Unravel Anomalies: An End-to-end Seasonal-Trend Decomposition Approach for Time Series Anomaly Detection

Zhenwei Zhang, Ruiqi Wang, Ran Ding, Yuantao Gu

Department of Electronic Engineering

Tsinghua University, Beijing, China

2024.04.19

TASK: Time series anomaly detection (TAD)

- **Time Series Anomaly:** The deviation of anomalous points from the distribution of normal samples.

- **Mostly used method:** Modeling the anomaly score

$$\mathcal{X} = [X_1, X_2, \dots, X_T], X_i \in \mathbb{R}^D$$

↓ Input: time series

$$\mathcal{S} = [s_1, s_2, \dots, s_T], s_i \in \mathbb{R}$$

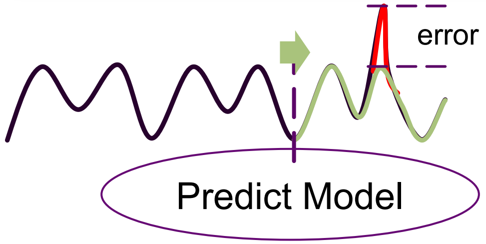
If $S_i \geq \epsilon$ ↓ Anomaly score series

$Y_i = 1$

$$\mathcal{Y} = [y_1, y_2, \dots, y_T], y_i \in \{0,1\}$$

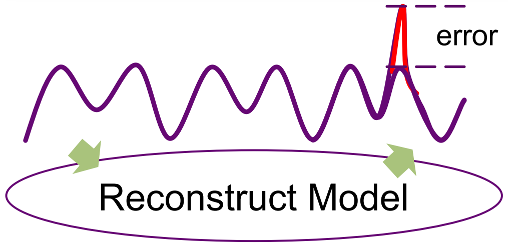
Output: anomalies

① Prediction-based



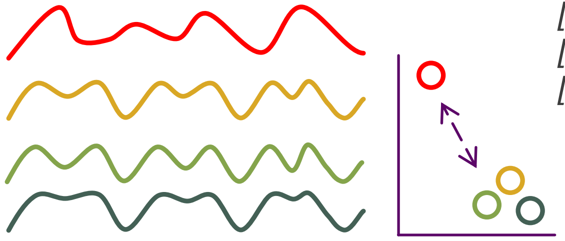
[ARIMA @1990]
[LSTM @KDD'18]
[HTM @Neurocomputing'18]

② Reconstruction-based



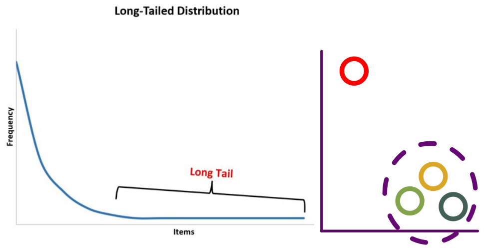
[DAGMM @ICLR'18]
[LSTM-VAE @IEEE RAL'18]
[OmniAnomaly @KDD'19]

③ Distance-based



[DBStream @TKDE'16]
[Series2Graph @PVLDB'20]
[SAND @PVLDB'21]

④ Distribution-based



[COPOD @ICDM'20]
[DSPOT @KDD'17]
[DWT-MLEAD @ITISE'17]

Motivation

Challenges in TAD task:

- **Variability and Complexity**: Time-series data can exhibit complex patterns such as seasonality, trends, and noise, making anomaly detection challenging.
- **Data Quality**: Issues with data quality, such as missing values or noise, can complicate the detection process.

Need for a Method that Can Handle Diverse Anomalies

- **Robustness**: robust against noise, providing reliable results across diverse data origins
- **Adaptability**: adapting to diverse nature of anomalies and the complex patterns in time-series data [1,2]

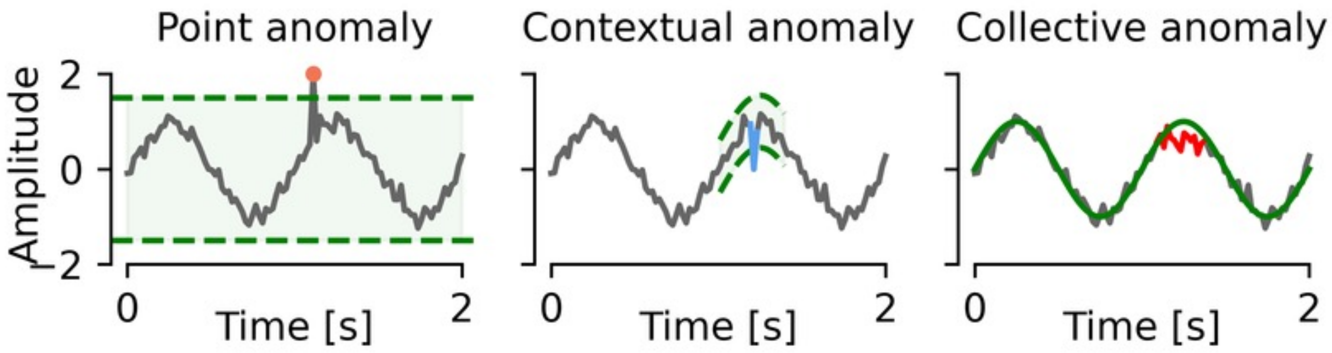


Figure from [3]

[1] Kwei-Herng Lai, Daochen Zha, Junjie Xu, Yue Zhao, Guanchu Wang, and Xia Hu, "Revisiting time series outlier detection: Definitions and benchmarks," in Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track, 2021.
[2] Jiehui Xu, Haixu Wu, Jianmin Wang, and Mingsheng Long, "Anomaly transformer: Time series anomaly detection with association discrepancy," in International Conference on Learning Representations, 2021.
[3] Yan, Peng, et al. "A comprehensive survey of deep transfer learning for anomaly detection in industrial time series: Methods, applications, and directions." *IEEE Access* (2024).

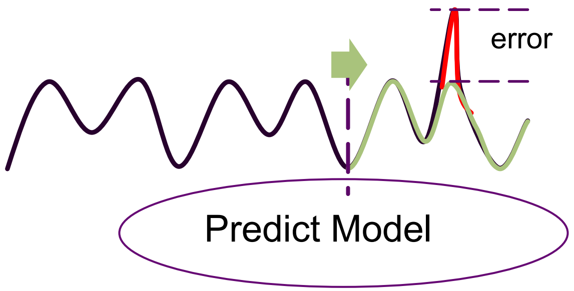
Related works

Q: Is this statement Real?

In recent years, deep learning models have surpassed classical techniques in TAD tasks.

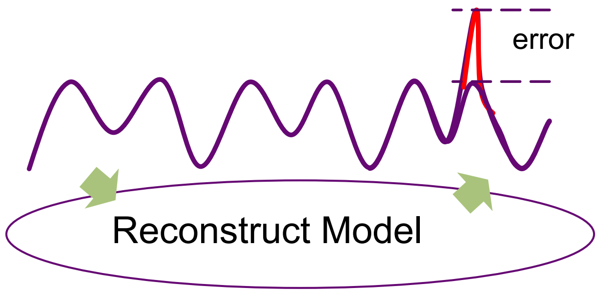
These models generally fall into two categories:

① Prediction-based



[ARIMA @1990]
[LSTM /SMAP @KDD'18]
[HTM @Neurocomputing'18]

② Reconstruction-based



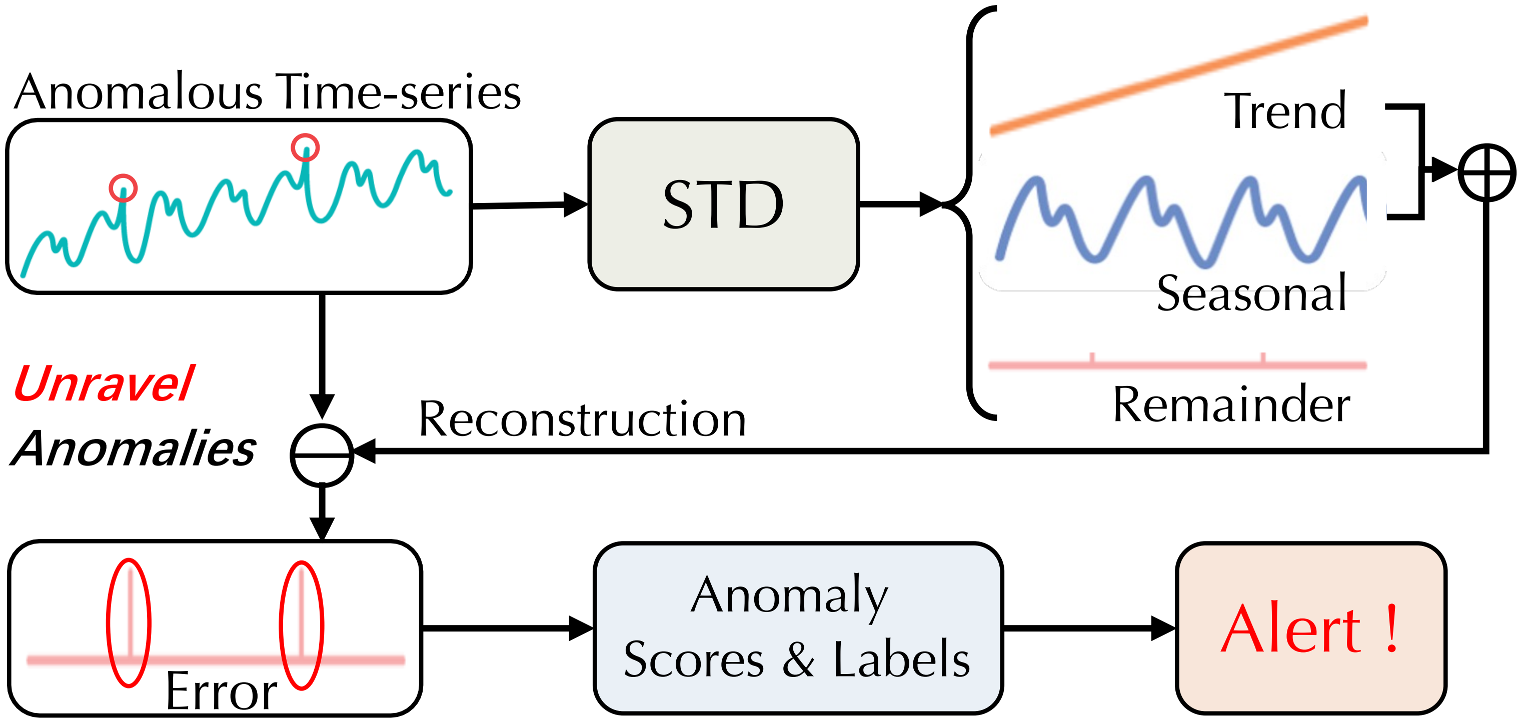
[DAGMM @ICLR'18]
[LSTM-VAE @IEEE RAL'18]
[OmniAnomaly @KDD'19]
[InterFusion @KDD'21]

Despite their overall accuracy, many of these models **fail to account for the complex compositional nature of patterns in time-series data** or **distinguish between different types of anomalies**.

The proposed solution: TADNet

Leveraging **decomposition**, our approach uniquely **break down** these complex composite patterns.

Furthermore, different types of anomalies can be systematically associated with their respective components: **seasonal anomalies** with the **seasonal component**, **trend anomalies** with the **trend component**, and **point anomalies** with the **remainder component**.



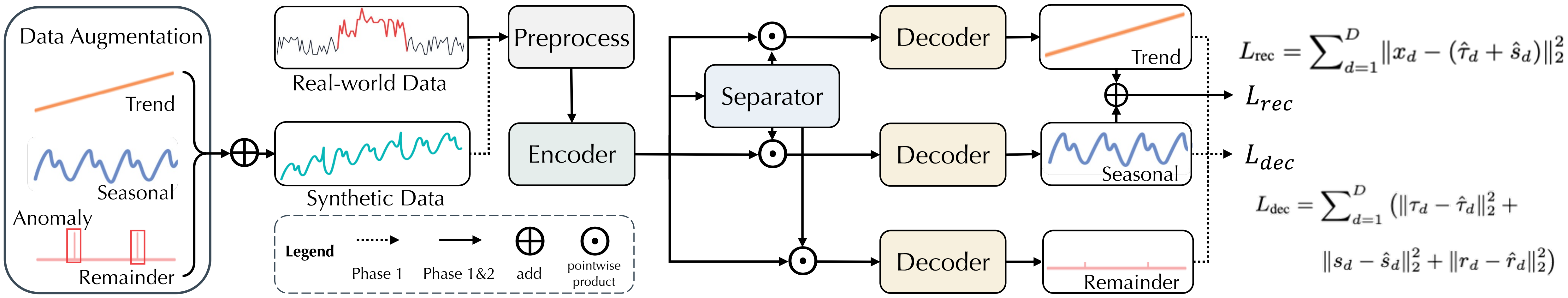
Schematic of the STD and TAD workflow

Two-Phase Training Strategy

Although existing research has **incorporated time-series decomposition into TAD task**[1, 2], these approaches **do not follow an end-to-end training manner**. Specifically, they either depend on *predefined decomposition algorithms with complex parameter tuning*[2], or *employ decomposition only for data preprocessing*[1].

To overcome the lack of supervised signals for end-to-end training, we introduce **a novel two-step training approach**:

1. **Pre-training** on Synthetic Datasets
2. **Fine-tuning** for Precise Anomaly Detection

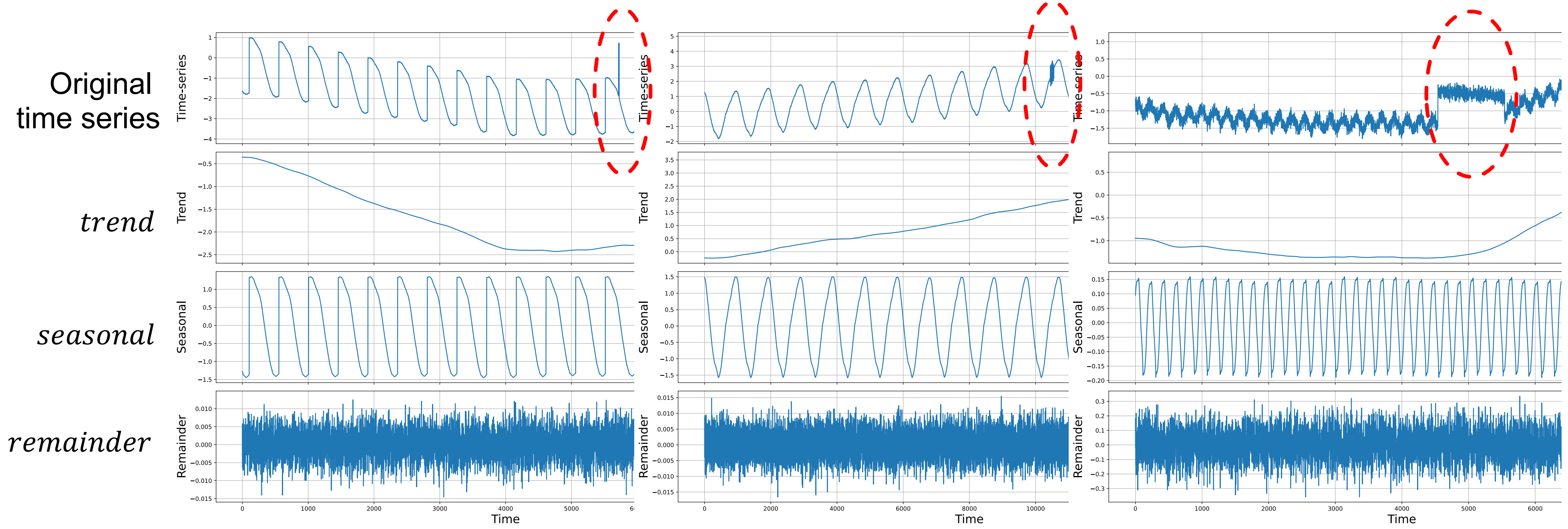


Overall workflow of TADNet.

[1] Shuxin Qin, Jing Zhu, Dan Wang, Liang Ou, Hongxin Gui, and Gaofeng Tao, “Decomposed transformer with frequency attention for multivariate time series anomaly detection,” in 2022 IEEE International Conference on Big Data (Big Data). IEEE, 2022, pp. 1090–1098.
[2] Jingkun Gao, Xiaomin Song, Qingsong Wen, Pichao Wang, Liang Sun, and Huan Xu, “Robusttad: Robust time series anomaly detection via decomposition and convolutional neural networks,” arXiv preprint arXiv:2002.09545, 2020.

Synthetic Datasets

	trend	*seasonal*	*remainder*
Generate process	Deterministic linear trend and random $ARIMA(0,2,0)$ model $\tau_t^{(d)} = \beta_0 + \beta_1 \cdot t$ $\tau_t^{(s)} = \sum_{n=1}^{t} nX_n$ $\quad \Delta^2 \tau_t^{(s)} = X_t$	Deterministic periodic terms $s_t^{(s)} = \tau_{mod(t+\phi, T_0)}^{(s)}$	White noise process $r_t \sim \mathcal{N}(0, \sigma^2)$
Details	β_0 and β_1 are tunable parameters, white noise X_t follows a normal distribution	Period T_0 and phase ϕ with gradual adjustments over time	Adjust the noise variance σ^2 to fit different data distributions



Synthetic Datasets Visualization

Clearer classification

Commonly used but somehow confusing

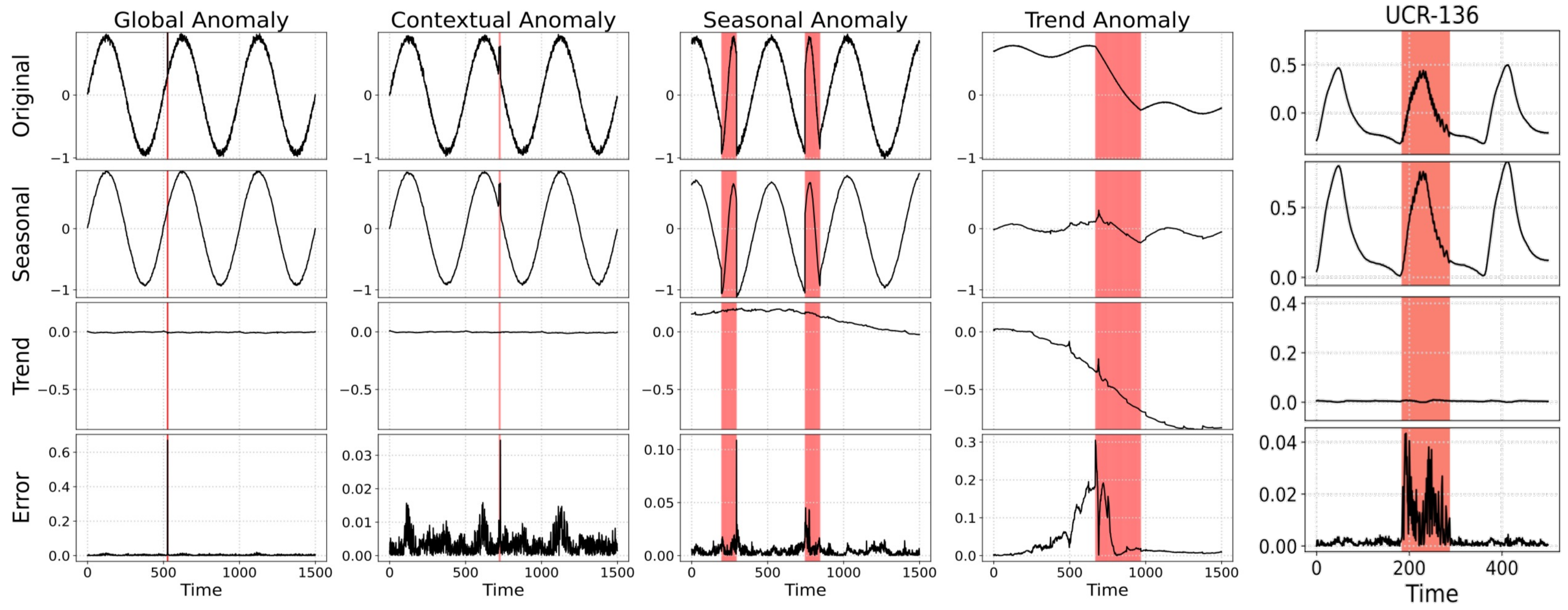
[CompEval @PVLDB'22]

Combining time series decomposition and anomaly detection for finer-grained classification of anomalies

Time-series Anomalies → { point, subsequence, correlation } ?

✅ *[NIPS-TS @NeurIPS'21]* *[TFAD @CIKM'22]*

- Time-series Anomalies
  - Point
    - Global
    - Contextual
  - Pattern
    - Seasonal
    - Trend
    - Shapelet



Experimental Validation

ICASSP 2024 KOREA

Table 2. Details of the dataset.

	Dataset	#Entities	#Dim	#Train	#Test(labeled)	Anomaly%
[KDD' 21]	UCR	4	1	1,200-3,000	4,500-6,301	1.9
[KDD' 19]	SMD	28	38	23.6K-28.7K	23.6K-28.7K	4.2
[CySWater' 16]	WADI	1	127	789,371	172,801	5.9
[CySWater' 16]	SWaT	1	51	496,800	449,919	12.1
[KDD'21]	PSM	1	25	132,481	87,841	27.8

without finetuning on synthetic datasets

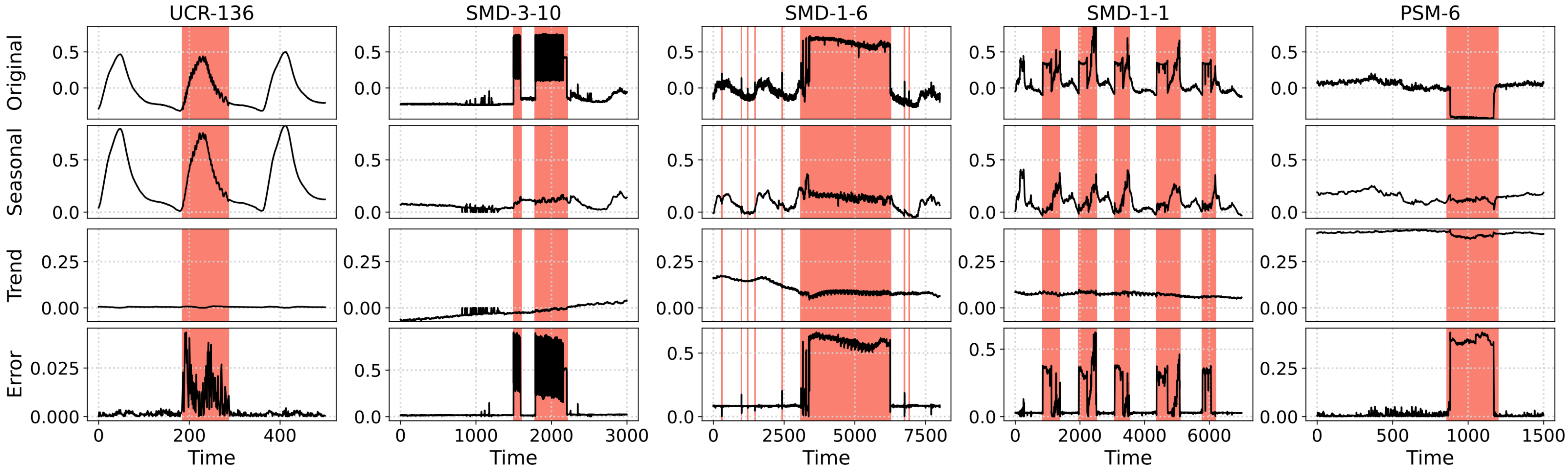
Ablation	UCR	SMD	SWaT	PSM	WADI
TADNet	98.74	**93.35**	**90.21**	**98.66**	88.15
w/o Sep	32.68	66.24	76.89	83.28	47.66
w/o Decomp	48.69	84.12	88.41	95.57	65.72
w/o Augment	40.12	74.17	83.26	98.01	62.15
Iterative	**99.12**	92.14	86.55	96.58	**92.06**

Table 1. Quantitative results for TADNet across five real-world datasets use metrics *P*, *R*, and *F1* for precision, recall, and F1-score (%). Higher values indicate better performance. Best and second-best results are in bold and underlined, respectively. Dataset are followed by brackets, where *u* indicates univariate and *m* multivariate.

	Dataset	**UCR** (*u*)			**SMD** (*m*)			**SWaT** (*m*)			**PSM** (*m*)			**WADI** (*m*)		
	Metric	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
[NeurComput.'01]	OCSVM	41.14	94.00	57.23	44.34	76.72	56.19	45.39	49.22	47.23	62.75	80.89	70.67	61.89	62.31	62.10
[IJCAI'19]	BeatGAN	45.20	88.42	59.82	72.90	84.09	78.10	64.01	87.46	73.92	90.30	93.84	92.04	65.13	38.32	48.25
[SIGKDD'19]	OmniAnomaly	64.21	86.93	73.86	83.34	94.49	88.57	86.33	76.94	81.36	91.61	71.36	80.23	31.58	65.41	42.60
[SIGKDD'21]	InterFusion	60.74	95.20	74.16	87.02	85.43	86.22	80.59	85.58	83.01	83.61	83.45	83.52	80.26	30.38	44.08
[ICLR'21]	AnomalyTran	72.80	99.60	84.12	89.40	95.45	<u>92.33</u>	91.55	96.73	**94.07**	96.91	98.90	<u>97.89</u>	80.30	79.23	79.76
[PVLDB'22]	TranAD	94.07	100.00	<u>96.94</u>	88.03	89.42	88.72	97.60	69.97	81.51	96.44	87.37	91.68	35.29	82.96	49.51
[IEEEBigData'22]	DecompTran	71.58	96.83	82.31	89.32	93.94	91.57	95.17	80.30	87.10	97.65	87.21	92.14	79.40	81.01	<u>80.20</u>
	TADNet(Ours)	97.51	100.00	**98.74**	94.81	91.93	**93.35**	92.15	88.35	<u>90.21</u>	98.12	99.21	**98.66**	94.03	82.96	**88.15**

Results Visualization

Visualization of decomposition and detection results in UCR and SMD.



The first row shows the raw time series with anomalies, the second and third rows display the seasonal and trend components, respectively, and the final row depicts the reconstruction error. Anomalies are marked with a red background.

Back to the question

Is this Real?

In recent years, deep learning models have surpassed classical techniques in TAD task.

Flaws on Datasets

Simple methods

Series difference → Peak Detect

Detect peak density

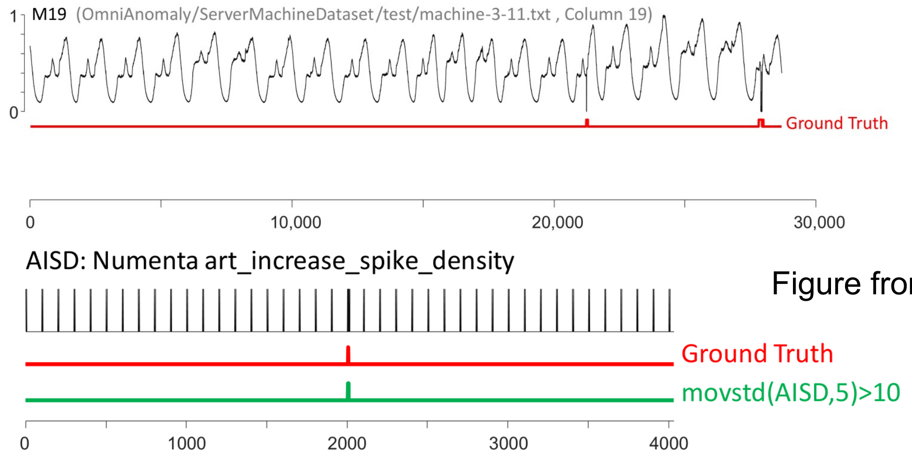


Figure from [1]

[1] Wu, Renjie, and Eamonn J. Keogh. "Current time series anomaly detection benchmarks are flawed and are creating the illusion of progress." *IEEE transactions on knowledge and data engineering* 35.3 (2021): 2421-2429.

Flaws on Evaluation Metrics

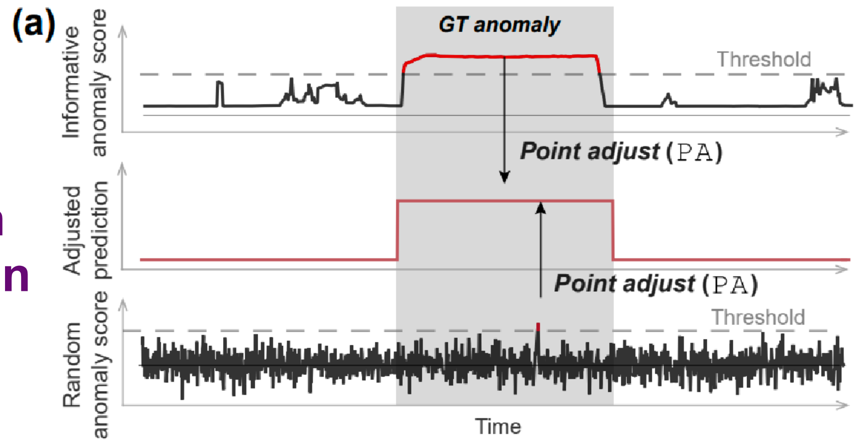


Figure from [2]

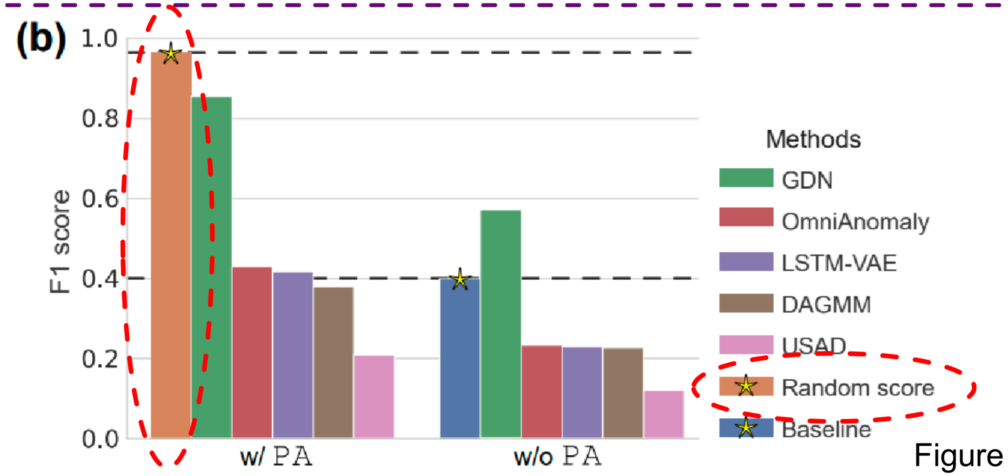


Figure from [2]

[2] Kim, Siwon, et al. "Towards a rigorous evaluation of time-series anomaly detection." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 7. 2022.
[3] Emmanouil Sylligardos, Paul Boniol, et al. 2023. Choose Wisely: An Extensive Evaluation of Model Selection for Anomaly Detection in Time Series. Proc. VLDB Endow. (July 2023).

清華大學 Tsinghua University

ICASSP 2024 KOREA

Thank you! Q & A

Zhenwei Zhang

2024.04.19

zzw20@mails.tsinghua.edu.cn

Github Repo