TOWARDS SPATIOTEMPORAL FOUNDATION MODEL IN CYBERSPACE

DAN PEI

TSINGHUA UNIVERSITY

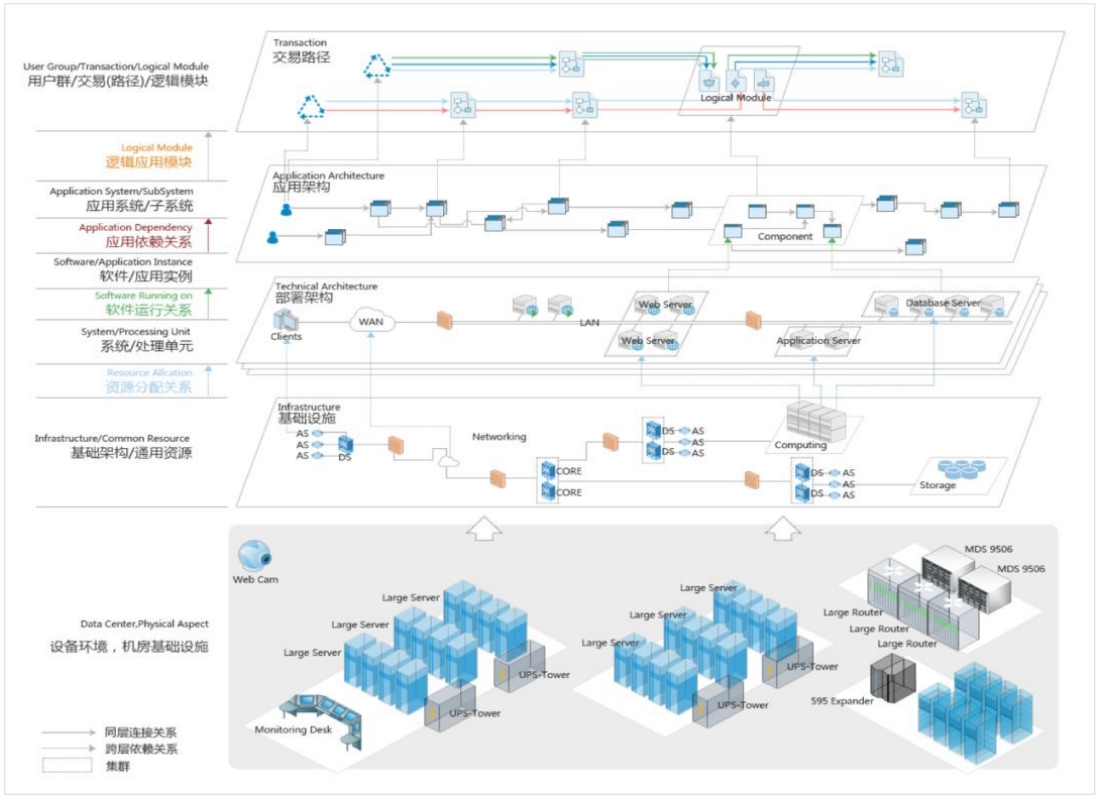
2024/04/26

OUTLINE

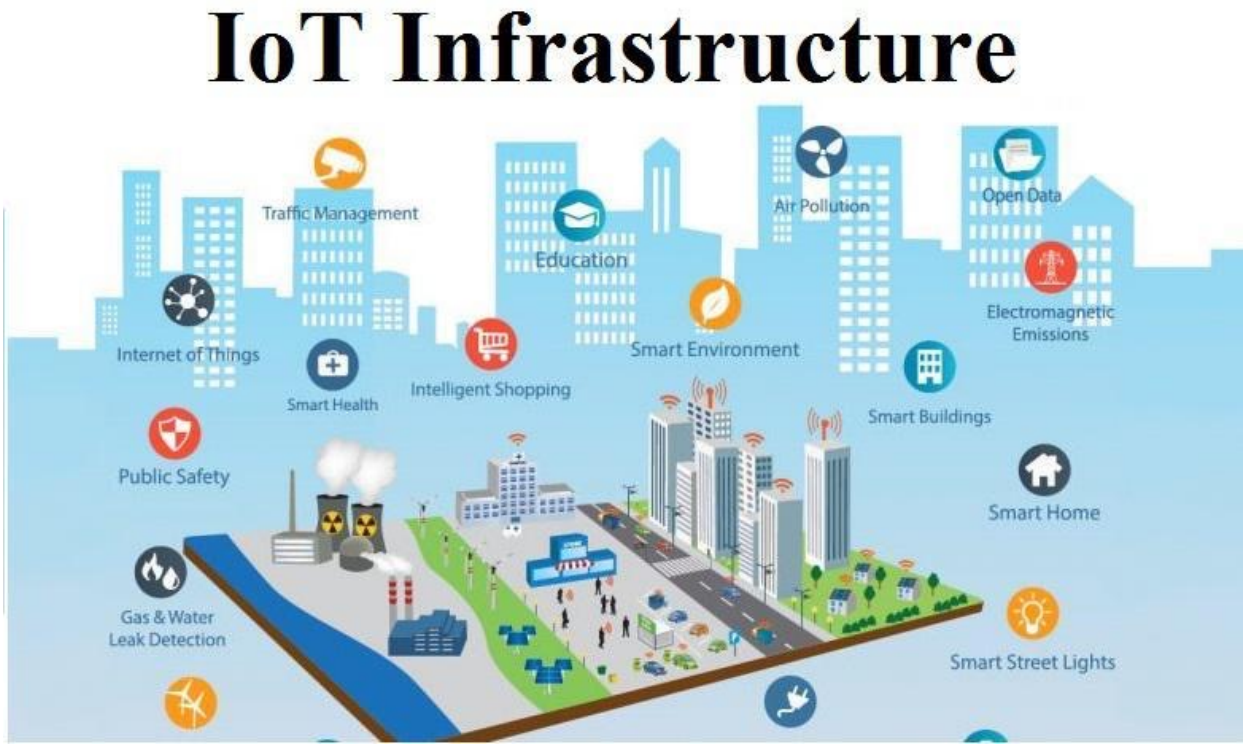
- The Definition of Cyberspace
- Spatiotemporal Data in Cyberspace
- Spatiotemporal Tasks in Cyberspace
- The necessity of explicitly modeling both spatial and temporal information
- NetMan's past efforts
- Outlook on Spatiotemporal Foundation Models in Cyberspace

DEFINITION OF CYBERSPACE

Cyber system



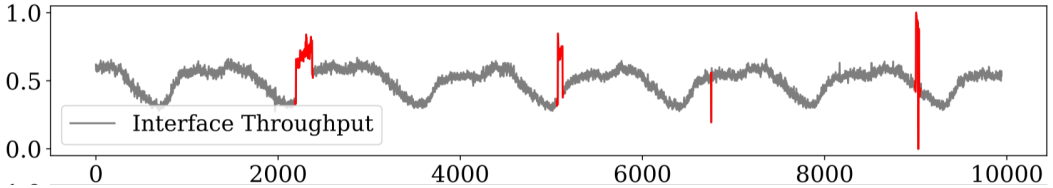
Cyber Physical System



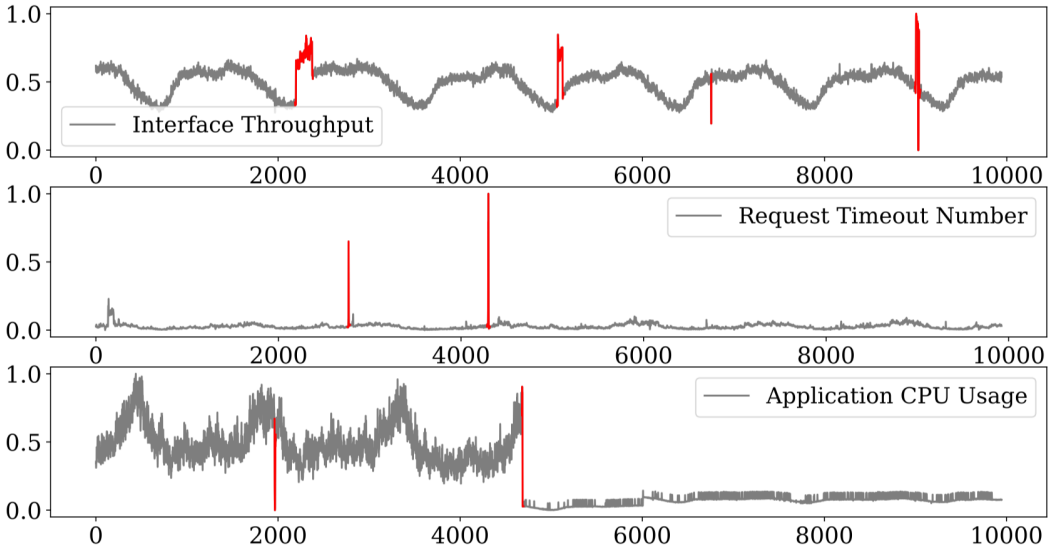
TIME SERIES DATA IN CYBERSPACE

- Time series data is a collection of data points arranged in temporal order.
- Time series data has significant temporal dependence, and both the value and time of data points affect their physical meaning.

- Univariate Time Series (UTS):
  - There is only one data point at a time
- Multivariate Time Series (MTS):
  - Multiple data points at the same time

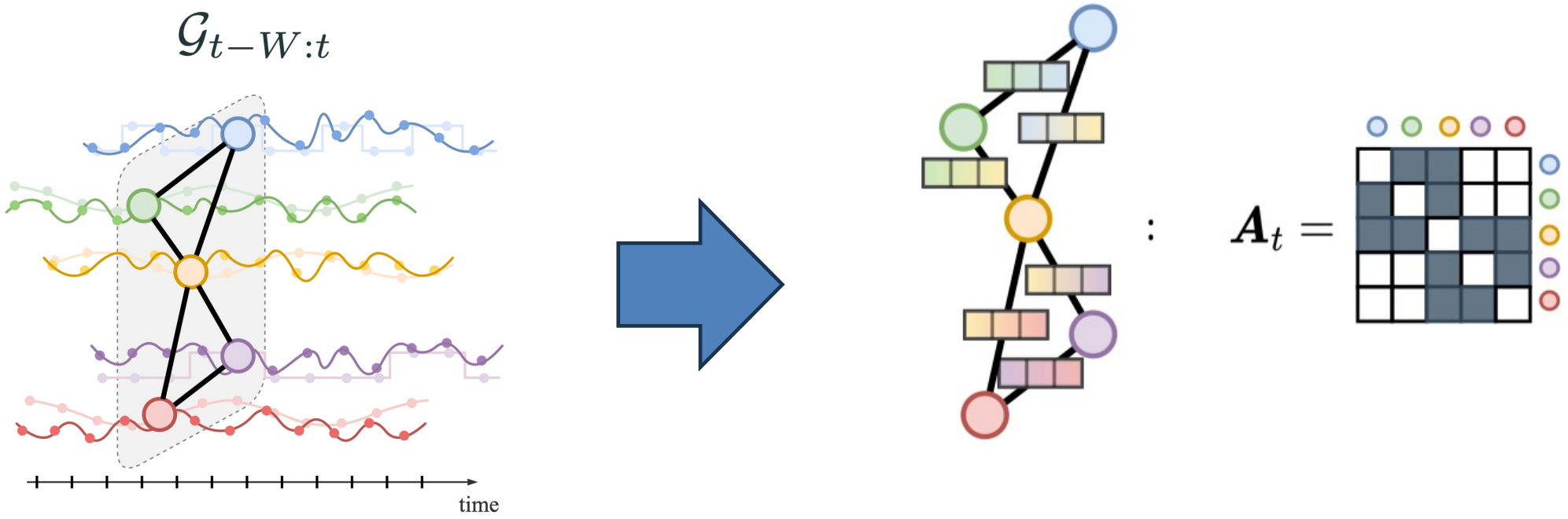


UTS



MTS

DEFINITION OF SPATIOTEMPORAL DATA IN CYBERSPACE



- The temporal information and spatial relationship between time series
- Taking server clusters as an example:
  - Temporal information: monitoring metrics
  - Spatial information: e.g. topology connections between servers

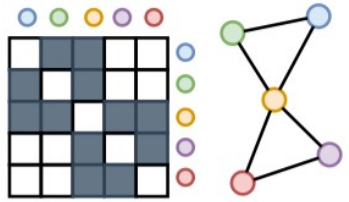
GENERALIZED SPATIOTEMPORAL DATA IN CYBERSPACE: GRAPH-STRUCTURED TIME SERIES DATA

Nodes

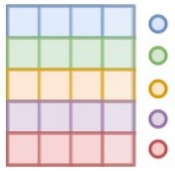
$\mathcal{V} = \{$ ○ ○ ○ ○ ○ $\}$

Adjacency Matrix

$\mathbf{A} =$

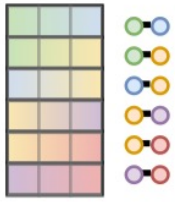


$\mathbf{V} =$

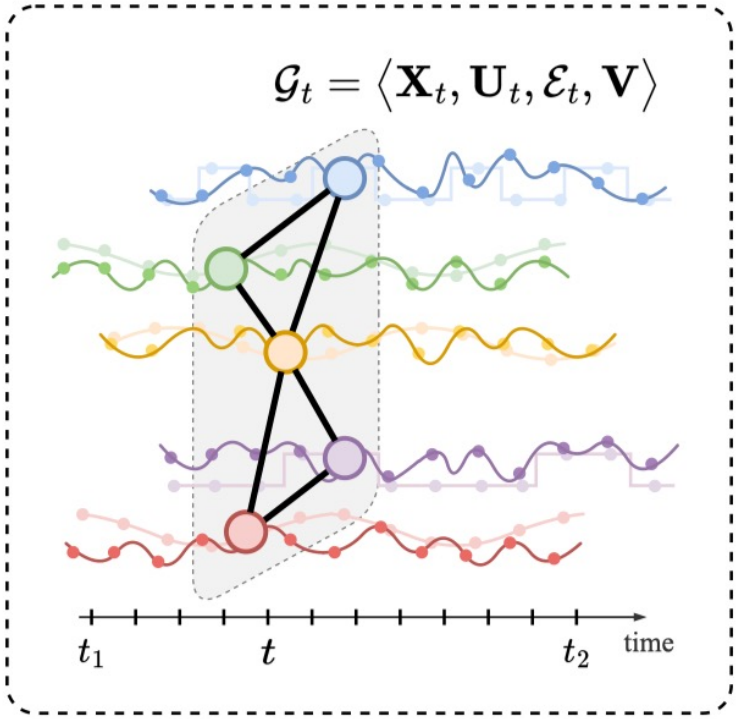


Static (time-independent) Attributes

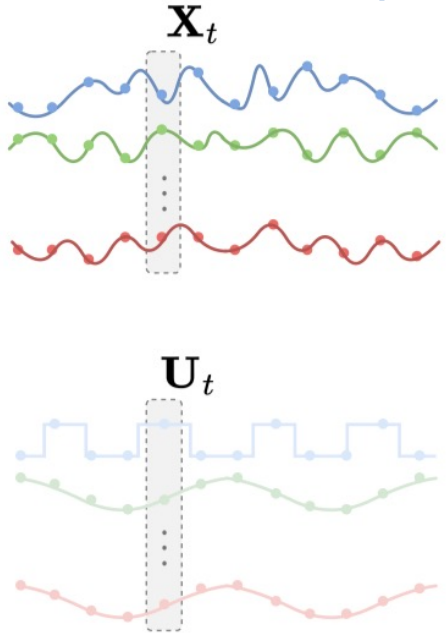
$\mathcal{E}_t =$



Edge Attributes



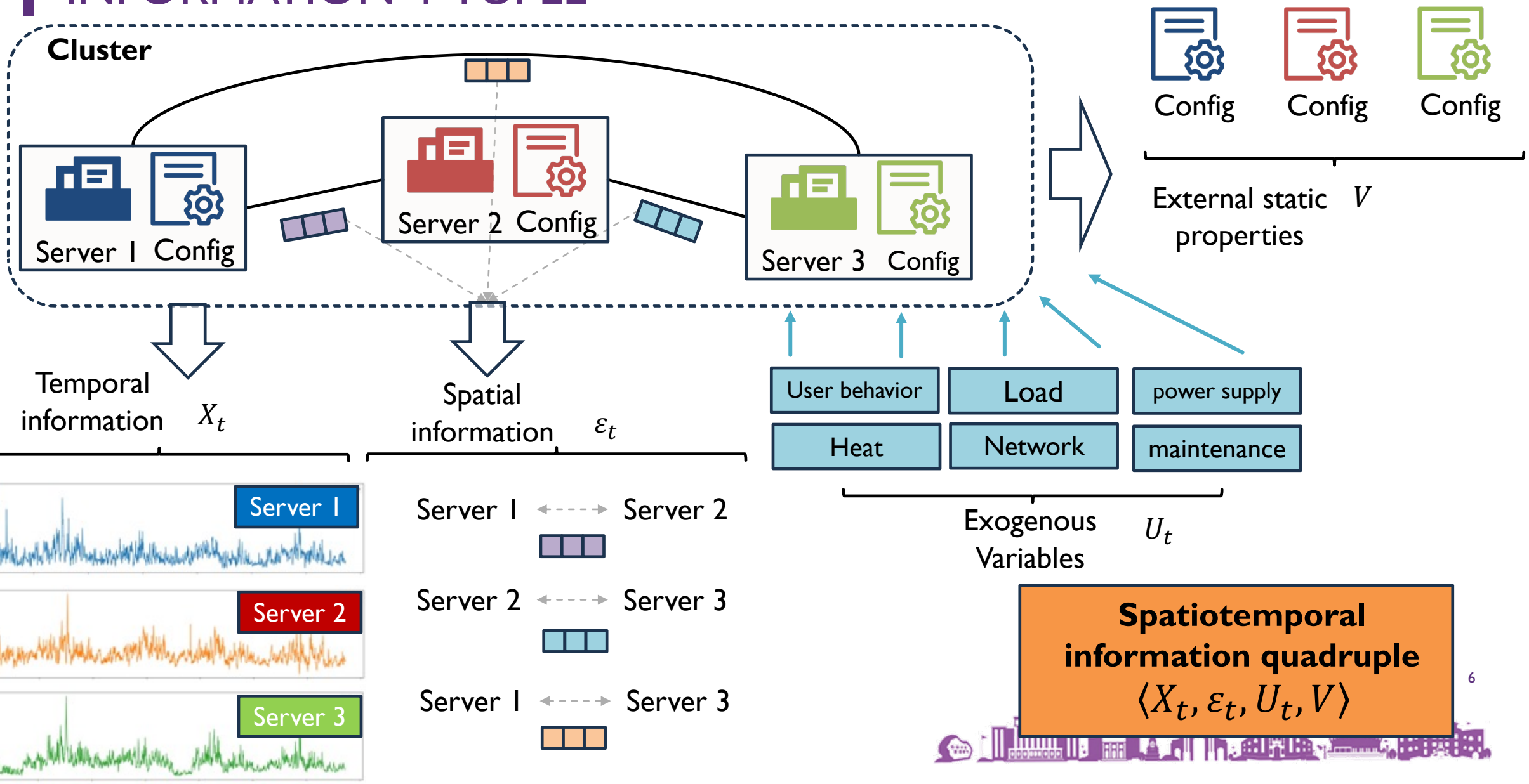
Observation Vectors (Time Series)



Exogenous Variables

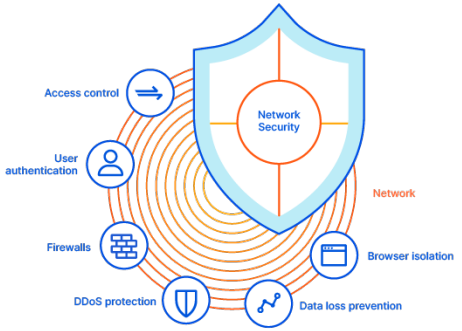
- Not limited to temporal and spatial information itself, but also includes external information (dynamic and static)

TAKING AIOPS AS AN EXAMPLE: SPATIO-TEMPORAL INFORMATION 4-TUPLE

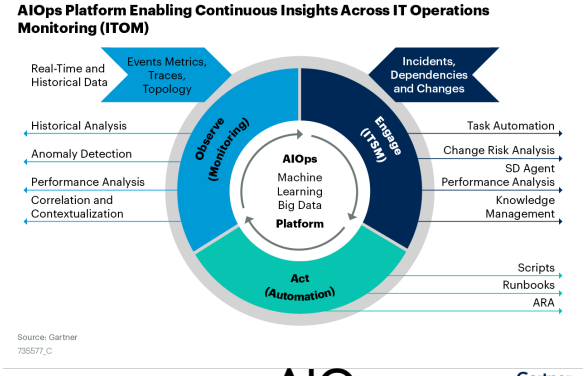


THE DOMAINS COVERED BY SPATIOTEMPORAL DATA IN CYBERSPACE

Cyber System



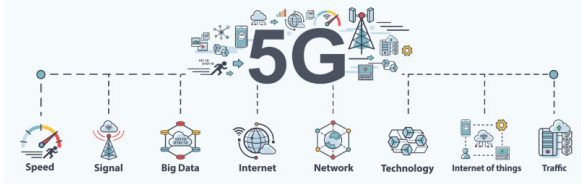
Network Security



AIOps

Cloud Computing

Streaming Media



Wireless Network

Social Networks

Cyber Physical System

IoT

Smart Transportation

Smart Healthcare

Industrial Internet

SPATIOTEMPORAL TASKS IN CYBERSPACE

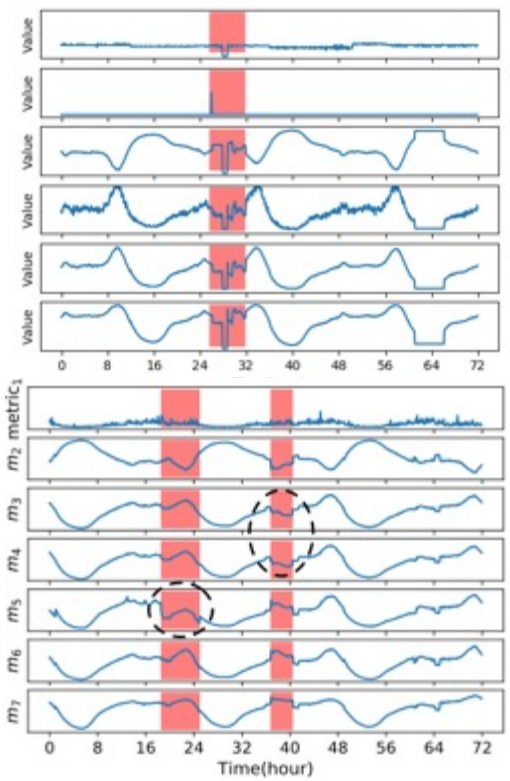
Classify spatiotemporal tasks into the three categories

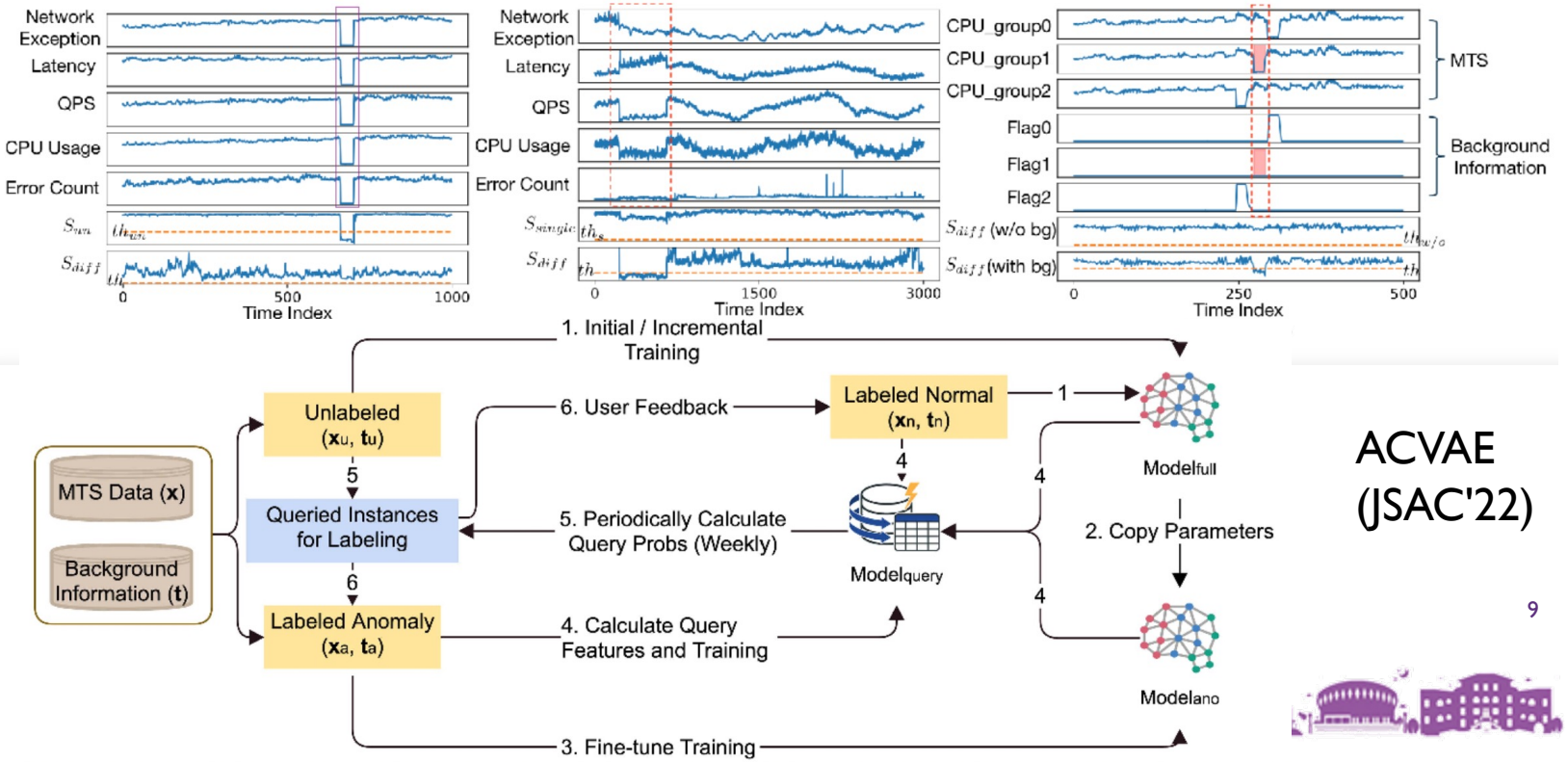
Temporal Analysis	Time series prediction, anomaly detection, data imputation, classification, concept drift detection, what-if analysis…
Spatial Analysis	Clustering, outlier detection, causal discovery, causal inference, relationship prediction…
Spatiotemporal Anaysi	Diagnostic analysis, root cause localization, behavioral analysis…

EXAMPLES OF CYBERSPACE TASKS FOCUSED ON TEMPORAL ANALYSIS

Taking anomaly detection as an example

- The conventional approach is to learn normal patterns from the time series data itself to determine anomalies
- In some cases, anomalies in time series are difficult to determine, and relying solely on temporal information cannot cover all scenarios → The need for exogenous variables



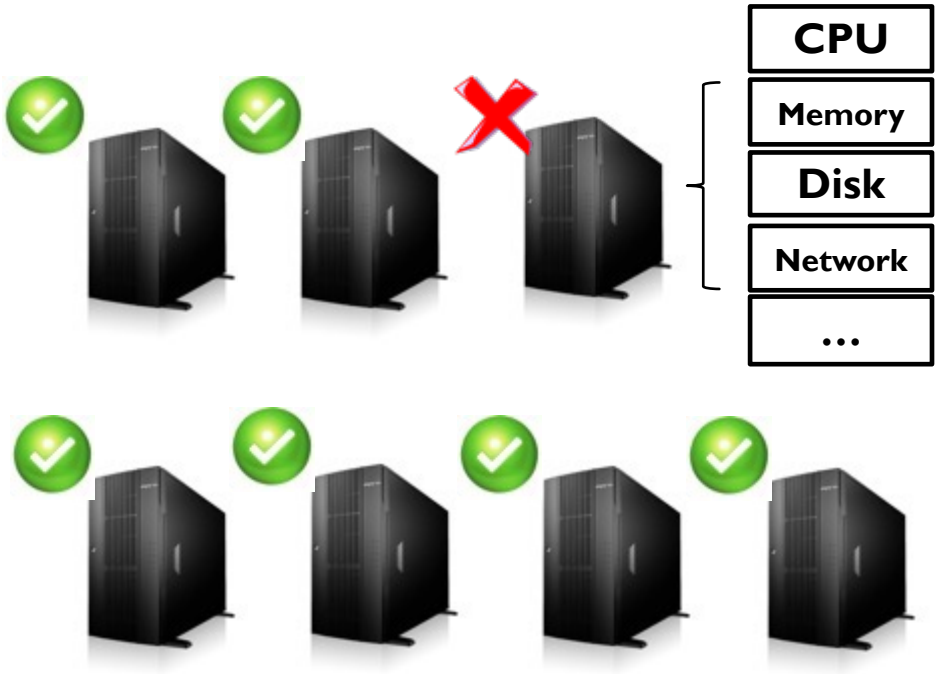


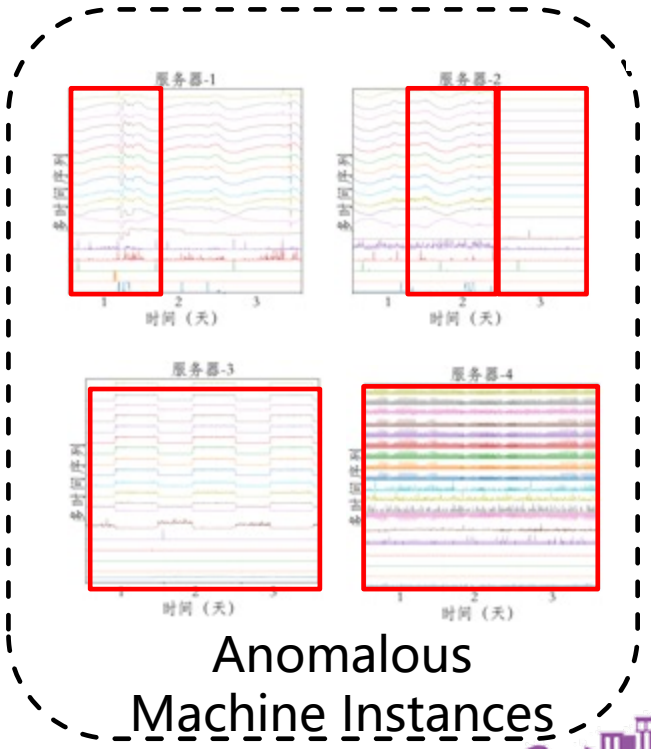
ACVAE (JSAC'22)

EXAMPLES OF CYBERSPACE TASKS OF FOCUSED ON SPATIAL ANALYSIS

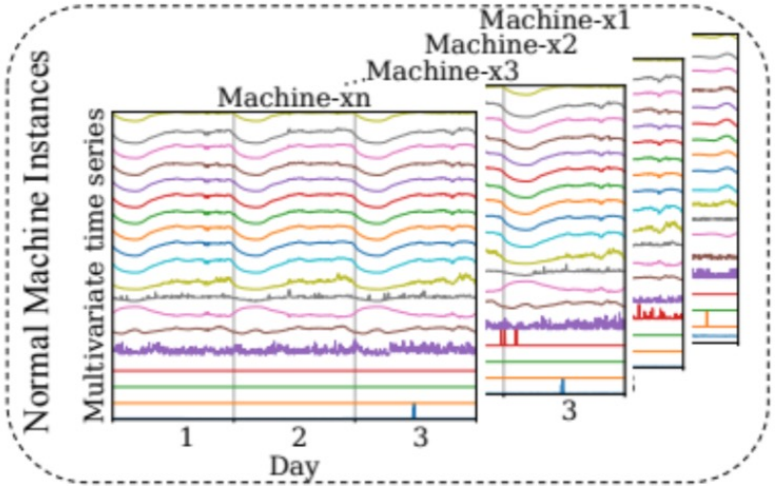
Taking outlier detection as an example

- Detected instances that do not conform to normal mode
- Judging through evaluation metrics such as the shape, similarity, and distan of monitoring data is often not comprehensive → The need for temporal information





Anomalous Machine Instances



Normal Machine Instances

1) Similar in shape but with long-term anomalies

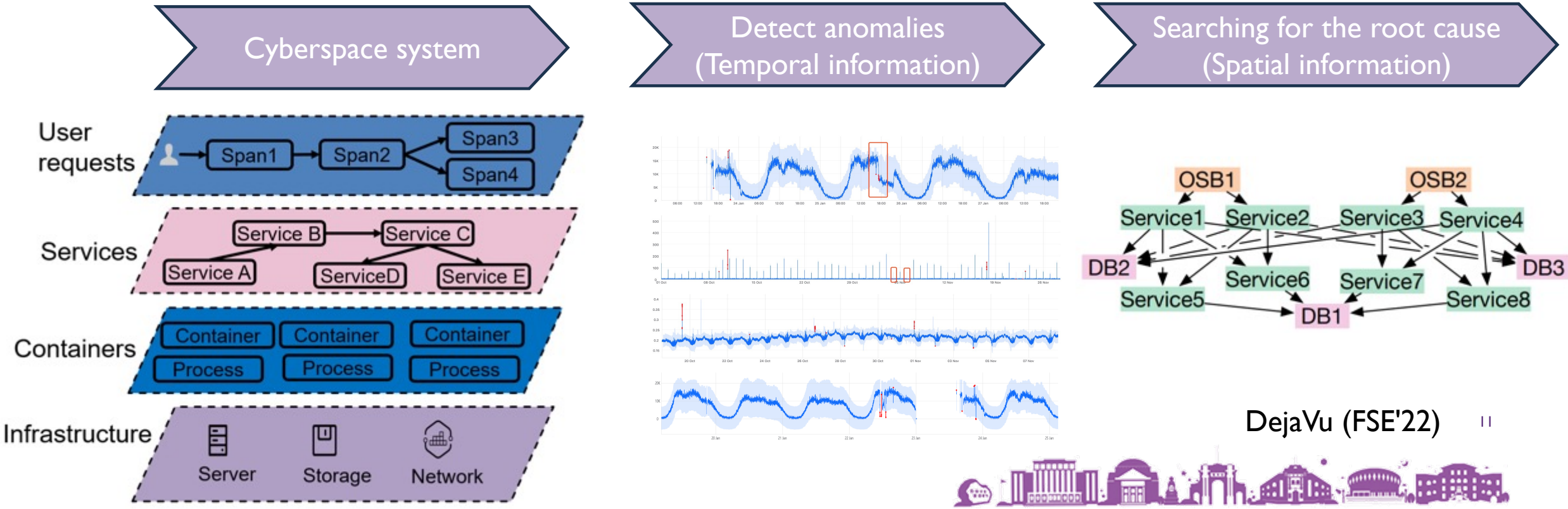
2) Having regular shapes but not similar

3) Irregular shape

AN EXAMPLE CYBERSPACE TASK FOCUSED ON SPATIOTEMPORAL ANALYSIS

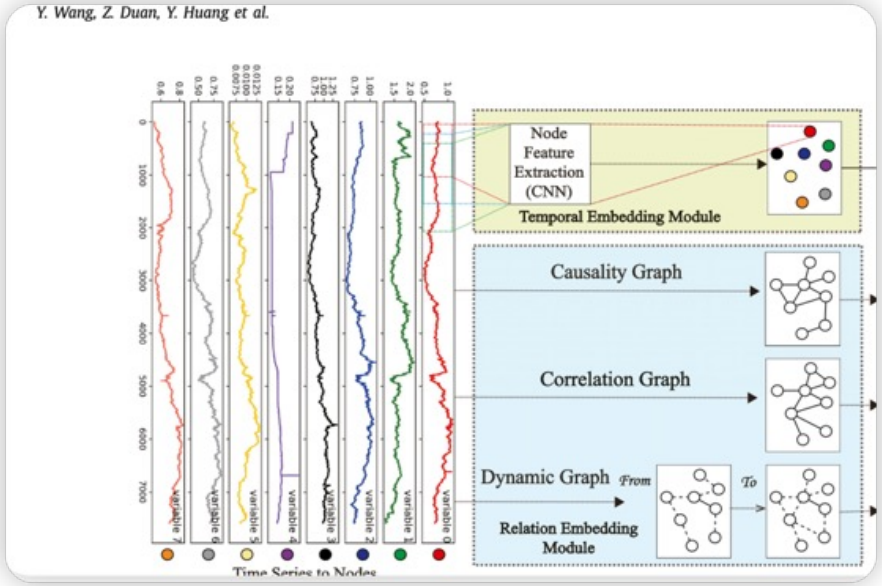
Taking Root Cause Analysis as an Example (Temporal Then Spatial)

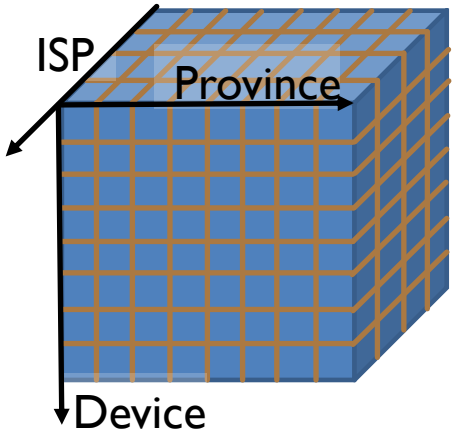
- RCA in Cyberspace systems:
  - Usually, it is necessary to first perform anomaly detection on the time series
  - Using anomaly results as the input to root cause analysis

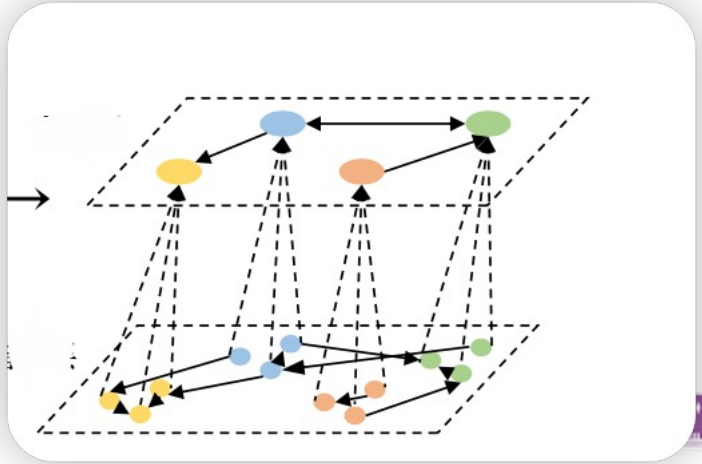


SPATIAL INFORMATION IN CYBERSPACE

- The complex spatial information in Cyberspace can be formulated as graph
- In Cyberspace, graphs are usually heterogeneous:
  - Topology Graph
  - Call Graph
  - Correlation Graph
  - Causal Graph
  - Knowledge Graph
- Graph can be hierarchical or Multi-attribute (Cubiod)
- The graph is usually incomplete, probabilistic, and dynamic





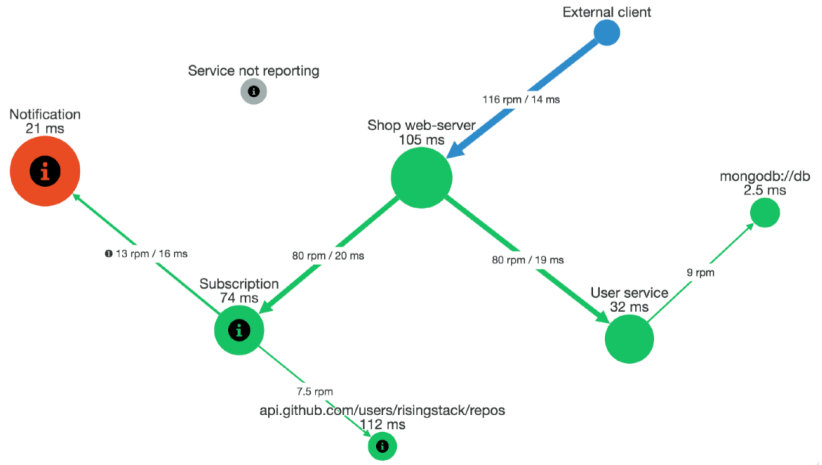


SUMMARY OF SPATIOTEMPORAL TASKS IN CYBERSPACE

Task \ Focus	Temporal Analysis	Spatial Analysis	Time=>Space	Time+Space
Time series forecasting (TS forecasting)	✅			
Univariate anomaly detection (UTS AD)	✅			
Multivariate anomaly detection (MTS AD)	✅			
What if Analysis	✅			
Causal Discovery		✅		
Causal Inference		✅		
Outlier Detection		✅		
Time series clustering (TS Cluster)		✅		
Trace anomaly detection (Trace AD)				✅
Root cause analysis (RCA)			✅	

THE CHALLENGES FACED WHEN USING ONLY TEMPORAL INFORMATION

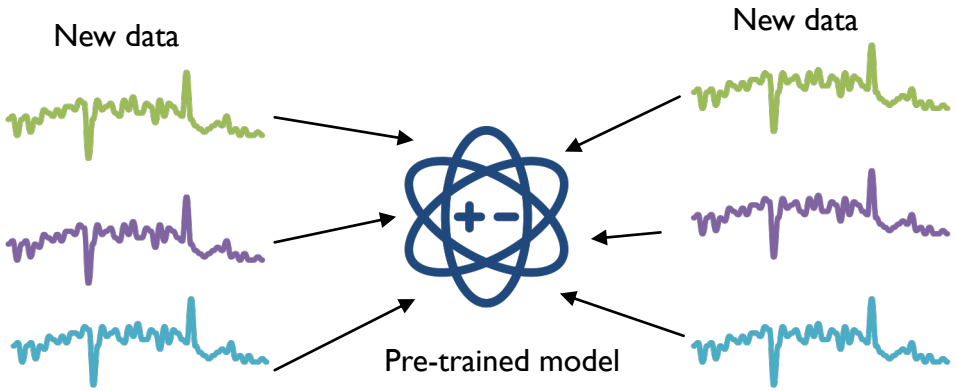
- Availability challenge:
  - There is a lot of room for improvement in performance



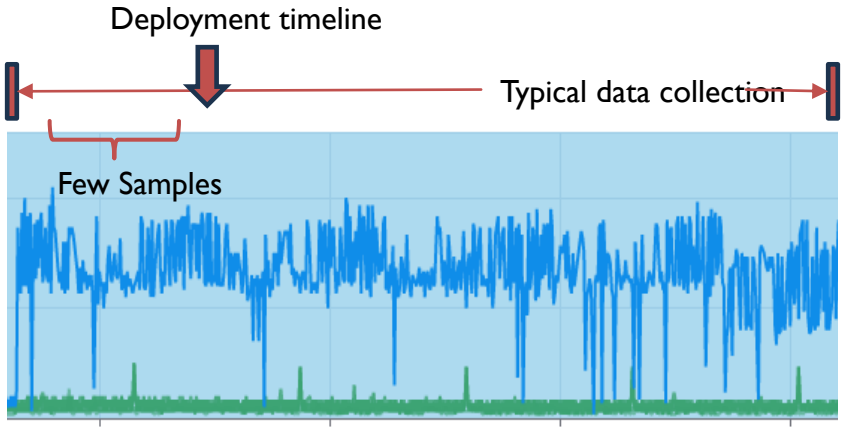
- Practical challenges:
  - Interpretability
  - Interactive

- Universal challenge:
  - Cross-task (prediction, anomaly detection, classification, root cause localization...)
  - Cross-domain (AIOps, network security, Internet of Things...)
  - Few-shot, zero-shot

THE SIGNIFICANCE OF FEW-SHOT AND ZERO-SHOT IN CYBERSPACE

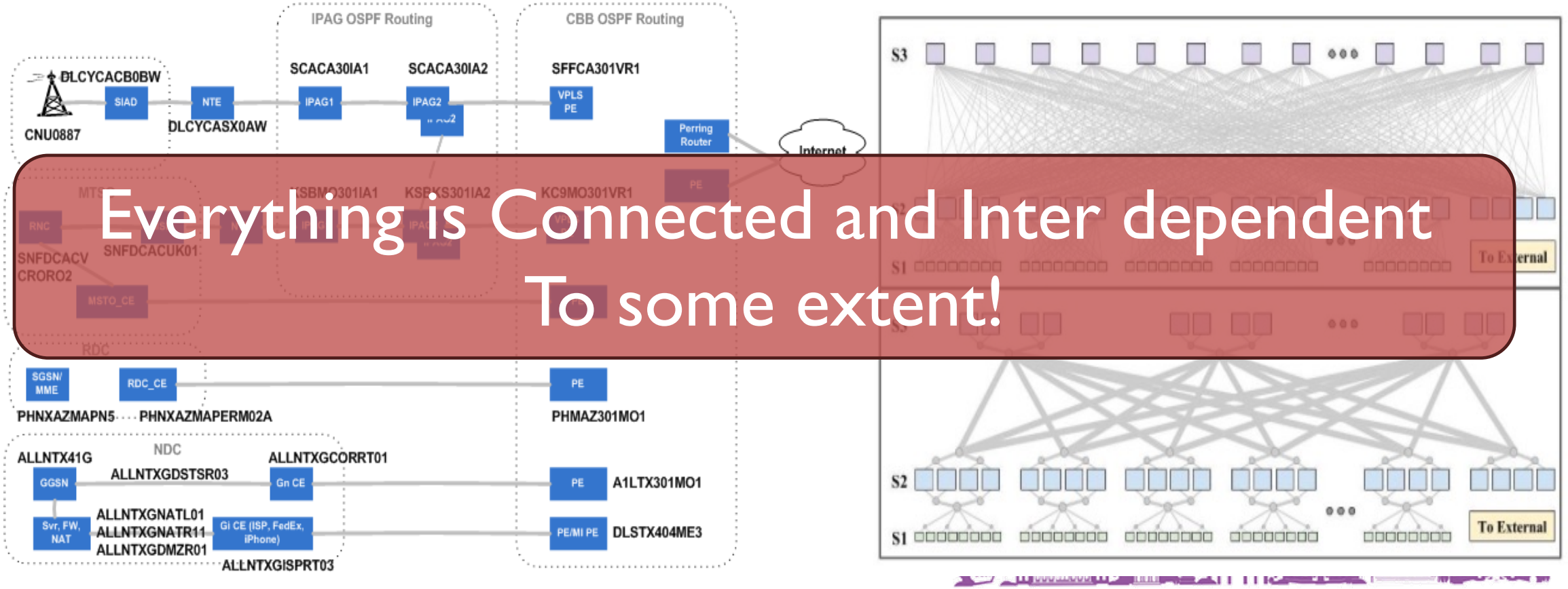


- The universality of testing models with few-shot or zero-shot samples
- Require the model to learn deep knowledge from the data
- How to define few-shot samples and zero-shot samples in the spatiotemporal domain?
  - Few-shot sample capability: Fast fine-tuning with a small amount of data can significantly improve performance
  - Zero-shot sample capability: Can achieve relatively satisfactory performance by directly working with new data without the need for additional fine-tuning
- The significance of few-shot or zero-shot samples for model application
  - Collecting data in a production environment requires a significant amount of time
  - Whether it can be "plug and play" affects whether the model can be quickly deployed



WHY DO WE NEED BOTH TIME AND SPACE IN CYBERSPACE

- Time series are usually not self-contained data
- The time series are more likely to be observation of the "effects", but the "cause" variable might not directly overserved.
- Therefore, it is necessary to keep adding more relevant time series, the their relationship with existing time sereies



Everything is Connected and Inter dependent

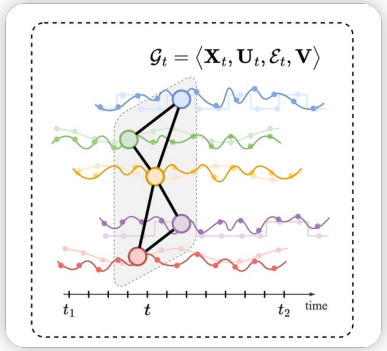
To some extent!

FEASIBILITY ANALYSIS IN CYBERSPACE

- Different tasks often focus on utilizing different temporal and spatial information
- Different tasks ultimately boils down to extracting knowledge from spatiotemporal information and utilizing it anyway. Then why not directly model it as a whole using the 4-tuple, graph-structed time series data?

- Time series forecasting (TS forecasting)
- Univariate anomaly detection (UTS AD)
- Multivariate anomaly detection (MTS AD)
- What if Analysis
- Causal Discovery
- Causal Inference
- Outlier Detection
- Time series clustering (TS Cluster)
- Trace anomaly detection (Trace AD)

- Time information
- spatial information
- Time=>Space
- Time+Space



Spatiotemporal information

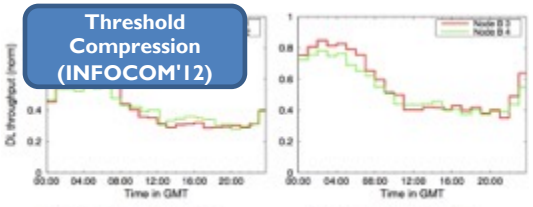
NETMAN'S PAST EFFORTS IN SPATIOTEMPORAL ANALYSIS IN CYBERSPACE

Focus on Temporal Analysis

- Mainly focused on anomaly detection

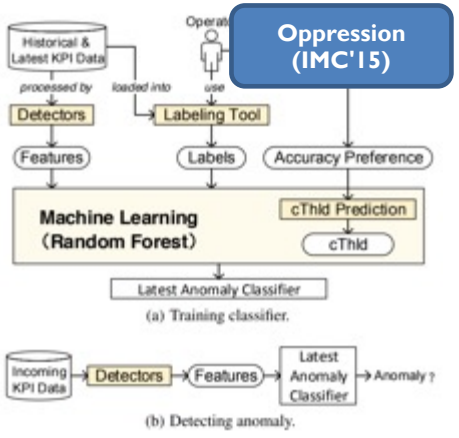
Single dimensional and multi-dimensional anomaly detection

Threshold Compression (INFOCOM'12)



(a) NodeB1 and NodeB2 (b) NodeB3 and NodeB4

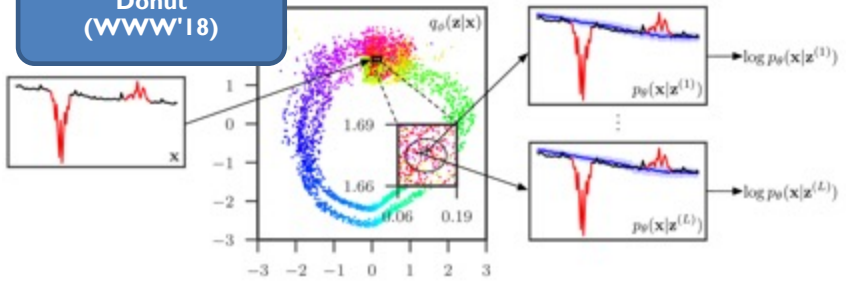
Oppression (IMC'15)



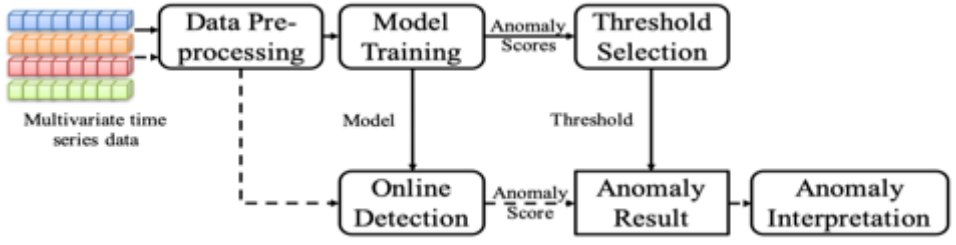
(a) Training classifier.

(b) Detecting anomaly.

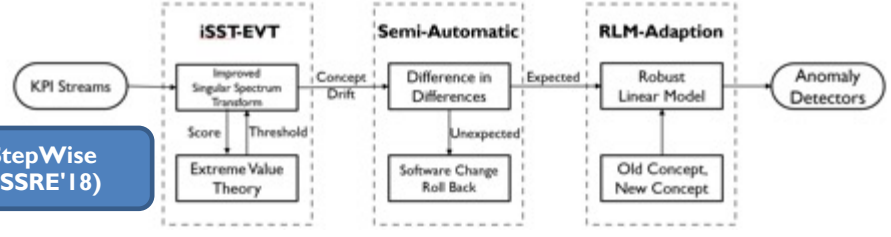
Donut (WWW'18)



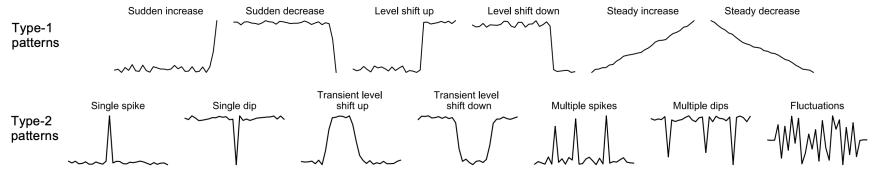
OmniAnomaly (KDD'19)



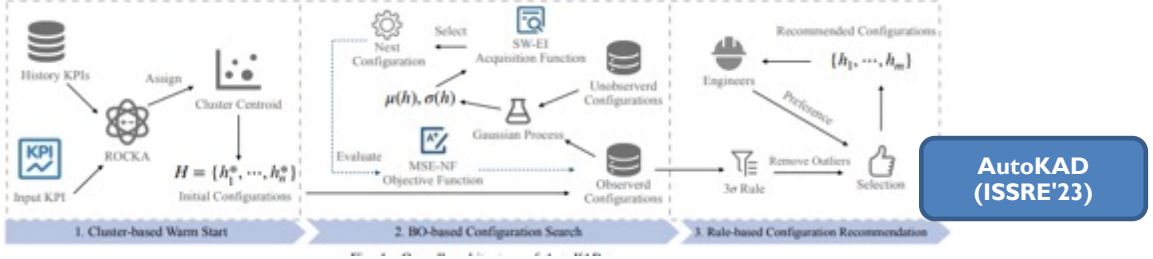
StepWise (ISSRE'18)



PatternMatcher (ISSRE'21)



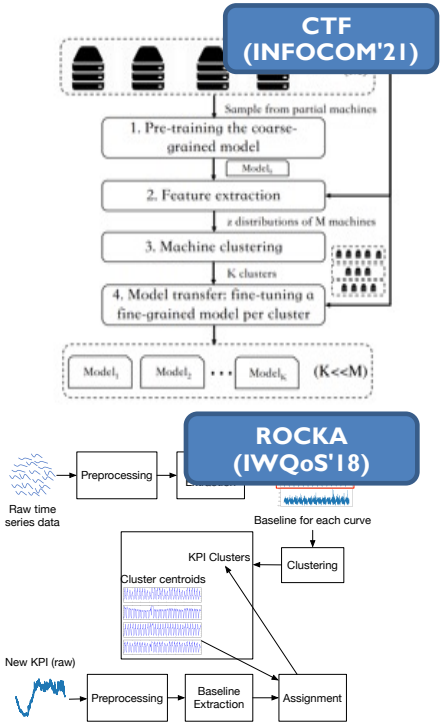
AutoKAD (ISSRE'23)



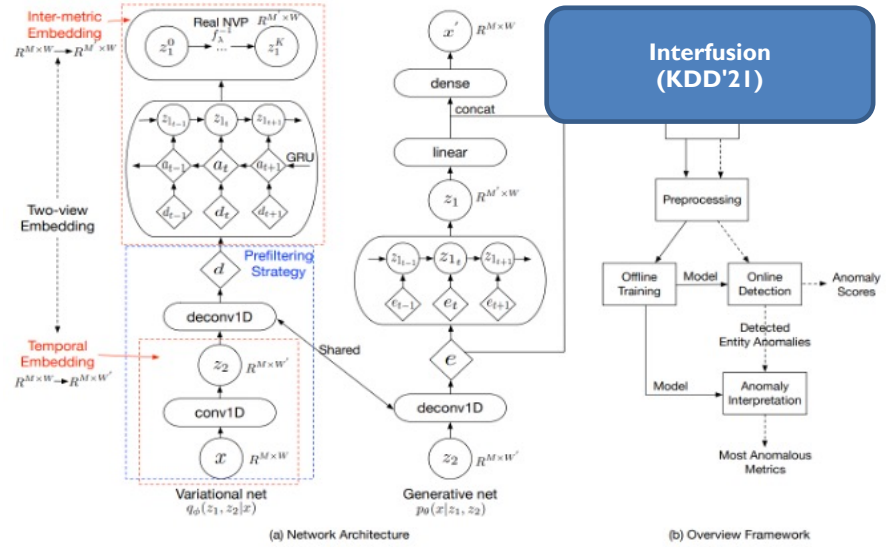
NETMAN'S PAST EFFORTS IN TEMPORAL ANALYSIS IN CYBERSPACE

Focus on temporal analysis

- Mainly focused on anomaly detection
- Explored transfer learning, SFT, utilizing more external information, and Human Feedback

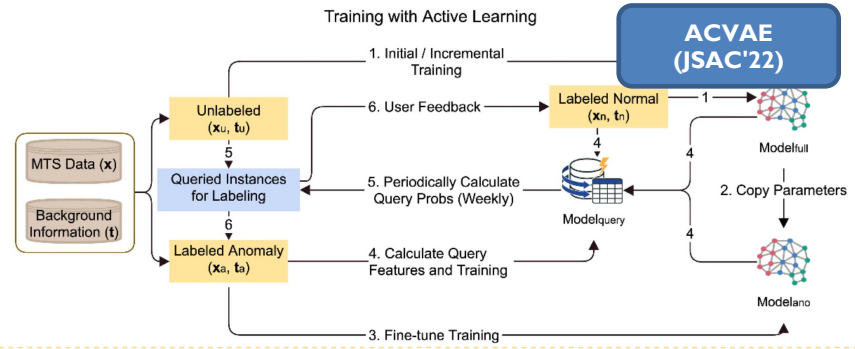


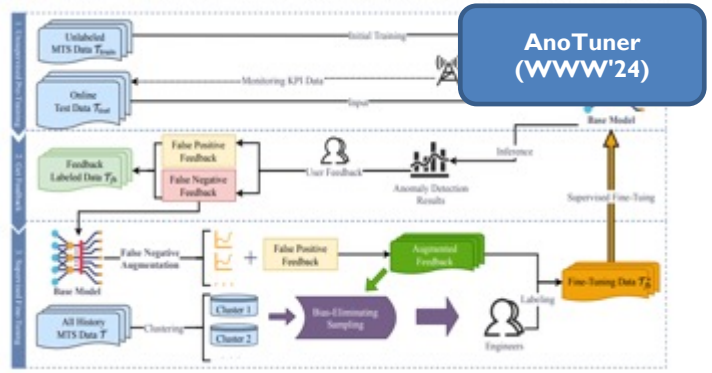
Transfer learning



Explicitly modeling spatial information

Explicitly Modeling Exogeneous information





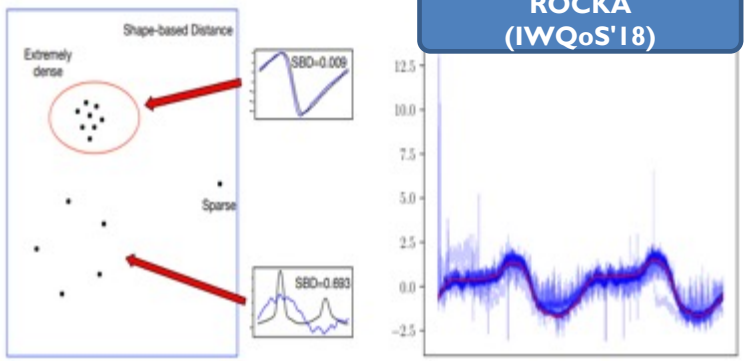
Adding Human Feedback

NETMAN'S PAST EFFORTS IN SPATIAL ANALYSIS IN CYBERSPACE

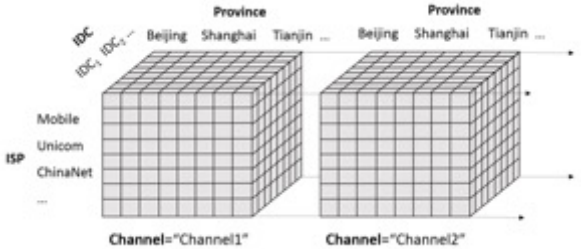
Focus on spatial analysis

- Time series clustering, causal discovery, causal inference
- Cannot be independent of time information

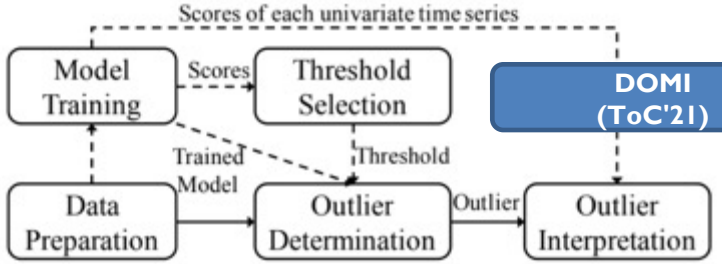
ROCKA (IWQoS'18)



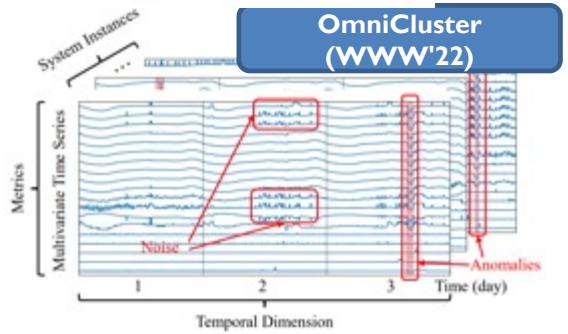
HotSpot (IEEE Acess'18)



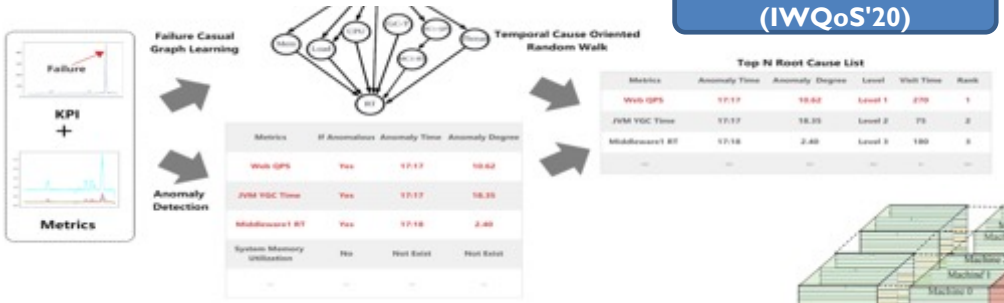
DOMI (ToC'21)



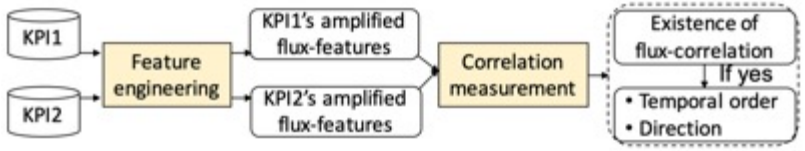
OmniCluster (WWW'22)

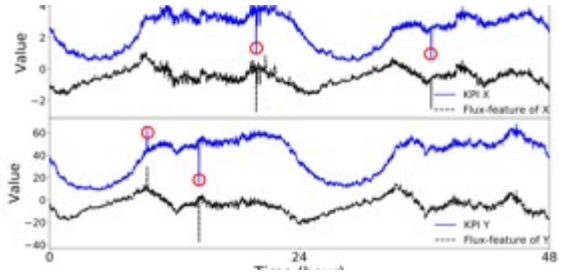


MicroCause (IWQoS'20)

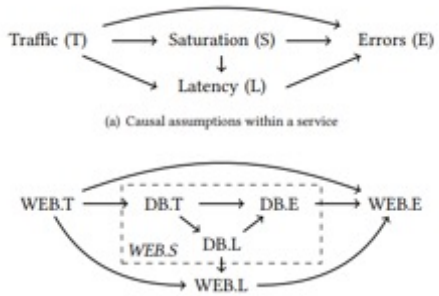


CoFlux (IWQoS'19)

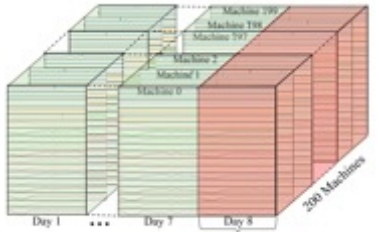


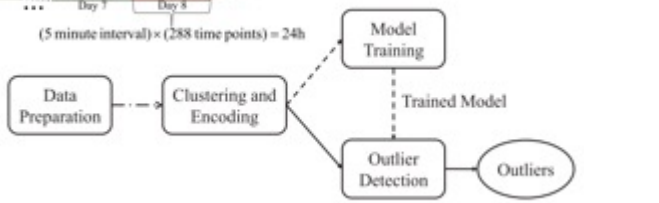


CIRCA (KDD'22)



OutSpot (ToC'23)

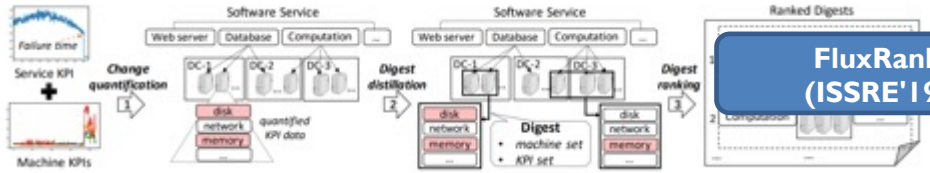




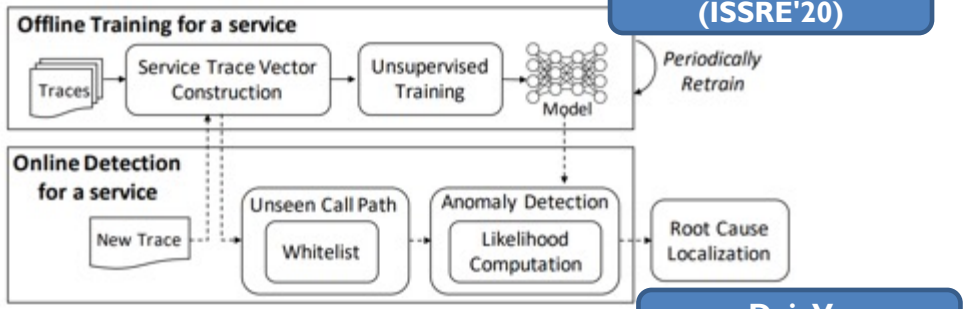
NETMAN'S PAST EFFORTS IN SPATIOTEMPORAL ANALYSIS IN CYBERSPA

Focus on spatiotemporal information

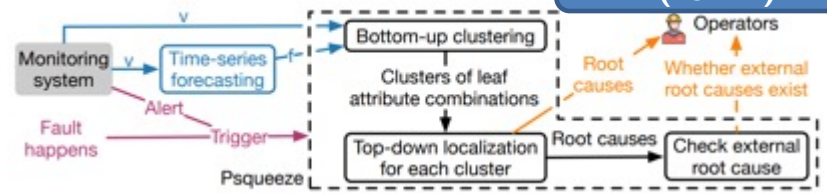
- Mainly focusing on root cause localization and trace anomaly detection
- Use temporal information first, then use spatial information
- Or utilize both temporal and spatial information simultaneously



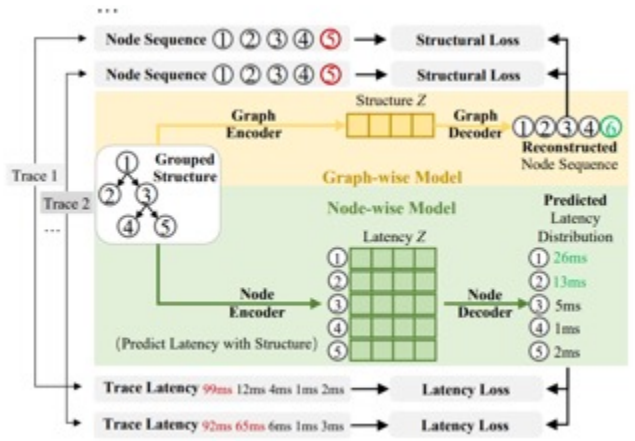
FluxRank (ISSRE'19)



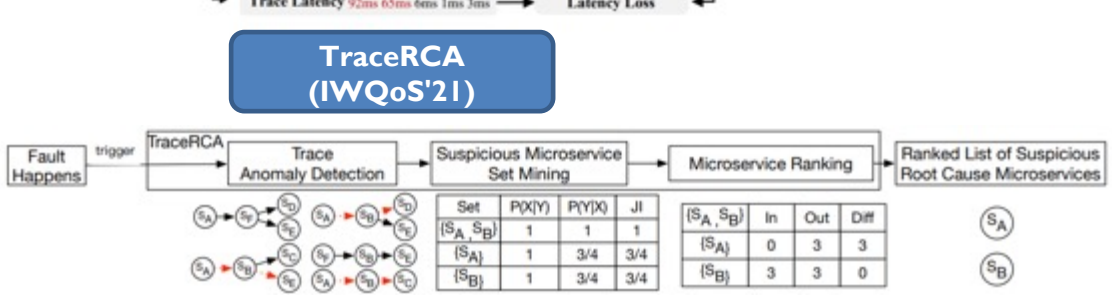
TraceAnomaly (ISSRE'20)



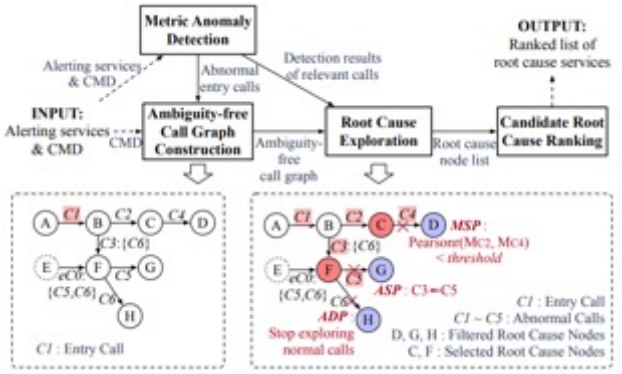
DejaVu (FSE'22)



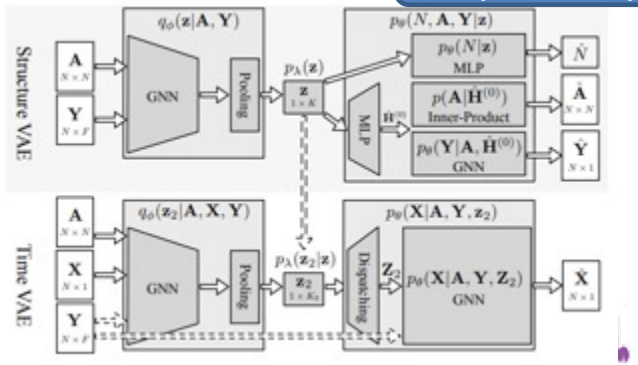
GTrace (FSE'23)



TraceRCA (IWQoS'21)



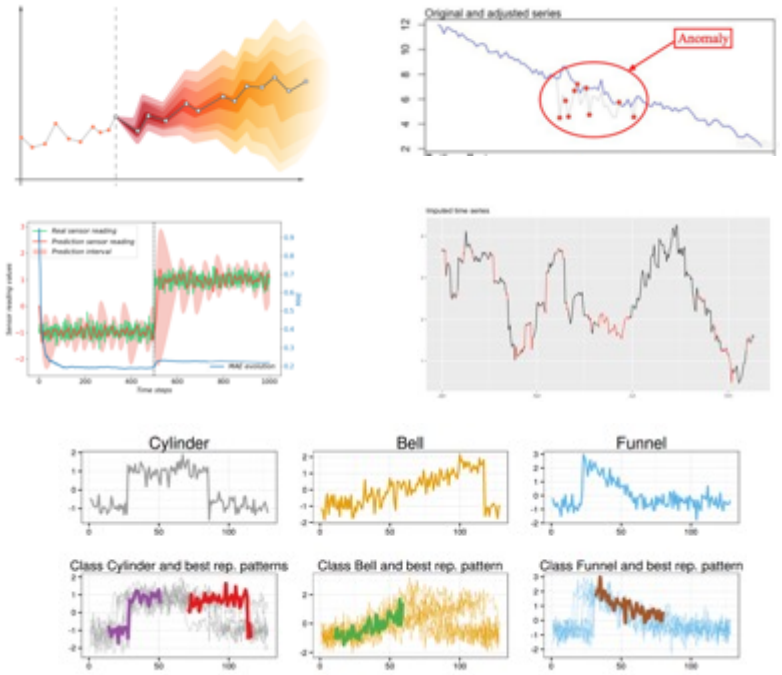
CMDiagnostor (WWW'23)



TraceVAE (WWW'23)

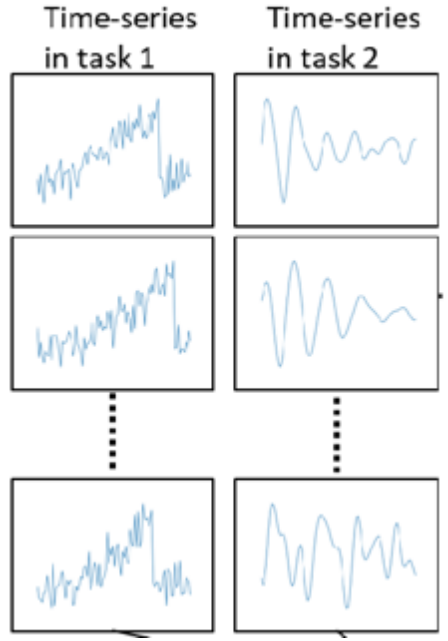
UNRESOLVED ISSUES

Cross-task



Cross-domain

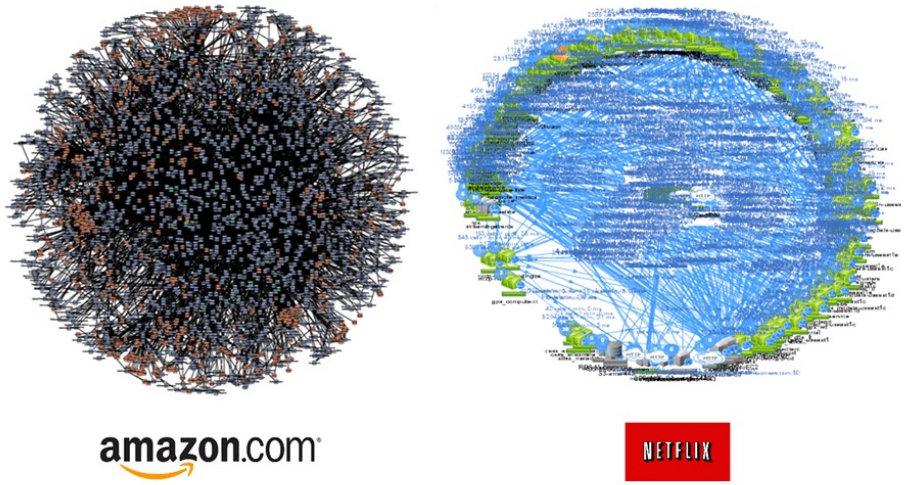
Few-shot zero-shot

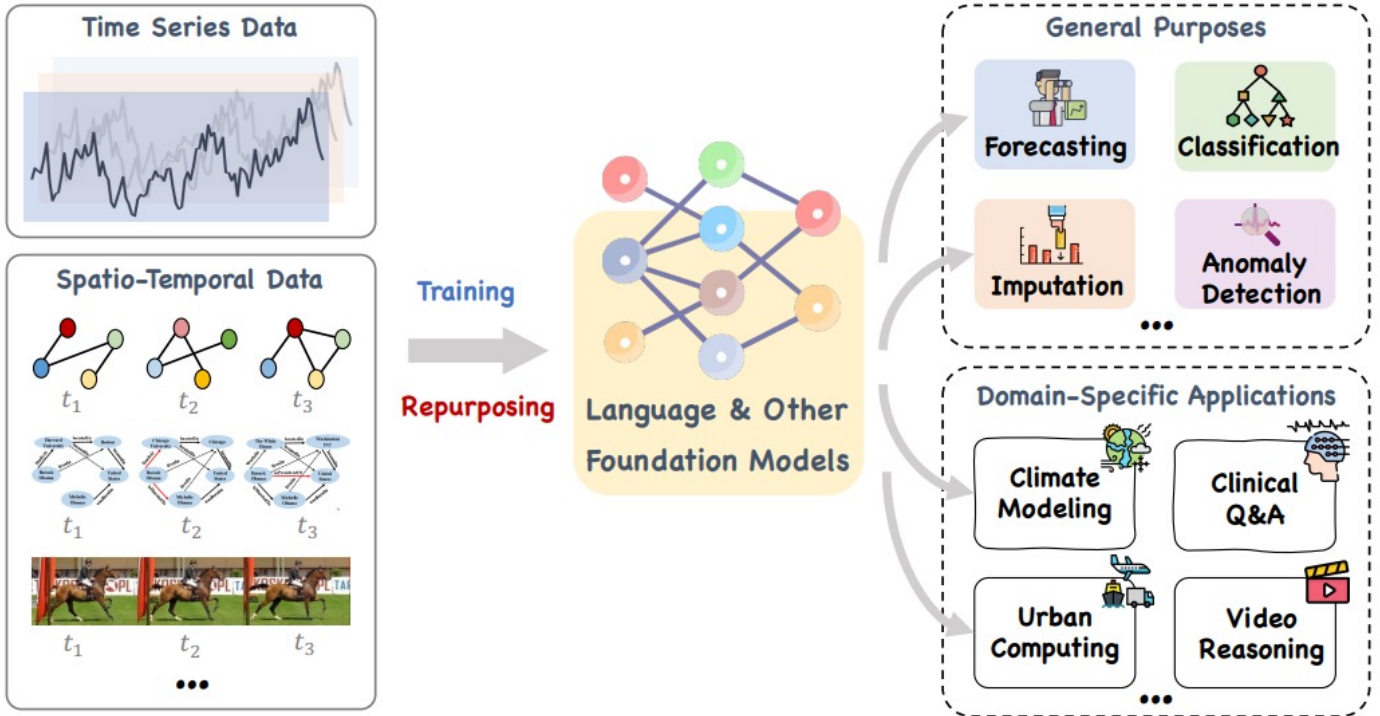


- Difficulty in balancing multiple tasks simultaneously (prediction, anomaly detection, classification, etc.)
- Difficult to consider multiple domains at the same time (AIOps, network security, Internet of Things, industrial Internet, etc.)
- There is still significant room for improvement in performance with few or zero samples
- Lack of universal and effective methods for utilizing external information

THE NECESSITY OF LARGE-SCALE MODELS IN THE SPATIOTEMPORAL DOMAIN

- Complex and diverse data
  - Different fields, different patterns, and different structures
  - At the temporal level
  - At the spatial level
  - External information level
- The tasks are complex and diverse:
  - Forecast
  - Anomaly detection
  - Root cause localization
  - …
- Disadvantages of small models:
  - Universal type
  - Scalability

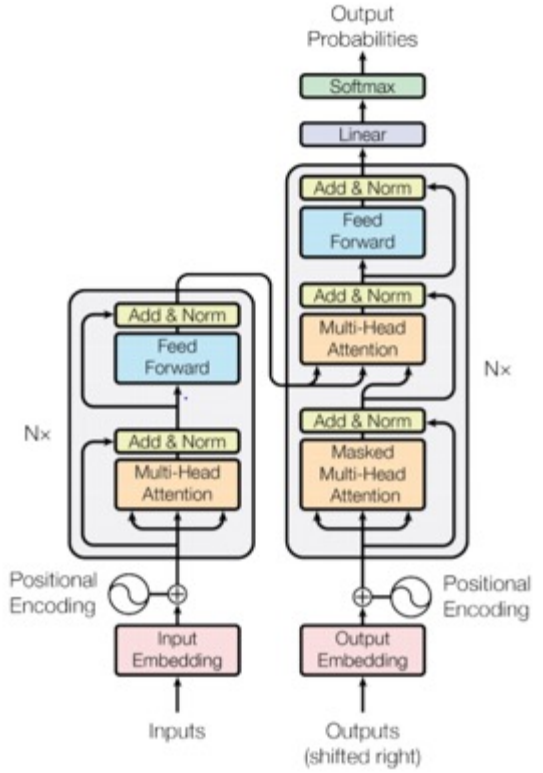


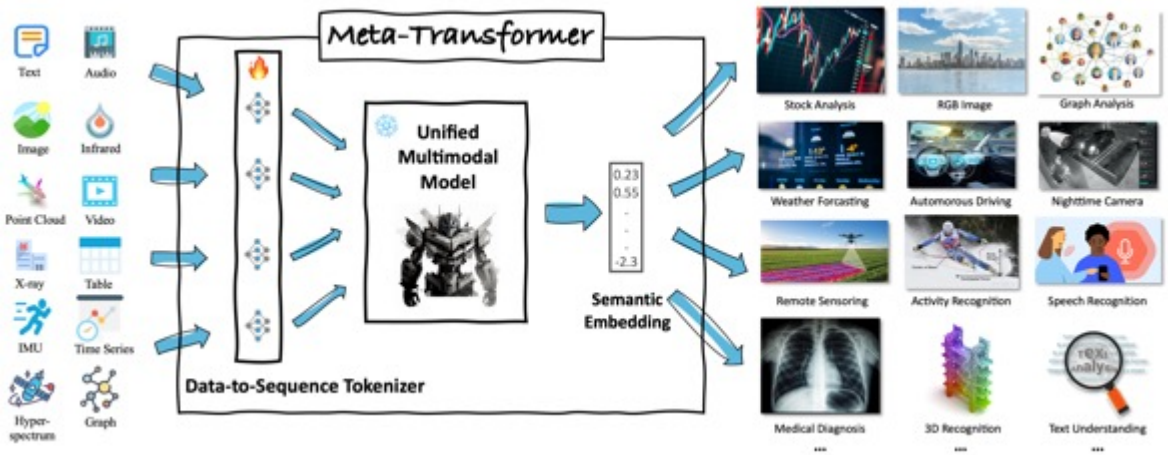


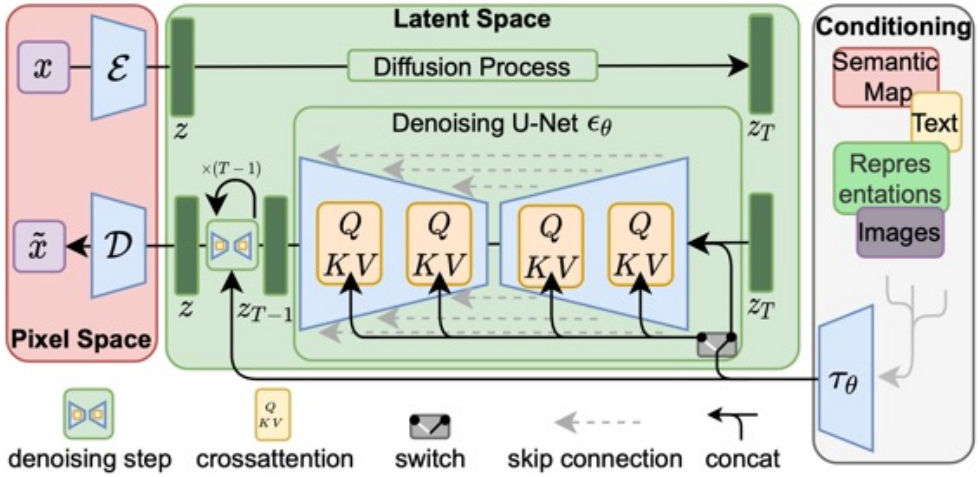
WHAT DOES THE SPATIOTEMPORAL MODEL NEED TO DO?

Powerful Backbone

- A Backbone with strong expression ability
- Can effectively compress and utilize a large amount of spatiotemporal knowledge, understanding ability, and reasoning ability
- Has to enable the scaling law to be effective (concise, capable of large-scale parallelization)
- Is Transformer suitable?
- Is the Diffusion structure effective?
- How to handle modality (Alerts, Logs) data?







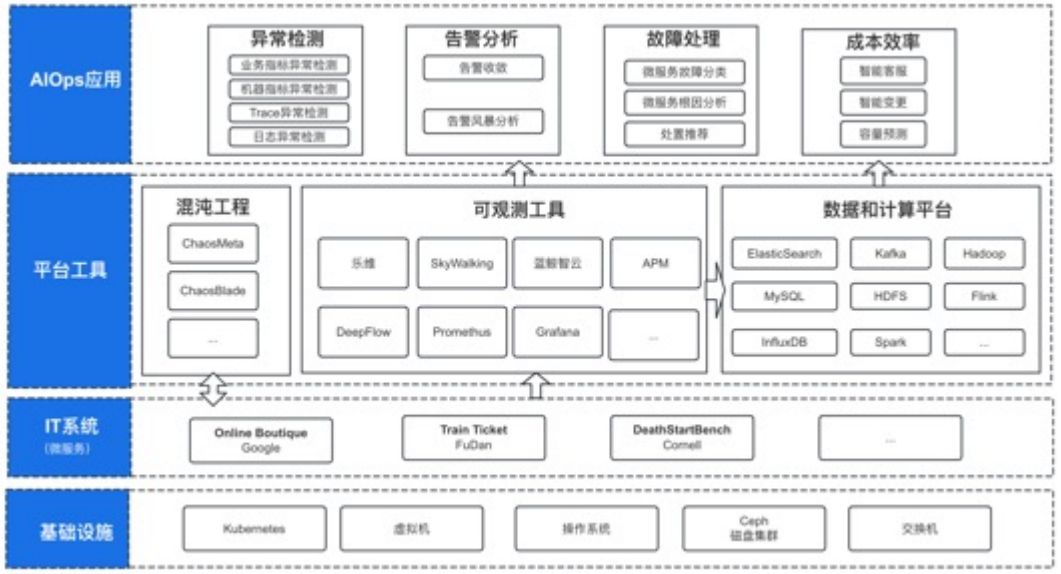
GOALS FOR THE SPATIOTEMPORAL FOUNDATION MODEL

General Pre-training Task

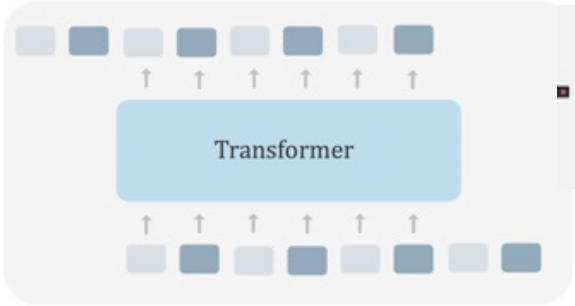
- Universal pre-raining tasks
- Can help the model fully learn knowledge from spatiotemporal data
  - Extract and compress spatiotemporal information comprehensively
  - Learn spatial reasoning skills for simultaneous use of time and space
  - Next Token Prediction, Masked Reconstruction, or something else?
- How to obtain high-quality self-contained data for pre-training?
  - The metrics and graphs have to be relatively complete, supplemented by spatiotemporal data governance

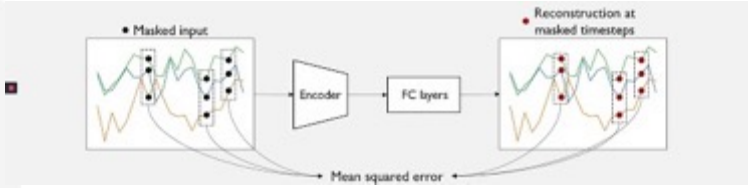
(data normalization, data imputation, multimodal alignment)

  - OpenAIOps Live Benchmark:
    - Data, Small Models, Chaos Engineering: Red Blue Confrontation
  - Full data in the data center?
  - Digital twin system?
  - Emulation system?



OpenAIOps Live Benchmark





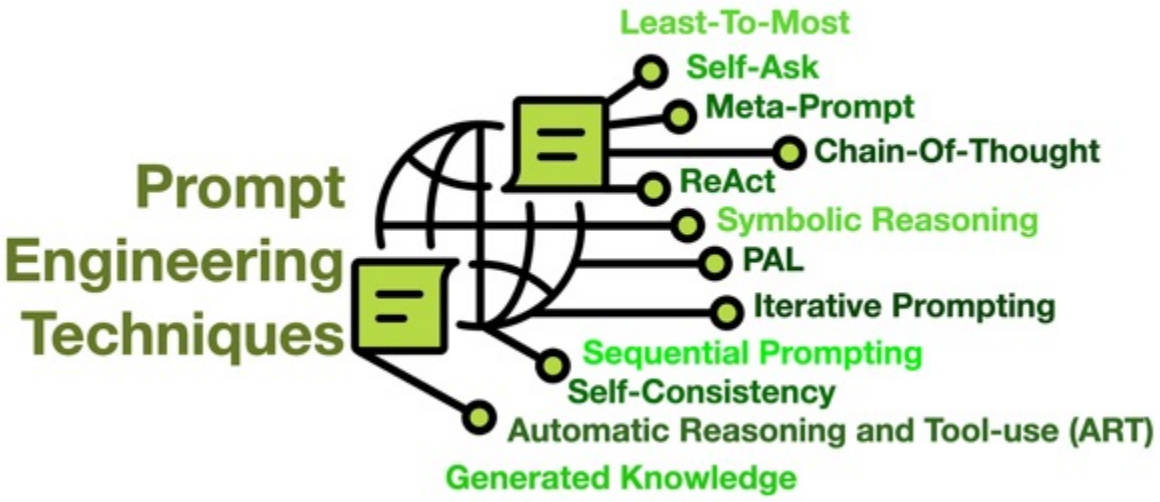
Masked Reconstruction

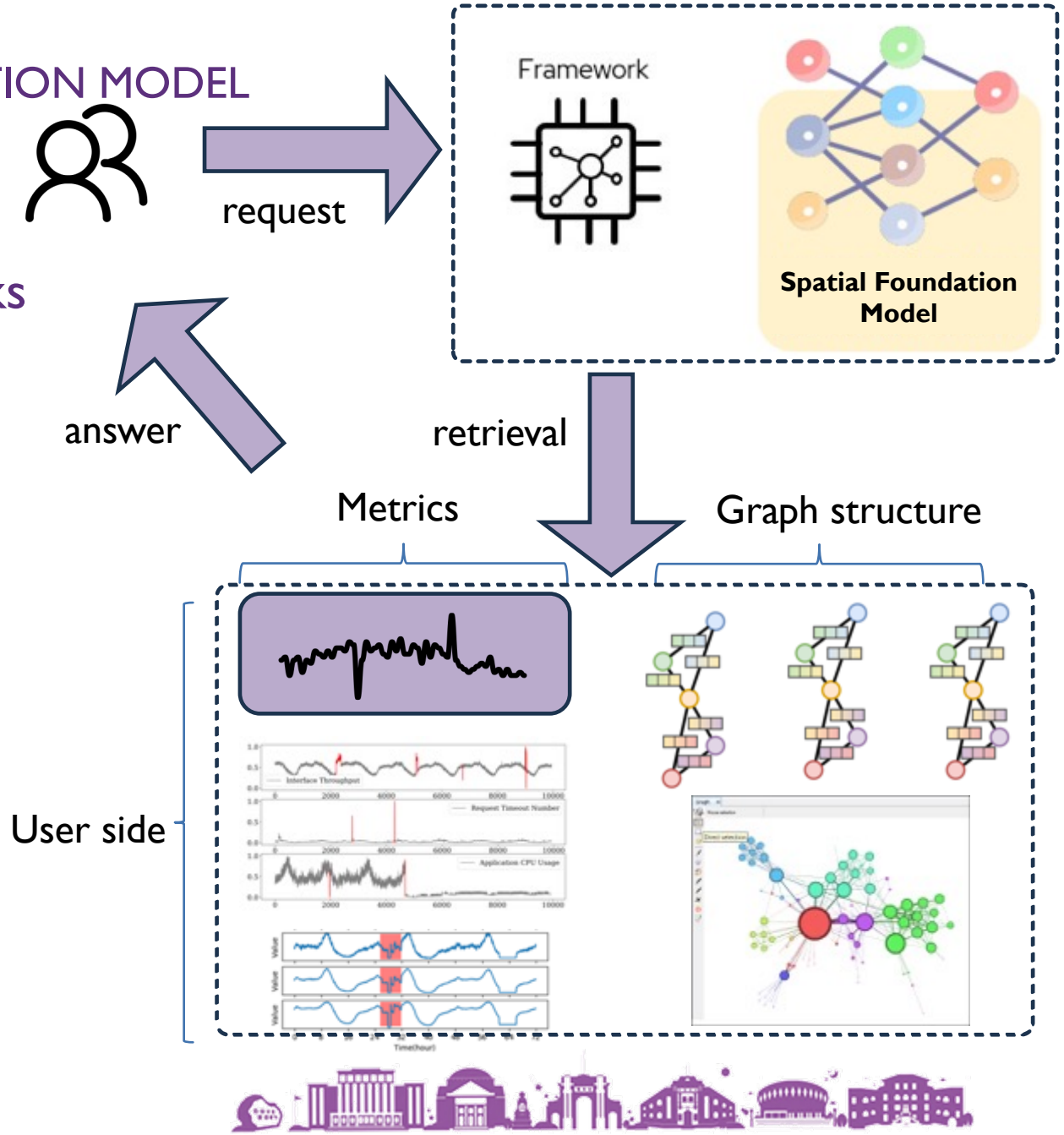
Next Token Prediction

GOALS FOR THE SPATIOTEMPORAL FOUNDATION MODEL

Task Adaptation

- Meet the needs of various downstream tasks
- Quickly adapt to user data
- How to efficiently fine tune?
- How to perform prompt engineering?
- RAG in the spatiotemporal domain?





CONCLUSION

- Everything is interconnected and interdependent in cyberspace
- Need for joint modeling of temporal and spatial information, as graph-structured time series (4-tuple)
- The spatiotemporal model is both necessary and feasible, and the best route is still an open question
- The prospect of the spatiotemporal model is promising, but there is still a long way to go.
- We will explore and advance step by step, eventually arriving at a powerful spatiotemporal foundation model

清華大學
Tsinghua University

THANK YOU!

Q&A