Event Detection and Classification in Temporal Networks from Scarce and Imprecise Labels

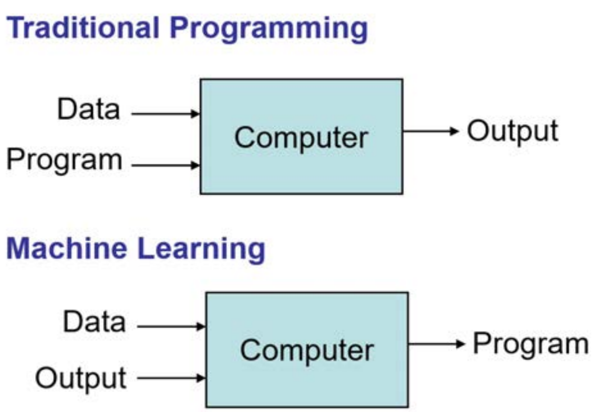
Zoran Obradovic

zoran.obradovic@temple.edu
https://dabi.temple.edu/zoran-obradovic/

Data Analytics Center
Department of Computer and Information Sciences
Statistics Department
Temple University, USA

Machine Learning in Temporal Networks

ML Objective: Learn a function f that maps the given input to the output



Some of common **ML assumptions:**

- Training and test instances are drawn independently from the same unknown distribution
- Function mapping explanatory variables (inputs/features**)** to response variables (outputs) is smooth almost everywhere
- Signal to noise ratio is large
- Ignored information has negligible effects
- Labels are precise and sufficiently large

- Some of **ML assumptions are often violated in temporal networks**

  Complex relationships ⇨ various modeling challenges arise

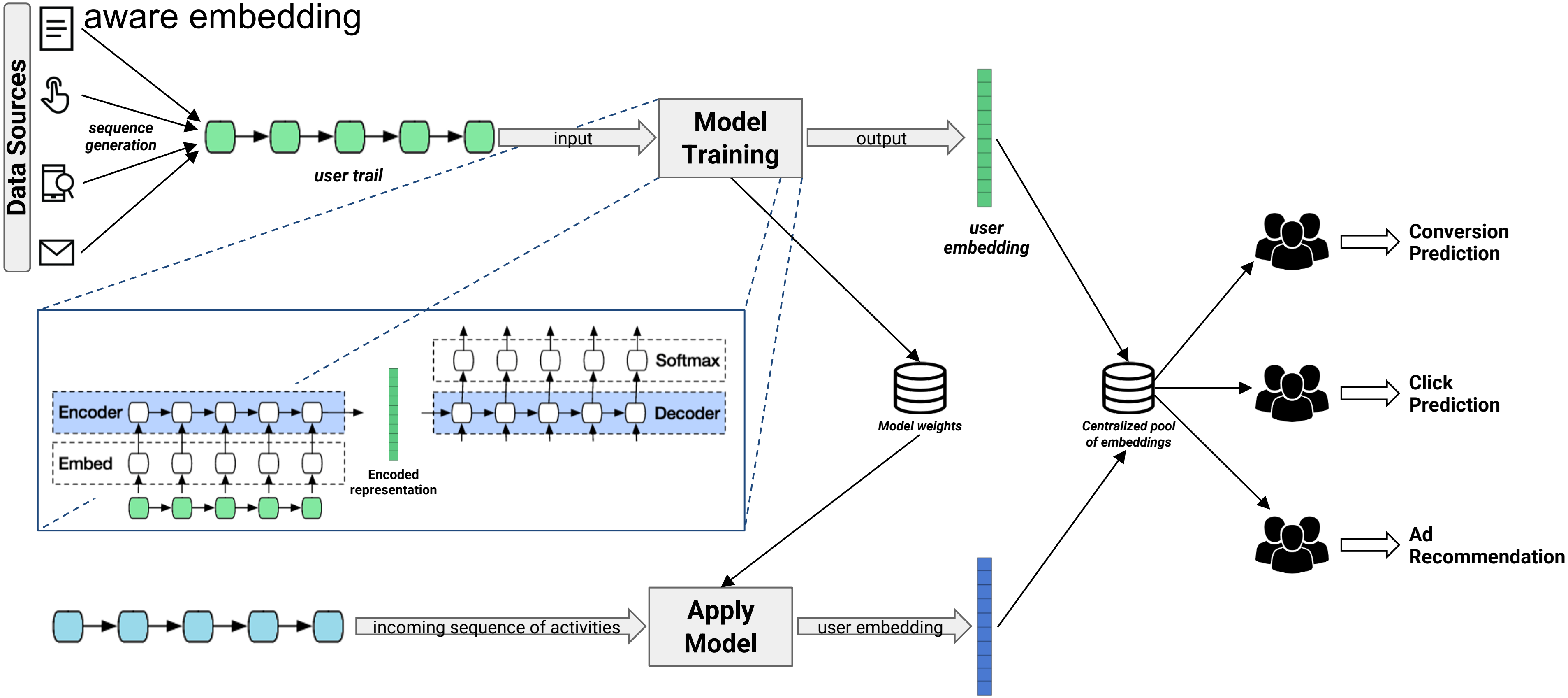
- <u>Challenges discussed today</u>:

  The ability of ML techniques to **detect, classify and anticipate events** when

1. **Data is anonymized**
2. **Observations are scarce**
3. **Labels are imprecise**

Example 1: Learning from Sequentially Structured Data

- **Given:** A sequence of user activities, along with their corresponding time stamps
.
- **Objective:** Learn a function that projects a user to a task-independent, compact, temporally aware embedding



Our Solution: Time-aware sequential autoencoder – achieved 25% lift in conversion prediction (advertising)

Pavlovski, M., Gligorijevic, J., Stojkovic, I., Komirishetty, S., Gligorijevic, Dj., Bhamidipati, N., Obradovic, Z. "Time-Aware User Embeddings as a Service," *ACM SIGKDD Conf. on Knowledge Discovery and Data Mining,* 2020

Example 2: Structured Learning in Temporal Networks: Predictive Electrical Outage Management

Outage Cost: $150B annually in US

Challenge: Heterogeneous spatio-temporal data (outage records, weather measurements and forecast)

Solution: Collaborative Logistic Ensemble Classifier

+ utilized distance correlation to **balance underfit/overfit** [Pavlovski et al, *IJCAI 2018*]

+ learned from **spatial substructures**

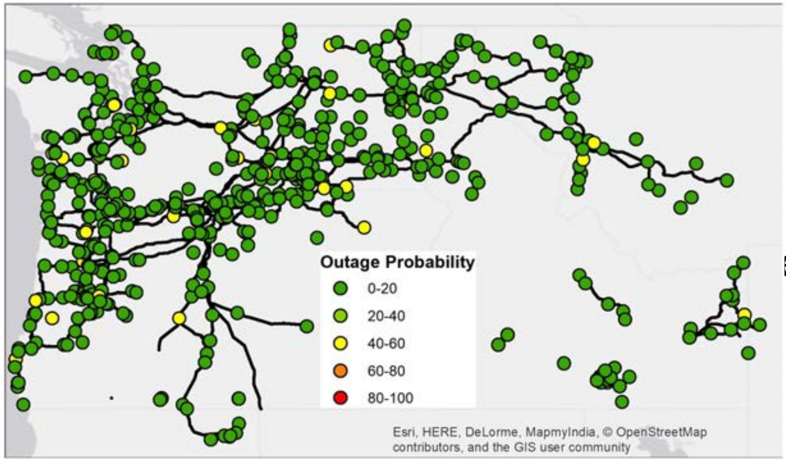
Train a **subbagging ensemble** of M components:

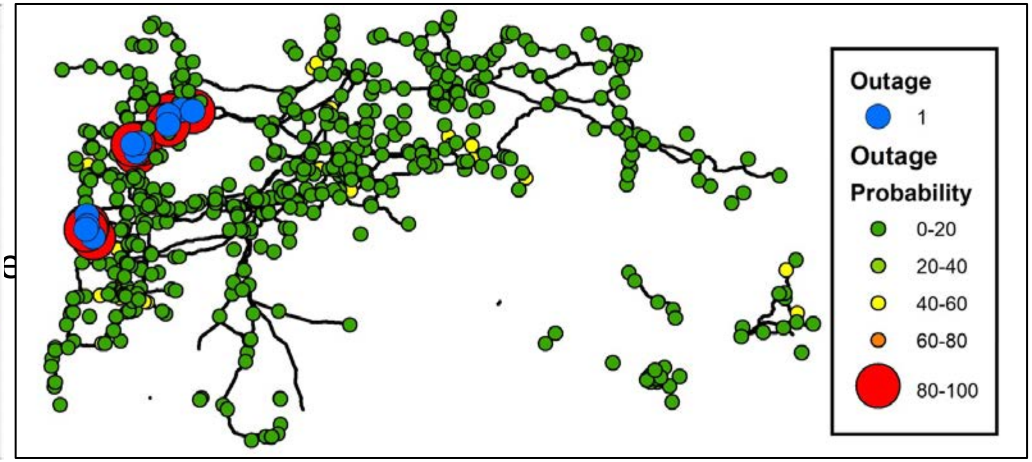
$$F_1, F_2, \dots, F_M$$

Iteratively **exchange examples to minimize:**

$$\sqrt{R_{emp}(h, \mathbf{X}_s, \mathbf{y})^2 + dCorr\big(\ell(\cdot, h), (\mathbf{x}_{s_{trn}}, y_{s_{trn}})\big)^2}$$

Prediction: $\Phi(\mathbf{x}_s) = \operatorname{sign}\left(\sum_{m=1}^{M} F_m(\mathbf{x}_s)\right)$





Probabilities of outages when: no outages occurred (left), and outages were caused by lightning (right).

- **No outages occurred** ⇨ outage probabilities are **smaller than 60%** for all substations
- **Outages occurred** ⇨ the area around the outages has points with probability **over 80%**

Dokic T, Pavlovski M, Gligorijevic Dj, Kezunovic M, Obradovic Z., "Spatially Aware Ensemble-Based Learning to Predict Weather-Related Outages in Transmission," *IEEE HICSS* 2019

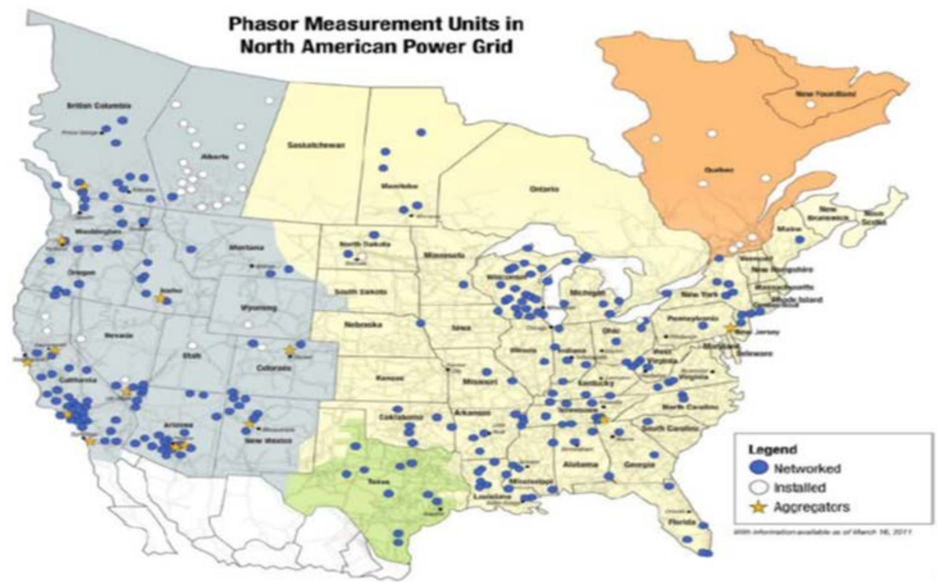
TODAY: Power System Events Detection and Classification from PMUs Across the US

PMUs (*Phasor measurement units*) - <u>hundreds of terabytes data</u>

- simultaneously register and records multiple variables in an electric grid
- record magnitude/phase angle of a phasor quantity (voltage, current), frequency, ROCOF
- use GPS for synchronization
- <u>sparsely located</u> (approximately 2,000 deployed, covering < 5% of the electrical buses) across *Eastern* Interconnection, *Western* Interconnection and *Texas* (ERCOT) Interconnection

Objective:

- **<u>Pro-active</u>** approach to improving the *reliability* and *situational awareness* of power systems based on **detection** and further **classification of *<u>local</u>* and *<u>global</u>* events from scarce PMUs based on non-representative and imprecise labels**

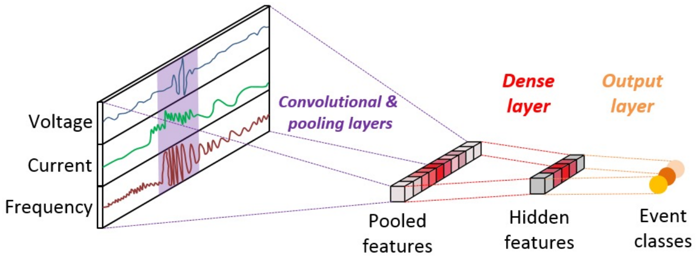


<u>Source</u>: Black, Clifton "Synchrophasors: Improving Reliability & Situational Awareness", *Research & Technology Management,* October 2011

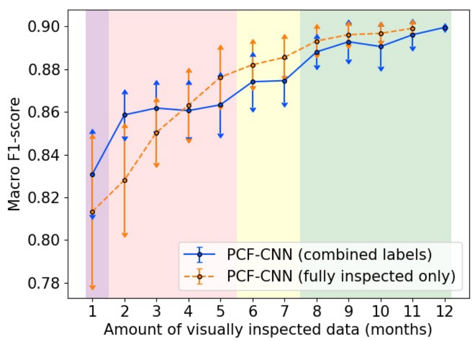
TODAY: Events Detection and Classification from PMU Data

Problem formulation: Given a **signal segment** $\mathbf{s}(t-\Delta, t+\Delta) = [\mathbf{s}^{(1)}(t-\Delta, t+\Delta), \ldots, \mathbf{s}^{(M)}(t-\Delta, t+\Delta)]$, from *multiple anonymized PMUs* (removed grid topology) **predict event type** $y \in \{0, \ldots, C\}$ that occurred at $[t-\Delta, t+\Delta]$ by learning from *scarce observations* and *low precision labels*.

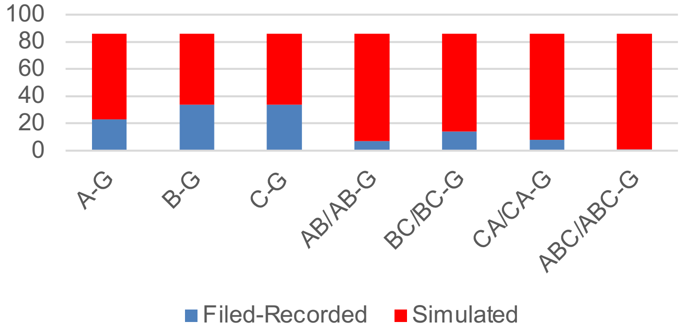
Q1: Can we automate feature learning?
Yes – Multi-channel filtering by CNN



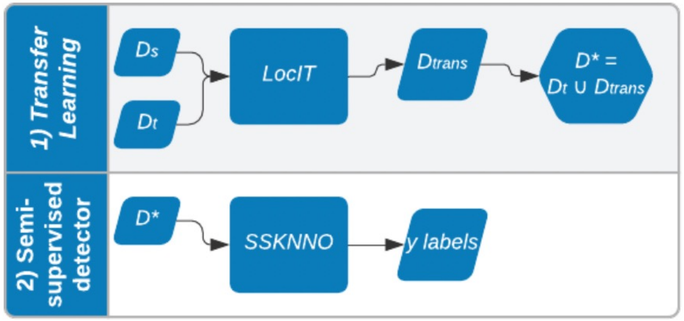
Q2: Should we learn from more data or from better data?
Use both if data is small



Q3: Can we enhance PMU data through simulations?
Yes, but need 3 phase data



Q4: Can we use relevant labeled PMU data from a related task? Yes - transfer learning



Details: Described in Our 2021 Publications

M. Pavlovski, M. Alqudah, T. Dokic, A. Abdel Hai, M. Kezunovic, Z. Obradovic, "Use of Hierarchical Convolutional Neural Networks for Event Classification on PMU Data," *IEEE Trans. on Instrumentation and Measurement*, In Press.

A. Abdel Hai, T. Dokic, M. Pavlovski, M. Alqudah, M. Kezunovic, Z. Obradovic, "Transfer Learning for Event Detection from PMU Measurements with Scarce Labels," *IEEE Access*, vol. 9, 1274420-127432, 10 Sept. 2021 Preprint.

H. Otudi, T. Dokic, T. Mohamed,Y. Hi, M. Kezunovic, Z. Obradovic, "Line Faults Classification Using Machine Learning On Three phases Voltages Extracted from Large Dataset of PMU Measurements," *IEEE HICSS-55*, January 2022, In Press.

R. Baembitov, T. Dokic, M. Kezunovic, Z. Obradovic, "Fast Extraction and Characterization of Fundamental Frequency Events from a Large PMU Dataset Using Big Data Analytics," *IEEE HICSS-54*, January 2021.

M. Alqudah, M. Pavlovski, T. Dokic, M. Kezunovic, Y. Hu, Z. Obradovic, "Convolution-based Event Detection Utilizing Timeseries Data Streams from Phasor Measurement Units Sparsely Located Across Electric Power Systems," *In Review*.

A. Abdel Hai, M. Pavlovski, T. Dokic, T. Mohammed, S. Saranovic, M. Kezunovic, Z. Obradovic, "Transfer Learning for Detection of Events using Phasor Measurement Data from Another Grid," *In Review.*

Problem Formulation

Problem formulation: Given a **signal segment**

$$\mathbf{s}(t-\Delta, t+\Delta) = [\mathbf{s}^{(1)}(t-\Delta, t+\Delta), \ldots, \mathbf{s}^{(M)}(t-\Delta, t+\Delta)] \,,$$

from *multiple anonymized PMUs*, **predict** $y \in \{0, \ldots, C\}$ which indicates the **type of event** that occurred during $[t-\Delta, t+\Delta]$ by learning from *scarce observations* and *low precision labels*.

Existing traditional ML methods: aim to increase the value of collected PMU data for improved situational awareness and predictive decision-making

(1) Face many challenges such as *high dimensionality*, *autocorrelation*, *adaptation, evaluation*

(2) Rely on *missing, unreliable*, or *imprecise* labels

(3) Do not focus on characterizing *local* and *global* events

(4) Typically utilize only a single PMU variable

Our objectives: Consider more advanced methods (capable of modeling temporal network data)

- Employ automated feature learning
- Utilize all available channels, as well as each channel separately
- Analyze the effect of different mechanisms for improving labels on *local/global* event characterization

This work is a part of **"Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART)",** an ongoing project funded by the ***US Department of Energy (DOE)***.

- Pavlovski, M., Alqudah, M., Dokic, T., Abdel Hai, A., Kezunovic, M., Obradovic, Z. "Hierarchical Convolutional Neural Networks for Event Classification on PMU Measurements," *IEEE Transactions on Instrumentation & Measurement*. In press.

Automated Feature Learning for Events Predictions in PMUs

Model Variants

Traditional

- Decision Tree **(DT)**
- Logistic Regression **(LR)**
- Multilayer Perceptron **(MLP)**
- Support vector machine **(SVM)**

Single-channel

- Single-channel Convolutional Neural Network **(SC-CNN)**

Multi-channel

- Parallel Channel Filtering CNN **(PCF-CNN)**
- Simultaneous Channel Filtering CNN **(SCF-CNN)**

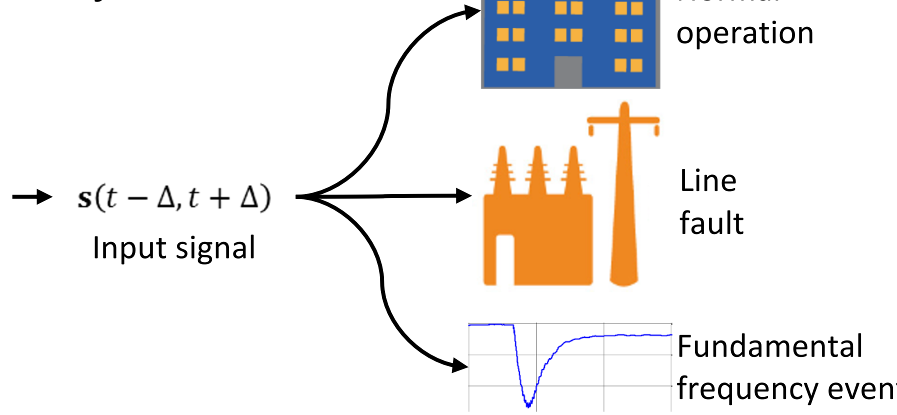
Classification modes

o *Standard* (multi-class)

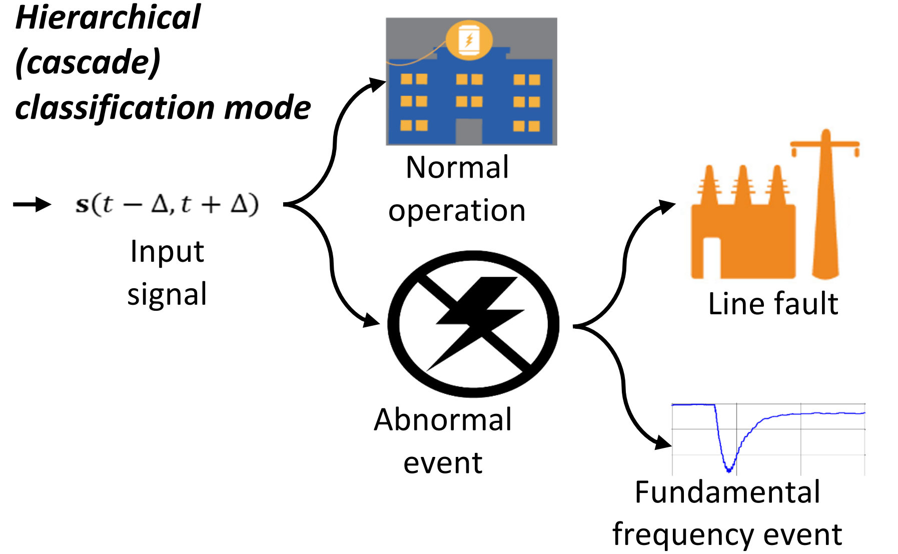
o *Hierarchical* (cascade)

- Detected events are classified into line or freq. events

Standard (multi-class) classification mode

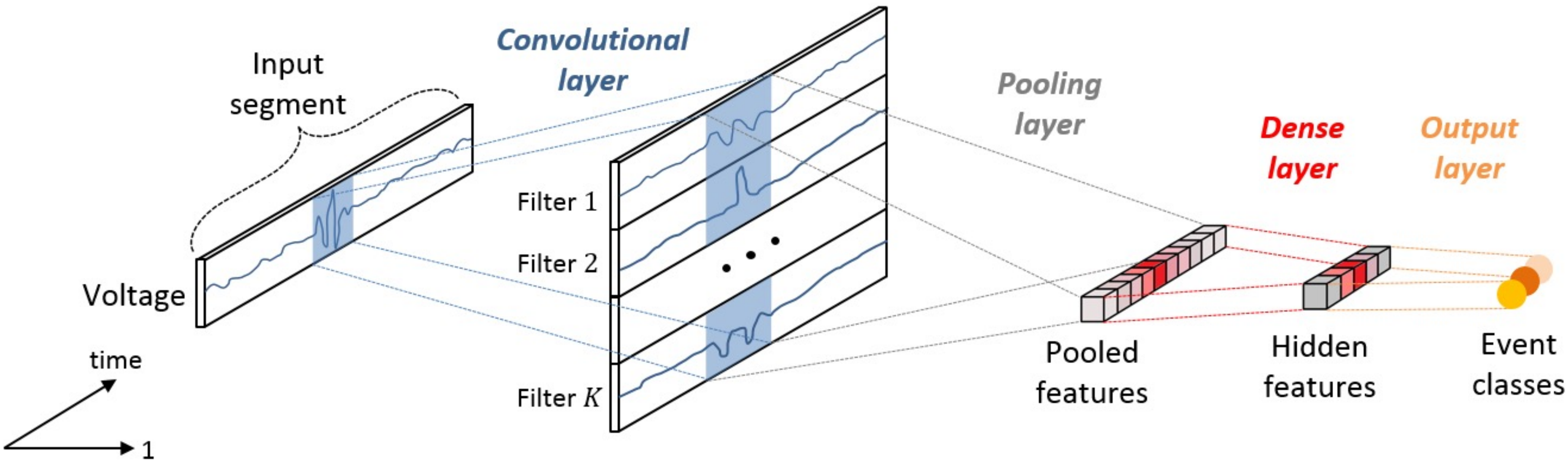


Hierarchical (cascade) classification mode



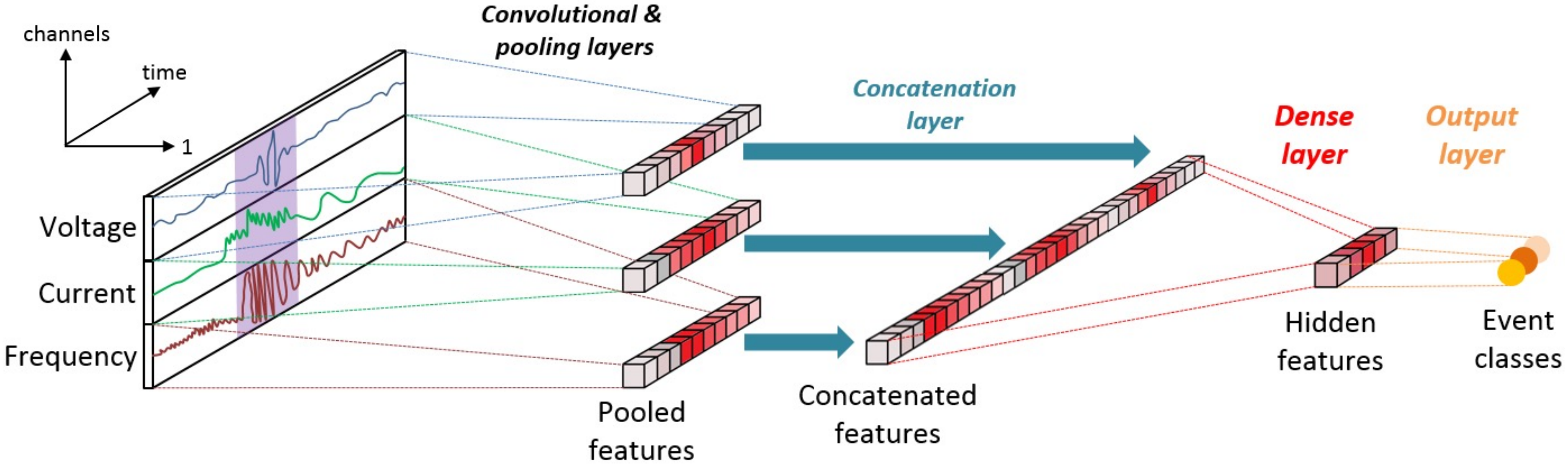
Single-Channel CNN Architecture

Single-Channel (SC) CNN **utilizing *voltage* signal segments**

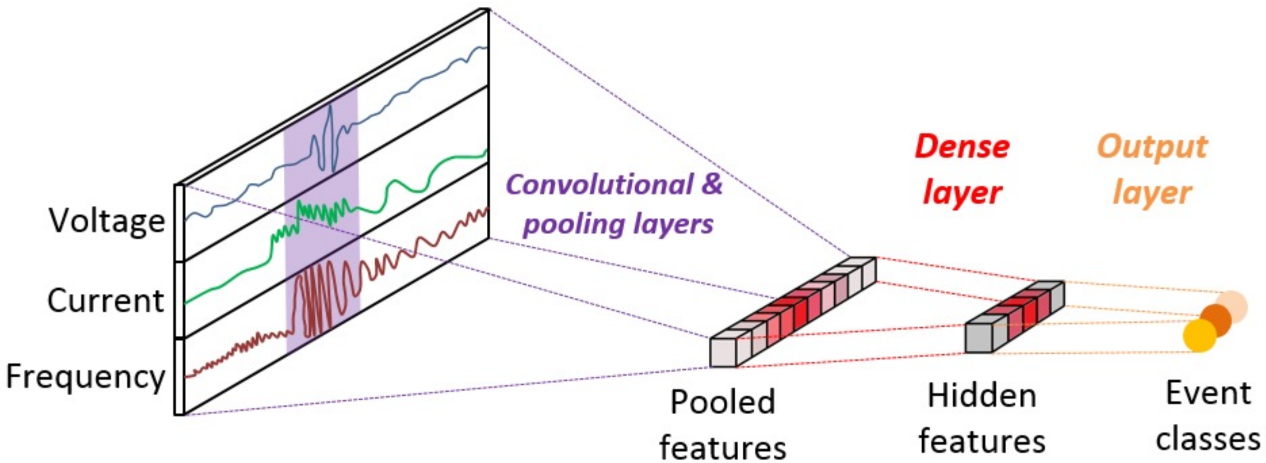


Multi-Channel CNN Variants

Parallel Channel Filtering (PCF) CNN



Simultaneous Channel Filtering (SCF) CNN

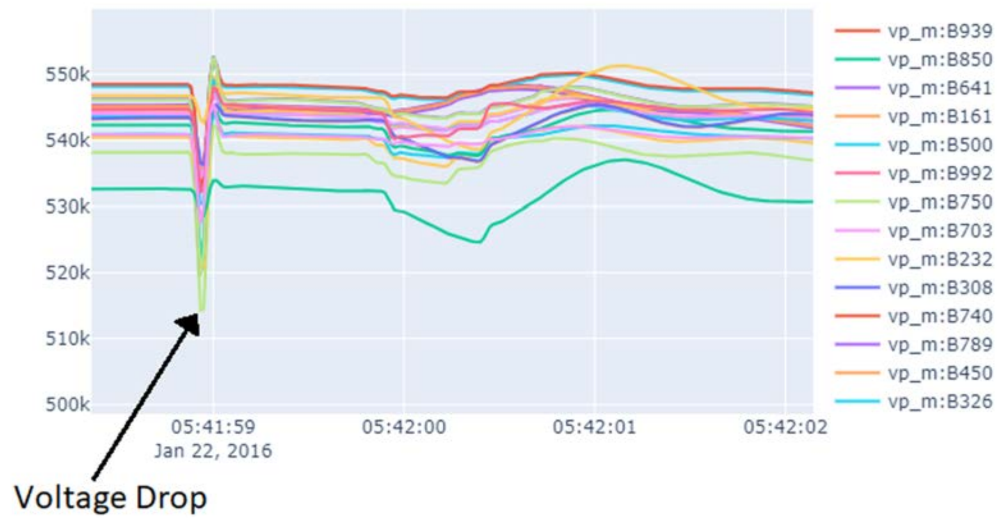


Data

- **38 PMU devices** from *unknown locations* at Western U.S. Interconnection
- **Signal types:** *vp_m, ip_m, f*
- **2 years** of data (2016, 2017)

2016-01-22 05:41:45 -- 05:42:15

PMUs' voltage signals (an event detected by majority of the PMUs).



Preprocessing

- **Downsampling:** 30 samples/seconds
- Resulting in 180 values in ***1 minute*** for each PMU
- ***Aggregation*** of all PMUs' sub-signals
  - **Soft-DTW** (Dynamic Time Warping) [*]

Event Logs

** The domain expert was working ~3 hours/day.*
+ Additional (secondary) visual inspection

Event Log / Advantages	*Rapidly refined*	*Partially inspected*	*Fully inspected*
Handpicked normal operation segments	✓	✗	✓
Narrower time intervals	✓	✓	✓
Single event per interval	✗	✓	✓
Precise intervals (centered events)	✗	✓	✓
Visually & inspected events	✗	✓	✓+
Labeling time *	38 hrs.	~2 months (120 hrs.)	2.5 months (150 hrs.)

Descriptive summary of train/test datasets.

Event Log	Number of labeled segments		
	Total	Grouped by type	
Rapidly refined ('16)	1170	Normal	467
		Line	454
		Frequency	249
Partially inspected ('16)	1748	Normal	1311
		Line	227
		Frequency	210
Fully inspected ('16)	921	Normal	481
		Line	229
		Frequency	211
Holdout ('17)	879	Normal	426
		Line	273
		Frequency	180

[*] Cuturi, M., et al. "Soft-DTW: a differentiable loss function for time-series," *ICML, 2017.*

Overall Effect of Event Labeling on Event Classification

Model variant \ Metrics		Macro AUPRC	Macro Precision	Macro Recall	Macro F1-score
Rapidly refined					
Traditional	DT	0.622	0.725	0.740	0.726
	MLR	0.802	0.783	0.758	0.754
	FFNN	0.845	0.778	0.769	0.758
	MCSVM	0.816	0.770	0.764	0.759
Single channel	SC-CNN (*V*)	0.803	0.803	0.797	0.800
	SC-CNN (*I*)	0.670	0.618	0.668	0.610
	SC-CNN (*f*)	0.650	0.598	0.647	0.592
Multi channel	PCF-CNN	0.804	0.833	0.831	0.818
	SCF-CNN	**0.854**	**0.841**	**0.845**	**0.836**
	HPCF-CNN	0.874	0.824	0.840	0.827
	HSCF-CNN	**0.897**	**0.856**	**0.861**	**0.857**
Partially inspected					
Traditional	DT	0.624	0.744	0.713	0.721
	MLR	0.848	0.820	0.765	0.773
	FFNN	0.824	0.833	0.749	0.769
	MCSVM	0.859	0.798	0.772	0.779
Single channel	SC-CNN (*V*)	0.867	0.863	0.803	0.824
	SC-CNN (*I*)	0.726	0.783	0.731	0.715
	SC-CNN (*f*)	0.700	0.769	0.687	0.695
Multi channel	PCF-CNN	**0.909**	**0.884**	**0.839**	**0.855**
	SCF-CNN	0.876	0.863	0.793	0.816
	HPCF-CNN	**0.915**	**0.898**	**0.854**	**0.871**
	HSCF-CNN	0.912	0.865	0.812	0.827
Fully inspected					
Traditional	DT	0.690	0.785	0.793	0.788
	MLR	0.856	0.821	0.807	0.808
	FFNN	0.838	0.836	0.830	0.832
	MCSVM	0.909	0.837	0.834	0.828
Single channel	SC-CNN (*V*)	0.906	0.871	0.865	0.861
	SC-CNN (*I*)	0.682	0.795	0.766	0.753
	SC-CNN (*f*)	0.696	0.774	0.714	0.731
Multi channel	PCF-CNN	**0.929**	**0.901**	**0.879**	**0.888**
	SCF-CNN	0.922	0.885	0.843	0.858
	HPCF-CNN	**0.940**	**0.911**	**0.891**	**0.900**
	HSCF-CNN	0.938	0.894	0.878	0.885

Observations

- **CNNs outperform traditional** models in most cases
- **Voltage** seems more **relevant** than current/ frequency
- **Hierarchical** models consistently **outperform** the **standard** multiclass variants
- In general, the **multi-channel CNNs** are outperforming the other model alternatives
  - HSCF-CNN on the refined event log
  - HPCF-CNN in the other cases
- **Increase** in classification **performance** as more **curated event logs** are used

Gradual Effect of Event Labeling on Event Classification

Four distinctive performance regions:

(1) **Sudden improvement** obtained by **fully inspecting 2 months** of rapidly refined segments

(2) Performance stabilizes up until **5 months** of fully inspected data are used

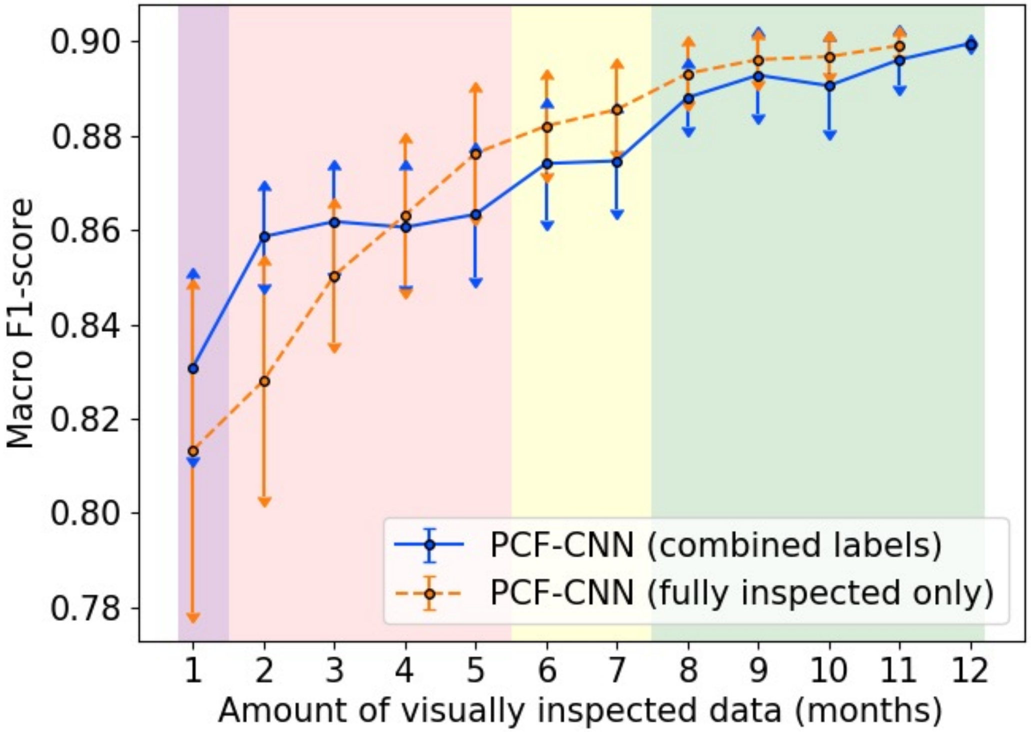
- o requires **16-28 days of inspection time**

(2) **Lift** in performance once **6 months** of segments are fully inspected

- o translates into **12% more inspection time**
- o no significant change for 7th month

(3) Inspecting ≥ 8 months: performance is similar to the one of the best-performing HPCF-CNN

- o requires the expert to devote **2-2.5 months**
- o ~20%-60% more inspection time



<u>Observations:</u>

- Domain expert's time is **extremely limited** ⇨ labeling **at least 2 months** of data is suggested
- **Otherwise** ⇨ inspection of ≥ **8 months** is needed to achieve satisfactory performance

* Using smaller fractions of expert-inspected labels alone yields greater performance than using them in addition to labels that were not fully inspected by a domain expert

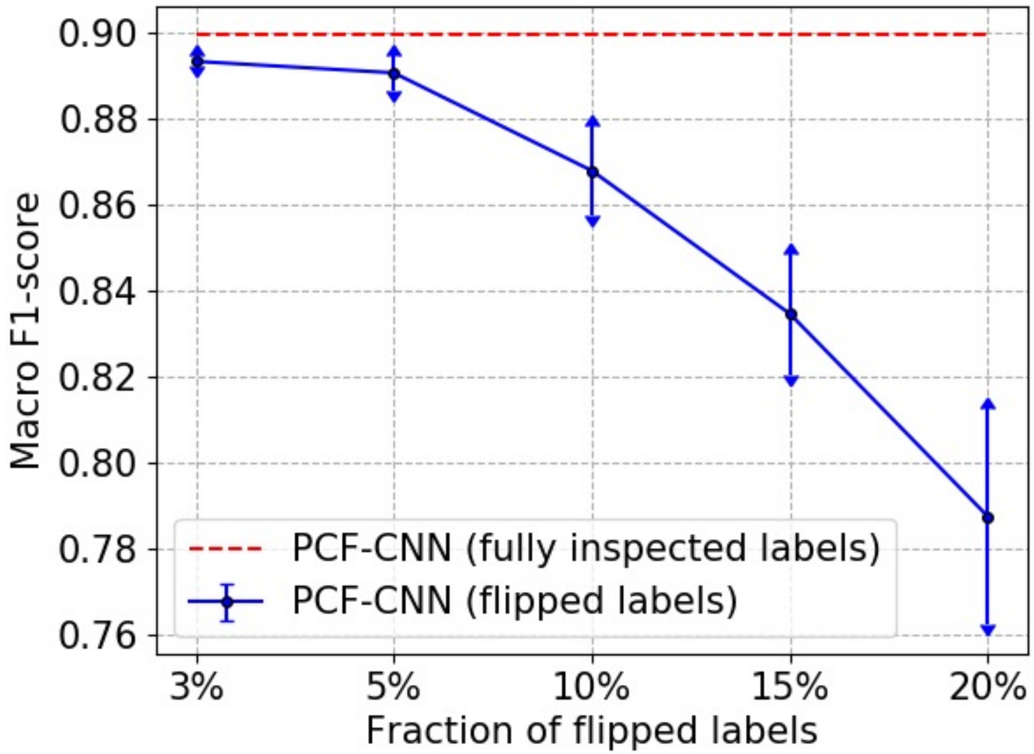
Sensitivity Analysis

- Aims to quantitatively inspect the ‘expertise’ of the domain expert

> Can we achieve event classification performance similar to the best-performing HPCF-CNN in case a less-experienced domain expert labeled the data?

- *Label noise* is injected by ***flipping labels*** to simulate a ***less-experienced domain expert***

- For each fraction of flipped labels, the experiment was repeated 10 times



<u>Observations:</u>

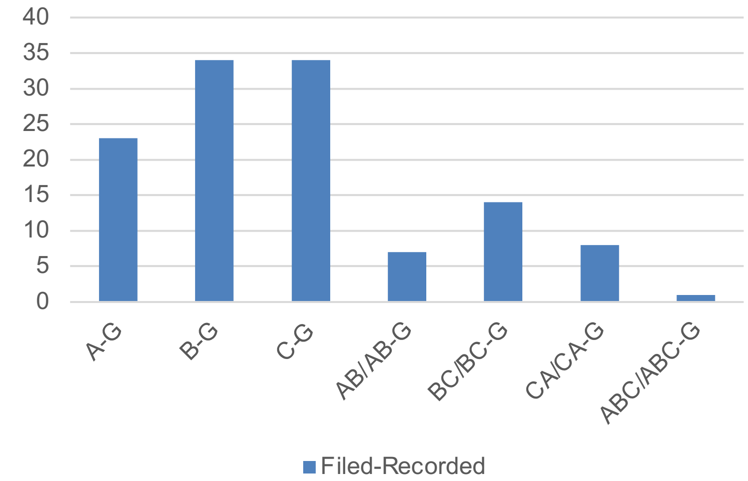
- As expected, the **performance drops** with the increase of the flipping fraction
- Larger fractions enforce more ‘randomness’ in selecting which labels will be flipped
    - o leads to **larger fluctuations**
- Significant drops in performance after flipping more than 5% of the labels

> Similar performance may be achieved with a less experienced curator as long as **< 5%** of the event annotations are mislabeled

TOPIC 2: Line Faults Classification Using Machine Learning on 3-Phase Voltages Extracted from Large PMU Measurements

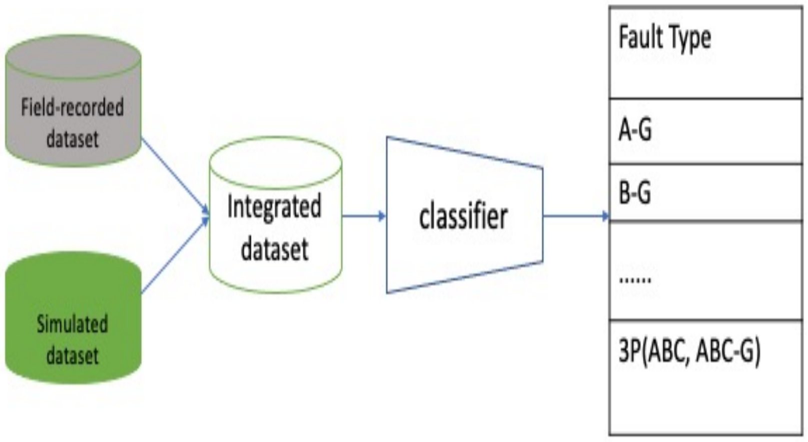
Objective:
Transmission *line faults classification* based on supervised learning

Data:
2-year field-recorded synchronized measurements of 3-phase voltages recorded by 38 phasor measurement units (PMUs) sparsely located in the US Western Grid interconnection.



Challenge 1:
It is difficult to separate PP (or 3P) from PP-G (or 3P-G) faults using field-recorded data.

Challenge 2:
There are fewer examples of PP, PPG,3P, and 3P-G faults compared to P-G faults in the field-recorded dataset.

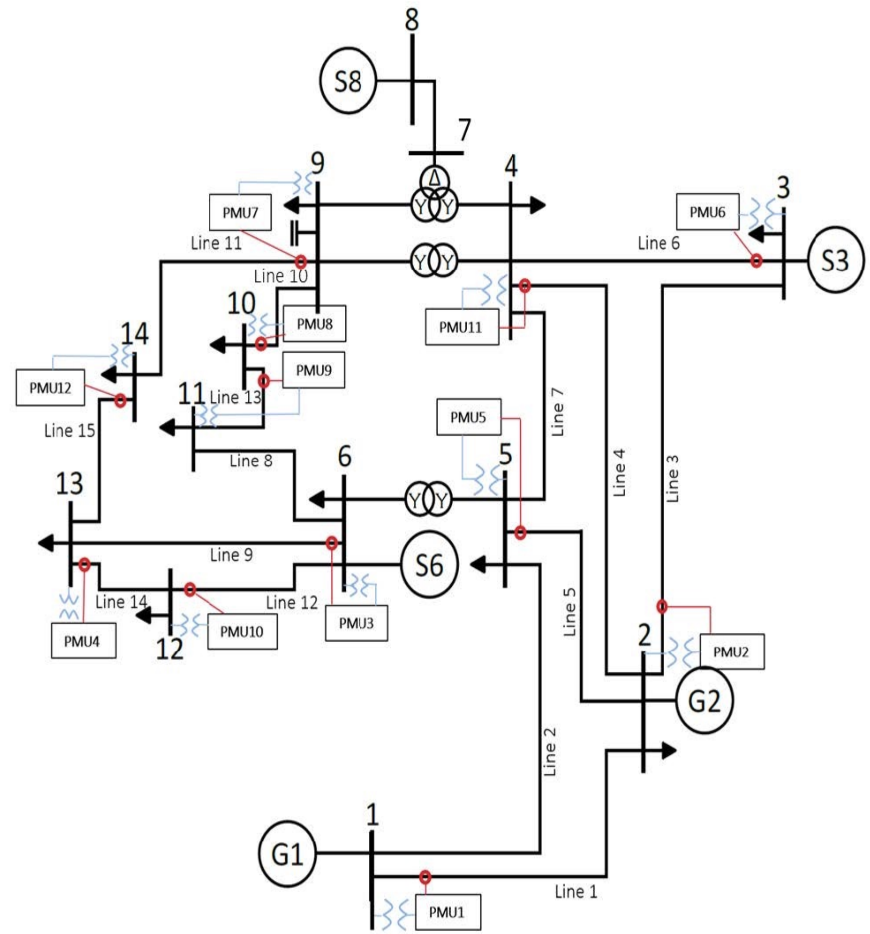


This work is a part of **"Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART)",** an ongoing project funded by the ***US Department of Energy (DOE)***.

• Otudi, H., Dokic, D., Mohamed, T., Kezunovic, M., Hu, Y., Obradovic, Z., *IEEE HICSS,* Jan. 2022.

PMU Measurement Data

- **Field-recorded data:** collected from 38 PMUs over 2 years at the US West Interconnection
- **Simulations:** based on 14-bus power system that includes 12 PMUs
- **Extracted from both sources:** three-phase measurements of the voltage magnitude
- **Features:** computed based on a statistical analysis of 2-second data windows (this window provides high accuracy of line fault-type classification)

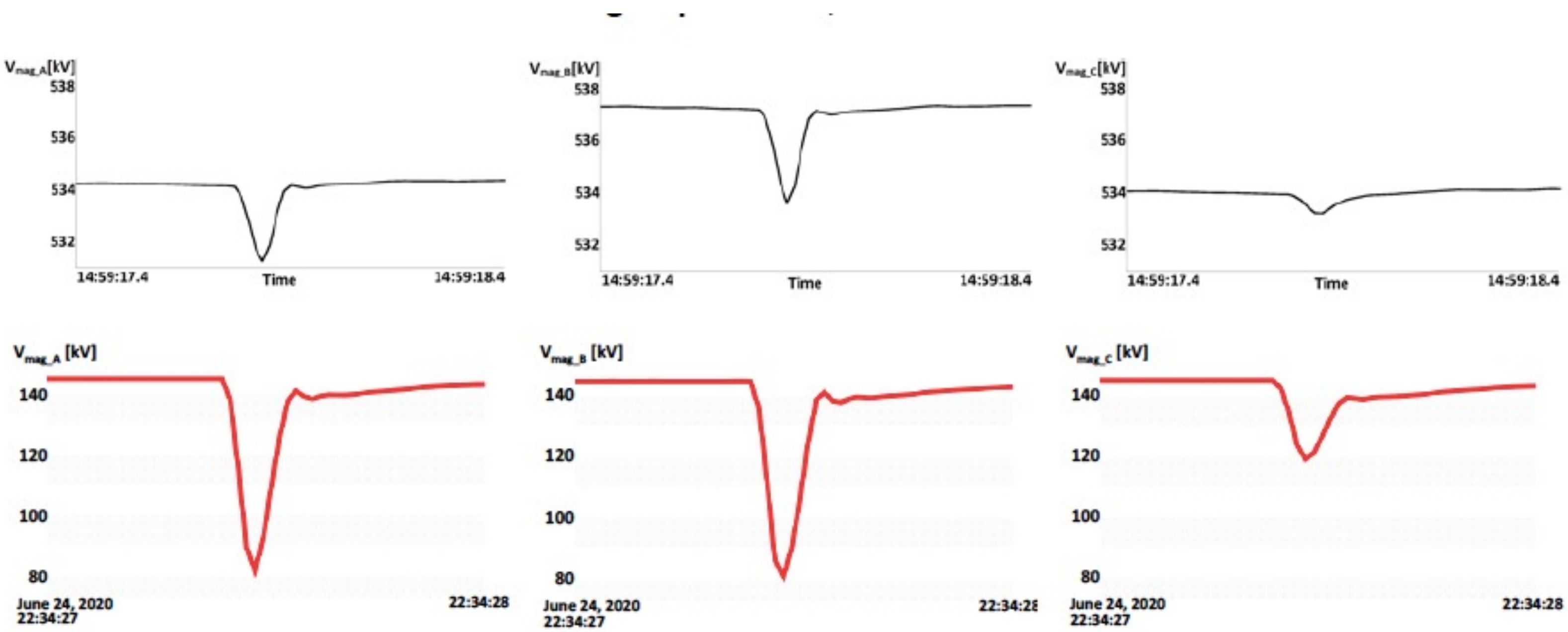


PMU placement in the synthetic IEEE 14-bus power system

Example:

Field-Recorded vs Simulated AB-G Fault

Top: Field-recorded data
Bottom: Simulated PMU data
(left to right: phases A, B and C)



Feature Extraction

- The range of voltage for each PMU is determined as

$$V_{range}(\phi_i) = \frac{\max\left(V_{mag}(\phi_i)\right) - \min\left(V_{mag}(\phi_i)\right)}{Avg\left(V_{mag}(\phi_i)\right)} \quad (1)$$

- The range of voltage is aggregated for all PMUs as

$$SUM\ (V_\phi) = \sum_{i=1}^{\text{number of PMUs}} \frac{V_{range}(\phi_i)}{number\ of\ PMUs} \quad (2)$$

- Then, difference between each two phases is determined as

$$A_{to}B = SUM\ (V_A) - SUM\ (V_B) \quad (3)$$

$$B_{to}C = SUM\ (V_B) - SUM\ (V_C) \quad (4)$$

$$C_{to}A = SUM\ (V_C) - SUM\ (V_A) \quad (5)$$

- Finally, the ratio between the differences in the voltage range is determined such that the larger value is always divided by the smaller value

Fault Type Labels

Automatic labeling was performed for multiclass line faults, as shown in Table 1.

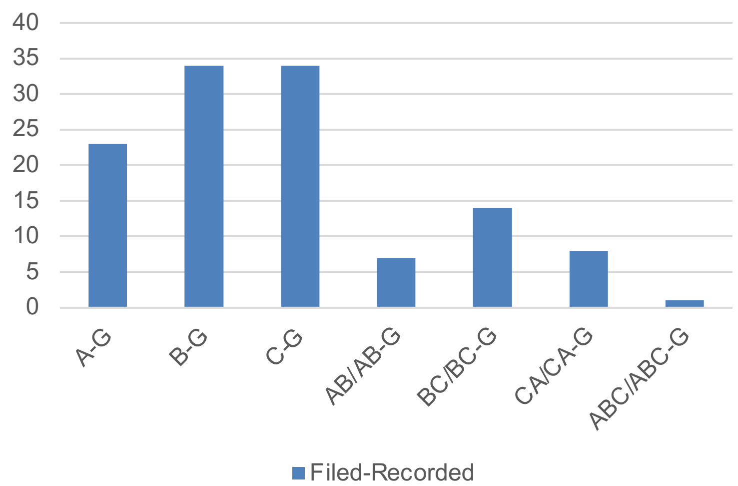
- Each label represents the occurrence of line fault type. E.g., a phase-to to ground fault (i.e., A-G, B-G, or C-G) is a combination of four labels.

- The labels are assigned as follows: for a phase-to-ground fault, if ABdiff = 1 and CAdiff = -1 and YZdiff = 1 and ZXdiff = -1, then the label will be "A-G"

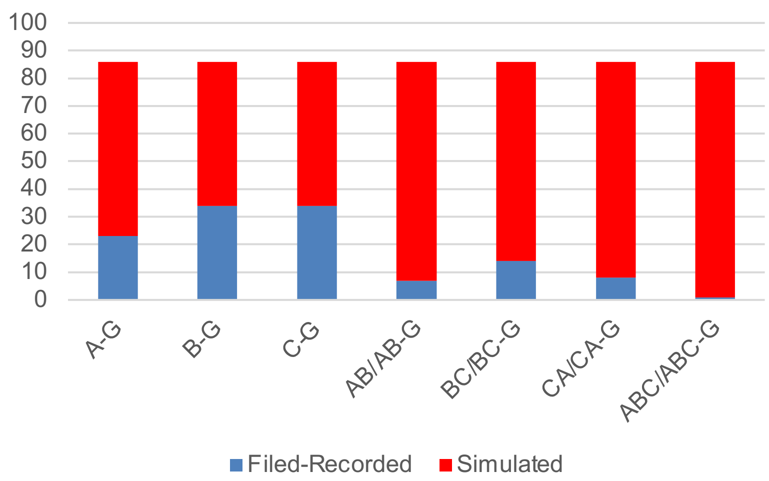
Extracted Features	Type of faults
AB_{diff} = 1 and CA_{diff} = -1 and YZ_{diff} = 1 and ZX_{diff} = -1	"A-G"
AB_{diff} = -1 and BC_{diff} = 1 and XY_{diff} = -1 and ZX_{diff} = 1	"B-G"
BC_{diff} = -1 and CA_{diff} = 1 and XY_{diff} = 1 and YZ_{diff} = -1	"C-G"
BC_{diff} = 1 and CA_{diff} = -1 and XY_{diff} = 1 and YZ_{diff} = -1	"AB/AB-G"
AB_{diff} = -1 and CA_{diff} = 1 and YZ_{diff} = 1 and ZX_{diff} = -1)	"BC/BC-G"
AB_{diff} = 1 and BC_{diff} = -1 and XY_{diff} = -1 and ZX_{diff} = 1	"CA/CA-G"
"ABC/ABC-G" will be assigned if the line fault is not one from all previous combinations	"ABC/ABC-G"

Table 1. Automatic labeling for seven types of line fault

Faults Distribution Before and After Data Integration



Field recorded data distribution(2016)



Integrated data distribution

Benefits of Data Integration: Results

Fault classification accuracy on unseen year 2017 field-recorded data across multiple metrics was **significantly improved by learning from integrated field-recorded and simulated data**

Field-Recorded Data			
Models	**Weighted Precision**	**Weighted Recall**	**F1-score**
SVM	83.25%	91.03%	86.87.%
RF	83.31%	91.03%	86.89%
XGBoost	84.13%	91.03%	87.17%
Micro_average_of_precision_recall			94.90%
Integrated Data			
SVM	**98.69%**	98.62%	**98.58**%*
RF	98.08%	97.93%	97.83%
XGBoost	98.25%	97.93%	97.88%
Micro_average_of_precision_recall			99.20%

TOPIC 3: Transfer Learning for Event Detection from PMU Measurements with Scarce Labels

Motivation

- Event detection tasks are often done using unsupervised approaches since assigning labels manually can be time-consuming and costly
- Unsupervised detectors do not benefit from labelled data that provide the possibility of correcting errors made by unsupervised detectors
- Supervised learning algorithms rely on a sufficient number of precisely labelled data
- Thus, both unsupervised and supervised learning algorithms are infeasible for event detection tasks when labels are scarce and temporally imprecise

This work is a part of **"Big Data Synchrophasor Monitoring and Analytics for Resiliency Tracking (BDSMART)",** an ongoing project funded by the ***US Department of Energy (DOE)***.

- Abdel Hai, A., Dokic, T., Pavlovski, M., Mohamed, T., Saranovic, D., Alqudah, M., Kezunovic, M., Obradovic, Z. "Transfer Learning for Event Detection from PMU Measurements with Scarce Labels," *IEEE Access, 2021*.

Problem Formulation

Objective

- Line fault, frequency, and transformer event detection based on *transfer learning* techniques to detect events based on minimal labeled time windows

Problem Formulation

Given a signal segment originating from multiple PMU devices, predict $y \in \{0, 1\}$, which indicates whether an event occurred during $[t - \Delta, t + \Delta]$

Transfer Learning

- Leverage a small number of relevant labeled data instances from a task related to the target task
- Often, it is used in conjunction with semi-supervised algorithms since semi-supervised algorithms assume only a limited amount of labeled data is available

Challenges: violates traditional machine learning assumptions:

(1) dimensionality of the feature space of the source and target might be different;
(2) marginal distributions could differ (*covariate shift*); and
(3) the same behaviour might have a different meaning in two domains (*concept shift*)

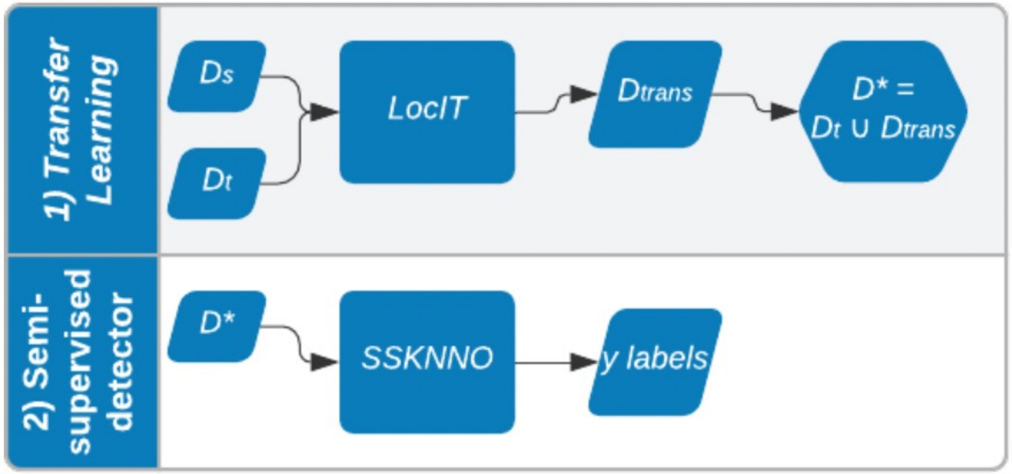
Transfer Learning for Event Detection: Methodology

Localized and unsupervised instance selection (LocIT):

- Transfer an annotated instance $\boldsymbol{x_t}$ from source domain **Ds** to target domain **Dt** if local structure of x_t is similar in Ds and Dt
- Measure distance between the centroids of the nearest neighbourhood N1 in Ds from N2 in Dt
- Also measure the relative distance between the covariance matrices of N1 and N2
- Use these distances to learn the instance-transfer function by training SVM on the target distribution

Semi-supervised anomaly detection (SSKNNO):

- Assign anomaly score $y \in \{0, 1\}$ by a semi-supervised nearest neighbor
  Input:
  - transferred labeled instances from Ds
  - unlabeled target dataset Dt
- Consider the local distribution of Dt when computing the score
- Weight this score by comparing the neighborhoods of transferred instances from Ds



Data

- **38** Phasor Measurement Unit **(PMU) devices** from the Western Interconnection of the U.S.A
- **Signal types:** *vp_m, ip_m, f*
- **2 years** of data (2016, 2017)
- **Geographical locations** of the PMUs and the **network topology** are not made available
- Data are collected with two fps: **30 fps**, and **60 fps**
- **Duplicates** and **outliers** were observed, but did not have a significant impact on our method

Pre-processing

For each time window TW and a specific PMU calculate the Rectangle Area using frequency and positive-sequence voltage magnitude:

$$RA_{PMU,TW} = (f_{max} - f_{min}) * (V_{max} - V_{min})$$

Time windows considered (West).

TABLE I
NUMBER OF LABELS PER CATEGORY AND WINDOW SELECTION METHOD

Event Log	# Event Labels	# Normal Labels	Event Start Time	Event End Time
1-min Labels	1033	923	$ST_{VI} - 5$ sec	$ST_{VI} + 55$ sec
30-sec Labels	1038	1846	$ST_{VI} - 2$ sec	$ST_{VI} + 28$ sec
10-sec Labels	1038	1846	$ST_{VI} - 1$ sec	$ST_{VI} + 9$ sec
5-sec Labels	1038	1846	ST_{VI}	$ST_{VI} + 5$ sec
2-sec Labels	1038	1846	ST_{VI}	$ST_{VI} + 2$ sec

ST_{VI} - start time of the event based on visual inspection.

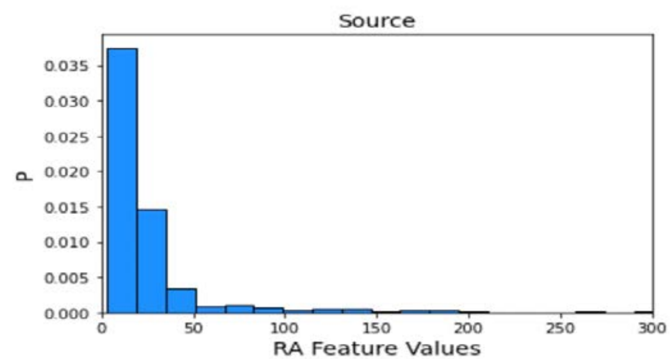
Experiments

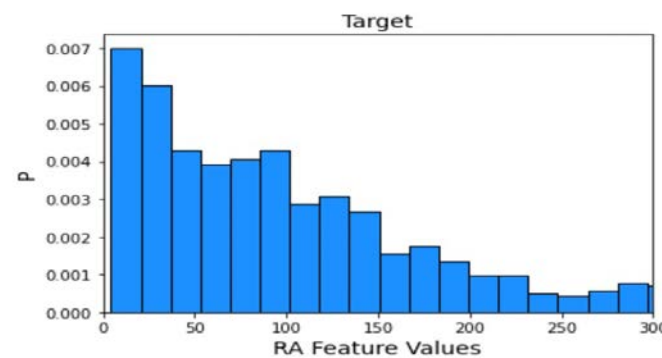
Experiments are based on

- *Temporal Split:*
  - *D_s:* leveraging labeled data from time windows collected from 2016
  - *D_t*: target domain containing unlabeled time windows collected from 2017
  - Feature Vectors containing 38 RA features
- *PMUs Split:*
  - a subset of PMUs was used for D_s and remaining PMUs for D_t
  - Feature Vectors containing 19 RA features

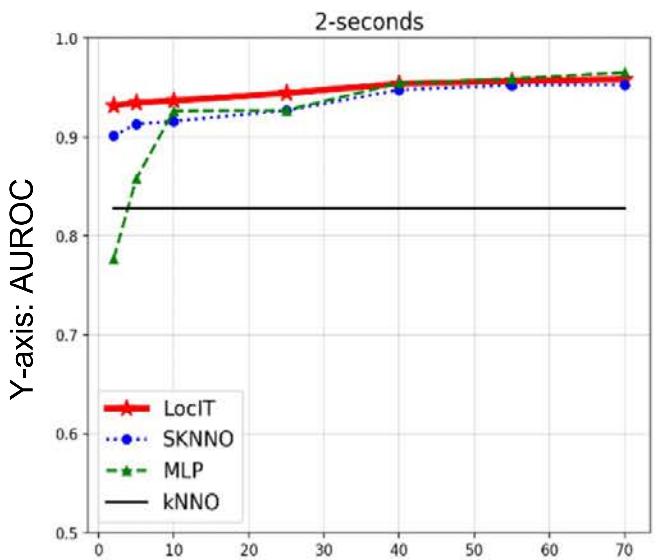
Distributional Difference between Source and Target Datasets

- ***covariate shift assumption** & **concept shift assumption*** validation
- Kolmogorov-Smirnov test was used to check whether the source and target distributions are identical by comparing the underlying distributions *F(x)* and *G(x)* of two independent samples
- Null hypothesis: $F = G$
- Obtained p-values for all PMUs. The maximum p-value was $\mathbf{3.9e^{-15}}$
- Hence, we can safely reject the null hypothesis





Events Detection by Transfer Learning vs. Alternatives



X-axis: percentage of the labeled source data 2%=20 examples, 5%=51 examples, 10%=103 examples

LocIT: transfer learning; SKNNO: semi-supervised; MLP: supervised; kNNO: unsupervised

THE AVERAGE AUROC AND THEIR CORRESPONDING TWO-SIDED CONFIDENCE INTERVAL, CALCULATED AT 90% CONFIDENCE LEVEL

Learning Type	Model	Average AUROC and Confidence Interval
Transfer Learning	LocIT	**0.94 ± 0.0032**
Semi-supervised	SSKNNO	0.90 ± 0.0036
	SSDO	0.81 ± 0.0078
Supervised	MLP	0.74 ± 0.0058
	LR	0.72 ± 0.0074
	KNN	0.76 ± 0.0042
	SVM	0.72 ± 0.0199
Unsupervised	kNNO	0.75 ± 0.0037
	iNNE	0.85 ± 0.0113

<u>Observations: (experiments conducted on a temporal split)</u>

- Performances improve with more labeled data added to the source
- LocIT obtained an average AUROC of 0.94 with a confidence interval width of 0.0032 (**±0.0032**) outperforming baselines with high confidence (p-value vs the second-best SSKNNO method was **$1.4e^{-8}$**)
- When learning from limited labelled data (using <10% of labeled source data from a related task) **transfer learning outperformed unsupervised, semi-supervised, and fully supervised algorithms**

Conclusions

- The ability of machine learning techniques to detect, classify and anticipate events is significantly reduced by
  (1) data anonymization (removal of the grid topology);
  (2) scarce observations;
  (3) low precision labels (log file).

- Manual labeling and precise time stamping of events in big PMU data is not cost-effective and could be infeasible, but our results provide evidence that **monitoring and analytics for resiliency tracking** can be significantly improved by
  – automated feature learning,
  – rapid refinement and partial inspection of labels,
  – events simulations, and
  – transfer learning

- The technology for development and implementation of automated means for characterizing the events is readily available, but related data processing standards and practices are needed to support the technology

Acknowledgements

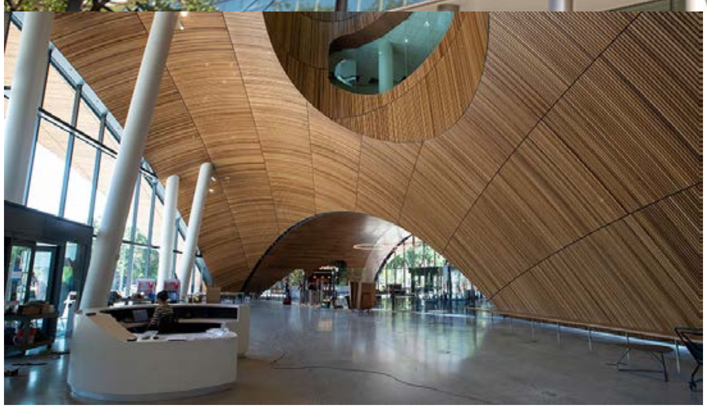
- Funding and Data: **Department of Energy**, Award Number DE-OE0000913
- Collaborators:

  Texas A&M University:
  - Prof. Mladen Kezunovic
  - Rashid Baembitov
  - Dr. Tatjana Dokic
  - Taif Mohamed

  Quanta Technology Inc.:
  - Dr. Yi Hu
  - Dr. Zheyuan Cheng

  Temple University:
  - Ameen Abdel Hai
  - Mohammad Alqudah
  - Hussain Otudi
  - Dr. Martin Pavlovski
  - Daniel Saranovic

Questions?



Research collaboration opportunities:

- **partnerships**
- **visiting scholars**
- **postdocs and Ph.D. students**

Zoran Obradovic
zoran.obradovic@temple.edu
www.dabi.temple.edu/~zoran