FOA 1861 FINAL PROJECT BREIFING
BIG DATA ANALYSIS OF SYNCHROPHASOR DATA

Discovery of Signatures, Anomalies, and Precursors in Synchrophasor Data with Matrix Profile and Deep Neural Networks

Nanpeng Yu
University of California, Riverside
nyu@ece.ucr.edu

March 7, 2022

Outline

- Project Overview
- Technical Accomplishments
- Summary of Experimental Results on Testing Dataset
- Value of Work
- Readiness for Commercialization
- Readiness for ML & BD Analytics
- Lessons Learned & Next Steps
- Publications and Presentations

Project Overview

- Project Team
  - UCR, EPG, Michigan Tech
- Project Objectives
  - Develop physics-informed machine learning algorithms using real-world PMU data to enhance power system reliability.
  - Accelerate future research and development of data-driven algorithms by creating a synthetic PMU dataset and an event signature library of bulk power systems.
  - Create prototype systems that implement the proposed machine learning algorithms.
- Significance and Impacts
  - Developed a suite of physics-informed machine learning algorithms for PMU data analytics
  - Enhanced the reliability of bulk power systems with data-driven algorithms and prototype systems
  - Advanced scientific knowledge of physics-based machine learning

Technical Accomplishments

- PMU Data Quality Improvement
  - Online PMU Missing Value Replacement via Event-Participation Decomposition
- Power System Event Detection
  - Graph Signal Processing-based Event Detection
  - Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms
  - Power System Event Detection with Bidirectional Generative Adversarial Network
- Power System Event Classification
  - Deep Neural Network-based Power System Event Classification
  - Classify Power System Event with a Small Number of Training Labels with Transfer Learning
- Power System Dynamic Parameter Estimation
  - Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations
- Synthetic Power System Event Data Creation
  - pmuBAGE: The Benchmarking Assortment of Generated PMU Events
- Power System Event Signature Library
  - A Dynamic Behavior-based Bulk Power System Event Signature Library with Empirical Clustering

PMU Data Quality Improvement

B. Foggo and N. Yu, "Online PMU Missing Value Replacement via Event-Participation Decomposition," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 488-496, Jan. 2022.

- Motivation
  - Failures in PMUs, phasor data concentrators, and communication links lead to missing PMU data
  - Problematic for data-driven applications that depend on streaming PMU data
- Main Idea
  - Event-participation decomposition model that decomposes an event into
    - A non-dynamic component represents the amount of participation of PMUs have in various disturbances.
    - A dynamic component represents the magnitudes of various disturbances.
- Key Results
  - The online SPIKE-P algorithm has substantially lower errors than the state-of-the-art algorithms.

BUS	Base	EnCorr	OLAP	SPIKE-P
Time (s)	**0.005**	0.303	0.067	0.038
VM (%)	2.449	2.449	1.309	**1.082**
IM (%)	68.792	56.531	21.393	**18.995**
P (%)	27.413	22.692	7.888	**7.225**
Q (%)	28.594	24.442	17.580	**10.937**

TABLE II: Average MAPEs over Bus Event data.

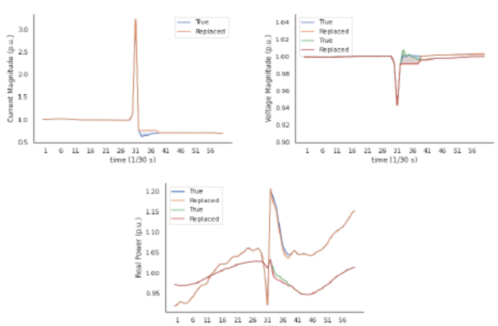


Fig. 4: SPIKE-P replacement on a Generator Event Sample.

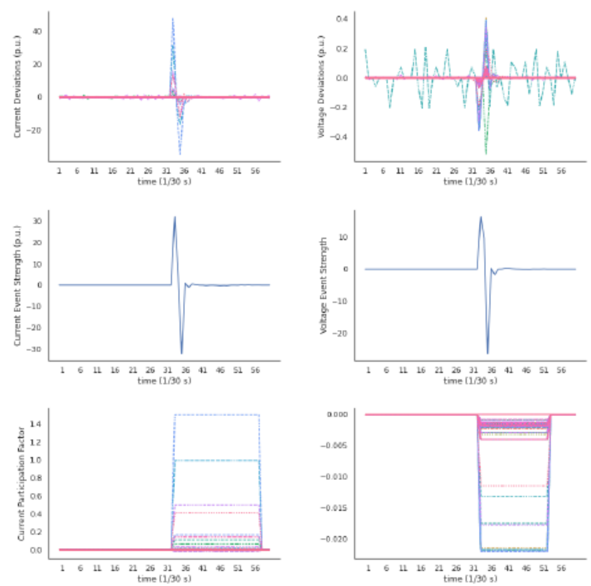


Fig. 1: Bus event sample decomposition

Power System Event Detection
Graph Signal Processing

J. Shi, B. Foggo, X. Kong, Y. Cheng, N. Yu, and K. Yamashita "Online Event Detection in Synchrophasor Data with Graph Signal Processing," *IEEE SmartGridComm*, pp. 1-7, 2020.

- Motivation
  - Timely detection of abnormal events can help operators take corrective control actions.
- Main Idea
  - Encode spatial and temporal correlations of streaming PMU data in the weighted adjacency matrix and graph Laplacian of the product graph.
  - Detect abnormal events using graph signal processing techniques.
- Key Results
  - Scalable and computationally efficient with linear time complexity and decent accuracy.

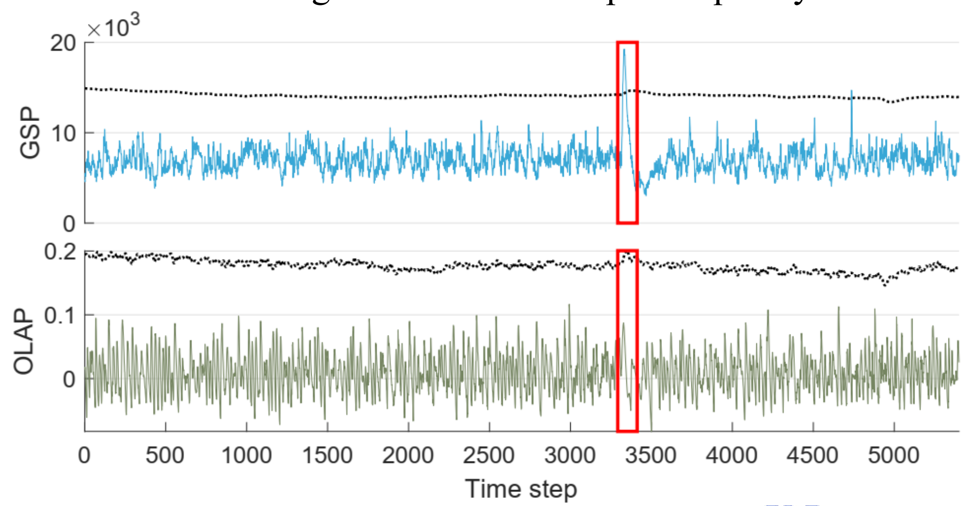
Comparison of F1 Scores

Method	GSP	OLAP
Category 1	0.7692	0.9
Category 2	1	0.8889
Category 3	0.8889	0.75
All Events	0.8750	0.8519

Scalability Test

Number of PMUs	30	60	90	120
Runtime	3.01 s	4.45 s	5.80 s	6.95 s

Abnormal event indicators of the GSP based approach and the OLAP algorithm for the sample frequency event



Voltage Event Detection
Optimization with Structured Sparsity-Inducing Norms

X. Kong, B, Foggo, and N. Yu, “Online Voltage Event Detection Using Synchrophasor Data with Structured Sparsity-Inducing Norms,“ to appear in *IEEE Transactions on Power Systems*, 2022.

- Motivation
  - Real-world PMU data matrices exhibit a unique row-sparse structure when the low-rank component is stripped away during voltage events.
- Main Idea
  - Decompose PMU data matrix: low-rank matrix, row-sparse event-pattern matrix & a noise matrix
  - Extract anomaly features from the low-rank matrix and the row-sparse event-pattern matrix.
- Key Results
  - Online, lower computation time, higher accuracy and scalability than state-of-the-art benchmark.

TABLE III
F SCORES OF THREE ALGORITHMS ON THE TESTING DATASET

Statistics	OLAP	HOLAP	P-BRP
Precision	**0.8889**	0.8824	0.8881
Recall	0.8955	0.8955	**0.9478**
F1 Score	0.8922	0.8889	**0.9170**
Precision	0.8089	**0.8571**	0.8000
Recall	0.9478	0.9403	**0.9851**
F2 Score	0.9163	0.9224	**0.9415**

TABLE IV
AVERAGE COMPUTATION TIME OF EVENT DETECTION ALGORITHMS OVER THREE-MINUTE TIME PERIOD

Number of PMUs		50	100	150
Computation Time (s) (partial/total)	HOLAP	61.78/68.46	181.50/189.25	336.27/344.58
	OLAP	7.53/15.01	9.58/17.33	16.99/24.79
	P-BRP	2.18/8.46	3.13/9.40	4.29/10.53

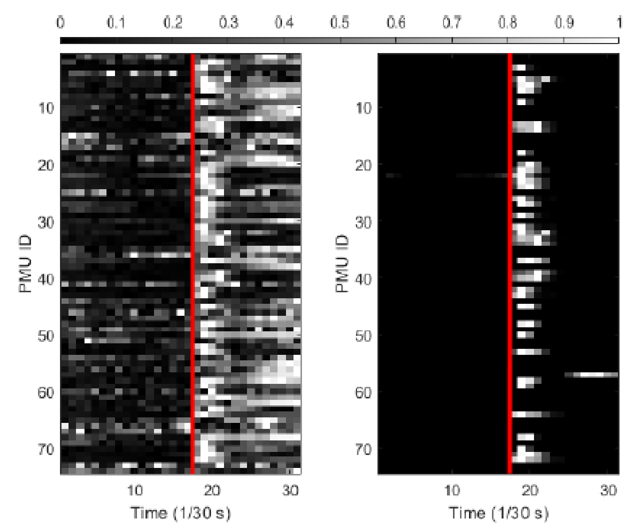


Fig. 2. The heatmap of “$X-L$” (left) and “$X-L-G$” (right) for normalized active power data (scaled from 0 to 1). The event happens approximately at the red line.

Power System Event Detection
Bidirectional Generative Adversarial Network

Y. Cheng, N. Yu, B. Foggo, and K. Yamashita, "Online Power System Event Detection via Bidirectional Generative Adversarial Networks," in *IEEE Transactions on Power Systems*, 2022.

- Motivation
  - Existing ML-based event detection algorithm requires thousands of confirmed events as training labels. Event detection accuracy drops quickly as the number of training label reduces.
- Main Idea
  - Learn two mapping functions that project PMU data samples during normal operating conditions to the noise space and then back to the data space.
  - If there is a large difference between PMU data and its reconstructed version → Likely an event
- Key Results
  - Beat state-of-the-art algorithms in accuracy and computational efficiency without event labels.

TABLE IV
ACCURACY OF DETECTION FOR VOLTAGE-RELATED EVENTS

	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
True Positive	**584**	534	561	512
False Positive	**42**	67	138	77
False Negative	**23**	73	46	95
Precision	**93.29%**	88.89%	80.26%	86.92%
Recall	**96.21%**	89.55%	92.42%	84.34%
F_1 Score	**94.73%**	89.22%	85.91%	85.61%

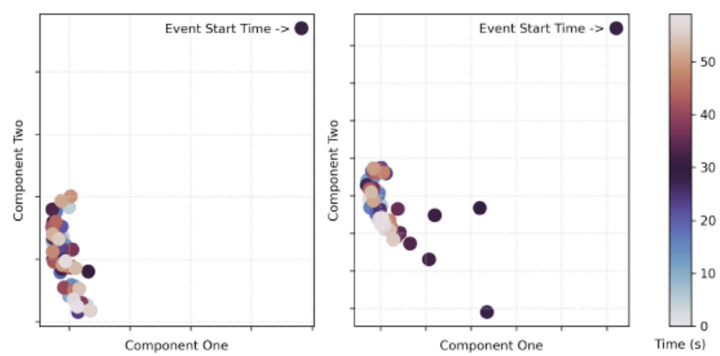
TABLE V
ACCURACY OF DETECTION FOR FREQUENCY-RELATED EVENTS

	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
True Positive	**82**	72	71	75
False Positive	**5**	56	19	46
False Negative	**0**	10	11	7
Precision	**94.25%**	53.33%	78.89%	61.98%
Recall	**100%**	88.89%	86.59%	91.46%
F_1 Score	**97.04%**	66.67%	82.56%	73.89%

TABLE VII
AVERAGE RUNTIME OF DIFFERENT ALGORITHMS FOR EVENT DETECTION

	Bi-AnoGAN	OLAP	GSP-based	AnoGAN
Voltage Events	**13.59 s**	21.75 s	7.16 s	876.58 s
Frequency Events	**13.47 s**	20.98 s	7.31 s	843.89 s

<2 milliseconds process each snapshot of incoming PMU data sample



Noise space of voltage and frequency events.

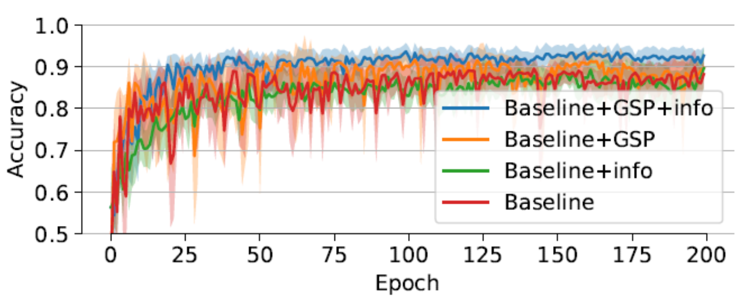
Power System Event Classification Information Loading and PMU Sorting

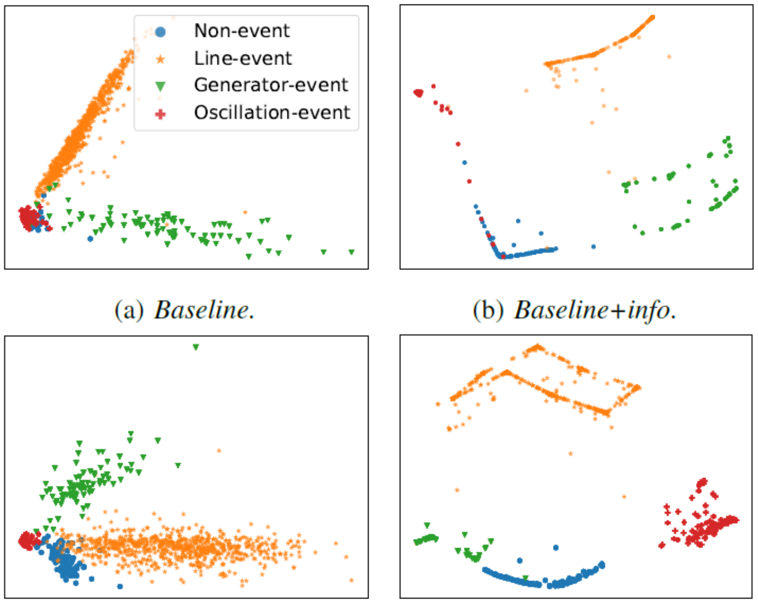
J. Shi, B. Foggo, and N. Yu, "Power System Event Identification based on Deep Neural Network with Information Loading," *IEEE Trans on Power Syst.* vol. 36, no. 6, pp. 5622-5632, Nov. 2021.

- Motivation
  - Online power system event classification is crucial to improving system reliability.
- Main Idea
  - Make parameter sharing scheme of convolutional neural network more effective with a graph signal processing-based PMU sorting algorithm.
  - Control the information compression in deep neural network with information loading.
- Key Results
  - Online event classification algorithm achieves higher F1-score with interpretable result

F1 Scores on Testing Dataset

	Non-event	*Line-event*	*Generator event*	*Oscillation event*
Baseline	0.554	0.879	0.881	0.208
Baseline+info	0.596	0.928	0.924	0.205
Baseline+GSP	0.894	0.937	0.907	0.922
Baseline+GSP+info	0.973	0.971	0.962	0.986





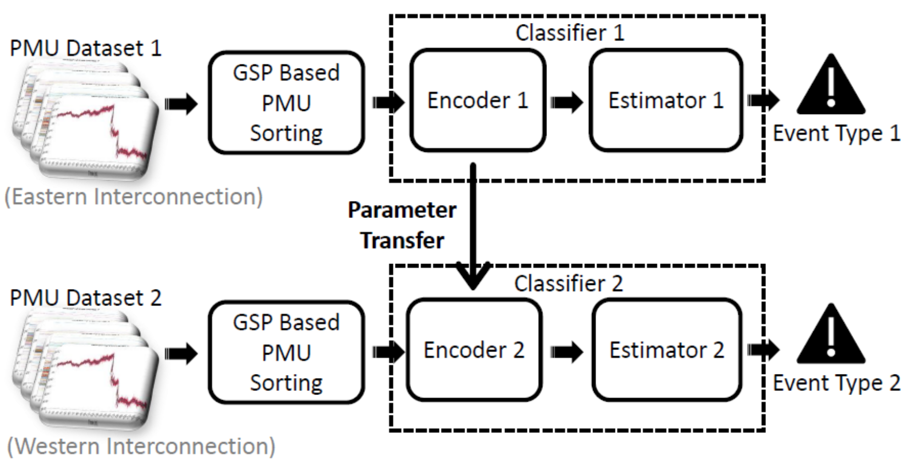
(a) *Baseline.* (b) *Baseline+info.*

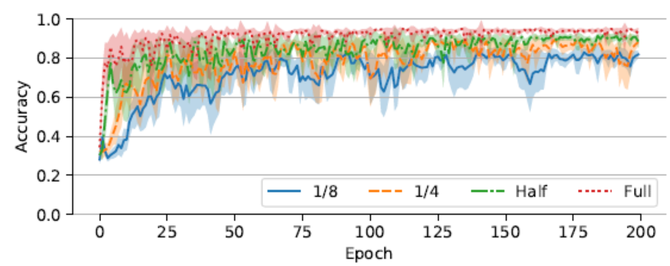
(c) *Baseline+GSP.* (d) *Baseline+GSP+info.*

Power System Event Classification Transfer Learning

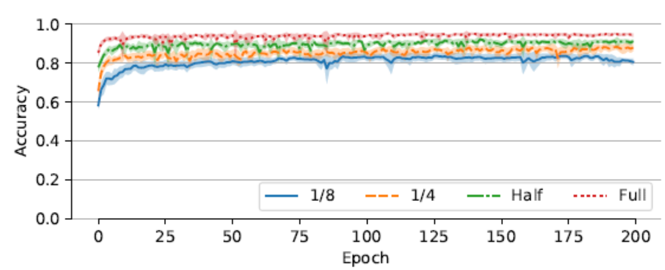
J. Shi, K. Yamashita, and N. Yu, "Power System Event Identification with Transfer Learning Using Large-scale Real-world Synchrophasor Data in the United States," in *IEEE ISGT NA*, 2022.

- Motivation
  - Deep neural network-based event classification algorithms require a large amount of power system event training labels. A single transmission grid operator has limited event labels.
- Main Idea
  - Accelerate the training of a deep learning model for a new transmission grid by exploiting the information from a previously trained model for a different transmission network.
- Key Results
  - Deep neural network trained to identify system events for the Eastern Interconnection provides useful information when building the event identification engine for Western Interconnection.





(a) Without transfer learning.



(b) With transfer learning.

Power System Dynamic Parameter Estimation

X. Kong, K. Yamashita, B. Foggo, and N. Yu, "Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations," *IEEE PES General Meeting*, pp. 1-5, 2022.

- Motivation
  - High fidelity power system dynamic models are critical to both dynamic studies and reliable operation of the power system.
- Main Idea
  - Convert the forward solvers of ordinary differential equations representing power system dynamics into physics-informed neural networks.
  - Derive the gradient of the loss function w.r.t. dynamics parameters based on the adjoint method.
  - Update the dynamic parameters with a quasi-Newton method.
- Key Results
  - Capable of accurately estimating the dynamic parameters with 3 seconds of noisy PMU data.

TABLE III
RELATIVE ESTIMATION ERROR (%) OF BASELINE AND NEURALODE-BASED METHOD

		Data Length			Initial
		1s	3s	5s	REE
Parameters	P_{m_1}	2.36 / **1.50**	7.70 / **1.49**	7.47 / **1.26**	4.21
	P_{m_2}	0.76 / **0.01**	5.05 / **0.12**	5.01 / **0.18**	5.77
	P_{m_3}	1.05 / **0.41**	6.27 / **0.32**	6.52 / **0.46**	4.76
	M_{0_1}	5.80 / **2.86**	5.52 / **2.06**	6.31 / **3.09**	4.80
	M_{0_2}	4.36 / **3.19**	5.92 / **2.18**	5.88 / **3.67**	6.10
	M_{0_3}	**4.58** / 8.37	5.50 / **4.82**	**5.49** /10.51	5.33
Average		3.15 / **2.72**	5.99 / **1.83**	6.11 / **3.20**	5.16

Note: Baseline / Physics-based Neural ODE Algorithm.

TABLE IV
RUNNING TIME (S) OF BASELINE AND NEURAL ODE-BASED METHOD.

Data Length	Running Time (second)		Learning Rate
	Baseline	Neural ODE-based	
1s	8.38	3.78	0.5 / 0.5
3s	38.55	4.82	0.05 / 0.5
5s	100.25	4.67	0.05 / 0.5

Note: Baseline / Physics-based Neural ODE Algorithm.

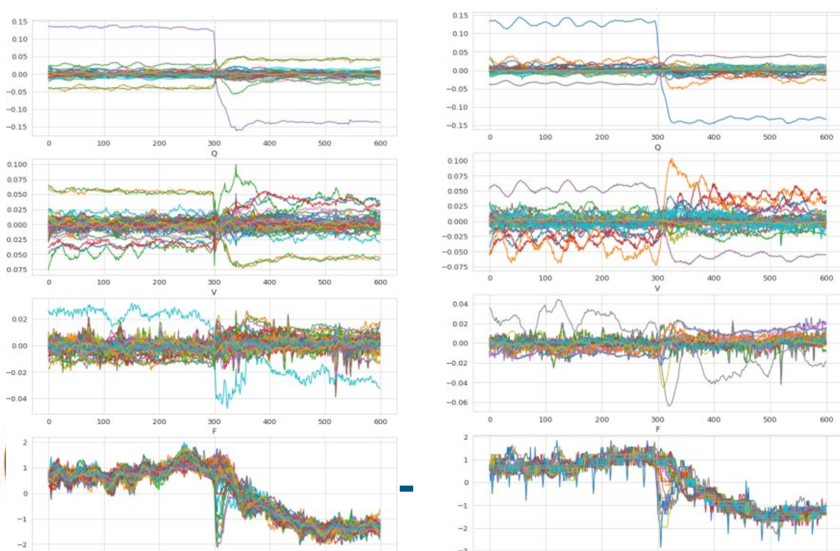
NETL NATIONAL ENERGY TECHNOLOGY LABORATORY

Synthetic Power System Event Data Creation: pmuBAGE The Benchmarking Assortment of Generated PMU Events

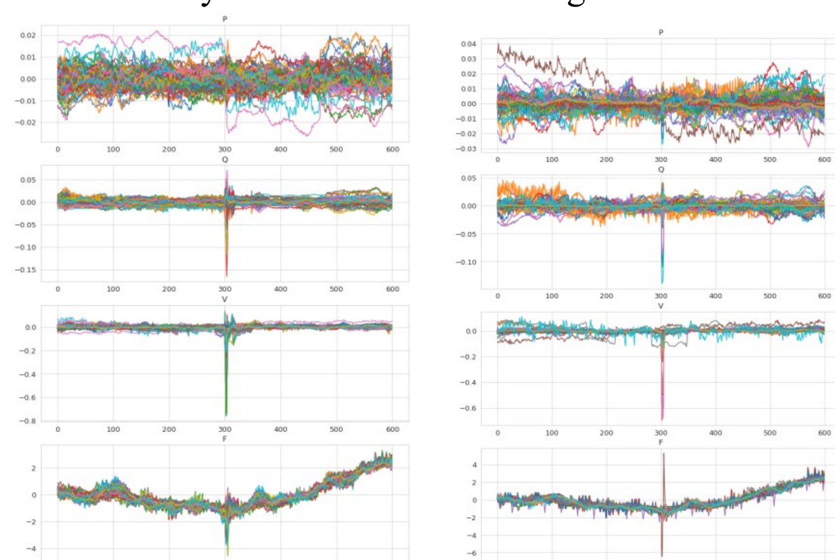
pmuBAGE: The Benchmarking Assortment of Generated PMU Events – Part I and II under review. Paper: https://arxiv.org/abs/2204.01095. Data: https://github.com/NanpengYu/pmuBAGE

- Motivation
  - The development of machine learning-based algorithm needs a reliable PMU data source.
  - Benchmarking across algorithms is hard when they are all tested on different data sets.
- Main Idea
  - Physical event signatures are PMU private and are used directly.
  - Statistical participation factors are synthesized with novel generative model.
  - Event signatures can be decomposed into two types: inter-event and intra-event.
- Key Results
  - A set of synthetic PMU event data set that maintain an unprecedented level of realism.

synthetic and actual frequency event



synthetic and actual voltage event

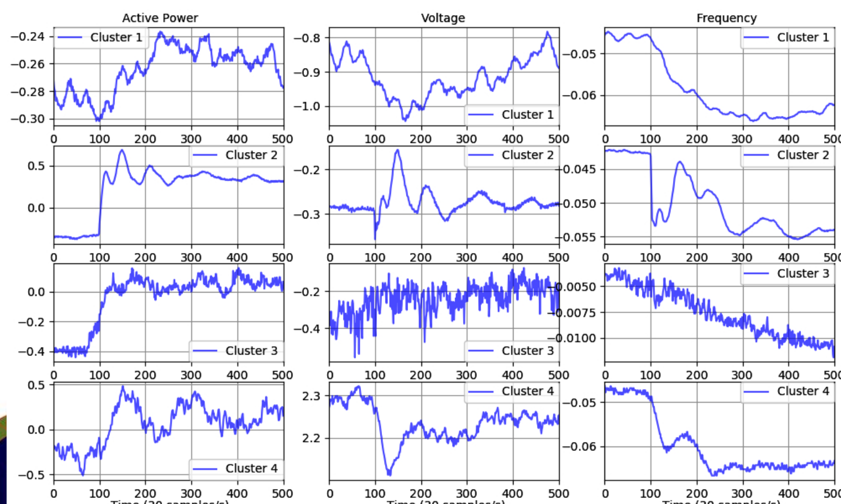


Power System Event Signature Library

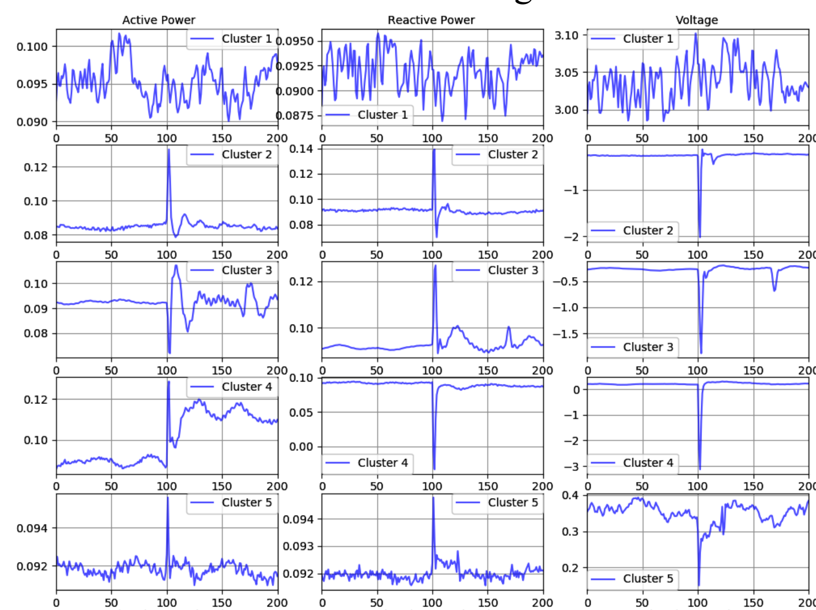
A dynamic Behavior-based Bulk Power System Event Signature Library with Empirical Clustering, under preparation

- Motivation
  - A dynamic behavior-based event signature library with PMU data could help discover new power system events or classify event into an existing category.
- Main Idea
  - The main categories and subcategories are determined empirically with expert knowledge.
  - Verified with classification model and information entropy (smaller entropy, better categorization).
- Key Results
  - A library of power system events in 3 groups and 14 sub-groups

Centroid PMU: frequency event



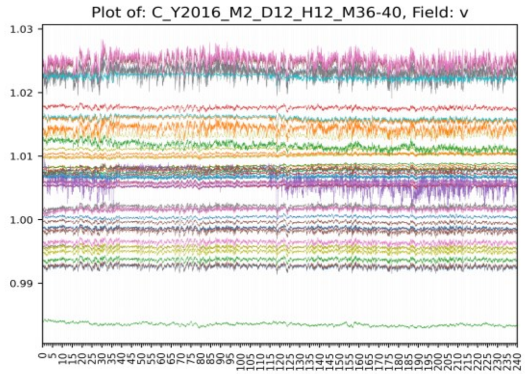
Centroid PMU: voltage event

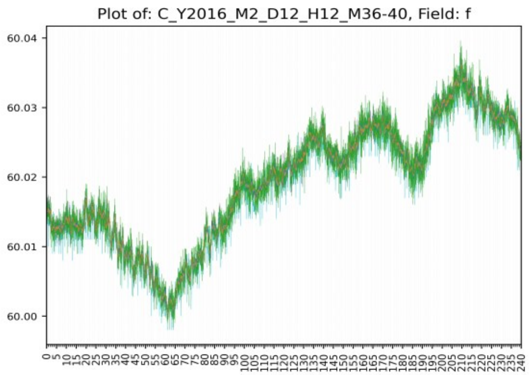


Summary of Results on Testing Dataset

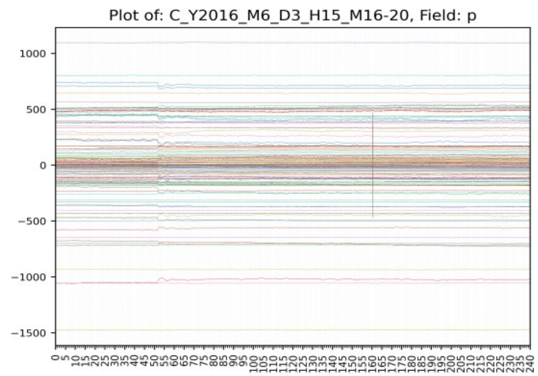
- Thousands of events are detected and classified by the proposed physics-informed machine learning algorithms in all three interconnections.

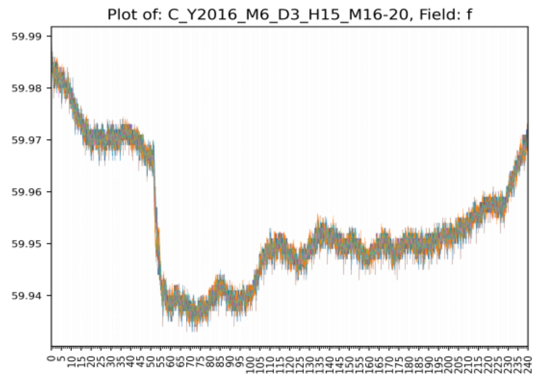
Sample oscillation event IC-C
02/12/2016 12:36-12:40 PM



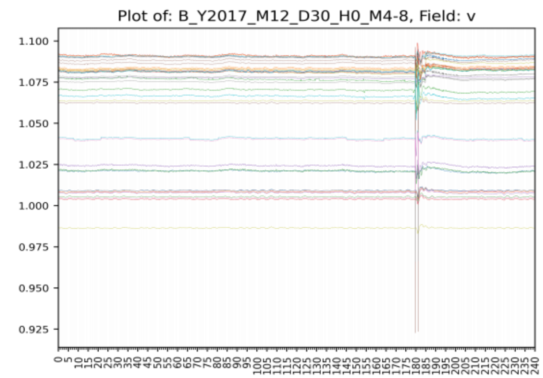


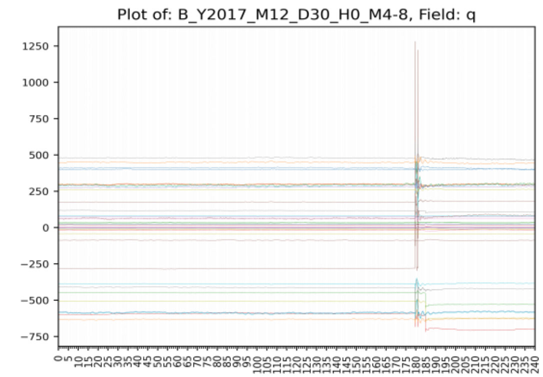
Sample frequency event IC-C
06/03/2016 03:16-03-20 PM





Sample voltage event IC-B
12/30/2017 00:04-00-08 AM



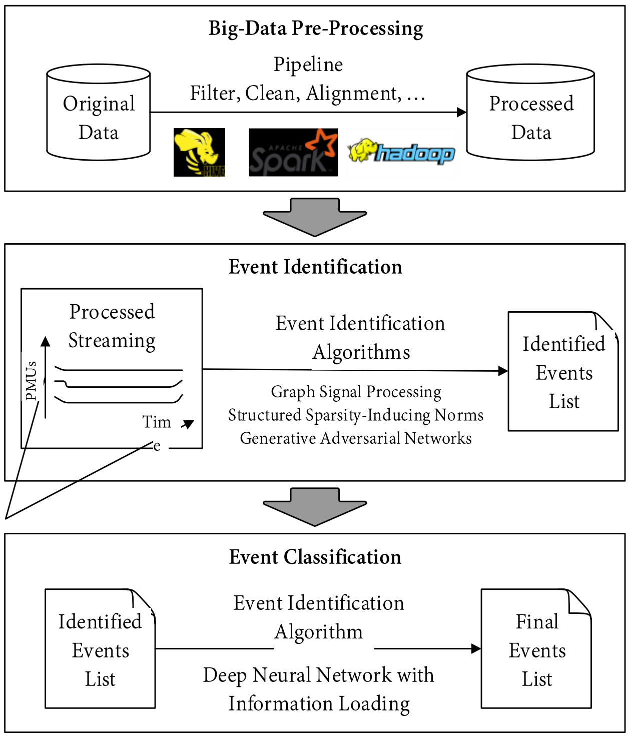


Value of Work

- Benefits for the Electric Power Utilities and Transmission System Operators
  - Improvement in data quality: missing value replacement and dynamic parameter estimation technologies
  - Boost in situational awareness: power system event detection algorithms and event signature library
  - Enhancement in reliability: power system event classification tools
- Benefits for Researchers and Developers
  - Access to a large-scale synthetic PMU event dataset for algorithm development and performance benchmarking
- Benefits for the Broader Scientific Community
  - Information theoretic machine learning: performance bounds for ML algorithms
  - An enlightening glimpse of how to embed power system dynamics and/or domain knowledge into machine learning algorithms

Readiness for Commercialization

- Modular software components of PMU data analytics and a complete data pipeline for machine learning.
- The technology readiness level (TRL) of all algorithms (missing value replacement, power system event detection, power system event classification, and dynamic parameter estimation) reached TRL 6.
- A prototype system has been built using a commercial big data platform with Hadoop and Spark.
- Commercialization Path
  - Partnering with vendors (e.g., EPG, EPRI)
    - Integration of software modules in commercial products
  - Pilot demonstration with electric utilities
    - Demonstration of entire software platform or module(s)



Overview of prototype system

Readiness for ML & BD Analytics

- Off-the-shelf machine learning models could not achieve reasonable performance for PMU data analytics.
- Challenges for AI/ML in the Context of Power System Data
  - Bad data quality (consecutive missing data, inaccurate event timestamps, outliers, inaccurate dynamic parameters)
  - Insufficient data (event labels, cause of event, network topology, dynamic system model, unbalanced dataset)
  - Interpretability of machine models for PMU data analytics
  - Trade-off between speed and accuracy
  - Safety of ML in critical infrastructure system (bulk power system)
  - Investment in physical hardware and human capital
  - How to turn data analytics into actionable intelligence?
  - Collaboration between operators and artificial intelligence

Lessons Learned and Next Steps

- Lessons Learned
  - Physics-based machine learning is the key to developing breakthrough technology in power system data analytics.
  - The availability of real-world (synthetic) power system data is crucial to the accelerated development and benchmarking of data-driven algorithms.
  - Crucial to have a collaborative team with both deep power system domain knowledge and theoretical understanding of advanced machine learning.
- Next Steps
  - Deeper integration of physical power system model with machine learning algorithms
  - Interpretable machine learning models for PMU data analytics
  - Making artificial intelligence algorithms actionable in bulk power system (safe and sample efficient deep reinforcement Learning and imitation learning)
  - A closer collaboration between artificial and operator intelligence

Publications

Accepted and Published

1. J. Shi, B. Foggo, X. Kong, Y. Cheng, N. Yu, and K. Yamashita "Online Event Detection in Synchrophasor Data with Graph Signal Processing," *IEEE SmartGridComm*, pp. 1-7, 2020.
2. J. Shi, B. Foggo, and N. Yu, "Power System Event Identification based on Deep Neural Network with Information Loading," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5622-5632, Nov. 2021.
3. B. Foggo and N. Yu, "Online PMU Missing Value Replacement via Event-Participation Decomposition," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 488-496, Jan. 2022.
4. X. Kong, B, Foggo, and N. Yu, "Power System Event Detection Using Optimization with Structured Sparsity-Inducing Norms," to appear in *IEEE Transactions on Power Systems*, 2022. DOI: 10.1109/TPWRS.2021.3134945.
5. J. Shi, K. Yamashita, and N. Yu, "Power System Event Identification with Transfer Learning Using Large-scale Real-world Synchrophasor Data in the United States," to appear in *IEEE ISGT North America*, pp. 1-5, 2022.
6. Y. Cheng, N. Yu, B. Foggo, and K. Yamashita, "Online Power System Event Detection via Bidirectional Generative Adversarial Networks," to appear in *IEEE Transactions on Power Systems*, 2022.
7. X. Kong, K. Yamashita, B. Foggo, and N. Yu, "Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations," to appear in *IEEE Power and Energy Society General Meeting*, pp. 1-5, 2022.

Under Review

1. pmuBAGE: The Benchmarking Assortment of Generate PMU Events – Part I and II
2. A dynamic Behavior-based Bulk Power System Event Signature Library with Empirical Clustering

Under Preparation

1. Short-term Forecasting of PMU Data by Attentional Seq2Seq LSTM with Prior Knowledge Matrix and Magnitude Direction Coupling

Panel Sessions and Presentations

Organized and Delivered

1. Presentation: 2021 NASPI Spring Work Group Meeting, “Online Power System Event Detection and Identification with PMU Data”, April 13, 2021.
2. Presentation: The 2nd IEEE International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA),“Physics-Informed Machine Learning for Power System Event Detection and identification with Synchrophasor Data”, May 24, 2021.
3. Presentation: 2021 IEEE PES GM, “Power System Event Identification based on Deep Neural Network with Information Loading”, July 25, 2021.
4. Presentation: Dominion Energy, “Phasor Measurement Unit Data Analytics”, June 23, 2020.
5. Presentation: Electric Power Group, “Machine Learning and Big Data Analytics Technologies for PMU Data”, April 21, 2021.
6. Presentation: LLNL and ORNL, “Extracting Useful Information from Terabytes of PMU Data with Machine Learning and Data Mining Techniques”, September 8, 2021.
7. Panel Session: 2021 IEEE PES GM, “Big Data Analytics of Synchrophasor Data – Experience from the U.S. (Academic Track)”, July 25, 2021.
8. Panel Session: 2021 IEEE PES GM, “Big Data Analytics of Synchrophasor Data- Experience from the U.S. (Industry Track)”, July 25, 2021.

Upcoming

1. Presentation: 2022 IEEE PES GM, “Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations”, July 2022.
2. Presentation: 2022 IEEE PES GM, “Development of Deep Learning and Data Mining Techniques Using Terabytes of PMU Data from the U.S.”, July 2022.
3. Presentation: 2022 IEEE ISGT North America, “Power System Event Identification with Transfer Learning Using Large-scale Real-world Synchrophasor Data in the United States”, April 2022.
4. Panel Session: 2022 IEEE PES GM, “Synchrophasor Data Analytics for Power System Monitoring, Operation and Planning”, July 2022.

Contract Information:
Nanpeng Yu, nyu@ece.ucr.edu

Thank You

DOE: Sandra Jenkins, Brian Mollohan, Carol Painter

PNNL: James Follum, Jeffery Banning

UCR: Brandon Foggo, Koji Yamashita, Jie Shi, Yuanbin Cheng, Xianghao Kong, Eamonn Keogh

EPG: Neeraj Nayak, Song Xue, Vikram Budhraja

MTU: Chee-Wooi Ten