Anomaly Detection Using Optimally-Placed μPMU Sensors in Distribution Grids

Mahdi Jamei, Anna Scaglione, Ciaran Roberts, Emma Stewart,
Sean Peisert, Chuck McParland, Alex McEachern

Abstract—As the distribution grid moves toward a tightly-monitored network, it is important to automate the analysis of the enormous amount of data produced by the sensors to increase the operators situational awareness about the system. In this paper, focusing on Micro-Phasor Measurement Unit (μPMU) data, we propose a hierarchical architecture for monitoring the grid and establish a set of analytics and sensor fusion primitives for the detection of abnormal behavior in the control perimeter. Due to the key role of the μPMU devices in our architecture, an optimal μPMU placement with limited number of sensors is also described that finds the best location of the devices with respect to our rules. The effectiveness of the proposed methods are tested through the synthetic and real μPMU data.

Index Terms—Distribution Grid, Micro-Phasor Measurement Unit (μPMU), Anomaly Detection, Optimal Placement.

I. Introduction

Historically, Distribution System Operators (DSOs) have lacked information about the real-time operation of the grid. However, this situation is changing. Similar to the Phasor Measurement Units (PMUs) at the transmission level, Micro-Phasor Measurement Units (μPMUs) are designed to measure real-time voltage and current phasors in a synchronized manner at specific locations on the distribution system. These measurements can be utilized by the DSO, but useful translation from raw data to information is first required [2], [3]. This information can provide a level of insight that is not attainable by Distribution Supervisory Control and Data Acquisition (DSCADA) data. Visualizing and interpreting raw sensor streams can be overwhelming for DSOs, considering the large quantity of data that could flow in from different parts of the grid [4]. Therefore, it is essential to mine the data collected with analytic tools that can derive informative measurements and form automated reports.

A. Literature Review

Recent work has focused on the PMU data utilization at the transmission level to improve the Wide-Area Monitoring, Protection, and Control [5], [6]. Distribution grids, however, are still lagging in that respect, since tools for the transmission grid may not be directly applicable to the distribution grid due

to a different operation environment, such as load imbalances, untransposed lines, and the existence of single-phase and two-phase laterals. Transmission operations and system wide analysis are concerned with large imbalances in load and generation, and as a result frequency, whereas distribution operations are concerned with localized, but frequent events such as voltage imbalance, overloading, and outage management. The goal of this paper is to fill the distribution systems analysis gap by defining *physics-aware* algorithms that process μPMU data for the automatic detection of any grid behavior that does not comply with the (quasi) steady-state regime of operation, as well as the physical limits of the grid. Most of the prior research on sensor data analytics (including SCADA and PMU measurements) is concerned with detecting events on the grid transmission level. For example, Pan et al., [7] use data mining techniques on PMU measurements and audit logs for event classification. A linear basis expansion for the PMU data is described by Chen et al., [8], [9] for event detection application. A similar approach, based on Principal Component Analysis (PCA), is used in [10], [11] for event detection and data archival. Allen et al., [12] describe the use of voltage phasor angle differences between different PMU readings to detect events. Biswal et al., [13] use the strongest signatures of event in PMU data for situational awareness enhancement. In our previous work [14] we describe an architecture for a cyber-physical distribution grid, and show how event detection using μPMU measurements can help in reasoning about the security status of the grid. A study by Brahma et al., [15] describes the real-time dynamic event identification in power system using PMU data based on a data-driven and also physics-based method.

B. Our Contribution

Our approach for event detection is a combination of the data-driven methods, as well as criteria resulting from analyzing the underlying physical model of the system. We first define a hierarchical *"anomaly detection architecture"* to host a set of anomaly detection rules that are proposed for the analysis and fusion of μPMU sensor measurements. Unlike model agnostic machine learning methods that look for statistical anomalies in a feature space that is often heuristic, the *anomalies* that algorithms in this paper identify are defined in the context of power quality and protection, in addition to what is imposed by the grid governing physical equations (i.e. Kirchhoff Voltage and Current Laws). This, in turn, gives a DSO much greater insight and help in the forensic analysis.

Because of the important role played by μPMU data in the framework, a μPMU placement optimization with limited number of the sensors is described, to achieve the maximum sensitivity in detecting a change[1].

C. Notation

The following notations are used throughout the paper:

j	Imaginary unit.
$\mathcal{I}_N$	$N \times N$ identity matrix.
$\mathbf{A}^T, \mathbf{A}^*,$	Transpose, conjugate,
$\mathbf{A}^H$	and conjugate transpose of matrix $\mathbf{A}$.
$\|\mathbf{A}\|, \|\mathbf{A}\|_F$	2-norm and F-norm of matrix $\mathbf{A}$.
$\mathbf{1}_{m \times n}$	$m \times n$ size matrix with entries 1.
$\mathbf{A}^\dagger$	Pseudo-inverse of matrix $\mathbf{A}$.
$\otimes$	Kronecker product.

II. ANOMALY DETECTION ARCHITECTURE

As shown in Fig. 1 the first stage of the *"anomaly detection architecture"* hierarchy is next to each μPMU sensor in the field and the second stage is at a central node that can be hosted in the Distribution System Operator (DSO) control room. The algorithms for anomaly detection applied next to each μPMU sensor are referred to as *"local rules"* and those that aggregate readings of multiple μPMUs are referred as *"central rules"*. For large grid sizes, the aggregation can occur
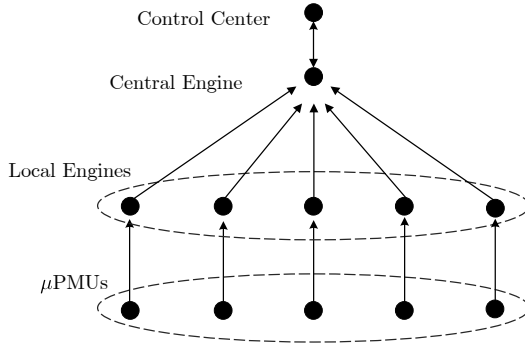

Fig. 1. Anomaly Detection Architecture Using μPMU Measurements.

in multiple steps, where mid-level stages analyze part of the data and forward them upward.

This design has the following benefits: (1) it is robust due to its hierarchical, redundant nature; (2) it can be inexpensively deployed at existing utilities; (3) it analyzes the data close to real-time; as opposed to the common practice of post hoc auditing of logged data; (4) the employed μPMUs are placed in separate protection domains—including a separate communication network—from the DSCADA systems, so that attacks against the DSCADA network will not compromise the operation of the architecture. In addition, μPMUs can be seamlessly added to the network via wireless modems. Further addressing security, the μPMUs are not only designed to be read-only via the network, but are both placed

---

[1]Note that PMU placement studies for state estimation are not applicable here since the state is not directly observable by solely depending on the μPMUs due to the scarcity of sensors.

behind networked "bastion hosts" and can communicate bi-directionally via encrypted protocols, thereby fortifying the devices against tampering without physical access; (5) finally, even if a data injection attack can be launched against some μPMUs, depending on level that is compromised (from field devices up to the control center), some μPMUs and some of our rules may still be effective, due to the fact that every device in a certain level processes information independently. Thus, in our threat model, it is assumed that our detection rules are less-vulnerable to data injection attacks than attacks against typical DSCADA networks.

The suite of algorithms that are proposed for the *anomaly detection architecture* 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, thus 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 due to poor modeling; (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. These modeling aspects are clarified in our discussion next.

III. THE μPMU DATA IN A DISTRIBUTION GRID

Fig. 2 shows the π model of a distribution line that connects bus m to n. Assuming normal conditions, the μPMUs are
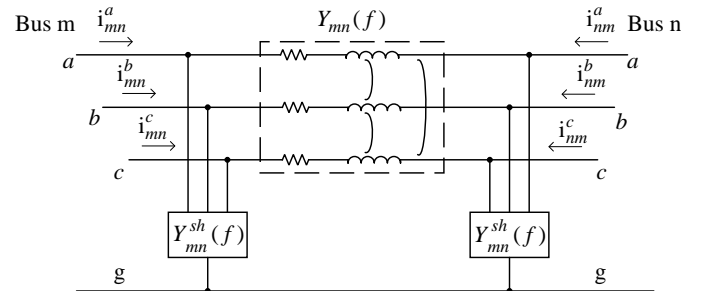

Fig. 2. π Model of a Distribution Line

designed to capture samples of the three phase voltage phasor, which is denoted by $\mathbf{v}[k] \in \mathbb{C}^{3 \times 1}$, and of the current phasor $\mathbf{i}[k] \in \mathbb{C}^{3 \times 1}$ in specific sites of a distribution grid. Next we apply Kirchhoff's and Ohm's law for a three-phase line in a quasi steady-state condition.

A well-known fact from signals and systems theory is that the relationship between voltage and current through a passive circuit with a certain admittance matrix can be represented as a multiplication in the frequency domain and as a convolution in the time domain. Because the circuit is three phase, these will be represented by a Multi-Input Multi-Output system. This fact also holds for the phasor of the signals. We first define $\boldsymbol{y}_{mn}(t)$ and $\boldsymbol{y}_{mn}^{sh}(t)$ to denote the time domain equivalents of the matrices $\boldsymbol{Y}_{mn}(f + f_0)$ and $\boldsymbol{Y}_{mn}^{sh}(f + f_0)$, respectively,

where f_0 is the main frequency of the grid. In discrete time, the convolution relationship is as follows:

$$\mathbf{i}_{mn}[k] = \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]\mathbf{v}_m[k-r] - \boldsymbol{y}_{mn}[r]\mathbf{v}_n[k-r] \quad (1)$$

where $\overline{\boldsymbol{y}}_{mn}[r] = \boldsymbol{y}_{mn}[r] + \boldsymbol{y}_{mn}^{sh}[r]$ and it is assumed that $\boldsymbol{y}_{mn}^{sh}[r]$ and $\boldsymbol{y}_{mn}[r]$ are the samples of $\boldsymbol{y}_{mn}^{sh}(t) \star h(t)$ and $\boldsymbol{y}_{mn}(t) \star h(t)$, respectively, and have finite support N and are causal[2].

Next, the form of this relationship during the quasi steady-state is shown, since the steady-state in reality never happens, which in turn represents the governing equations during a regime of operation that is referred to as *"normal"*.

The fundamental frequency of the voltage and current signals are always varying, even in a normal state (although slowly and over a very small range), because of load-generation imbalances, active power demand interactions, large generators inertia, and the automatic speed controllers of the generators [16]. The effect observable is a change of grid operation regime from steady-state to quasi steady-state. The off-nominal frequency therefore affects the phase angle captured by phasor measurement devices. To highlight that mathematically, the phasor readings $\mathbf{v}_m[k]$ and $\mathbf{i}_{mn}[k]$ are decomposed as follows:

$$\mathbf{v}_m[k] = \hat{\mathbf{v}}_m[k]e^{j\beta_m[k]k}, \;\; \mathbf{i}_{mn}[k] = \hat{\mathbf{i}}_{mn}[k]e^{j\beta_m[k]k} \quad (2)$$

where $\hat{\mathbf{v}}_m[k]$ is the voltage phasor that is observed at nominal frequency, and $\hat{\mathbf{i}}_{mn}[k]$ is the current phasor after removing the exponential term due to the off-nominal frequency at bus m, $\beta_m[k]$, that represents the (time-varying) drift in the frequency induced by the above-mentioned reasons.

Considering the Discrete Time Fourier Transform, we have:

$$\boldsymbol{Y}_{mn}(f+f_0)H(f) = rect(T_s f)T_s \sum_{r=0}^{N-1} \boldsymbol{y}_{mn}[r]e^{-j2\pi r T_s f} \quad (3)$$

where $\frac{1}{T_s} = 120$Hz is the μPMU output rate, and $H(f)$ is the frequency response of the filter $h(t)$. Introducing:

$$\boldsymbol{Y}_{mn}(f_0,k) \triangleq \frac{1}{T_s}\boldsymbol{Y}_{mn}\left(f_0 + \frac{\beta_n[k]}{2\pi T_s}\right)H\left(\frac{\beta_n[k]}{2\pi T_s}\right), (4)$$

we have that

$$\boldsymbol{Y}_{mn}(f_0,k) = \sum_{r=0}^{N-1} \boldsymbol{y}_{mn}[r]e^{-j\beta_n[k]r}, \quad (5)$$

and similarly $\overline{\boldsymbol{Y}}_{mn}(f_0,k) = \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]e^{-j\beta_m[k]r}$.

During the quasi steady-state the variations in $\hat{\mathbf{v}}_m[k]$ and $\beta_m[k]$ are negligible over a window of N samples, which

---

[2]$h(t)$ denotes the low-pass filter implemented in the μPMU to extract the base-band signal and $H(f)$ is its corresponding frequency response. It should be noted that the outputs of μPMU, and their corresponding functions, are not the exact phasors if the bandwidth of the signal (voltage/current) is greater than $2f_0$, and are aliased. However, what we are interested in is to track any data abnormality (instead of the exact grid state during the event). Therefore, as long as it is not suppressing the anomaly, having aliasing is not an issue for our rules.

means $\hat{\mathbf{v}}_m[k-r] \approx \hat{\mathbf{v}}_m[k]$ and $\beta_m[k-r] \approx \beta_m[k]$. Using this approximation, (1) can be re-written as:

$$\begin{aligned}\mathbf{i}_{mn}[k] &\approx \sum_{r=0}^{N-1} \overline{\boldsymbol{y}}_{mn}[r]\hat{\mathbf{v}}_m[k]e^{j\beta_m[k](k-r)}\\ &- \boldsymbol{y}_{mn}[r]\hat{\mathbf{v}}_n[k]e^{j\beta_n[k](k-r)}\\ &= \overline{\boldsymbol{Y}}_{mn}(f_0,k)\mathbf{v}_m[k] - \boldsymbol{Y}_{mn}(f_0,k)\mathbf{v}_n[k]\end{aligned} \quad (6)$$

Equation (6) is Ohm's law in the phasor domain and is the cornerstone for our transients detection algorithm derived in the following section. The analysis above explains through (6) why in the phasor domain, the equivalent effect of the quasi-steady state conditions is that the admittances (4) fluctuate. The effect is usually modest, because $\beta[k]$ is small. However, during a severe transient or frequency in the order of 10Hz the relationship (6) with the matrices in (4) does not hold, due to the manifestation of the full dynamic behavior in (1).

In the following, the *"local"* and *"central"* rules are defined leveraging these insights and the knowledge about power system operation. Note that our rules are set up in a way that all the local engines are agnostic about $\overline{\boldsymbol{Y}}_{mn}(f_0,k)$ and $\boldsymbol{Y}_{mn}(f_0,k)$, and the sensor siting. However, it is assumed that the central engine knows $\boldsymbol{Y}_{mn}^{sh}(f_0,k)|_{\beta[k]=0}$ and $\boldsymbol{Y}_{mn}(f_0,k)|_{\beta[k]=0}$ for the lines within its monitoring region and the difference from (4) is treated as equivalent to noise in the observation model. Therefore, when dealing with central rules, we will simply use $\boldsymbol{Y}_{mn}^{sh}$ and $\boldsymbol{Y}_{mn}$ to refer to $\boldsymbol{Y}_{mn}^{sh}(f_0,k)|_{\beta[k]=0}$ and $\boldsymbol{Y}_{mn}(f_0,k)|_{\beta[k]=0}$.

IV. DATA ANALYSIS

Our rules monitor for abnormalities in the following quantities 1) voltage magnitude, 2) current magnitude, 3) active power, 4) reactive power, 5) instantaneous frequency drift, and 6) the validity of quasi steady-state equations. The *"local rules"* just require the stream of phasors from a single μPMU, while the *"central rules"* combine multiple streams across μPMUs.

A. Local Rules

The *"local rules"* are applied at the lowest layer of the *"anomaly detection architecture"*, and run on systems adjacent to the μPMUs. Their common feature is that they require no specific prior knowledge of the grid network parameters.

TABLE I
VOLTAGE MAGNITUDE ANOMALIES

anomaly	signature [3]
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]The voltage magnitude is in p.u.

1) Voltage Magnitude Changes: The magnitude of the voltage varies within a small range that power quality standards enforce during the normal operations [17]. Therefore, any large deviation from that range indicates an abnormal condition. Table. I lists the anomalies that can be observed in the voltage magnitude labeled by their severity and duration, denoted by $|\mathbf{v}|$ and τ, respectively.

2) Current Magnitude, Active, and Reactive Power Changes: Even when the voltage magnitude is within the safe range discussed previously, changes in active and reactive power can still happen due to the change of the load, affecting current magnitude and the phase angle between current and voltage phasors. Therefore, it is also of interest to track fast-changes of these quantities using the method described in Section. IV-C.

The three phase apparent power can simply be computed as:

$$\mathbf{S}_{mn}[k] = \mathbf{P}_{mn}[k] + j\mathbf{Q}_{mn}[k] = \mathrm{diag}(\mathbf{v}_m[k])\mathbf{i}^*_{mn}[k] \quad (7)$$

where $\mathbf{P}_{mn}[k]$ and $\mathbf{Q}_{mn}[k]$ are the three-phase active and reactive powers, respectively. Note that for the sake of tracking a power flow change in the distribution grid, tracking the active power and reactive power is preferable over monitoring the phase angle difference, since the resistance of the lines is not negligible, and therefore the angle difference does not necessarily indicate the direction of the power flow. We observed this fact in our partner utility grid, when the voltage phase angle at one end of the line was less than the angle at the other end, though the direction of the power flow was not from the higher angle to the lower one.

Because physical changes in this class of data on the distribution grid can be potentially indicative of negative behavior, it is important to determine the direction (upward trend, downward trend or oscillation) of the change. The anomalies related to fast changes are labeled 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 have primarily introduced 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 exceeded leading to accelerated failure. During the quasi steady-state, the three phase current phasor magnitude flowing in each line, i.e., $|\mathbf{i}_{mn}[k]|$, should be less than or equal the line rated current, $|\mathbf{i}_{mn}|_{\max}$. This constraint is imposed as feeder limit, and the violation is flagged as *overcurrent*.

3) Instantaneous Frequency Changes: For a μPMU at bus m, we propose to estimate adaptively the instantaneous local frequency deviation from the nominal frequency during the quasi-steady state, using the approach for instance in [18] that is tailored to three-phase distribution lines, to isolate abnormal changes in the estimated frequency.

4) Quasi Steady-State Regime Validity: As previously noted, when the grid is not in the normal quasi steady-state conditions, the relationship between voltage and the current phasors represents its full dynamic behavior, i.e. the grid is no

longer well-approximated by the set of memory-less algebraic equations. Therefore, it is proposed to check the validity of the quasi steady-state regime to flag the presence of transients in the grid. At the local engines, (6) provides the basis for our rule. For the line in Fig. 2, assuming that a μPMU is installed at bus m means that $\mathbf{i}_{mn}[k]$, and $\mathbf{v}_m[k]$ are both available. Let $\boldsymbol{\alpha}[k]$ be the diagonal matrix such that voltage phasors of bus m and bus n, connected via a power line, are related through:

$$\mathbf{v}_n[k] = \boldsymbol{\alpha}[k]\mathbf{v}_m[k], \quad (8)$$

Defining:

$$\boldsymbol{R}^{(mn)}_{\mathrm{iv}}[k] = \frac{1}{M-1}\sum_{r=0}^{M-1}\mathbf{i}_{mn}[k-r]\mathbf{v}^H_m[k-r], \quad (9)$$

$$\boldsymbol{R}^{(nm)}_{\mathrm{vv}}[k] = \frac{1}{M-1}\sum_{r=0}^{M-1}\mathbf{v}_n[k-r]\mathbf{v}^H_m[k-r], \quad (10)$$

Assuming that $\boldsymbol{\alpha}[k]$ remains constant over a window of M samples in the quasi-steady state, one can write:

$$\boldsymbol{R}^{(nm)}_{\mathrm{vv}}[k] \approx \boldsymbol{\alpha}[k]\boldsymbol{R}^{(mm)}_{\mathrm{vv}}[k] \quad (11)$$

It can be assumed that the variation of $\overline{\boldsymbol{Y}}_{mn}(f_0,k)$ is negligible over M samples in normal operation and use (6) to write:

$$\left(\ \mathcal{I}_3\ \vdots\ -\overline{\boldsymbol{Y}}_{mn}(f_0,k) + \boldsymbol{Y}_{mn}(f_0,k)\boldsymbol{\alpha}[k]\ \right)\underbrace{\begin{pmatrix}\boldsymbol{R}^{(mn)}_{\mathrm{iv}}[k]\\ \boldsymbol{R}^{(mm)}_{\mathrm{vv}}[k]\end{pmatrix}}_{\boldsymbol{R}^{(mn)}_k} \approx \mathbf{0}$$

Proposition 1. *Correlation matrix $\boldsymbol{R}^{(mn)}_k$ is approximately rank-1 during the quasi steady-state.*

Proof. During the quasi-steady state along a distribution line, the following assumptions hold with a very good approximation for $r = 0, 1, ..., M-1$:

$$\hat{\mathbf{v}}_m[k-r] \approx \hat{\mathbf{v}}_m[k], \quad \beta_m[k-r] \approx \beta_m[k] \quad (12)$$

Therefore, we can write:

$$\boldsymbol{R}^{(mm)}_{\mathrm{vv}}[k] = \frac{1}{M-1}(\mathbf{v}_m[k]\otimes\mathbf{E}_m[k])(\mathbf{v}^H_m[k]\otimes\mathbf{E}^H_m[k])$$
$$= \frac{1}{M-1}(\mathbf{v}_m[k]\mathbf{v}^H_m[k])\otimes(\mathbf{E}_m[k]\mathbf{E}^H_m[k]) \quad (13)$$

where $\mathbf{E}_m[k]$ is defined as follows and represents the variations due to the off-nominal frequency:

$$\mathbf{E}_m[k] = \mathbf{1}_{3\times1}\otimes\left(e^{-j\beta_m[k](M-1)}\ \ \ldots\ \ e^{-j\beta_m[k]}\ \ 1\right) \quad (14)$$

which we can then write:

$$\mathbf{E}_m[k]\mathbf{E}^H_m[k] = (\mathbf{1}_{3\times1}\mathbf{1}_{1\times3})\otimes(M) = M\mathbf{1}_{3\times3} \quad (15)$$

and therefore:

$$\boldsymbol{R}^{(mm)}_{\mathrm{vv}}[k] = \frac{M}{M-1}(\mathbf{v}_m[k]\mathbf{v}^H_m[k])\otimes(\mathbf{1}_{3\times3}) \quad (16)$$

which accordingly means that:

$$rank(\boldsymbol{R}^{(mm)}_{\mathrm{vv}}[k]) = rank(\mathbf{v}_m[k]\mathbf{v}^H_m[k])\times rank(\mathbf{1}_{3\times3}) = 1$$

Because:

$$rank(\boldsymbol{R}_k^{(mn)}) = rank((\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)})$$

we analyze the rank of $(\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)}$ here, where:

$$(\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)} = (\boldsymbol{R}_{\mathrm{iv}}^{(mn)}[k])^H \boldsymbol{R}_{\mathrm{iv}}^{(mn)}[k] + \\ (\boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k])^H \boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k] \tag{17}$$

From the structure of (12) during the quasi-steady state, we have:

$$\boldsymbol{R}_{\mathrm{iv}}^{(mn)}[k] = \tilde{\mathbf{Y}}_{mn}(f_0, k) \boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k] \\ \tilde{\mathbf{Y}}_{mn}(f_0, k) = \overline{\boldsymbol{Y}}_{mn}(f_0, k) - \boldsymbol{Y}_{mn}(f_0, k)\mathrm{diag}(\boldsymbol{\alpha}[k]) \tag{18}$$

Substituting (18) in (17), we have:

$$(\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)} = (\boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k])^H \boldsymbol{\mathcal{Y}}_{mn}(f_0, k) \boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k] \tag{19}$$

where:

$$\boldsymbol{\mathcal{Y}}_{mn}(f_0, k) = \tilde{\mathbf{Y}}_{mn}^H(f_0, k)\tilde{\mathbf{Y}}_{mn}(f_0, k) + \mathcal{I}$$

Since the linear transformation of $\boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k]$ does not increase its rank and since it has already been shown that $\boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k]$ is of rank-1 during the quasi-steady state, one can conclude that:

$$rank(\boldsymbol{R}_k^{(mn)}) = rank((\boldsymbol{R}_k^{(mn)})^H \boldsymbol{R}_k^{(mn)}) \leq \\ rank(\boldsymbol{R}_{\mathrm{vv}}^{(mm)}[k]) \rightarrow rank(\boldsymbol{R}_k^{(mn)}) = 1 \tag{20}$$

$\square$

Hence:

$$\boldsymbol{R}_k^{(mn)} \approx \sigma_1^{(mn)}[k]\boldsymbol{u}_1^{(mn)}[k](\boldsymbol{\nu}_1^{(mn)}[k])^H \rightarrow \\ \boldsymbol{R}_k^{(mn)}(\boldsymbol{R}_k^{(mn)})^H \approx (\sigma_1^{(mn)}[k])^2 \boldsymbol{u}_1^{(mn)}[k](\boldsymbol{u}_1^{(mn)}[k])^H \tag{21}$$

where $\sigma_1^{(mn)}$ is the largest singular value of $\boldsymbol{R}_k^{(mn)}$, and $\boldsymbol{u}_1^{(mn)}$ and $\boldsymbol{\nu}_1^{(mn)}$ are the corresponding left and right singular vectors to that, respectively. Deviation from this subspace structure can indicate that the line is experiencing a transient. We can automate the detection of anomaly using this criterion by computing the following cost minimization and tracking the fast changes in $x[k]$ for each incidental line to a bus with μPMU:

$$x[k] = \min_{\boldsymbol{u}} ||(\mathcal{I}_6 - \boldsymbol{u}\boldsymbol{u}^H)\boldsymbol{R}_k^{(mn)}(\boldsymbol{R}_k^{(mn)})^H||_F \\ \text{s.t. } ||\boldsymbol{u}|| = 1 \tag{22}$$

In other words, $x[k]$ measures the size of the residual that $\boldsymbol{R}_k^{(mn)}$ has in the space orthogonal to the optimal $\boldsymbol{u}$, which should be zero in the quasi steady-state and non-zero otherwise. Since it has already been shown that $\boldsymbol{R}_k^{(mn)}$ is of rank-1 during the quasi steady-state, the left singular vector corresponding to the largest singular value of $\boldsymbol{R}_k^{(mn)}$ is the only quantity of interest to compute the metric, instead of computing all the singular vectors. Therefore, also owing it to the small size of the matrix, the method is not computationally very expensive.

To conclude the local rules, a flowchart of the analysis performed at the local engine next to μPMU at bus m is presented in Fig. 3. At each instant of time, the phasor readings
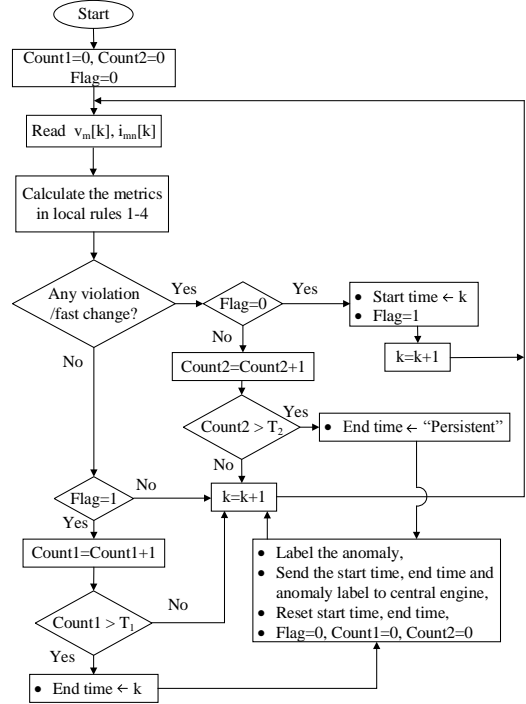

Fig. 3. Local Engine Analysis Flowchart Next to μPMU at Bus m.

are received by the local engine and the introduced metrics above are calculated. The pre-processed data are then passed to the local rules to check for any violation. A violation could be trespassing pre-defined limits (e.g., the voltage magnitude rule, or maximum current magnitude limit) or fast changes in a metric with smooth behavior during the normal condition (e.g., quasi steady-state validity rule). Once a violation is found in one of the metrics, the start time is recorded. The search for anomaly continues until no new violation is found for a certain window of time (*"Count1 > T_1"*), and that specifies the end time of the anomaly. The type of anomaly is then determined based on the behavior of the data between the start time and the end time (e.g., active power surge, voltage interruption,...). The start time, end time and the anomaly label is then sent upstream for further analysis/visualization, and the parameters are reset for next event. If the number of detected violations related to a certain event passes a pre-defined threshold (*"Count2 > T_2"*), the end time is replaced with a *"Persistent"* label, and the results are sent to the central engine, without waiting for the end of the event to arrive. The reason is to be able to inform the operator about the anomaly in time, and not waiting too long before something more damaging happens.

B. Central Rule

At higher levels of aggregation, the central engine in our case, the natural way to relate the measurements is through the grid interconnection. For a grid with B buses, let $\mathbf{I}[k]$ denote the vector of three-phase current injection phasors with size $3B$, and $\mathbf{V}[k]$ represent the vector of three-phase voltage phasor at all the buses, which contains $3B$ elements. We define

the measurement vector $\mathbf{d}[k] = (\mathbf{I}^T[k], \mathbf{V}^T[k])^T$. During the steady-state, the following algebraic equation should hold:

$$\mathbf{H}\mathbf{d}[k] = \mathbf{0}, \quad \mathbf{H} = \left(\begin{array}{c|c} \mathcal{I}_{3B} & -\boldsymbol{Y}_{3(B \times B)} \end{array} \right) \quad (23)$$

where $\boldsymbol{Y}$ is the admittance matrix of the grid that connects the current injection to the bus voltages. During the quasi steady-state these equations are close to be homogeneous. Since the distribution grid is unbalanced, and the lines are not transposed the set of equations that should be dealt with are three phase instead of working with positive sequence [19]. In this framework, we are also able to include the laterals in the admittance matrix by putting the entries corresponding to the phases that do not exist equal to zero.

It is assumed that $\mathbf{H}$ is known, K denotes the number of μPMUs that are available and that each device has enough channels to measure the voltage and all incidental current measurements of the bus on which it is placed. Therefore, having a μPMU at bus m means that the entries in $\mathbf{V}[k]$ corresponding to $\mathbf{v}_m[k]$ and entries in $\mathbf{I}[k]$ corresponding to $\sum_{n:m \sim n} \mathbf{i}_{mn}[k]$ are both available, where $m \sim n$ denotes that bus m and n are connected through a line. Let $\mathbf{T}$ denote a matrix parsing the vector $\mathbf{d}[k]$ into two parts corresponding to the unavailable measurements, $\mathbf{d}_u[k]$, and the available measurements, $\mathbf{d}_a[k]$:

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_u \\ \mathbf{T}_a \end{pmatrix} \;\rightarrow\; \mathbf{T}\mathbf{d} = \begin{pmatrix} \mathbf{d}_u \\ \mathbf{d}_a \end{pmatrix}, \; \mathbf{H}\mathbf{T}^T = \left(\begin{array}{c|c} \mathbf{H}_u & \mathbf{H}_a \end{array} \right) \quad (24)$$

where $K' = B - K$ and

$$\mathbf{T}_u \in \{0,1\}^{6(K' \times B)}, \quad \mathbf{T}_a \in \{0,1\}^{6(K \times B)} \quad (25)$$

$\mathbf{T}_u$ and $\mathbf{T}_a$ here are block diagonal matrices of size $6(K' \times B)$ and $6(K \times B)$ with entries equal to 0 or 1. The former is supposed to select the unavailable current and voltage phasors by having entries equal to 1 at corresponding locations, and the latter is supposed to pick the available phasors. Since $\mathbf{T}^T \mathbf{T} = \mathcal{I}$, we can rewrite (23) as follows:

$$\mathbf{H}_u \mathbf{d}_u[k] + \mathbf{H}_a \mathbf{d}_a[k] = \mathbf{0} \rightarrow \quad (26a)$$
$$\mathbf{H}_a \mathbf{d}_a[k] = -\mathbf{H}_u \mathbf{d}_u[k]. \quad (26b)$$

Let first assume that we have enough μPMUs that satisfy $K > \frac{B}{2}$. Therefore, the matrix $\mathbf{H}_u$ would be a tall matrix and has a left null-space. Premultiplying both sides of (26b) by the projector on the left null-space of $\mathbf{H}_u$, we have:

$$(\mathcal{I} - \mathbf{H}_u \mathbf{H}_u^\dagger) \mathbf{H}_a \mathbf{d}_a[k] = \mathbf{0} \quad (27)$$

because $(\mathcal{I} - \mathbf{H}_u \mathbf{H}_u^\dagger)\mathbf{H}_u = \mathbf{0}$. The equality in (26b) only holds during the steady-state (and for quasi steady-state with a good approximation), which means that equation (27) is homogeneous only in the steady-state (nearly-homogeneous in the quasi steady-state), and non-homogeneous otherwise. The following metric should, therefore, be close to zero only during normal operation and if $\mathbf{H}$ is unchanged:

$$x[k] = \frac{||(\mathcal{I} - \mathbf{H}_u \mathbf{H}_u^\dagger) \mathbf{H}_a \mathbf{d}_a[k]||_2^2}{||\mathbf{d}_a[k]||_2^2} \quad (28)$$

However, in reality, the number of available μPMUs, $K \ll \frac{B}{2}$. In this case, $\mathbf{H}_u$ is a fat matrix and, in general, is of full

row rank, which in turn means that $(\mathcal{I} - \mathbf{H}_u \mathbf{H}_u^\dagger) = \mathbf{0}$, and our criterion becomes trivial. However, $\mathbf{H}_u \mathbf{H}_u^H$ has generally a high condition number, due to the weak grid connectivity. Therefore, both sides of (26b) can be projected on the sub-space spanned by the left singular vector, denoted by $\mathbf{u}_{u,s}$ corresponding to the smallest singular value of $\mathbf{H}_u$ expecting that $\mathbf{u}_{u,s}^H \mathbf{H}_u \mathbf{d}_u[k]$ to be small. Accordingly, when the quasi steady-state is the regime of operation, i.e., when the equality in (26b) holds, it is expected that the following function to be small and to vary smoothly:

$$x[k] = \frac{|\mathbf{u}_{u,s}^H \mathbf{H}_a \mathbf{d}_a[k]|^2}{||\mathbf{d}_a[k]||^2}. \quad (29)$$

The exit from this behavior is then marked as an anomaly in the central engine and $x[k]$ is the quantity that is proposed to be tracked for fast changes for this purpose.

The flowchart for the central engine is similar to the one for local engine to some extent except that all the μPMU readings are required to form the central metric, and the results from the local analysis are received by the central engine, and the analysis results from this stage are not shipped anywhere, and are ready to be displayed for the operator.

C. Fast Change Detection Method

As explained above, some of the criteria defined in the local and central rules require tracking fast changes in the quantities that are defined, because severe variations in $x[k]$ are signatures of an anomaly. From real data and simulations we have verified that variations in the mean value for these quantities during the quasi steady-state regime are extremely smooth. This observation prompted us to consider changes in their mean value as the common statistical trademark of anomalies in all of these quantities and to use the sequential two-sided Cumulative Sum (CUSUM) algorithm [20], [21] as a heuristic. The use of this algorithm amounts to approximating the samples, for all the aforementioned quantities, as outcomes of a Gaussian non-zero mean process with independent observations. Although the observations are in fact temporally correlated, our objective (i.e. fast change detection in mean) justifies the relaxation that the random process has independent observation samples. The algorithm decides between two hypotheses $\mathcal{H}_0$: *no change in the mean*, or $\mathcal{H}_1$: *change in the mean*, at time k.

A change is detected when the decision function of CUSUM crosses a pre-defined threshold. Therefore, choosing a proper threshold would be essential for practical applications. In this sense, our intrusion detection architecture falls into "semi-supervised behavior-based detection" category, where a collection of real data points containing a variety of anomalies are required to train the detector and find the proper threshold in the learning phase. Several metrics are proposed to evaluate the performance of the CUSUM. The *"average run length"* proposed in [20] and the *"worst case detection delay"* proposed in [22] are the two most popular ones. Having the approximation of a Gaussian distribution, these metrics can be computed accordingly.

We expect to see multiple change points during an event. Detection of multiple change points is achieved by resetting

the decision functions and cumulative sums to zero after the change is detected, and continuing the inspection of upcoming samples. The fast change anomaly is completed if no new changes are detected for a defined window of time.

V. OPTIMAL μPMU PLACEMENT

In tandem with the anomaly detection, the criterion that was described for the central engine can be the basis to determine an optimal placement for the μPMUs. As it was mentioned, the challenge here is that we cannot feasibly nor practically deploy μPMUs at all nodes in the distribution system, therefore we consider a limited deployment, where $K \ll \frac{B}{2}$.

Ideally, we want the matrix $\mathbf{H}_u$ to have a left null-space for the criterion in (29) to be zero in the steady-sate. Considering the high condition number of $\mathbf{H}_u \mathbf{H}_u^{\dagger}$, a reasonable approach to find the optimal configuration is to minimize the norm in (29) over all the possible placement configurations. Let first rewrite the defined metric in (29):

$$x[k] = \frac{\mathbf{d}_a^H[k] \mathbf{H}_a^H \mathbf{u}_{u,s} \mathbf{u}_{u,s}^H \mathbf{H}_a \mathbf{d}_a[k]}{||\mathbf{d}_a||^2}. \quad (30)$$

We desire our formulation to be only topology-dependent and not to be a function of measurements. Therefore, the optimal placement problem is formulated as a min-max optimization with the following structure:

$$
\begin{aligned}
\mathbf{\Pi}^{\text{opt}} = \underset{\mathbf{\Pi}}{\arg\min} \quad & \lambda_{\max}(\mathbf{W}) \\
\text{s.t.} \quad & \left(\mathbf{H}_u \mid \mathbf{H}_a \right) = \mathbf{H} \left(\mathbf{T}_u^T \mid \mathbf{T}_a^T \right) \\
& \mathbf{T} = \mathcal{I}_2 \otimes (\mathbf{\Pi} \otimes \mathcal{I}_3), \\
& \mathbf{W} = \mathbf{H}_a^H \mathbf{u}_{u,s} \mathbf{u}_{u,s}^H \mathbf{H}_a, \quad [\mathbf{\Pi}]_{i,j} \in \{0,1\} \\
& \sum_j [\mathbf{\Pi}]_{i,j} = 1, \quad \sum_i [\mathbf{\Pi}]_{i,j} = 1
\end{aligned} \quad (31)
$$

Since $\lambda_{\max}(\mathbf{W}) = \max_{\mathbf{d}_a[k]} \frac{\mathbf{d}_a^H[k] \mathbf{W} \mathbf{d}_a[k]}{||\mathbf{d}_a[k]||^2}$, essentially we are choosing a placement that minimizes the maximum value that our objective function can take over possible set of available measurement vectors $\mathbf{d}_a[k]$.

An exhaustive search is required to find the global optimum of the optimization problem in (31), which is exponentially complex, and therefore does not scale well. This becomes a barrier when the size of the grid is large, i.e., in most of the real grids. Therefore, we propose to employ a *"Greedy Search"* as an alternative to reduce the time complexity to *polylog*, while accepting to be near-optimal. The pseudo-code of the employed greedy search is illustrated in Algorithm. 1.

VI. NUMERICAL RESULTS

In this section, we first find the optimal placement for our μPMUs based on (31) and then test the effectiveness of the proposed anomaly detection criteria through simulated data and real data, provided from the μPMUs that are installed in our partner utility medium voltage (12.47 kV) grid.

Algorithm 1: Greedy Search Pseudo-Code for Optimal μPMU Placement.

Initialization
  K := Number of μPMUs;
  $\mathcal{P} := \emptyset$, // Set of selected placement locations;
  $\mathcal{L}$:= Set of candidate placement buses;
begin
  for *n=1..K* **do**
    $Cost \leftarrow \inf$;
    for *each* $l \in \mathcal{L}$ **do**
      $\mathcal{P} := \mathcal{P} \cup \{l\}$;
      given $\mathcal{P}$, calculate $\lambda_{\max}(\mathbf{W})$;
      if $\lambda_{\max}(\mathbf{W}) < Cost$ **then**
        $l_{opt} \leftarrow l$;
        $Cost \leftarrow \lambda_{\max}(\mathbf{W})$;
      $\mathcal{P} := \mathcal{P} \setminus \{l\}$;
    $\mathcal{P} := \mathcal{P} \cup \{l_{opt}\}$;
    $\mathcal{L} := \mathcal{L} \setminus \{l_{opt}\}$;

A. Synthetic Data

The IEEE-34 bus test case [23] is simulated using the time-domain simulation environment of DIgSILENT [24] that deals with differential equations rather than memory-less equations. The sampling rate is selected to be equal to the sampling rate of the Analog-to-Digital Converter (ADC) in a real μPMU, which is 512×60 Hz =30720 samples per sec.. We then processed these time-domain data through our phasor estimation algorithm, that emulates a two-cycle, P-class algorithm based on the IEEE C37.118.1 [25] producing phasor samples at a rate of 120 Hz. The single-line diagram of the test case is shown in Fig. 4. This case includes single-phase laterals, voltage regulators, and untransposed lines, which all are modeled exactly in our admittance matrix.
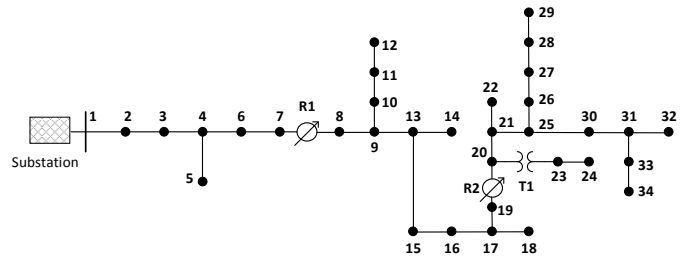

Fig. 4. IEEE 34-Bus Test Feeder Single-Line Diagram

Table. II 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 close to each other, while the run time of the *"Greedy Search"* is 102.206 times faster. Hence, the *"Greedy Search"* can be a very good choice to solve our optimal placement problem.

As expected, the placement rule tries to scatter the available μPMUs all over the grid, in order to achieve the maximum

TABLE II
IEEE-34 Case Optimal μPMU Placement Result for $K = 3$

	Random	Greedy	Exhaustive
Optimum Cost	1.7085	0.51477	0.51477
Buses with μPMUs	{1,3,9}	{7,19,31}	{9,19,31}
Run Time	–	2.84 s	290.266 s

possible coverage.

In order to investigate how the time complexity grows, and also analyze the results of the placement criterion, the 123 standard test case in [23] was used, considering 20 μPMUs available (i.e., $K = 20$). Without loss of generality, the 123 test feeder was reduced to 70 buses to only include three-phase lines, and roll up all the laterals (single-phase and two-phase lines). The reasoning behind this reduction is that visibility on the main feeder is more important for us than the the visibility on the laterals, considering the limited number of μPMUs. Using a machine with 60 Intel(R) Xeon(R) CPU E7-4870 v2 @ 2.30GHz cores, it took 11.56 Sec for the algorithm to place 20 μPMUs over 70 buses. Fig. 5 illustrates the location of the optimally-placed μPMUs in the reduced grid. It can be observed from the figure that the μPMUs are scattered over the grid to achieve the maximum sensitivity with respect to different locations of anomalies.
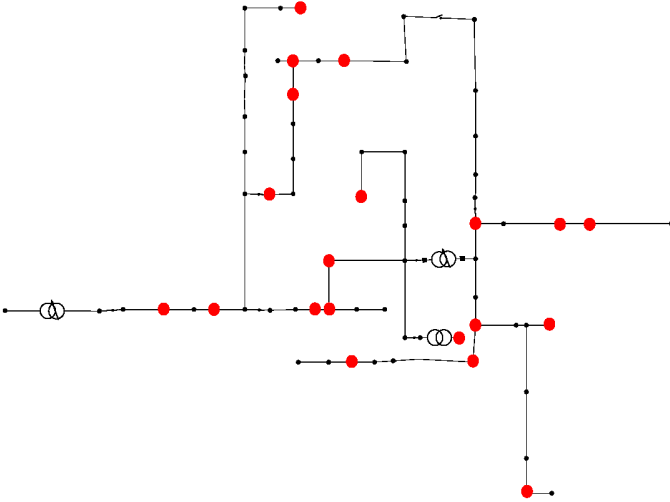
Fig. 5. Optimally-Placed μPMUs on Reduced IEEE 123 Test Feeder.

1) Single-Line to Ground Fault-IEEE 34 Bus: In this simulation, the rules are tested for detecting 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. A SLGF was introduced on "Phase a" of line (25,26), which then caused the fuse placed on the phase a of this line near bus 25 to melt down. Our three μPMUs are placed on buses, 7, 19, and 31 based on the "greedy search" result for the optimal placement criterion.

The results of the *"voltage magnitude change"* rule are shown in Fig. 6 for μPMUs 7, and 19 for instance. Simply, the rule inspects the data for large deviations, label them

accordingly, and find the start time and the end time of the event, marked with blue and red stars.
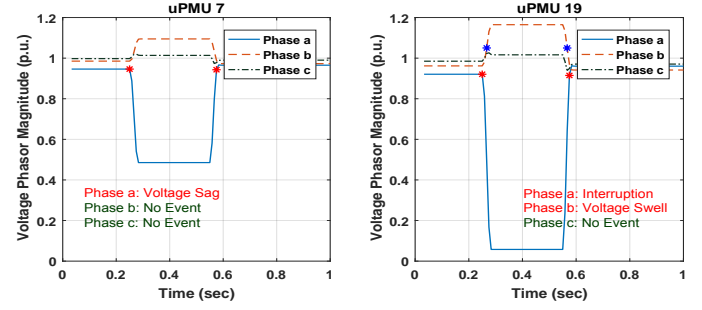

Fig. 6. Voltage Magnitude Local Rule Result for SLGF.

Fig. 7 illustrates the metric value in (22) for the specified lines that is inspected with $M = 12$ to check whether the grid is in the quasi steady-state or not, where the voltage and current data are first converted to per-unit system assuming $S_b = 1$ MVA. The start time of the detected changes are also marked, setting the CUSUM detector parameters fixed in all the three local engines corresponding to each μPMU. As it 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 large 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
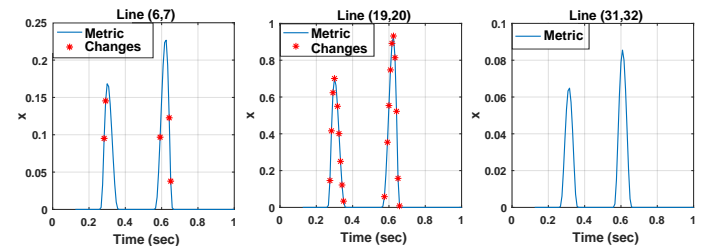

Fig. 7. Quasi Steady-State Validity Checking for SLGF.

with different severity and behavior (we just show some of the results here for lack of space). 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 these time-tags for eventful segments of data helps understanding their relationship.

The metric defined for our central engine is also illustrated in Fig. 8. The delay in appearing and disappearing of the event in here is solely due to the two-cycle phasor calculation.
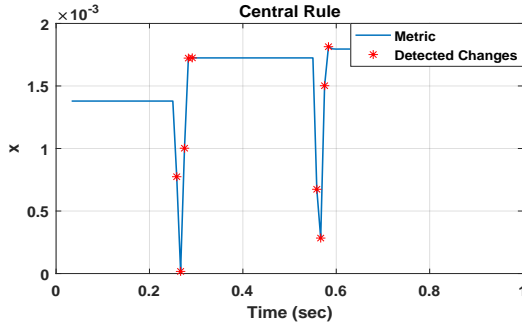

Fig. 8. Central Rule Inspection for SLGF.

Fig. 9. Voltage Magnitude and Quasi Steady-State Validity Check Rule for Feeder Reconfiguration.

It should be noted that the value of the metric in (29) during the quasi steady-state is highly dependent on the number of μPMUs and the topology of the grid that together determine the structure of $\mathbf{H}_u$ and $\mathbf{H}_a$. However, the best placement of the μPMUs for a certain topology and number of μPMUs that makes the metric to be as close as possible to zero, returns a certain objective value that can be used as the baseline to determine what is a normal value for $x[k]$ and what would be an anomaly.

2) Feeder Reconfiguration-IEEE 34 Bus: In this case, it is assumed that there is a normally-open line that connects node 13 to 21. This line is closed as part of a hypothetical feeder reconfiguration scheme when the line 13-15 is opened. We run the simulation for 1.5 sec, and the feeder reconfiguration happens at t = 0.75 sec. Fig. 9 shows the voltage phasor magnitude and the metric related to the quasi steady-state validity check at the local level for the reconfiguration event. As it can be seen the event does not trigger any anomaly in the voltage phasor magnitude. For example, if we wanted to solely rely on the raw phasor data coming from μPMU 7 and not checking other metrics, this event would have been missed because the current phasors measured by this μPMU also do not see significant change, mainly because the downstream load of bus 7 has not changed. However, checking the validity of quasi steady-state for line 6-7 using the measured phasor data by μPMU 7 clearly reveals the existence of an anomaly in the system.

3) Discussion–Compromised Data: We illustrate here the resilience of the architecture to data injection attacks. For this purpose, three data attack scenarios are investigated happening concurrently with the SLGF event discussed previously. We consider the case of the attacker manipulating the data of μPMU 7 in the first case, and μPMU 19 in the second case, and finally μPMU 7 and 19 at the same time in the third case on their way to the central engine. 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 7 compromised will not affect our central rule significantly (case-1). However, when the μPMU 19 is compromised, the
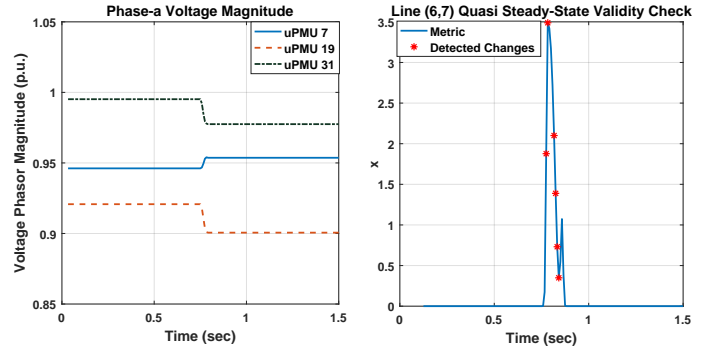
number of detected changes reduce significantly (case-2), and when both μPMU 7 and 19 are compromised and the only healthy data is coming from μPMU 31, the detector does not pick any fast changes based on the set parameters. This also reveals the importance of tuning the detector thresholds to have a certain "false alarm," while maximizing the "detection" probability. We wish to remark that in all the three cases, the
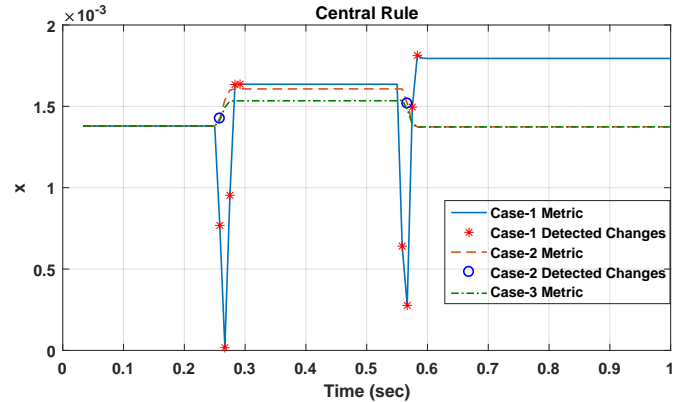

Fig. 10. Central Rule for SLGF with Manipulated μPMU Data.

local analytics that directly draw data from μPMUs will still flag the alarm, so buffering locally at the site of the event these data can be an important way of helping understand what communications were compromised in an ex-post analysis.

4) Discussion–Optimal vs Non-Optimal Placement: As it was mentioned, the placement criterion tries to scatter the available μPMUs over the grid to achieve the maximum coverage, and therefore make the central rule more sensitive to anomalies.

In order to compare the performance of an optimal versus non-optimal placement, a load loss event is created on bus 24 of IEEE-34 test case at $t = 0.4$s. The μPMUs are placed based on the random placement and greedy search result given in Table. II, corresponding to non-optimal and optimal placement, accordingly. Fig. 11 shows the central metric for these two cases. Since the relative change is what matters to our detector, the metrics corresponding to two placement are brought on the same scale. As it can be observed, since the μPMUs in the random placement are concentrated at a certain

area, events like this would not be very pronounceable in the central metric, which could possibly lead to a "false-negative".
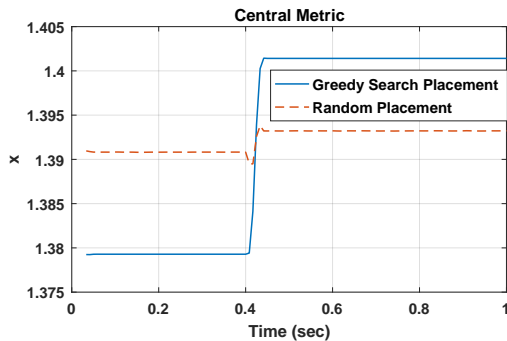

Fig. 11. Central Metric Change of a Load Loss Event; Optimal versus Non-optimal Placement.

B. Real Data

Fig. 12 shows the abstract one-line diagram of the partner utility grid and the location of the installed μPMUs. The installed μPMUs sample the voltage and current with a rate of 512×60 Hz, and output the estimated phasors at 120 Hz rate. These devices achieve an accuracy of 0.001 deg resolution for phasor angle, 0.0002% for phasor magnitude, and 0.01% for Total Vector of Error (TVE) [26]. The two feeders here are connected through the subtransmission grid.
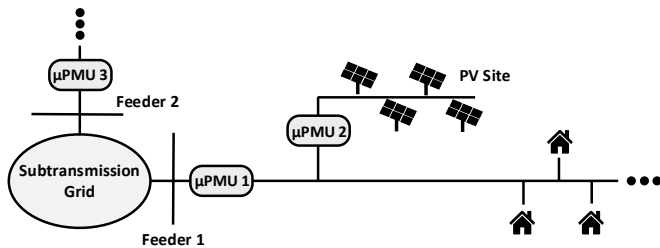

Fig. 12. Location of Installed μPMUs in Our Partner Utility Grid.

Fig. 13 shows the voltage magnitude change rule inspected on the data from these μPMUs over a certain time. 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 also illustrated in Fig. 14(a) and Fig. 14(b), respectively. Observing the results of analysis, the 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 due to the space limit but other rules can also flag the existence of anomaly on feeder 2.

All the metrics introduced and tested above are designed considering the specifications of a distribution grid. It should be noted that not all the proposed methods in the literature for transmission grid are applicable in the distribution side. For example, the phase angle difference in the distribution grid is
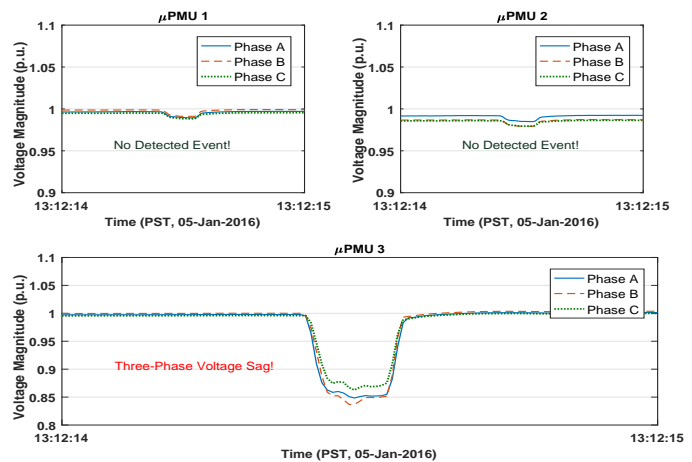

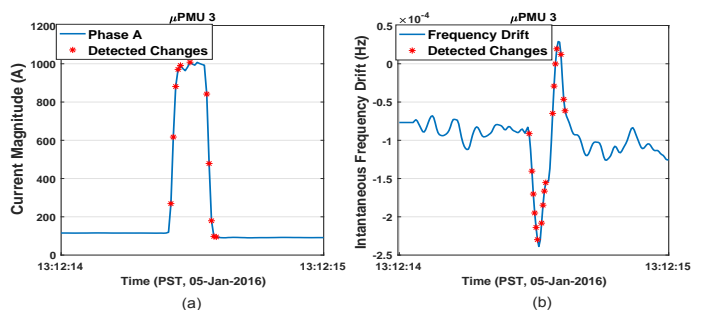Fig. 13. Voltage Magnitude Change Rule for Real Data.

Fig. 14. Fast Change Tracking of Current Phasor Magnitude and Bus Instantaneous Frequency Drift for Real Data.

known to be much smaller than that in the transmission level. Therefore, as opposed to transmission grid that this metric would work well for event detection [12], it might not be a proper metric to look at in the distribution grid, since the signal to noise ratio might be small. The example next illustrates how the voltage angle difference metric could have failed if it was used as a "local rule". Using the data from two μPMUs installed at two ends of a line at a second utility grid (not the one in Fig. 12), the voltage magnitude captured by the μPMU at one end of the line is shown in Fig. 15(a) and the voltage phasor angle difference between the two μPMUs at two ends of the line is shown in Fig. 15(b). As it can be observed, the important event in this period corresponds to the two voltage sags. However, the angle difference shows a significant number of spikes without clearly marking these two events with the same significance. Passing this metric to our detector, we would pick too many fast changes that do not represent any specific event of interest, which therefore means an increase in the number of "false positives". All streams of interest were examined during this period, including both the active and reactive power, and all anomalies detected were in agreement with those visible in Fig. 15(a).

1) Case Study-Robustness Against Volatility: The increasing presence of renewable resources, like wind and solar, in the distribution grid may raise some concerns about the robustness of the proposed rules to "false alarms" due to their inherent volatility. However, our method is based on an
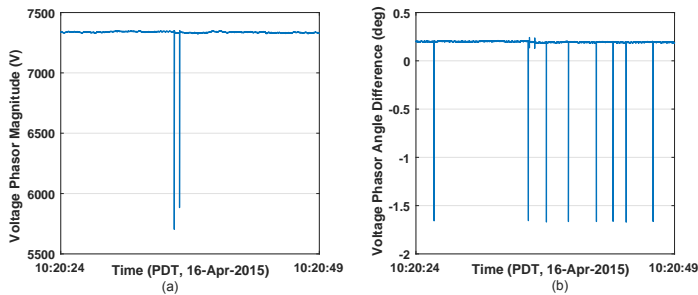
Fig. 15. Voltage Phasor Angle Difference Between Two μPMUs at Distribution Grid.

Fig. 17. Fast Change Tracking of Current Magnitude for a Non-Linear Load.

adaptive estimation of the data mean using an exponential window and, as shown in the following numerical example, this makes the algorithm capable to only pick the fast changes over a very short period of time, while being insensitive to changes in the time order of normal renewable resources fluctuations. To show it numerically, we refer to two real events captured by μPMU 2 in the grid in Fig.12. Fig. 16(a) shows a dramatic change of about -18.25% over 9 seconds in the PV site active power injection due to the cloud effects, and Fig. 16(b) shows a step change in the active power after the PV site went out of service. Setting the detection parameters to be the same for both events, it can be seen that the detector flags no event for the first case and finds multiple of fast changes in the second. This is ideal since the events of our interest in the this context are of the second type, which should be discriminated from the normal quasi steady-state behavior in the first case.
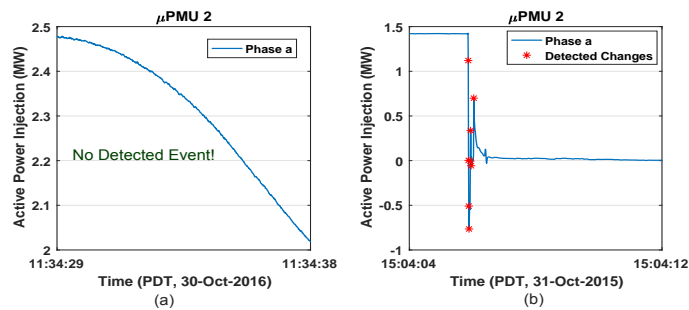

Fig. 16. Fast Change Tracking of Injected Active Power in the PV Site.

Considering the existence of different types of loads in the distribution grid, the detector should also be able to differentiate between possible normal fluctuations in the load profile and those that are caused by a rare event. To show the performance of CUSUM for this case, real data of a μPMU installed behind a building with a non-linear load is used. As it can be seen in Fig. 17, the fast change tracker of the current magnitude does not pick the fluctuations due to the non-linear load behavior as an event and only flags a segment of data that corresponds to a voltage sag in the grid as an eventful segment.

VII. CONCLUSION AND FUTURE DIRECTION

In this paper, a hierarchical *"anomaly detection architecture"* has been described that integrates high resolution, syn-
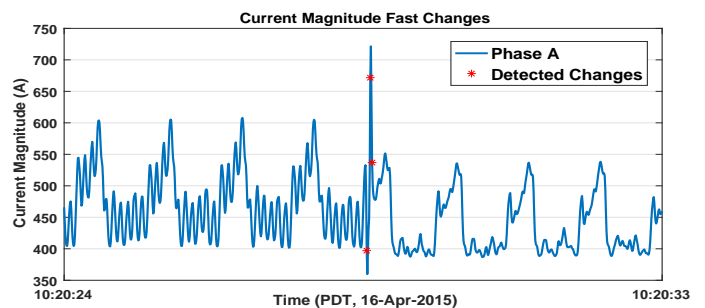
chronous μPMU data to attain high accuracy in the detection of the physical effects of an event. A set of rules has been formulated to process the μPMUs' data at the local and central level of our hierarchical architecture. Exiting from the quasi steady-state regime is a common signature of many anomalies, and this insight is exploited to design a key part of the proposed metrics. To ensure scalability of our architecture, the rules at the local stage are designed to be agnostic about the grid interconnection and physical location of the sensor. Depending on the type of event, the effects on different metrics has different severity and, therefore, inspecting different metrics at the local level reduces the miss-detection rate. The analysis in the higher stages is designed to bind the readings from different μPMUs using the knowledge about the grid topology and sensor locations in search for eventful segments.

Using real data and simulations, we have shown how processing real-time physical measurements from μPMUs enables us to detect effectively anomalies in the system and inform the operator about the grid status. Having a limited number of μPMUs, an optimal placement criterion is also designed with respect to the central rule that aims to scatter μPMUs over the grid to achieve maximum coverage and therefore to increase the sensitivity of the rule to different event source locations.

Having an anomaly identified, our future efforts will focus on identifying and localizing the source of the anomaly using the underlying physical model when possible or, in general, at least confine the search region to a smaller set of components/lines. This include the identification of any change in the admittance matrix, due to loss of a line, reconfiguration of the grid, and etc.

REFERENCES

[1] M. Jamei, A. Scaglione, C. Roberts, E. Stewart, S. Peisert, C. McParland, and A. McEachern, "Automated anomaly detection in distribution grids using upmu measurements," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.

[2] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Microsynchrophasors for distribution systems," in *Proc. IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, 2014, pp. 1–5.

[3] J. H. Eto, E. M. Stewart, T. Smith, M. Buckner, H. Kirkham, F. Tuffner, and D. Schoenwald, *Scoping Study on Research and Priorities for Distribution-System Phasor Measurement Units*. Lawrence Berkeley National Laboratory, 2015.

[4] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid*, Aug 2013, pp. 1–9.

[5] A. Phadke and R. M. de Moraes, "The wide world of wide-area measurement," *Power and Energy Magazine, IEEE*, vol. 6, no. 5, pp. 52–65, 2008.

[6] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 80–93, 2011.

[7] S. Pan, T. Morris, and U. Adhikari, "Developing a hybrid intrusion detection system using data mining for power systems," *Smart Grid, IEEE Transactions on*, vol. 6, no. 6, pp. 3104–3113, 2015.

[8] Y. Chen, L. Xie, and P. R. Kumar, "Dimensionality reduction and early event detection using online synchrophasor data," in *Power and Energy Society General Meeting (PES), 2013 IEEE*. IEEE, 2013, pp. 1–5.

[9] L. Xie, Y. Chen, and P. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784–2794, 2014.

[10] J. Valenzuela, J. Wang, and N. Bissinger, "Real-time intrusion detection in power system operations," *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1052–1062, 2013.

[11] Y. Ge, A. J. Flueck, D.-K. Kim, J.-B. Ahn, J.-D. Lee, and D.-Y. Kwon, "Power system real-time event detection and associated data archival reduction based on synchrophasors," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 2088–2097, 2015.

[12] A. Allen, M. Singh, E. Muljadi, and S. Santoso, *PMU Data Event Detection: A User Guide for Power Engineers*. National Renewable Energy Laboratory, 2014.

[13] M. Biswal, S. M. Brahma, and H. Cao, "Supervisory protection and automated event diagnosis using pmu data," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1855–1863, Aug 2016.

[14] 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, 2016.

[15] S. Brahma, R. Kavasseri, H. Cao, N. R. Chaudhuri, T. Alexopoulos, and Y. Cui, "Real time identification of dynamic events in power systems using pmu data, and potential applications - models, promises, and challenges," *IEEE Transactions on Power Delivery*, vol. PP, no. 99, pp. 1–1, 2016.

[16] A. G. Phadke and J. S. Thorp, *Synchronized phasor measurements and their applications*. Springer Science & Business Media, 2008.

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

[18] Y. Xia and D. P. Mandic, "Widely linear adaptive frequency estimation of unbalanced three-phase power systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 1, pp. 74–83, 2012.

[19] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2012.

[20] E. Page, "Continuous inspection schemes," *Biometrika*, pp. 100–115, 1954.

[21] M. Basseville, I. V. Nikiforov *et al.*, *Detection of abrupt changes: theory and application*. Prentice Hall Englewood Cliffs, 1993, vol. 104.

[22] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.

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

[24] D. P. F. Manual and D. PowerFactory, "Version 14.0," *DIgSILENT GmbH, Gomaringen, Germany*, 2009.

[25] "IEEE Standard for Synchrophasor Measurements for Power Systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, Dec 2011.

[26] "Micro-phasor measurement unit brochure," http://www.powersensorsltd.com/microPMU-brochure.php, accessed: 2016-01-15.