Nankai University

Tsinghua University

An Empirical Analysis of Anomaly Detection Methods for Multivariate Time Series

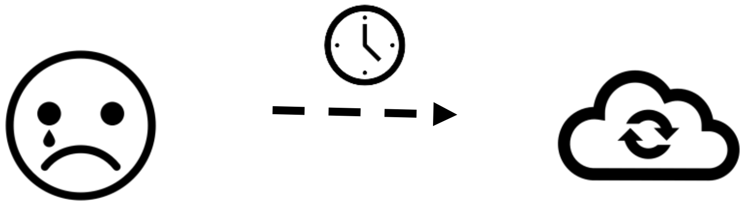
Dongwen Li, Shenglin Zhang, **Yongqian Sun*,** Yang Guo, Zeyu Che, Shiqi Chen, Zhenyu Zhong, Minghan Liang, Minyi Shao, Mingjie Li, Shuyang Liu, Yuzhi Zhang, Dan Pei

Name: **Yongqian Sun**

Date: **2023/10/16**

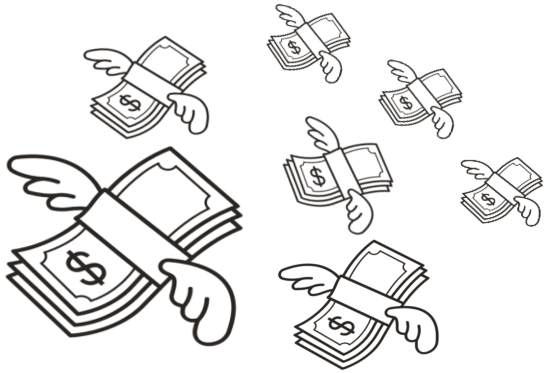
Content

The Reliability of Service Systems is Important



High Latency

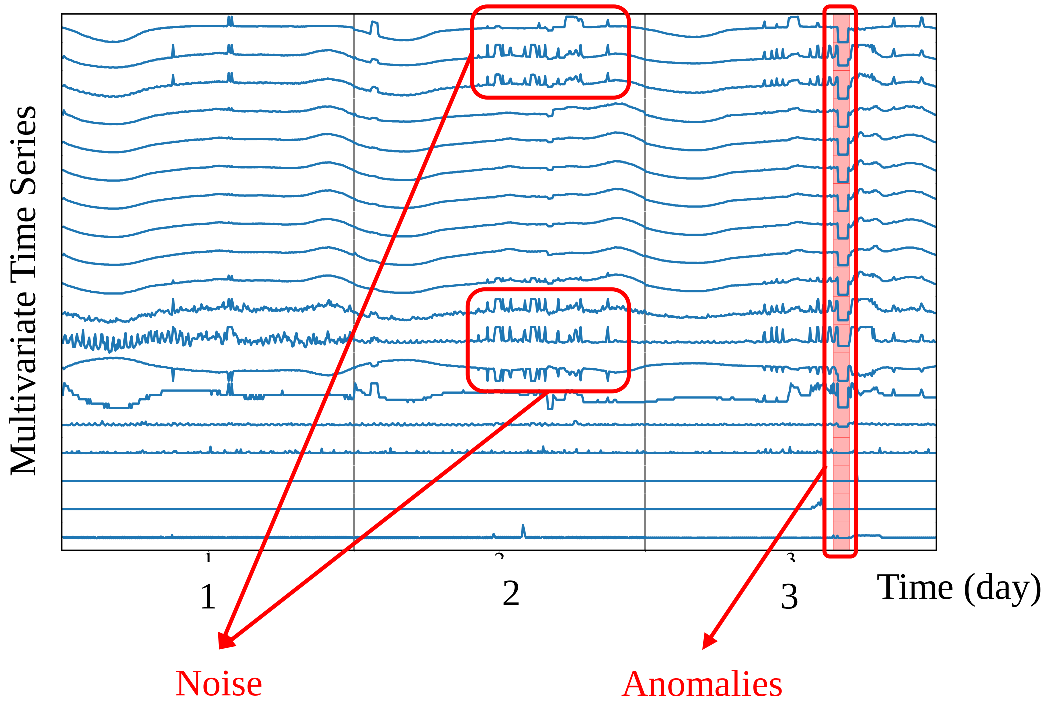
Users



Service Providers

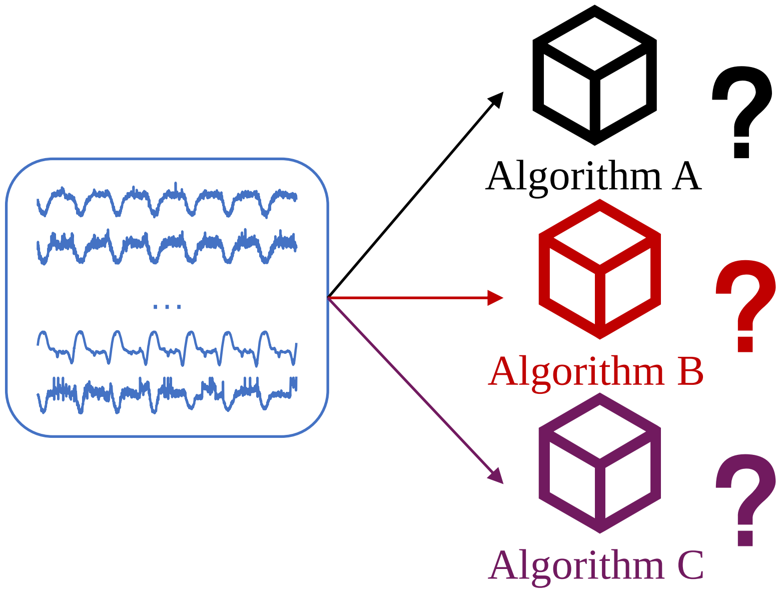
MTS Anomaly Detection

- Many well-performed multivariate time series (MTS) anomaly detection algorithms are proposed.
  - DAGMM[ICLR 2018], USAD[ACM 2020], OmniAnomaly [ACM 2019], DOMI [TON 2022], SDFVAE [WWW 2021], InterFusion [ACM 2021], JumpStarter [ATC 2021], JumpStarter [ATC 2021], GDN [AAAI 2021].
- Determine whether the behaviors of system instances deviate from the normal patterns.



Algorithm Selection

- None of the algorithms is adaptable to all scenarios.
- There is currently **a lack of comprehensive analysis work** of these algorithms to guide operators in selecting the appropriate one in practice.



No objective guidance!

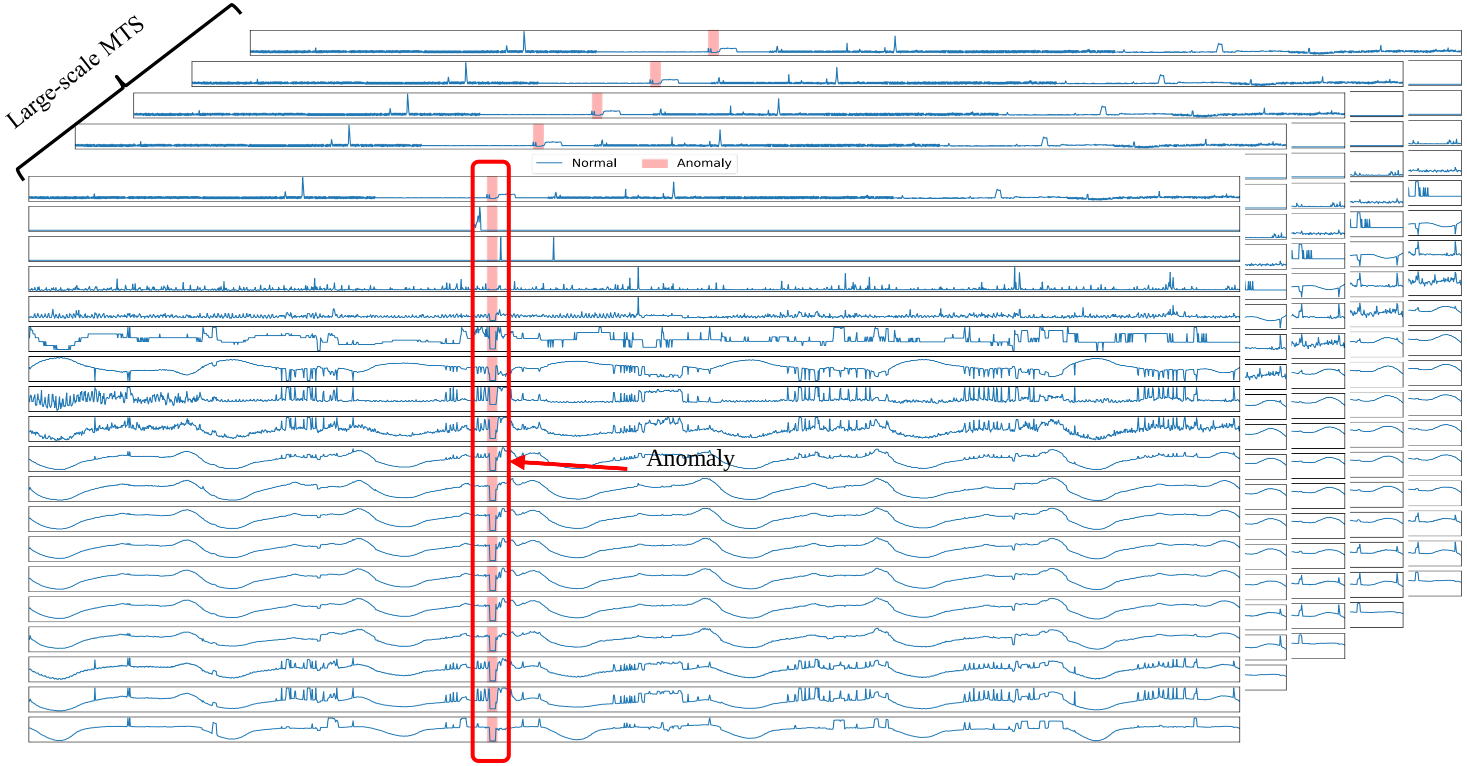
- Goals:
  - **Gain comprehensive understanding of SOTA anomaly detection algorithms**
  - **provide tailored recommendations for algorithm selection**

Content

Challenge 1: Large-Scale MTS

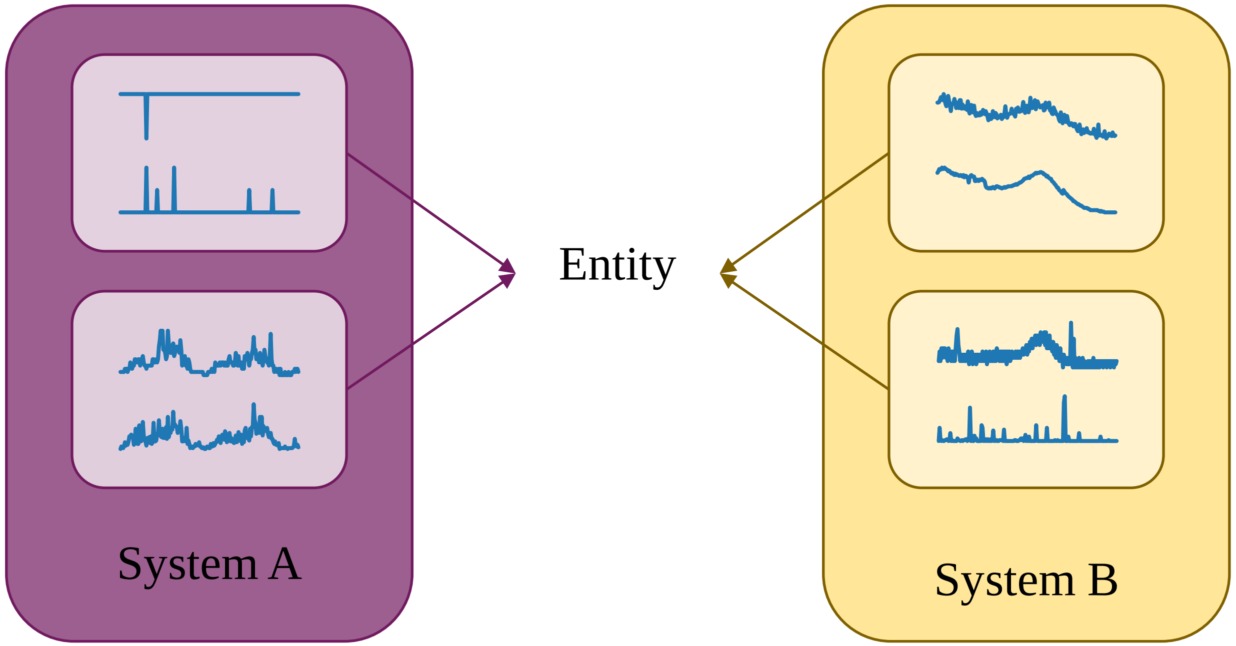
南開大學 Nankai University

- A vast number of system instances resulting in massive MTS generated daily.
- Deep learning-based models often possess complex structures and require high training resources.



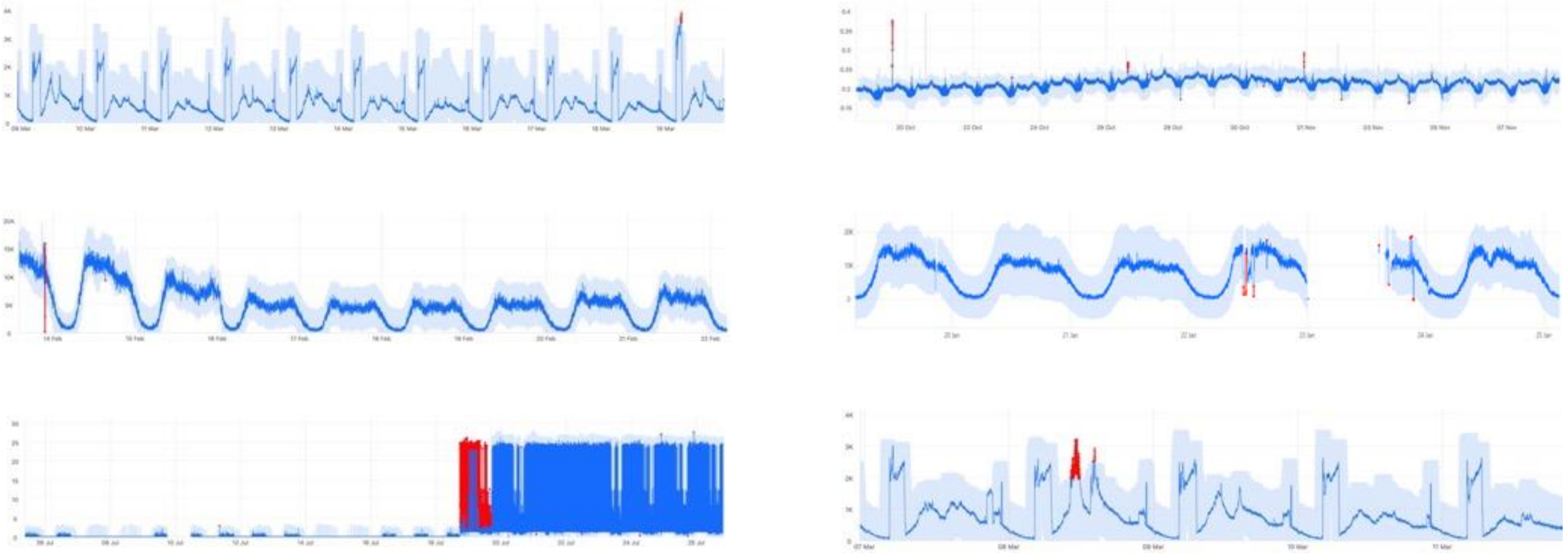
Challenge 2: Various MTS Characteristics

- Data collected from various systems and scenarios often display distinct characteristics.
- Even in the same system, different components can exhibit different MTS characteristics.
- Service systems experience changes, leading to changes in the underlying patterns of MTS.



Challenge 3: Various Anomaly Patterns

- ➢MTS exhibits various types of anomalies when different failures occur or when it is subjected to different attacks.
- ➢A single algorithm, whether supervised or unsupervised, is usually insufficient to detect all types of anomalies.



Content

Research Questions

Nankai University

- RQ1: What data characteristics and anomaly types exist in MTS?
- RQ2: What are the characteristics of the most popular unsupervised algorithms?
- RQ3: How do the existing algorithms work in practice?

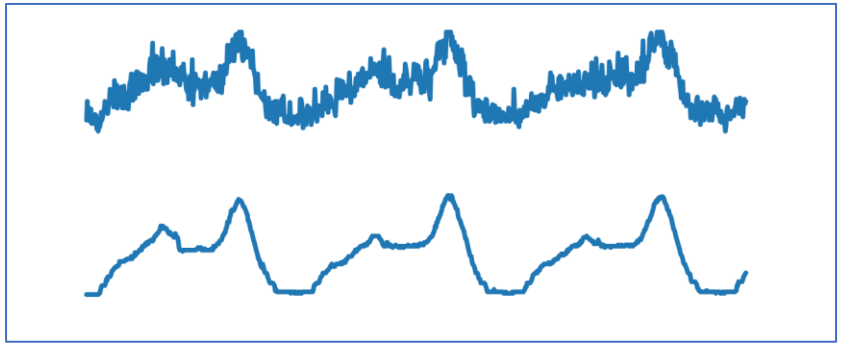
Datasets

- Two private datasets (D1 and D2) from our partner companies.
- Six public datasets from various practical scenarios.

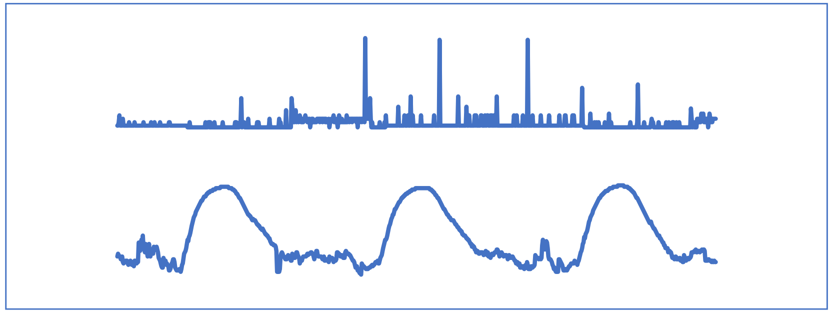
Dataset	**Source**	**Scenarios**	**#Entities**	**#Metrics**	**Time Interval**	**#Train**	**#Test**	**Anomalies (%)**
D1	A global content service provider	Web services.	26	49	30 sec	14400	23040	0.05
D2	An Internet service provider	Network operation service.	107	22	15 min	672	672	0.02
SMD [ACM 2019]	An Internet company	/	28	38	1 min	28479	28479	0.04
ASD [ACM 2021]	An Internet company	/	12	19	5 min	8640	4320	0.05
SMAP [ACM 2018]	NASA	Global measurements of soil moisture and its freeze-thaw status.	54	25	1 min	2818	7331	0.13
MSL [ACM 2018]	NASA	The Mars rover Curiosity's operations.	27	55	1 min	4308	6100	0.11
SWaT [IEEE 2016]	A water treatment plant	The real-world industrial water treatment plant operation status.	1	51	1 sec	496800	449919	0.12

Data Characteristics

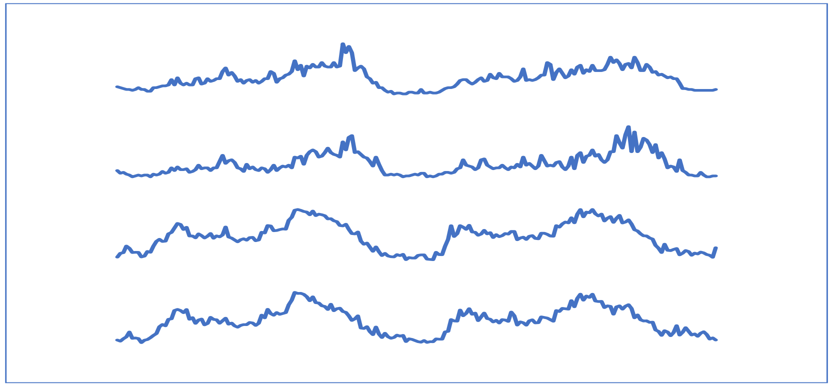
- Smoothness
  - The level of fluctuation between adjacent data points.



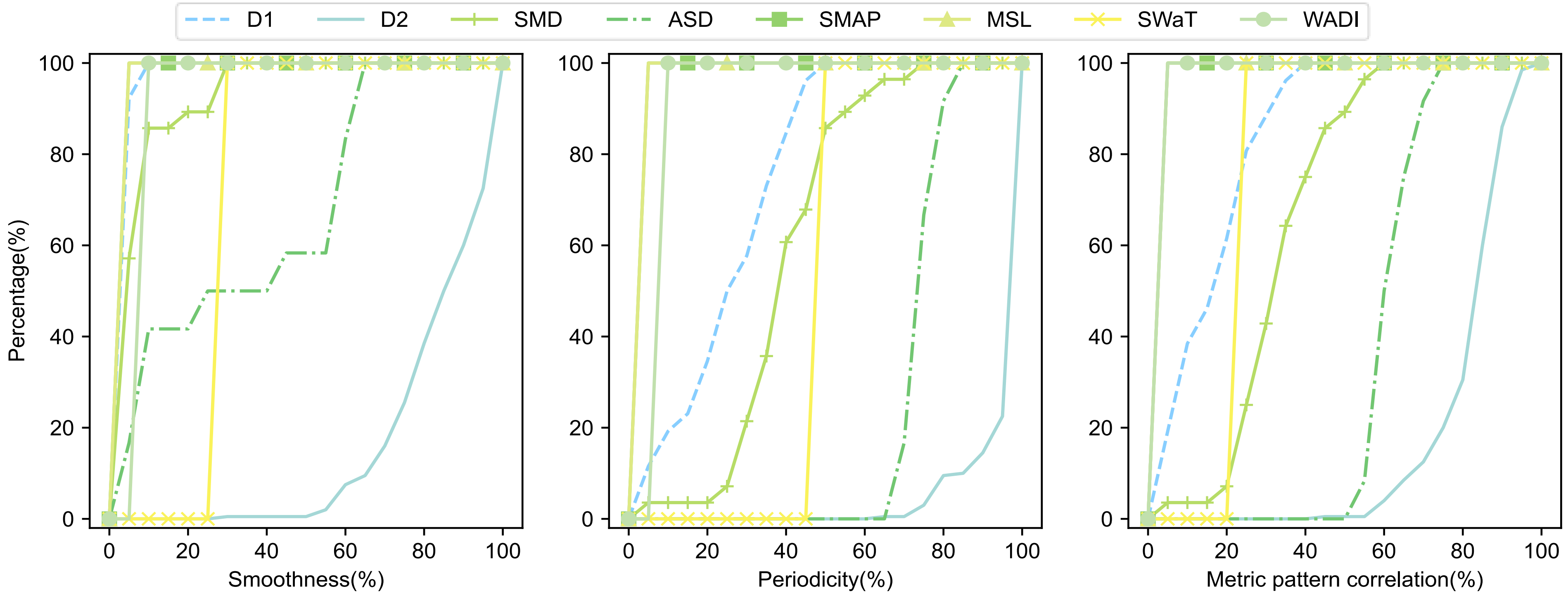
- Periodicity
  - Provide useful historical data.



- Metric correlation
  - Provide complementary information from multiple metrics.



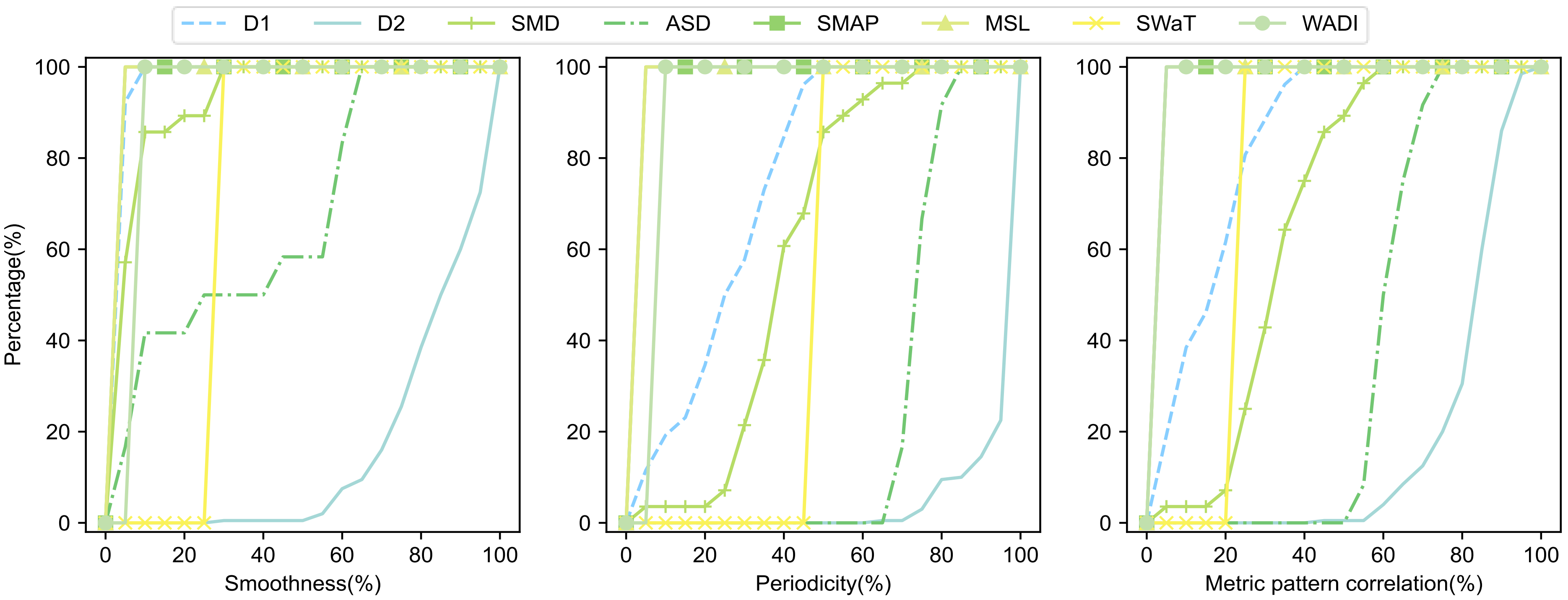
- Smoothness
  - Most datasets exhibit a relatively unsmooth characteristic.
  - In the same dataset, each MTS exhibits unique characteristics.



Data Characteristics

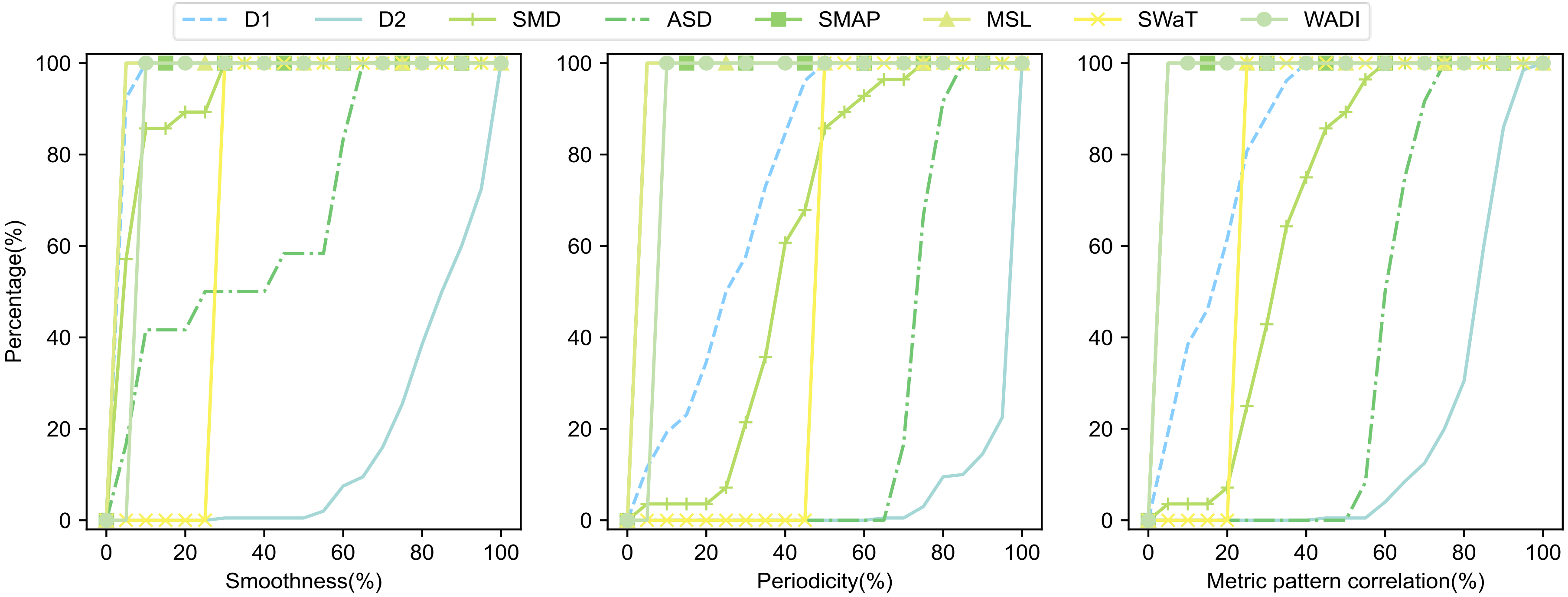
Nankai University

- Periodicity
  - Varies across the datasets.
  - SMAP, MSL, and WADI predominantly comprise non-periodic MTS.
  - D2 exhibits relatively strong periodicity.



Nankai University

- Metric correlation
  - Few correlation in SMAP, SWaT, and WADI.
  - Others exhibit some level of metric correlation.



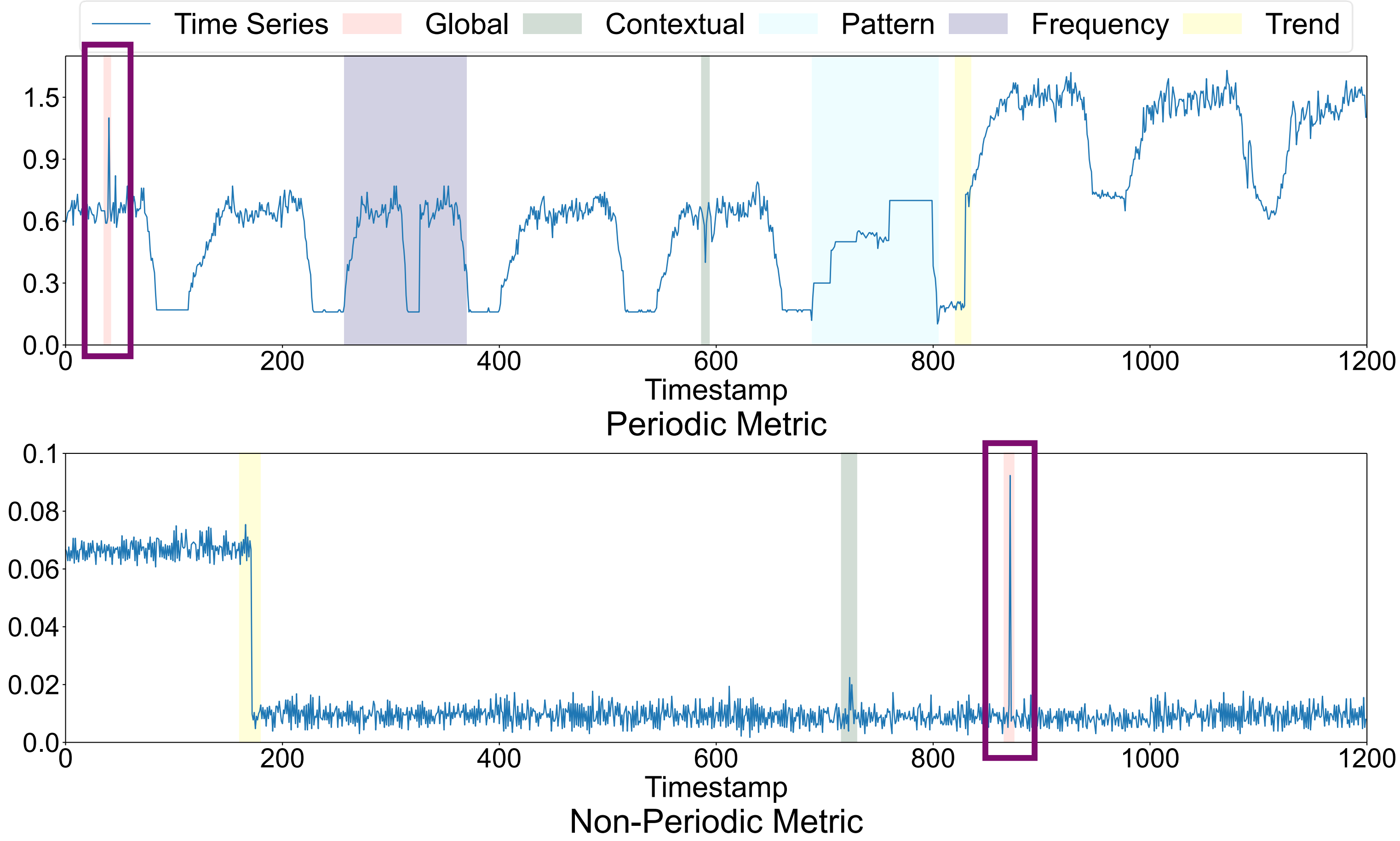
Anomaly Types

南開大學 Nankai University

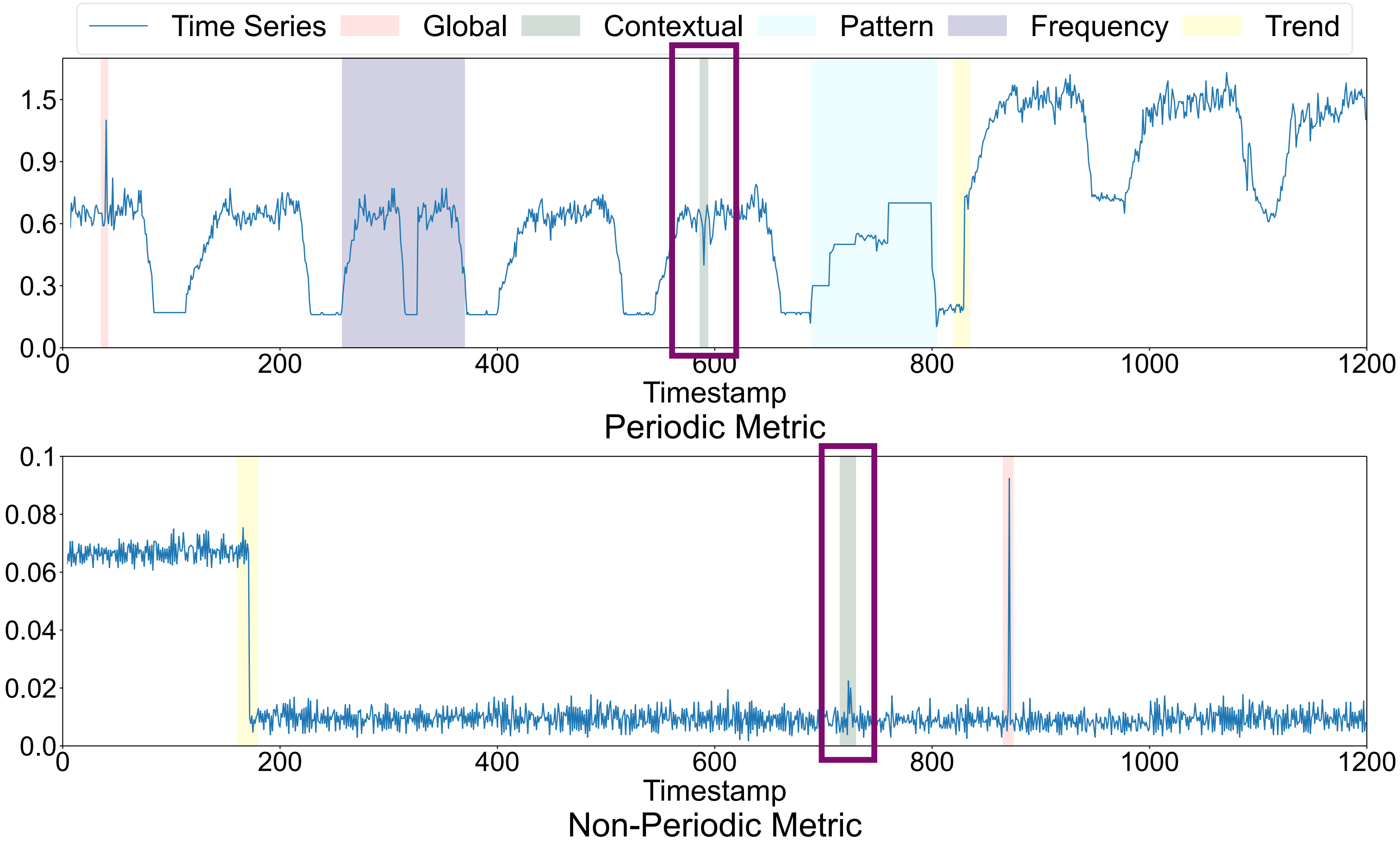
- Global anomalies
- Contextual anomalies
- Pattern anomalies
- Frequency anomalies
- Trend anomalies
- Others

Anomaly Types

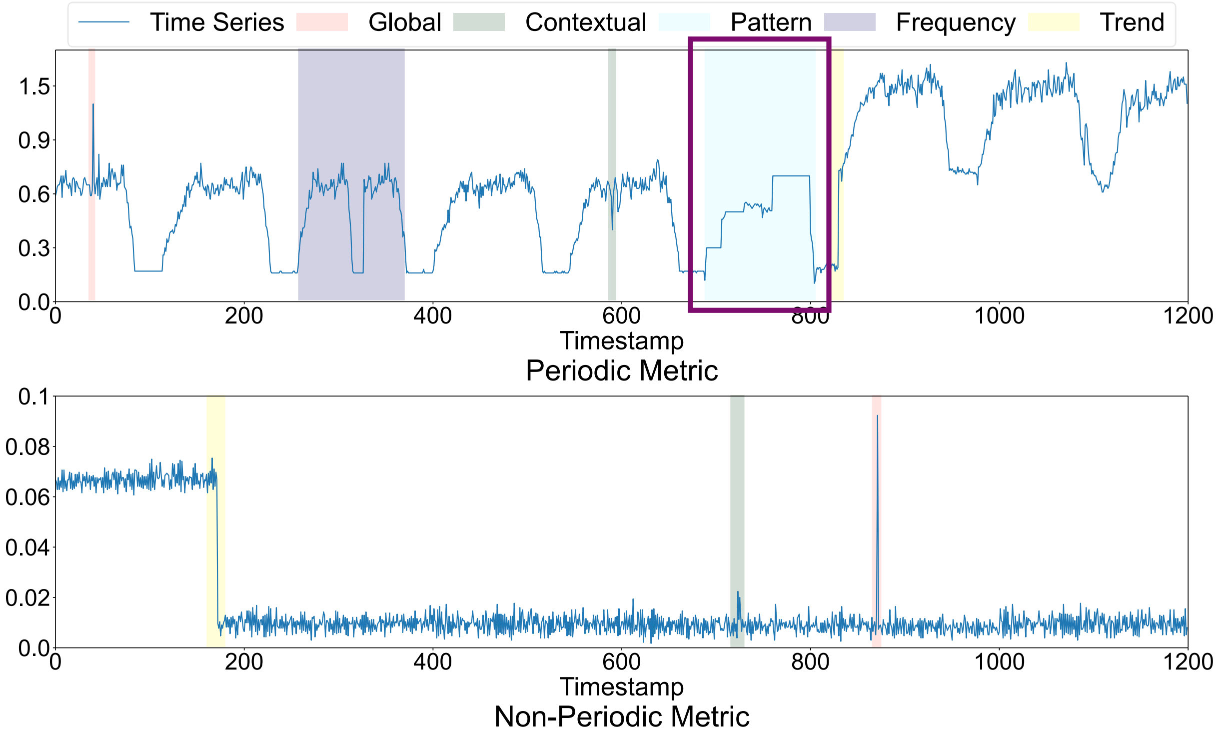
- Global anomalies
  - Short-lived
  - Extreme values



- Contextual anomalies
  - Short-lived
  - Deviate from the neighboring time points or differ from corresponding time points in other cycles.



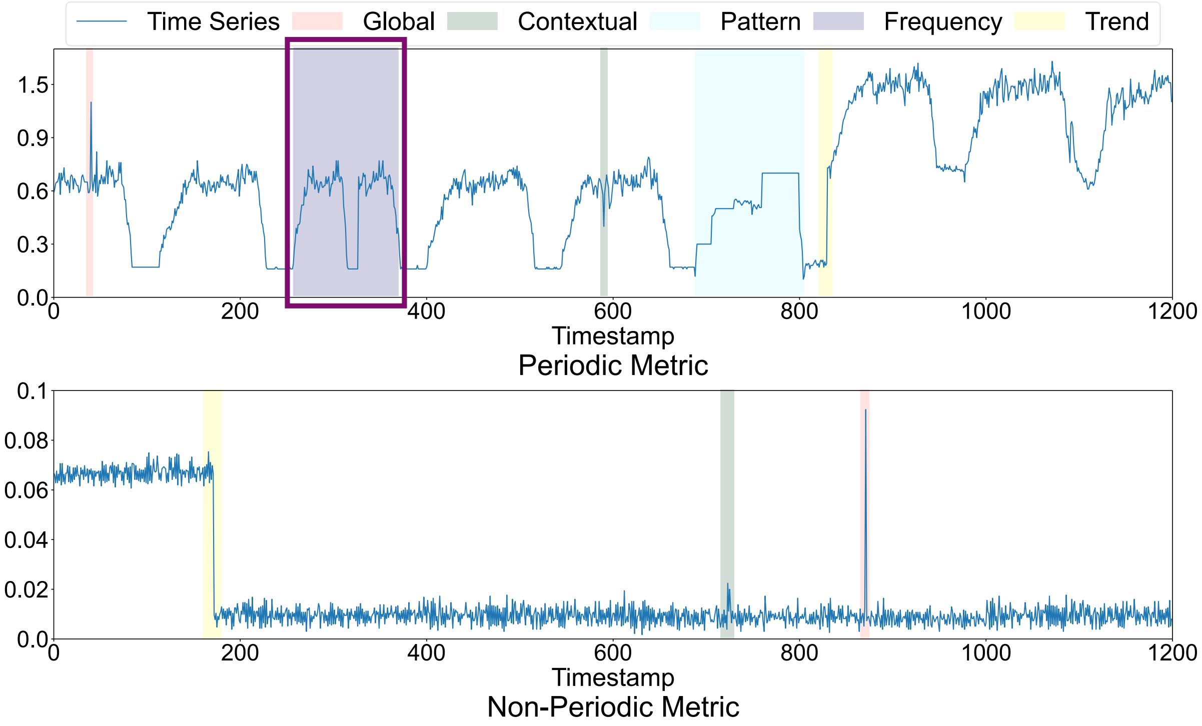
- Pattern anomalies
  - Longer segments equal to or greater than one period.
  - Occur in periodic metrics.
  - Segments with different basic patterns compared to normal patterns.



Anomaly Types

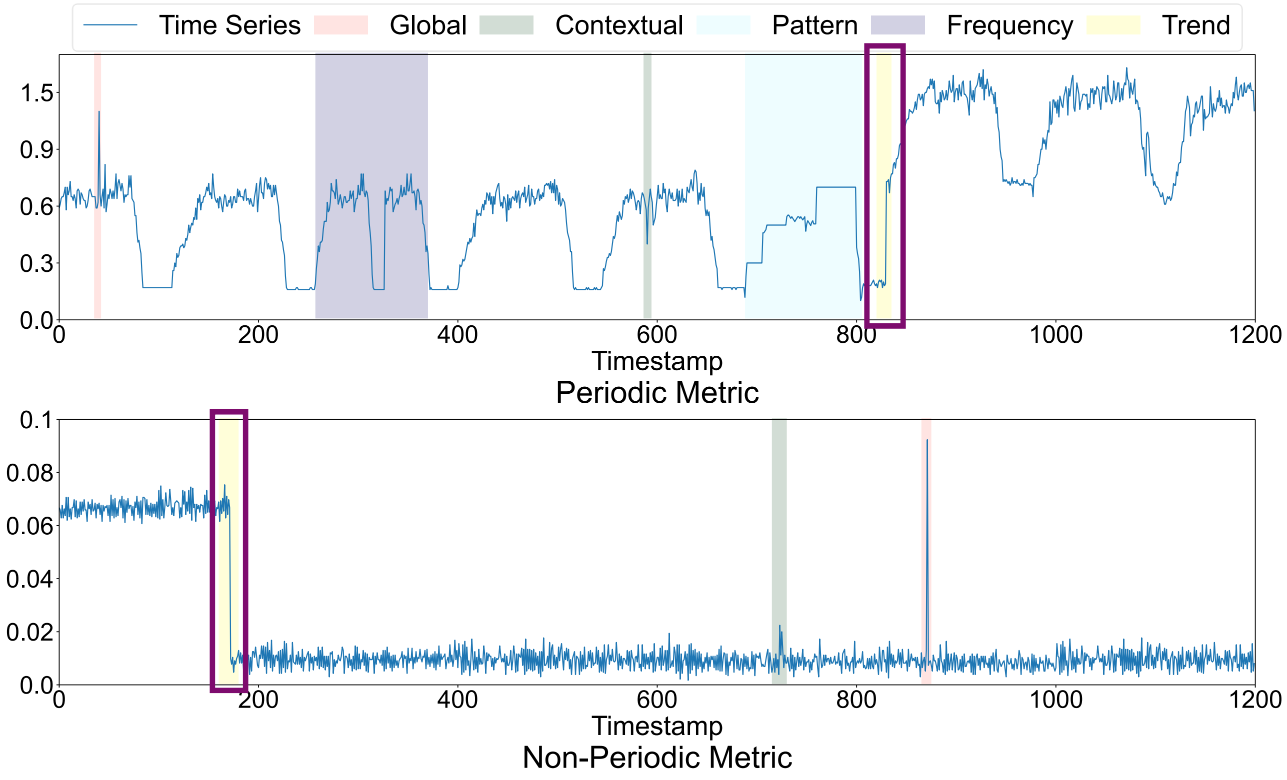
Nankai University

- Frequency anomalies
  - Longer segments equal to or greater than one period.
  - Occur in periodic metrics.
  - Display unusual frequency compared to the overall frequency.



Anomaly Types

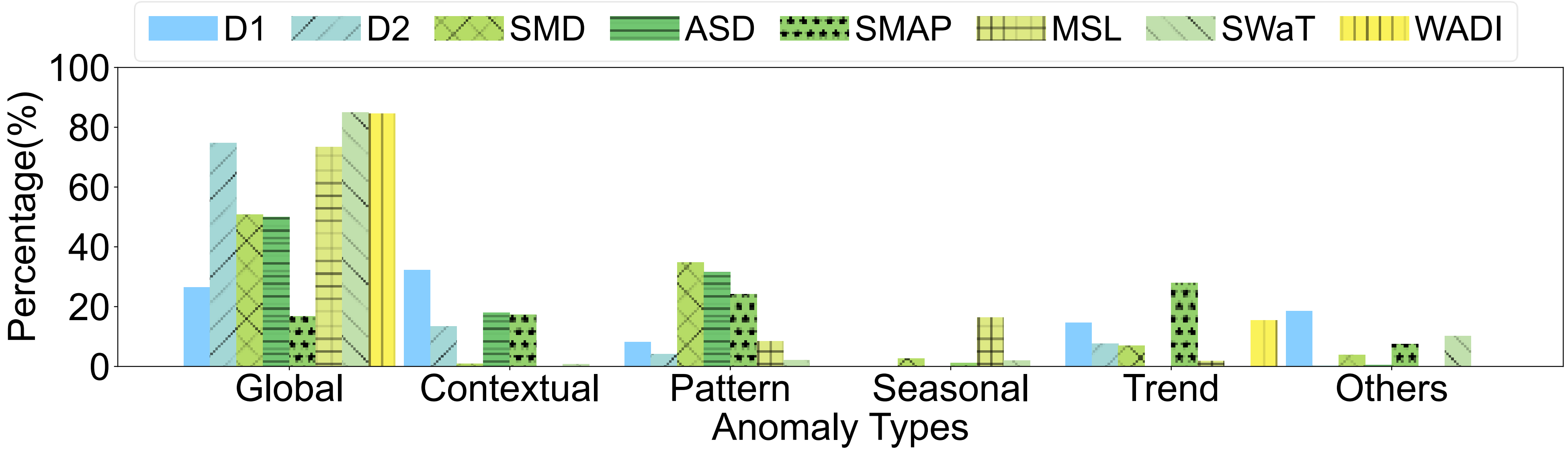
- Trend anomalies
  - Longer segments equal to or greater than one period.
  - Significantly deviate from the underlying trend of the time series.



Anomaly Types

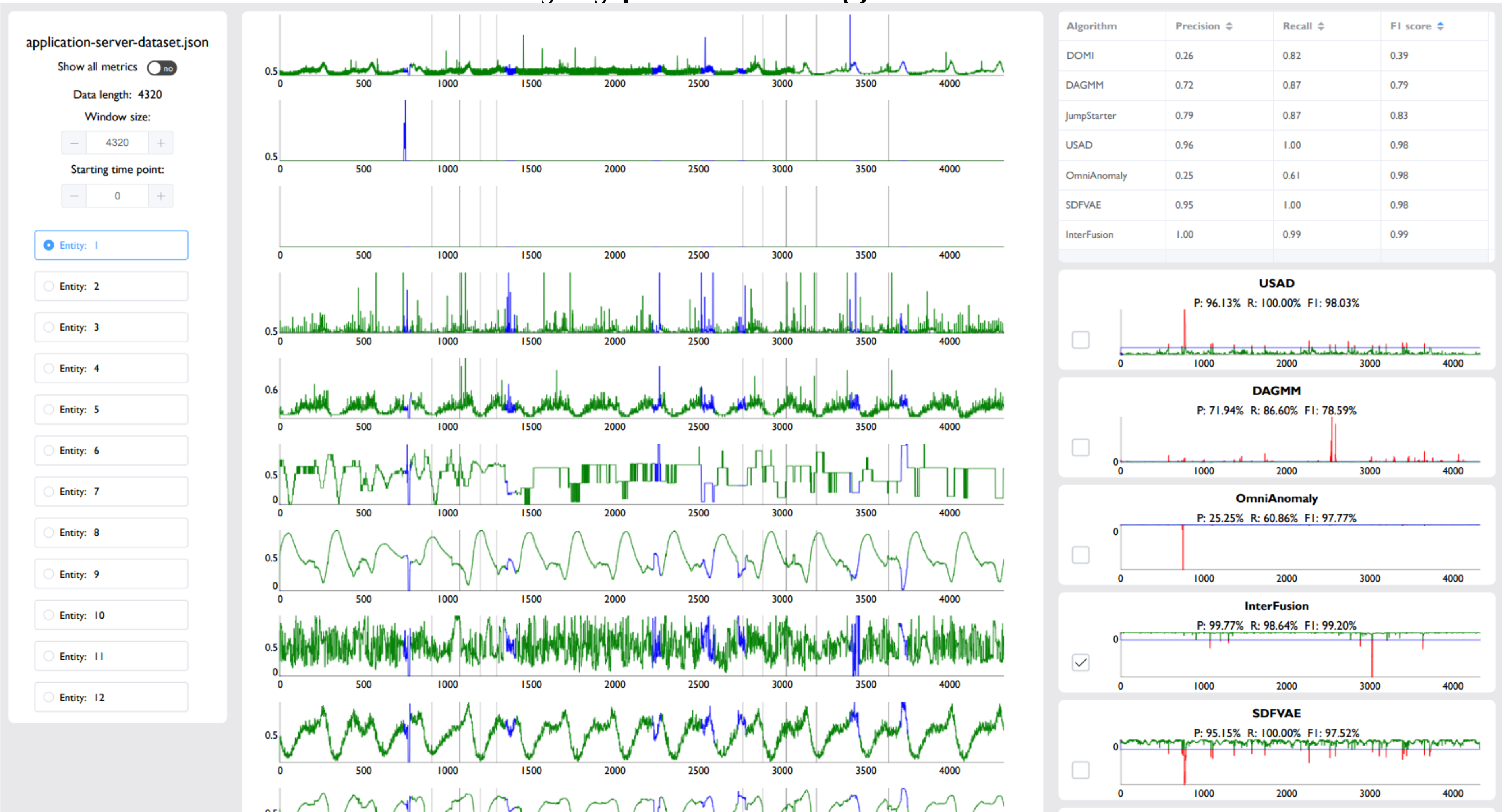
➢Global anomaly is the predominant type across all datasets.

➢Each dataset has noticeable variations in the proportion of anomaly types.



Labeling & GUI Visualization Tool

- Anomaly labeling based on data changes and incident tickets.
- Data characteristics and anomaly types labeling.



➢Eight popular unsupervised models.

Model	Advantages	Model Structures
DAGMM [ICLR 2018]	• Based on time point. • Preserves the low-dimensional features and reconstruction error for anomaly detection.	AE + GMM
USAD [ACM 2020]	• Leverages the advantages of AE and adversarial training • A straightforward model structure and a limited number of parameters.	AE + GAN
OmniAnomaly [ACM 2019]	• Models the explicit temporal dependence. • Employs a VAE to map input observations to stochastic variables.	RNN + VAE
DOMI [TON 2022]	• Simultaneously extracts both categorical variables and low-dimensional data features. • Works better with MTS data that exhibits multiple normal patterns.	1D-CNN + GMVAE
SDFVAE [WWW 2021]	• Be capable of explicitly learning the representations of time-invariant and time-varying characteristics.	CNN + RNN + VAE
InterFusion [ACM 2021]	• Employs a hierarchical VAE (HVAE) to learn different features independently. • Learns both low-dimensional inter-metric and temporal embeddings.	1D-CNN + RNN+ VAE
JumpStarter [ATC 2021]	• Clusters univariate time series in MTS. • Reconstructs MTS based on compressed sensing. • Effectively reduces initialization time.	Clustering + Compressed Sensing
GDN [AAAI 2021]	• Uses an attention-based GNN to learn the inter-metric dependence.	Attention + GNN

Content

Research Questions

Nankai University

- RQ4: How to select the most appropriate algorithms for different scenarios?

Recommendations

- General recommendations based on the computational resources, and effectiveness requirements.

Scenario	**Recommended Algorithms**
Effectiveness	SDFVAE, InterFusion, GDN
Efficiency for train	JumpStarter
Efficiency for test	USAD
Balance for both	SDFVAE, GDN

Recommendations

- Tailored recommendations based on data characteristics.

Scenario	Recommended Algorithms
Low smoothness	DAGMM, SDFVAE, InterFusion, GDN
High smoothness	USAD, OmniAnomaly, InterFusion
Low periodicity	DAGMM, SDFVAE, InterFusion, GDN
High periodicity	DAGMM, USAD, OmniAnomaly, DFVAE
Low metric correlation	DAGMM, InterFusion, GDN
High metric correlation	USAD, OmniAnomaly, SDFVAE

Recommendations

- Tailored recommendations based on anomaly types.

Scenario	**Recommended Algorithms**
Global	InterFusion
Contextual	GDN
Pattern	SDFVAE
Frequency	DOMI
Trend	SDFVAE

Evaluation

- The recommended SDFVAE, InterFusion, and GDN are among the top three algorithms.
- JumpStarter, and USAD, which can reduce model training overhead and detection overhead, achieved scores of 0.7855 and 0.8269, respectively.
- SDFVAE, and GDN, offer satisfactory performance while maintaining low training and detection overhead.

Model/Strategy	$\boldsymbol{F_1}$
DAGMM	0.8499
USAD	0.8269
OmniAnomely	0.738
DOMI	0.8372
SDFVAE	0.8657
InterFusion	0.8878
JumpStarter	0.7855
GDN	0.8823

Evaluation

Nankai University

- Our data characteristics-based solutions consistently outperform the best results achieved using a single algorithm, with performance gains of up to 1.3%.
- Our recommended solution effectively addresses this overhead concern by combining multiple algorithms with lower overhead.

Model/Strategy	F_1
DAGMM	0.8499
USAD	0.8269
OmniAnomely	0.738
DOMI	0.8372
SDFVAE	0.8657
InterFusion	0.8878
JumpStarter	0.7855
GDN	0.8823
Smoothness	0.9008
Periodicity	0.8906
Metric correlation	0.8889
Anomaly types	0.8624

Evaluation

Nankai University

- Our anomaly types-based solution achieves a score of 0.8624 on all MTS data.
- Operators can easily utilize our recommended models based on their practical application without validating all algorithms.

Model/Strategy	$\boldsymbol{F_1}$
DAGMM	0.8499
USAD	0.8269
OmniAnomely	0.738
DOMI	0.8372
SDFVAE	0.8657
InterFusion	0.8878
JumpStarter	0.7855
GDN	0.8823
Smoothness	0.9008
Periodicity	0.8906
Metric correlation	0.8889
Anomaly types	0.8624

Content

Conclusion

- Identify common data characteristics and types of anomalies that are crucial for effective MTS anomaly detection.
- Analyzes current practices and provides recommendations for selecting appropriate algorithms in specific scenarios.
- The evaluation results show that most of our suggestions can achieve better than any algorithm alone.
- Overall, we derive key findings and valuable insights that aim to guide and advance future research in anomaly detection.

Nankai University

Tsinghua University

Q&A

Thank you！

Name: Yongqian Sun

Email: sunyongqian@nankai.edu.cn

Date: 2023/10/16