The L-CAPE Project at FNAL

Milan Jain*, Jan Strube, Vinay C Amatya, Gihan U Panapitiya
Pacific Northwest National Laboratory, Richland, WA

Beau Harrison, Kyle Hazelwood, Bill Pellico, Brian A. Schupbach, Kiyomi Seiya, Jason St. John
Fermi National Accelerator Laboratory, Batavia, IL

Summary

- Controls System at FNAL
  - Records data asynchronously from several thousand Linac (Linear Accelerator) devices
  - Frequency ranging from 15 Hz down to once per minute
- Current Operations:
  - Mostly Reactive: Investigating cause after the effect
  - On beam interruption, operators use the data to investigate the source of unplanned beam outage from the FNAL Main Control Room.
- Goal
  - Apply data-analytic methods to improve information available to the operators
  - Use machine learning to automate the detection & labeling of outages and
  - Discover patterns in the data that could lead to prediction of outages

Table-1: Energy Use During Beam Outages

Event Length	Event Count	Duration (hours)	Energy (MWh)
< 5s	1626	1.04	0.182
< 10s	321	0.59	0.104
< 60s	205	2.01	0.351
< 120s	201	4.70	0.822
< 10m	169	11.69	2.05
≥ 10m	47	111.35	19.49

Data Sources

- Data Logger
  - Accelerator control system's data logger nodes record data streams into circular buffers
  - Pipeline Created by the developers at the FNAL's Controls Department using modern tools
  - Solves a common problem and is being used on other ML projects
- Data Stats
  - L-CAPE makes 5567 requests over 4292 control system parameters and stores each request in an HDF5 group
  - The HDF5 output is collected by the hour with an average file size of 644MB per period

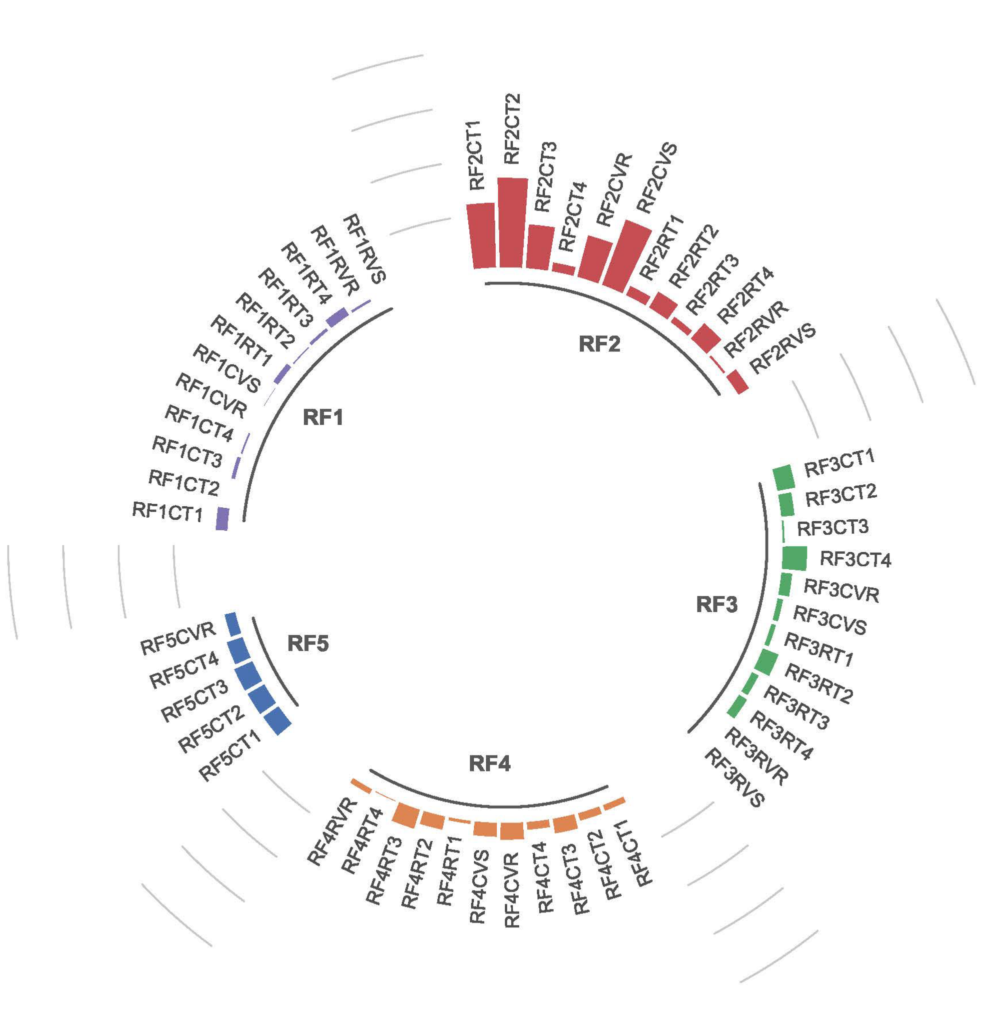

Figure-3: LRF2 Driver Anode OL

Data Preprocessing Procedures

- Challenges with Raw Data
  - Hourly raw data files (in HDF5 format) are sampled from devices
    - at their respective cadences **(different frequencies)**, and
    - timestamped by independent front-end nodes' clocks **(misaligned timestamps)**
  - High disk utilization and read/write time **(bulky and slow)**
  - Requires preprocessing for data analytics & visualization **(inefficient)**
- Preprocessing Procedures
  - Reference Clock: Timestamps from the 15Hz devices (devices with highest frequency)
  - DataFrame: Combined data using reference clock for quick analytics & visualization
  - Parquet+Snappy: Reduced disk utilization by 20x and accelerates read/write ops

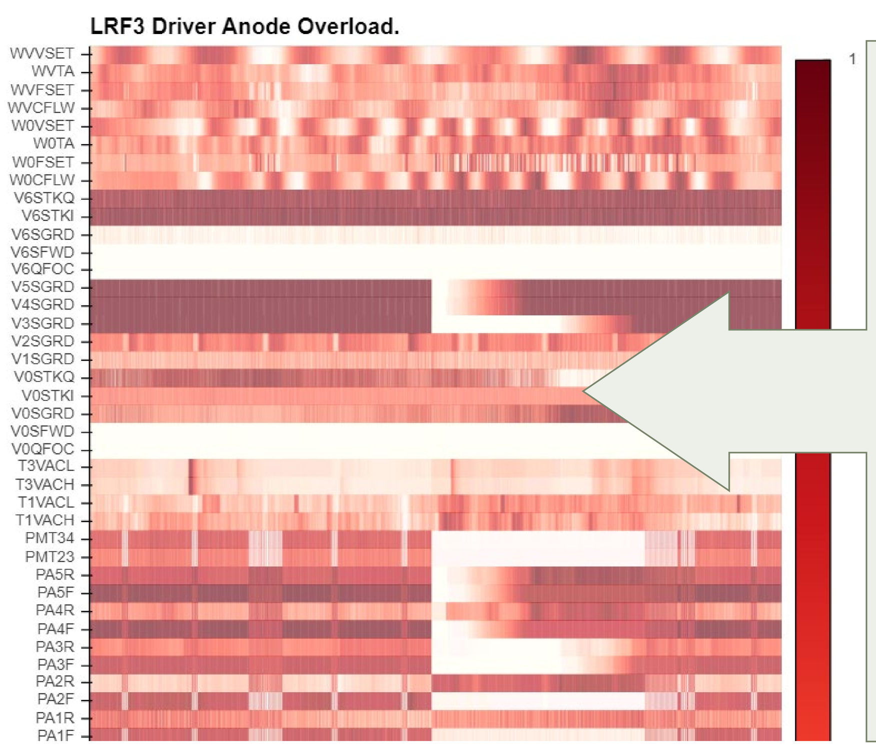

Figure-1: Heatmap Visualization of a Fault

Visualizing Processed Data

- Subset of devices for a single fault instance – "LRF3 Driver Anode Overload"
- For all the devices, data normalized between 0-1
- Varying Start Time and Duration: Different devices show different patterns during this fault
- Pattern can vary:
  - from one instance of a fault to another; and
  - from one fault type to another

Outage Labelling

- Operator Labeling (also referred to as "Ground Truth Labels")
  - List of outages with labels and times reported by the operators
  - Required for analytics, visualization, and model verification and validation
  - No label if the downtime duration is <1 min
- Bit Permits
  - Permits raised by the operators to report downtimes
  - Bit-27 (downstream permit) and Bit-28 (upstream permit)

Modelling & Evaluation

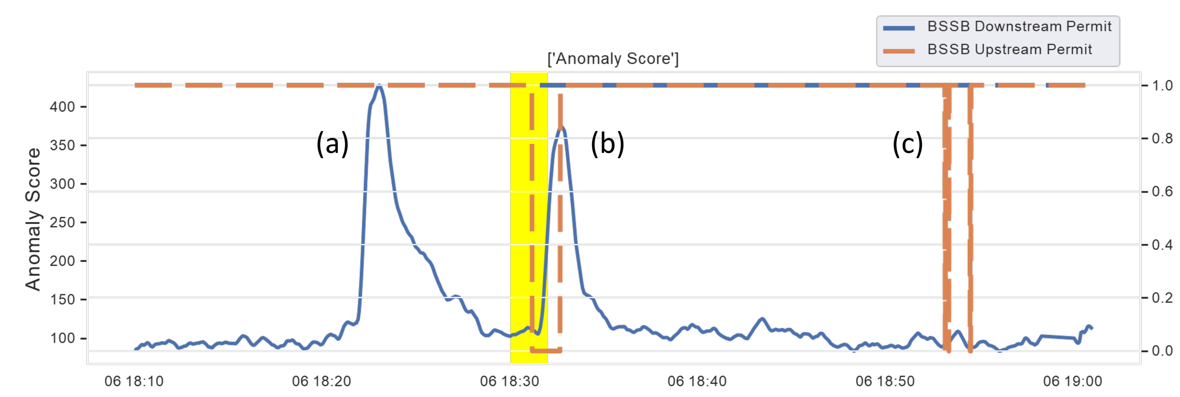

Figure-2: Comparison of anomaly score (TADGAN), operators' labels, and bit permits.
(a) High anomaly score but no reported fault **(False Positive)**
(b) Anomaly score aligns well with ground truth and bit permits **(True Positive)**
(c) No operator label but bit permits report a fault

- #Devices: 2081 (status bits, settings, and noisy devices were filtered)
- Only 53 devices with complex data patterns were trained using the TADGAN

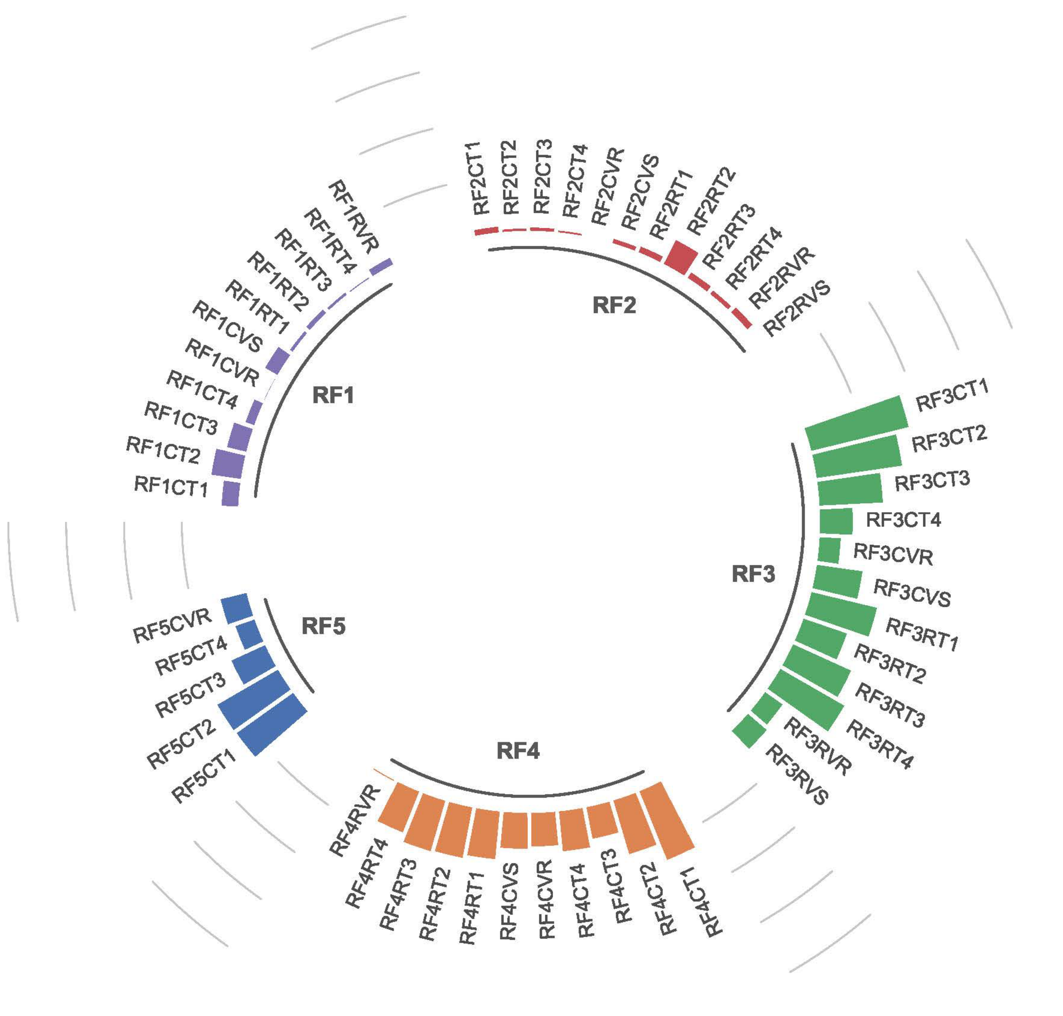

Figure-4: LRF3 Driver Anode OL

Thresholding
Assumes device value lower than the 5th percentile or higher than the 95th percentile is anomalous

Filtering
Data passes through a low-pass filter to filter high-frequency noise followed by a high-pass filter to filter brief anomalies/spikes

TADGAN
An unsupervised anomaly detection approach built on Generative Adversarial Networks (GANs) by Geiger et al. [1]

Table-2: Performance Comparison - Thresholding, Filtering, and TADGAN. Data for March 2021

	#Devices	True Positives	False Negatives	False Positives	Precision	Recall	F1-Score
Thresholding	2081	99	13	210	0.32	0.88	0.47
Filtering	2081	105	7	279	0.27	0.94	0.42
TADGAN	53	51	61	443	0.1	0.45	0.17

- Overall, statistical techniques better (high recall) than TADGAN
- However, statistical techniques tend to perform poorly for:
  - ❑ Devices with multi-modalities
  - ❑ Real-time inference (smaller windows)
- High false positive rate: no operators label for faults <1min
- Ensemble of statistical and ML models required for a scalable and accurate fault detection framework
- Variation in patterns can be exploited for unsupervised labeling (see Figure 3 and 4)

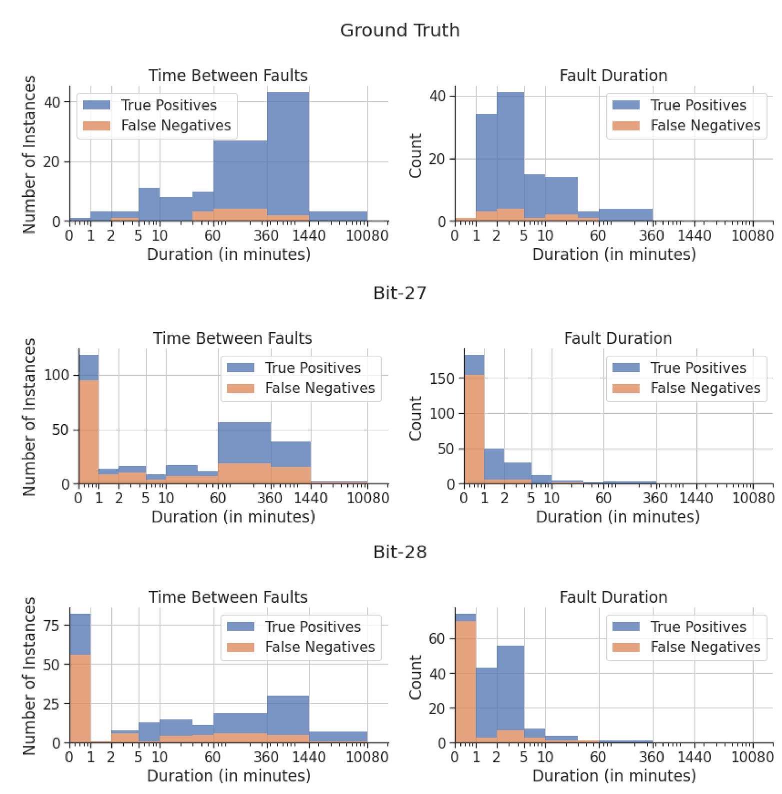

Figure-5: Evaluating the False Negatives

Acknowledgements

This work was supported in part by the U.S. DOE Office of Science, Office of High Energy Physics, under award 76651: "Machine learning for Accelerator Operations". Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC05-76RL01830. It was also partly supported by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

*milan.jain@pnnl.gov

[1] Alexander Geiger, Dongyu Liu, Sarah Alnegheimish, Alfredo Cuesta-Infante, and Kalyan Veeramachaneni. Tadgan: Time series anomaly detection using generative adversarial networks. In 2020 IEEE International Conference on Big Data (Big Data), pages 33–43. IEEE, 2020.

NAPAC 2022 Pacific Northwest NATIONAL LABORATORY Fermilab