Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

RICE UNIVERSITY
School of Engineering
Department of Computer Science

PennState
College of Information Sciences and Technology

Revisiting Time Series Outlier Detection: Definition and Benchmark

[*]Kwei-Herng Lai, [*]Daochen Zha, [$]Junjie Xu,
[#]Yue Zhao, [*]Guanchu Wang, [*]Xia Hu

[*]Rice University
[#]Carnegie Mellon University
[$]Pennsylvania State University

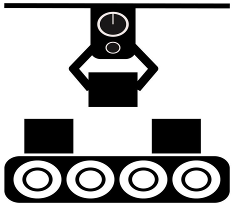
NEURAL INFORMATION PROCESSING SYSTEMS

Detecting Outliers From Time Series Data

- **Outlier Detection**
  - Find the weirdest/abnormal part of the data
  - Unsupervised/Semi-supervised training

- **Time Series Applications**



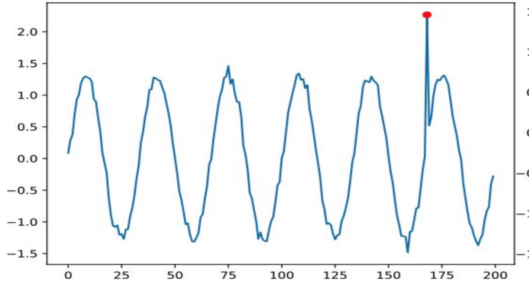


- **Main Challenge: Anomalies are Rare**
  - Only attribute of the historical anomalies are accessible.
  - Context of future anomalies are unknown.

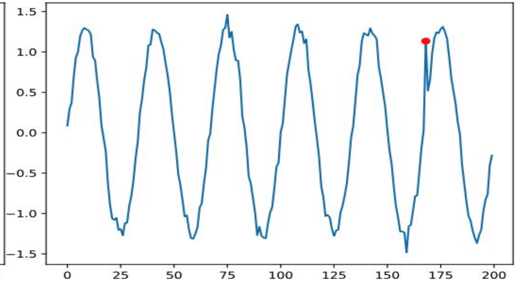
Daochen Zha (daochen.zha@rice.edu)

Outliers Definitions

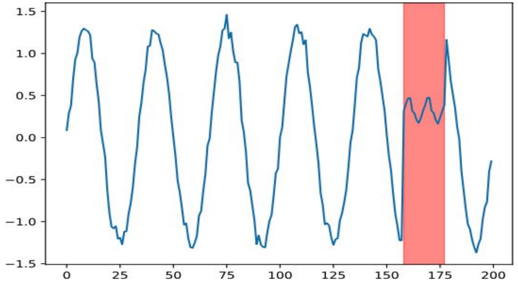
- **Outlier Definition:**
  - **Point (left figure):**
    - Data point that is anomalous comparing to the rest of the data.
    - Extreme values lead to serious consequences —> Easy to detect
  - **Contextual (middle figure):**
    - Data point that is anomalous under certain context.
    - Context is often defined unclearly —> Hard to detect.
  - **Collective (right figure):**
    - Collection of data points that is anomalous with respect to entire dataset.
    - Individual points of the outlier may not be outlier and the scenario of collection is unknown —> Hard to detect.



Point Outlier



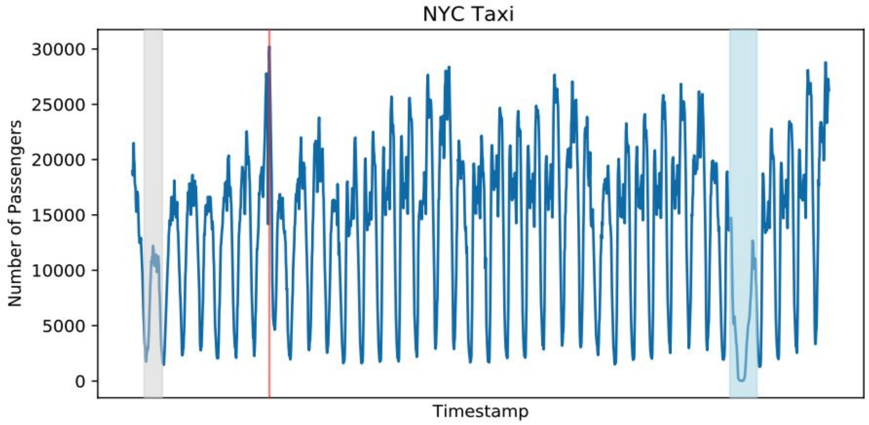
Contextual Outlier

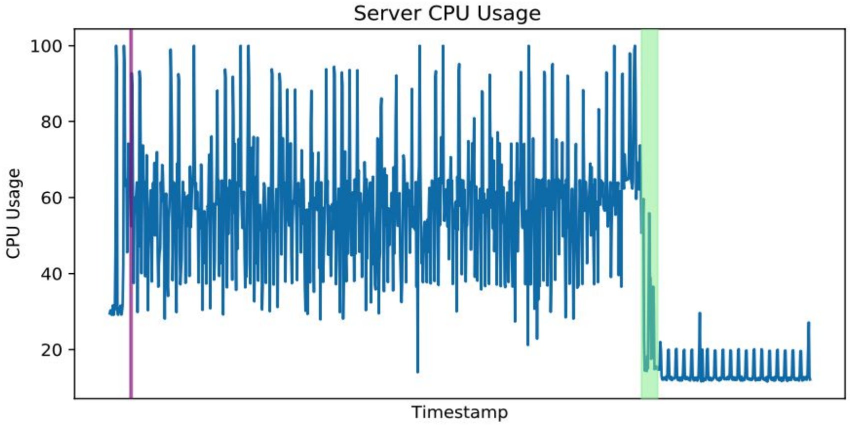


Collective Outlier

Time Series Outliers are Complex

- **Misleading labeled benchmark datasets create illusion for the progress of time series outlier detection problem [1].**





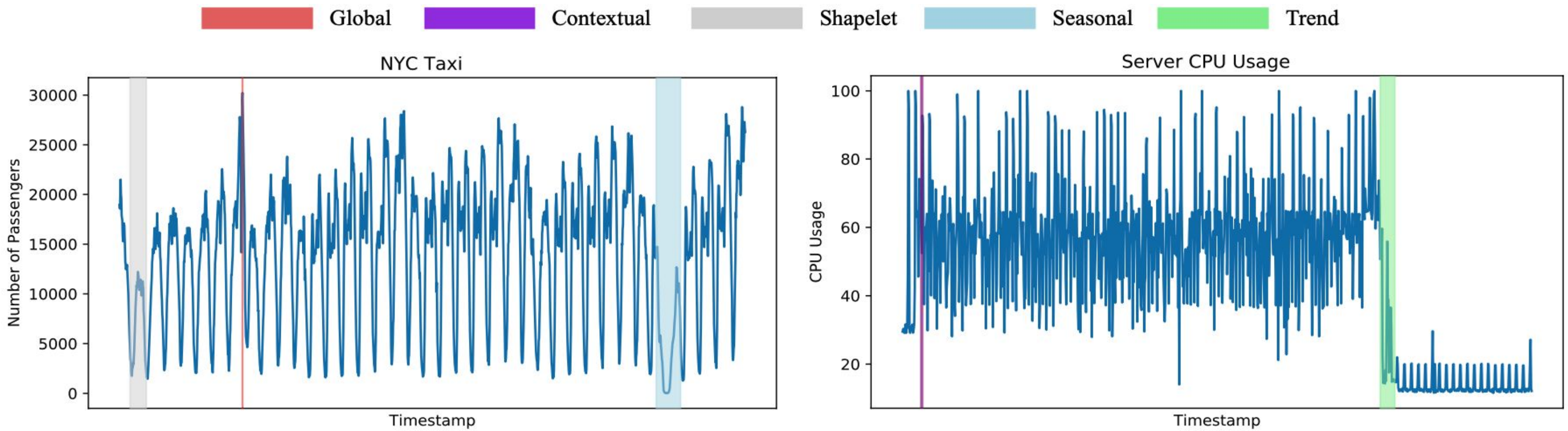
Synthetic Dataset

- **Using synthetic datasets for evaluation becomes a common practice.**
  - **In this way, we can easily isolate different types of outliers.**
- **Existing outlier definitions are confusing for synthesizing data**

  Contextual —> Unknown context

  Collective —> No specific scenario

  Can we refine the outlier definitions for better benchmarking algorithms?



[1] Renjie Wu, Eamon Keogh, Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress, IEEE Transactions on Knowledge and Data Engineering, 2021

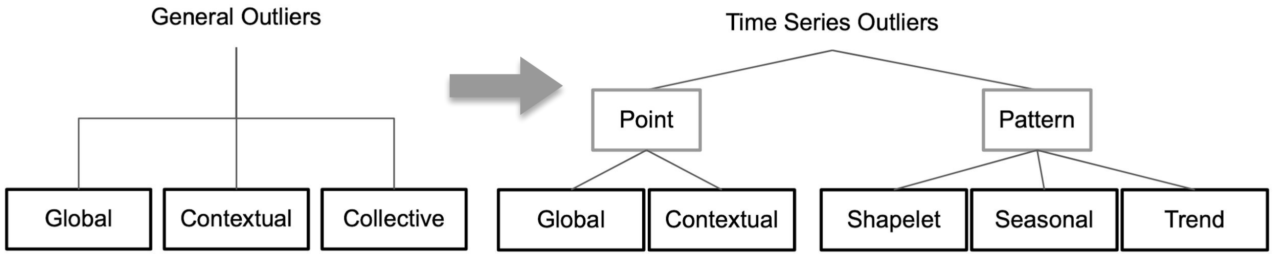
Our Proposal: Behavior-driven Outlier Taxonomy

- **Time Series Behaviors:**

$$X = \underbrace{(x_1, x_2, \cdots, x_t)}_{\text{Point-wise Behavior}} = \underbrace{\boxed{\rho(2\pi\omega T)} + \tau(T)}_{\text{Pattern-wise Behavior}}$$

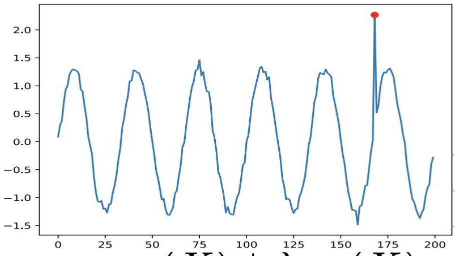
$$: \sum_n [Asin(2\pi\omega_n T) + Bcos(2\pi\omega_n T)]$$

  - **Point-wise Outliers:** $|x_t - \hat{x}_t| > \delta$
    - Global Outlier: $\delta = \lambda \cdot \sigma(X)$
    - Contextual Outlier: $\delta = \lambda \cdot \sigma(X_{t-k,t+k})$
  - **Pattern-wise Outliers:** $s(X_{i,j}, \hat{X}_{i,j}) > \delta$
    - Shapelet Outlier: $s(\rho(.), \hat{\rho}(.)) > \delta$
    - Seasonal Outlier: $s(\omega, \hat{\omega}) > \delta$
    - Trend Outlier: $s(\tau(.), \hat{\tau}(.)) > \delta$

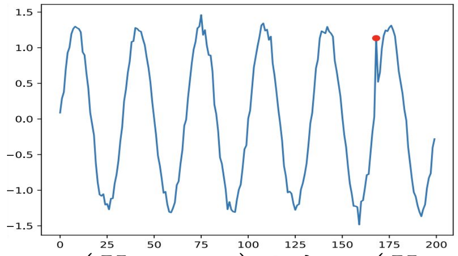


Synthetic Criteria for Time Series Outliers

- **Synthetic Criteria**
  - **Point-wise Outlier: Defining range of context.**

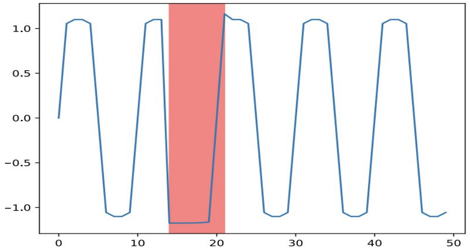


$$x_t = \mu(X) \pm \lambda \cdot \sigma(X)$$

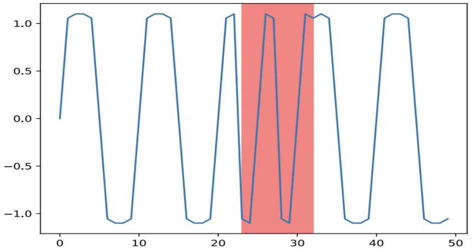


$$x_t = \mu(X_{t-k,t+k}) \pm \lambda \cdot \sigma(X_{t,t+k})$$

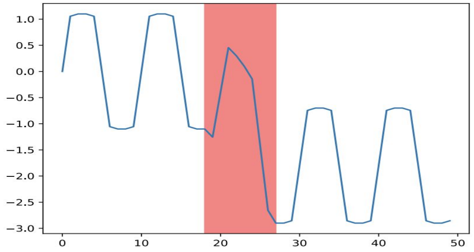
  - **Pattern-wise Outliers: Selecting** ρ, ω, τ



$$\rho = sgn(sin2\pi\omega t)$$



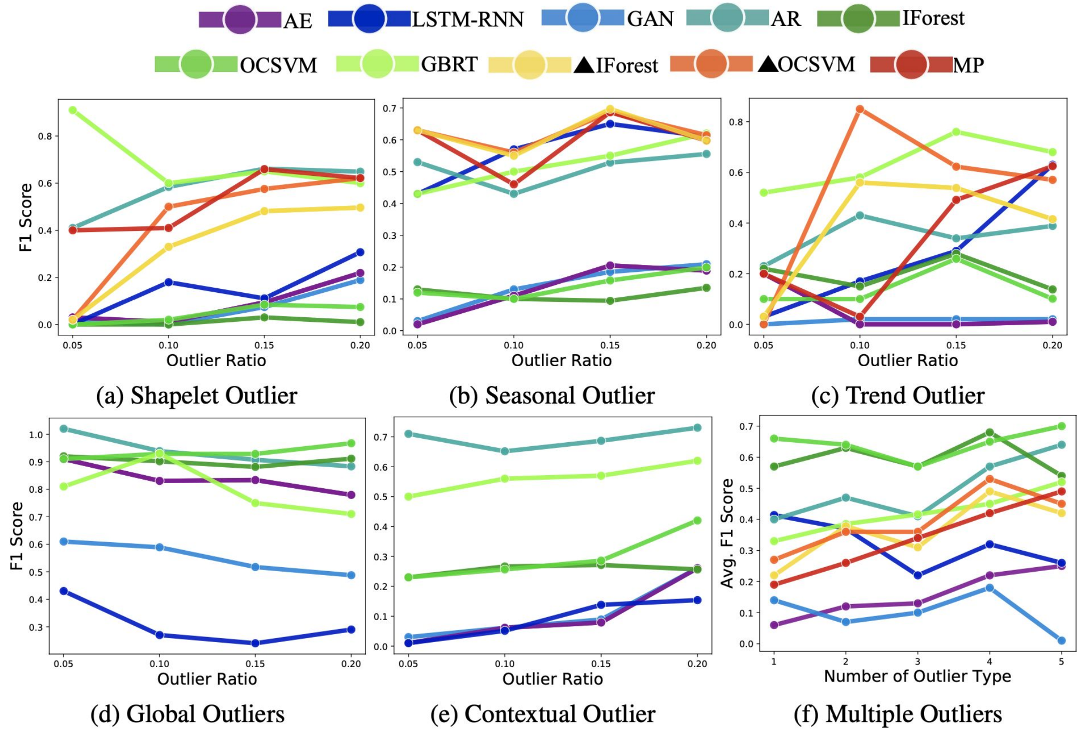
$$\omega = 2\hat{\omega}$$



$$\tau(T_{i,j}) = \{-1, \cdots, -(j-i+1)\}$$

Benchmarking Existing Algorithms

- **Benchmark on Synthetic Datasets**



- **Key Observations:**
  - Classical Algorithms > Deep Algorithms.
  - Difficulty in Point-wise:
    Contextual >> Global.
  - Difficulty in Pattern-wise:
    Seasonal > Trend > Shapelet.

Benchmarking Existing Algorithms

- **Benchmark on Real-world Datasets**

	Dataset (Best)	Credit Card			CICIDS			GECCO			SWAN-SF		
	Metrics	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Prediction-based	AR	0.113	0.652	0.192	0.016	0.310	0.030	0.392	0.314	0.349	0.421	0.354	0.385
	GBRT	**0.113**	**0.657**	**0.193**	0.018	0.351	0.034	0.175	0.140	0.156	0.447	0.375	0.408
	LSTM-RNN	0.004	0.110	0.007	**0.024**	**0.383**	**0.046**	0.343	0.275	0.305	0.527	0.221	0.312
Majority Modeling	IForest	0.098	0.569	0.168	0.010	0.040	0.016	**0.439**	0.353	**0.391**	**0.569**	**0.598**	**0.583**
	OCSVM	0.107	0.620	0.183	0.004	0.046	0.007	0.185	**0.743**	0.296	0.474	0.498	0.485
	AutoEncoder	0.103	0.598	0.176	0.011	0.042	0.017	0.424	0.340	0.377	0.497	0.522	0.509
Discords Analysis	▲IForest	0.039	0.226	0.066	0.011	0.168	0.020	0.392	0.315	0.390	0.406	0.425	0.416
	▲OCSVM	0.002	0.305	0.004	0.000	0.000	0.000	0.021	0.341	0.040	0.193	0.001	0.001
	MatrixProfile	0.006	0.514	0.012	0.007	0.080	0.013	0.046	0.185	0.074	0.167	0.175	0.171

- **Key Observations:**
  - Classical Algorithms > Deep Algorithms.
  - Subsequence clustering (▲) is not robust under concentration assumption.

Take Home Messages

- **In this work, we identify that**
  - Existing taxonomy of outlier definition is not applicable for time series data.
  - Unclear outlier definition may lead to confusing labeled data.
- **To tackle the problem**
  - A new taxonomy by observing the data from **point** and **pattern** of views.
  - Categorize time series outliers into **global**, **contextual**, **trend**, **seasonal** and **shapelet** outliers.
  - Formal definition of individual types of outliers to create **synthetic criterias** for benchmarking existing algorithms and future data labeling.
- **From the benchmark, we further find out that**
  - Classical algorithms generally outperform Deep Learning algorithms.
  - The difficulty of individual types of outliers can be ranked as
    - Point-wise: Contextual >> Global
    - Pattern-wise: Seasonal > Trend > Shapelet.

Carnegie Mellon University
HeinzCollege
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

RICE UNIVERSITY
School of Engineering
Department of Computer Science

PennState
College of Information Sciences and Technology

Thank You!

Paper

GitHub

Contact

Kwei-Herng Lai (khlai@rice.edu)
Daochen Zha (daochen.zha@rice.edu)
Junjie Xu (jmx5097@psu.edu)
Yue Zhao (zhaoy@cmu.edu)
Guanchu Wang (hegsns@rice.edu)
Xia Hu (xiahu@rice.edu)

NEURAL INFORMATION PROCESSING SYSTEMS