Physics-Informed Deep Neural Network Method for Limited Observability State Estimation

Jonathan Ostrometzky[1], Konstantin Berestizshevsky[1],
Andrey Bernstein[2], Gil Zussman[3]

[1] Tel-Aviv University; [2] National Renewable Energy Lab (NREL); [3] Columbia University

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

TEL AVIV UNIVERSITY

NREL NATIONAL RENEWABLE ENERGY LABORATORY

Power Grid Resilience

Physical Attacks/Failures

Cyber Attacks/Failures

Control

Data
Power

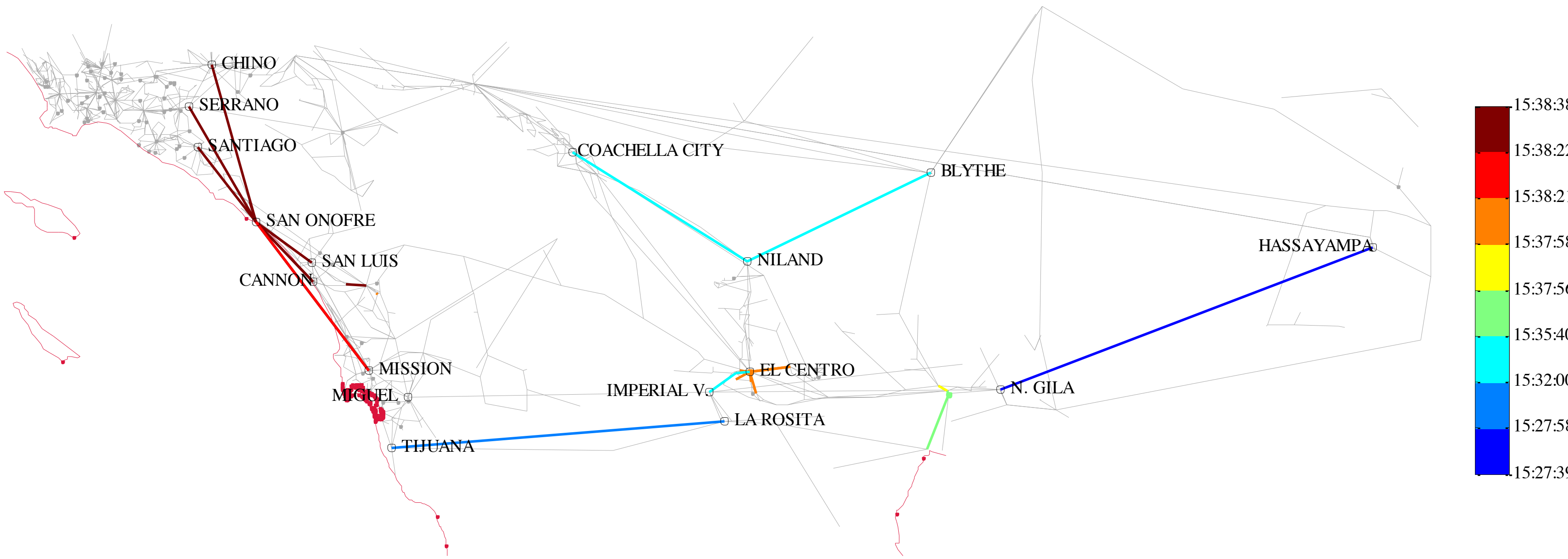
Power Grid
Physical Infrastructure

Communication networks

Supervisory Control and Data
Acquisition (SCADA) system

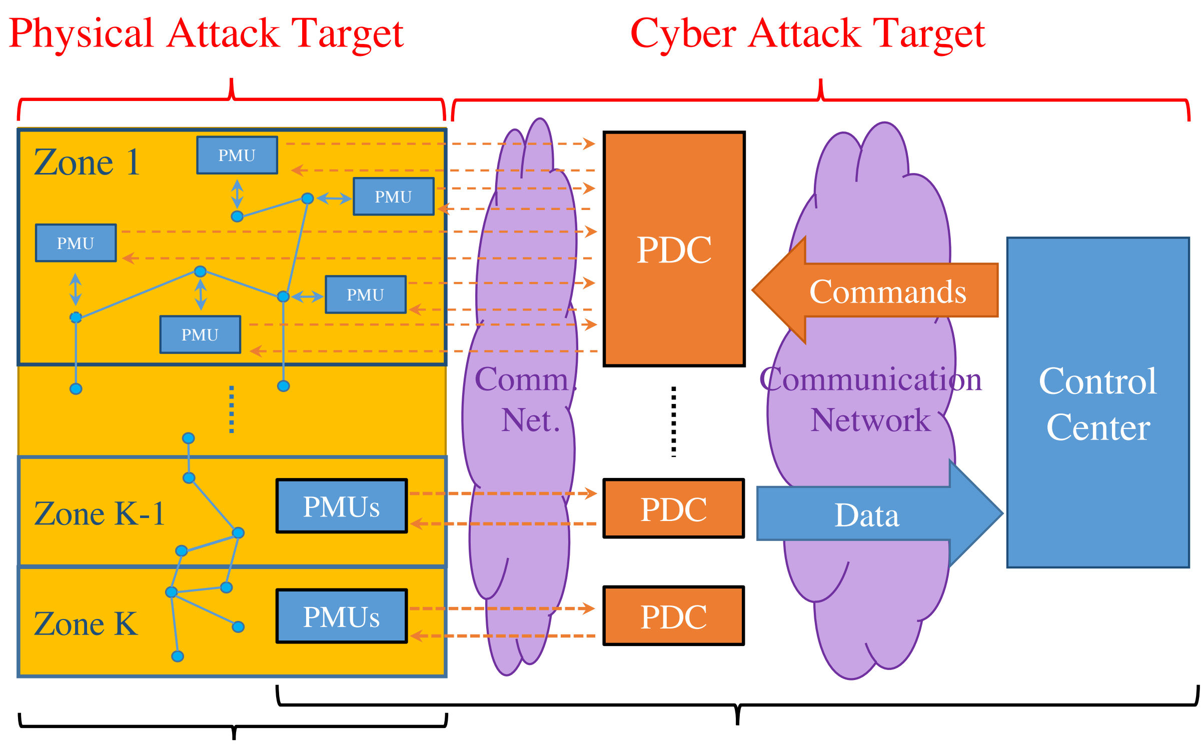
Due to the potential for cascading failures
a clever cyber-attack can be amplified by the grid operators

San Diego Blackout, Sept. 2011 – Human Error



"Ideally" a cyber attack would cause the operators to make a human error

Simplistic view of a Power Grids



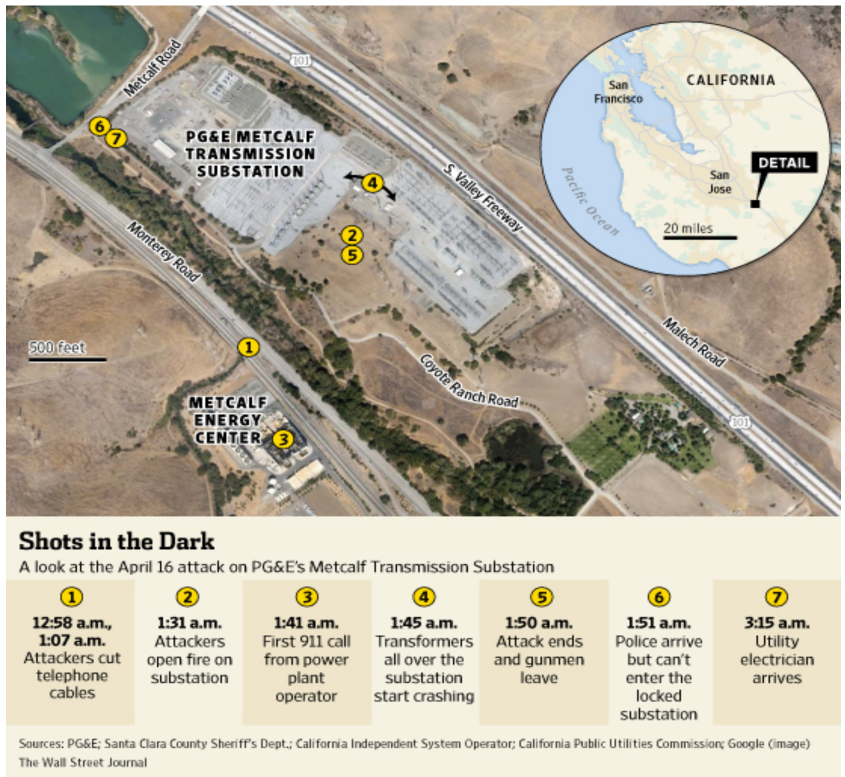
Power Grid
Physical Infrastructure

Supervisory Control and Data
Acquisition (SCADA) system

PMU: Phasor Measurement Unit
PDC: Phasor Data Concentrators

Physical Attack in San Jose (Apr. 2014)

“A sniper attack in April 2014 that knocked out an electrical substation near San Jose, Calif., has raised fears that the country's power grid is vulnerable to terrorism. ” –The Wall Street Journal



Cyber Attack in Ukraine (Dec. 2015)

Unplugged 225,000 people from the Ukrainian electricity grid

Before June 2015	June – December 2015	23 December 2015
• Extensive reconnaissance of distribution utilities' corporate networks • Spear phishing emails to executives to implant a variant of the Black Energy malware • Theft of credentials for accessing SCADA systems	• Exploration of SCADA systems and attack planning • Development of malicious firmware for substation equipment	• Synchronized, remote operation of substation breakers causes blackout • Control-room backup power supplies are remotely disconnected • Phone jamming attack keeps operators unaware • Malware destroys data needed to operate equipment

Source: ICS-CERT, SANS Institute

Cyber Attack in Ukraine (Dec. 2015)

Unplugged 225,000 people from the Ukrainian electricity grid

Before June 2015	June – December 2015	23 December 2015

Source: ICS-CERT, SANS Institute

Transmission Grid - State Recovery after a Cyber-Physical Attack

- State recovery under the DC model
- State recovery in the presence of measurement noise and uncertainty
- State recovery under the AC model
- Attack identification when the affected area is unknown

[1] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, "REACT to Cyber Attacks on Power Grids," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 3, pp. 459–473, Sept. 2019.

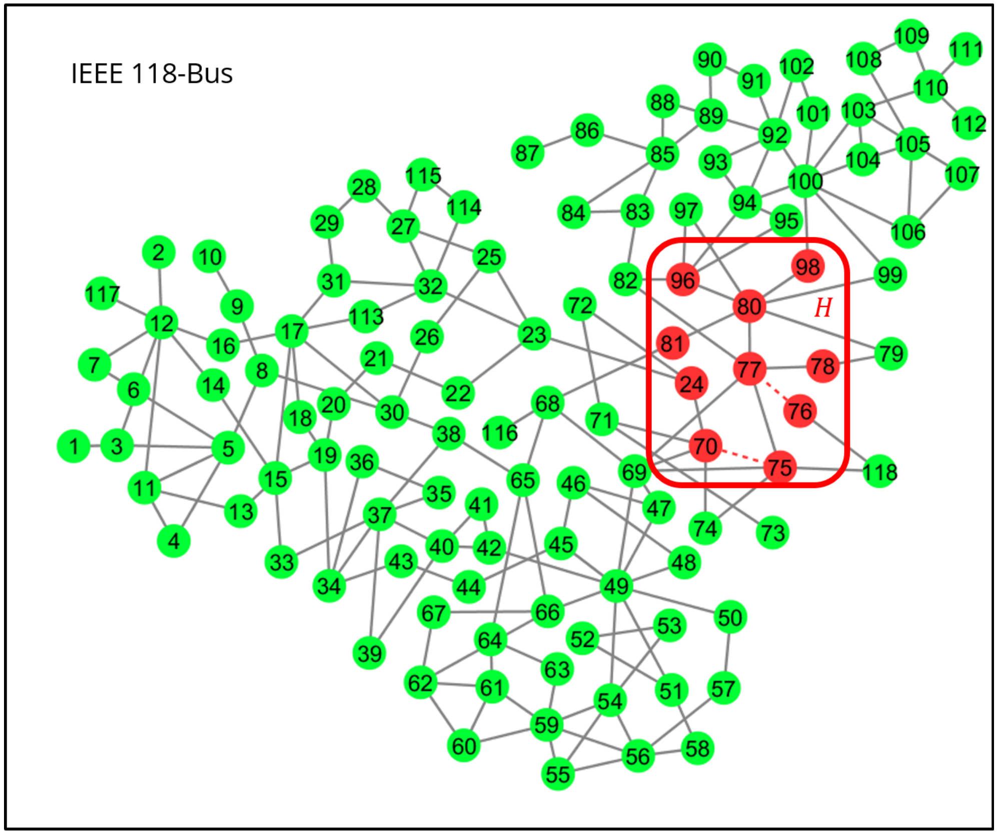
[2] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, "EXPOSE the Line Failures following a Cyber-Physical Attack on the Power Grid ," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 451–461, Mar. 2019.

[3] Saleh Soltan and Gil Zussman, "Power Grid State Estimation after a Cyber-Physical Attack under the AC Power Flow Model," *Proc. IEEE PES-GM'17*, 2017.

[4] Saleh Soltan, Mihalis Yannakakis, Gil Zussman, "Power grid state estimation following a joint cyber and physical attack," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 499–512, Mar. 2018.

Attack Identification when the Affected Area is Unknown

Detect the line failures as well as the attacked area H after a cyber-physical attack



[1] S. Soltan, M. Yannakakis, and G. Zussman, "REACT to cyber attacks on power grids," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 3, pp. 459–473, Sept. 2019.

Location Unknown - Cyber Attacks

Physical attack - some lines in the area fail

Cyber attack:

- Data distortion
- Data Replay

$\vec{\theta}^{\star}$ is the observed phase angles vector after the attack which is different from the actual $\vec{\theta}'$

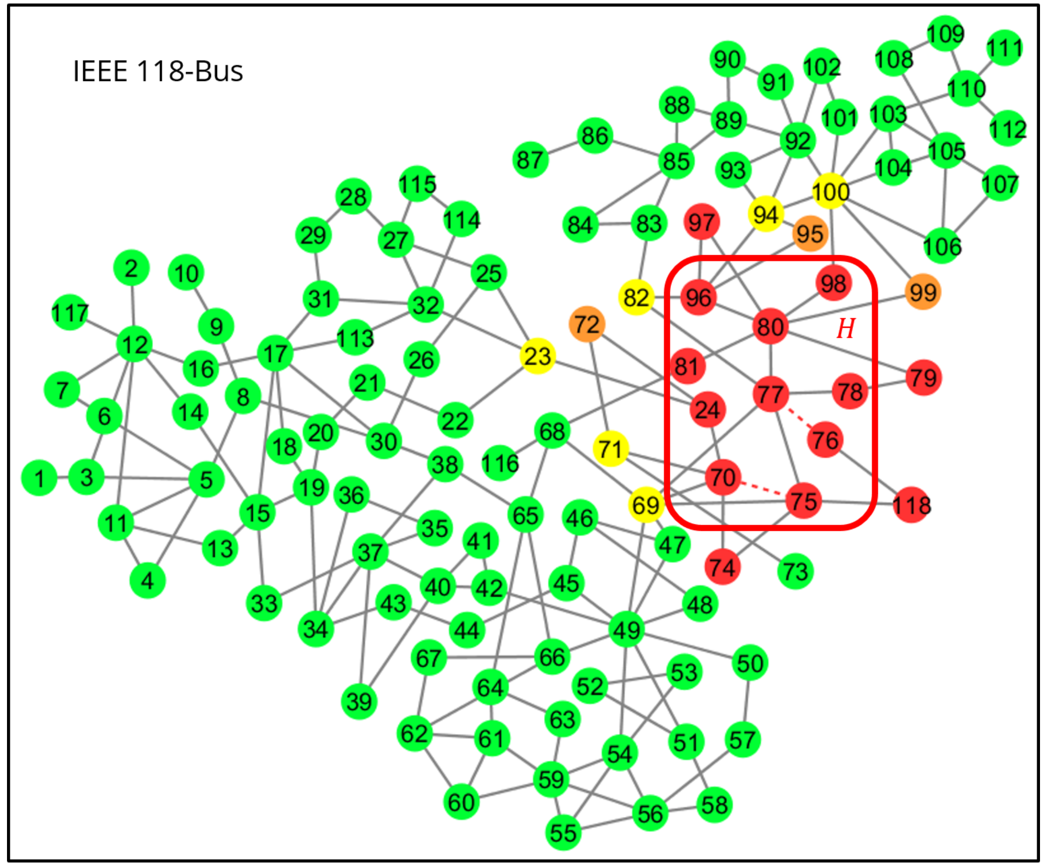
NP-Hard to detect the set of line failures (even if the attack area is known and even under the DC approximation)

Approximate solutions

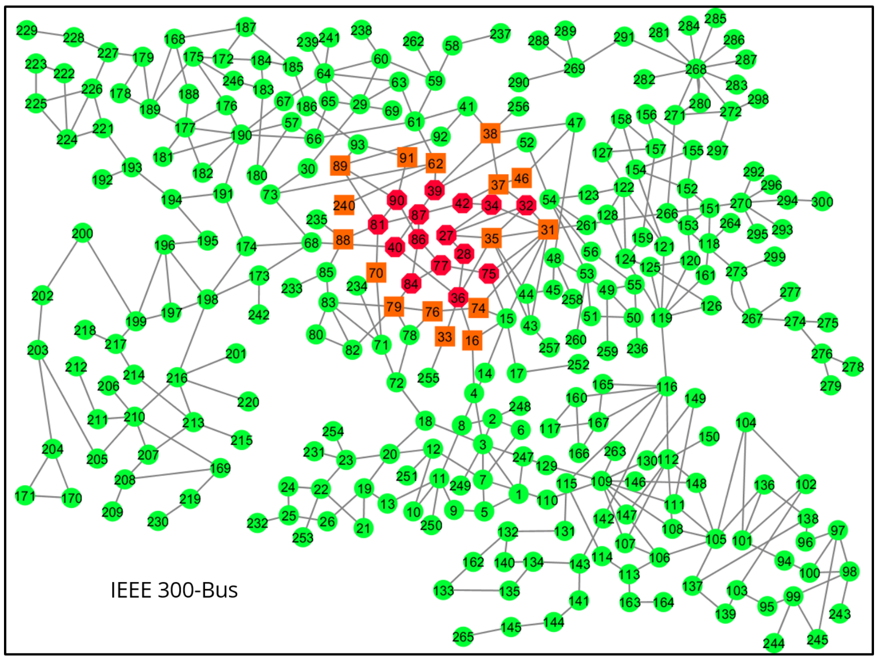
Example

Approximately detect the attacked area in 3 steps

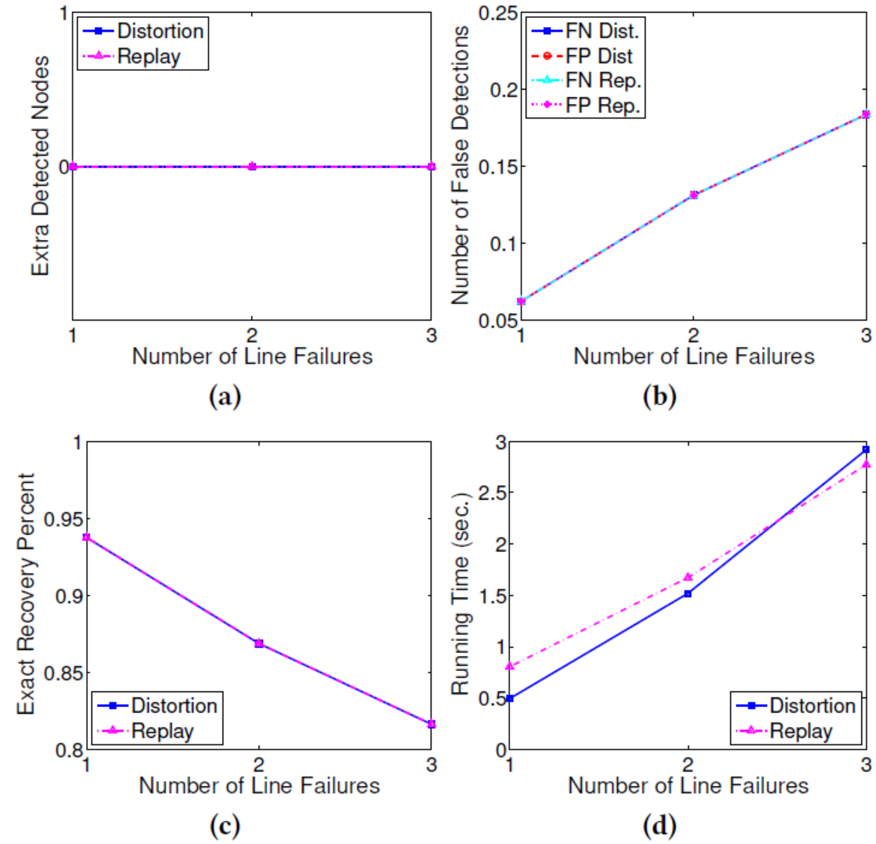
Identify line failures with some confidence



Performance - Small Area (15 nodes)

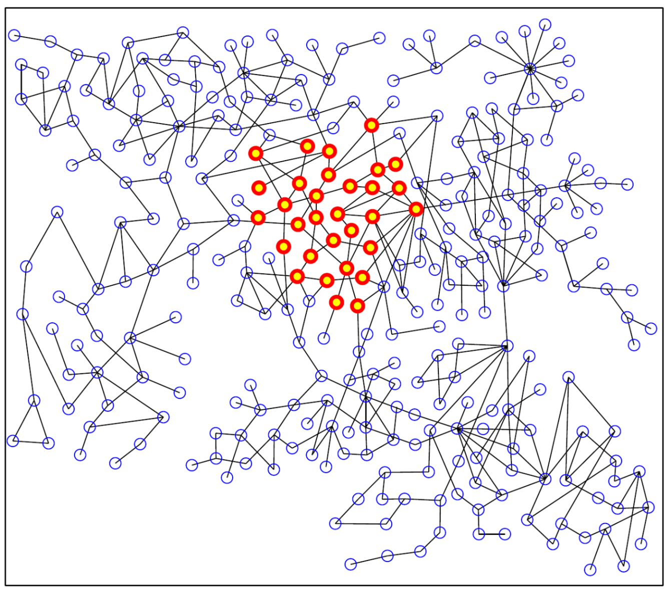


100 1,2,3-line failure samples

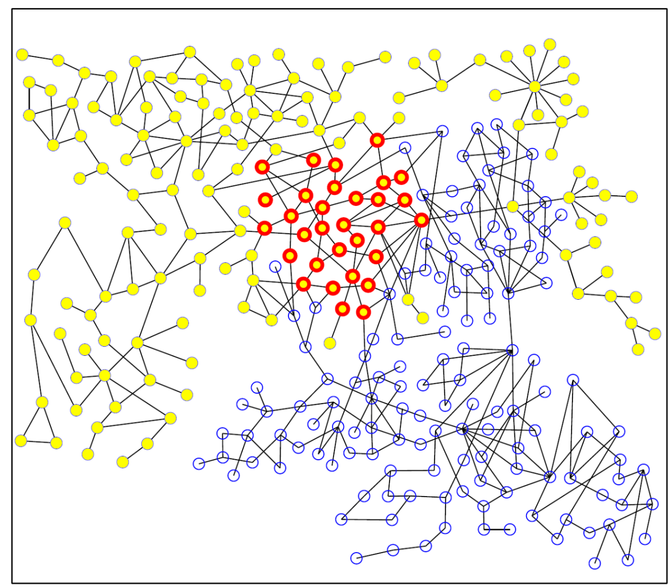


Data Distortion vs. Data Replay

Difficulty in detecting the attacked area after a data replay attack

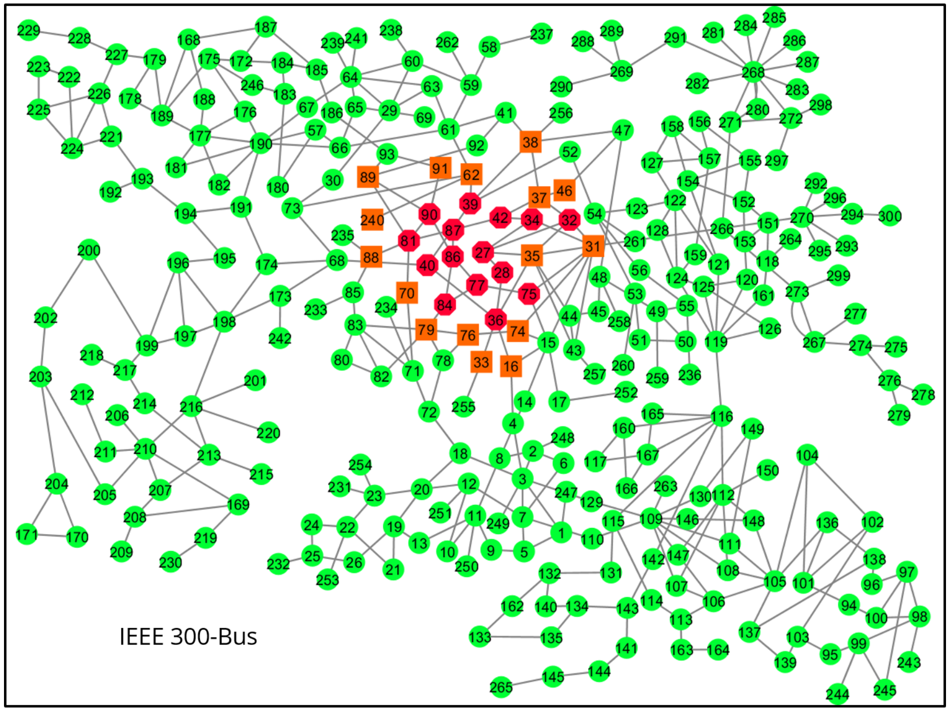


(a) Data Distortion Attack

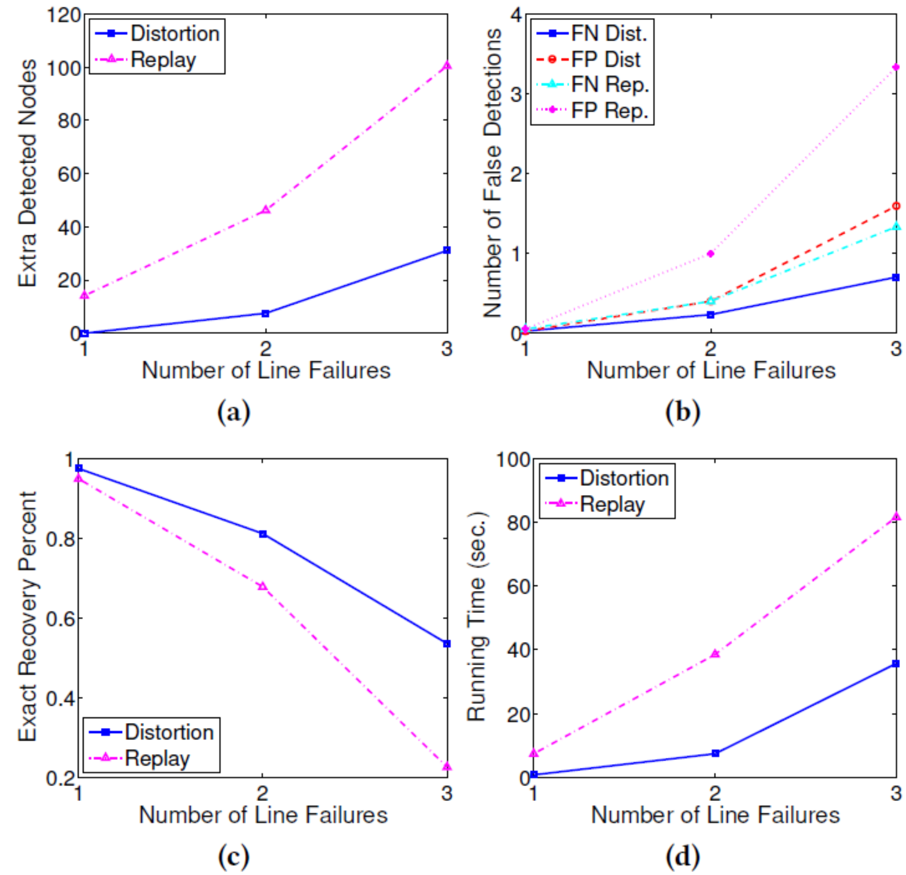


(b) Data Replay Attack

Performance - Large Area (31 nodes)

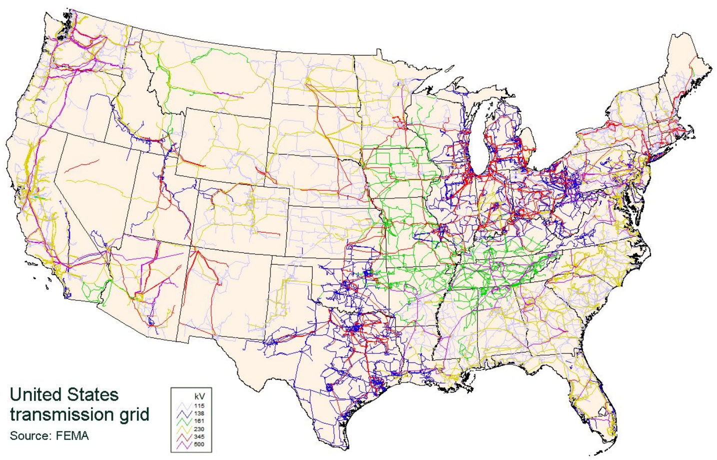


100 1,2,3-line failure samples



From Transmission to Distribution

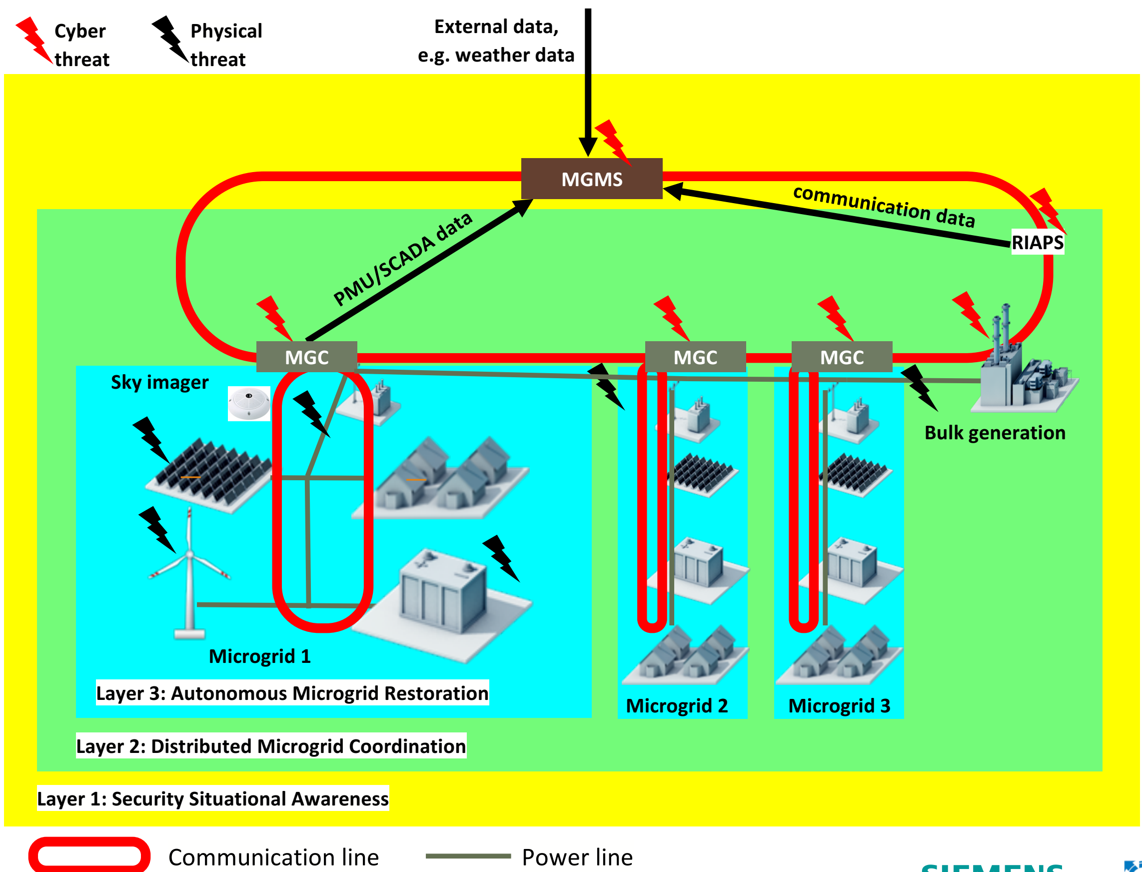
Most of the research in this field has focused on the Transmission grid



The Distribution grid, on the other hand, suffers from under-observability even when not attacked

AURORA (**AU**tonomous and **R**esilient **O**peration of energy systems with **R**enewa**A**bles), PI: Ulrich Muenz

Develop and demonstrate a 3-layer protection scheme against cyber and physical threats



Project objectives

Layer 1: **Security Situational Awareness**

1. Assess and optimize resiliency against physical threats
2. Detect and localize cyber attacks

Global restoration ↑ ↓ Loss of MGMS, Loss of communication

Layer 2: **Distributed Microgrid Coordination**

3. Continuity of service after attack on control center or communication system

Local restoration ↑ ↓ Blackout

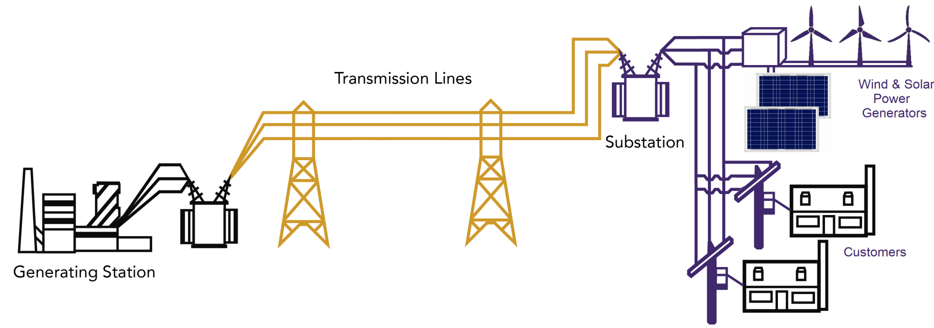
Layer 3: **Autonomous Microgrid Restoration**

4. Fast restoration after blackouts
5. Robust parallel grid-forming inverters

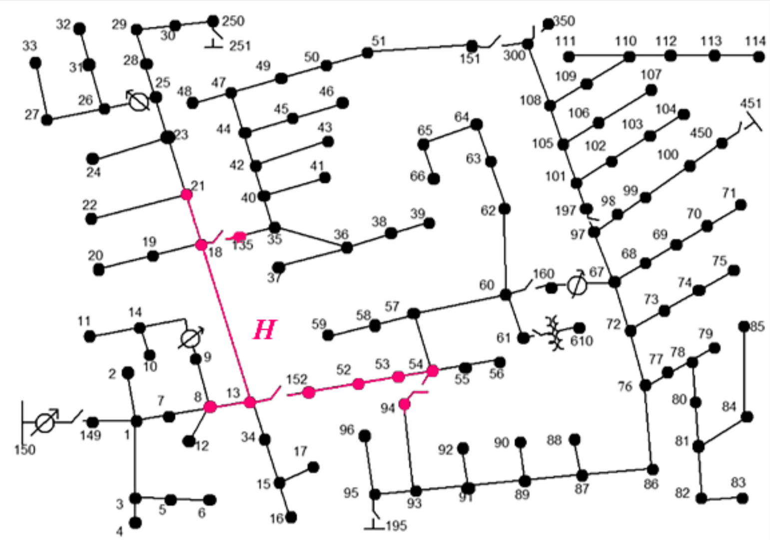
MGMS: Microgrid Management System; MGC: Microgrid Controller; RIAPS: Resilient Information Architecture Platform for the Smart Grid

Distribution Grid – Partial Observability

- Distribution grid
  - Natural fluctuations
  - Limited observability
  - Sensors are becoming more pervasive but still "fragile"
  - DC approximation does not hold



- Given:
  - Historical data on voltage and power
  - Partial real-time power measurements (e.g., due to cyber attacks)
- Power-flow equations may be under-determined
  - Model-driven approach may fail
- Objective: prediction of voltages
- Method: Incorporate the physical model of the power-flow equations into the Deep Learning training
  - Hybrid **model and data driven approach**

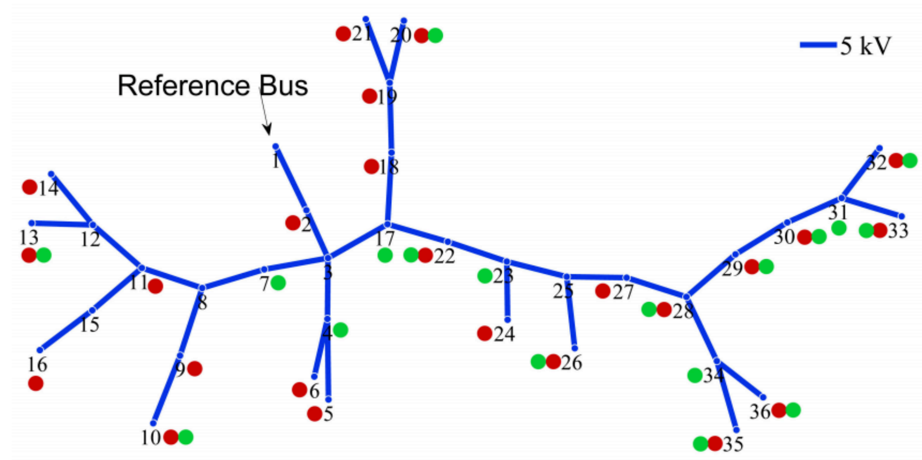


[1] Jonathan Ostrometzky, Konstantin Berestizshevsky, Andrey Bernstein, and Gil Zussman, "Physics-Informed Deep Neural Network Method for Limited Observability State Estimation", arXiv:1910.06401v2 [eess.SY], Feb. 2020

Objective and Assumptions

Goal:

- Accurate estimation of the distribution grid state



Assumptions:

- The distribution grid is affected, and becomes under-observable
- The Power-Flow Equations cannot be solved

Method:

Historical data: $\{\hat{v}(\tau), \hat{s}(\tau)\}_{\tau=t-T-1}^{\tau=t-1}$

Real-time measurements: $\hat{s}_j(t)$ for some j

Deep Learning

Power-flow

Predict: $\hat{v}_i(t)$ for all i

Evaluation: numerical

Related Work

- Distribution system state estimation [Chen et al. 2019], [Primadianto and Lu, 2017]
- Matrix completion techniques [Donti et al., 2018], [Genes et al., 2019], [Miao et al., 2019]
- Machine learning tools for distribution system state estimation [Bhela et al., 2018], [Jiang and Zhang, 2016]
- Physics-informed deep learning methods [Zamzam and Sidiropoulos, 2019], [Hu et al., 2020], [Singh et al., 2020], [Zhang et al., 2019]
- Hybrid machine learning models in other domains [Zhu et al., 2020]

Sudden Failure State Estimation (SFSE)

Problem formulation

Specification 1. ***SFSE***$(T, t, N_s(t), N_v(t))$

Inputs:

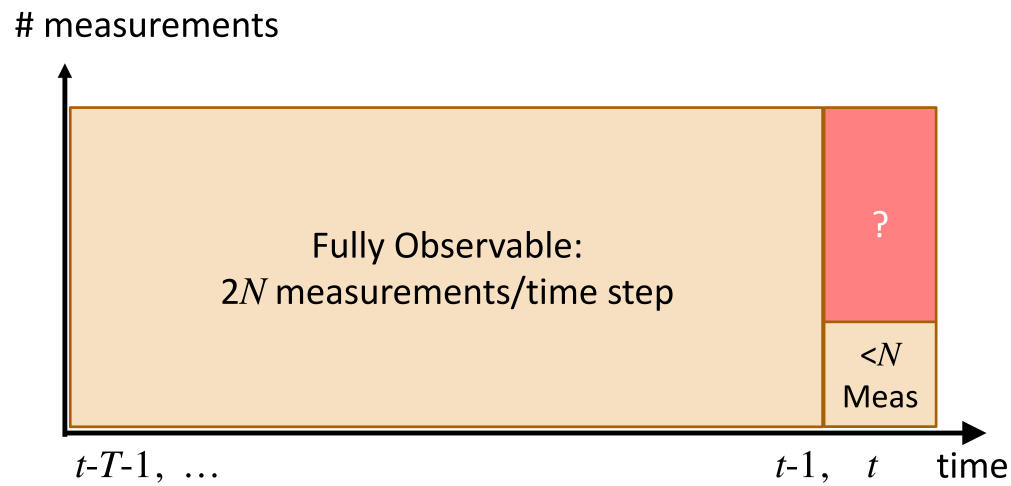
full history before time t: $\{\underline{s}(\tau), \underline{v}(\tau)\}_{\tau=t-1-T}^{t-1}$

partial measurements set at time t:

$[s_1(t), \ldots, s_{N_s(t)}(t)]; [v_1(t), \ldots, v_{N_v(t)}(t)]$.

Output:

voltage estimation at time t: $\hat{\underline{v}}(t)$.



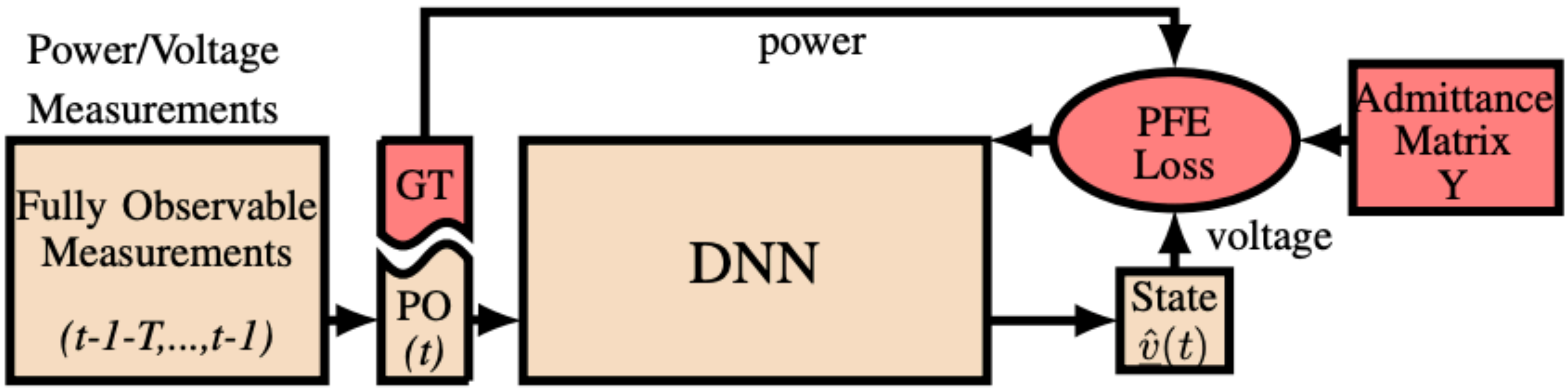
For different levels of Observability at time (t), defined as $\mathcal{O}(N_s(t), N_v(t))$ for a distribution network of N nodes:

For any $N_s(t) \in \{0, \ldots, N\}$; $N_v(t) \in \{0, \ldots, N\}$, let $\mathcal{O}(N_s(t), N_v(t)) \triangleq \dfrac{N_s(t) + N_v(t)}{2N}$

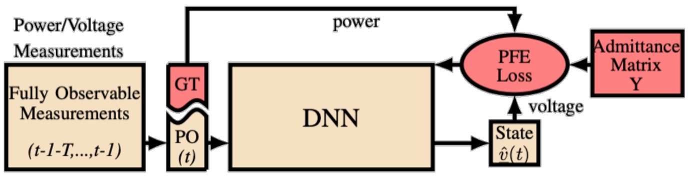
The Power-Flow Eqautions cannot be directly solved if the observability level drops below 50%

→ $\mathcal{O}(N_s(t), N_v(t)) < 50\%$ defines a low-observable, under-determined scenario

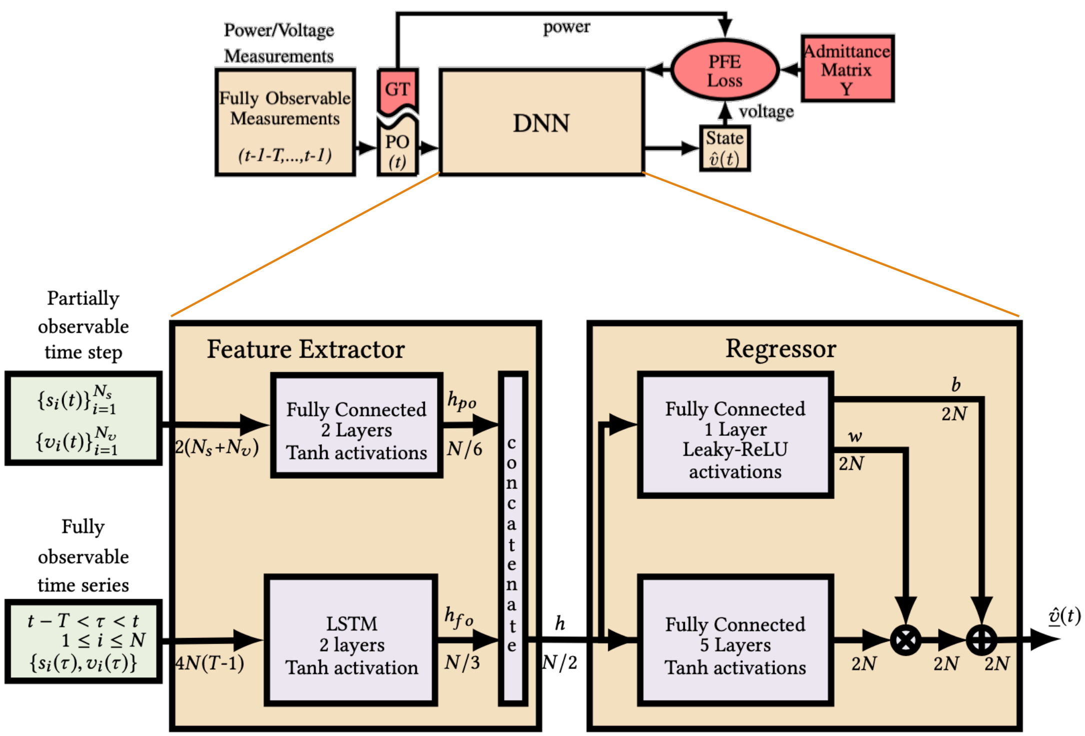
Power Flow-informed Deep Neural Network (DNN)



Power Flow-informed Deep Neural Network (DNN)



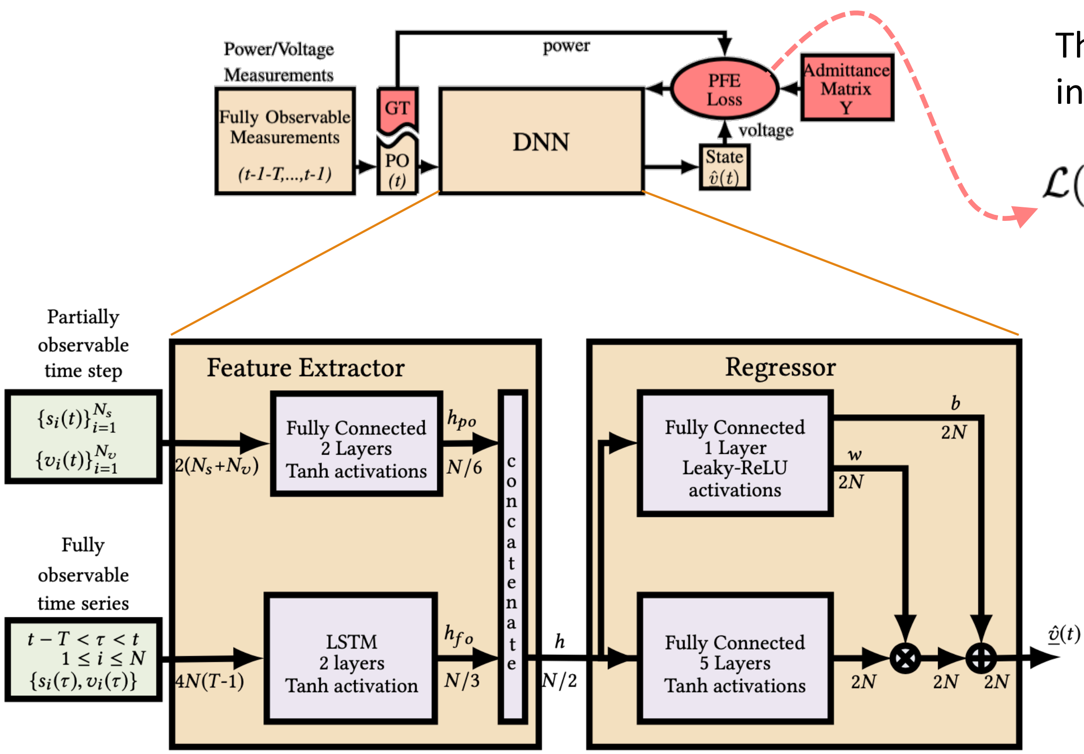
Power Flow-informed Deep Neural Network (DNN)



N – The number of nodes
N_s – The number of nodes that report the complex power values
N_v – The number of nodes that report the complex voltage values

Inputs:
N time-series [t-T,…t-1] of the complex voltage values
N time-series [t-T,…t-1] of the complex power values
N_s < N complex power values (for time index t)
N_v < N complex power values (for time index t)

Power Flow-informed Deep Neural Network (DNN)



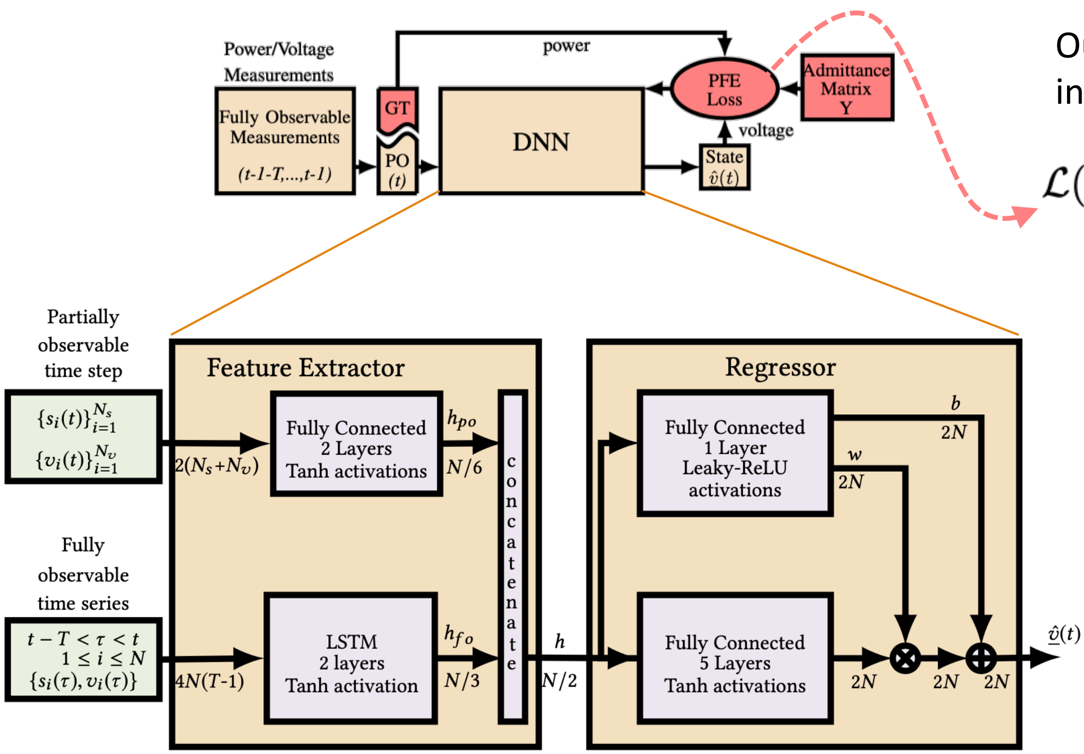
The Loss function acts as a regularizer for the DNN, incorporating the AC Power-Flow Equations

$$\mathcal{L}(\underline{s},\underline{v},\hat{\underline{v}},Y,\lambda) = ||\underline{v}-\hat{\underline{v}}||^2 + \lambda||\underline{s} - \text{diag}(\hat{\underline{v}})Y^*\hat{\underline{v}}^*||^2$$

N – The number of nodes
N_s – The number of nodes that report the complex power values
N_v – The number of nodes that report the complex voltage values

Inputs:
N time-series [t-T,…t-1] of the complex voltage values
N time-series [t-T,…t-1] of the complex power values
$N_s < N$ complex power values (for time index t)
$N_v < N$ complex power values (for time index t)

Power Flow-informed Deep Neural Network (DNN)

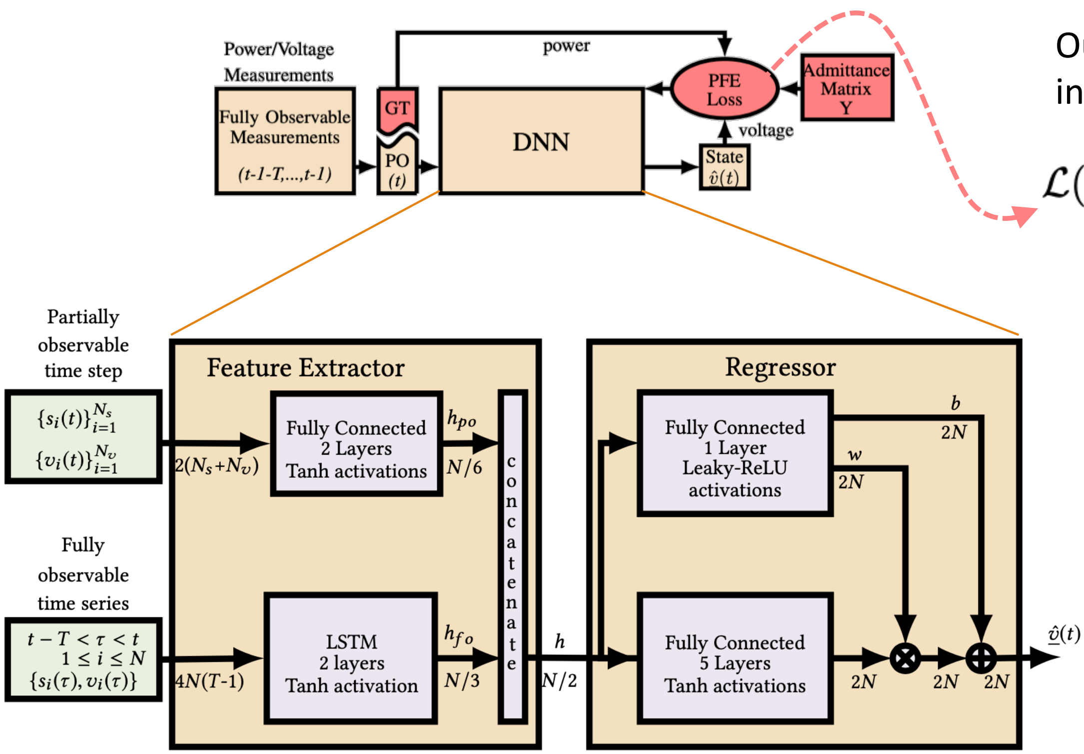


Our Loss function acts as a regularizer for the DNN, incorporating the AC Power-Flow Equations

$$\mathcal{L}(\underline{s},\underline{v},\hat{\underline{v}},Y,\lambda) = \boxed{||\underline{v}-\hat{\underline{v}}||^2} + \lambda||\underline{s} - \mathrm{diag}(\hat{\underline{v}})Y^*\hat{\underline{v}}^*||^2$$

MSE Term

Power Flow-informed Deep Neural Network (DNN)



Our Loss function acts as a regularizer for the DNN, incorporating the AC Power-Flow Equations

$$\mathcal{L}(\underline{s}, \underline{v}, \hat{\underline{v}}, Y, \lambda) = \boxed{||\underline{v} - \hat{\underline{v}}||^2} + \boxed{\lambda||\underline{s} - \text{diag}(\hat{\underline{v}})Y^*\hat{\underline{v}}^*||^2}$$

MSE Term

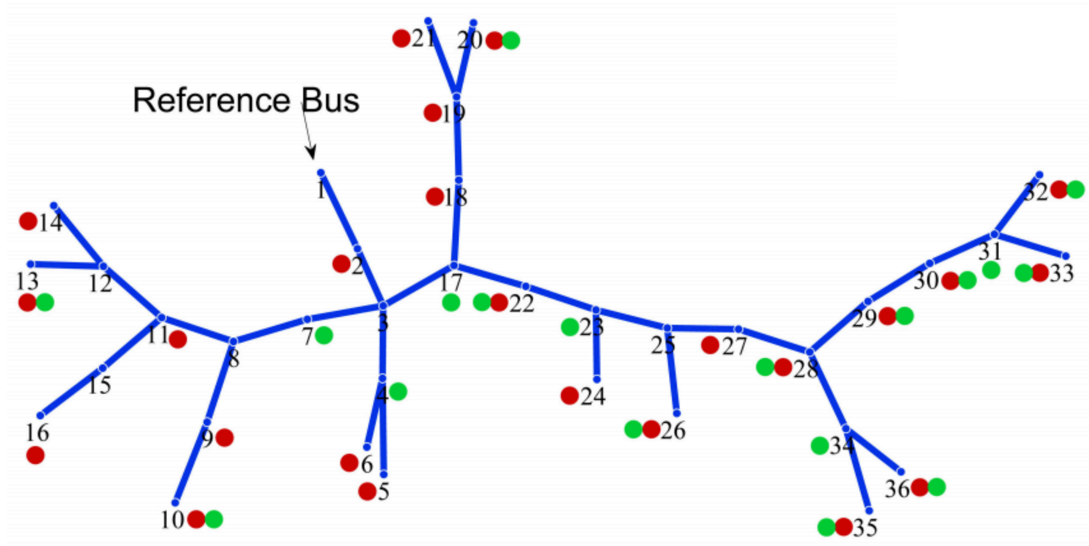
$0 \le \lambda \in \mathbb{R}$

+ Penalize infeasible power flow

Evaluation

Based on the IEEE-37 bus feeder

~50% of the buses inject power

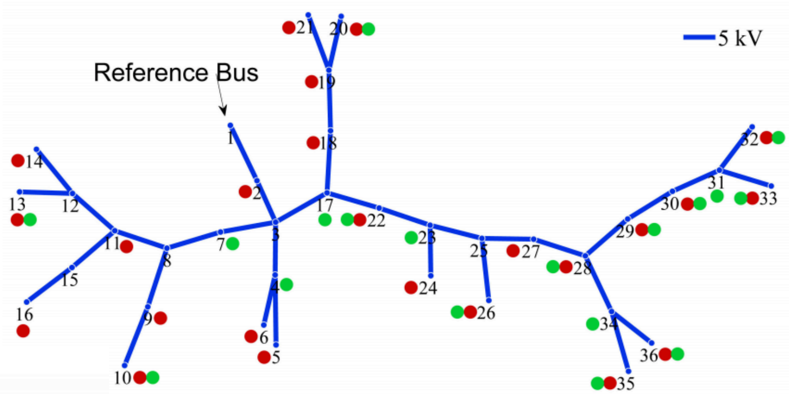


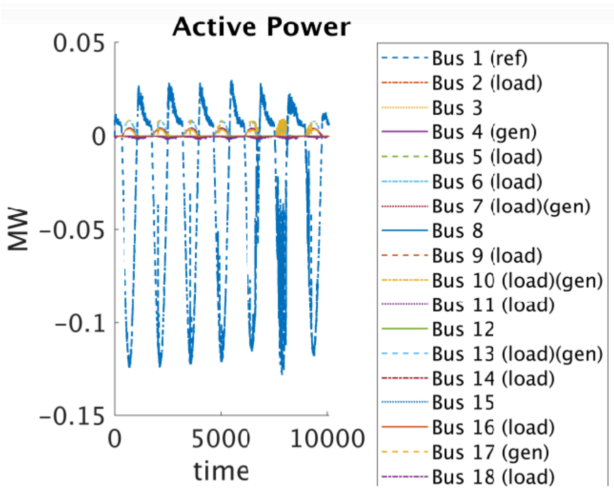
Available Data

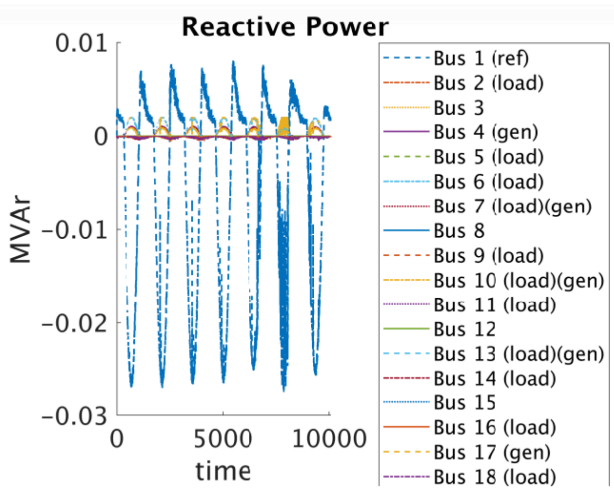
- NREL Provided us with **real distribution grid data**:
  - One photovoltaic panel production (active power) – sampling rate of 1 Hz
  - Eight real usage of houses (active power) – sampling rate of 1 Hz
- Processing
  - Randomly allocated to buses
  - Generated corresponding reactive power
  - Smoothed the data, using a moving-average 60-second window, and down-sampled
  - Used MATPOWER to solve the Power Flow Equations (AC model) and obtain voltages
- Overall, acquired a full week of data (~10,080 time-steps per time-series)
  - 90% of the T-long sequences used for training
  - The rest used for validation

Available Data - Power

- Arbitrarily assigned different nodes with data based on the real-world measurements provided by NREL



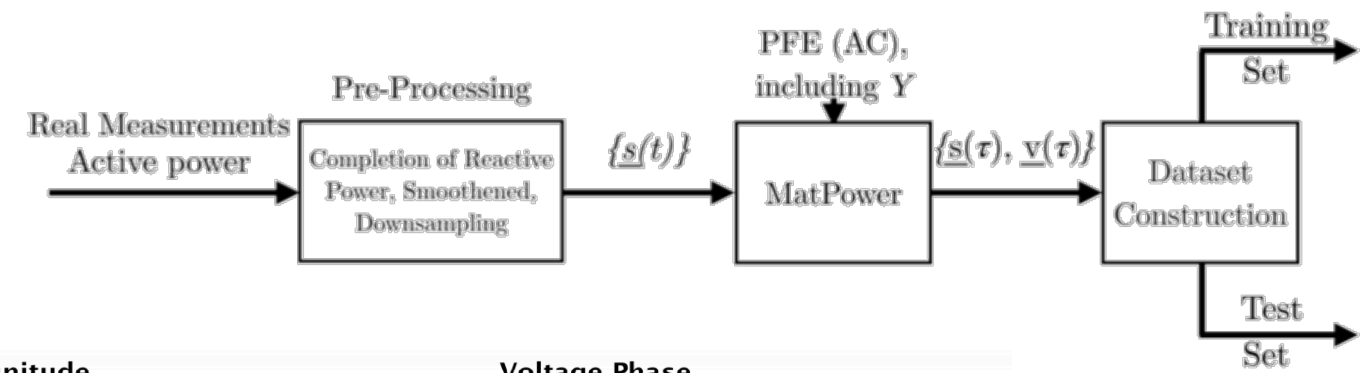


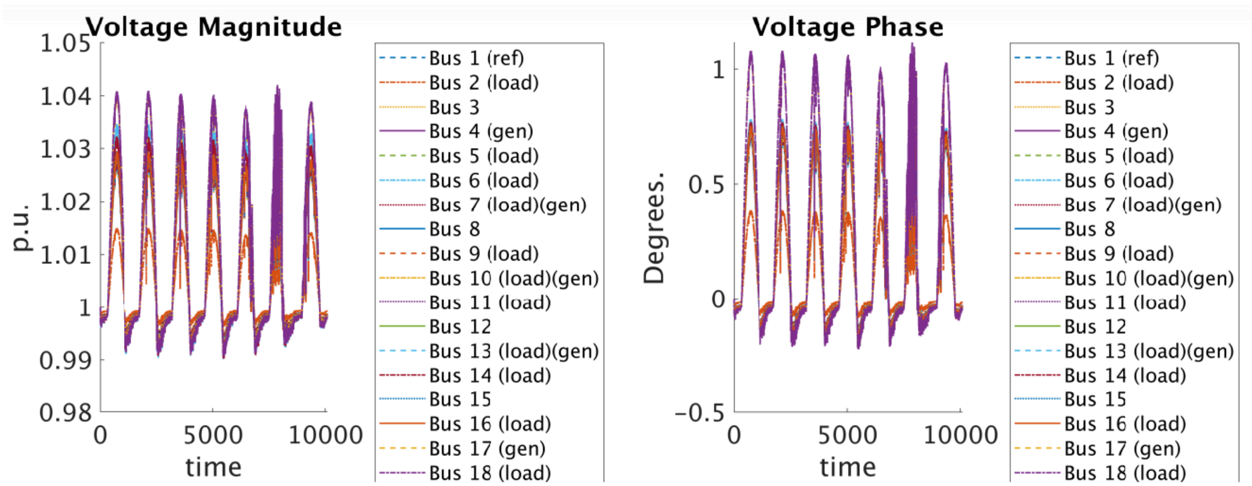


IEEE 37-node test feeder distribution grid
● represents a generator-node
● represents a load-node

Available Data

- Used MATPOWER to calculate the time-series of the complex voltages, which satisfies the Power-Flow Equations, to complete the dataset needed for training and validation

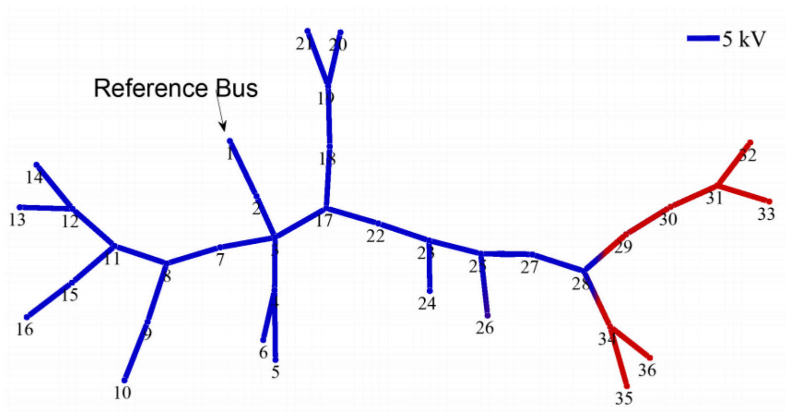




Training

- We trained the setup for **different levels of observability:** 49%, 39%, 25%, 17%, and 8%
- This mimics actual **attacks/malfunctions**

$$\mathcal{O}\big(N_s(t), N_v(t)\big) = \frac{N_s(t) + N_v(t)}{2N} = \frac{28 + 0}{36 \cdot 2} = 0.39$$

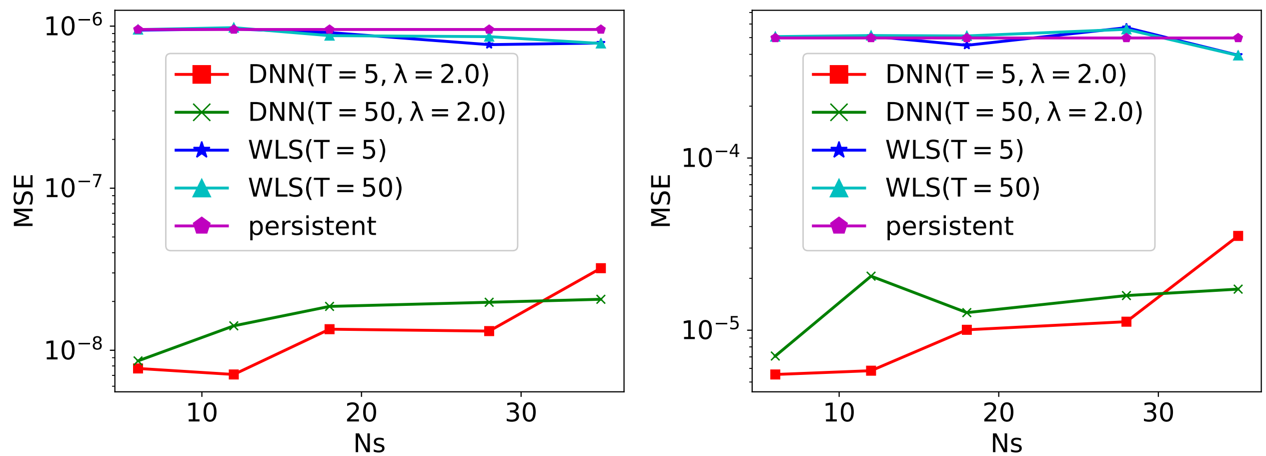


Example of an observability value of 39%:
- 0/36 voltages are known at time (t),
- 28/36 power-values are known at time (t).

- 90% used for training
- 10% used for validation

* We use 36 instead of 37 nodes since one of the nodes is a behind a transformator.

Numerical Results – Comparison with WLS and Sensitivity to T



(a) Magnitude$\{\underline{v}\}$ **(b) Angle$\{\underline{v}\}$**

The magnitude and the angle of the normalized Mean-Square-Error for the complex voltages time-series, compared with the Weighted Least Square Estimation

$$W_i \triangleq 1/\mathrm{std}\Big(\{s_i(\tau)\}_{\tau=t-1}^{t-1-T}\Big)$$

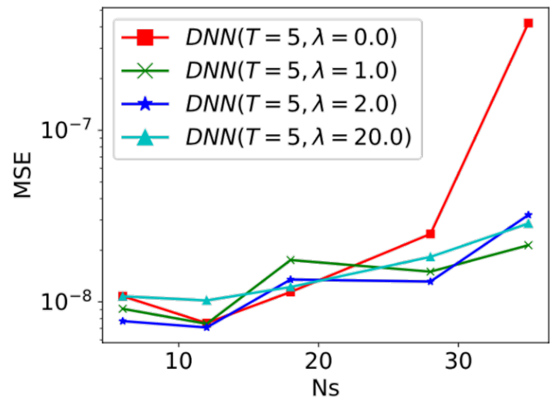
$$f_{Y,\hat{\underline{s}}}(\hat{\underline{v}}, i) \triangleq \hat{s}_i - \sum_{j=1}^{N} \hat{v}_i \cdot Y_{i,j} \cdot \hat{v}_j^*$$

$$\min_{\hat{\underline{v}}} F(\hat{\underline{v}}) := \frac{1}{2}\sum_{i=1}^{N} W_i(\Re\{f_{Y,\hat{\underline{s}}}(\hat{\underline{v}}, i)\}^2 + \Im\{f_{Y,\hat{\underline{s}}}(\hat{\underline{v}}, i)\}^2)$$

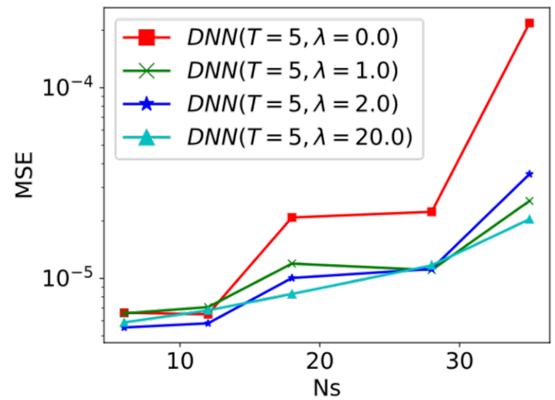
Numerical Results – Sensitivity to λ

$$\mathcal{L}(\underline{s}, \underline{v}, \hat{\underline{v}}, Y, \lambda) =$$

$$= \underbrace{||\underline{v} - \hat{\underline{v}}||^2}_{\text{MSE term}} + \underbrace{\lambda||\underline{s} - \mathsf{diag}(\hat{\underline{v}})Y^*\hat{\underline{v}}^*||^2}_{\text{Power-flow Equations regularizer term}}$$



$\mathbf{Magnitude}\{\underline{v}(t)\}$



$\mathbf{Angle}\{\underline{v}(t)\}$

The magnitude and the angle of the normalized Mean-Square-Error for the complex voltages time-series

Quick Detour – COSMOS – Potential Testbed for Studying Interdependencies

- NSF Platforms for Advanced Wireless Research (PAWR) - City Scale Wireless Testbed
- COSMOS (Rutgers, Columbia, NYU, NYC) - A community infrastructure in Upper Manhattan
- Potential testbed for power/communication interdependencies

Project Vision

Latency and **compute power** are two important dimensions

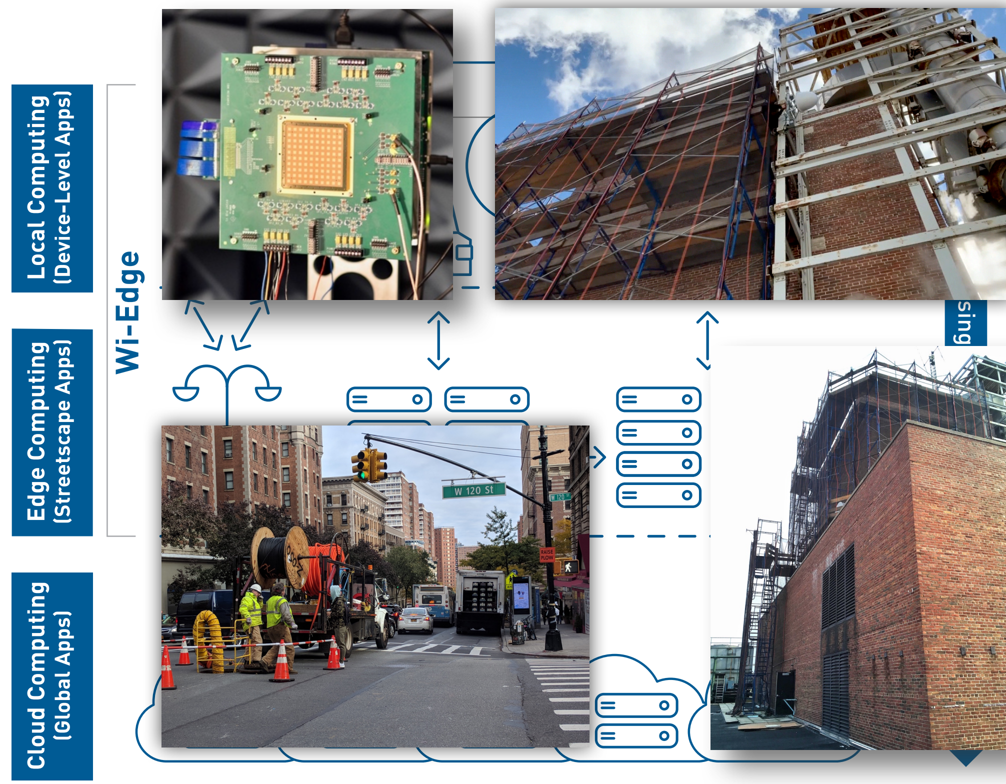
Edge computing is an enabler for real-time applications (autonomous vehicle, etc.)

Objective: Real-world investigation of urban environments with

- Ultra-high bandwidth (~Gb/s)
- Low latency (<5 ms)

Enablers:

- Antennas on lightpoles
- 10s of 64 element mmWave arrays
- 10s of miles of Manhattan dark fiber
- B5G edge cloud base stations
- Programmable

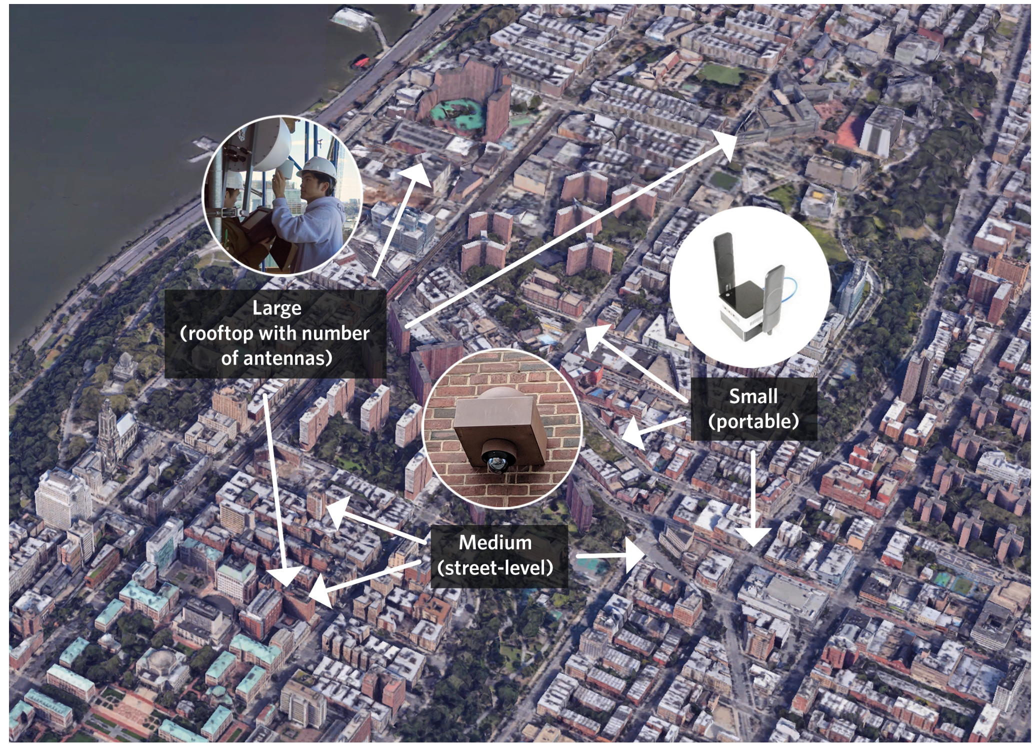


- Ultra-high bandwidth, low latency, and powerful edge computing will enable new classes of real time applications
- NSF supplements to run experiments in the testbed (DCL-20-046) + EC NGIAtlantic grants

Columbia Electrical Engineering – Postdoc Positions

Power grid resilience (with Prof. James Anderson)

Beyond 5G wireless – design and evaluation in the NSF PAWR COSMOS city-scale testbed



gil.zussman@columbia.edu wimnet.ee.columbia.edu

Summary and Ongoing Work

- Expanded previous work on transmission systems and static model to distribution system with streaming data
- Developed a hybrid model and data driven approach to recover missing data in distribution grid
- Has a “black box” nature but takes the power flow equations and system parameters into account
- Showed that it works well with real-world data
- Future/ongoing work:
  - Improve the DNN to accommodate a general training set, rather than a training set per scenario
  - Evaluate the method with the Holly Cross Energy distribution grid as part of the AURORA project
  - Extend to false data injection

[1] Jonathan Ostrometzky, Konstantin Berestizshevsky, Andrey Bernstein, and Gil Zussman, “Physics-Informed Deep Neural Network Method for Limited Observability State Estimation”, arXiv:1910.06401v2 [eess.SY], Feb. 2020

Acknowledgments

[1] Jonathan Ostrometzky, Konstantin Berestizshevsky, Andrey Bernstein, and Gil Zussman, "Physics-Informed Deep Neural Network Method for Limited Observability State Estimation", arXiv:1910.06401v2 [eess.SY], Feb. 2020

This work was authored in part by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

This work was supported in part by the Laboratory Directed Research and Development (LDRD) Program at NREL, by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technology Office, ASSIST Initiative award number DE-EE0008769, and by DTRA grant HDTRA1-13-1-0021.