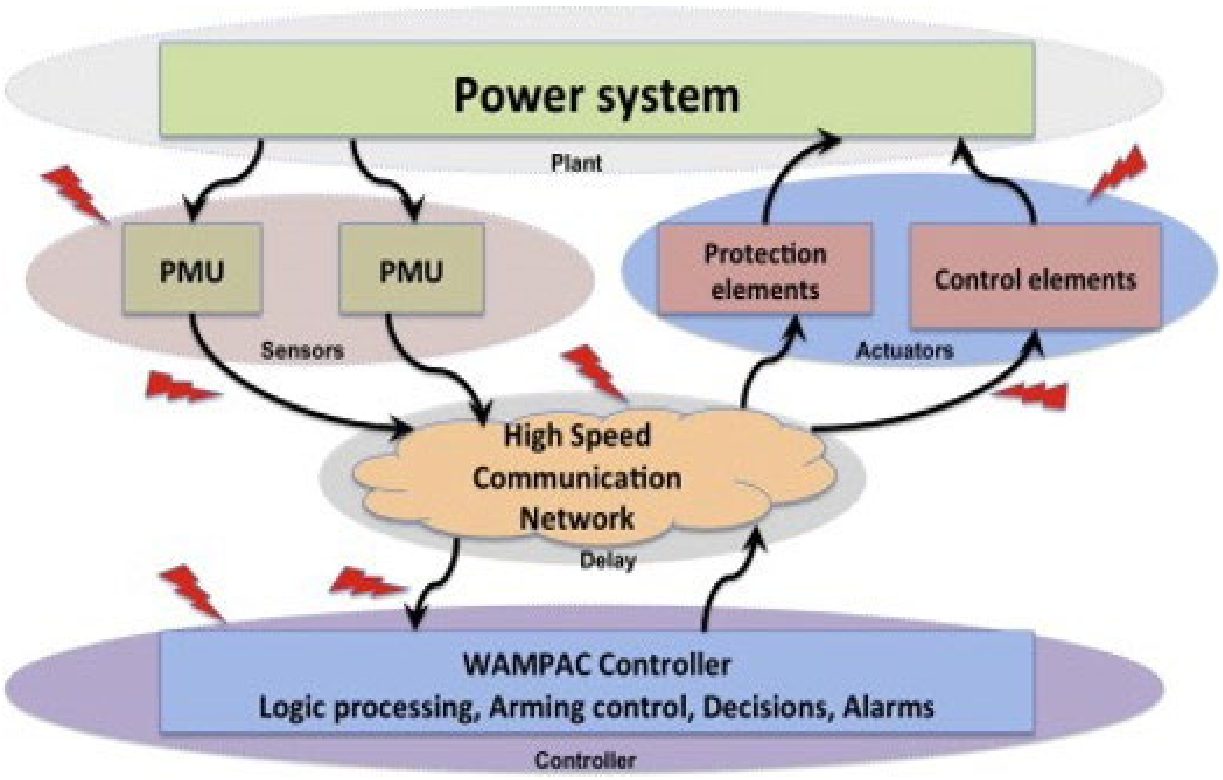
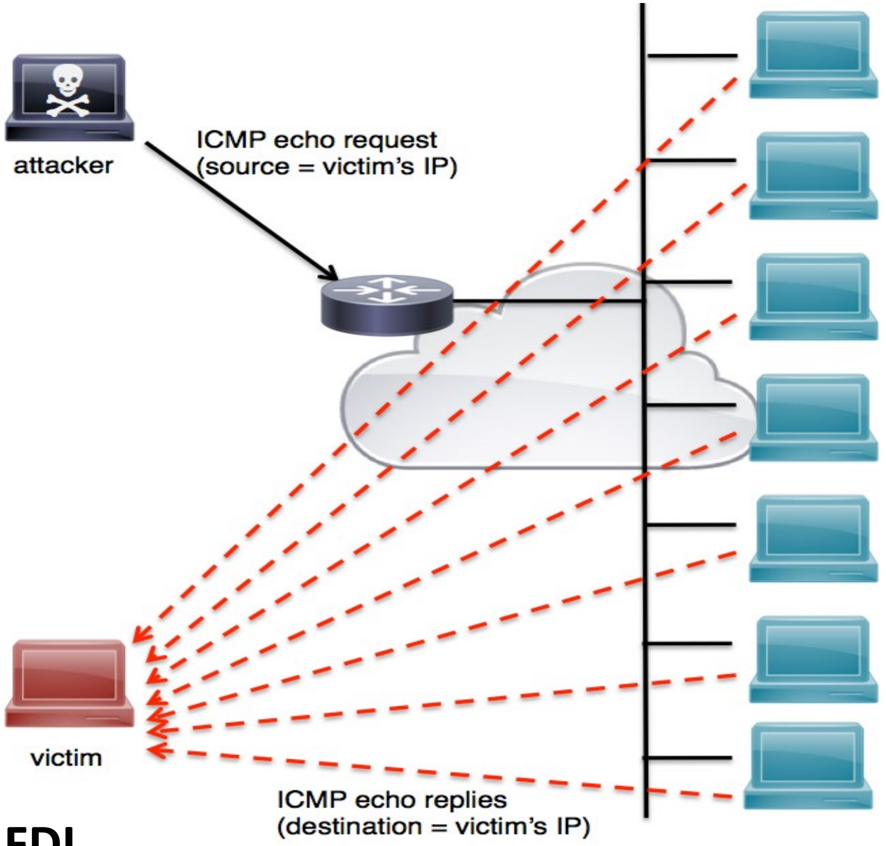
An Integrated Generative Adversarial Network for Identification and Mitigation of Cyber-Attacks in Wide-Area Control

Jishnudeep Kar
PhD Student
North Carolina State University
(jkar@ncsu.edu)

Aranya Chakrabortty
Professor
North Carolina State University
(achakra2@ncsu.edu)

About DoS/FDI attacks





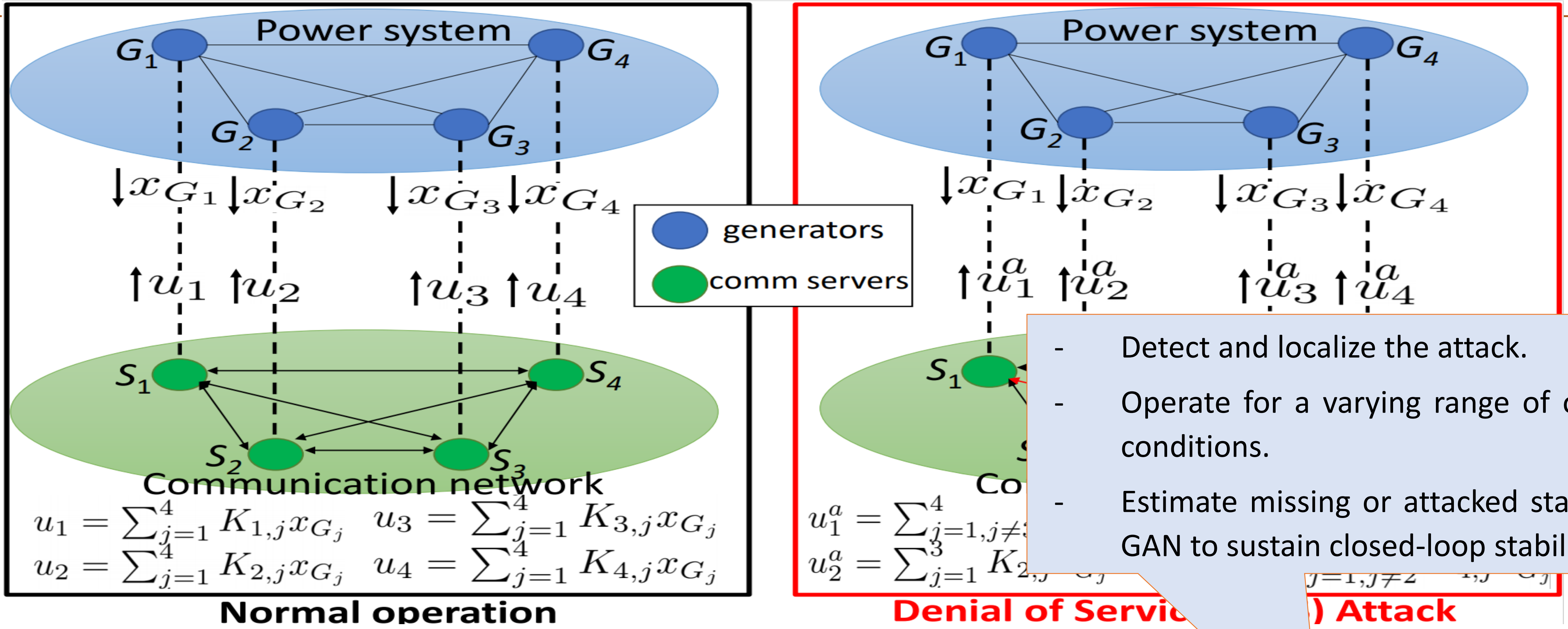
About DoS

- The communicating servers are jammed with malicious request
- Server becomes unable to respond to legitimate requests.

About FDI

- The data shared over the network is corrupted by adding bias.
- This could cause controllers and operators to actuate wrong control actions which may cause closed-loop instability..

About DoS/FDI attacks



- Detect and localize the attack.
- Operate for a varying range of operating conditions.
- Estimate missing or attacked states using GAN to sustain closed-loop stability.

Fig. 2 : Cyber-physical architecture for WAC

$$\Delta \dot{x} = A\Delta x + Bu \qquad u = -K\Delta x$$

$$J = \int_0^\infty (\Delta x^T Q \Delta x + u^T R u)dt\,, \quad Q \geq 0\,,\ R > 0$$

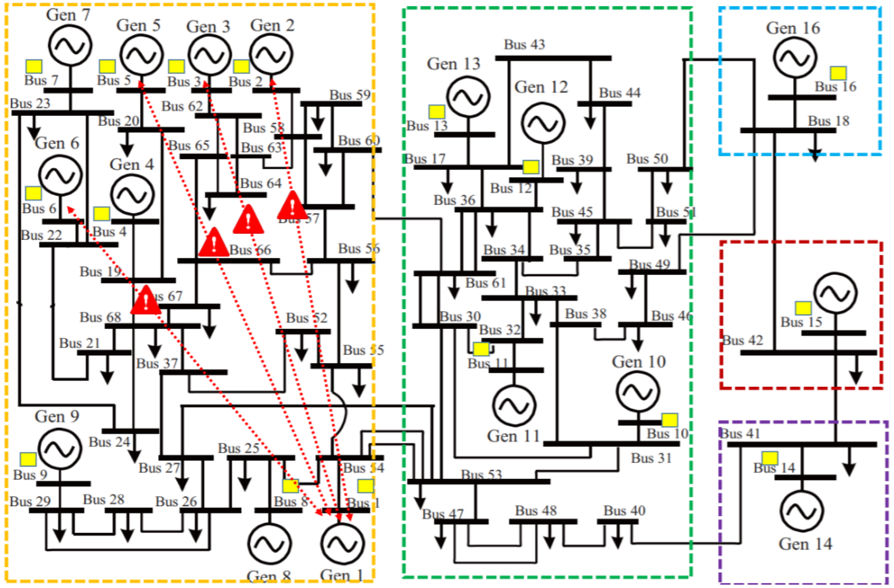
$$\Delta u^a_{G_i}(t) = -\Big(\sum_{G_j \in \mathcal{S}_i} K_{i,j}(x_{G_j}(t) - x_{o,G_j}) + \sum_{G_k \in \mathcal{A}_i} K_{i,j}(x^a_{G_k}(t) - x_{o,G_k}) \Big)$$

Why is resiliency needed ?

Bus (.) : indicates a PMU at bus (.)

Area 1 : $G_{1,..,9}$
Area 2 : $G_{10,..,13}$
Area 3 : G_{14}
Area 4 : G_{15}
Area 5 : G_{16}

Attacked link



IEEE 68-bus 16 m/c system

Small-signal power grid model

$$\Delta\dot{x} = A\Delta x + Bu$$

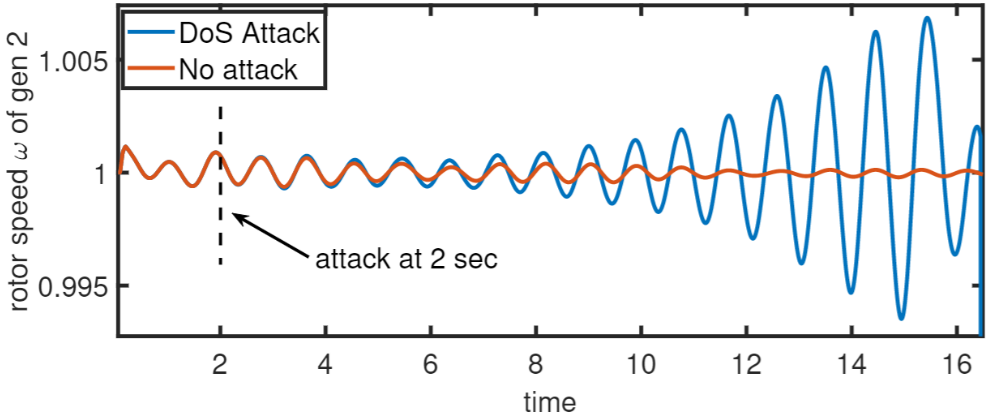
Design a **damping** control input

$$u = -K\Delta x$$

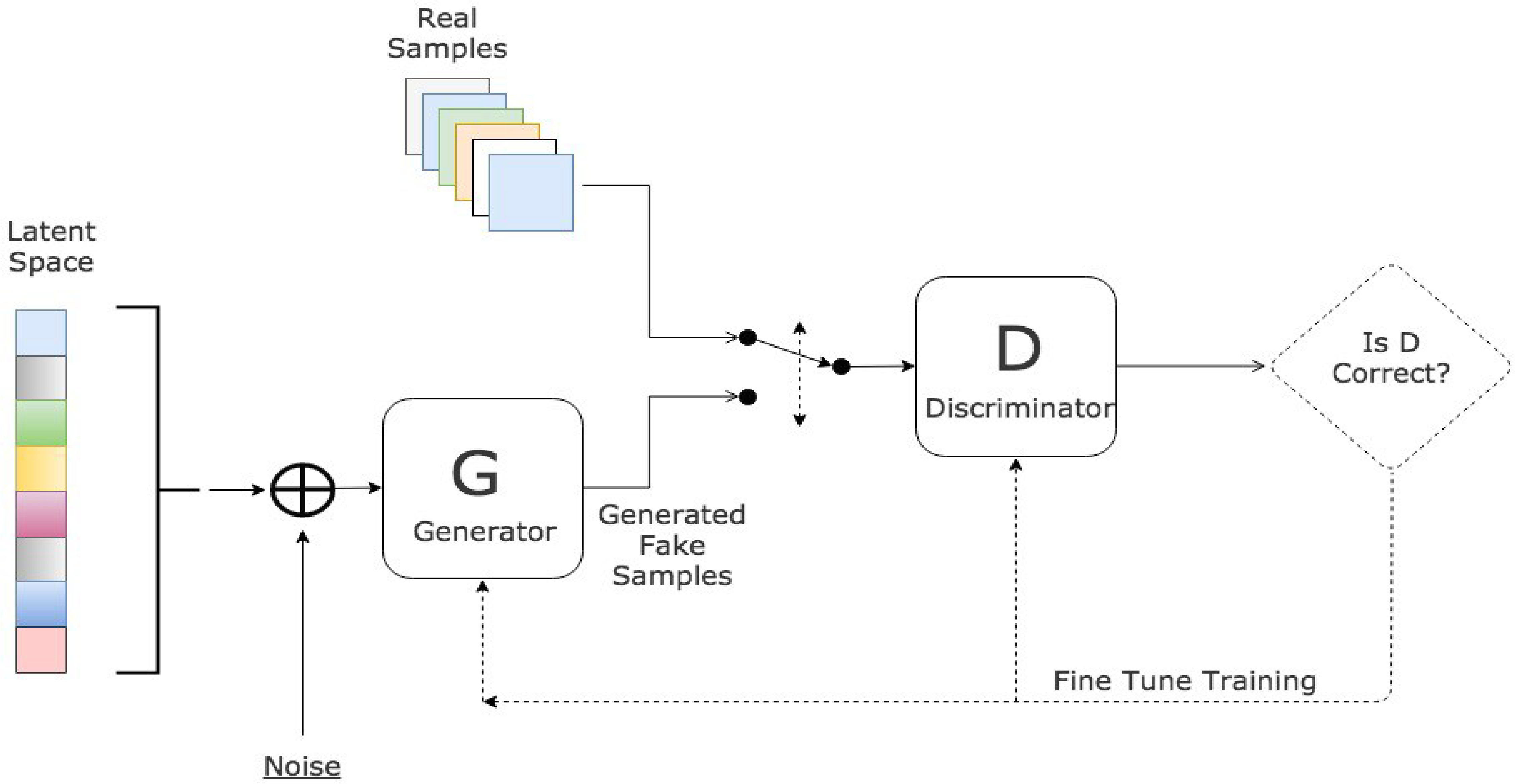
Minimizing LQR cost

$$J = \int_0^{\infty} (\Delta x^T Q \Delta x + u^T R u)dt \ , \ Q \geq 0 \ , \ R > 0$$

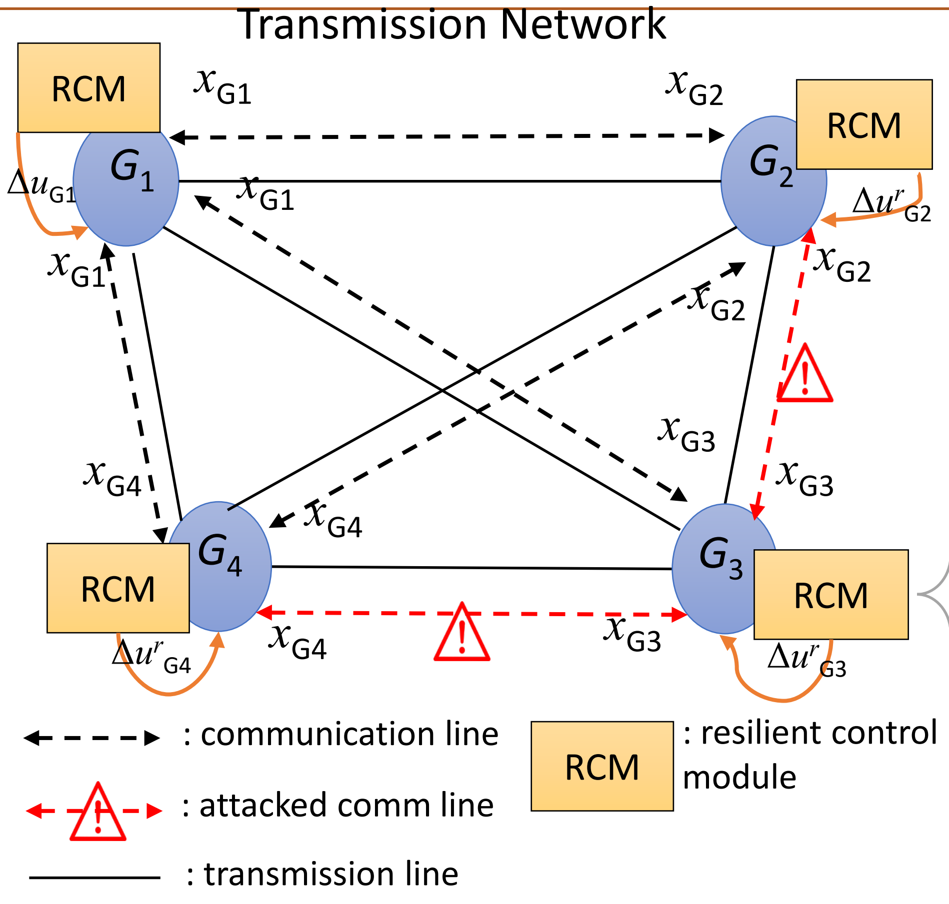
$$K = \begin{bmatrix} K_{1,1} & K_{1,2} & \dots & K_{1,n} \\ \vdots & \vdots & \vdots & \vdots \\ K_{n,1} & K_{n,2} & \dots & K_{n,n} \end{bmatrix}$$

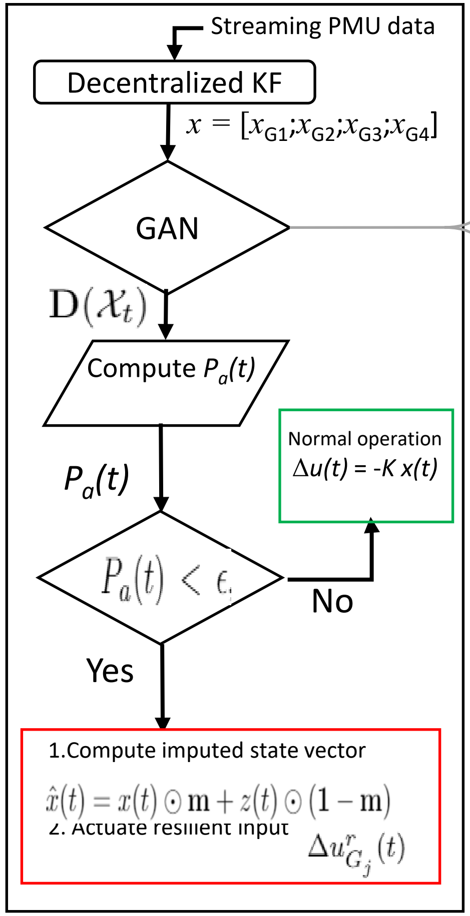


Generative Adversarial Network



Our proposed GAN



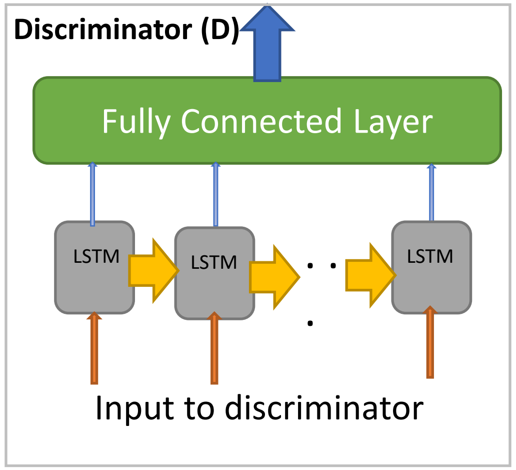


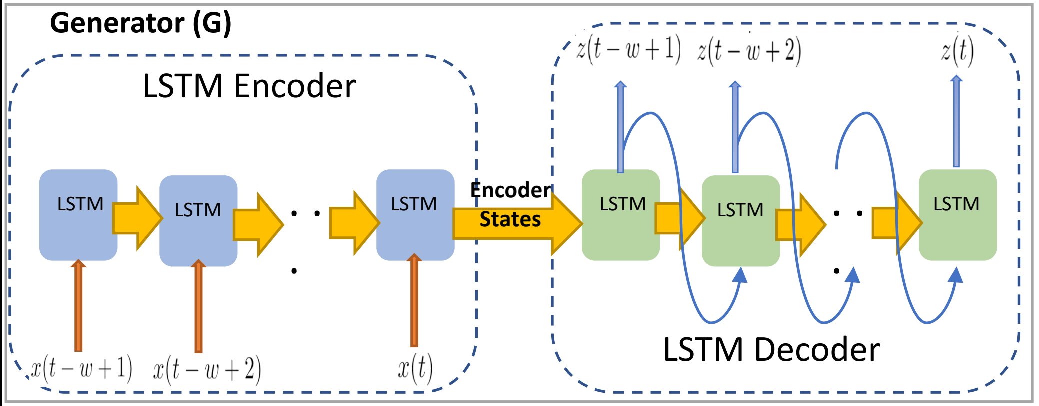
Training

- Input to discriminator is a concatenation of real data and generator output.
- Input to LSTM encoder is attacked state vector from moving data window.

Discrimination

- Input to discriminator is data from moving window for attack detection and identification
- Output of LSTM decoder is imputed healthy states to form augmented state vector.





$$P_a^{G_j}(t) = \left(\mathbf{D}(\mathcal{X}_t^{G_j}) + \mathbf{D}(\mathcal{X}_{t-1}^{G_j}) + \ldots + \mathbf{D}(\mathcal{X}_{t-d+1}^{G_j})\right)/d.$$

Moving average – WHY ?

- GANs cannot be trained to 100% accuracy
- Instances of anomalous scores.
- Averaging removes anomaly.

NC STATE UNIVERSITY

NASPI

Simulation

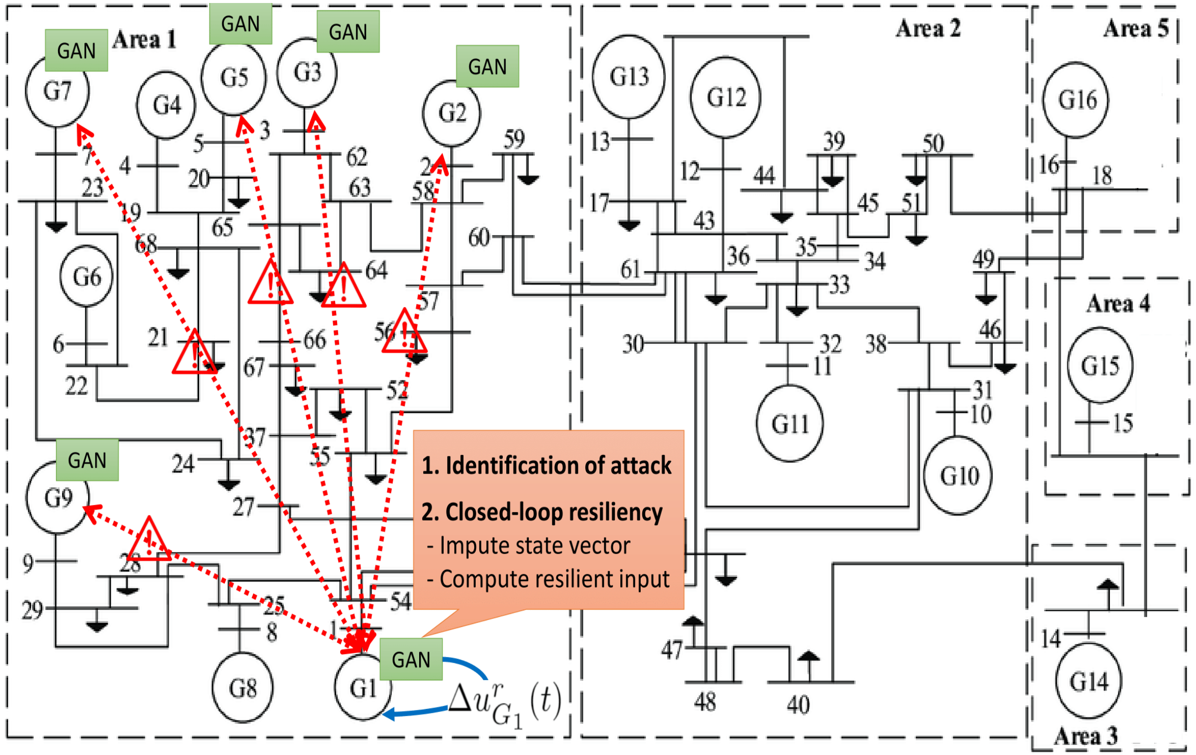
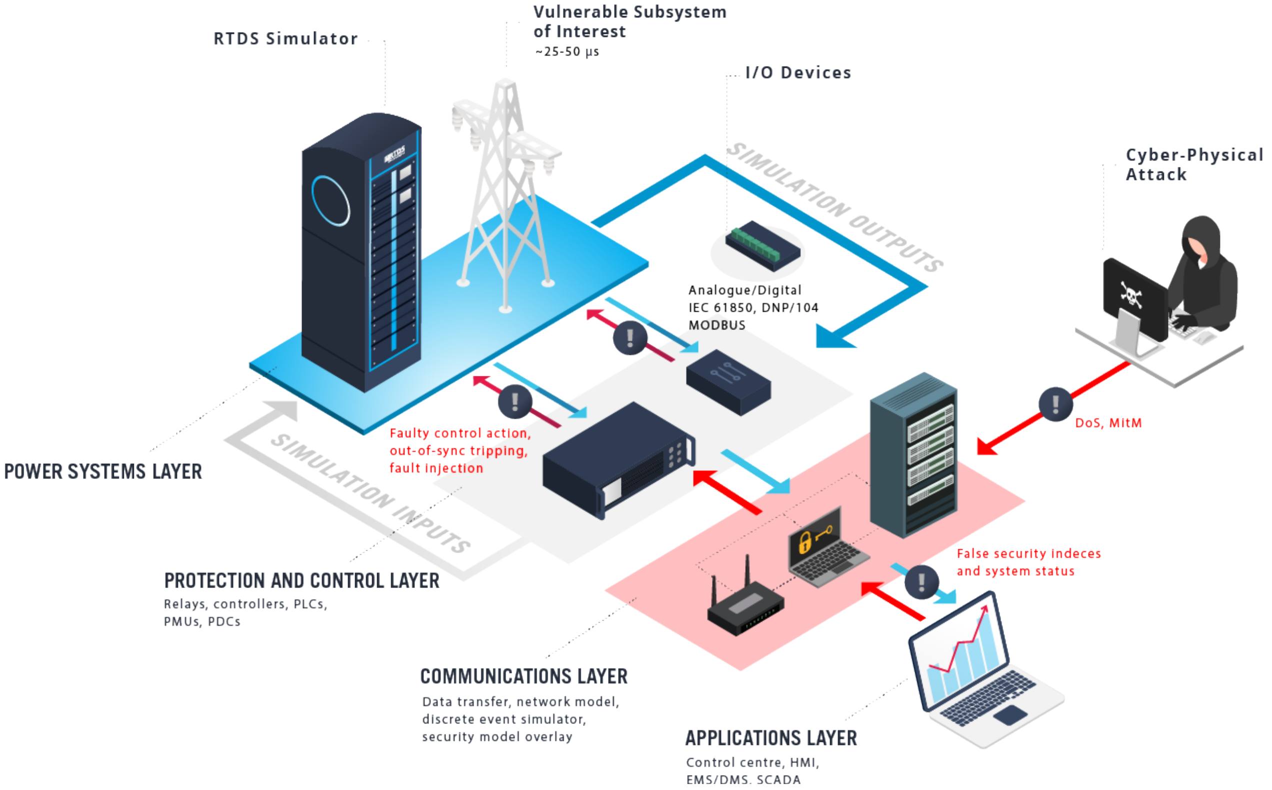


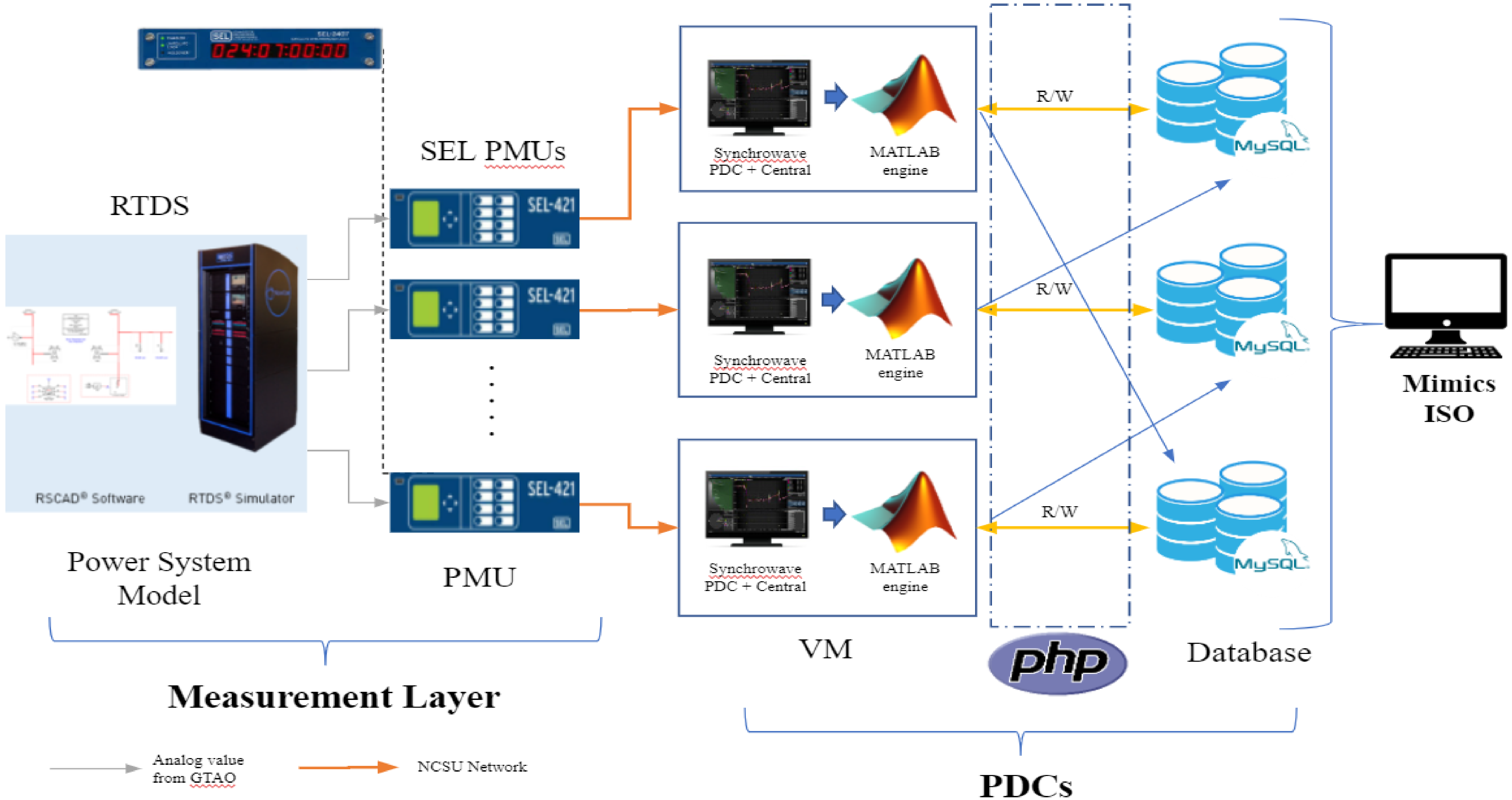
Fig. 4 : IEEE 68-bus system



- Attacked links are shown by red lines.
- Training data consists of 5000 operating points.

- Communication delays are :
  - *Intra-area = 30ms, Inter-area = 60ms.*
  - *Deviation = +/- 10%*

Testbed



SEL PMUs
RTDS
SEL-421
RSCAD® Software
RTDS® Simulator
Power System Model
PMU
Measurement Layer
Synchrowave PDC + Central
MATLAB engine
R/W
MySQL
Mimics ISO
VM
php
Database
PDCs
Analog value from GTAO
NCSU Network

Detection and identification (FDI)

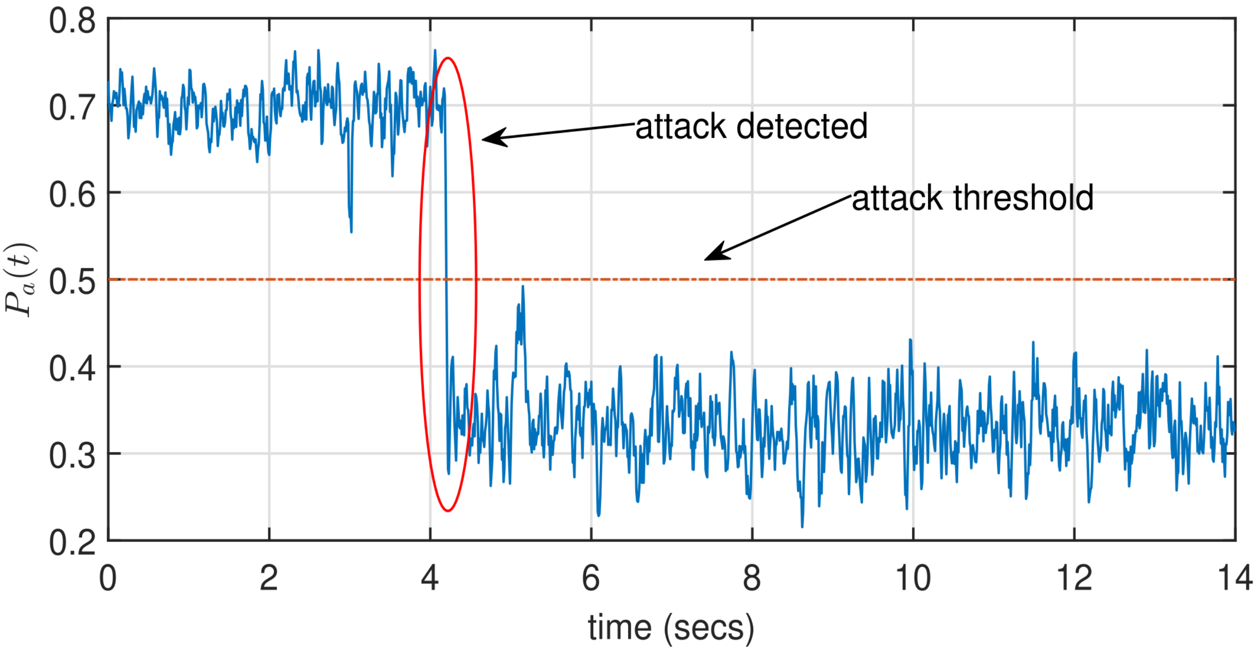


Fig. 5a : Detection of FDI attack

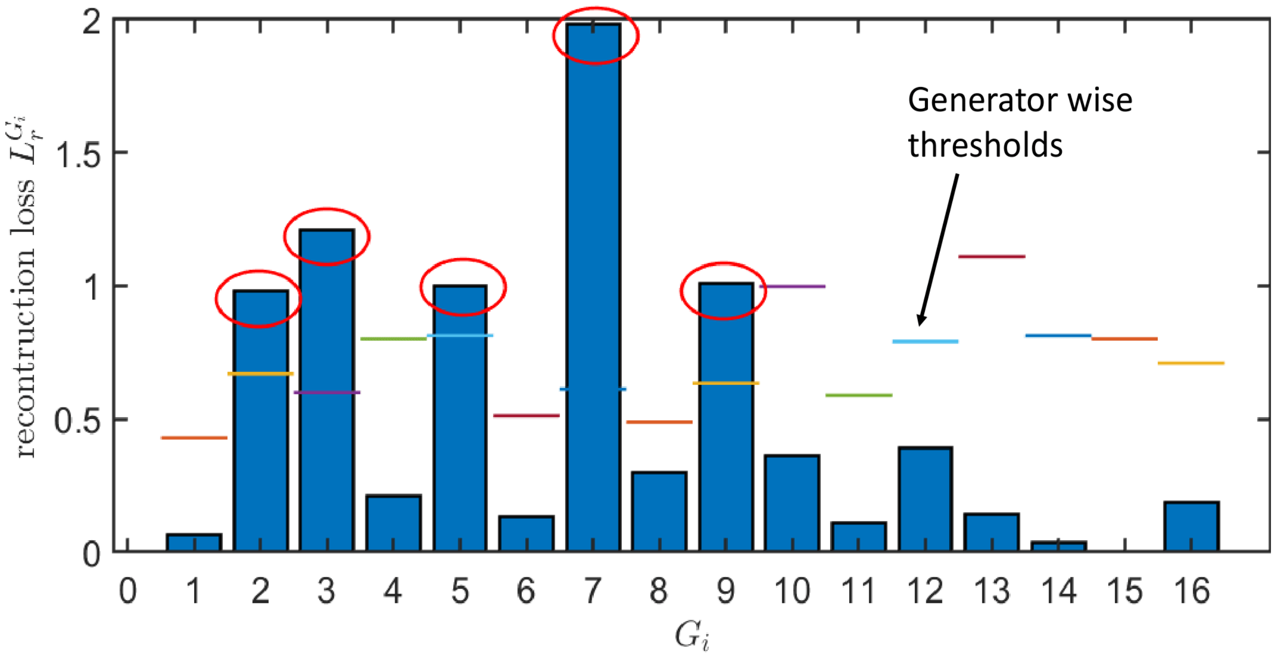


Fig. 5b : Localization of FDI attack

- It is seen that during a FDI attack, the average discriminator Pa shows a sudden drop.
- The threshold can be estimated based on the best score obtained during training phase.
- Generator wise reconstruction error is computed between received and predicted states.

$$L_r^{G_j}(t) = ||(\mathcal{X}_t - \mathbf{G}(\mathcal{X}_t)) \odot \Omega_{G_j}||$$

Closed-loop results

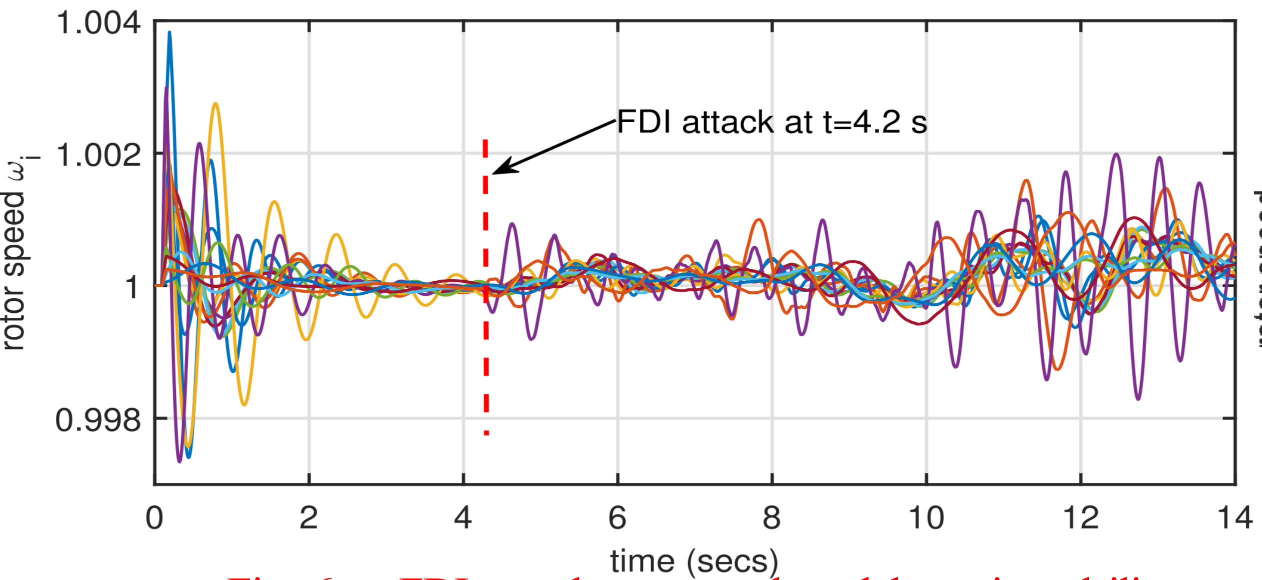


Fig. 6a : FDI attack causes closed-loop instability

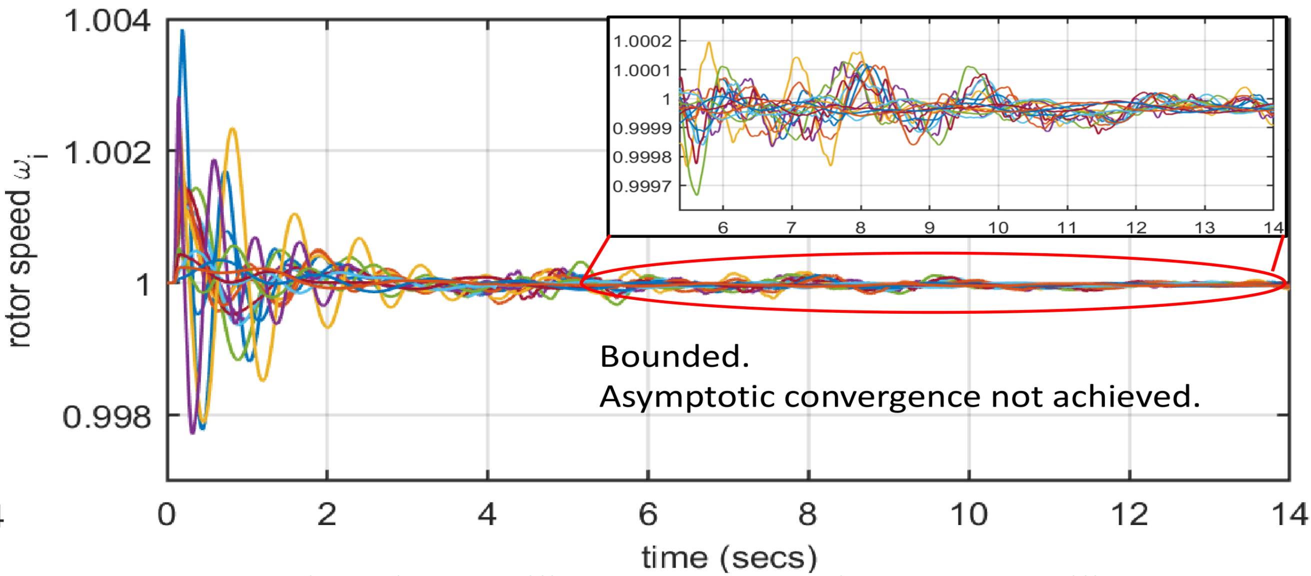


Fig. 6b : Resiliency to FDI using GAN resiliency

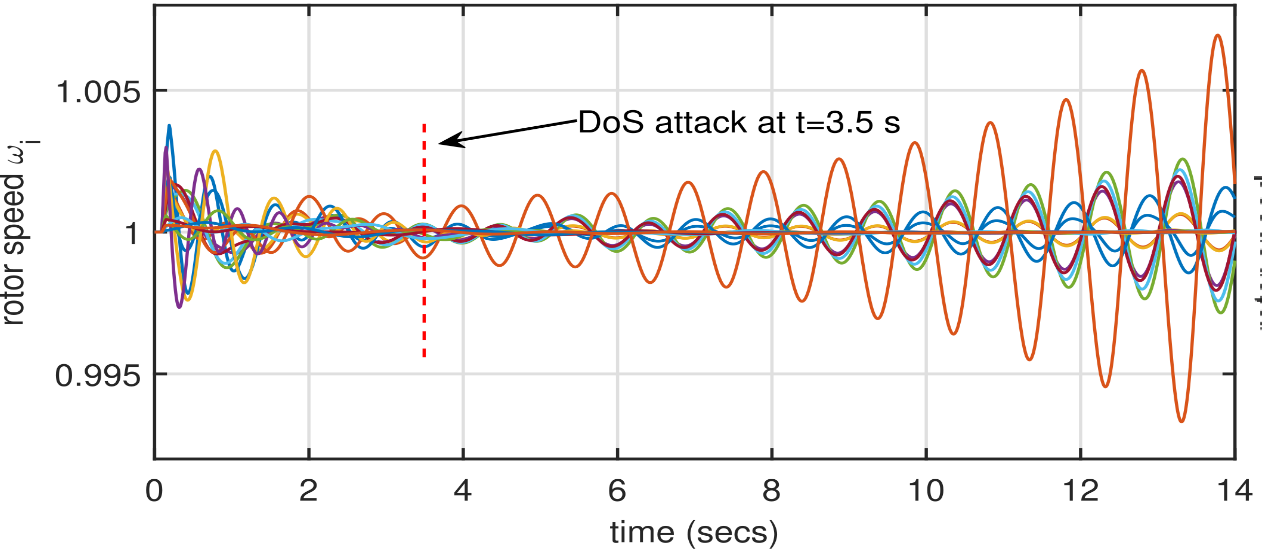


Fig. 7a : DoS causes closed-loop instability

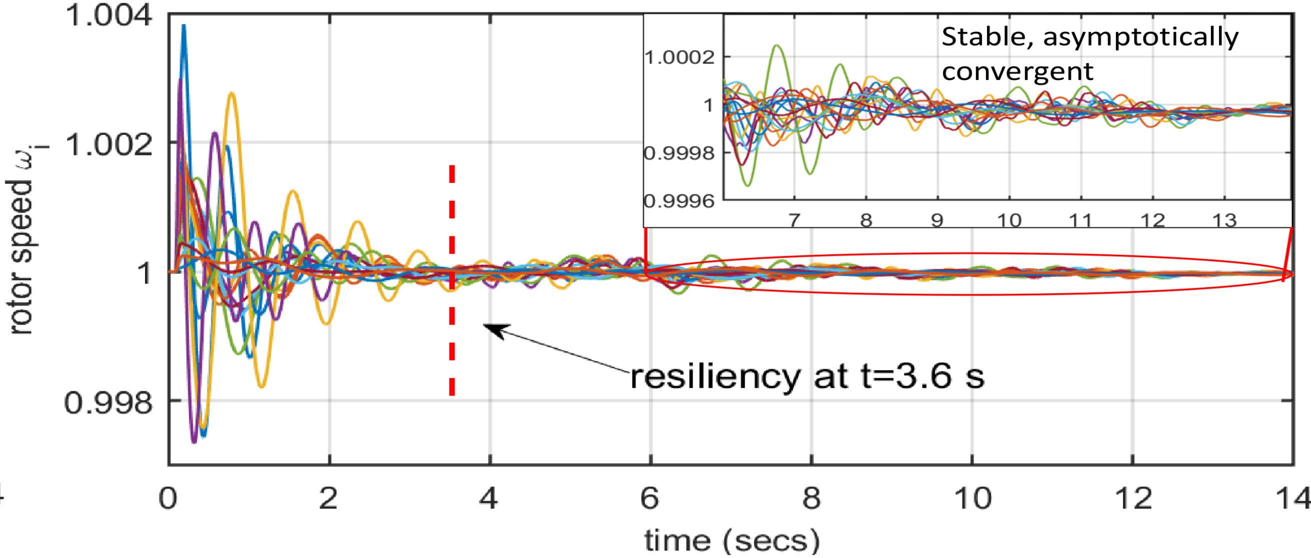


Fig. 7b : Resiliency to DoS using GAN

Conclusions and references

- Neural network-based methods benefit in not requiring the actual model to ensure resiliency during a cyber-attack.
- Can be implemented in a decentralized manner ensuring model privacy.
- Proposed GAN based method work effectively to both localize and mitigate both FDI and DoS cyber-attacks.
- **Future Work** : Large changes in operating points, non-linear controller, IBRs

References

1. G. Liang, J. Zhao, F. Luo, S. R. Weller and Z. Y. Dong, "A Review of False Data Injection Attacks Against Modern Power Systems," in IEEE Transactions on Smart Grid, vol. 8(4), July 2017
2. I. Goodfellow, et al., Deep Learning. The MIT Press, 2016.
3. S.M. Dibaji, et al., "Delay-Aware Control Designs of Wide-Area Power Networks", IFAC-PapersOnLine, vol. 50(1), 2017.
4. X. Deng et al., "Deep Learning Model to Detect Various Synchrophasor Data Anomalies", IET Generation, Transmission &; Distribution, 2020.

Published at

IEEE PES Power & Energy Society®

Partially funded by

NSF