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
  o UCR, EPG, Michigan Tech

• Project Objectives
  o Develop physics-informed machine learning algorithms using real-world PMU data to enhance power system reliability.
  o Accelerate future research and development of data-driven algorithms by creating a synthetic PMU dataset and an event signature library of bulk power systems.
  o Create prototype systems that implement the proposed machine learning algorithms.

• Significance and Impacts
  o Developed a suite of physics-informed machine learning algorithms for PMU data analytics
  o Enhanced the reliability of bulk power systems with data-driven algorithms and prototype systems
  o Advanced scientific knowledge of physics-based machine learning
Technical Accomplishments

• PMU Data Quality Improvement
  o Online PMU Missing Value Replacement via Event-Participation Decomposition

• Power System Event Detection
  o Graph Signal Processing-based Event Detection
  o Voltage Event Detection Using Optimization with Structured Sparsity-Inducing Norms
  o Power System Event Detection with Bidirectional Generative Adversarial Network

• Power System Event Classification
  o Deep Neural Network-based Power System Event Classification
  o Classify Power System Event with a Small Number of Training Labels with Transfer Learning

• Power System Dynamic Parameter Estimation
  o Dynamic Parameter Estimation with Physics-based Neural Ordinary Differential Equations

• Synthetic Power System Event Data Creation
  o 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


- **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.

### Table II: Average MAPEs over Bus Event data.

<table>
<thead>
<tr>
<th>BUS</th>
<th>Base</th>
<th>EnCorr</th>
<th>OLAP</th>
<th>SPIKE-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (s)</td>
<td>0.005</td>
<td>0.303</td>
<td>0.067</td>
<td>0.038</td>
</tr>
<tr>
<td>VM (%)</td>
<td>2.449</td>
<td>2.449</td>
<td>1.309</td>
<td>1.082</td>
</tr>
<tr>
<td>IM (%)</td>
<td>68.792</td>
<td>56.531</td>
<td>21.393</td>
<td>18.995</td>
</tr>
<tr>
<td>P (%)</td>
<td>27.413</td>
<td>22.692</td>
<td>7.888</td>
<td>7.225</td>
</tr>
<tr>
<td>Q (%)</td>
<td>28.594</td>
<td>24.442</td>
<td>17.580</td>
<td>10.937</td>
</tr>
</tbody>
</table>

Fig. 1: Bus event sample decomposition

Fig. 4: SPIKE-P replacement on a Generator Event Sample.
Power System Event Detection
Graph Signal Processing


- **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

<table>
<thead>
<tr>
<th>Method</th>
<th>GSP</th>
<th>OLAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td>0.7692</td>
<td>0.9</td>
</tr>
<tr>
<td>Category 2</td>
<td>1</td>
<td>0.8889</td>
</tr>
<tr>
<td>Category 3</td>
<td>0.8889</td>
<td>0.75</td>
</tr>
<tr>
<td>All Events</td>
<td>0.8750</td>
<td>0.8519</td>
</tr>
</tbody>
</table>

### Scalability Test

<table>
<thead>
<tr>
<th>Number of PMUs</th>
<th>30</th>
<th>60</th>
<th>90</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runtime</td>
<td>3.01 s</td>
<td>4.45 s</td>
<td>5.80 s</td>
<td>6.95 s</td>
</tr>
</tbody>
</table>

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


- **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

<table>
<thead>
<tr>
<th>Statistics</th>
<th>OLAP</th>
<th>HOLAP</th>
<th>P-BRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>0.8889</td>
<td>0.8824</td>
<td>0.8881</td>
</tr>
<tr>
<td>Recall</td>
<td>0.8955</td>
<td>0.8955</td>
<td>0.9478</td>
</tr>
<tr>
<td>F1 Score</td>
<td>0.8922</td>
<td>0.8889</td>
<td>0.9170</td>
</tr>
<tr>
<td>Precision</td>
<td>0.8089</td>
<td>0.8571</td>
<td>0.8000</td>
</tr>
<tr>
<td>Recall</td>
<td>0.9478</td>
<td>0.9403</td>
<td>0.9851</td>
</tr>
<tr>
<td>F2 Score</td>
<td>0.9163</td>
<td>0.9224</td>
<td>0.9415</td>
</tr>
</tbody>
</table>

### TABLE IV
Average computation time of event detection algorithms over three-minute time period

<table>
<thead>
<tr>
<th>Number of PMUs</th>
<th>50</th>
<th>100</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation Time (s)</td>
<td>HOLAP</td>
<td>61.78/68.46</td>
<td>181.50/189.25</td>
</tr>
<tr>
<td></td>
<td>OLAP</td>
<td>7.53/15.01</td>
<td>9.58/17.33</td>
</tr>
<tr>
<td></td>
<td>P-BRP</td>
<td>2.18/8.46</td>
<td>3.13/9.40</td>
</tr>
</tbody>
</table>

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


• Motivation
  o 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
  o Learn two mapping functions that project PMU data samples during normal operating conditions to the noise space and then back to the data space.
  o If there is a large difference between PMU data and its reconstructed version → Likely an event

• Key Results
  o Beat state-of-the-art algorithms in accuracy and computational efficiency without event labels.

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>ACCURACY OF DETECTION FOR VOLTAGE-RELATED EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bi-AnoGAN</td>
</tr>
<tr>
<td>True Positive</td>
<td>584</td>
</tr>
<tr>
<td>False Positive</td>
<td>42</td>
</tr>
<tr>
<td>False Negative</td>
<td>23</td>
</tr>
<tr>
<td>Precision</td>
<td>93.29%</td>
</tr>
<tr>
<td>Recall</td>
<td>96.21%</td>
</tr>
<tr>
<td>$F_1$ Score</td>
<td>94.73%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE V</th>
<th>ACCURACY OF DETECTION FOR FREQUENCY-RELATED EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bi-AnoGAN</td>
</tr>
<tr>
<td>True Positive</td>
<td>82</td>
</tr>
<tr>
<td>False Positive</td>
<td>5</td>
</tr>
<tr>
<td>False Negative</td>
<td>0</td>
</tr>
<tr>
<td>Precision</td>
<td>94.25%</td>
</tr>
<tr>
<td>Recall</td>
<td>100%</td>
</tr>
<tr>
<td>$F_1$ Score</td>
<td>97.04%</td>
</tr>
</tbody>
</table>

<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


**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

<table>
<thead>
<tr>
<th></th>
<th>Non-event</th>
<th>Line-event</th>
<th>Generator event</th>
<th>Oscillation event</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>0.554</td>
<td>0.879</td>
<td>0.881</td>
<td>0.208</td>
</tr>
<tr>
<td><strong>Baseline+info</strong></td>
<td>0.596</td>
<td>0.928</td>
<td>0.924</td>
<td>0.205</td>
</tr>
<tr>
<td><strong>Baseline+GSP</strong></td>
<td>0.894</td>
<td>0.937</td>
<td>0.907</td>
<td>0.922</td>
</tr>
<tr>
<td><strong>Baseline+GSP+info</strong></td>
<td>0.973</td>
<td>0.971</td>
<td>0.962</td>
<td>0.986</td>
</tr>
</tbody>
</table>

---

F1 Scores on Testing Dataset

---

![Accuracy vs Epoch](image1)

![Non-event Line-event Generator-event Oscillation-event](image2)

(a) **Baseline**

(b) **Baseline+info**

(c) **Baseline+GSP**

(d) **Baseline+GSP+info**
**Power System Event Classification**

**Transfer Learning**


- **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.

![Diagram](image)
Power System Dynamic Parameter Estimation


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

• Motivation
  o The development of machine learning-based algorithm needs a reliable PMU data source.
  o Benchmarking across algorithms is hard when they are all tested on different data sets.

• Main Idea
  o Physical event signatures are PMU private and are used directly.
  o Statistical participation factors are synthesized with novel generative model.
  o Event signatures can be decomposed into two types: inter-event and intra-event.

• Key Results
  o 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
  o 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
  o The main categories and subcategories are determined empirically with expert knowledge.
  o Verified with classification model and information entropy (smaller entropy, better categorization).

• Key Results
  o 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
  o Improvement in data quality: missing value replacement and dynamic parameter estimation technologies
  o Boost in situational awareness: power system event detection algorithms and event signature library
  o Enhancement in reliability: power system event classification tools

• Benefits for Researchers and Developers
  o Access to a large-scale synthetic PMU event dataset for algorithm development and performance benchmarking

• Benefits for the Broader Scientific Community
  o Information theoretic machine learning: performance bounds for ML algorithms
  o 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
  o Physics-based machine learning is the key to developing breakthrough technology in power system data analytics.
  o The availability of real-world (synthetic) power system data is crucial to the accelerated development and benchmarking of data-driven algorithms.
  o Crucial to have a collaborative team with both deep power system domain knowledge and theoretical understanding of advanced machine learning.

• Next Steps
  o Deeper integration of physical power system model with machine learning algorithms
  o Interpretable machine learning models for PMU data analytics
  o Making artificial intelligence algorithms actionable in bulk power system (safe and sample efficient deep reinforcement Learning and imitation learning)
  o A closer collaboration between artificial and operator intelligence
Publications

Accepted and Published


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

Upcoming
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