Tsinghua University

Robust Anomaly Detection for Multivariate Time Series through Stochastic Recurrent Neural Network

Ya Su, Youjian Zhao, Chenhao Niu,
Rong Liu, Wei Sun, Dan Pei

SIGKDD 2019

Outline

Background

Algorithm

Evaluation

Conclusion

Outline

Background

Algorithm

Evaluation

Conclusion

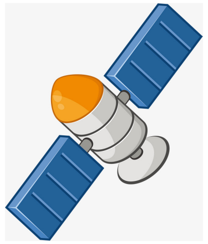
Anomaly Detection

- Graph [SIGKDD 2018, AI Magazine 2014]
- Log Messages [SIGKDD 2016, SIGKDD 2017]
- Time Series [SIGKDD 2015, SIGKDD 2017, SIGKDD 2018]
  - Univariate Time Series
  - Mutivariate Time Series

Entities with monitored multivariate time series

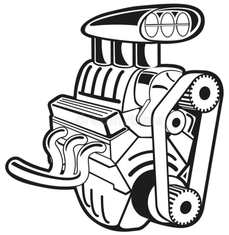
Entities

Server Machine



SpaceCraft

Robot-assisted System



Engine

. . .

Multi-metrics

Server Machine	SpaceCraft	Robot-assisted System	Engine
CPU Load Network Usage Memory Usage ...	Radiation Temperature Power ...	Kinematic Visual Haptic ...	Accelerator Torque Temperature ...

. . .

Entities with monitored multivariate time series

Entities

Server Machine

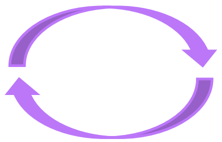
More intuitive

More effective

More efficient

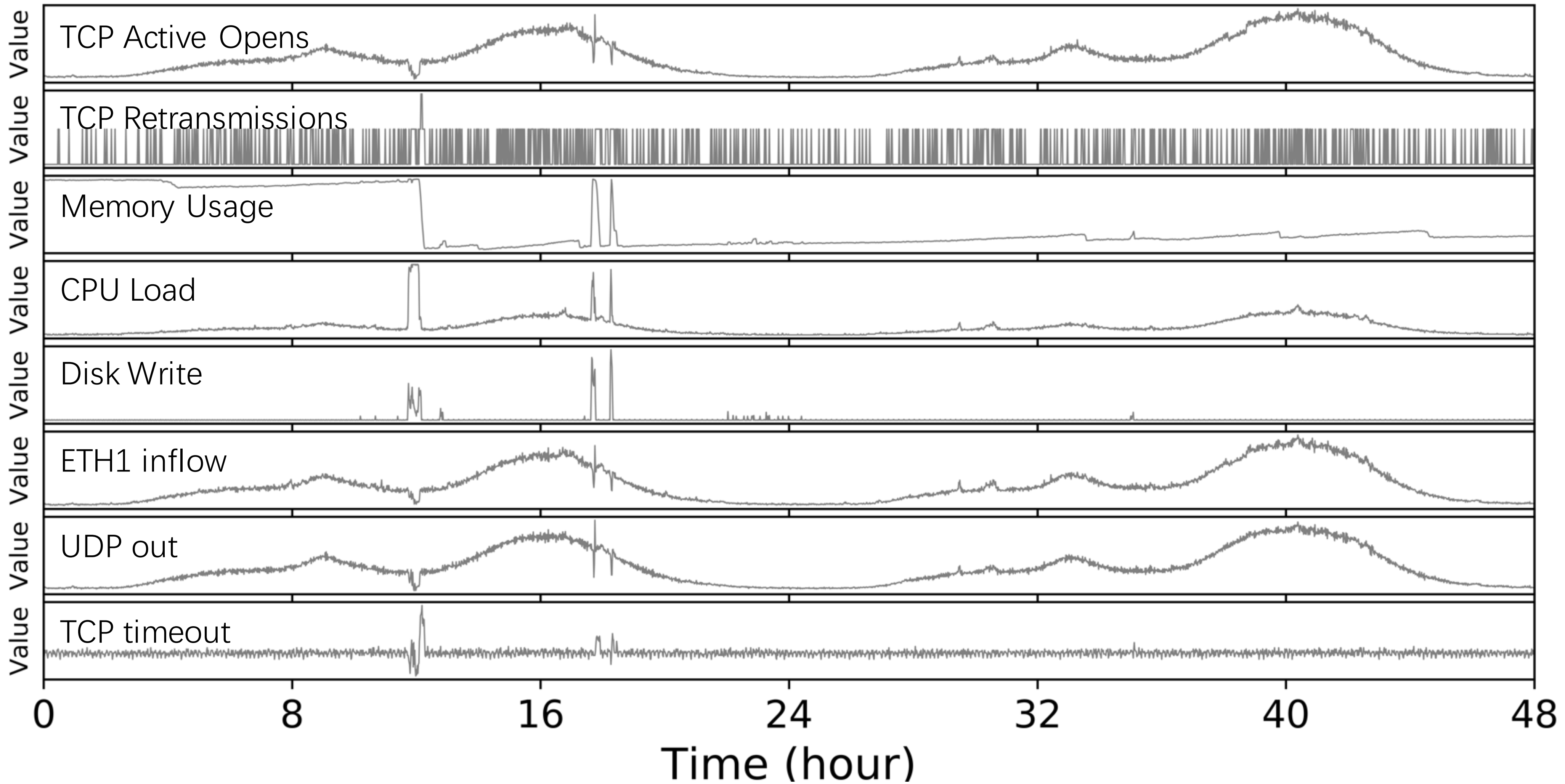
Multi-metrics

CPU Load
Network Usage
Memory Usage
...

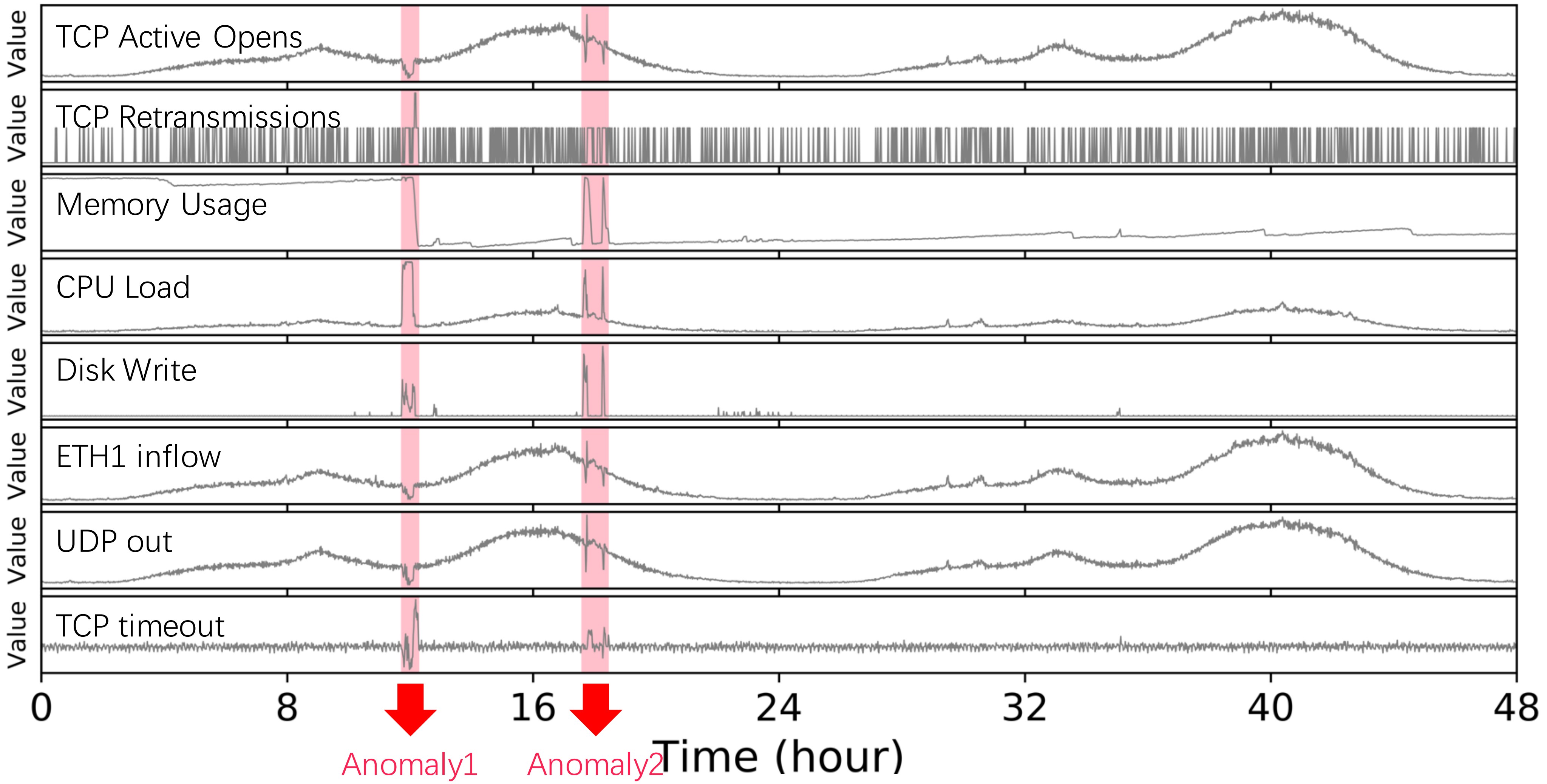


Multivariate time series

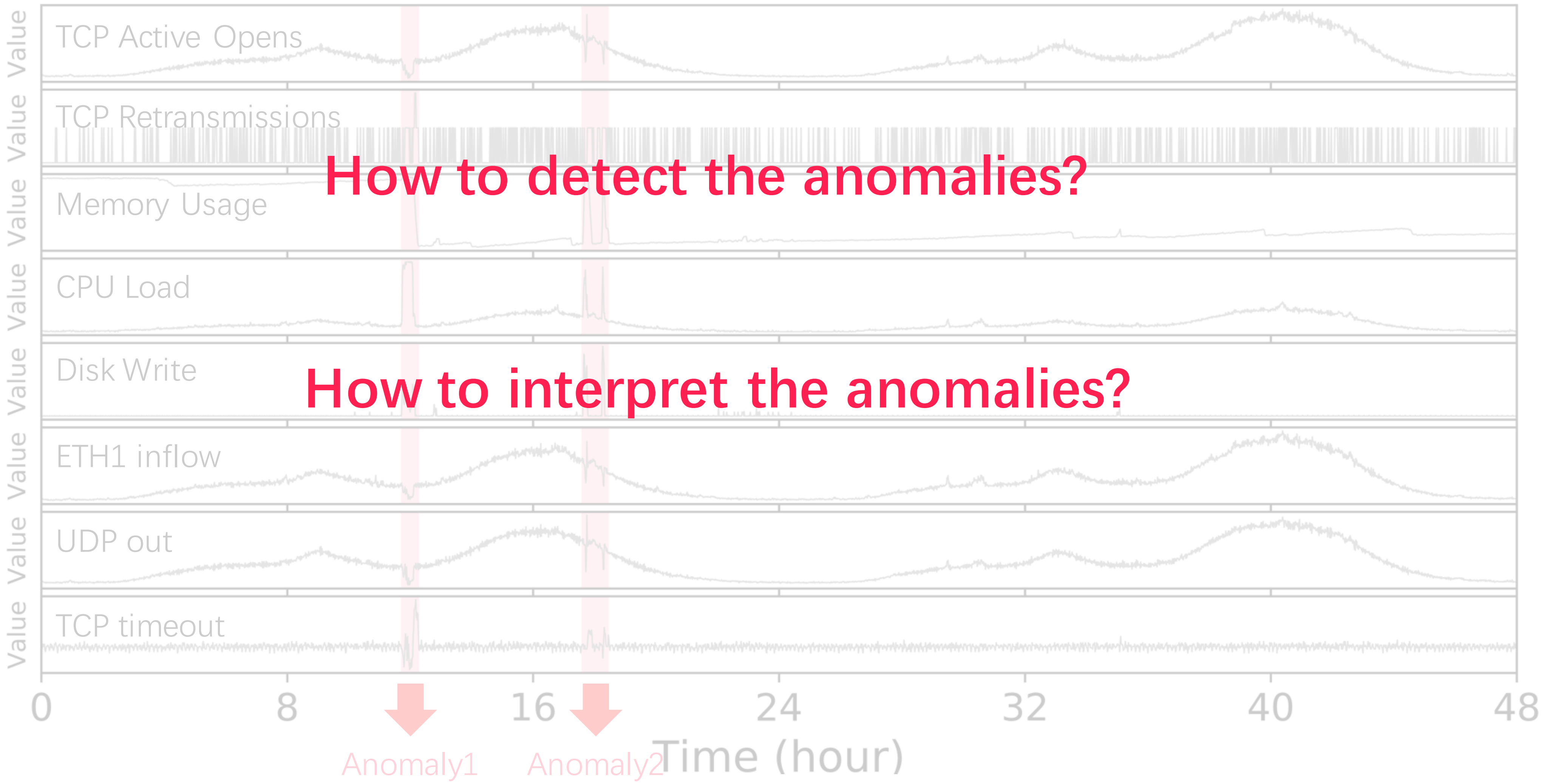
Machine with monitored multivariate time series



Machine with monitored multivariate time series



Motivations



Challenges

- How to deal with the temporal dependence of multivariate time series ?
- How to deal with the stochasticity of multivariate time series ?
- How to provide interpretation to the detected entity-level anomalies ?

Related work

Deterministic models	Stochastic based models
LSTM、 LSTM-based Encoder-Decoder [SIGKDD2018, ICML workshop 2016, NIPS 2016]	DAGMM、LSTM-VAE [IEEE Robotics and Automation Letters 2018, ICLR 2018]
Deterministic models without stochastic variables	Ignore the dependence of time series or stochastic variables.

Outline

Background

Algorithm

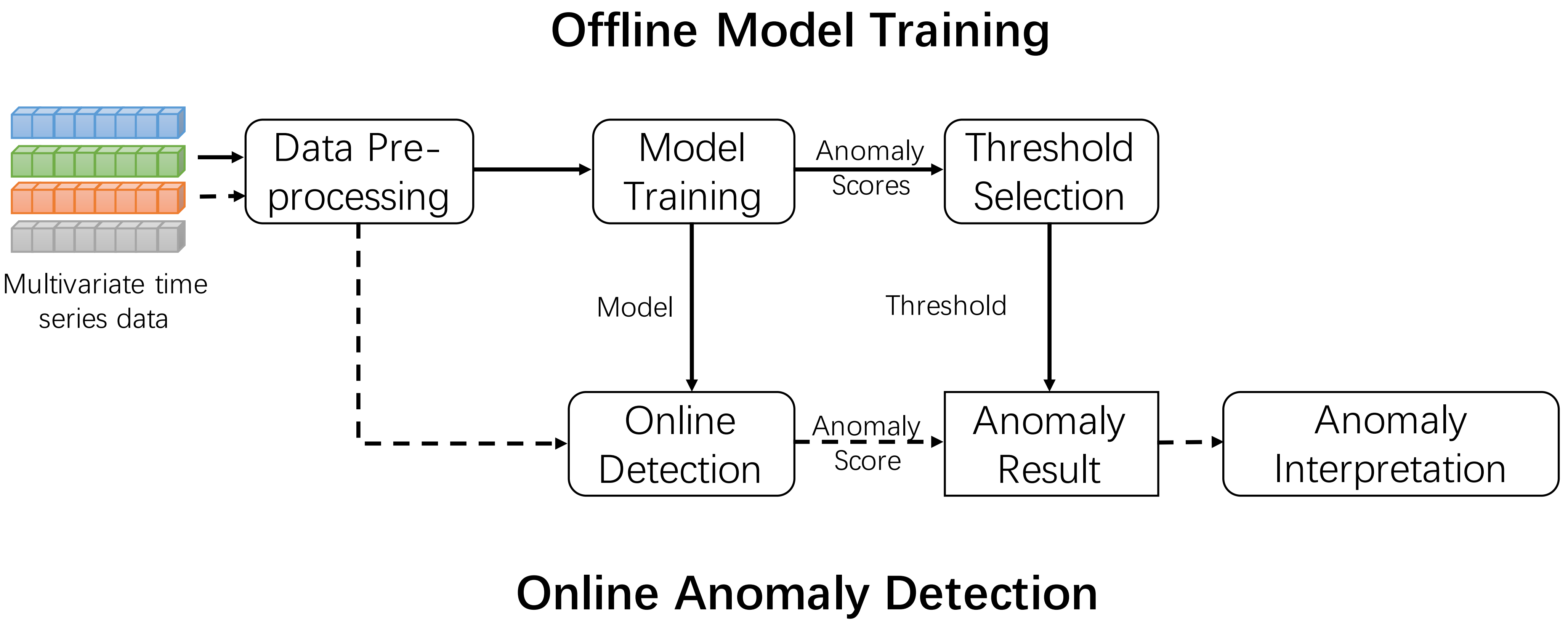
Evaluation

Conclusion

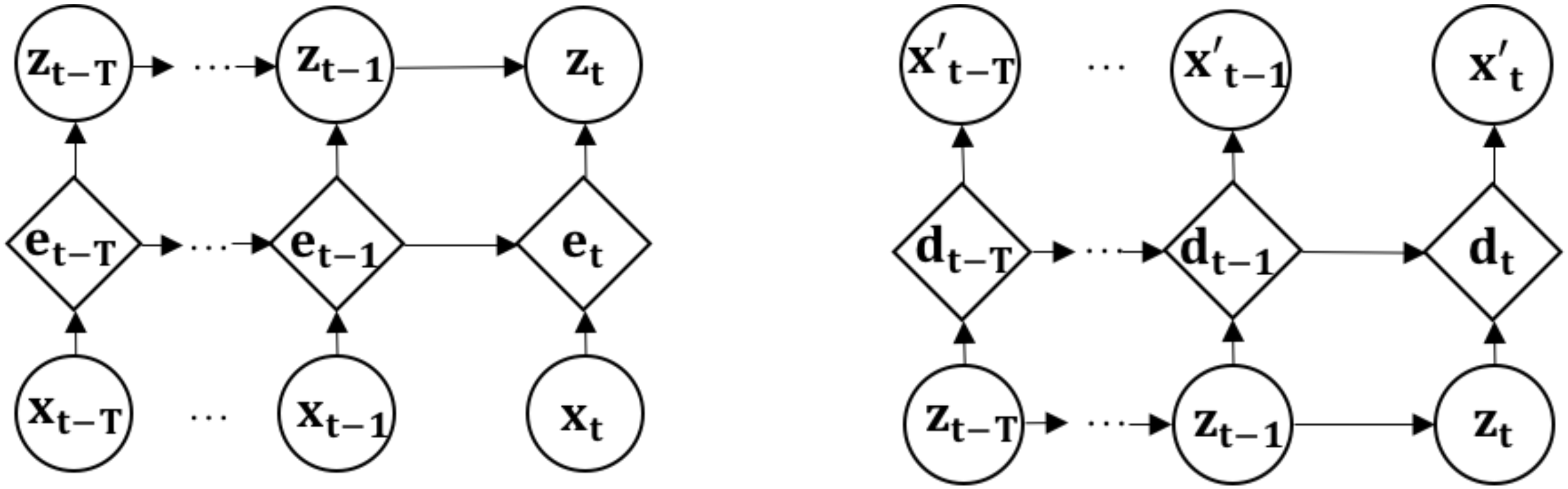
OmniAnomaly

Helps answer the questions

Structure of OmniAnomaly



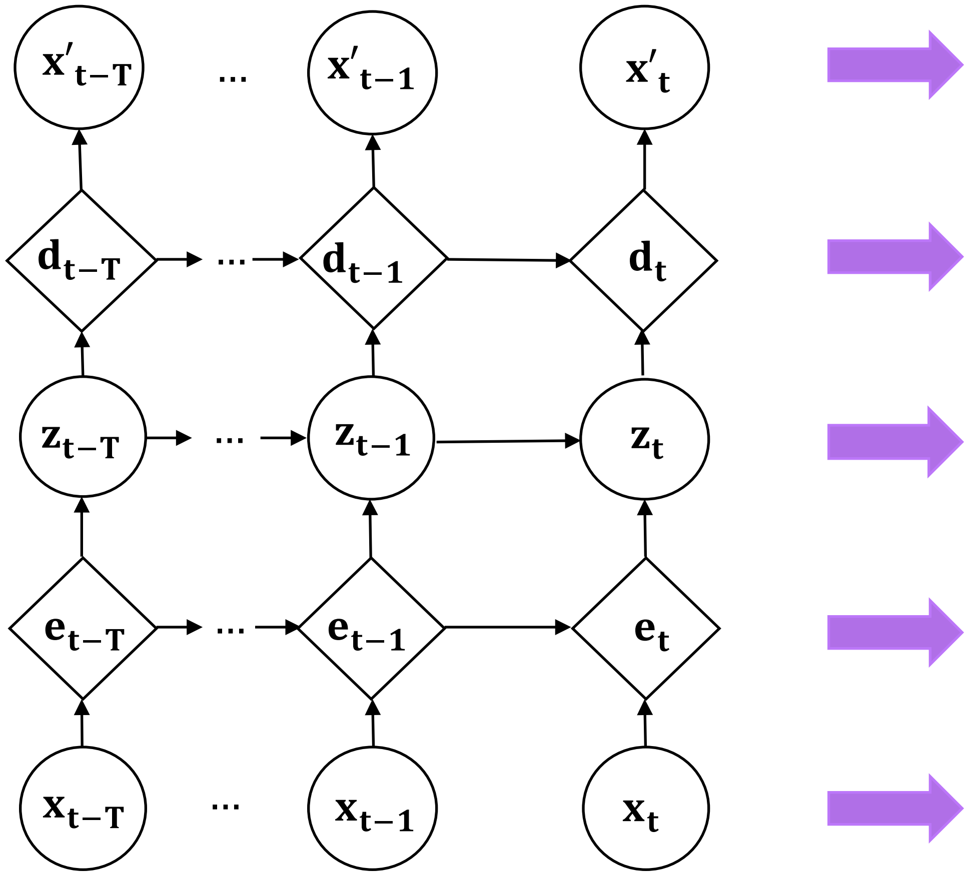
Model Architecture of OmniAnomaly



(a1) qnet

(a2) pnet

Model Architecture of OmniAnomaly



Reconstructed data

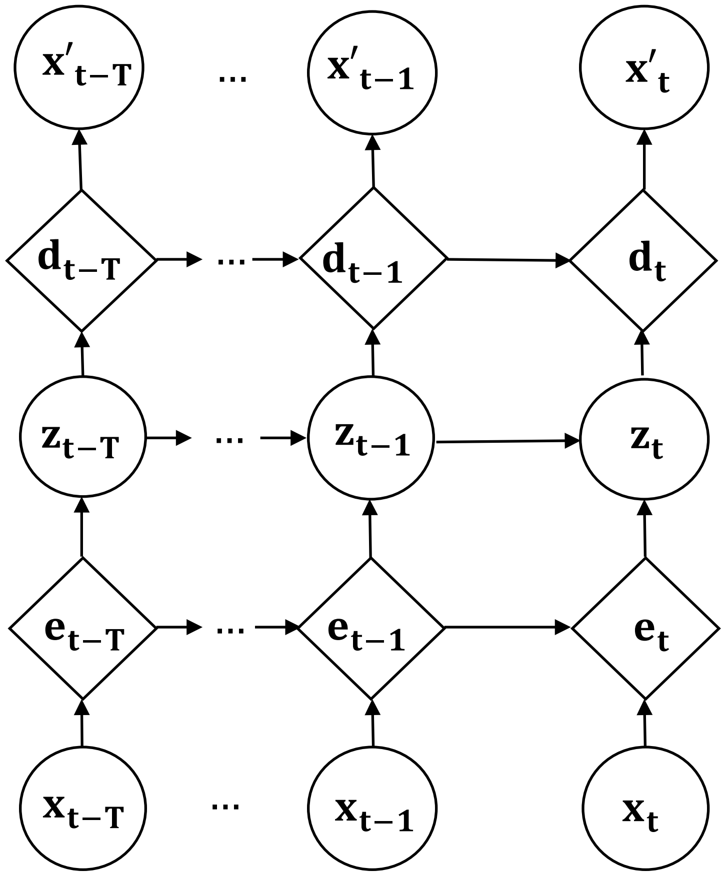
GRU cells for capturing temporal dependence

Stochastic cells for modeling data distribution

GRU cells for capturing temporal dependence

Input Sequence data

Core idea of OmniAnomaly



A good z_t can represent x_t well no matter x_t is anomalous or not.

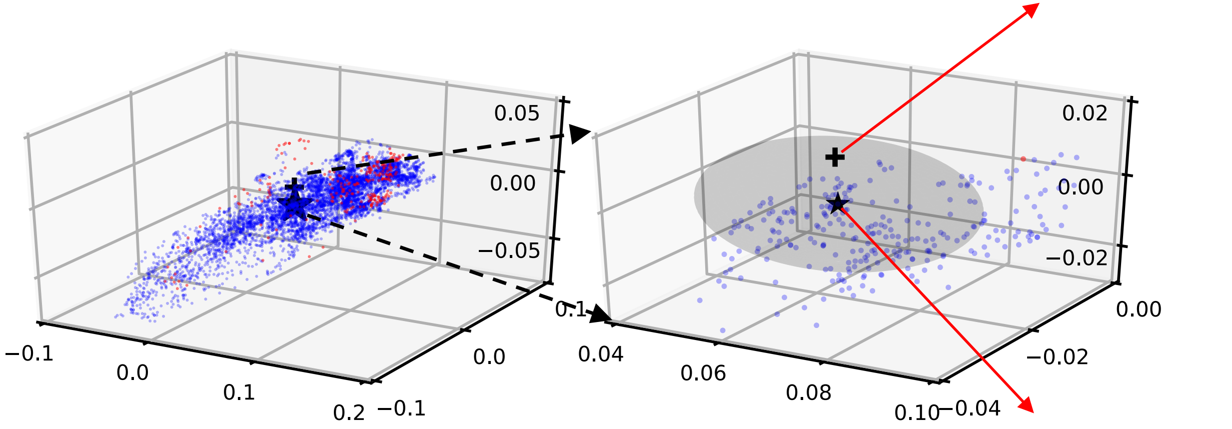
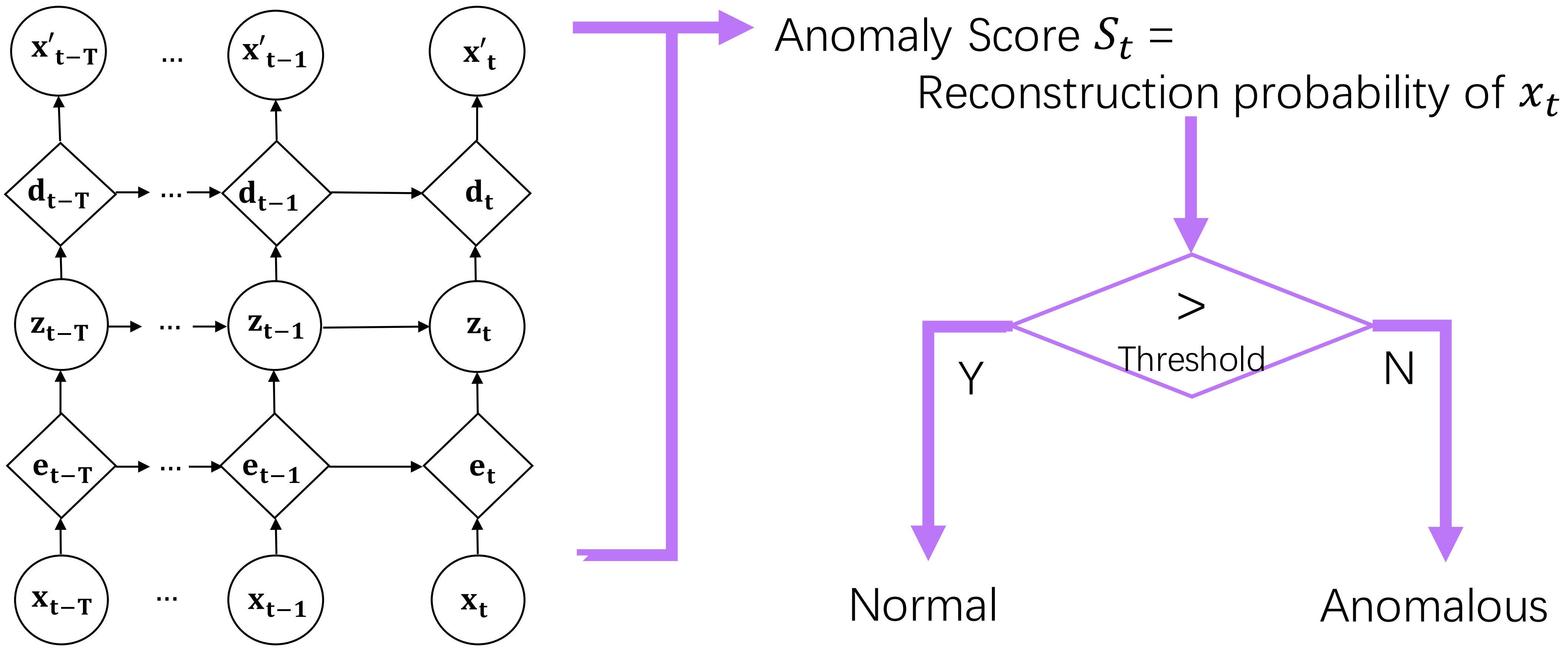


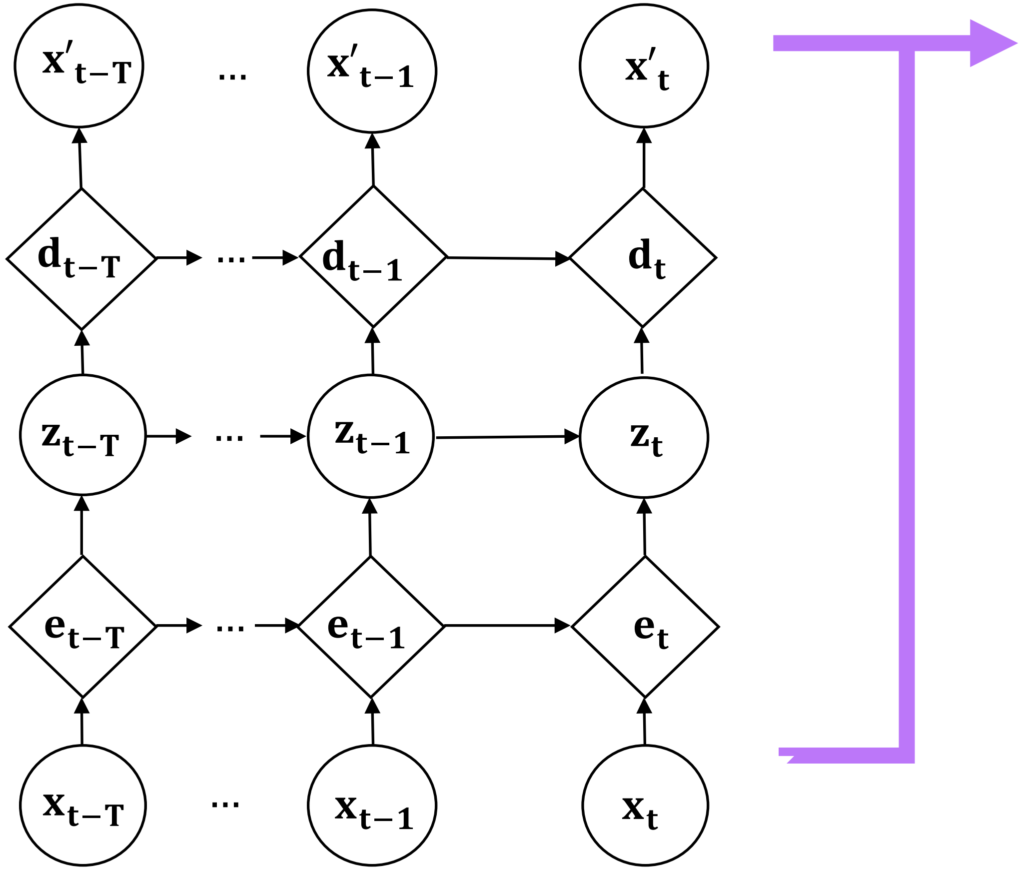
Fig: 3-dimensional z_t of x_t

When x_t is anomalous, its z_t can still represent its normal pattern and x_t' will be normal too.

Anomaly detection of OmniAnomaly



Anomaly detection of OmniAnomaly



Anomaly Score S_t = Reconstruction probability of x_t

$x_t = [x_t^1, x_t^2, \ldots, x_t^M]$, M is the dimension

$$S_t = \sum_{i=1}^{M} S_t^i$$

Sort the $[S_t^1, S_t^2, \ldots, S_t^M]$ in ascending order, and the Top K dimensions can interpret the anomaly.

Outline

Background

Algorithm

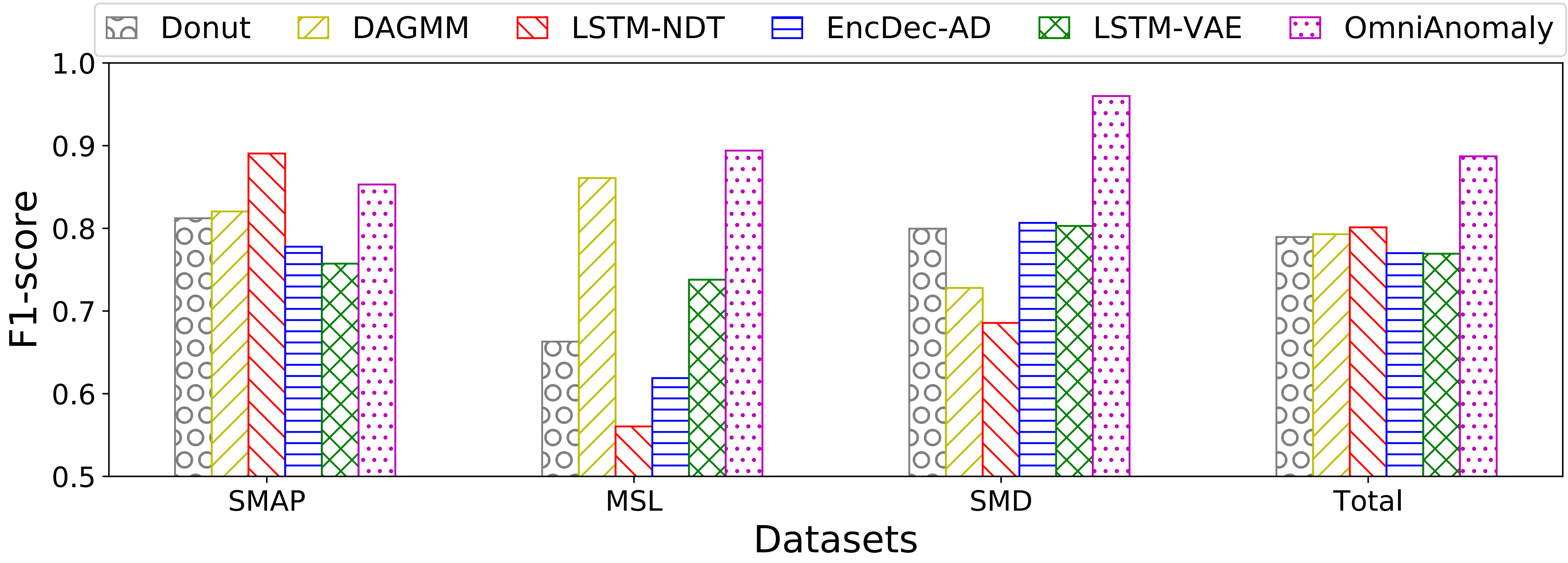
Evaluation

Conclusion

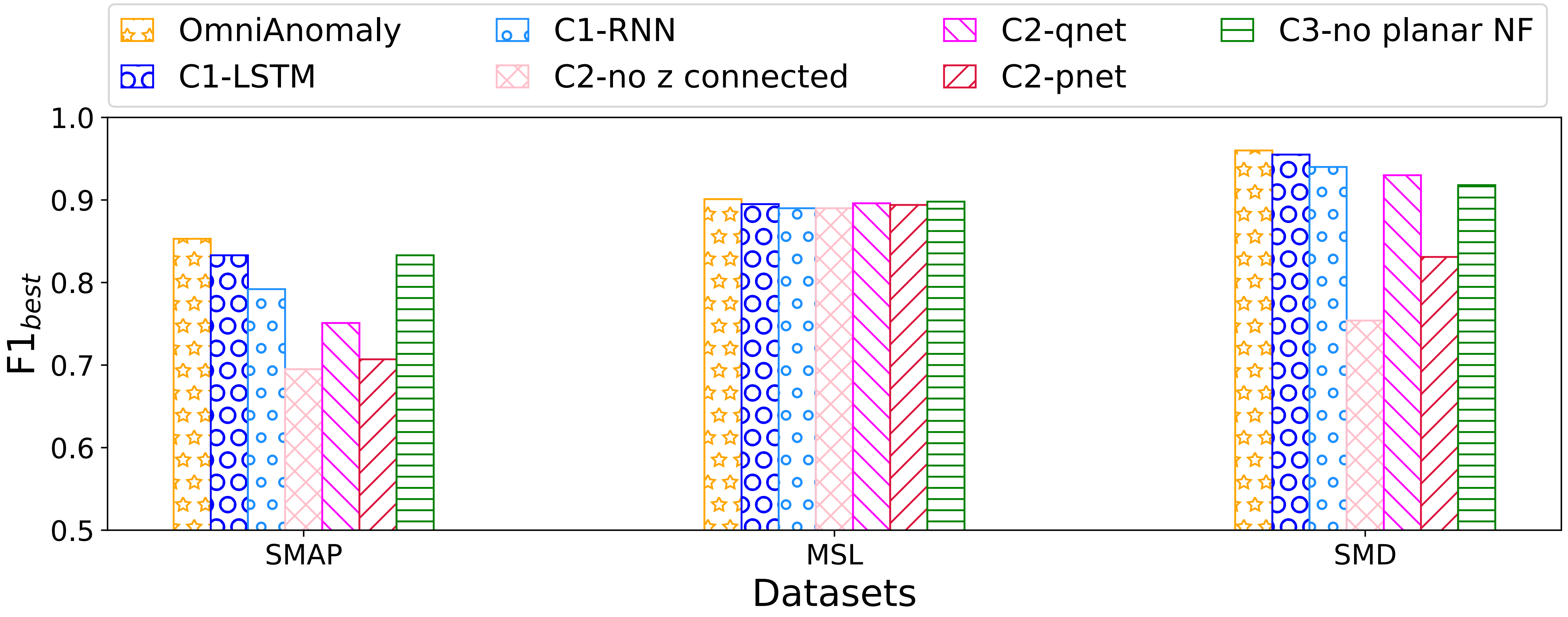
Datasets

DataSet name	Number of entities	Number of dimensions	Training set size	Testing set size	Anomaly ratio(%)
SMAP	55	25	135183	427617	13.13
MSL	27	55	58317	73729	10.72
SMD	28	38	708405	708420	4.16

F1-best of OmniAnomaly and baselines



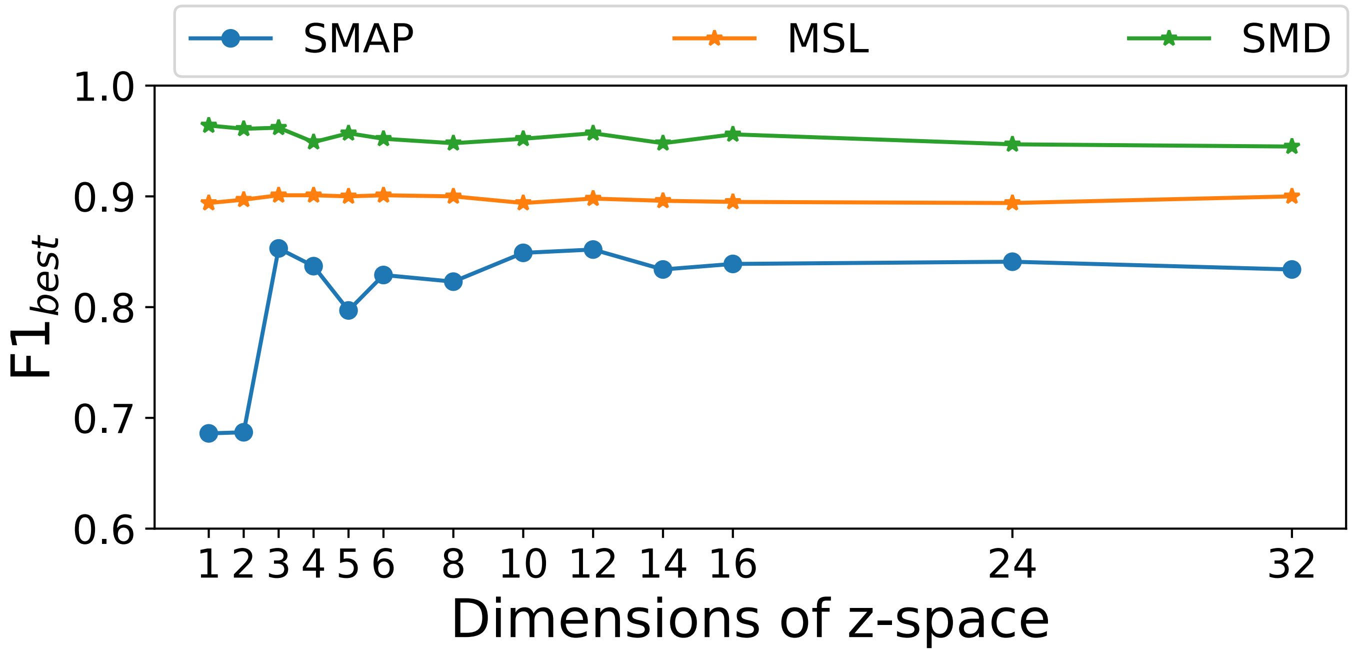
F1-best of OmniAnomaly and variants



F1 obtained through POT vs. F1-best

Evaluation metrics for OmniAnomaly	SMAP	MSL	SMD
F1 obtained through POT	0.8434	0.8989	0.8857
F1-best	0.8535	0.9014	0.9620

F1-best of OmniAnomaly with different z dimension



Outline

Background

Algorithm

Evaluation

Conclusion

OmniAnomaly

- The first multivariate time series anomaly detection method that deal with explicit temporal dependence among stochastic variables
- The first anomaly interpretation approach for stochastic based multivariate time series anomaly detection algorithms
- Achieve an overall F1-score of 0.86 in three real world datasets.
- The interpretation accuracy is up to 0.89.

Lessons for time series data learning

- A combination of stochastic deep Bayesian model and deterministic RNN model is necessary
- The connection of stochastic variables is necessary and effective
- It is necessary to assume non-Gaussian distributions in z-space

Lessons for for multivariate time series anomaly detection

- Reconstruction-based models are more robust than prediction-based models

- It is critical to obtain robust latent representations which can accurately capture the normal patterns of time series

- Reconstruction-based stochastic approaches offer an opportunity to interpret the anomalies

Thanks

su-y16@mails.tsinghua.edu.cn

清華大學
Tsinghua University