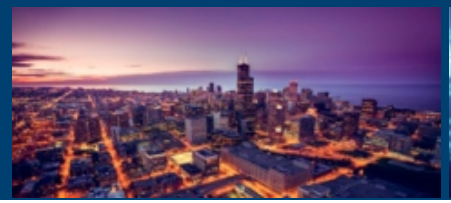
SAND2021-13181C

Sandia National Laboratories

RESILIENT ENERGY SYSTEMS

Mission Campaign



Adaptive, Cyber-Physical Special Protection Schemes to Defend the Electric Grid Against Predictable and Unpredictable Disturbances

Shamina Hossain-McKenzie (PI), Daniel Calzada, Chris Goes, Nicholas Jacobs, Adam Summers, Katherine Davis, Hanyue Li, Zeyu Mao, Thomas Overbye, and Komal Shetye

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.

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.

Problem Statement

Goal of the project

- **HARMONIE Special Protection Scheme** (SPS) is a methodology to process cyber-physical data and provide effective, automated responses to defend against grid disturbances

1) Defend against unpredictable disturbances that do not fit predefined abnormal conditions,

2) process digital relay measurements and incorporate out-of-band (OOB) data for increased situational awareness, and

3) proactively respond to compromises by deploying cyber-physical corrective actions to reduce/eliminate system impact

Impact

- Defend and increase resilience of our nation's critical infrastructure
- HARMONIE-SPS would enable adaptive, fast, and proactive response to both predictable and unpredictable cyber-physical disturbances, reduce/eliminate cascading impact

Proposed Solution

Traditional SPSs

- SPSs are leveraged by utilities to maintain stability, acceptable voltages, and loading limits during disturbances in the electric grid
  - Detect predefined abnormal conditions and deploy predefined corrective actions; operate in playbook manner
  - Can take actions beyond the isolation of a fault and include changes to demand, generation, and system configuration
- **Traditional SPSs are unable to defend against unpredictable disturbances**
  - Resilience and volatile disturbances such as EMPs, extreme weather, and malicious events threatening national security must be considered; cyber-attacks targeting grid operations are increasing in frequency and intensity

A next-generation SPS with the following attributes is needed:

1) A SPS that can adapt to unpredictable events and effectively respond to limit/eliminate the disruption quickly;

2) A SPS that is cyber-physical in analyzing collected data and taking response actions;

3) A SPS that extends the use of protective relays from fault isolation to also adaptively learning system conditions, preventing cyber-attack propagation, and taking proactive actions to prevent compromise within the relay set itself

Return: Technical Review

Proposed Solution

HARMONIE-SPS is a strategy to respond to cyber-physical grid disturbances, both predictable and unpredictable, that can learn system conditions using relay measurements and OOB data, detect abnormal events, and deploy proactive response

- HARMONIE-SPS will be validated using high-fidelity, cyber-physical testing using both virtual and hardware relays within dynamic transmission system models (e.g., synthetic Texas 2000-bus system)

Deployable cyber-physical SPS

Coordinates relays to prioritize selectivity, speed, and/or security based on ML algorithms and deploys cyber-physical corrective controls

ML algorithms process a combination of relay physical measurements, relay host data, and OOB data to classify actual system conditions, adapting to any disturbances

Next-generation relay voting scheme

Considers inter-relay relationships and OOB data, including both full and reduced-order versions, to provide confidence in relay actions

Emulytics™ testing

Using RTDS using HIL and simulated relays within dynamic grid models paired with realistic communication infrastructure; the environment can benefit various existing and future projects.

Emulytics™ Testing of HARMONIE-SPS

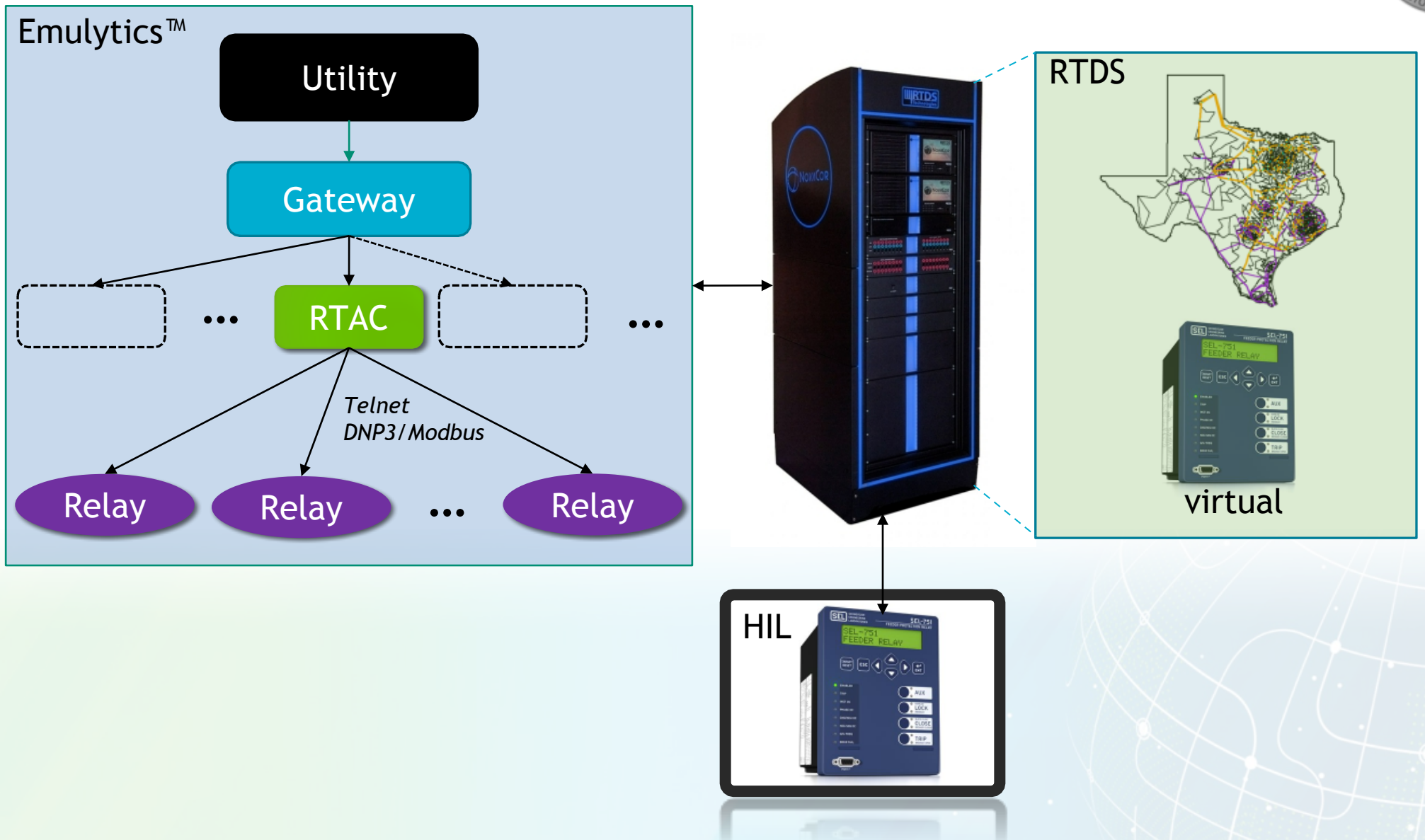
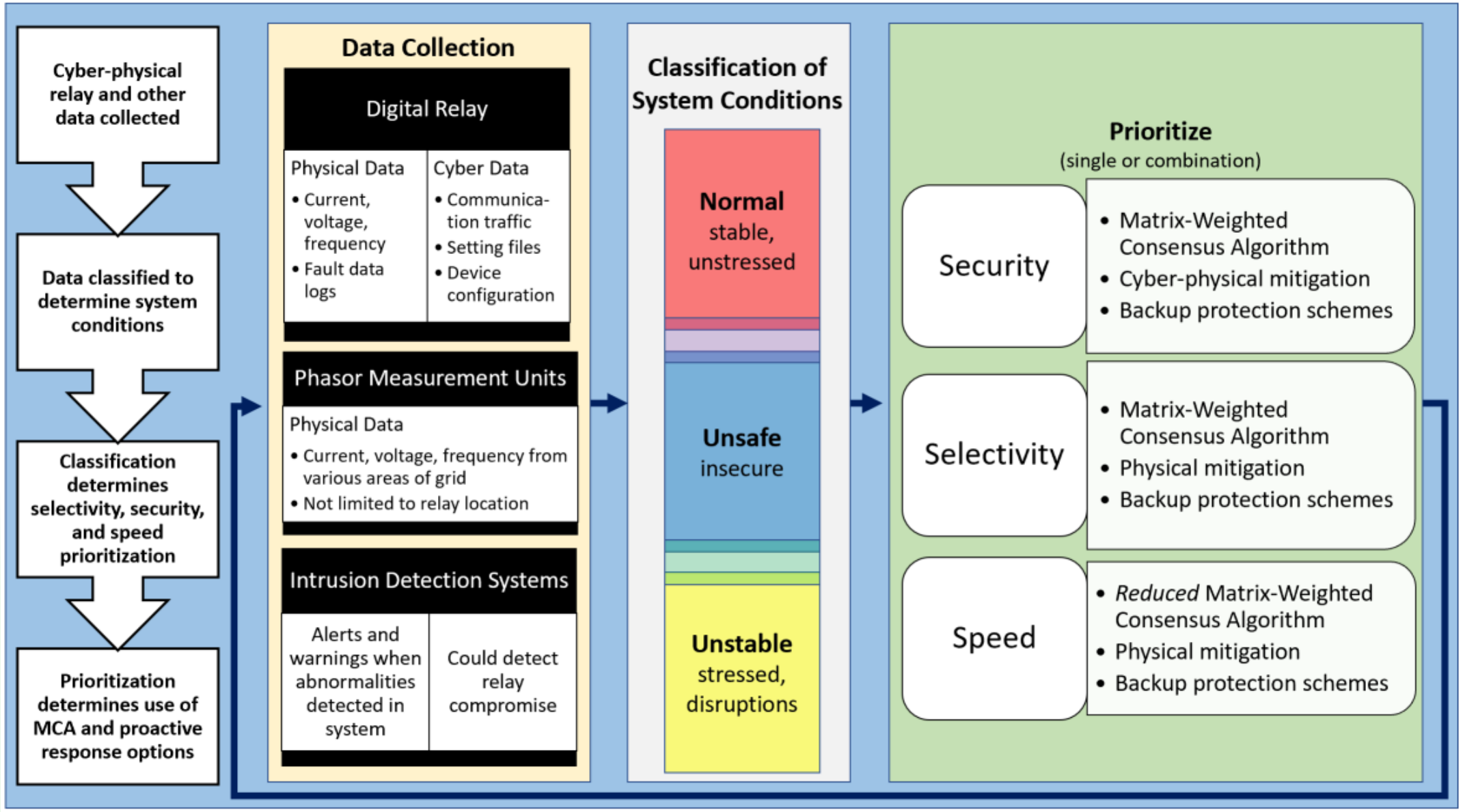


Image Sources: https://knowledge.rtds.com/hc/en-us/articles/360034290474-NovaCor-; https://selinc.com/products/751/; https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000

Overview of HARMONIE-SPS Approach



Machine Learning Approach

Approach converts incoming cyber-physical data into a graph of interconnected nodes, where each edge is a flow of information with an associated timestamp

After the whole capture is split into subgraphs using 24-second sliding windows, the algorithm relies upon two deep learning architectures to obtain an overall representation of the system state in each window:

- A Graph Convolutional Neural Network (GNN/GCN/GCNN), which applies deep learning to the structure of interconnected nodes in the subgraph, and
- A Recurrent Neural Network (RNN), which applies deep learning to the temporal ordering of the edges in the subgraph

A classification layer was added onto the network that predicts two binary labels

- Labels were whether a cyber disturbance is occurring and whether a physical disturbance is occurring
- This combination of two binary labels allows our model to categorize the system state into four categories:
  - 1) normal operations, 2) cyber-only disturbances, 3) physical-only disturbances, and 4) cyber-physical disturbances

We have demonstrated that such a network can be trained using 50 network and physical data captures of various 2-minute scenarios, along with ground truth labels for when a cyber and/or physical anomaly was occurring Tested 4 different disturbances:

1) denial of service (cyber-only),
2) single-line-to-ground fault (physical-only),
3) tripping command injection (cyber-physical), and
4) time-delay attack (cyber-physical)

Initial Machine Learning Results

For the experiments, we partitioned all scenarios into 30 for training, 10 for validation and model selection, and 10 for testing --- these were then split into their respective sliding windows

We ran experiments varying the size of the training data and comparing the results when using a model that has already been pretrained using some basic predefined perturbations versus a model that had not been pretrained.

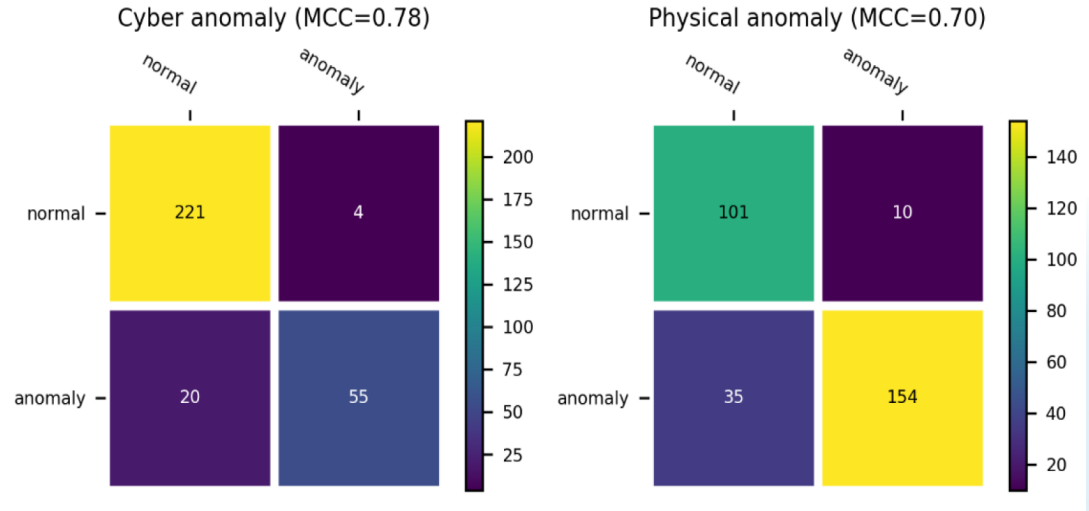
- Used the area under the receiver operator curve (AUC) as our metric because it identifies how well a model's predictions split the two classes apart and does not require a predefined threshold to convert real-valued confidence scores into a discrete class prediction

Table 1-1. Preliminary Results for HARMONIE-SPS ML Model (Cyber anomaly AUC / Physical anomaly AUC)

	With Pretraining	No Pretraining
Training with all 900 windows (30 scenarios)	0.74 / **0.92**	**0.95 / 0.92**
Training with 100 windows (~3.3 scenarios)	0.49 / 0.64	0.52 / 0.60

Initial Machine Learning Results

- From the initial results, we can see that using the full training data, our model can differentiate between disturbances and normal behavior
  - The pretraining step seems to either add nothing to or even mildly hiders the performance of the model, especially when identifying cyber anomalies
  - Could attribute this to a domain shift between the inputs during the pretraining step, where perturbed graphs are given to the model, and the training step, where unmodified graphs are given to the model
  - The current approach is closer to transfer learning than pretraining

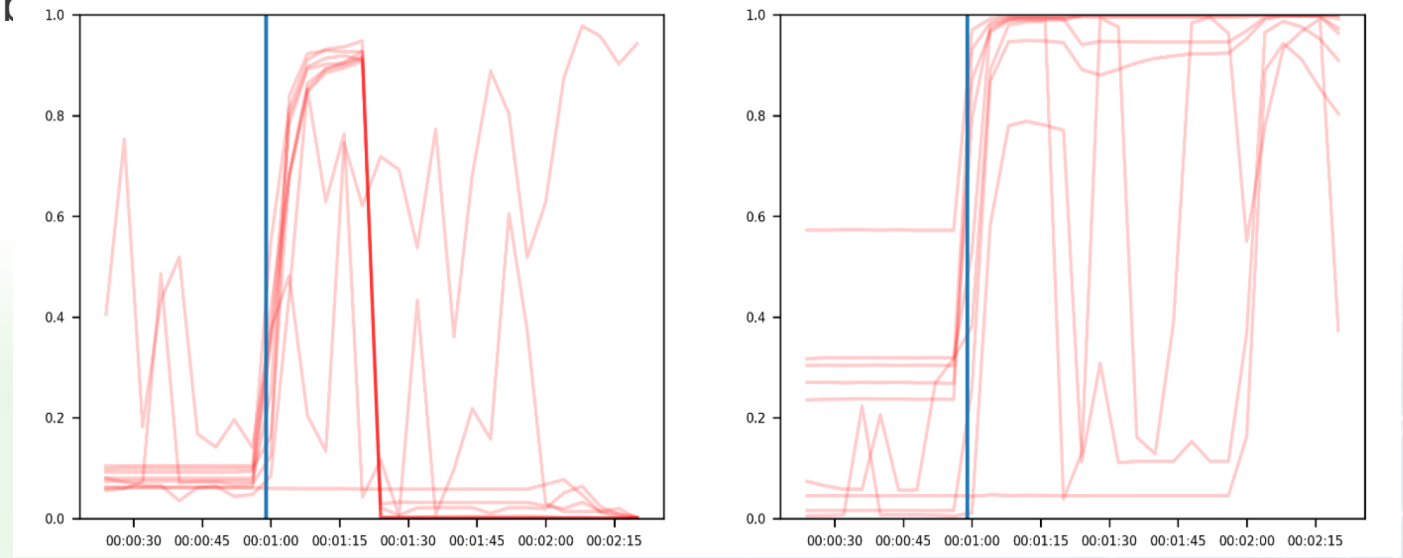


900 training windows with no pretraining

Confusion matrices for identifying cyber and physical disturbances on the test data using a threshold of 0.5. Matthew's Correlation Coefficient (MCC) is used to assess the quality of the predictions. Rows correspond to actual classes and columns correspond to predicted classes.

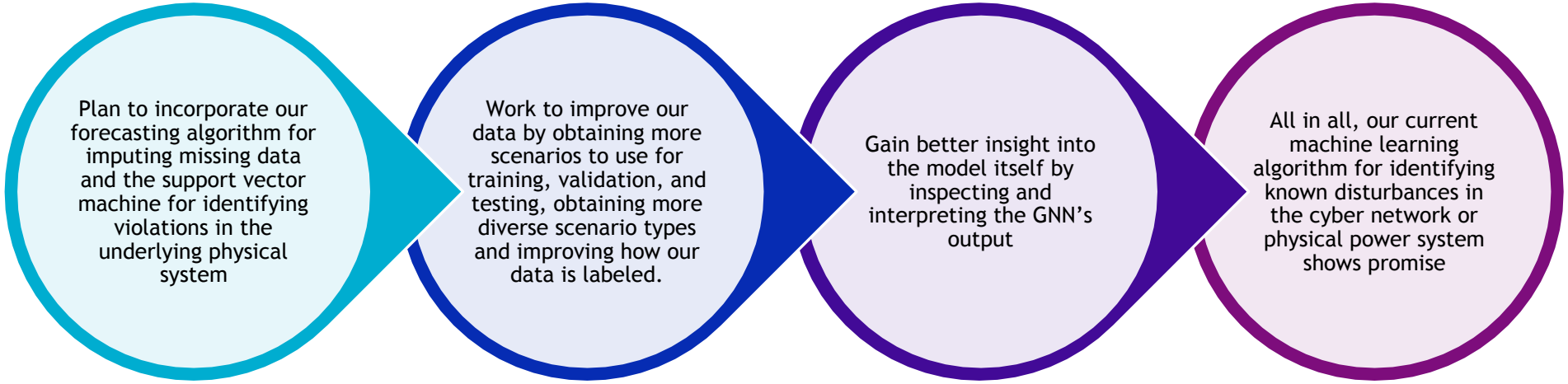
Initial Machine Learning Results

- Also used our best model (trained on all 30 scenarios with no pretraining) to plot the predicted anomaly scores for each scenario in the test set.
- Since our approach uses 24-second sliding windows, all windows ending between 00:00:59 and 00:01:23 will contain the disturbance which occurs at 00:00:59 (blue vertical line).
  - Note that some scenarios have cyber disturbances only in the middle of the capture, which is why some cyber anomaly scores drop to nearly 0 after 00:01:24



Reported anomaly scores over time for the 10 test scenarios. A value of 1 indicates confidence in an anomaly, and a value of 0 indicates the confidence of normal operations. Left: Cyber anomaly score. Right: Physical anomaly score.

Next Steps for Machine Learning Framework



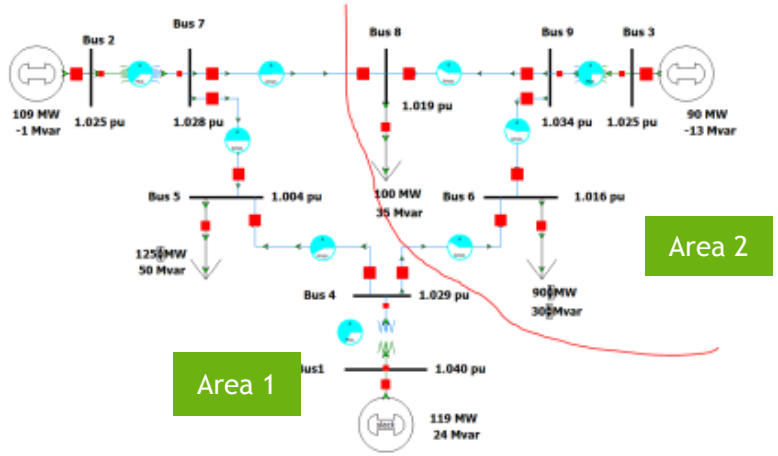
- Our data is complex and noisy, and understanding what our network is learning will be a critical step in refining it

- By continuing to improve upon our existing approach, we believe this will be a viable solution to the problem of using machine learning to understand the wholistic state of a power system and recommend action for our SPS

Machine Learning Approach: Power System Scenarios

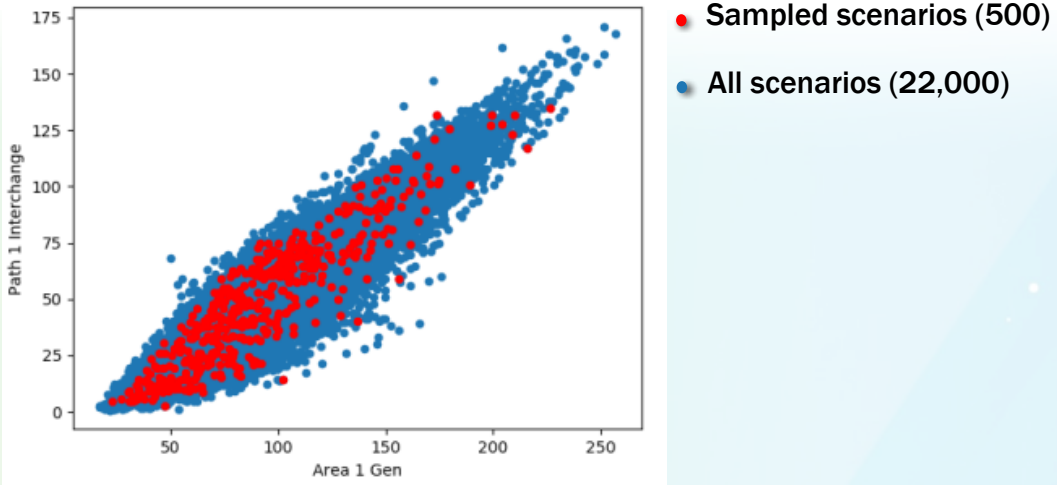
For each test case, many scenarios are created to represent a wide spectrum of physical system conditions

- WSCC 9 bus system:
  - Over 22,000 scenarios representing different combinations of 5 operational variables (area 1 load/gen, area 2 load/gen, area interface flow)



Contingency Analysis for sampled scenarios

- Unrealistic to execute power system studies on all 22,000 scenarios
- Instead, subset of scenarios is sampled and analyzed to reduce the computation time while maintain the results within acceptable accuracy range
- Distributed computing is utilized for the contingency analysis of each sampled scenario



	A	B	C	D	E	F
1	Scenario Name	CTGLabel	LimViolCat	LimViolID	LimViolValue	LimViolLimit
2	WSCC9_0001_8591_tr33.pwb	G_000002Bus2U	Bus High Volts	Bus 2 (2)	1.1016705	1.10000002
3	WSCC9_0001_8591_tr33.pwb	G_000002Bus2U	Bus High Volts	Bus 2 (2)	1.11696506	1.10000002
4	WSCC9_0001_8591_tr33.pwb	G_000003Bus3U	Bus High Volts	Bus 3 (3)	1.11528087	1.10000002
5	WSCC9_0001_8591_tr33.pwb	G_000003Bus3U	Bus High Volts	Bus 3 (3)	1.12898779	1.10000002
6	WSCC9_0001_8591_tr33.pwb	G_000002Bus2U	Bus High Volts	Bus 5 (5)	1.10541546	1.10000002
7	WSCC9_0001_8591_tr33.pwb	G_000002Bus2U	Bus High Volts	Bus 7 (7)	1.10167038	1.10000002
8	WSCC9_0001_8591_tr33.pwb	G_000002Bus2U	Bus High Volts	Bus 7 (7)	1.11693692	1.10000002
9	WSCC9_0001_8591_tr33.pwb	G_000003Bus3U	Bus High Volts	Bus 8 (8)	1.10760462	1.10000002
10	WSCC9_0001_8591_tr33.pwb	G_000003Bus3U	Bus High Volts	Bus 9 (9)	1.11528039	1.10000002
11	WSCC9_0001_8591_tr33.pwb	G_000003Bus3U	Bus High Volts	Bus 9 (9)	1.12898767	1.10000002
12	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 3 (3)	1.10521531	1.10000002
13	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 3 (3)	1.11844742	1.10000002
14	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 3 (3)	1.12450445	1.10000002
15	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 6 (6)	1.10190713	1.10000002
16	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 8 (8)	1.11241782	1.10000002
17	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 9 (9)	1.105214	1.10000002
18	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 9 (9)	1.11844718	1.10000002
19	WSCC9_0001_7809_tr40.pwb	G_000003Bus3U	Bus High Volts	Bus 9 (9)	1.12445295	1.10000002
20	WSCC9_0001_3536_tr72.pwb	G_000002Bus2U	Branch MVA	Bus 4 (4) -> Bus 6	123.1743774	110
21	WSCC9_0001_3536_tr72.pwb	G_000001Bus1U	Branch MVA	Bus 7 (7) -> Bus 5	124.8182449	110
22	WSCC9_0001_3536_tr72.pwb	G_000001Bus1U	Branch MVA	Bus 7 (7) -> Bus 8	121.5330582	110

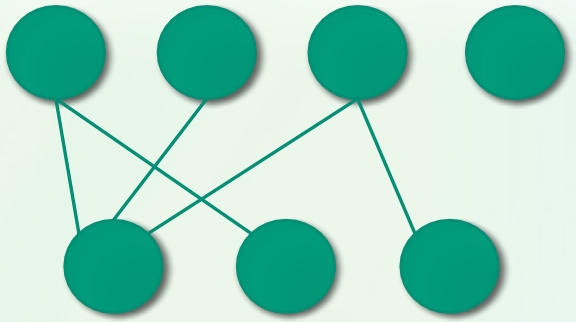
Machine Learning Approach: Automated Corrective Actions

CTG violation elements are clustered as scheme groups; violations within the same cluster can be addressed within one corrective action

- The violation elements from the contingency analysis are presented in graphs
  - *Nodes are unique scenarios and contingencies*
  - *Links are combinations of scenario and contingency that will result in the CTG violation*
- Graphic embedding is utilized to compress the graph information as vectors
- The compressed vectors are then used as input to a hierarchical clustering algorithm that determines the violations being addressed in one single remedial action scheme

S1		S2	S3	
C1	C2	C1	C1	C3

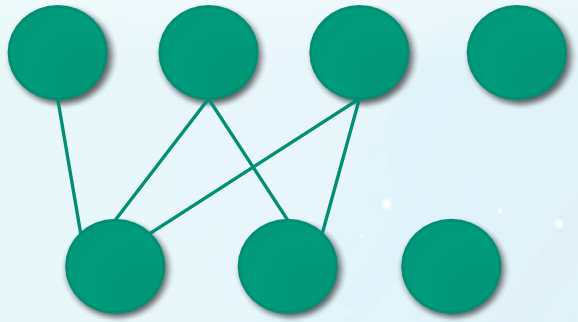
Violation Element 1



[0.12,0.45,0.72,...]

S1	S2		S3	
C1	C1	C2	C1	C2

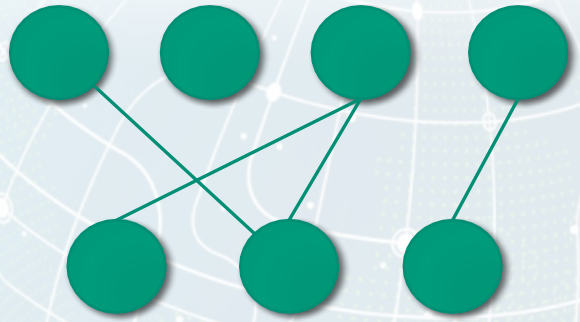
Violation Element 2



[0.62,0.45,0.33,...]

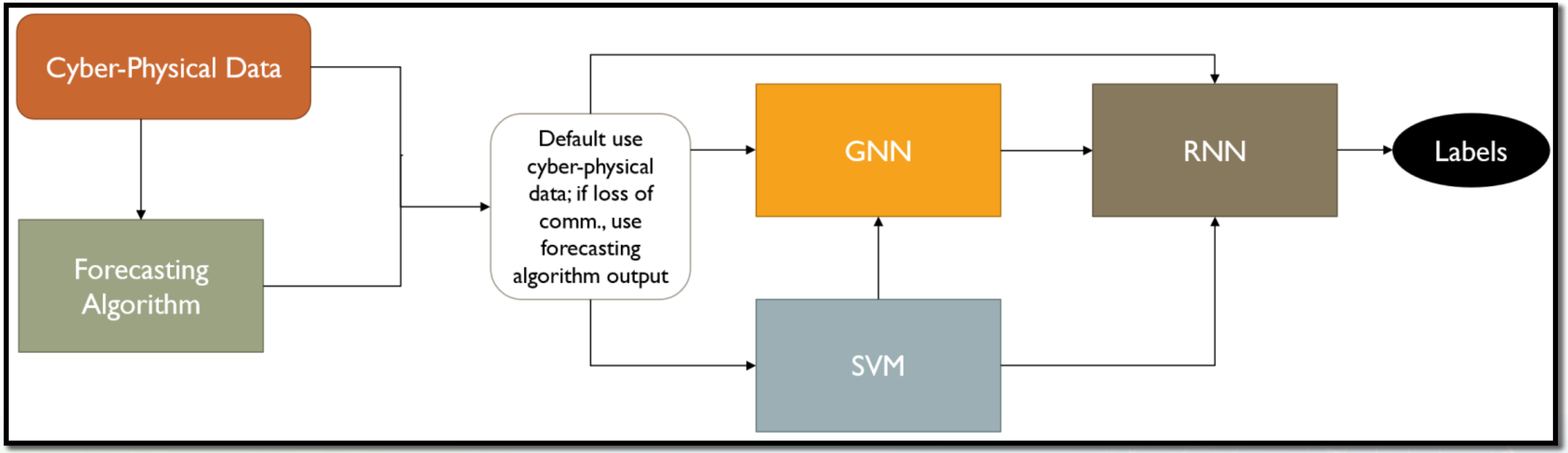
S1	S3		S4
C2	C1	C2	C3

Violation Element 3



[0.12,0.25,0.03,...]

Overall Machine Learning Architecture



GNN: processes cyber-physical structural information
RNN: processes cyber-physical temporal information
SVM: processes power system violations, automates corrective action deployment
Forecasting algorithm: predicts data flows if comm. lost

Next-Generation Relay Voting Scheme

Examining how to formulate next-gen relay voting schemes with far more connectivity and communications enabled devices

- Examining use of communication-enabled digital relays
- Allows new ways to incorporate relays into situational awareness and new protection scheme designs that require fast response times

Using consensus algorithms to enable distributed computation and voting for increased security and resilience

- Not a central point of failure
- Allows us to better analyze impact and potential failures, as well as algorithmic issues and potential challenges like Byzantine failures

Next-Generation Relay Voting Scheme

Examined existing state-of-the-art for relay voting schemes and introduced new designs for communication enabled relays to use consensus algorithms to securely agree on protection actions

- This extends current voting designs such as 2/3 voting to a distributed system
- Final design combines several features, including finding agreement on system state and voting on response actions to take
- Enables faster response to system failures, as interacting with relays in practice is typically done by hand today

Enabling distributed computation for relay voting helps prevent common failures from centralized failure points

Algorithm 1 Relay voting with BFT

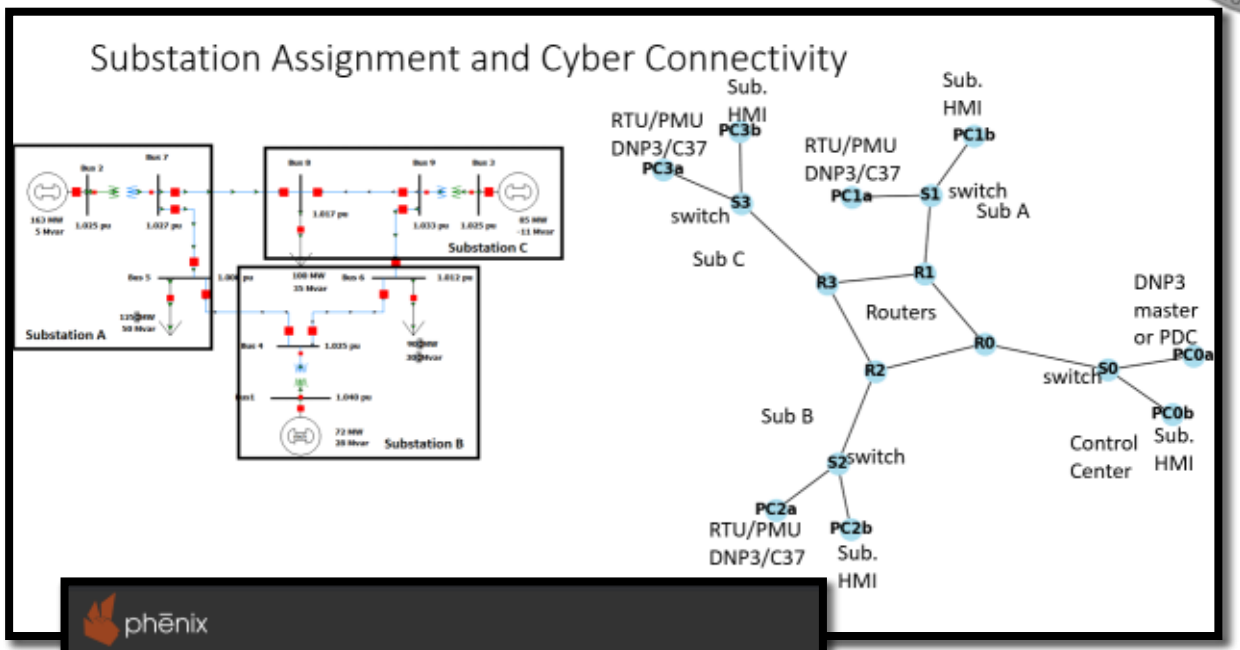
1) Relay i detects under frequency conditions
2) Relay i initiates request
3) Request for voting multicast to all other relays
4) All relays compute protection scheme calculations, determine load to shed
5) Each relay multicasts result to all other relays in group
6) Each relay waits for $f + 1$ replies, saves result.
7) Relay j that needs to shed load acts accordingly

Jacobs, Nicholas, et al. "Next-Generation Relay Voting Scheme Design Leveraging Consensus Algorithms." *2021 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2021.

Cyber-Physical Emulation Environment

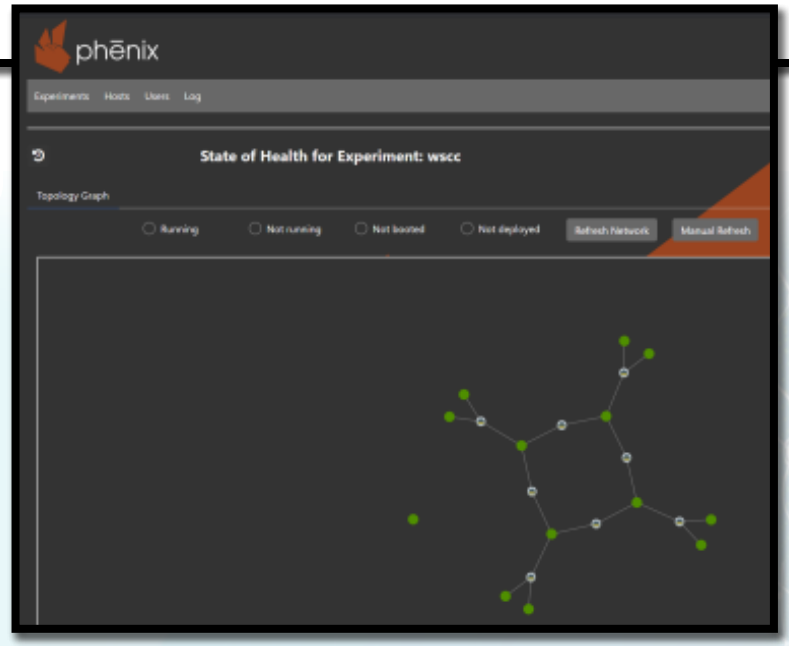
Have WSCC 9-bus system model running in RTDS

- Interested in collection of data from the RTDS at different sampling rates due to non-contingency (low sampling) and contingency events (high sampling)
- This data will be used to both train and analyze the HARMONIE ML algorithms
  - Also leveraging TAMU testbed and related project datasets

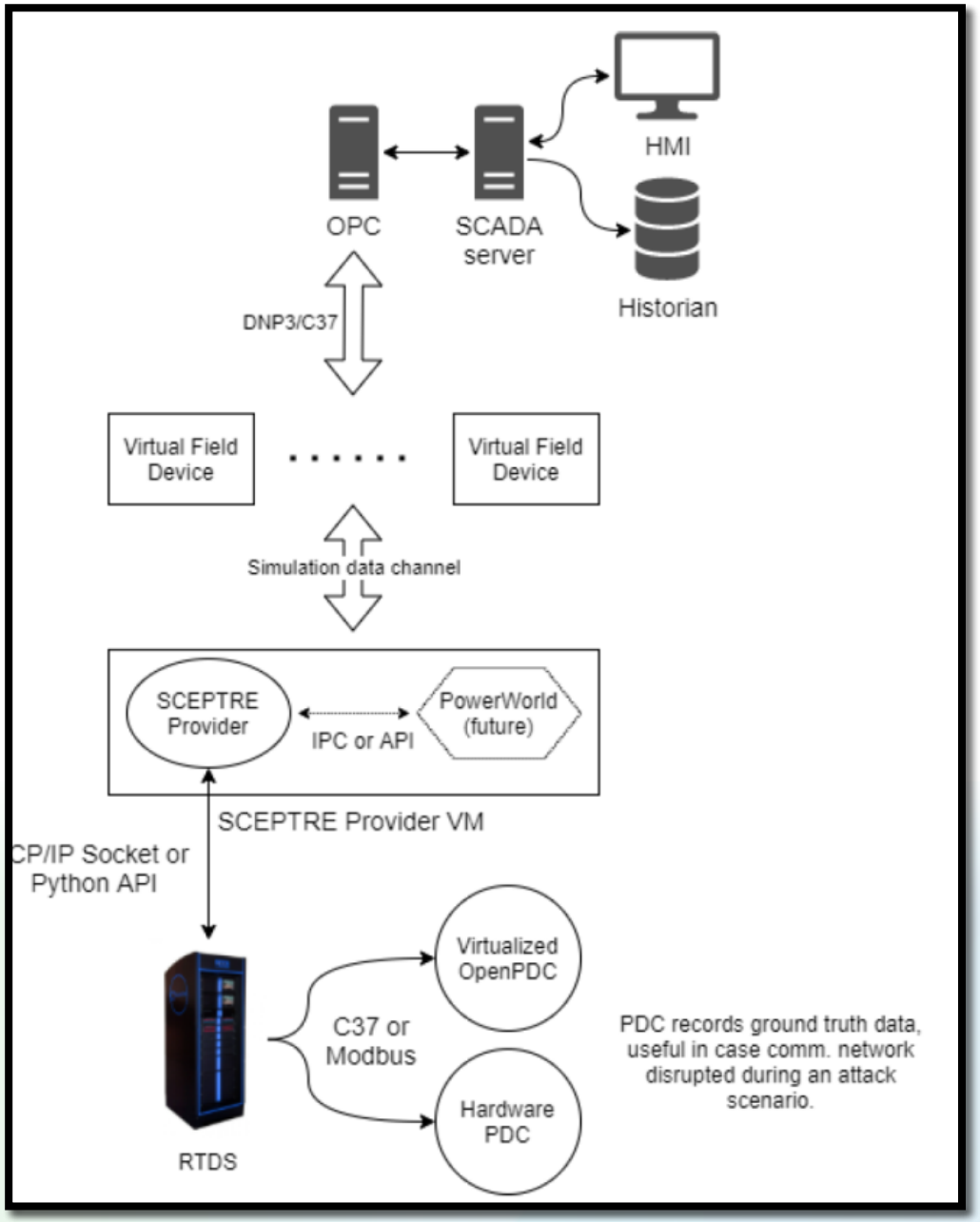


SCEPTRE plan

- Communication network design by TAMU
- Integration currently in-progress
- 9-bus emulation will be used as a testbed to prototype and test the RTDS integration
  - Leveraging virtual PDC
  - Upgraded RTDS to stream PMU data



Cyber-Physical Emulation Environment



Overall HARMONIE-SPS emulation environment architecture including SCEPTRE™,
RTDS, and both virtualized and HIL components

Current Efforts and Future Work

HARMONIE-SPS Approach progress:

- Cyber-physical machine learning framework for classifying disturbances
- Automated corrective action deployment (physical-side)
- Next-generation relay voting scheme
- Cyber-physical emulation environment

Future Work

- Extending SVM framework to deploy cyber corrective actions
- Collecting disturbance data for ML testing from cyber-physical emulation environment
- Cohesive HARMONIE-SPS deployment (e.g., as a tool)

Thanks! Questions?

Cyber-Physical Disturbance Scenarios

Exploring different scenarios with varying cyber-physical impact

Command Injection (MITM & Data Injection): tripping command injection cause single failure and multiple failures

Implementation: Tripping command could be sent from the substation-HMI, or from other nodes through the routing path

Data:
- Cyber: flow data
- Physical: Before and after event measurements (steady state) or transient data (5 seconds)

Relay setting change (Data Injection): changing settings or disabling cause single and multiple failures

Implementation: Change setting command could be sent from the substation-HMI, or from other nodes through the routing path

Data:
- Cyber: flow data
- Physical: Before and after event measurements

DDOS-1: Block the normal communication for collecting measurements

Implementation: Missing PMU data

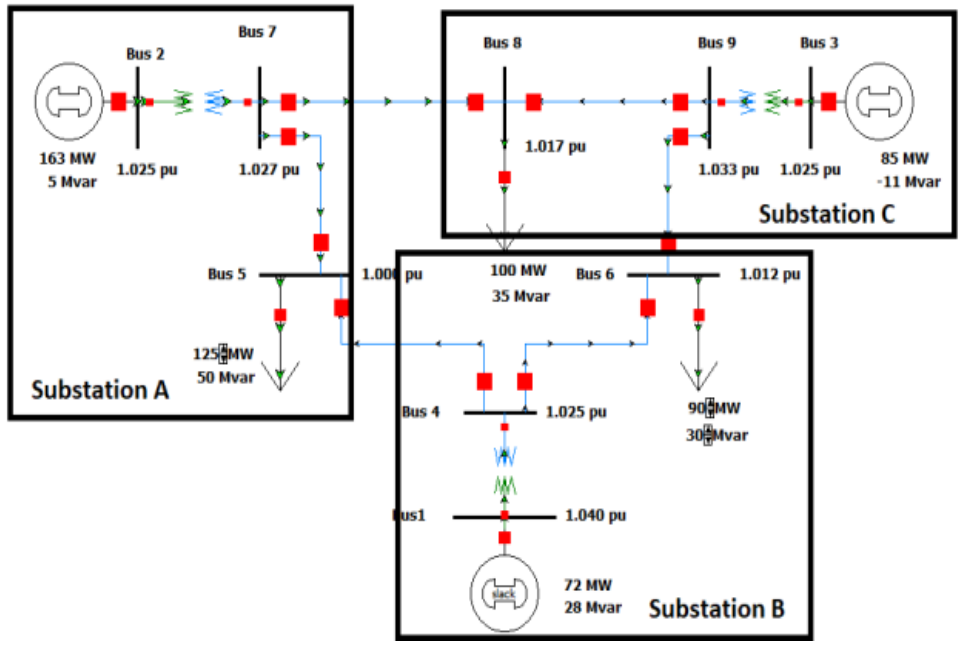
Data:
- Cyber: synthetized flow data (from CIC-DDoS2019)
- Physical: Before and after event measurements: unavailable data

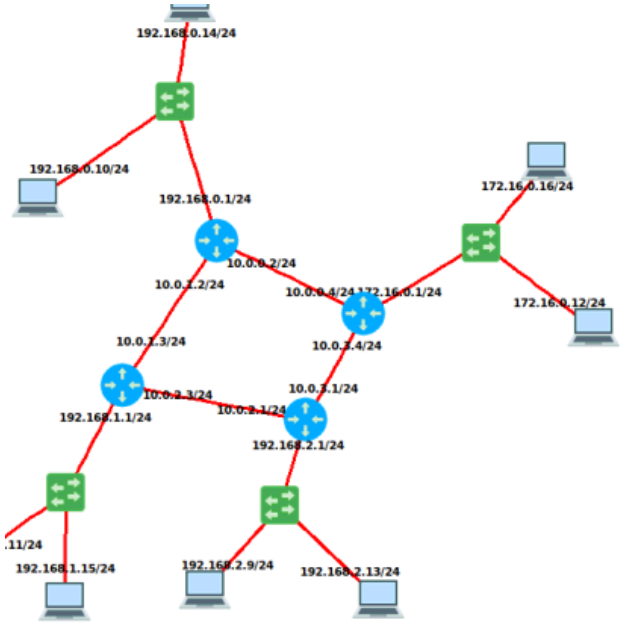
DDOS-2: Block the normal communication for control (normal and mitigation)

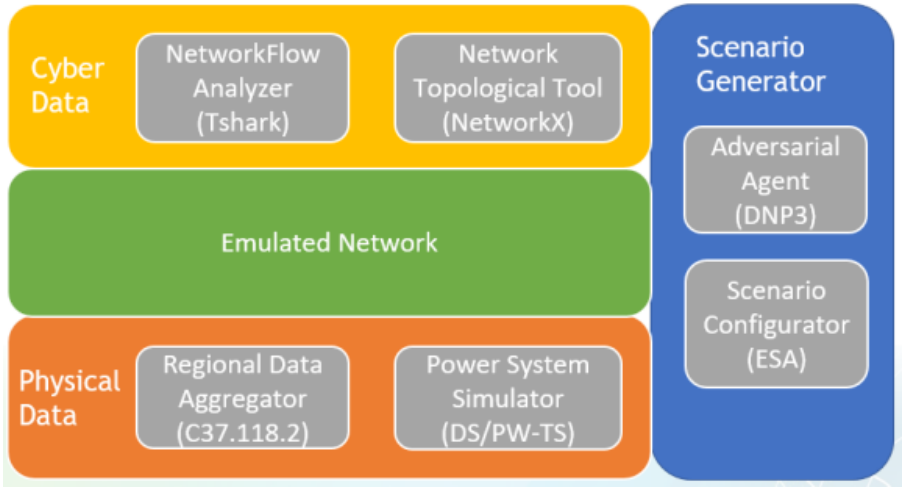
Implementation: DDOS in SCEPTRE (single location/multiple location) + physical control command

Data:
- Cyber: synthetized flow data (from CIC-DDoS2019)
- Physical: Before and after event measurement
- Note: based on the DDOS location, the physical mitigation command may or may not be received by the targets

Cyber-Physical Disturbance Scenarios



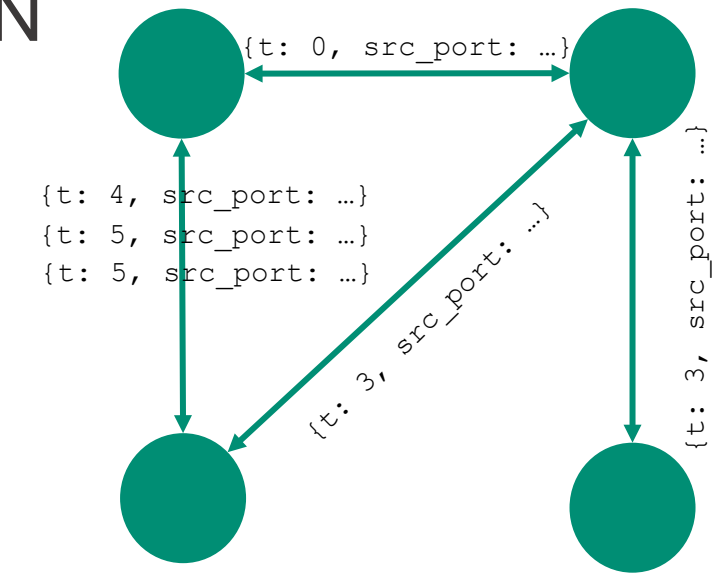




Tripping command injection

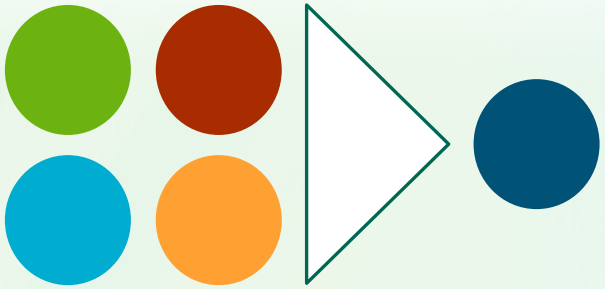
- Each substation has 5 IEDs (connected to breakers) and 1 HMI connected to the substation relay
- Base scenarios: 140
- Single Injection (in-substation): 15
- Multiple Injection (in-substation): 30
- Total scenarios: 6300
- Cyber data: flow information (source ip, source port, destination ip, destination port, time, protocol)
- Physical data: 1 steady-state data 2 transient data

GNN + RNN

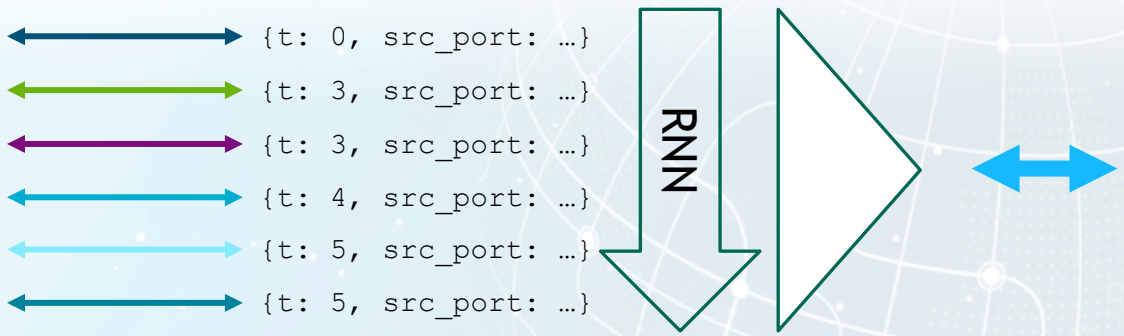


After n iterations of message passing, each node and edge has its own vector

A weighted mean of node vectors encodes structural information



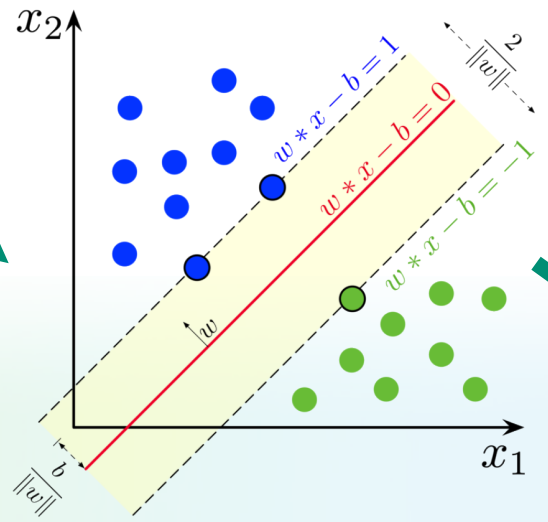
A weighted mean of edge vectors encodes temporal information



Edge Transformations

Use the fault classifier SVM's output to add alert edges when violation conditions are met

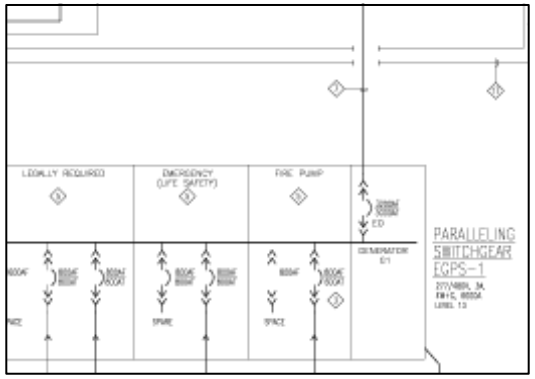
```
{t: 4, device: XYZ, voltage: 123, …}
```




```
{t: 4, device: XYZ, voltage: 123}, …}
{t: 4, device: XYZ, type: voltage-violation-alert, …}
```

Forecasting

Impute missing edges using forecasting with Facebook's Prophet



PROPHET

{t: 1, v: 123, …}
{t: 2, v: 123, …}
{t: 3, v: 124, …}
{t: 4, v: 122, …}
{t: 5, v: 124, …}
{t: 6, v: 124, …}
{t: 7, v: 125, …}
{t: 8, v: 124, …}
{t: 9, v: 124, …}
{t: 10, v: 121, …}
…

{t: 1, v: 123, …}
{t: 2, v: 124, …}
{t: 3, v: 123, …}
{t: 4, v: 122, …}
{t: 5, v: 124, …}
?
{t: 7, v: 125, …}
{t: 8, v: 123, …}
{t: 9, v: 124, …}
{t: 10, v: 122, …}
…