SAND2022-2543C

Sandia National Laboratories

Deep Learning Architecture for Processing Cyber-Physical Data in the Electric Grid

Daniel Calzada, Shamina Hossain-McKenzie, Zeyu Mao

Presenter: Zeyu Mao

March 2022

Power and Energy Conference at Illinois

U.S. DEPARTMENT OF ENERGY

NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Outline

Background and problem statement

Data collection

Model architecture

Results

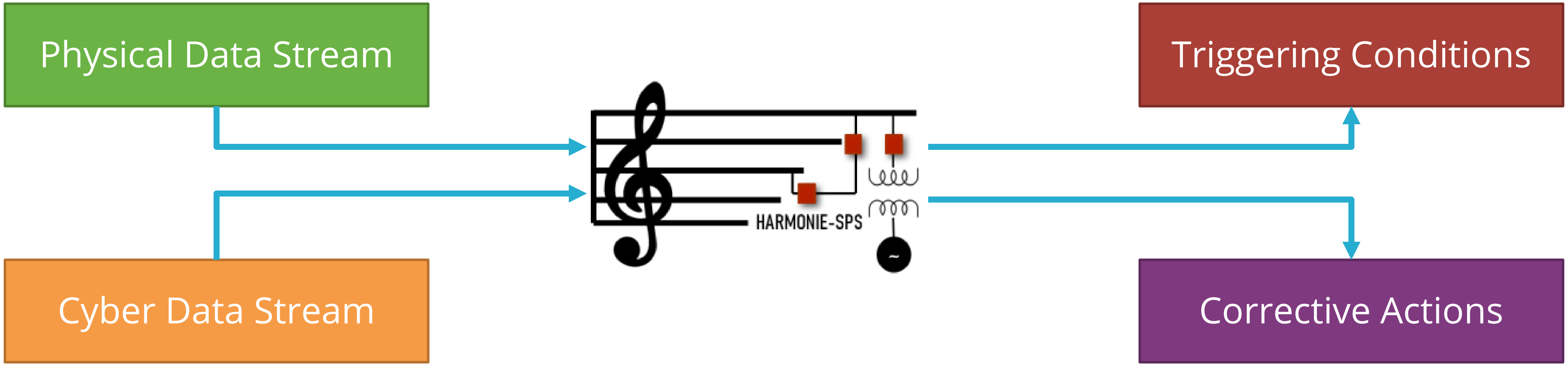
Conclusions

Background

Harmonized Automatic Relay Mitigation of Nefarious Intentional Events (HARMONIE)

Towards meeting the need for an SPS to adapt to quickly unpredictable events

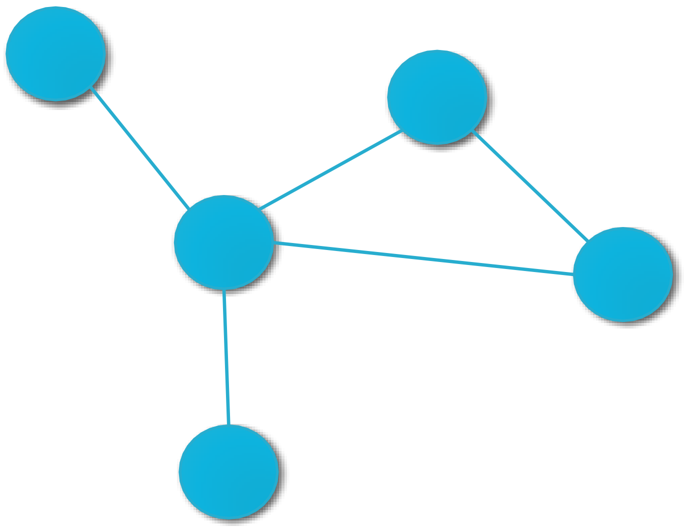
This paper: We are investigating the feasibility of processing physical and network data jointly to identify disturbances and gain insights to eventually deploy a corrective action



Background

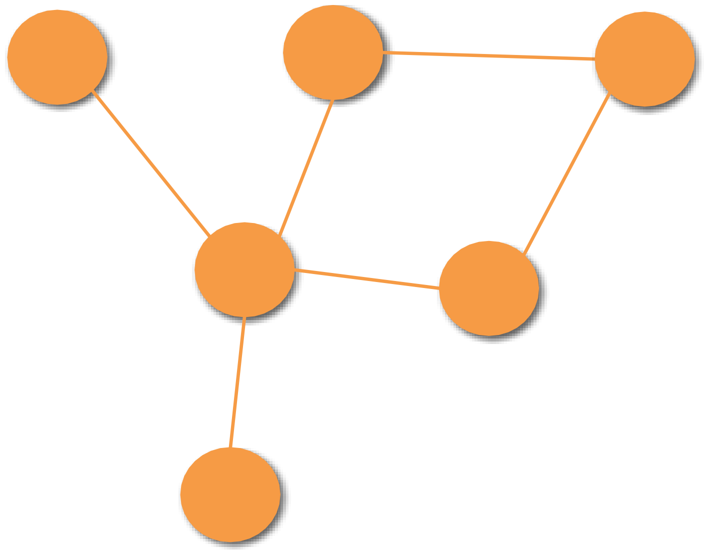
All **cyber** data can be passed through a graph convolutional neural network (GNN) to perform supervised learning to identify **unsafe** states

- Vertices are devices on the network
- Edges are network traffic flows



All **physical** data can be passed through a separate graph convolutional neural network to perform supervised learning to identify **unstable** states

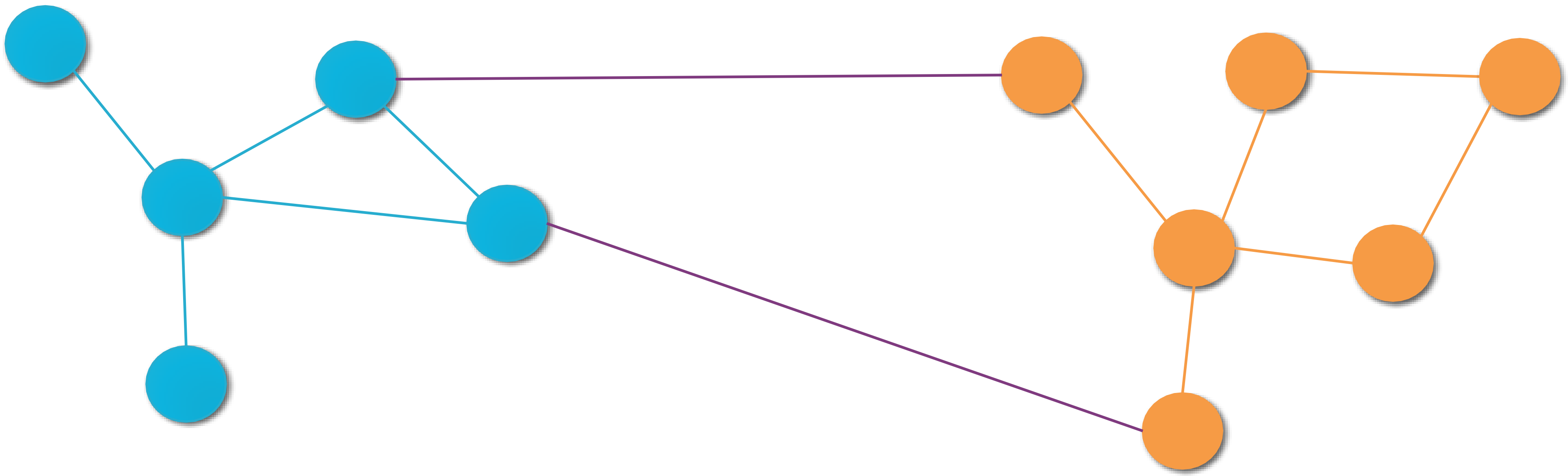
- Vertices are PMUs, relays, HMIs, out-of-band sensors, buses, etc.
- Edges are information flows



Background

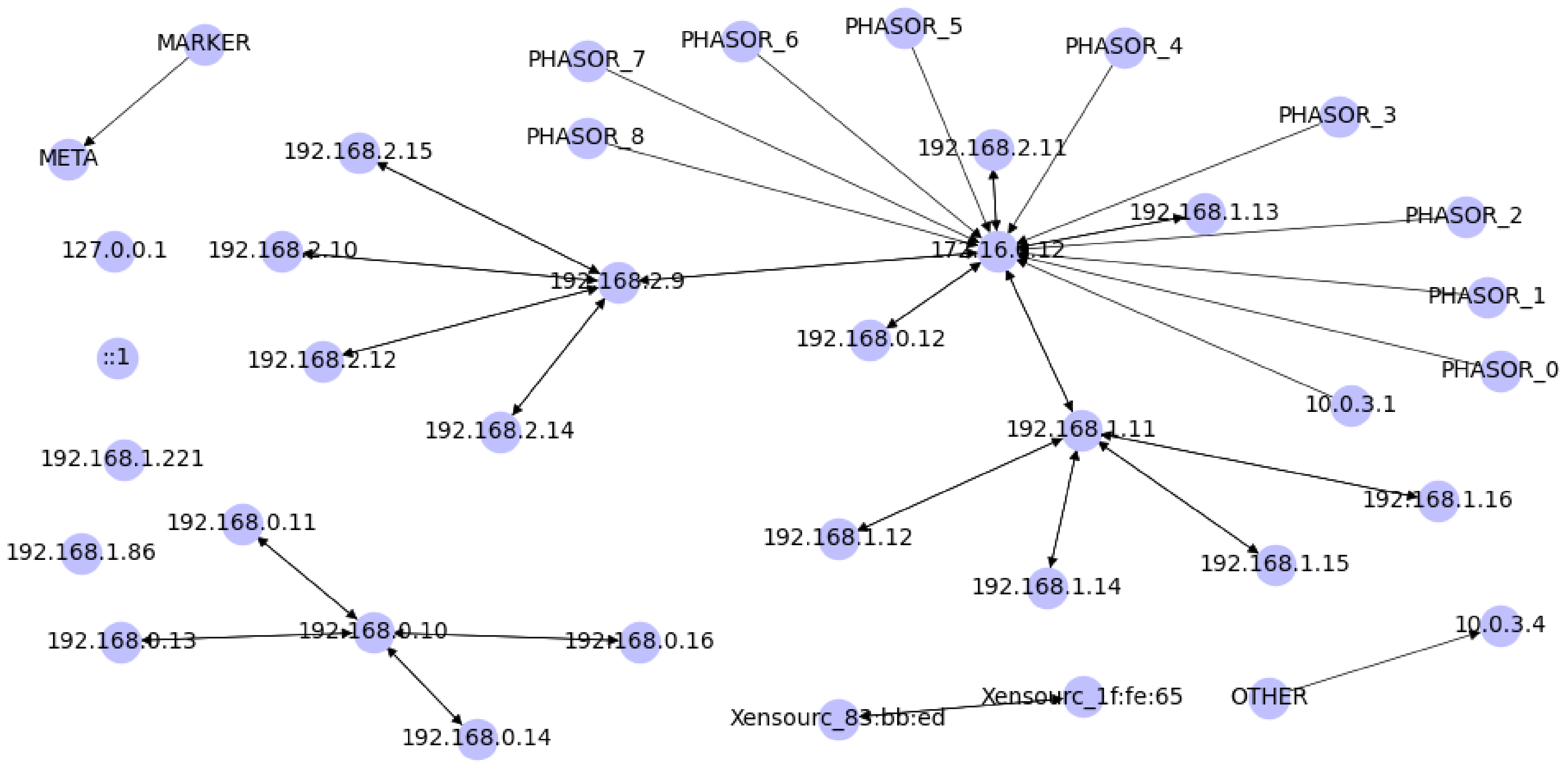
The cyber and physical graphs contain overlapping nodes

Our deep learning architecture will process these graphs jointly



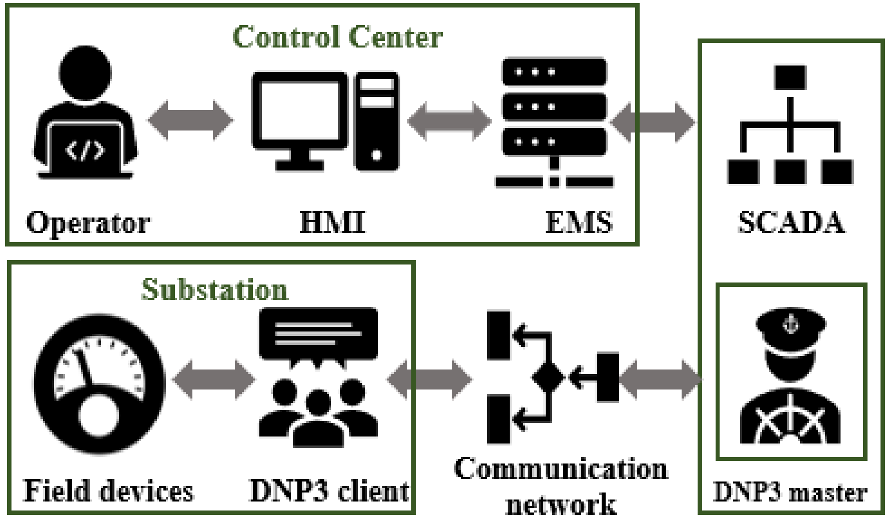
Background

Network traffic and phasor measurements are combined into the same graph

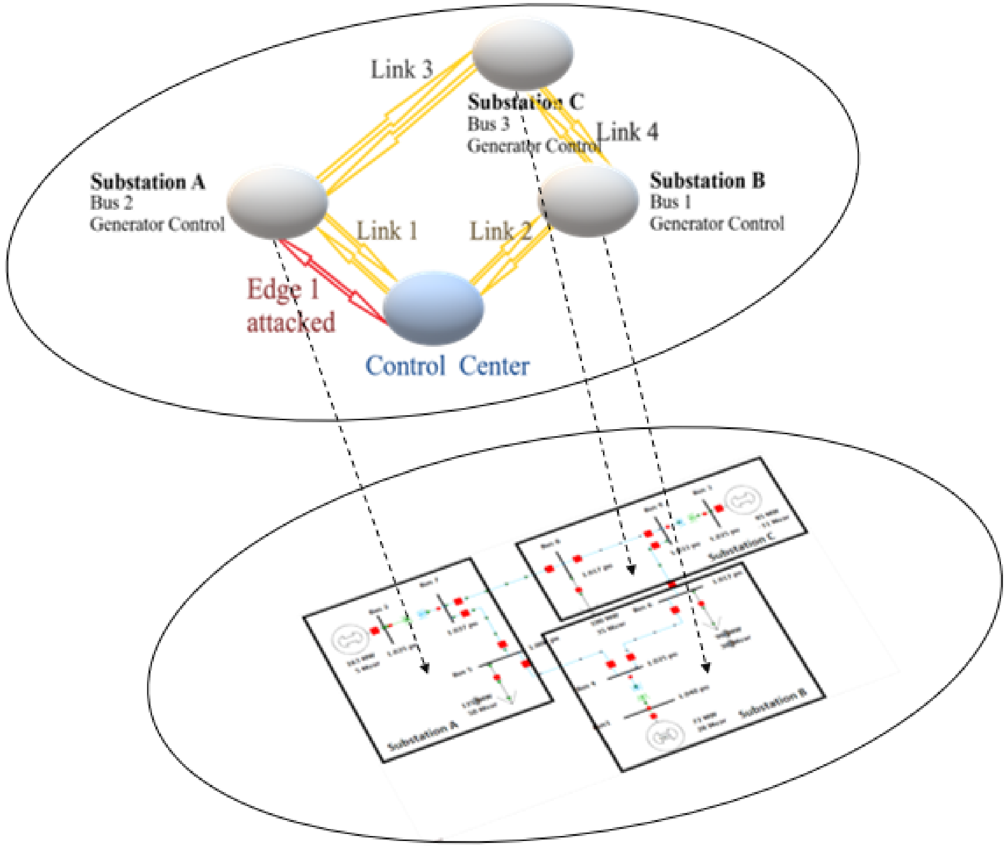


Data Collection

Using the Western Systems Coordinating Council (WSCC) 9-bus power system



The diagram of hierarchies for the simulated grid



The cyber-physical model for the WSCC 9-bus case

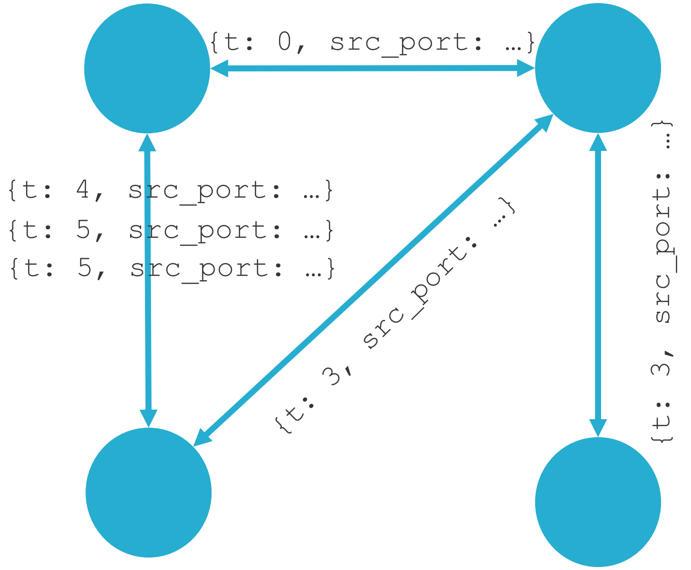
Data Collection

- 2 minute captures broken into 30-second sliding windows
  - Disturbances, when present, happen at around 1 minute
  - Models will be identifying if a disturbance occurs within a 30-second sliding window
- 50 total scenarios
  - Normal operations
  - Denial of Service (DoS) attacks
  - False command injection (FCI) attacks
  - Time delay (TD) attacks
  - Single-line-to-ground (SLG) faults
- Cyber and physical data are interleaved and represented as JSON

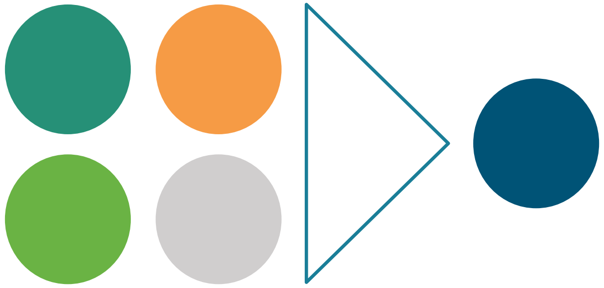
Model Architecture: Graph Neural Network

To process spatially-structured data (particularly useful for network traffic), we employ a Graph Convolutional Neural Network (GCNN/GNN) [1]

Neural message passing: Each vertex starts with a learned state, states are adjusted using the edges between the vertices



After n iterations of message passing, each node and edge has its own vector



A weighted mean of node vectors encodes structural information

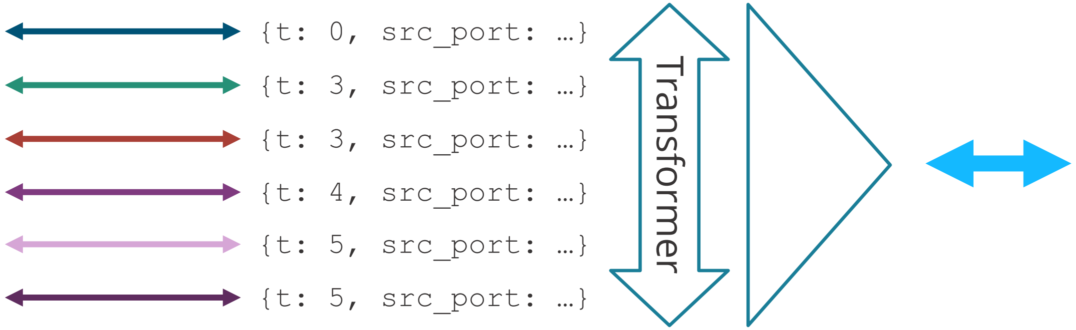
Model Architecture: Transformer

To process temporal data (particularly useful for physical data), we employ a Transformer model [2]

Popular in natural language processing, can be applied to physical systems [3]

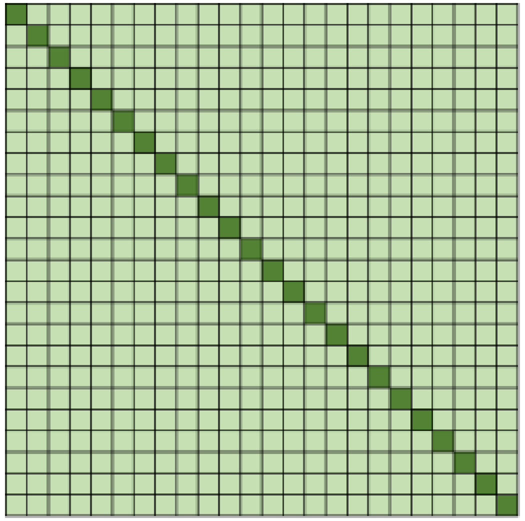
Repeatedly transforms a sequence to another sequence

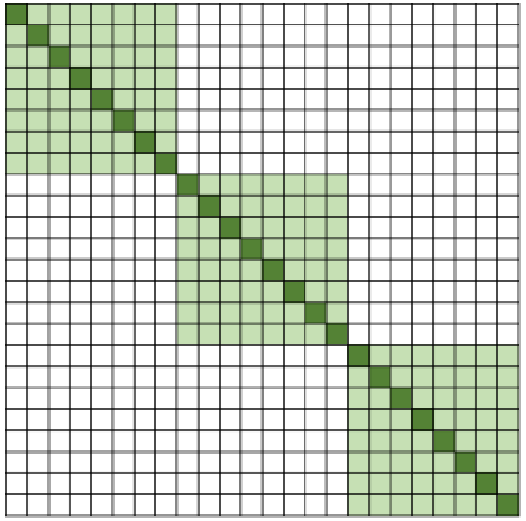
A weighted mean of edge vectors encodes temporal information

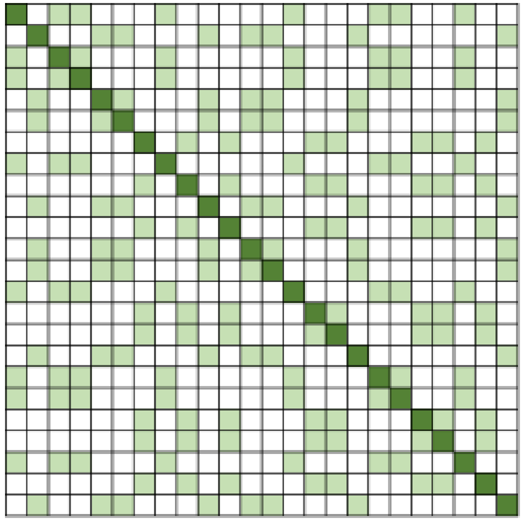


Model Architecture: Random Windowed Transformer

- By default, Transformers use $O(n^2)$ memory
- Splitting the sequence into fixed-size windows reduces memory complexity
  - Loses long-term dependencies
- We try splitting the data into random fixed-size windows to maintain long term dependencies


Dense attention matrix


Windowed attention matrix


Random-windowed attention matrix

Model Architecture: Rationales

Goal: Move towards identifying the cause of a disturbance

We use Rationale Neural Networks [4] to learn to mask irrelevant timesteps (packets or sensor measurements)

The step before the GNN/Transformer in the model architecture

Can later be interpreted as the “rationale” for predicting the existence of a disturbance

Results

- For each architecture:
  - Split the data into 5 folds and trained 20 models, one for each combination of validation and test folds
  - Generated and aggregated predictions for each test fold, then aggregated the results below

Architecture	Rationale %	Cyber Disturbance Detection		Physical Disturbance Detection	
		MCC	AUC	MCC	AUC
Traditional Transformer	39.3%	0.77	**0.98**	0.57	0.85
Random-windowed Transformer	46.1%	0.70	0.95	**0.63**	**0.87**
GNN	48.0%	**0.85**	0.96	0.18	0.68
GNN + Transformer	N/A	0.74	0.97	0.30	0.77

Conclusions

- The GNN is most effective with cyber disturbances, the Transformers are most effective with physical disturbances
  - Logical given the relative strengths of each of these architectures
- Cyber disturbances are easier than physical disturbances for the model to detect
  - Understandable since cyber disturbances are often a single or multiple packets
- Combining the GNN and Transformer did not outperform either independently as expected
- The Rationale Neural Network component kept ~40-50% of the edges
  - Whether an edge is kept or masked is largely determined by whether the edge is a packet or phasor measurement
  - Good starting point, need more downsampling to be useful

Thank you!

Questions?

References

[1] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, “The graph neural network model,” IEEE transactions on neural networks, vol. 20, no. 1, pp. 61–80, 2008.

[2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” in Advances in neural information processing systems, 2017, pp. 5998–600

[3] N. Geneva and N. Zabaras, “Transformers for modeling physical systems,” arXiv preprint arXiv:2010.03957, 2020.

[4] T. Lei, R. Barzilay, and T. Jaakkola, “Rationalizing neural predictions,” arXiv preprint arXiv:1606.04155, 2016.