Title

Supporting Cyber Security of Power Distribution Systems by Detecting Differences Between Real-time Micro-Synchrophasor Measurements and Cyber-Reported SCADA - Final Report

Authors

Peisert, Sean
Roberts, Ciaran
Scaglione, Anna
et al.

Publication Date

2018-10-15

Peer reviewed

Supporting Cyber Security of Power Distribution Systems by Detecting Differences Between Real-time Micro-Synchrophasor Measurements and Cyber-Reported SCADA

Final Report

https://dst.lbl.gov/security/project/ceds-upmu/

https://powerdata.lbl.gov

Principal Investigator: Sean Peisert (sppeisert@lbl.gov)
Co-Principal Investigators: Ciaran Roberts (cmroberts@lbl.gov), Anna Scaglione (ASU)
Senior Personnel: Reinhard Gentz (LBNL), Charles McParland (LBNL),
Alex McEachren (PSL), Galen Rasche (EPRI), Aaron Snyder (EnerNex)
Graduate Students: Mahdi Jamei (ASU)

Lawrence Berkeley National Laboratory (LBNL),
Arizona State University (ASU),
Electric Power Research Institute (EPRI),
EnerNex,
Power Standards Lab (PSL)

Original Version: October 15, 2018
Last Revised December 20, 2020

1 Introduction

As modern power grids tend towards greater levels of automation and communication, the challenges of identifying and mitigating vulnerabilities to cyber-attacks are ones that are increasingly demanding attention. Today's power system has evolved to form the foundational bedrock of modern society, and an attack on this infrastructure could prove disastrous. In this project we were tasked to investigate the use of distribution synchrophasors as an independent isolated sensor network with which we can corroborate, or flag potentially spoofed, Supervisory Control And Data Acquisition (SCADA) data. We adapted an approach to marry the underlying physical properties of power systems with the network communications used by power systems in order to offer insights unattainable by either data stream in isolation. While the concept of intrusion detection systems (IDS) is well understood for monitoring network traffic and traditional IT computing systems, the approach discussed in this report is motivated by several key notions: first, current SCADA communications alone presents an incomplete view of the grid. Second, the power grid, and the equipment controlling it, is grounded by laws of physics. Given this, we leverage high-frequency physical grid measurements to understand the physical condition of the grid, and combine this with SCADA. While high-frequency physical grid measurements and SCADA communication over Internet Protocol (IP) networks are fundamentally disparate information sources, when collectively examined through appropriate lenses, they offer a much more nuanced depiction of the grid. This concept follows our earlier work of leveraging insights regarding the physical behavior and limitations of network-connected control systems to detect cyber-attacks [MPS14].

Lawrence Berkeley National Laboratory (LBNL), in collaboration with Arizona State University (ASU), the Electric Power Research Institute (EPRI), EnerNex, Power Standards Lab (PSL), and its utility partners, have been investigating such an approach whereby physical grid measurements are analyzed in conjunction with SCADA captured traffic in order to determine whether an adversary may have infiltrated a utility's SCADA system. Within the scope of this work, we have focused on using distribution phasor measurement unit (PMU) data with high output sample rate alongside SCADA traffic to examine the state of the distribution grid. The PMUs and computing infrastructure that analyze this data would ideally be isolated from the potentially infiltrated SCADA network local to distribution systems, and would offer system operators and operations engineers a redundant and independent view of the network.

Section 2 begins with an overview of the system architecture that facilitates the data acquisition and processing. Section 3 outlines the specific algorithms developed and their validation on both simulated and field data while Section 4 concludes with a summary of the work performed and our industry engagement activities.

2 System Architecture Overview

An overview of our system architecture, the LBNL Stream-Processing Architecture for near-Real-time Cyber-physical Security (SPARCS), is shown in Fig. 1.

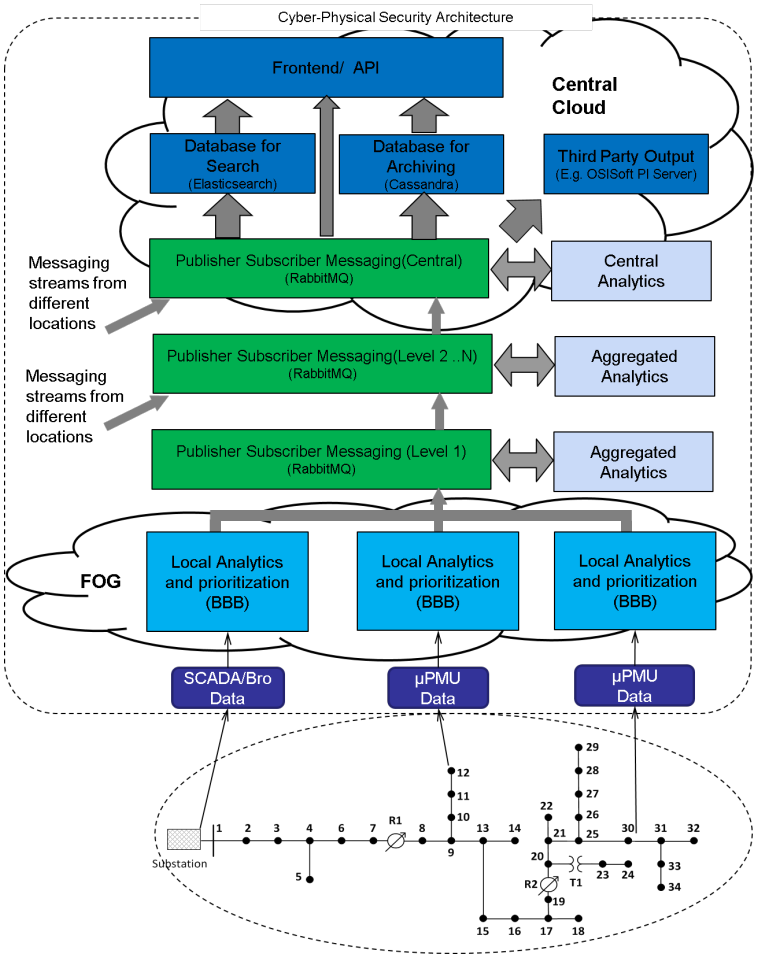


Figure 1: Stream-Processing Architecture Overview

The SPARCS architecture, which leverages open source tools, and is itself also open source,[1] is designed to support the collection, analysis, and storage of the high-frequency power data and SCADA data that we use to detect cyberattacks against distribution grid equipment.

The main novelty of our architecture lies in the hierarchical design, and the application of "edge" computation where possible, so that unnecessary data need not be transported across low-bandwidth (e.g., cellular) network connections from potentially very remote locations in

[1]SPARCS: `https://github.com/lbnl-cybersecurity/sparcs`

the power distribution grid. This approach has multiple benefits: first, the local portion of the data analysis allows for minimal latency to react on local events, even in situations where wide area communications are not available, and could have protected against the attacks described in, for instance [Smi14]. Furthermore, the paradigm of "edge" computing also allows prioritization and annotation of data going upstream when anomalies are detected, over the transmission of raw sensor data and SCADA heartbeats. If desired, data reduction mechanisms can be applied, which is especially of interest when network capacity is limited or when pay-per-use connections apply, such as over cellular. Additionally, since the "edge" computing is directly attached to the sensor, it can effectively shield the sensor from the network and therefore allows the sensor to become less reliant on security patches. This is important, as sensors typically live longer than they are supported by the vendor with timely (security) updates [HMCP04]. Instead, we only have a single, unified, local "edge" computing device, that is inexpensive and easy to maintain and upgrade if needed.

The "edge" or "local" computing device that we currently use is the BeagleBone Black ("BBB") [Col13], although a variety of low-power, low-cost computing devices could be used as long as the device has two network interface cards (NICs) to function as a bastion host between the μPMU and the network connection to the Internet.

2.1 Data Acquisition

2.1.1 SCADA Data

In order to generate synthetic SCADA traffic, We use the openDNP3 library that implements the DNP3 communication protocol for generating and communicating simulated SCADA data [ope]. We then utilize the Bro network security monitor [Pax99] to analyze captured DNP3 packets (e.g., via SPAN ports, port mirroring, etc..), and forward them to a SCADA data analytic module that we have developed. This module leverages the DNP3 parser built into the Bro Network Security Monitor. The analytic results from this module are then correlated with μPMU data analytic results for anomaly detection in the power grid.

2.1.2 Synchrophasor data from μPMUs

Two methods are used to transfer synchrophasor data from a μPMU to a directly connected BBB. The first method is using the IEEE C37.118.2 protocol that is commonly used for synchrophasor data transfer for power systems. This method is used for live-streaming of data; hence it cannot make up for missing data in the event of loss of communication. Therefore, the second method (reading from the internal memory of the μPMU using FTP) serves as a work around for missing data. In this method, the μPMU stores its sampled data worth a certain time duration in its internal memory. When the BBB (any device consuming

synchrophasor data from the μPMU) detects that data is missing, it retrieves the missing data from the μPMU's internal memory using the FTP protocol.

2.2 Messaging system & Field Site Computations

The result of local computations (and raw data if needed) are sent upstream to a decentralized publish-subscribe messaging system (RabbitMQ), that distributes data to its individual destinations, which can include other layers of edge/local-computing or central systems. In our previous work [JSR+17a] we described a number of local data analyses currently running. We term this collection of analyses the "μPMU fog" as it represents a collection of local, edge computing and central and potentially cloud computing resources. These inference rules are agnostic about infrastructure parameters, and use only local information to process and label the data.

In our implementation of SPARCS, we did not adopt a fully-distributed messaging system, such as ZeroMQ [Zer18], in order to maintain low network overhead. Instead, we deployed one messaging system per field-site/area, each with their associated local security analytics using the corresponding sensor feeds. We do this analysis of the data locally for the same reasons as indicated earlier (latency, resilience, prioritization, data reduction), only now with data from multiple sources available to the analytic process, which is required by some analysis techniques. Next, data is sent to a central messaging system with multiple systems attached to it as follows. These algorithms are described in section 3.1.

2.3 Central System

The central messaging system contains all historical data for both short and long term storage. The access to historical data is a key feature of the central analysis, which allows one to (re)evaluate historical measurements, which is especially useful for newly-developed algorithms or configuration updates that can be tested and evaluated on a substantial corpus of real data. And, for machine learning applications, historical data can be used to train the necessary machine learning models.

For scalability, multiple intermediate aggregation systems can be deployed, which have the same architecture as the field-site messaging systems. We archive the data in both short and long term databases. The short-term database is intended to enable rapid searches and complete queries, and stores the (compact) analysis reports from the various subsystems and a set raw data for easy visualization. The long-term database stores raw data loss-lessly encoded using Google's Protocol Buffers [Var08] format, and compressed, using Cassandra's [Cas18] built-in compression, to 91% of its original raw data size. This database includes descriptive metadata added to describe the raw data such as the unique identifier

of the data source (e.g., sensor or network security monitor) and timestamps.

From this central messaging system we also allow export of data to third party systems and software. We have deployed an OSISoft PI Server historian, a software commonly used in both power utility companies and sensor control environments in general, and have also tested the export of data to the Splunk SIEM, an industry standard for operational security teams for storing and analyzing security data. The system supports the use of Elasticsearch as a SIEM by default, because we use that as one of our native data storage systems already.

Finally, we have deployed a graphical front-end for public use at `https://powerdata.lbl.gov` that allows visualization a web browser and also API access to both real-time data feeds and also historical data.

3 Cyber-Physical Intrusion Detection Algorithms

The proposed architecture in section 2 is designed to receive physical data from μPMUs and cyber packets mirrored by the packet sniffers from the SCADA to analyze and cross-correlate the data based on a set of defined rules and decide whether SCADA data has been compromised. The objective of these rules can classified as having three distinct approaches:

- The first of these approaches is the development of a suite of algorithms that, once an attack has been initiated, rapidly detect an anomaly that has put the grid in the transient regime, i.e., a a state whereby the grid cannot be characterized by a set of memory-less algebraic equations or has placed it in an unallowed zone of operation. These algorithms begin with localized analytics at the sensor level that flag potential anomaly signatures. These data packets are then prioritized for communication to a centralized layer where analysis is done across devices to corroborate behavior and understand the attack. The aim of these algorithms is to filter out the deviation from what is considered as "*normal*" in the power distribution grid. The cause of an anomaly can be a cyber attack, or it can be a failure or fault. In either case, we desire to spot this anomaly and further verify the root cause by looking at the monitored SCADA cyber traffic. The detailed description of these algorithms will be covered in section 3.1.

- The second of these approaches is to identify signatures in the data which may indicate that an intruder is performing a set of "reconnaissance" activities by testing its controllability and reachability of the network. These types of attacks would be structured such that the actions of the attacker would be verifiable in SCADA data by the intruder but would not manifest themselves as transient regime in the data or will not put the grid in an unallowed operating zone for the operator. An example of such an attack could be altering the time delay on a tap changing transformer as long as the

voltage is within the allowed limit. Another example is altering the Normally Open (N.O.) switch between two transformers in a substation. The details of the associated algorithms are given in section 3.2.

- In the third approach, we investigate the localization of faults using μPMU data, in an operating regime where the number of μPMU sensors available to carry out the inference is insufficient to have observability. We then show its application in the forensic analysis of cyber-physical attacks to a distribution Fault Location, Isolation, and Service Restoration (FLISR) system.

3.1 Anomaly Detection

The suite of algorithms we have developed for the anomaly detection layers have the following advantages in comparison to the present state of the art: (1) due to the near-real-time analysis, analytic results can be used to prioritize the traffic flow from the lower to higher layer, pushing forward reports of anomalies faster than data that do not raise a flag and need to simply be accrued for historical purposes; (2) it employs three-phase distribution grid equations rather than the more commonly-used positive sequence solution, thus avoiding the errors arising from this assumption; (3) a quasi steady-state condition is considered as the *normal regime* of operation rather than the idealistic steady-state which assumes there is no frequency drift.

3.1.1 Voltage Magnitude Changes:

During quasi steady-state operations, the magnitude of the voltage varies within a small range as outlined in power quality standards [IEE09]. Therefore, any large deviation from that range indicates an abnormal condition. Table. 1 lists the anomalies that can be observed in the voltage magnitude labeled by their severity and duration, denoted by |v| and τ, respectively and T_o denotes the period of a cycle.

Table 1: Voltage Magnitude Anomalies

anomaly	signature [2]
voltage sag	$0.1 \leq \|v\| \leq 0.9,\ T_0/2 \leq \tau \leq 60s$
voltage swell	$1.1 \leq \|v\| \leq 1.8,\ T_0/2 \leq \tau \leq 60s$
interruption	$\|v\| < 0.1,\ T_0/2 \leq \tau \leq 60s$
sustained interruption	$\|v\| < 0.1,\ \tau > 60s$
undervoltage	$0.1 \leq \|v\| \leq 0.9,\ \tau > 60s$
overvoltage	$1.1 \leq \|v\| \leq 1.8,\ \tau > 60s$

3.1.2 Current Magnitude, Active, and Reactive Power Changes

In addition to the voltage magnitude, suitable quantities to monitor for changes than can be computed locally from the μPMU voltage and current phasor measurements are current magnitude, active and reactive powers. Even when the voltage magnitude is within the safe range discussed previously, changes in current and/or active and reactive power can still happen due to the change of the load, affecting the magnitude of the current and/or the phase angle between current and voltage phasors. Therefore, it is also of interest to track fast-changes of these time-varying quantities. We are using the two-sided Cumulative Sum (CUSUSM) in these algorithms when tracking a fast change is required. The details of this method can be found in [Pag54, Lai95]. We label the anomalies related to fast changes in this class of data with *surge*, *drop*, and *oscillation* for increasing, decreasing, and swinging trends respectively by estimating the slope of the signal around the time of change.

While we are primarily concerned with fast state changing events in both the dynamic and transient realms, we must also consider events in the steady state time frame, slower changing yet also potentially critical. An example of this could include a line rating or transformer load being slowly, but consistently, pushed beyond its maximum rating, leading to accelerated aging and eventful failure. During the quasi steady-state, the three phase current phasor magnitude flowing in each line should be less than or equal the line rated current. We impose this constraint as feeder limit, and the violation is flagged as *overcurrent*.

3.1.3 Instantaneous Frequency Changes

For a μPMU at bus i, we estimate adaptively the instantaneous local frequency deviation from the nominal frequency during the quasi-steady state, using the approach outlined in [XDM12,XM12] that is tailored to three-phase distribution lines, to isolate abnormal changes in the estimated frequency.

3.1.4 Quasi Steady-State Regime Validity

Fig. 2 shows the abstract one-line diagram of a circuit from one of our partner utilities and the location of the installed μPMUs. The two feeders here are connected through the subtransmission grid.

[2]The voltage magnitude is in p.u.

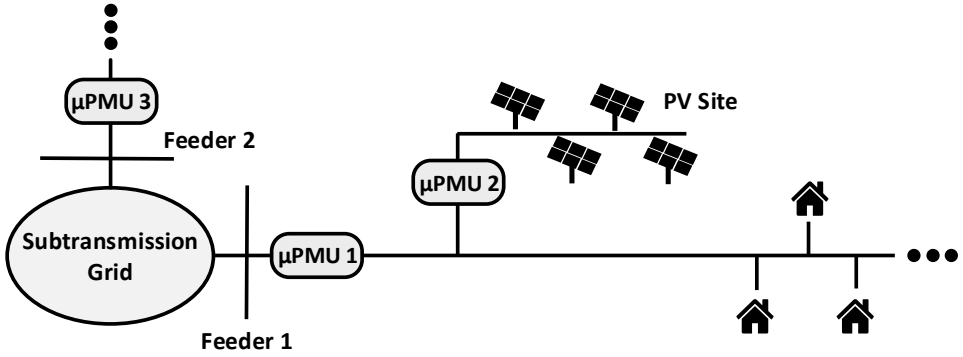


Figure 2: Location of Installed μPMUs in Our Partner Utility Grid.

Fig. 3 shows the voltage magnitude change rule inspected on the data from these μPMUs over a period of interest. The results of the fast change inspection on the current magnitude of phase a and instantaneous local frequency at the substation bus of feeder 2 are illustrated in Fig. 4 (a) and Fig. 4 (b), respectively. Observing the results of analysis, the Distribution System Operator (DSO) can deduce that the cause of the event is most probably located on feeder 2. Also, from the pre- and post-anomaly value of the current magnitude, the DSO can conclude that some of the loads on this feeder tripped due to the voltage sag. We just showed the results from some of the rules but other rules can also flag the existence of anomaly on feeder 2. Now, we wish to demonstrate also some results obtained from simulating an IEEE standard test case. We simulated the IEEE-34 bus test case [IEE] using then time-domain simulation environment of DIgSILENT [MP09]. The sampling rate is selected to be equal to that of Analog-to-Digital Converter (ADC) in a real μPMU. We then passed the time-domain data to our phasor estimation algorithm that emulates a two-cycle, P-class algorithm based on the IEEE C37.118.1 [IEE11]. The single-line diagram of the test case is shown in Fig. 5. This case includes single-phase laterals, voltage regulators, and untransposed lines, which all are modeled exactly in our admittance matrix. In this simulation, we tested our rules for detecting an anomaly with respect to a Single-Line to Ground Fault (SLGF), which is a very common type of short-circuit fault in the distribution grid. We introduce a SLGF on "Phase a" of line (25,26), which then blows the fuse placed on the phase a of this line near bus 25. Our three μPMUs are placed on buses, 9, 19, and 31 based on the optimal placement criterion.

The results of the *"voltage magnitude change"* rule is shown in Fig. 6 for μPMUs 9, and 19. The rule inspects the data for large deviations, labels them accordingly, and finds the start time and the end time of the event, marked with blue and red stars.

Fig. 7 illustrates the metric value for the specified lines that is inspected with $M = 12$ to check whether the grid is in the quasi steady-state or not.[3] The start time of the detected changes are also marked, setting the CUSUM detector parameters fixed in all the three local

[3]The voltage and current data are first converted to per-unit system assuming $S_b = 1$ MVA.

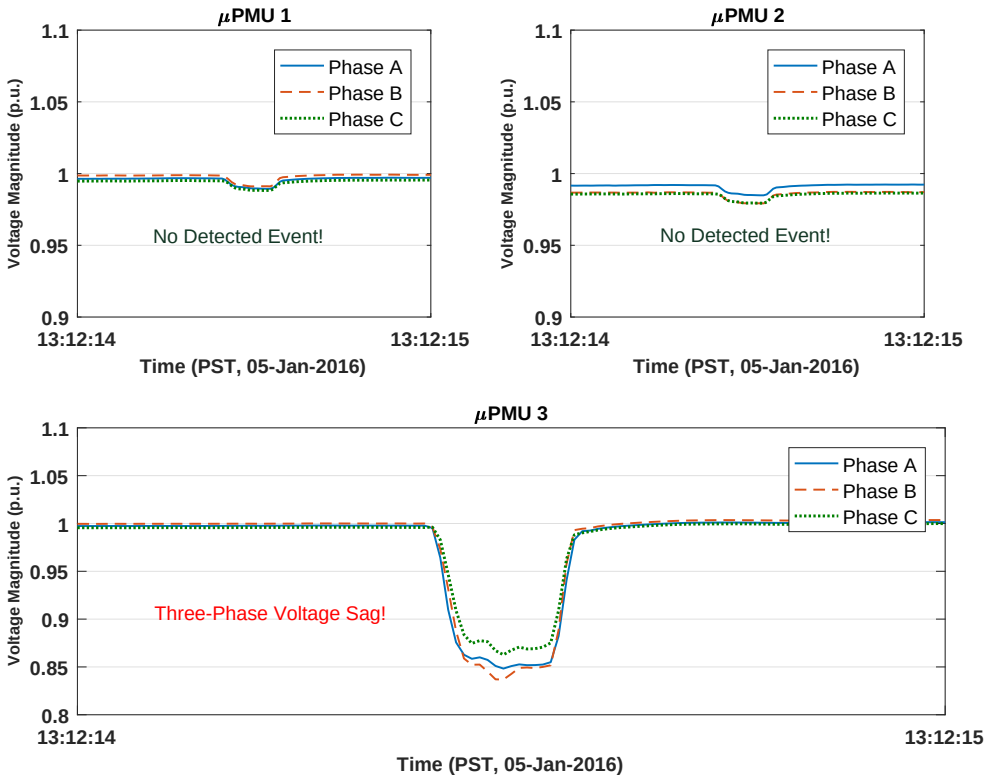


Figure 3: Voltage Magnitude Change Rule for Real Data.

engines corresponding to each μPMU. As can be observed, there are two periods in which the grid manifests its dynamic: the first one corresponds to the occurrence of the fault and the second matches with the fuse meltdown. In addition, based on the severity of the transient that each μPMU measures, the number of detected changes via CUSUM varies. In this case, considering the location of the μPMUs and the location and type of the fault, the most severe change appears in the metric corresponding to measurements from line (19,20), while the changes in metric for line (31,32) is very small. Therefore, based on the defined parameters on the detector, CUSUM finds quite a number of change points in the former, while it is not set to be sensitive to the changes in the order that appears in the latter. In fact, if the detector is set to be too sensitive, it can increase "*false alarms*" in the system. Also note that due to the two-cycle calculation of the phasor, and use of M samples to calculate the correlation matrix, the event appears and disappears with a systematic delay. The other local rules also capture the anomaly, though with different severity and behavior. In fact, many of the local rules may detect the same event, though some rules are more informative than others depending on the cause. Each triggered rule reports a start and an end time for the event. Storing theses time-tags for eventful segments of data helps understanding their relationship.

Although the corresponding rule at the local level is formulated irrespective to the grid

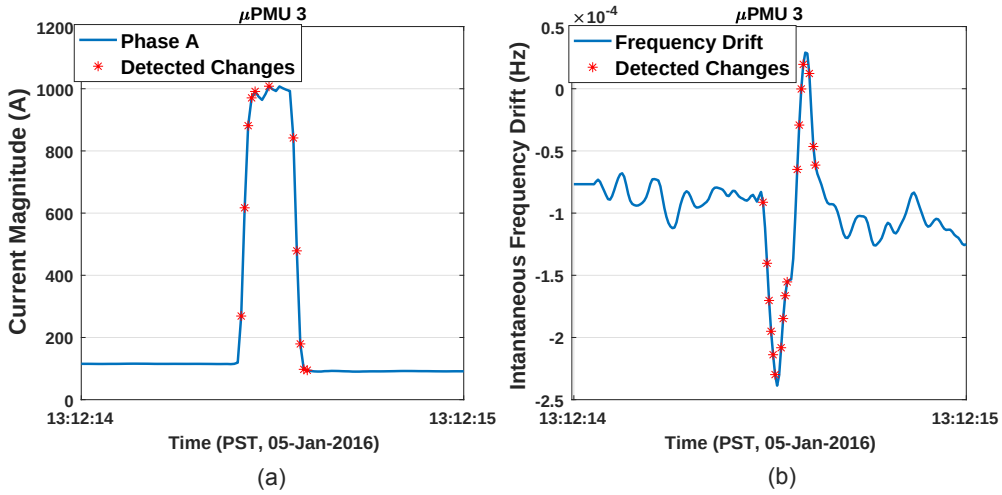


Figure 4: Fast Change Tracking of Current Phasor Magnitude and Bus Instantaneous Frequency Drift for Real Data.

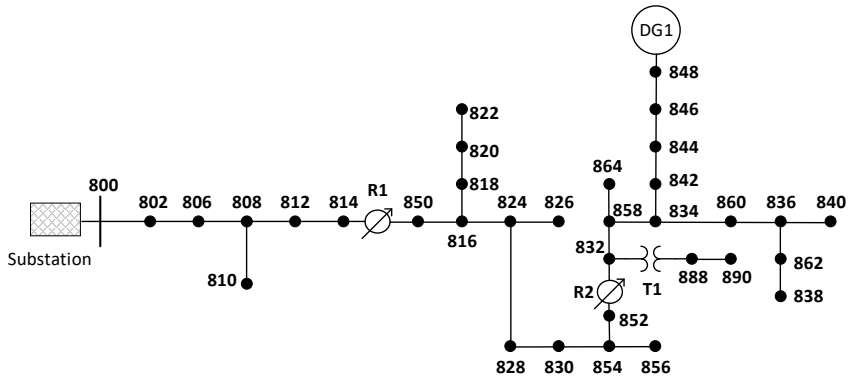


Figure 5: IEEE 34-Bus Test Feeder Single-Line Diagram

parameters and the geographical location of the sensor for scalability purposes, the rules at higher stages are set up incorporating the grid parameters. In particular, we use the *Kirchhoff's Voltage and Current Law* as the cornerstone of our analysis to establish our rules in these stages. As we have previously shown [JSR+17b], the equation relating the vector of three-phase voltage phasors to the vector of three-phase current injection phasors over the grid through the admittance matrix only takes form of a memory-less algebraic system of equations when the grid is in the steady-state, which can also be used for quasi steady-state condition with a very good approximation. However, when a severe-enough transient is induced by an event, these equations are no longer memory-less, and are in the form of *"differential algebraic equations"* instead. We use this interpretation as the basis for our rule to detect anomalies, which means that at higher levels of aggregations, deviation from the quasi steady-state regime (memory-less algebraic equation) is a signature of an anomaly for the perimeter that is monitored. If we have a large number of μPMUs

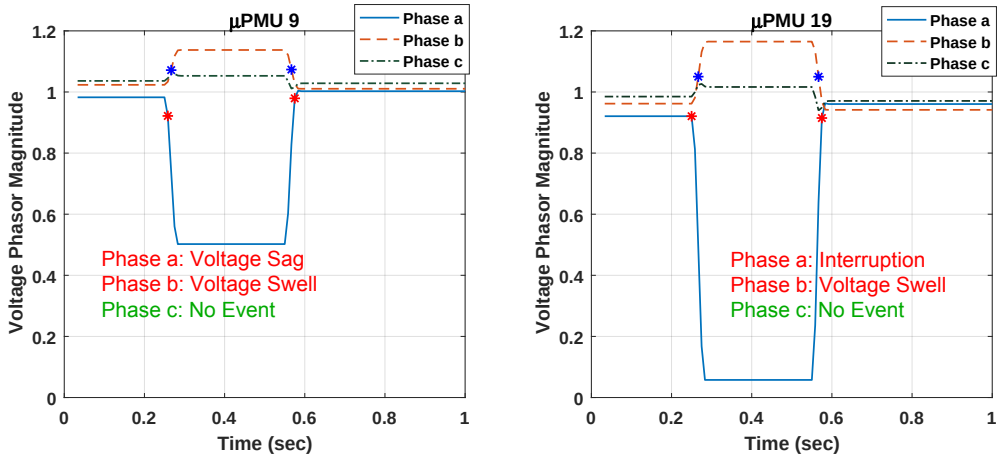


Figure 6: Voltage Magnitude Local Rule Result for SLGF.

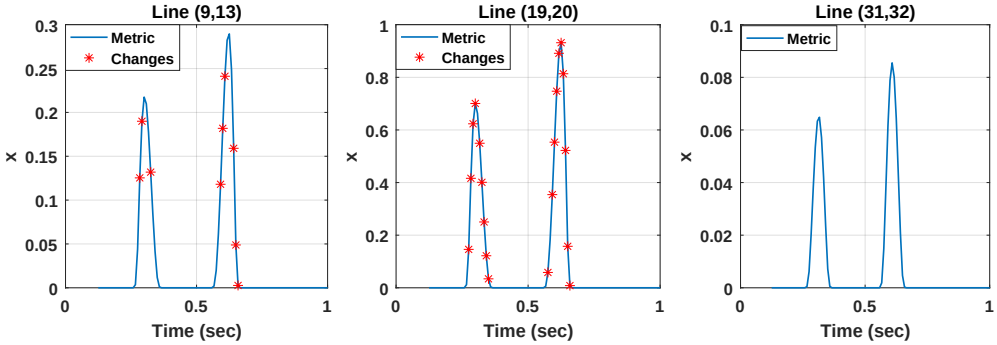


Figure 7: Quasi Steady-State Validity Checking for SLGF.

(more than half the number of nodes), we can formulate a metric that is zero in the quasi steady-state and is non-zero otherwise so we can simply spot the transient. However, as we mentioned earlier, the challenge here is the limited number of μPMUs. We can still use a similar metric, that approximates zero and varies smoothly during the quasi steady-state and non-zero otherwise with fast changes otherwise. The schematic diagram of this rule is illustrated in Fig. 8. The metric defined for our central engine is also illustrated in Fig. 9 for the SLGF scenario, which indicate the successful performance of the metric in appearing the anomaly and the detector to track these changes. The delay in appearing and disappearing of the event in here is solely due to the two-cycle phasor calculation.

3.1.5 Discussion–Compromised Data

In this experiment we illustrate the resilience of our anomaly detection architecture to data injection attacks. For this purpose, we investigate three data attacks scenarios happening concurrently with the SLGF event discussed previously. We consider the case of the attacker

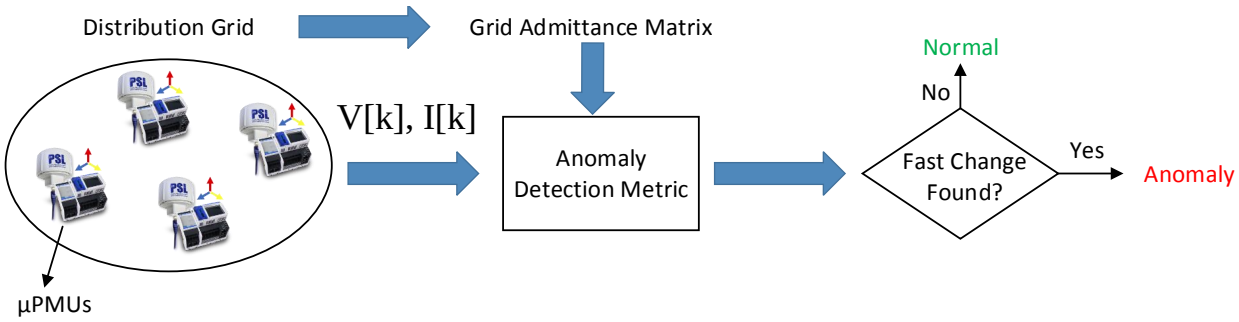


Figure 8: The Schematic Diagram of the Anomaly Detection Rule at Higher Levels of Data Aggregation.

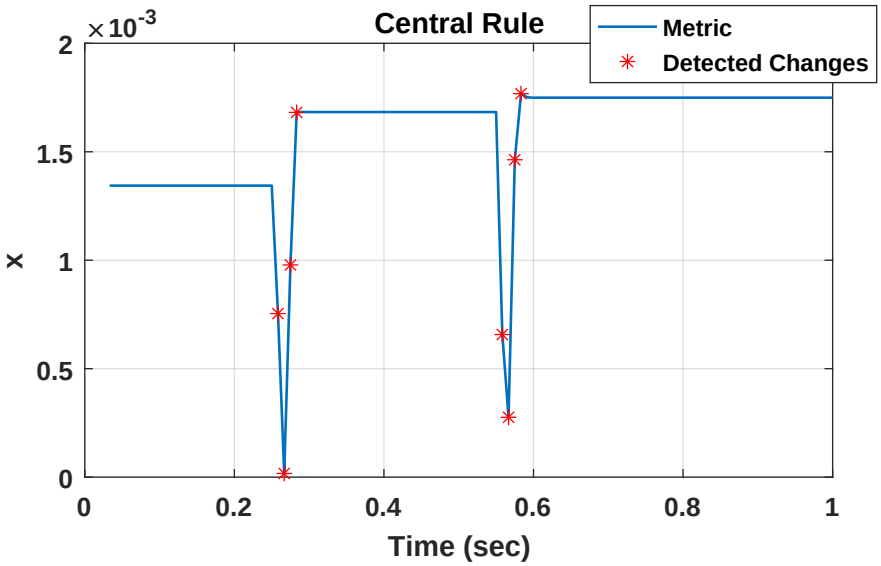


Figure 9: Central Rule Inspection for SLGF.

manipulating the data of μPMU 9 in the first case, and μPMU 19 in the second case, and finally μPMU 9 and 19 at the same time in the third case. In all cases the data injected are a *replay* of the last available data set before the anomaly starts. Setting the change detector parameters fixed for all the three cases, Fig. 10 shows, for each case, the central rule and the start time of the detected changes. As can be seen, since μPMU 19 is playing an important role for this event, having μPMU 9 compromised will not affect our central rule significantly (case-1). However, when the μPMU 19 is compromised, the number of detected changes reduce significantly (case-2), and when both μPMU 9 and 19 are compromised and the only healthy data is coming from μPMU 31, the detector does not pick any fast changes based on the set parameters. This also reveals the importance of tuning the detector thresholds to have a certain "false alarm," while maximizing the "detection" probability. We note that the local analytics that directly draw data from μPMU 19 will still flag the alarm, so buffering locally at the site of the event these data can be an important way of helping understand

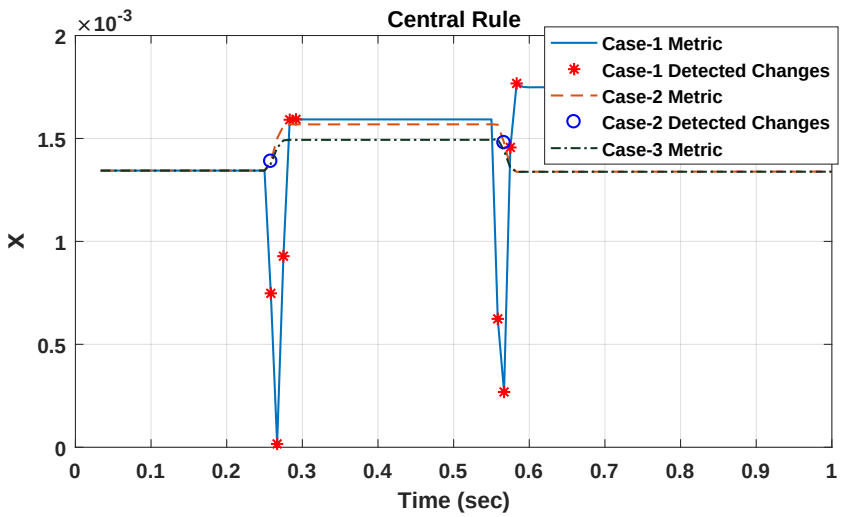


Figure 10: Central Rule for SLGF with Manipulated μPMU Data.

what communications were compromised in an ex-post analysis.

3.1.6 Optimal μPMU Placement for Event Detection

We define our placement criteria with respect to our above-mentioned rule at higher levels of data aggregation. Our criterion in this case is the minimum number of sensors as possible during the quasi steady-state. This in turn means that the sensitivity of the rule with respect to the anomalies increases. Therefore, we formulate our optimization as a min-max problem aiming to minimize our metric for the maximum value that it can take over all the possible available measurement set. The advantage of the inner maximization is that we are no longer dependent on the measurements, and we are only dependent on the grid parameters.

To test our placement criterion, we first started with IEEE-34 bus standard test case. To find the optimum set in here, we used an "*exhaustive search*" algorithm that evaluates the objective function for all the possible combinations. Since this search is exponentially complex, it therefore becomes intractable for larger cases with higher numbers of μPMUs. Thus, we employ a *"greedy search"* algorithm as an alternative to reduce the time complexity to *polylog*, but accept the results to be sub-optimal.

Table. 2 compares the objective value for a random placement, *"Greedy Search"* and *"Exhaustive Search"*, and the time complexity of each solver, assuming that $K = 3$ μPMUs are available. The objective value of the *"Greedy Search"* and the *"Exhaustive Search,"* and the set of the selected buses are very similar, while the runtime of the *"Greedy Search"* is 102.206 times faster. Overall, it indicates that the *"Greedy Search"* can be a very good choice to solve our optimal placement problem.

Table 2: IEEE-34 Case Optimal μPMU Placement Result for $K = 3$

	Random	**Greedy**	**Exhaustive**
Optimum Cost	2.4035	0.51477	0.51477
Buses with μPMUs	{1,7,23}	{7,19,31}	{9,19,31}
Run Time	–	2.84 s	290.266 s

The result of the optimal μPMU placement for the LBNL utility grid is shown in Fig. 11 assuming that 4 μPMUs are available. The placement algorithm places the μPMUs on both

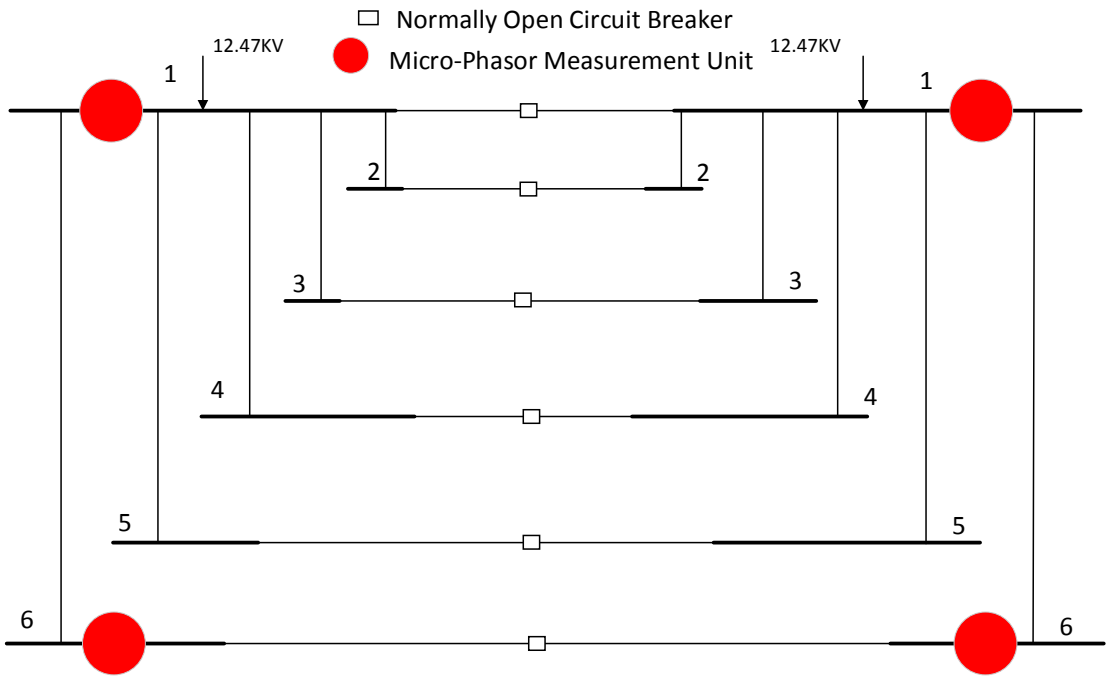


Figure 11: Optimally-Placed μPMUs in the LBNL Utility Grid.

feeders and on similar locations because they are approximately symmetric. Besides, it puts μPMUs at substation and at the end of the longest line (1,6) to give the maximum coverage.

While the formulation of the placement is general, we can put the constraints that μPMUs should be only placed at the three phase level not laterals since the events happening on single-phase and two-phase lines can be captured by the μPMUs nearby at the three-phase level or we can use the reduced model, where the laterals are all rolled up.

3.2 Reconnaissance Activity Identification

As a motivation for the consideration of "reconnaissance attacks," we consider the Ukrainian power grid attack report [LAC16]. This report suggests that the attackers may have been passively monitoring and potentially testing their controllability over the network for up to 6 months before initiating the main attack. This suggests the need to identify the points in the grid that an attacker could test its controllability over switches in the network without causing any damage to the grid or raising any alarms. Along with the identification of such potential "reconnaissance attacks," we have developed the tools to be able to detect such activities in the grid before the adversaries move to the next stage, a potential large scale attack. In particular, focusing on the security monitoring of the On Load Tap Changers, auto-switched capacitor banks and N.O. substation switches as potential points that can potentially be subject to a 'reconnaissance' activity, we briefly go over the developed tools and accordingly describe how they can help us detecting such activities.

3.2.1 Learning Control Logic of Discrete Control Devices

Our work presented in [RSJ+19] gives an overview of the possible control logic schemes that may be operational for regulators and controllers as well as outlines a suite of algorithms for parameterizing these control schemes. To summarize how a regulator typically operates.

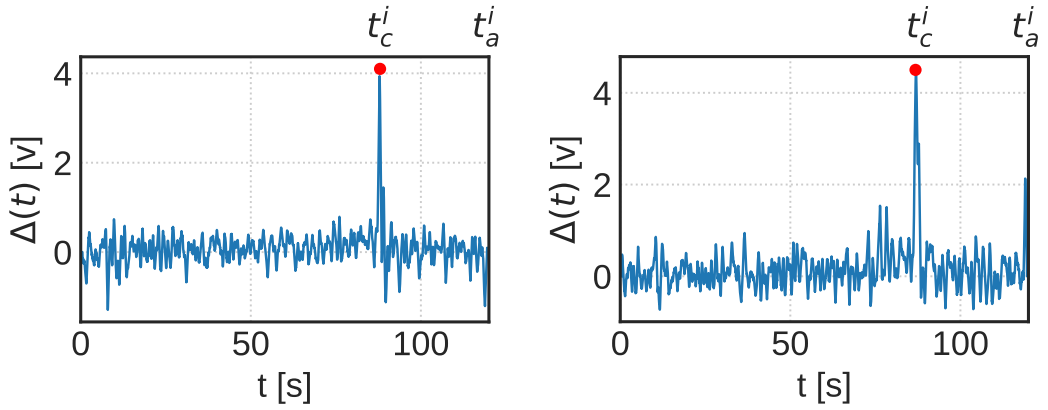
- It has an upper- and lower-threshold boundary outside of which the regulator will seek to perform a tap operation to bring the voltage back within allowable bounds.
- These thresholds can be static, i.e. not related to how much active and reactive power is being consumed, or they can be a function of the current power demand. The latter is called Line Drop Compensation (LDC) where the regulator attempts to estimate the voltage drop along the line to the main customer load center, and account for this offset. Intuitively this means that if the regulator expects a higher voltage drop along the line i will keep its local voltage higher to allow for the drop.
- Once the regulator has sensed that the voltage has exited its allowable range it begins a countdown timer. If the voltage re-enters its allowable range before this countdown timer expires then the timer is reset. Otherwise, once the timer has reached zero, the regulator executes a control action, i.e. it changes its tap position to raise or lower the voltage. This countdown timer can be static (i.e. 30 seconds for all tap change events) or it can be inversely proportional to how far outside its allowable range the voltage has deviated. The latter is called an inverse time delay. A simple example is that if the voltage is 1.25 V above its target it will execute a control action is 20 seconds. however if it is 5 volts above its target, it will execute a control action in 4 seconds.

Within this work we seek to estimate both the upper- and lower-thresholds and associated time delay. Then, our algorithms can run online and passively monitor the operational behavior of the device and alert the operator should it detect behavior inconsistent with its historical behavior. The voltage generally exits its allowable range in one of two manners: 1) either via a gradual ramp-up or ramp-down in the net-load of the feeder or 2)a discrete jump in the voltage caused by either a large change in net-load load or an event on the transmission/subtransmission grid, e.g. a switching action or the actuation of subtransmission regulation equipment. It is the former which we seek to exploit in order to estimate a regulators time delay. We will examine the difference between the individual PMU voltage measurements using an approach similar to early-late gate from signal processing. This approach take a sample at time t and computes the difference between the mean of the next n samples and the mean of the previous n samples. In this application we take n to be 01 second in order to filter out high frequency voltage fluctuations due to motor starts and other short term voltage behavior. We denote the time instant immediately before the actuation of a switching event i as t_a^i and the time which the voltage time series exited its allowable range as t_c^i. Therefore, if the control device has a fixed time delay, this time delay is given by $t_a^i - t_c^i$. If a device has a dynamic time delay, as a function of the deviation of the voltage from the target voltage, we adapt a iterative approach to converge on this user set delay and control scheme parameters. This methodology is detailed in [RSJ+19].

In order to validate our approach using real data we consider two OLTC transformers at pilot site deployments whose secondary terminals are being measured my a distribution PMU reporting at 120 Hz. Both locations are independent distribution feeders electrically connected through a sub-transmission network. In order to identify operation of the OLTC, both sensors were used to classify events as either global, i.e. originating from the transmission or subtransmission network, or local, i.e. OLTC transformers or downstream switched capacitor banks. Following this, $k-$means clustering was employed to group events and distinguish between OLTC transformer actions and switched capacitor banks [ARAS17]. Once a set of tap operations has been identified the control logic is identified and parameterized. In both cases, the summation of the one second time series difference profiles exhibited one distinct peak, as can be seen in 12, and correspondingly a value for the regulator time was obtained and can be seen in Table 3.

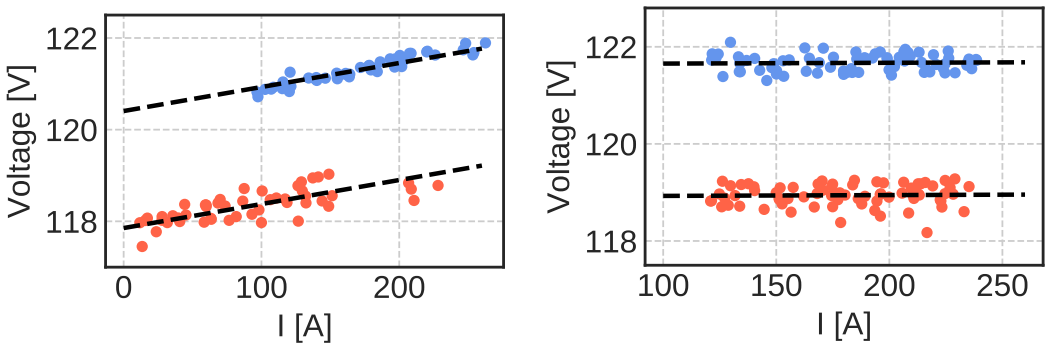
Table 3: COMPARING PARAMETER ESTIMATIONS FOR UTILITY MEASURED DATA

		Target Voltage	Controller Deadband	Time Delay
OLTC 1	**Estimate**	119.1 V	2.6 V	32.3 s
	Reported	119 V	3 V	30 s
OLTC 2	**Estimate**	120.13 V	2.7 V	33s
	Reported	120 V	3 V	30 s



(a) OLTC 1 (b) OLTC 2

Figure 12: Estimating the regualotr time delay via sum of one second first discrete difference



(a) OLTC 1 (b) OLTC 2

Figure 13: Estimated lower- and upper-thresholds

These estimation values along with those as reported by the utility are shown in Table 3. All estimated quantities for the optimal case were within the specified accuracy of the controller with any minor discrepancies attributable to systematic error of their individual instrumental transformers and controller measurement inaccuracy. In terms of monitoring a devices behavior for cyber-physical security purposes, it is less critical that estimated parameters match the reported parameters but rather that the observed behavior is consistent with the learned control logic. i.e. we are focused on detecting abnormalities in operational behavior rather than minor discrepancies between estimated and reported values.

Specific considerations for feeders with Distributed Energy Resources (DERs): An OLTC and/ or switched capacitor banks offers an attacker with a unique avenue through which to disrupt the network operations. OLTC transformers are typically capable of regulating the voltage at their secondary terminals with ± 10% of the voltage at their primary

terminals. As per the revised IEEE voltage ride-through requirements [15418], smart inverters are required to trip offline after a minimum time period for $V > 110\%$ after the onset of an event. Dependent on the voltage at the primary side of the transformer and the instantaneous tap position, an attacker may be capable of utilizing an OLTC and/or capacitor banks to push the voltage within these ranges as outlined in the revised IEEEE 1547 standrd for DER [15418]. If a coordinated attack was carried out across multiple substations, this may threaten the stability of the system. Therefore it is critical to monitor these assets and rapidly detect erroneous operation.

3.2.2 Substation Normally-Open Switch

In this use case we assume that the substation connected to the distribution feeder is equipped with a normally-open spare transformer parallel to the main transformer, to restore the delivery of the power to the feeder when the main transformer fails or requires maintenance. Inserting the spare transformer into service, while the main transformer is also in service, has no adverse effect on the delivery of the power to the feeder downstream. It is therefore appealing for an attacker to try and test its ability to control the network through a compromised SCADA network, by changing the service status of the backup transformer switch when the main transformer is in service. The attacker may then confirm the switching in of this transformer via intercepting SCADA traffic and mask this control action by spoofing the SCADA packets sent to the control center. The question we seek to answer is:"can we detect such an activity with μPMU data?"

Let us assume that we have a μPMU placed at the head of distribution feeder as shown in Fig. 14. In this circuit, the dominant term in the Thevenin source impedance would be due to transformer impedance. Therefore, when the spare transformer switch is toggled maliciously, while the main transformer is in service, the source impedance will roughly decrease by 50% relative to its normal value. We track the fast changes in the estimated positive sequence of the Thevenin source impedance. When a fast change is found, the magnitude of the change can be extracted to determine whether it can be attributed to the spare transformer switch or not. Fig. 15 shows the online tracking of the Thevenin impedance. As it can be observed, when the switch gets closed, the estimated impedance is almost half of what it was before. The dominant term in the source impedance is due to the transformer and such a change cannot be the consequence of change in the transmission grid topology, the most plausible cause is the event that the spare transformer switch was closed. Once such a change has been detected the control center would be notified to determine where this was a scheduled or unexpected switching operation, potentially a signature of malicious behavior.

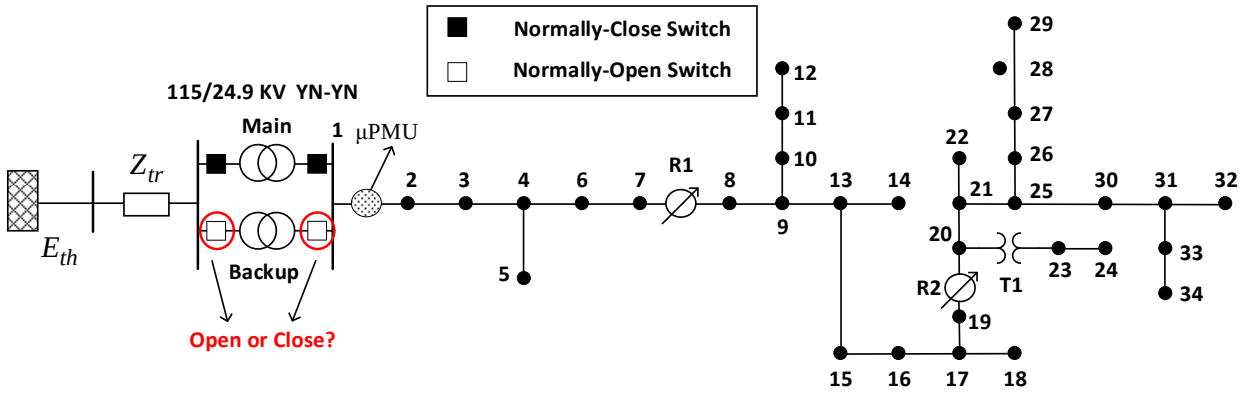


Figure 14: Modified IEEE-34 Test Case for Reconnaissance Attack Identification

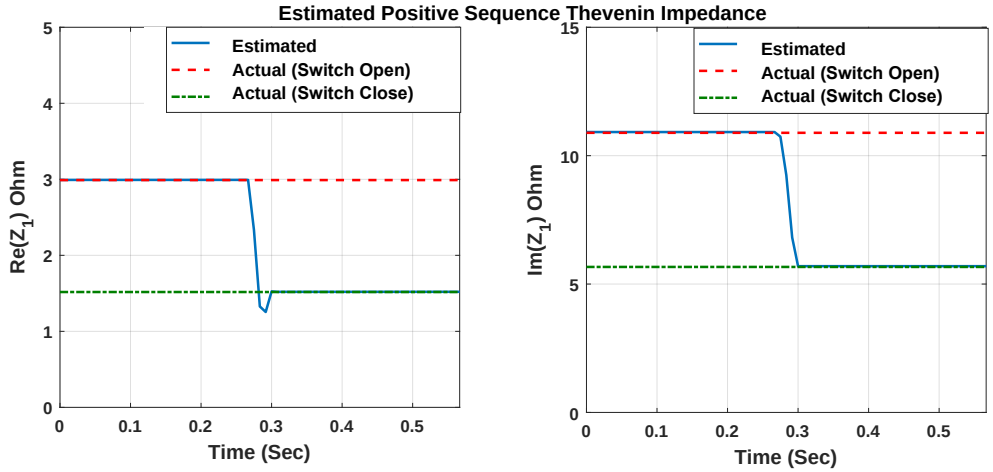


Figure 15: Online Tracking of the Thevenin Source Impedance using μPMU Measurements

3.3 Low-Resolution Fault Identification

Once an anomaly has been detected, the next step is determining the source of this anomaly in order to prioritize forensic analysis of the corresponding SCADA packets. Using a statistical approach, we first improve upon the state of the art and study the fault localization problem with measurements that are too scarce for observability (we refer to it as the *under-sampled grid* regime). Our main contribution is to show that in an *under-sampled grid* fault location can be resolved reliably at a level of *clusters* that typically form connected subgraphs. These clusters depend both on the properties of grid admittance matrix as well as the placement of the sensors. As a figure of merit, the *cluster-level fault localization* resolution can be measured in terms of the maximum size of all clusters in which the graph is divided. Leveraging the insights from our analysis we propose a placement strategy that achieves the highest localization resolution over the grid, by clustering the graph in a number of connected components with clusters' sizes as even as possible. In the paper, we also com-

pare our method with [Bra11] to better highlight the benefits of our statistical approach. We then apply the developed tool in the forensic analysis of cyber-physical attacks to a distribution Fault Location, Isolation, and Service Restoration (FLISR) system [Ulu12].

In this section, we first corroborate our technical discussion and then showcase its application in the forensic analysis of cyber attacks to FLISR systems.

3.3.1 Fault Localization Optimal Sensor Placement and Identified Communities

Before analyzing the results of the fault identification algorithm, we identify the optimal *greedy* placement for the PMUs based on our envisioned optimal sensor placement for fault localization. We highlight the clustering by showing the correlation metric among nodes to highlight the set of nodes that are highly correlated in the fault localization sense and form a cluster.

- IEEE-34 Bus Test Feeder: The one-line diagram of the test-case is shown in Fig. 16, where a 100 kW generator is added at bus 848. This grid is unbalanced and radial and has untransposed lines and one phase laterals. Also, without loss of generality, only the nodes at the same voltage level have been considered. The placement is

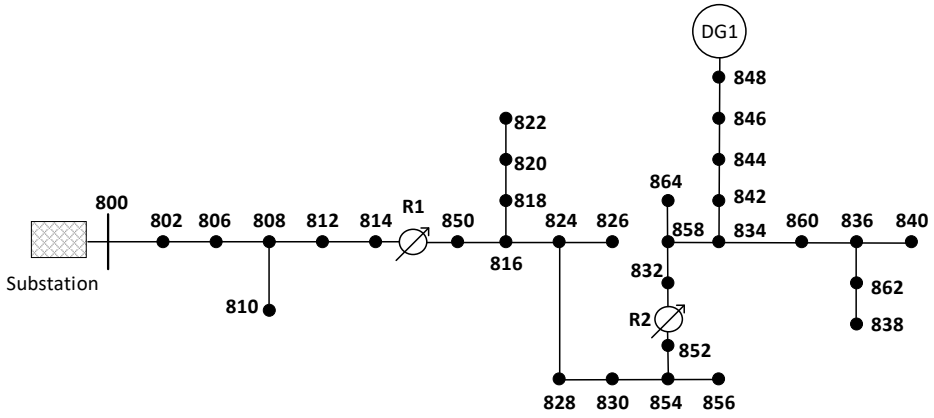


Figure 16: Reduced IEEE-34 Test Case with Added Generator.

done on the reduced network, where the laterals are rolled up to avoid putting a PMU on a single phase node, thus avoiding the case where we only use using 1/3 of its measurement channels. Assuming that there are 4 PMUs available i.e., $P = 4$, Table. 4 shows the optimal location of the sensors and the random placement used as a comparison. It is clear from the locations in Fig. 16 that are optimal according to Table. 4 that the sensors are placed to cover the grid. Fig. 17(a) shows the thresholded correlation coefficients corresponding to phase-A for the optimal placement in Table 4 and Fig. 17(b) is the same quantity that corresponds to the random choice. The clusters

Table 4: PMU Locations for IEEE-34 Bus

Test Case	#PMUs	Optimal sites	Random sites
IEEE-34	4	800-830-848-840	800-814-816-848

Table 5: Optimal Sensor Locations for IEEE-123 Bus

Test Case	#PMUs	Optimal Location
IEEE-123	10	149-81-61-56-105-250-86-151-72-57

of light gray values effectively represent fault locations that are hard to discriminate. The larger the cluster, the lower the resolution. Hence, Fig. 17(b) clearly illustrates the impact on the resolution of a bad sensor siting, when compared to the clustering that emerges in Fig. 17(a). To corroborate our conjecture that the placement is affected

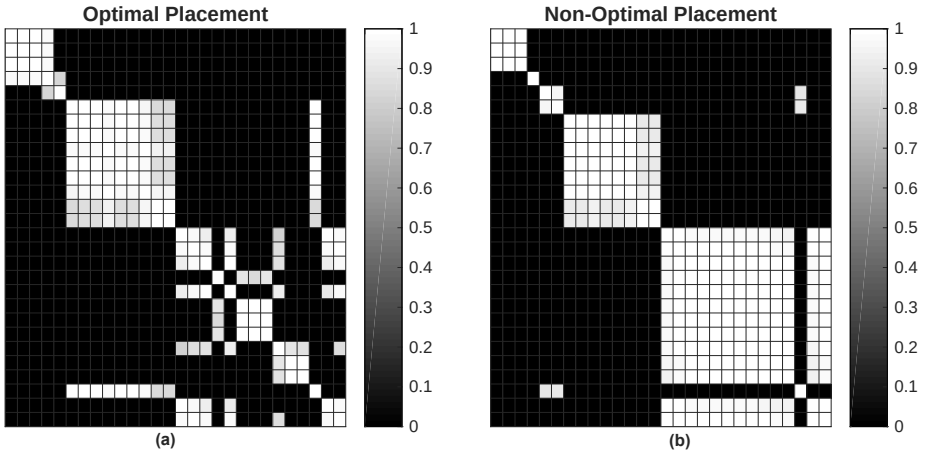


Figure 17: Thresholded Correlation of Nodes for IEEE-34 Bus Case Related to Phase-A for a) Optimal Location b) Non-Optimal Location

by the intrinsic topology of the graph and its clusters, in Fig. 18 we highlight the set of buses that exhibit high correlation based on the optimal placement, by building a graph adjacency matrix such that its entry is one if the corresponding correlation coefficient is greater than a threshold and otherwise is zero. As we predicted, the highly correlated nodes are those that are located in a neighborhood of each other.

- IEEE-123 Bus Test Feeder The analysis of the larger IEEE-123 test case, with an optimal assignment of 10 PMUs, returns the results in Table 5. Once again, the placement is done after rolling up the laterals and only nodes at the same voltage level are analyzed. The one-line diagram and node numbering of this case can be found in [Ker01]. The PMUs in this case also are spread over different areas of the

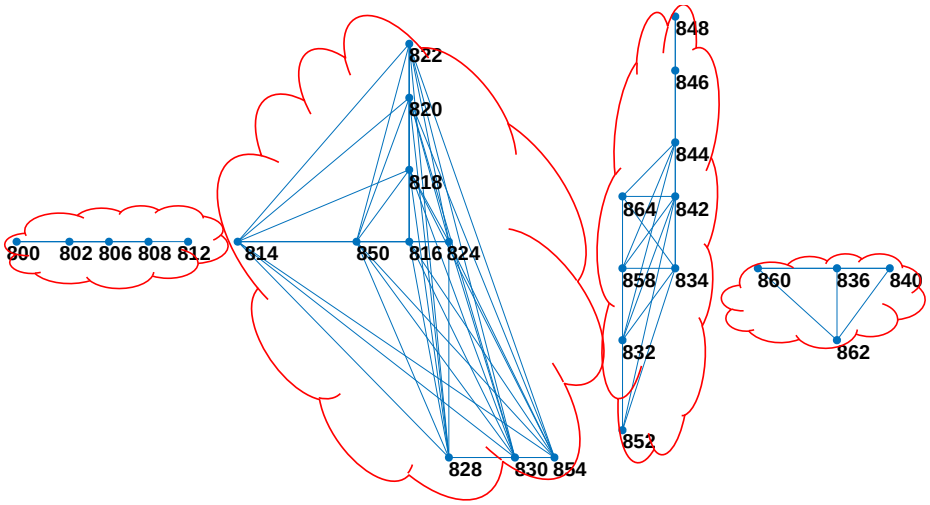


Figure 18: Adjacency Matrix Graph for Correlation Coefficients Corresponding to Phase-A.

grid topology to form communities of neighboring nodes of comparable sizes. The thresholded heat-map given in Fig. 19 based on the optimal placement of Table 5, highlights the balance across clusters.

3.3.2 Fault Localization Algorithm

The synthetic PMU data for this experiment were generated using the OpenDSS software [Dug12]. The pre-fault data were first recorded using a power-flow snapshot and then a fault was introduced in the dynamic mode to represent the behavior of the grid after a fault occurs. It should be noted that the tap changers usually have a delay for 15-30 seconds in order to respond to a change so the voltage and current data should be recorded before the tap value changes, so that the admittance matrix stays the same before and after the fault.

- IEEE-34 Bus Test Feeder: Fig. 20 shows the value of the proposed metric for different hypotheses when a single-phase to ground fault occurs at phase-A of bus 820. The higher the value of the metric corresponding to a node is, there is a higher possibility that the fault has happend there. As our analysis indicated, using the proposed metric shown in Fig. 18 it becomes readily apparent that one can narrow down the location of the fault up to a certain resolution, since a cluster of nodes return very close metric value. These are also the nodes that are highly correlated with respect to the actual location of the fault. To further test and verify the fault localization method, we introduce different types of fault and show the results of the fault localization in Table 6 under the optimal placement.
- IEEE-123 Bus Test Feeder: The analysis of the fault localization algorithm has been extended to IEEE-123 test case for different types of fault after the placement of

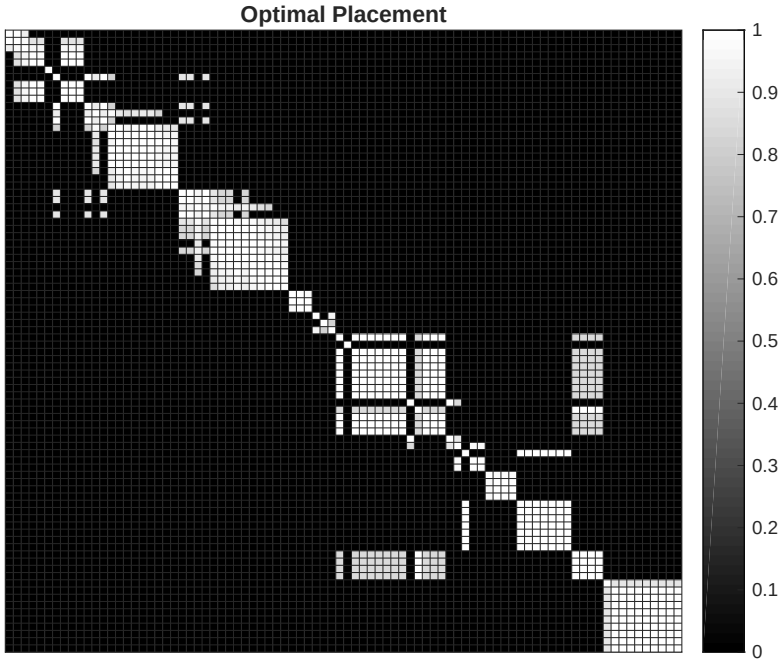


Figure 19: Thresholded Correlation of the Nodes for IEEE-123 Bus Case Related to Phase-A for Optimal Location.

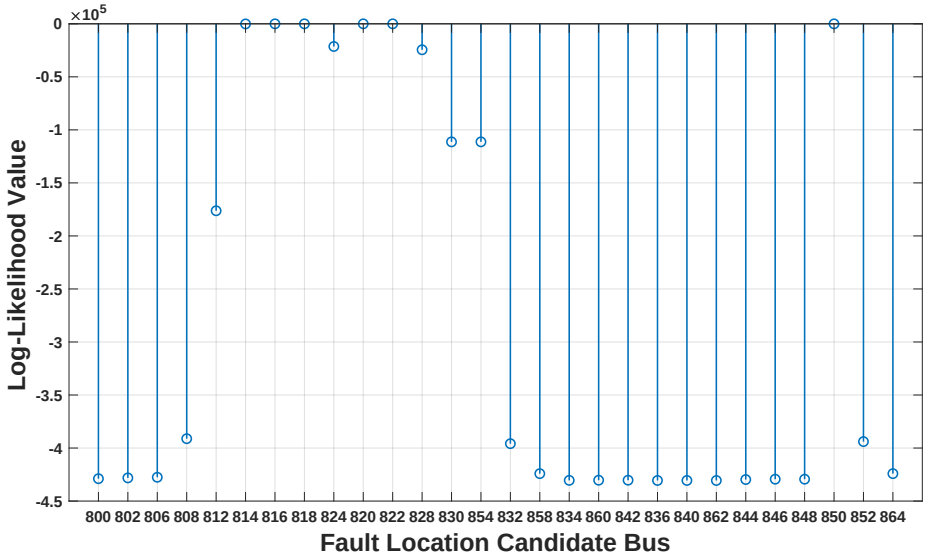


Figure 20: Metric Value for Fault at Bus 820-Phase-A

Table 6: Identified Fault Locations for IEEE-34 Bus Case

Fault Type	Exact Fault Location	Possible Fault Locations
LLL	816	814,816,850
A-G	822	814,816,818,820,822,850
BC-G	852	832,852,858
AC-G	836	836,840,862
AB-G	808	800,802,806,808

the sensors on the optimal locations in Table 5. Fig. 21(a) shows the metric value of different hypotheses for a three-phase fault at bus 160. Part of the correlation coefficients heat-map that illustrates the correlation of the columns that have high correlation with bus 160 is snipped out of Fig. 19 and is illustrated in Fig. 21(b). As expected, the neighboring nodes that also have high correlation with bus 160 are those, which have the closest metric value and can be mis-identified as the fault location.

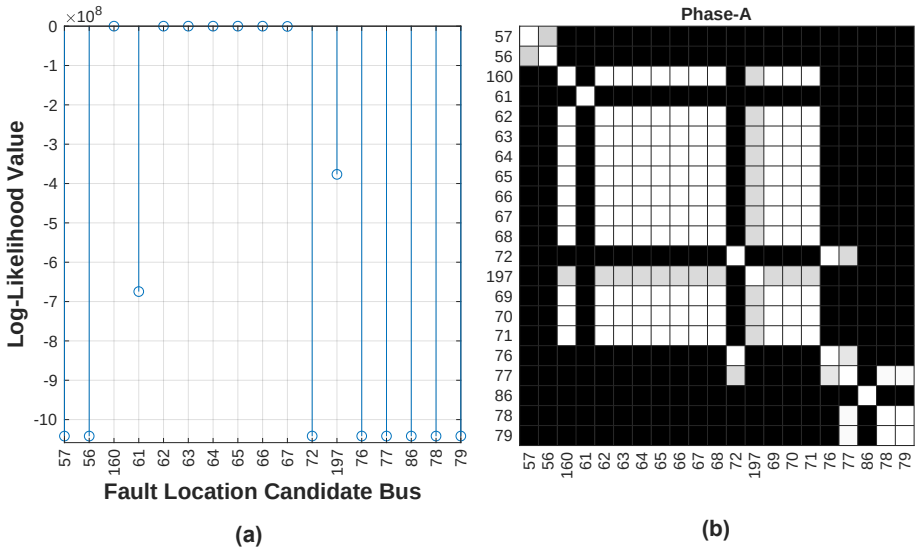


Figure 21: a) Metric Value for Three-Phase Fault at Bus 160 and b) Correlation of Nodes-phase A.

Other types of faults and the potential locations are summarized in Table 7.

The faults are introduced at different locations of the grid and the results show that the fault is identifiable up to the resolution of the clusters of neighboring nodes identified with the correlation coefficients. An interesting case is the two-phase fault at node 57 (last row in Table 7), where the fault is identifiable with a high resolution. This is consistent with our claims, since the correlation coefficients reveal that there is no other node with very high correlation with this specific node.

Table 7: Identified Fault Locations for IEEE-123 Bus Case

Fault Type	Exact Fault Location	Possible Fault Locations
LLL	42	40,42,44,47,48
A-G	108	105,108,109,300,110, 111,112,113,114
BC-G	89	86,87,89,91,93,95
AC-G	50	47,48,49,50,51,151
AB-G	57	57

- Comparison with the State of the Art: Our approach in using the pre and post-fault samples and treating the fault as a current injection improves upon the work in [Bra11] through its statistical underpinning. To show this, a two-phase fault is introduced at bus 836-A-C and the results are shown in Fig. 22. The original method in [Bra11] is designed as a minimization problem to find the fault location so we changed it to a maximization by adding a negative sign to make a better visual comparison with our method. The results clearly show that the proposed method can locate the fault more accurately than the algorithm in [Bra11].

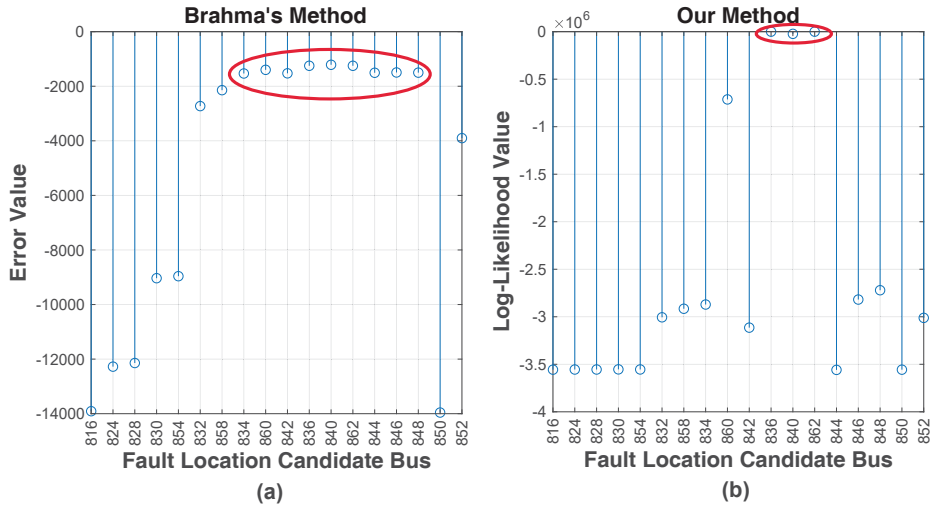


Figure 22: Fault Localization for Fault at Bus 836-A-C Using a) Method in [Bra11] b) Proposed Method.

Besides this improvement, investigating the reason behind this ambiguity is missing in [Bra11], whereas this matter has been investigated theoretically and analytically in our work. In addition, an optimal sensor placement strategy is proposed here to place the sensors at nodes that would return the highest fault localization resolution.

3.3.3 Application to Cyber-Physical Intrusion Detection

We leveraged the insights from this work to enhance our network intrusion detection system testbed hosted at LBNL [Gen17, §4] [PGB+17] and incorporate additional rules to monitor for cyber attacks that interfere with the normal operations of FLISR systems [Ulu12]. Our testbed (see Fig. 23) is defined as hierarchical architecture to fast-detect the presence of cyber-physical attacks by correlating analytic results from SCADA traffic (which includes traces from fault detector communications) with PMUs deployed sparsely over a distribution system [Gen17, §4] [PGB+17].

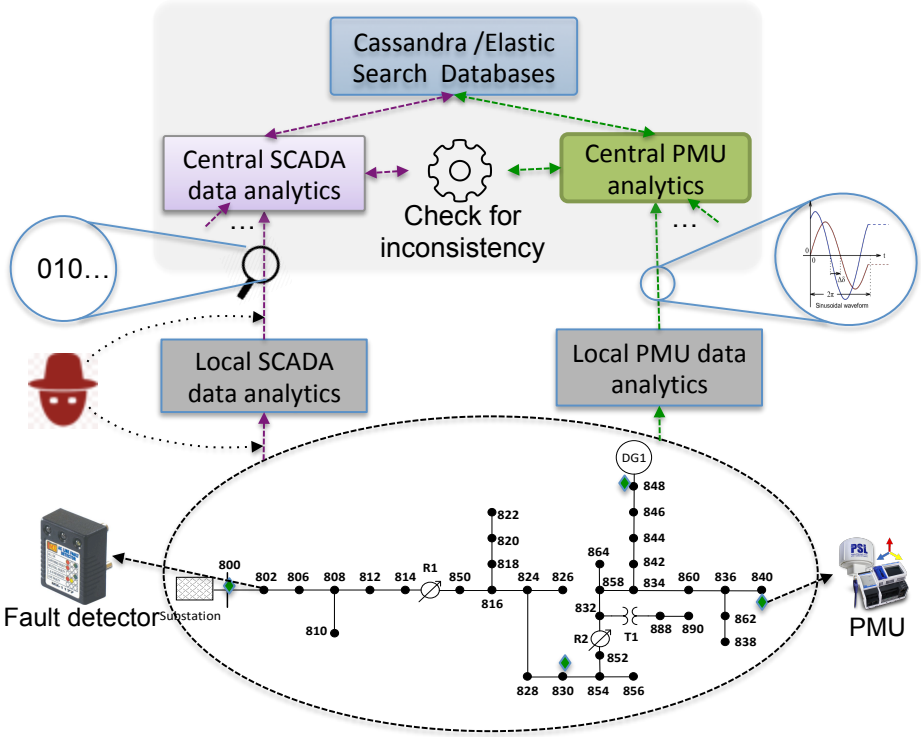


Figure 23: The LBNL SPARCS hierarchical intrusion detection system used in our experiments [Gen17, §4] [PGB+17]. Its components are: 1) a robust and open-source network monitoring framework called Bro; 2) a publisher-subscriber messaging system called RabbitMQ used to transfer PMU and SCADA analytics and data; 3) a NoSQL database system called Cassandra, used for permanent archiving of all data; 4) a NoSQL database system called Elasticsearch that records the analytics and receives events notifications in real time; 5) BeagleBoneBlack (BBB) devices that receive synchrophasor data at a rate of 120 samples/sec, and analyze 1 sec. worth of data to fast detect a cyber-physical event [JSP+16].

The FLISR System Operations and Vulnerabilities: FLISR systems detect the location of a permanent fault in a feeder and automatically restore service to customers in the healthy section of the feeder. To do so they employ directional fault detectors, installed

at every line in a distribution feeder, that communicate the occurrence of a fault and its direction to a distribution management system (DMS). The DMS analyzes these data and, once the faulted section is identified, issues commands to a predetermined set of switches to first isolate the faulted section and then restore service to the non-faulted areas [Ulu12], so as to minimize the service disruption. Consider the one-line diagram of a sample radial network shown in Fig. 24 to describe how FLISR works. The network is connected to a substation at bus 1 and a distributed generation is connected to bus 6. Fault detectors and normally-closed switches are connected at the end of every line. When a fault happens at line $3-4$, all the fault detectors detect the presence of this fault such that fault detectors at buses $1, 2$ and 3 indicate that the fault is in their right side, whereas fault detectors at buses $4, 5$ and 6 indicate the fault is to their left. Fault localization analysis at the DMS using this information identifies the faulted section to be line $3-4$. The DMS, therefore, sends control signals to automatically open switches to the right of bus 3 and to the left of bus 4 splitting the network into two parts and restoring service to customers on both sections.

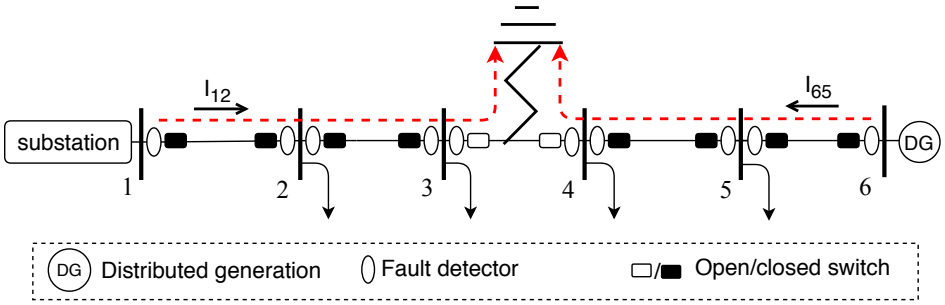


Figure 24: Diagram of a Sample Radial Network with a Fault in Line $3-4$.

Fault detector data are transmitted to the DMS through a wide area network using industrial protocols such as DNP3 [IEE12] or Modbus [Mod04] that are not secure by design. Therefore, a dedicated attacker can tamper with fault detector data either remotely, through a compromised network device (e.g., a router), or by physically connecting herself to an exposed section of the communication network. For instance, the tampering can lead the DMS analysis to pick the wrong location for an actual fault or completely hide the presence of a fault.

Although such attacks can be partially prevented by enabling photographically authenticated communication on top of the stated communication protocols, such a measure does not prevent the whole spectrum of possible cyber attacks. Therefore, a more robust approach, that does not merely rely on cryptographic solutions, is desired that is capable of detecting the presence of an attacker. The analysis we carried out in this paper shows that, even with a very limited number of sensors scattered in the systems, the PMU localization performed using proposed metric enhances an operator's confidence about the fault localization information (or the lack of thereof) extracted from fault detectors data, by giving an additional means to verify the trustworthiness of the SCADA messaging, albeit at a lower resolution. In the next section we illustrate the effectiveness of the forensic analysis on FLISR attacks

carried out by streaming simulated SCADA and PMU data to our LBNL testbed in Fig. 23.

Intrusion Detection on Fault Detectors:

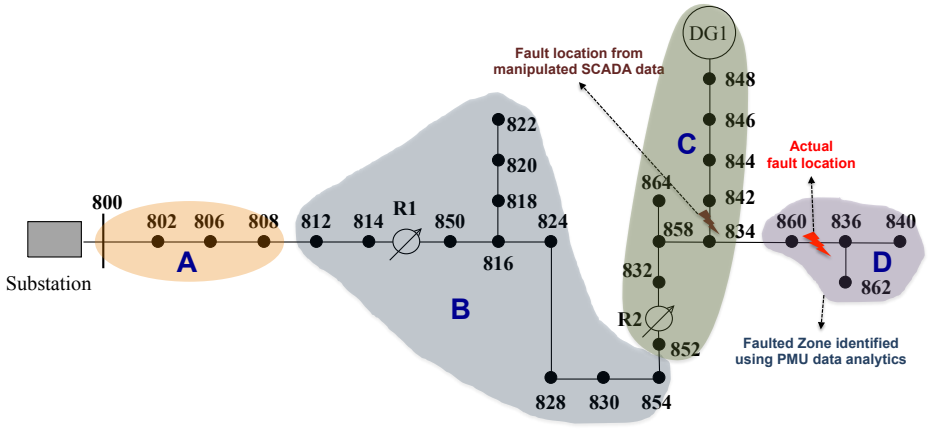


Figure 25: Output Mismatch as a Result of an Attack on SCADA data

In our simulation to demonstrate intrusion detection on fault detectors, we use the modified IEEE 34-bus system shown in Fig. 25. We introduce a two-phase AC-G fault at line $860 - 836$. Our PMU data analytic results correctly indicate the fault to be in Zone D. We generate simulated SCADA data for each fault detector in the network using the openDNP3 library [G+12] that implements the DNP3 communication protocol. In our simulation, we modify the packets from some of the fault detectors (those at buses 834 and 860) so that the SCADA data analytics indicate the fault is at line $834 - 842$, which is in zone C. The inconsistent results from the PMU data analytics and the SCADA data analytics raise an alarm about a possible cyber-attack on the fault detectors.

4 Summary

This project has developed a system for leveraging high-frequency power measurements, and, either on its own or in concert with SCADA traffic, is able to detect a variety of attacks on power distribution substation equipment, including transformer tap changes, capacitor bank switches, and "reconnaissance" attacks involving probing and surreptitious manipulation. This project builds on our earlier fundamental tenet around the notion that it is the *physics* of the power grid and the equipment that controls the grid that gives us the most useful insight in terms of how the grid is being manipulated, and what the consequences are.

This project has resulted in numerous peer-reviewed papers and academic conference presentations. However, beyond scholarly output, the primary goal of this work has always

been to develop this technology to the point that it could be sufficiently demonstrated to the power industry such that the power industry could adapt our work to enhance their own environments by implementing the security improvements that we have developed.

To that end, we have done several things. First, we have developed a robust platform, SPARCS, for collecting, ingesting, analyzing, visualizing and storing relevant power measurement and SCADA data, open sourced it, and made it publicly available online. Second, we have developed a robust set of analytics for detecting numerous classes of attacks against power distribution grid substations, SPARCS-Analytics, and also made that freely available for non-commercial use, and licenseable for commercial use.

Third, we have promoted our work significantly to industry by leveraging our industry partners to give talks at industry events, and present to and meet with potential users of our technology. The result of these discussions is that significant interest has been expressed. We note that one current limitation of our current implementation is that relatively few PSL μPMUs have been installed in U.S. distribution grid. However, adoption is growing, and international adoption, in particular, is very strong. Given this trend, and the interest in our approach we have strong reason to believe that either our technology in specific, or our technique in general will be adapted going forward as a method for detecting cyber attacks against distribution grids.

Acknowledgements

This research was supported in part by the Director, Office of Electricity Delivery and Energy Reliability, Cybersecurity for Energy Delivery Systems program, of the U.S. Department of Energy, under contract DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory. Any opinions, and findings expressed in this material are those of the authors and do not necessarily reflect those of the sponsors.

References

[15418] Ieee standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces. *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pages 1–138, April 2018.

[ARAS17] Daniel B Arnold, Ciaran Roberts, Omid Ardakanian, and Emma M Stewart. Synchrophasor data analytics in distribution grids. In *Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017 IEEE*, pages 1–5. IEEE, 2017.

[Bra11] Sukumar M Brahma. Fault location in power distribution system with penetration of distributed generation. *IEEE transactions on power delivery*, 26(3):1545–1553, 2011.

[Cas18] Cassandra. `https://cassandra.apache.org/`, 2018.

[Col13] Gerald Coley. Beaglebone black system reference manual. *Texas Instruments, Dallas*, 2013.

[Dug12] Roger C Dugan. Reference guide: The open distribution system simulator (opendss). *Electric Power Research Institute, Inc*, 7, 2012.

[G+12] Open DNP3 Group et al. Dnp3–distributed network protocol 3.0–google project hosting, 2012.

[Gen17] Reinhard Gentz. *Wireless Sensor Data Transport, Aggregation and Security*. PhD thesis, Arizona State University, 2017.

[HMCP04] Wendi B Heinzelman, Amy L Murphy, Hervaldo S Carvalho, and Mark A Perillo. Middleware to support sensor network applications. *IEEE Network*, 18(1):6–14, 2004.

[IEE] IEEE. http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html.

[IEE09] IEEE. IEEE Recommended Practice for Monitoring Electric Power Quality. *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, pages c1–81, June 2009.

[IEE11] IEEE. Ieee standard for synchrophasor measurements for power systems. *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pages 1–61, Dec 2011.

[IEE12] IEEE. IEEE Std 1815-2012 (Revision of IEEE Std 1815-2010) - IEEE Standard for Electric Power Systems Communications-Distributed Network Protocol (DNP3), 2012.

[JSP+16] Mahdi Jamei, Emma Stewart, Sean Peisert, Anna Scaglione, Chuck McParland, Ciaran Roberts, and Alex McEachern. Micro synchrophasor-based intrusion detection in automated distribution systems: Towards critical infrastructure security. *IEEE Internet Computing*, 20(5), 2016.

[JSR+17a] Mahdi Jamei, Anna Scaglione, Ciaran Roberts, Emma Stewart, Sean Peisert, Chuck McParland, and Alex McEachern. Anomaly Detection Using Optimally-Placed μPMU Sensors in Distribution Grids. *IEEE Trans. on Power Systems*, 2017.

[JSR+17b] Mahdi Jamei, Anna Scaglione, Ciaran Roberts, Emma Stewart, Sean Peisert, Chuck McParland, and Alex McEachern. Automated Anomaly Detection in Distribution Grids Using μPMU Measurements. In *Proceedings of the 50th Hawaii*

International Conference on System Sciences (HICSS), Electric Energy Systems Track, Resilient Networks Minitrack, January 2017.

[Ker01] William H Kersting. Radial distribution test feeders. In *Power Engineering Society Winter Meeting, 2001. IEEE*, volume 2, pages 908–912. IEEE, 2001.

[LAC16] Robert M Lee, Michael J Assante, and Tim Conway. Analysis of the Cyber Attack on the Ukrainian Power Grid. *SANS Industrial Control Systems*, 2016.

[Lai95] Tze Leung Lai. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 613–658, 1995.

[Mod04] IDA Modbus. Modbus messaging on tcp. *IP Implementation Guide v1. 0a, North Grafton, Massachusetts (www. modbus. org/specs. php)*, 2004.

[MP09] DIgSILENT Power Factory Manual and DIgSILENT PowerFactory. Version 14.0. *DIgSILENT GmbH, Gomaringen, Germany*, 2009.

[MPS14] Chuck McParland, Sean Peisert, and Anna Scaglione. Monitoring security of networked control systems: It's the physics. *Security & Privacy, IEEE*, 12(6):32–39, 2014.

[ope] Automatak. `https://www.automatak.com/opendnp3/`. Accessed: 08/03/2018.

[Pag54] ES Page. Continuous inspection schemes. *Biometrika*, pages 100–115, 1954.

[Pax99] Vern Paxson. Bro: a system for detecting network intruders in real-time. *Computer networks*, 31(23):2435–2463, 1999.

[PGB+17] Sean Peisert, Reinhard Gentz, Joshua Boverhof, Chuck McParland, Sophie Engle, Abdelrahman Elbashandy, and Daniel Gunter. Lbnl open power data. `http://powerdata-explore.lbl.gov`, 2017.

[RSJ+19] Ciaran Roberts, Anna Scaglione, Mahdi Jamei, Reinhard Gentz, Sean Peisert, Emma M Stewart, Chuck McParland, Alex McEachern, and Daniel Arnold. Learning behavior of distribution system discrete control devices for cyber-physical security. *IEEE Transactions on Smart Grid*, 11(1):749–761, 2019.

[Smi14] Rebecca Smith. Assault on california power station raises alarm on potential for terrorism. *Wall Street Journal*, 5, 2014.

[Ulu12] RW Uluski. Using distribution automation for a self-healing grid. In *Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES*, pages 1–5. IEEE, 2012.

[Var08] Kenton Varda. Protocol buffers: Google's data interchange format. *Google Open Source*, 72, 2008.

[XDM12] Yili Xia, Scott C Douglas, and Danilo P Mandic. Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators. *IEEE Signal Processing Magazine*, 29(5):44–54, 2012.

[XM12] Yili Xia and Danilo P Mandic. Widely linear adaptive frequency estimation of unbalanced three-phase power systems. *IEEE Transactions on Instrumentation and Measurement*, 61(1):74–83, 2012.

[Zer18] ZeroMQ. `https://www.zeromq.org/`, 2018.

Appendices

A Software Products

- Stream-Processing Architecture for near-Real-time Cyber-physical Security (SPARCS) `https://github.com/lbnl-cybersecurity/sparcs`

B List of Publications

Jamei, Mahdi, Raksha Ramakrishna, Teklemariam Tesfay, Reinhard Gentz, Ciaran Roberts, Anna Scaglione, and Sean Peisert. *"Phasor Measurement Units Optimal Placement and Performance Limits for Fault Localization" IEEE Journal on Selected Areas in Communications 38, no. 1 (2019): 180-192.*

Roberts, Ciaran, Anna Scaglione, Mahdi Jamei, Reinhard Gentz, Sean Peisert, Emma M. Stewart, Chuck McParland, Alex McEachern, and Daniel Arnold. *"Learning Behavior of Distribution System Discrete Control Devices for Cyber-Physical Security" IEEE Transactions on Smart Grid 11, no. 1 (2019): 749-761.*

Jamei, Mahdi, Anna Scaglione, Sean Peisert. *"Low-Resolution Fault Localization Using Phasor Measurement Units with Community Detection." IEEE International Conference on Smart Grid Communications (SmartGridComm) (2018).*

Jamei, Mahdi, Anna Scaglione, Ciaran Roberts, Emma Stewart, Sean Peisert, Chuck McParland, and Alex McEachern. *"Anomaly detection using optimally-placed μPMU sensors in distribution grids." IEEE Transactions on Power Systems (2017).*

Jamei, Mahdi, Anna Scaglione, Ciaran Roberts, Alex McEachern, Emma Stewart, Sean Peisert, and Chuck McParland. *"Online Thevenin Parameter Tracking Using*

Synchrophasor Data." In Power & Energy Society General Meeting, 2017 IEEE, pp. 1-5. IEEE, 2017.

Jamei, Mahdi, Anna Scaglione, Ciaran Roberts, Emma Stewart, Sean Peisert, Chuck McParland and Alex McEachern *"Automated Anomaly Detection in Distribution Grids Using micro-PMU Measurements."* In Proceedings of the 50^{th} Hawaii International Conference on System Sciences. 2017

Jamei, Mahdi, Emma Stewart, Sean Peisert, Anna Scaglione, Chuck McParland, Ciaran Roberts, and Alex McEachern. *"Micro Synchrophasor-Based Intrusion Detection in Automated Distribution Systems: Toward Critical Infrastructure Security."* IEEE Internet Computing 20, no. 5 (2016): 18-27.

Roberts, Ciaran, Anna Scaglione and Sean Peisert. *"A Holistic Approach to Distribution Grid Intrusion Detection Systems"* EnergyCentral, July 18, 2018

C Additional Presentations

"Computer Security & the Electric Power Grid", 15th Annual ON*VECTOR Photonics Workshop, UC San Diego, La Jolla, CA, March 1, 2016,

"Managing Energy: Role of Data and Security", Prospect Silicon Valley 2017 Innovation and Impact Symposium, Microsoft Research Silicon Valley, Mountain View, CA, June 14, 2017,

"Cybersecurity for the Electric Grid", Bits and Watts Annual Workshop, Hoover Institution, Stanford, CA, Nov. 6, 2017.

"Cyber Security of Power Distribution Systems Using Micro-Synchrophasor Measurements and Cyber-Reported SCADA", EPRI Power Delivery and Utilization (PDU) Program Advisory Meeting February 5-7, 2018 – Loews Coronado Bay Resort, Coronado, CA 92118

"Cyber Security of Power Distribution Systems Using Micro-Synchrophasor Measurements and Cyber-Reported SCADA", OSIsoft PI WORLD 2019, April 23-27, 2018 San Francisco, CA

D Featured Articles

"Combination of Old and New Yields Novel Power Grid Cybersecurity Tool" DOE/LBNL Press Release, 7 March 2018

"Cyber Defense Tool Is an Early Warning System for Grid Attacks" IEEE Spectrum Energywise, 27 March 2018