A Knowledge Representation and Information Fusion Framework for Decision Making in Complex Cyber-Physical Systems

PI: Soumik Sarkar, PhD (soumiks@iastate.edu)

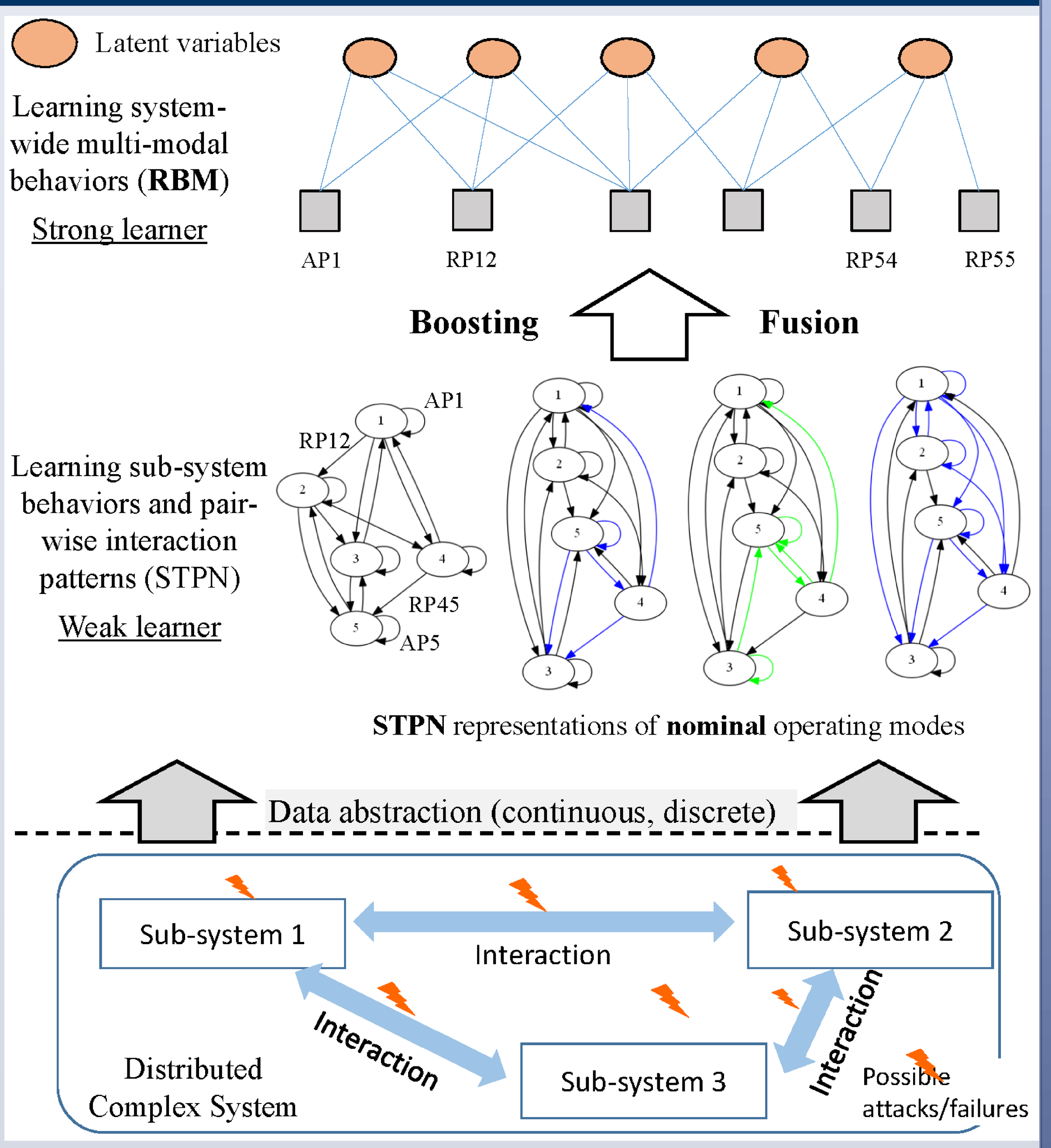
Department of Mechanical Engineering, Iowa State University, Ames, IA

IOWA STATE UNIVERSITY
Department of Mechanical Engineering

Project Objectives

- Develop a **data-driven modeling framework** for CPSs that reliably captures cyber and physical sub-system behaviors as well as their interaction characteristics.
- To address the need of **performance monitoring and fault detection & diagnostics** (FDD) in distributed CPSs (e.g., integrated building), with **cyber attacks and physical anomalies**.
- Challenge: **Inference** and **root cause analysis** in complex CPSs with **multiple (possibly unforeseen) anomalies** at the same time, **system wide impact estimation** in a large interconnected system.

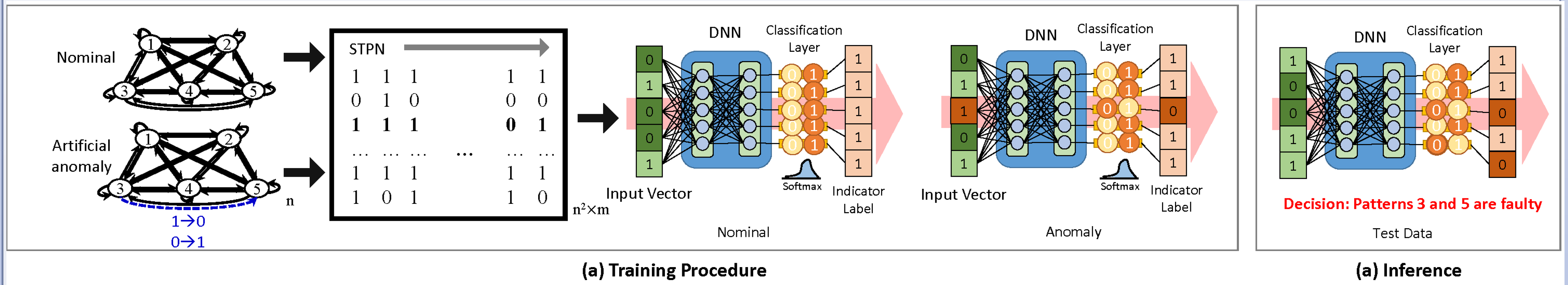
Previous work: Anomaly detection



- A data-driven framework for **system-wide anomaly detection** was proposed, noted as the *STPN+RBM* model, to implement ***unsupervised*** anomaly detection with ***spatiotemporal causal graphical modeling***.
- Validation on synthetic data and real system shows the proposed framework can handle **mixed data types, local and global anomalies**, and capture **multiple nominal modes**.

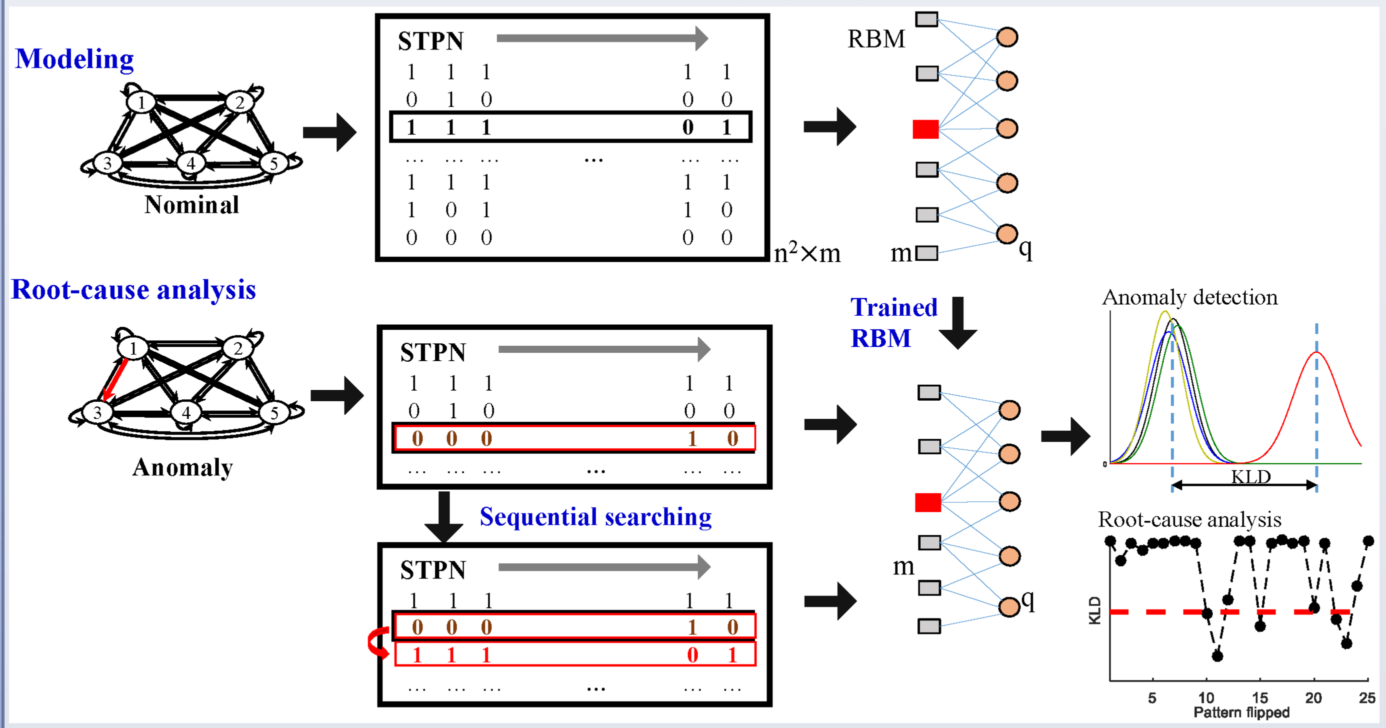
Root-cause analysis of complex CPSs via spatiotemporal causal graphical modeling

Artificial anomaly association (A^3)



(a) Training Procedure (a) Inference

Sequential state switching (S^3)



Accuracy metrics

$$\alpha_1 = \frac{\sum_{j=1}^{n^2}\sum_{i=1}^{m}\chi_1(T_{ij}=P_{ij})}{mn^2}$$

where T^{ij} denotes the ground truth state (nominal/ anomalous) of the j^{th} pattern of the i^{th} test sample. P^{ij} is the corresponding predicted state.

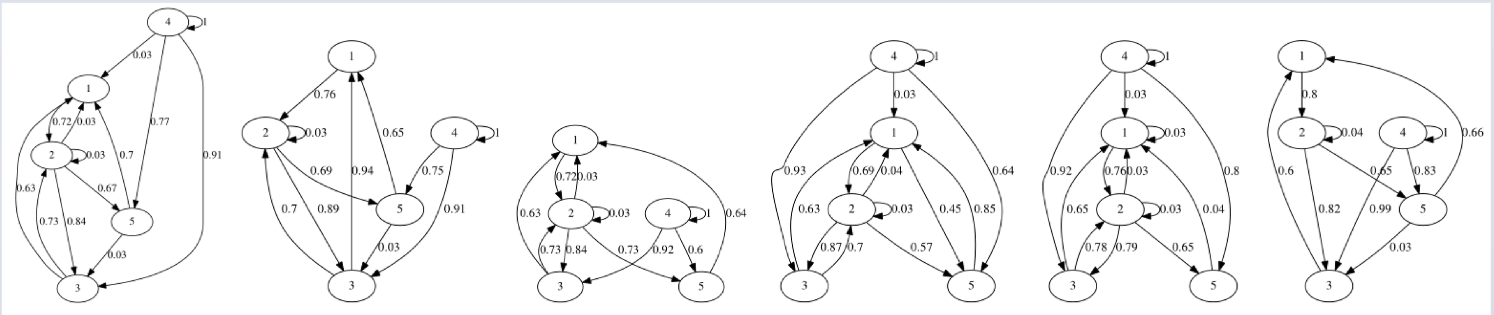
$$Recall = \frac{TP}{TP+FN} \quad Precision = \frac{TP}{TP+FP}$$

$$F-measure = \frac{2}{1/precision + 1/recall}$$

where *TP* is true positive rate, *FN* is false negative rate, and *FP* is false positive rate

Case study with synthetic data

Learning of multiple nominal modes and root-cause analysis of failed patterns



(a) Mode 1 (b) Mode 2 (c) Mode 3 (d) Mode 4 (e) Mode 5 (f) Mode 6

Table 1: Root-cause analysis results in S^3 method and A^3 method with synthetic data.

Approach	Training samples	Testing samples	Accuracy α_1 (%)	Recall (%)	Precision (%)	F-measure (%)
S^3	11,400	57,000	97.04	99.40	97.10	98.24
A^3	296,400	57,000	98.66	90.46	95.95	93.12

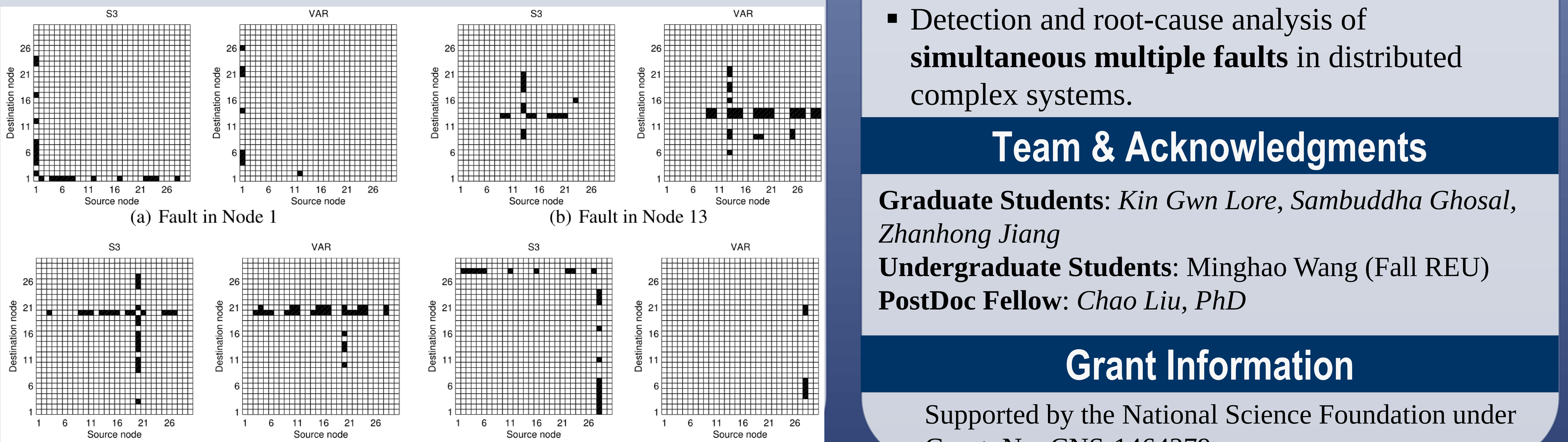
- **Dataset**: 5-node graphical model, 6 nominal operation modes, anomalies in 30 cases including 5 in one failed pattern, 10 in two failed patterns, 10 in three failed patterns, and 5 in four failed patterns.
- **Methods**: Artificial anomaly association (A^3) and Sequential state switching (S^3).

Root-cause analysis of failed node Scalability analysis

- Dataset: 5-node and 30-node systems, 5 and 30 anomalies via simulating every node failure respectively.
- Methods: Sequential state switching (S^3) and Vector autoregressive (VAR) model.

Table 2: Comparison of root-cause analysis results with S^3 and VAR.

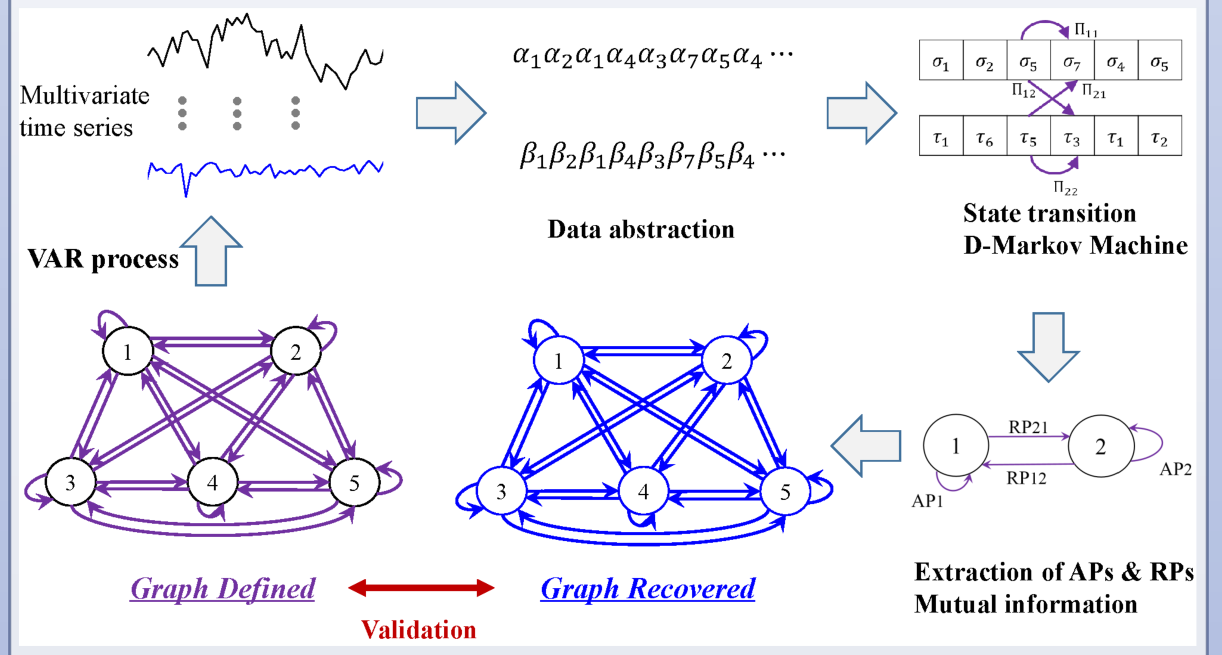
Approach	Dataset 2 (5 nodes)			Dataset 3 (30 nodes)		
	$\lvert\{\Lambda^{ano}\}\rvert$	$\lvert\{\Lambda^{\epsilon}\}\rvert$	ϵ (%)	$\lvert\{\Lambda^{ano}\}\rvert$	$\lvert\{\Lambda^{\epsilon}\}\rvert$	ϵ (%)
S^3	13	2	15.38	653	18	2.76
VAR	20	4	20.00	521	113	21.69

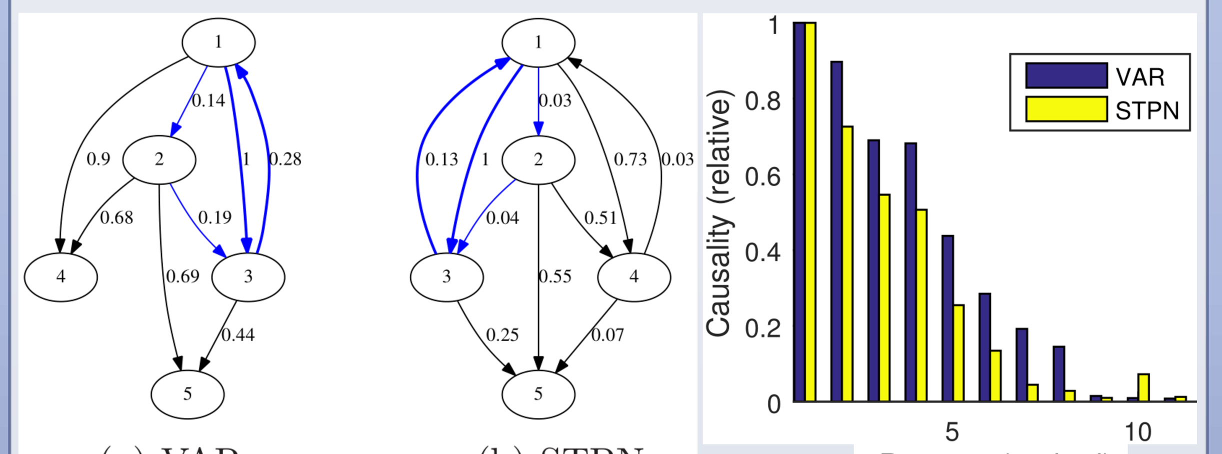


(a) Fault in Node 1 (b) Fault in Node 13 (c) Fault in Node 20 (d) Fault in Node 28

STPN for recovering graphical models

- To validate the efficacy of STPN in interpreting causality in graphical models, case studies are carried out and compared with VAR model.





(a) VAR (b) STPN

Conclusions & Future Work

- Sequential state switching (S^3) and artificial anomaly association (A^3)– are proposed for root-cause analysis in complex cyber-physical systems.
- With synthetic data, proposed approaches are validated and showed high accuracy in finding failed patterns and diagnose for anomalous node.

Further works will pursue the following:

- **Inference** approach in node failure including **single node and multiple nodes**
- Detection and root-cause analysis of **simultaneous multiple faults** in distributed complex systems.

Team & Acknowledgments

Graduate Students: *Kin Gwn Lore, Sambuddha Ghosal, Zhanhong Jiang*

Undergraduate Students: Minghao Wang (Fall REU)

PostDoc Fellow: *Chao Liu, PhD*

Grant Information

Supported by the National Science Foundation under Grant No. CNS-1464279.