

Event Detection in Micro-PMU Data: A Generative Adversarial Network Scoring Method

Armin Aligholian¹, Alireza Shahsavari², Ed Cortez³, Emma Stewart⁴ and Hamed Mohsenian-Rad¹

¹Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA

²Los Angeles Department of Water and Power, Los Angeles, CA, USA

³Energy Delivery Engineering Group, Riverside Public Utilities, Riverside, CA, USA

⁴Defense Infrastructure Group, Lawrence Livermore National Laboratory, Livermore, CA, USA

Abstract—A new data-driven method is proposed to detect events in the data streams from *distribution-level phasor measurement units*, a.k.a., micro-PMUs. The proposed method is developed by constructing *unsupervised deep learning anomaly detection models*; thus, providing event detection algorithms that require *no or minimal human knowledge*. First, we develop the core components of our approach based on a Generative Adversarial Network (GAN) model. We refer to this method as the *basic method*. It uses the same features that are often used in the literature to detect events in micro-PMU data. Next, we propose a second method, which we refer to as the *enhanced method*, which is enforced with additional feature analysis. Both methods can detect *point signatures* on single features and also *group signatures* on multiple features. This capability can address the unbalanced nature of power distribution circuits. The proposed methods are evaluated using *real-world* micro-PMU data. We show that both methods highly outperform a state-of-the-art statistical method in terms of the event detection accuracy. The enhanced method also outperforms the basic method.

Keywords: Micro-PMU data, power distribution, event detection, deep learning, generative adversarial network, feature analysis.

I. INTRODUCTION

The voltage and current measurements that are reported by distribution-level phasor measurement units, a.k.a., micro-PMUs, have high-resolution and high-precision. They significantly enhance our visibility into the distribution grid, [1], [2]. Applications of micro-PMU data include topology and phase identification [3], load modeling, [4], state estimation [5], asset monitoring [6], and distribution system cyber-security [7].

An important and emerging class of studies when it comes to micro-PMU data is to investigate the *events* in distribution systems. Here, an event is defined rather broadly and may refer to load switching, capacitor bank switching, connection or disconnection of distributed energy resources (DERs), inverter malfunction, a minor fault, a signature for an incipient fault, etc. [8]–[10]. Undoubtedly, the very first step to investigate the events in micro-PMU data is to *detect* them from the large volume of data that is being collected. Note that, each micro-PMU reports over *one hundred million data points* every day.

The literature on event detection in micro-PMU data can be divided into two broad classes; namely *statistical methods*, such as in [11]–[13], and *machine learning methods*, such as in [14], [15]. The statistical method in [10] uses the absolute

deviation around median, combined with dynamic window sizes. In [13], the analysis of the inverse power flow problem is combined with the turning point test method to detect events. In [11], the physical equations of the power distribution circuits are combined with techniques from statistical quality control in order to develop a hierarchical anomaly detection architecture that uses data from optimally placed micro-PMUs.

While we use the state-of-the-art statistical method in [10] as a *benchmark* for performance comparison in this paper, our approach here is rather based on machine learning. In [14], a machine learning method, called ensembles of bundle classifiers, is used to train multiple classifiers based on multiple instances of the same predetermined event, so that the patterns of that event are learned in order to detect more instances of that event in the micro-PMU data. In [15], a hidden structure semi-supervised machine learning model is established to combine micro-PMU data for both labeled and unlabeled events. A parametric dual optimization procedure is used to tackle the non-convex learning objective function.

The event detection method in [14] is based on *supervised* machine learning. Also, the method in [15] is based on *semi-supervised* machine learning. In both cases, full or partial *expert knowledge* is needed in order to establish the event detection tool. In contrast, in this paper, we develop an *unsupervised* method to detect events in micro-PMU data. This eliminates the need for human labor; which makes our proposed approach suitable for implementation in practice when we work with large volumes of micro-PMU data. It is worth adding that unsupervised learning is previously used in the analysis of micro-PMU data, but as a feature extraction tool for *event classification* [16]. Therefore, it has *not* been used for event detection, which is the focus of this paper.

The event detection methods that are proposed in this paper, work by constructing deep learning Generative Adversarial Network (GAN) models. The contributions are as follows:

- To the best of our knowledge, this is the first paper to develop event detection methods for micro-PMU data based on GAN models. Two such methods are proposed. The first one, called the basic method, trains a single GAN model. The second one, called the enhanced method, involves additional analysis of the features of the micro-PMU data; which leads to training two GAN models.
- Both methods are *unsupervised* deep learning methods, which require no or minimal human knowledge; which

makes them suitable for automated and scalable operation. Furthermore, they can detect both *point-signatures* and *group-signatures* in micro-PMU data. This is an important capability because of the unbalanced nature of power distribution circuits; where many events may affect only a subset of the features on only one or two phases.

- Real-world micro-PMU data is used to evaluate the proposed event detection methods. In order to create a reference, first, more than 1000 events of different kinds are extracted manually from the micro-PMU data within a given period of time. It is observed that both the basic and the enhanced methods highly outperform a prevalent statistical method. The advantage is particularly major for the events that cause *small changes in magnitude*. Also, the enhanced method outperforms the basic method; thus confirming the importance of the revised model.

II. METHODOLOGY

A. Basic Method

In its core, the proposed basic event detection method uses a GAN model which has two components, a *generator* and a *discriminator*. The generator is a deep neural network that tends to *produce* data samples that follow the distribution of the historical training data. The discriminator is a deep neural network that tends to *distinguish* between the data samples generated by the generator and the true historical data. By training the generator and the discriminator subsequently and iteratively, the GAN model can achieve an *equilibrium*, at which the discriminator can no longer distinguish between the distribution of the generated samples and the historical data, which means the training of the GAN model is complete.

1) *Features*: As in [10], [11], we use the following time-series as the features to train the GAN model in our basic method: 1) magnitude of voltage, i.e., V ; 2) magnitude of current, i.e., I ; 3) active power, i.e., P ; and 4) reactive power, i.e., Q . All these features are defined separately for each three phases. Therefore, in total, the GAN model is trained with 12 time-series. Note that, while micro-PMUs measure V and I directly, P and Q are obtained rather indirectly by combining V and I with the measurements on voltage phase angle and current phase angle, which are both provided by micro-PMUs.

2) *Generator*: It takes a noise vector z from a distribution function $p_z(z)$, such as $z \sim \mathcal{N}(\mu_z, \sigma_z^2)$, and tries to produce samples similar to the ones from the true sample distribution. We seek to train a neural network $G(z, \theta_g)$ to generate samples which follow the distribution of the historical data. Here, θ_g denotes weights of the generator network. Mathematically, we seek to minimize the following objective function [17]:

$$\frac{1}{N} \sum_{i=1}^N [\log(1 - D(G(z_i)))], \quad (1)$$

where N is the number of samples in each training batch, D is the discriminator function, G is the generator function, and z_i is the random vector for i th generated sample. In order to train the generator, after forward propagation, we need to update the generator parameters by calculating gradient and using a proper optimizer, such as Adam optimizer [18].

Algorithm 1 Event Detection - Basic Method

Input: Training data and test data: V, I, P and Q .

Output: Event Detection Flag F .

// Learning Phase

Train the *GAN* model.

Use the Discriminator as scoring function $D^*(\cdot)$.

Calculate the scores for the training data.

Fit a Normal PDF $\mathcal{N}(\mu, \sigma^2)$ to the obtained scores.

// Detection Phase

For each new micro-PMU test data **Do**

 Calculate the score s using $D^*(\cdot)$.

If $s \notin (\mu - z_p\delta, \mu + z_p\delta)$ **Then**

$F = 1$ // Event

Else

$F = 0$ // No Event

End

End

3) *Discriminator*: It is meant to distinguish between the fake data samples generated by the generator and the real measurements. Our goal is to train a neural network $D(x, \theta_d)$, which creates a single scalar value as its output. Here, x is the vector of the actual measurement data and θ_d is the weights of the discriminator network. The primary objective of the discriminator is to maximize the probability of distinguishing between the true measurement data and the data generated by the generator. Therefore, we seek to minimize:

$$\frac{1}{N} \sum_{i=1}^N [\log(D(x_i)) + \log(1 - D(G(z_i)))], \quad (2)$$

where x_i is the i th real sample and the second term is the same as the term in (1).

Together, the generator and the discriminator play a *min-max* game with the following value function:

$$V(G, D) = \mathbb{E}_x \sim p_{data}(x) [\log(D(x))] + \mathbb{E}_z \sim p_z(z) [\log(1 - D(G(z)))]. \quad (3)$$

4) *Training*: Both the generator and discriminator are formed with Long Short-Term Memory (LSTM) modules, which are connected back-to-back to capture the relationship between different features and their time dependencies. The micro-PMU data is normalized and segregated into sequences of training blocks. At the beginning, the samples generated by the generator are very different from the measurement data; thus, the discriminator can reject them with high confidence. As time goes by, the value of $V(G, D)$ can attain its global optimum by satisfying the following two conditions:

- **C1**: For any fixed G , the optimal discriminator D^* is:

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)}. \quad (4)$$

- **C2**: There exists a global solution such that:

$$\min_D (\max_G (V(G, D))) \iff p_g(x) = p_{data}(x). \quad (5)$$

If these conditions are not satisfied at the equilibrium, then the training is repeated with new random initial points. More details on the training mechanism can be found in [17].

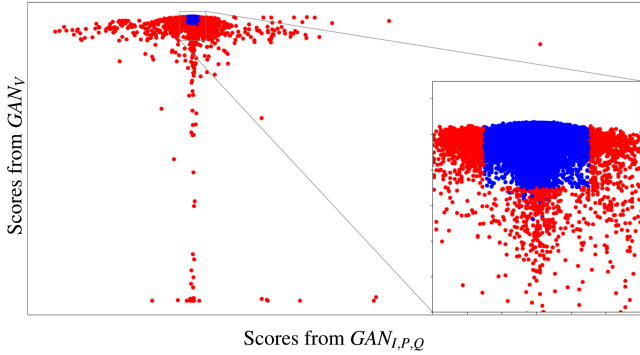


Fig. 1: The importance of using two GAN models in the enhanced method: while the scores from the $GAN_{I,P,Q}$ model can detect most events; there are events that are detected only if the scores from the GAN_V model are also considered. Blue dots denote normal data while red dots denote events.

5) *Event Scoring*: After training the basic model, the blocks of micro-PMU data stream are passed to the discriminator and the output is a scalar number which is defined as *score*. We pass the whole training set to the discriminator and calculate the scores. A normal probability distribution function (pdf) is fitted to the obtained scores, i.e., $scores \sim \mathcal{N}(\mu, \sigma^2)$, due to the fact that these scores must be very close to the global optimum, see (4) and (5). This is because of the infrequent nature of the events in power distribution systems.

6) *Algorithm*: The proposed basic event detection method is summarized in Algorithm 1. It works based on the fact that events in micro-PMU data are infrequent. In fact, our analysis of the real-world micro-PMU data shows that events occur at about 0.04% of the times. Thus, the default for the trained model must be the normal operation of the power distribution system. Accordingly, the generator is trained how to generate data that follow the distribution of measurements during normal operation, i.e., in the absence of any event. As a result, the discriminator is essentially trained to distinguish between the absence and the presence of the events, which is exactly what is needed in order to detect the events. It should be noted that, a common choice for parameter z_p in the threshold $\mu \pm z_p \sigma$ is 3, known as the three-sigma rule [19].

B. Enhanced Method

The basic method in Section II.A requires training a *single* GAN model, where the features are V, I, P , and Q . However, given the characteristics of the micro-PMU data, in this section, we propose to develop and train *two* separate GAN models, one for the voltage measurements V , and another one for the rest of the measurements, i.e., I, P , and Q . This results in developing a new event detection method that has a significantly better accuracy. Next, we explain the underlying reasons for the proposed changes and also new algorithm.

1) *Feature Analysis*: After applying the basic method to real-world micro-PMU data, we observed that Algorithm 1 sometimes fails to detect events that demonstrate signatures only in voltage magnitude. Such event cannot trigger the score to exceed the threshold. Further investigation revealed that this is because, in power distribution systems, voltage measurements are much less volatile than current measurements.

Algorithm 2 Event Detection - Enhanced Method

Input: Training data and test data: V, I, P and Q .

Output: Event Detection Flag F .

// **Learning Phase**

Train the $GAN_{I,P,Q}$ model.

Use the Discriminator as scoring function $D_{I,P,Q}^*(\cdot)$.

Calculate the scores for the training data.

Fit a Normal PDF $\mathcal{N}(\mu, \sigma^2)$ to the obtained scores.

Train the GAN_V model.

Use the Discriminator as scoring function $D_V^*(\cdot)$.

Calculate the scores for the training data.

Fit a Normal PDF $\mathcal{N}(\phi, \varphi^2)$ to the obtained scores.

// **Detection Phase**

For each new micro-PMU test data **Do**

 Calculate the score s_1 using $D_{I,P,Q}^*(\cdot)$.

 Calculate the score s_2 using $D_V^*(\cdot)$.

If $s_1 \notin (\mu - z_p \delta, \mu + z_p \delta)$ **or**
 $s_2 \notin (\phi - z_p \varphi, \phi + z_p \varphi)$ **Then**

$F = 1$ // Event

Else

$F = 0$ // No Event

End

End

Therefore, the GAN model sometimes cannot properly extract the characteristics of the voltage measurements.

2) *Training Multiple GAN Models*: To remedy the above issue, we propose to construct two separate GAN models that are trained in parallel. One GAN model, denoted by GAN_V , has 3 features as its input, which are the voltage magnitude measurements across the three phases. The other GAN model, denoted by $GAN_{I,P,Q}$, has 9 features as its input, which are current magnitude, active power, and reactive power measurements across the three phases. Importantly, we should not have more than two GAN models. In fact, it is observed that I has high correlations with P and even Q , which makes it desirable to combine I, P , and Q into one GAN model; as opposed to having four GAN models for V, I, P , and Q .

3) *Event Scoring*: Once each of the two GAN models is trained, the resulting Discriminator function is used to generate its own scores. An example for the scores that are generated by the two GAN models are shown in Fig. 1. The *blue dots* represent *normal data*. The red dots represent *events*. We can see that each of the two GAN models detects only a sub-set of events. The events that are scattered across x-axis are the ones that are detected by $GAN_{I,P,Q}$. They include the majority of the events. The events that are scattered across y-axis are the ones that are detected by GAN_V . Thus, both GAN models are both needed to enhance accuracy of event detection.

4) *Algorithm*: The proposed enhanced event detection method is summarized in Algorithm 2. It works by examining the scores of the two separate GANs; thus having a dedicated deep learning architecture to detect the events in voltage magnitude and another deep learning architecture to detect the events that involve the current, active power, and reactive power. The rest of the algorithm is similar to Algorithm 1.

TABLE I: Event Detection Accuracy

	Benchmark [10]	Basic Method	Enhanced Method
Accuracy	0.3640	0.6943	0.8805
F1-score	0.3614	0.7676	0.9023

III. EXPERIMENTAL RESULTS

The proposed event detection methods are applied to the *real-world* data from a distribution feeder in Riverside, CA [10]. The resolution of the data is 120 readings per second. In total, 1.8 billion measurement points are analyzed. In particular, two weeks of data are used to train the GAN models. One day of data is used to test the event detection methods. Event detection is applied on windows of size 40 data points. Each window has an overlap of size 20 data points with the next window in order to assure not missing any event.

A. Performance Comparison

The performance comparison in this paper is done not only between the basic and the advanced methods; but also with respect to the prevalent statistical event detection method in [10]. The effectiveness of the event detection methods is investigated over 1000 reference events in micro-PMU data, that are visually extracted within a specific period of time.

The summary of the results are shown in Table I. We can see that the basic method significantly outperforms the benchmark statistical event detection method in [10]. Furthermore, the enhanced method considerably outperforms the basic method. Next, we explain the underlying causes for these differences by going through several examples of the events that are detected.

B. Assessment of the Basic Method

Figs. 2 to 6 show five examples of the events that are detected by the basic method. Importantly, the prevalent statistical method in [10] detected only the first two of such events. Regarding the events in Figs. 4 and 5, they are not detected by the method in [10] because the changes in the magnitudes are relatively small and do not significantly affect the statistical measures, such as the absolute deviation around median. As for the event in Fig. 6, all the pieces of this long event are detected by the basic method at several subsequent windows of the data. However, the statistical method in [10] only captures the step change the beginning of this event; because the statistical characteristics remain the same afterwards.

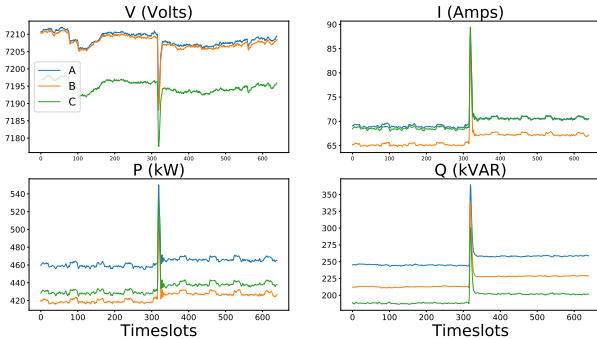


Fig. 2: Inrush current with impact on all features. This event is detected by all the three methods: statistical, basic, and enhanced.

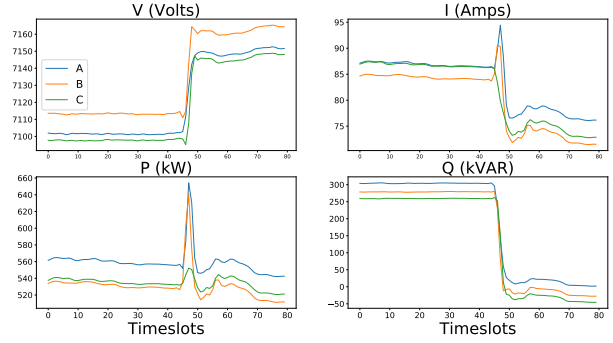


Fig. 3: Capacitor bank switching with impact on all features. This event is detected by all the three methods: statistical, basic, and enhanced.

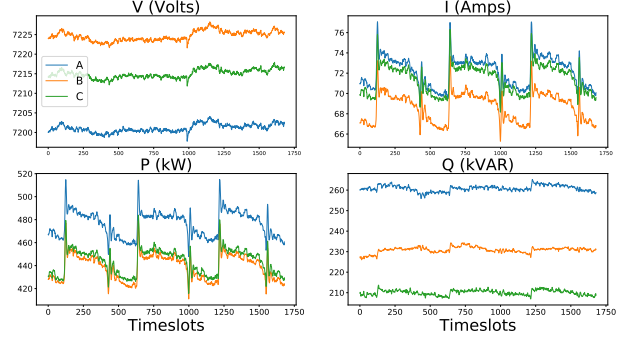


Fig. 4: An event with major impact only on current and active power. This event is detected by the basic method, but not by the statistical method.

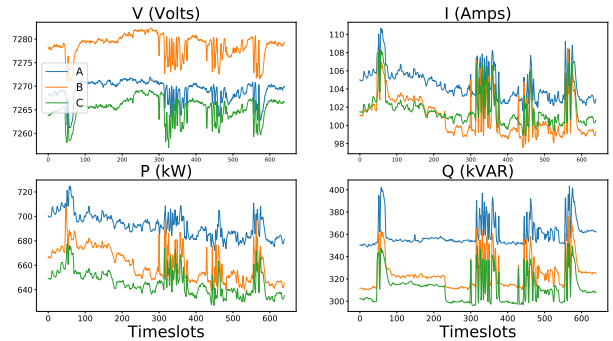


Fig. 5: An event involving oscillations. This event is detected by the basic method, but it is not detected by the statistical method.

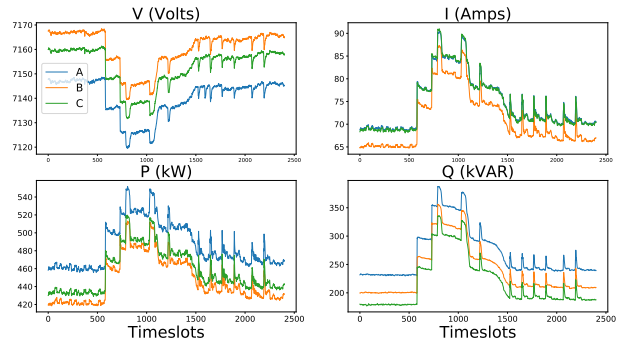


Fig. 6: A rare and long event with 20 seconds of transient signature. All pieces of this long event are detected and captured by the basic method. The statistical method only detects a step change at the beginning of this event.

C. Assessment of the Enhanced Method

Figures 7 and 8 show two events that are detected by the

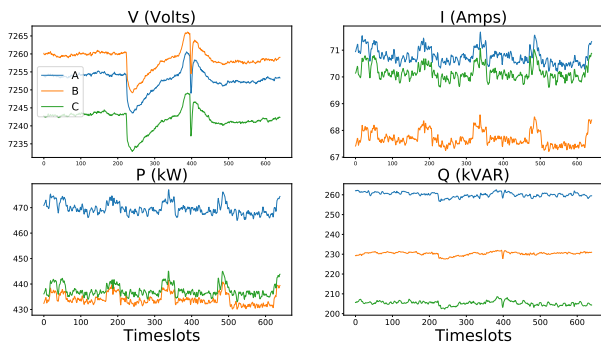


Fig. 7: An event with impact mainly on voltage. It is detected by the enhanced method. But it is not detected by the basic method or the statistical method.

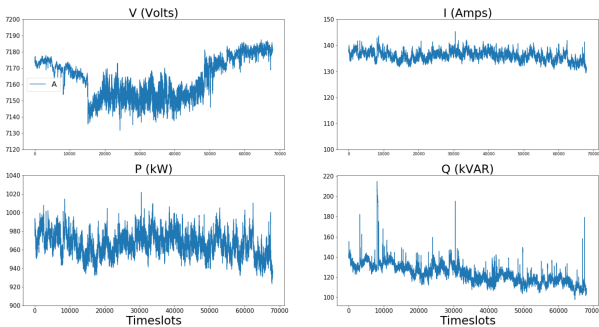


Fig. 8: An event with momentary and damping oscillations in voltage, shown on one phase only. This event is detected by the enhanced method. But it is not detected by the basic method or the statistical method.

enhanced method. But they are not detected by either the prevalent statistical method in [10] or even the basic method. The basic method fails to detect these two events because the main signatures are in voltage and they are relatively small in magnitude. Therefore, only the additional GAN model in the enhanced method can capture these events. This demonstrates the importance of the change in the model that was proposed in the enhanced method. Regarding the event in Fig. 8, it demonstrates momentary oscillations that started only after some sort of actions, possibly a tap changing event, where the oscillations damped after a short period of time. Events like this are important, for example, for asset monitoring. However, only the enhanced method was able to detect such event.

IV. CONCLUSIONS

Two novel unsupervised deep learning methods are proposed to detect events in micro-PMU data streams. They work by constructing Generative Adversarial Network (GAN) models. They are capable of extracting the characteristics of a wide variety of events in large volumes of micro-PMU data. The basic method involves a single GAN model. The enhanced method is equipped with additional analysis of features. It involves training two parallel GAN models. Both methods are capable of detecting events with point-signatures and group-signatures. They are particularly well-suited to detect the events in distribution systems where the event may impact only a subset of the features and only or two phases; in addition to the cases that all three phases are affected. Real-world data from micro-PMU field installation is used to evaluate

the performance of the proposed event detection methods. It is observed that the basic method significantly outperforms a prevalent statistical event detection method in the literature. Furthermore, the enhanced method considerably improves the performance over the basic method. Several examples of the events that detected by different methods are shown and discussed in order to understand the characteristics of the proposed unsupervised event detection methods.

REFERENCES

- [1] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov 2017.
- [2] H. Mohsenian-Rad, E. Stewart, and E. Cortez, "Distribution synchrophasors: Pairing big data with analytics to create actionable information," *IEEE Power and Energy Magazine*, vol. 16, no. 3, pp. 26–34, May 2018.
- [3] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, "Topology identification in distribution systems using line current sensors: An ml approach," *IEEE Trans. on Smart Grid*, pp. 1–1, 2019.
- [4] A. Shahsavari, M. Farajollahi, and H. Mohsenian-Rad, "Individual load model parameter estimation in distribution systems using load switching events," *IEEE Trans. on Power Systems*, vol. 34, no. 6, pp. 4652–4664, Nov 2019.
- [5] M. Farajollahi, A. Shahsavari, and H. Mohsenian-Rad, "Linear distribution system state estimation using synchrophasor data and pseudo-measurement," in *2019 International Conference on Smart Grid Synchronized Measurements and Analytics (SGSMA)*, May 2019, pp. 1–6.
- [6] A. Shahsavari, M. Farajollahi, E. Stewart, A. von Meier, L. Alvarez, E. Cortez, and H. Mohsenian-Rad, "A data-driven analysis of capacitor bank operation at a distribution feeder using micro-pmu data," in *Proc. of the IEEE PES ISGT*, April 2017, pp. 1–5.
- [7] M. Jamei, E. Stewart, S. Peisert, A. Scaglione, C. McParland, C. Roberts, and A. McEachern, "Micro synchrophasor-based intrusion detection in automated distribution systems: Toward critical infrastructure security," *IEEE Internet Computing*, vol. 20, no. 5, pp. 18–27, Sep. 2016.
- [8] A. von Meier, E. Stewart, A. McEachern, M. Andersen, and L. Mehrmanesh, "Precision micro-synchrophasors for distribution systems: A summary of applications," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2926–2936, Nov 2017.
- [9] M. F. McGranaghan and S. Santoso, "Challenges and trends in analyses of electric power quality measurement data," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, p. 057985, Dec 2007. [Online]. Available: <https://doi.org/10.1155/2007/57985>
- [10] A. Shahsavari, M. Farajollahi, E. Stewart, E. Cortez, and H. Mohsenian-Rad, "Situational awareness in distribution grid using micro-pmu data: A machine learning approach," *IEEE Trans. on Smart Grid*, vol. 10, pp. 6167–6177, Nov. 2019.
- [11] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Anomaly detection using optimally placed μ PMUsensors in distribution grids," *IEEE Trans. on Power Systems*, vol. 33, no. 4, pp. 3611–3623, July 2018.
- [12] A. Shahsavari, M. Farajollahi, E. Stewart, E. Cortez, and H. Mohsenian-Rad, "A machine learning approach to event analysis in distribution feeders using distribution synchrophasors," in *IEEE SGSMA*, May 2019.
- [13] O. Ardakanian, Y. Yuan, R. Dobbe, A. von Meier, S. Low, and C. Tomlin, "Event detection and localization in distribution grids with phasor measurement units," in *IEEE PES General Meeting*, July 2017.
- [14] Y. Zhou, R. Arghandeh, and C. Spanos, "Distribution network event detection with ensembles of bundle classifiers," in *IEEE PES General Meeting*, 06 2016.
- [15] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Partial knowledge data-driven event detection for power distribution networks," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 5152–5162, Sep. 2018.
- [16] I. Niazazari and H. Livani, "Disruptive event classification using pmu data in distribution networks," in *IEEE PES General Meeting*, July 2017.
- [17] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," *arXiv e-prints*, p. arXiv:1406.2661, Jun 2014.
- [18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.
- [19] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994. [Online]. Available: <http://www.jstor.org/stable/2684253>