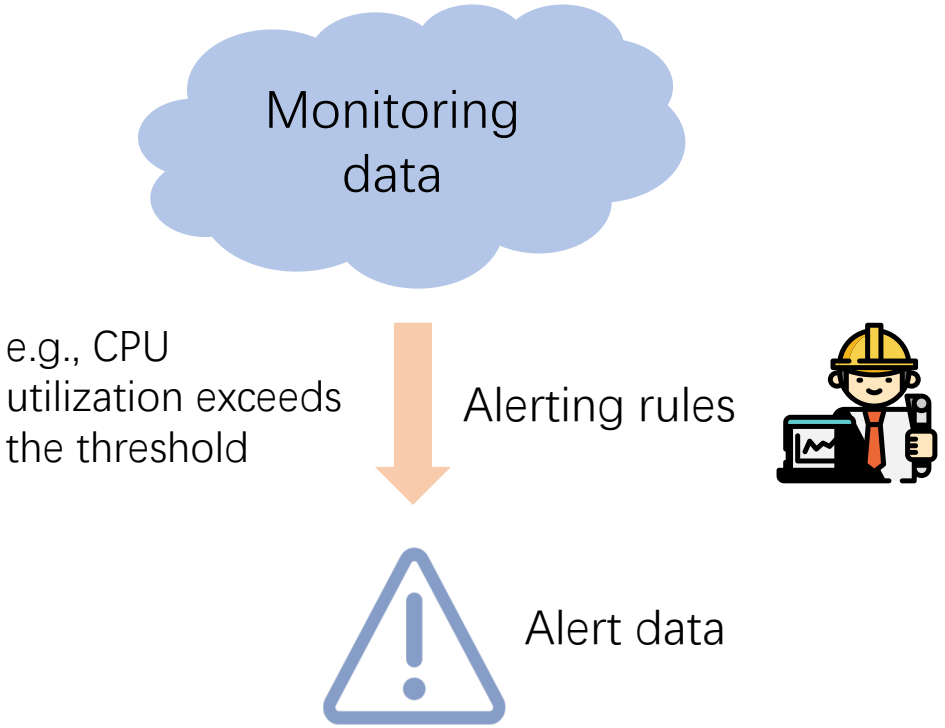
Understanding and Handling Alert Storm for Online Service Systems

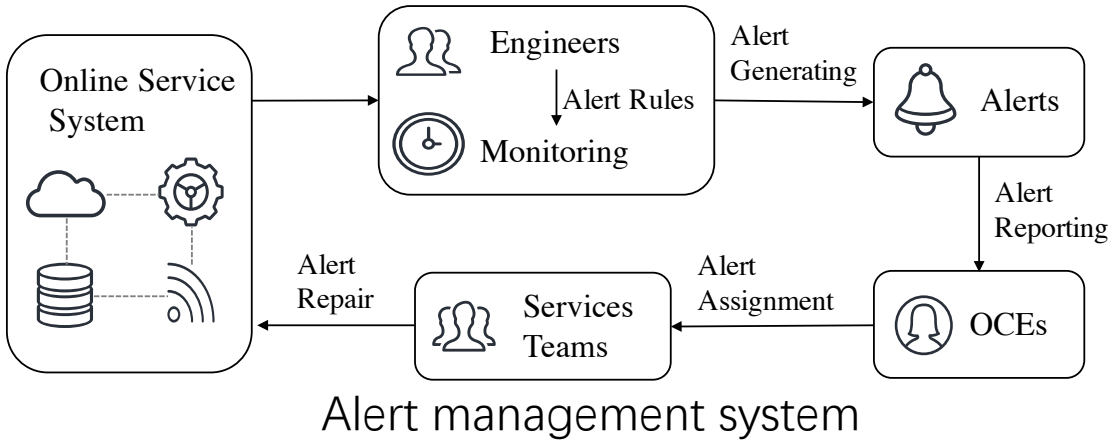
Nengwen Zhao, Junjie Chen, Xiao Peng, Honglin Wang, Xinya Wu, Yuanzong Zhang, Zikai Chen, Xiangzhong Zheng , Xiaohui Nie , Gang Wang , Yong Wu, Fang Zhou , Wenchi Zhang , Kaixin Sui , Dan Pei

ICSE SEIP 2020

Background

Service quality and user experience are vital for online service systems.





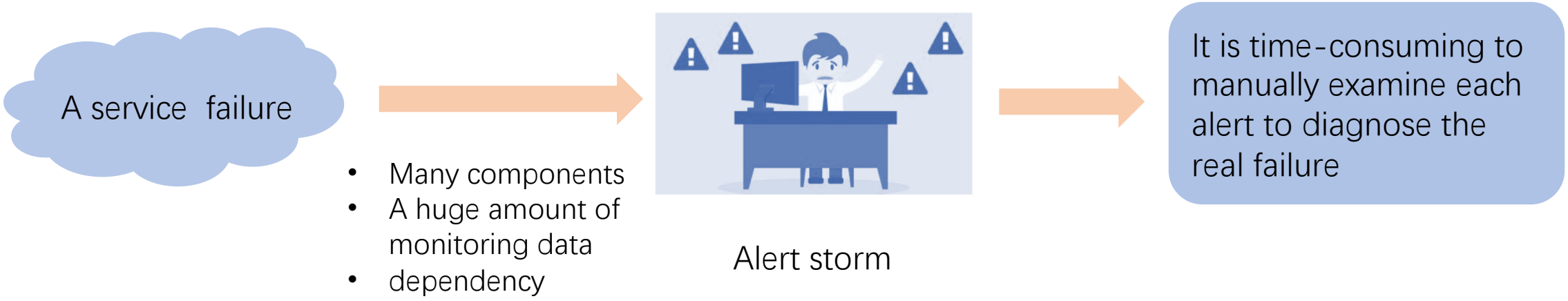
Alert management system

Time	Severity	Type
2019-02-20 10:04:32	P2-error	Memory
AppName	**Server**	**Close Time**
E-BANK	IP(*.*.*.*)	2019-02-20 10:19:45
Content		
Current memory utilization is 79% (Threshold is 60%).		
Resolution Record		
Contact the service engineers responsible for E-BANK and get a reply that there is no effect on business, then close the alert.		

Sample of alert data

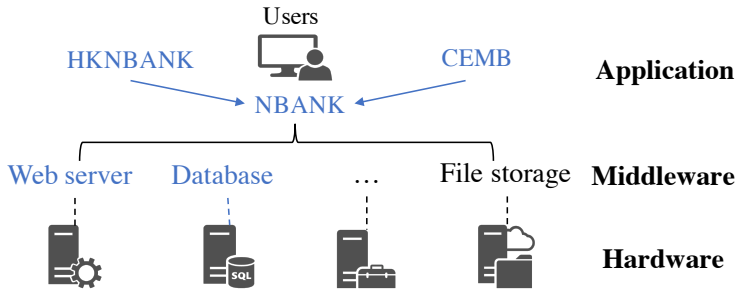
Background

Due to the large scale and complexity of online service systems, service failures are inevitable.

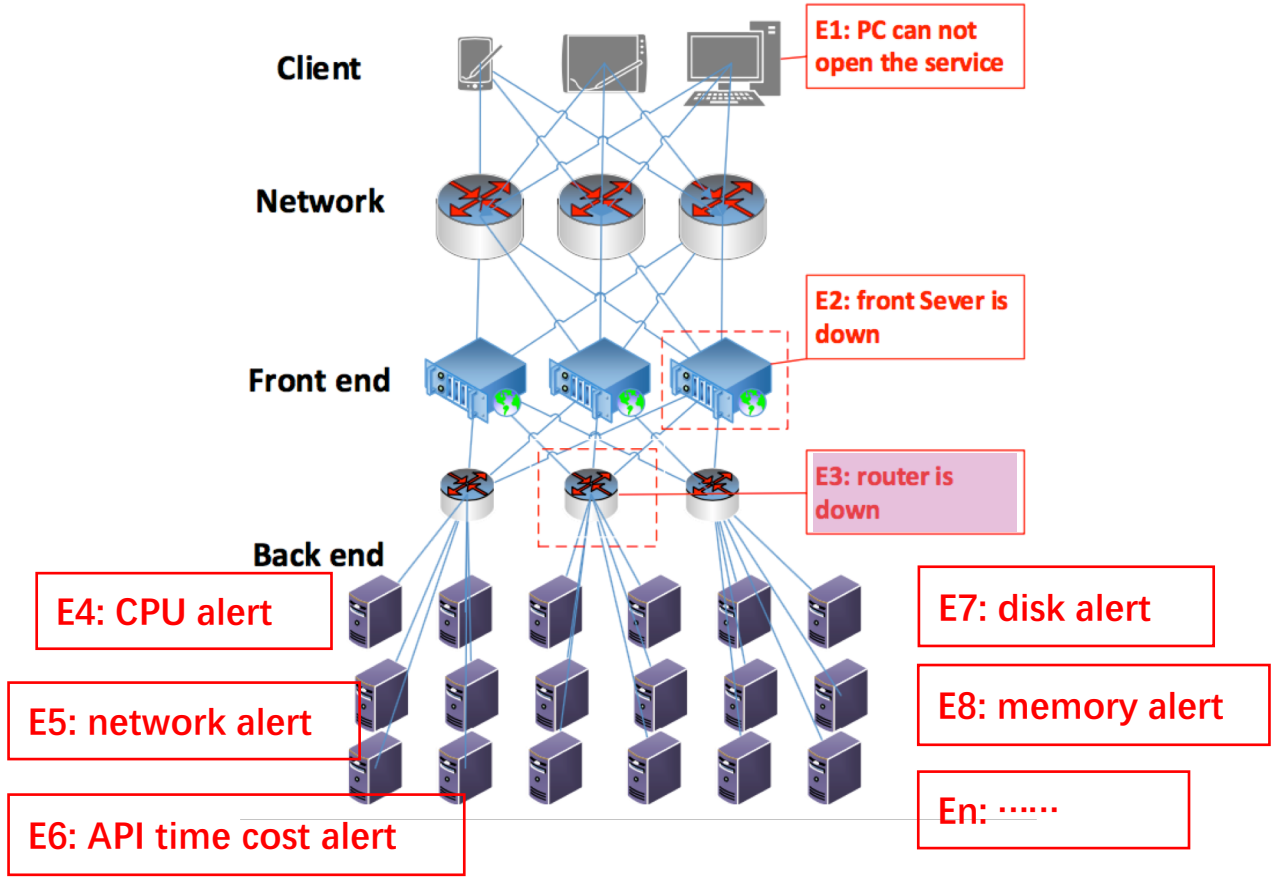


Therefore, to facilitate the assurance of the service quality, understanding and handling alert storm are very essential.

Case Study



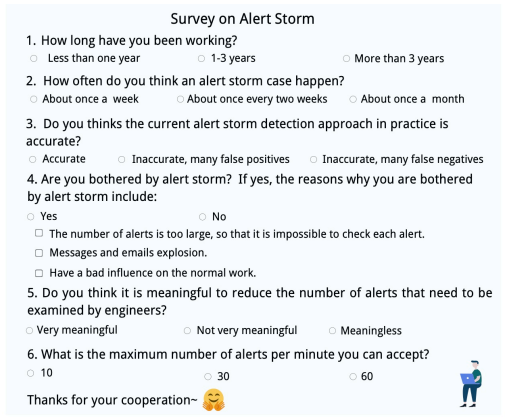
- Database server down.
- NBANK cannot receive data and generate alerts.
- Other services (HKNBANK, CEMB) calling it also generated alerts.



Router is down

Empirical Study

A large amount of alert data from a large commercial bank


Survey on Alert Storm

1. How long have you been working?
○ Less than one year ○ 1-3 years ○ More than 3 years

2. How often do you think an alert storm case happen?
○ About once a week ○ About once every two weeks ○ About once a month

3. Do you thinks the current alert storm detection approach in practice is accurate?
○ Accurate ○ Inaccurate, many false positives ○ Inaccurate, many false negatives

4. Are you bothered by alert storm? If yes, the reasons why you are bothered by alert storm include:
○ Yes ○ No
☐ The number of alerts is too large, so that it is impossible to check each alert.
☐ Messages and emails explosion.
☐ Have a bad influence on the normal work.

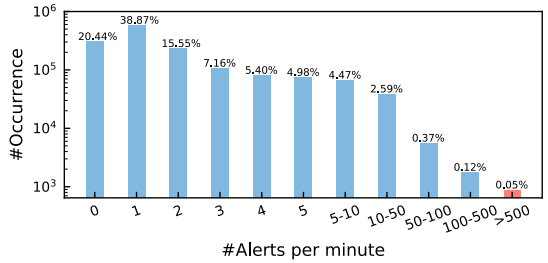
5. Do you think it is meaningful to reduce the number of alerts that need to be examined by engineers?
○ Very meaningful ○ Not very meaningful ○ Meaningless

6. What is the maximum number of alerts per minute you can accept?
○ 10 ○ 30 ○ 60

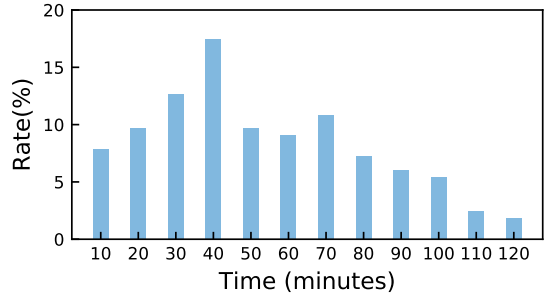
Thanks for your cooperation~ 🤗

Observations

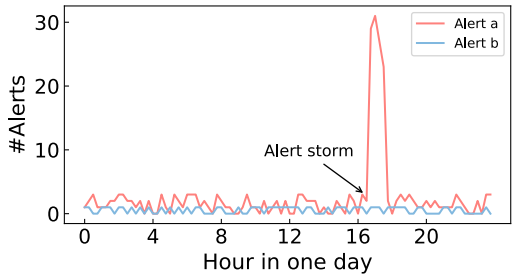
1. Alert storm occurs frequently (about once a week) and brings great trouble to engineers in practice.



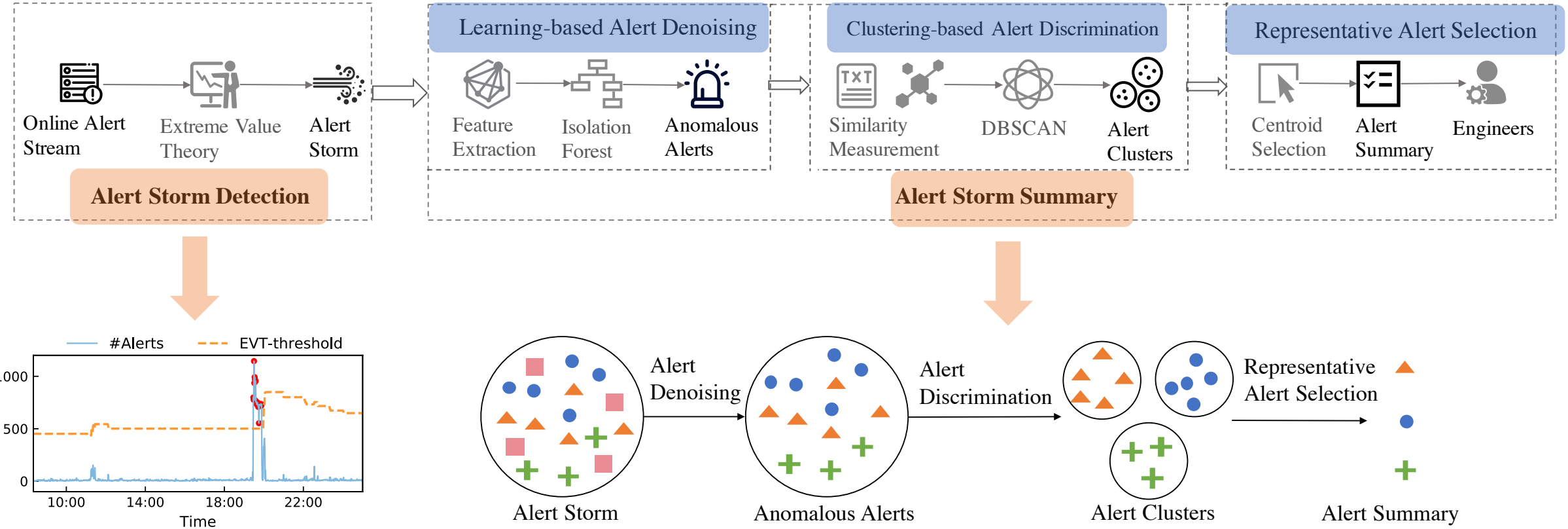
2. The current practice of identifying alert storm is just to set a fixed threshold, which cannot fit the dynamic environment.



3. Some alerts in alert storm are irrelevant to the failure, and also many alerts relevant to the failure have certain correlation



Approach


Online Alert Stream
Extreme Value Theory
Alert Storm
Alert Storm Detection
Learning-based Alert Denoising
Feature Extraction
Isolation Forest
Anomalous Alerts
Clustering-based Alert Discrimination
Similarity Measurement
DBSCAN
Alert Clusters
Representative Alert Selection
Centroid Selection
Alert Summary
Engineers
Alert Storm Summary
#Alerts
EVT-threshold
1000
500
0
10:00
14:00
18:00
22:00
Time
Alert Storm
Alert Denoising
Anomalous Alerts
Alert Discrimination
Alert Clusters
Representative Alert Selection
Alert Summary

Approach

Alert Storm Summary

1

Learning-based alert denoising

- Anomaly detection problem
- Features: alert attributes
- Isolation forest

2

Clustering-based Alert Discrimination

- Similarity measurement
  - Textual similarity: Jaccard distance
  - Topological similarity: graph path
- Clustering

3

Representative Alert Selection

- Pick the centroid of each cluster

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

Evaluation

Dataset: 166 alert storm cases from the real world

1. EVT-based approach can detect alert storm more accurately compared with the traditional threshold-based method, achieving the F1-score larger than 0.9.

Datasets	A			B			C		
Methods	P	R	F1	P	R	F1	P	R	F1
EVT	**0.92**	**0.96**	**0.94**	**0.90**	**0.97**	**0.93**	**0.95**	**0.96**	**0.95**
Threshold	0.82	0.99	0.90	0.75	0.92	0.83	0.59	0.91	0.72

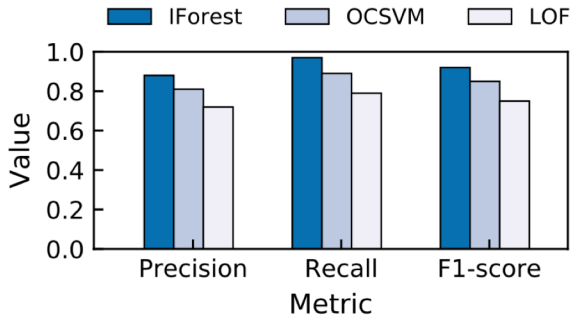
2. Alert storm can reduce the number of alerts that engineers need to examine by larger than 98%, and accurately recommend representative alerts related to the failure.

Datasets	Raw	Severity	Denoising	**Summary**
A	0%	88.7%	6.9%	**98.8%**
B	0%	85.6%	5.1%	**98.2%**
C	0%	84.1%	8.4%	**99.1%**

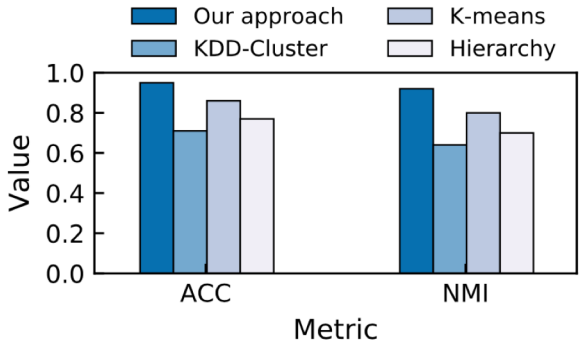
Method	Raw	Severity	W/o denoising	**Summary**
Precision	0.08	0.42	0.64	**0.75**

Evaluation

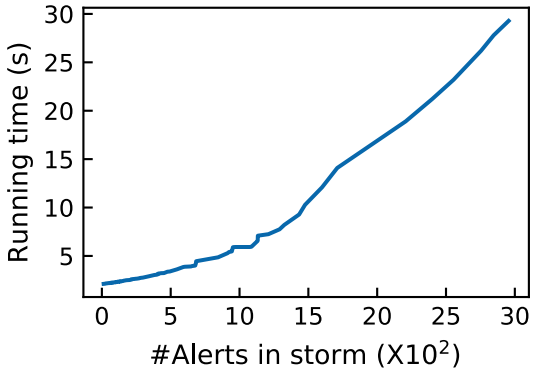
3 Two major components of alert summary are effective and outperform other baseline methods.



(a) Alert denoising



(b) Alert clustering



4 Our alert summary approach can achieve a relative short response time, which is user-friendly for engineers.

Conclusion

1. To better understand the alert storm, we conduct the first empirical study to investigate alert storm based on the large-scale real-world alert data.

2. We propose a novel approach to handling alert storm, which can detect the alert storm accurately and recommend a small set of typical alerts to engineers from the numerous alerts.

3. The experimental results show that our approach is indeed effective in both alert storm detection and alert storm summary.

Thank you!

Q&A

znw17@mails.tsinghua.edu.cn

ICSE SEIP 2020