A Sensitivity Study of PMU-Based Fault Detection on Smart Grid

Richard Barella¹, Duc Nguyen¹, Ryan Winter¹, Kuei-Ti Lu¹, Scott Wallace¹, Xinghui Zhao¹, and Eduardo Cotilla-Sanchez²

¹School of Engineering and Computer Science, Washington State University, 14204 NE Salmon Creek Ave., Vancouver, WA 98686
²School of Electrical Engineering and Computer Science, Oregon State University, 3023 Kelley Engineering Center, Corvallis, OR 97331

{richard.t.barella, duc.nguyen, kuei-ti.lu, ryan.j.winter, wallaces, x.zhao}@wsu.edu, ecs@eecs.oregonstate.edu

Keywords: Smart Grid, Sensitivity, Fault Detection, Accuracy, PMU Deployment

Abstract: Phasor measurement units (PMUs) are widely used in power transmission systems to provide synchronized measurements for the purpose of fault detection. However, how to efficiently deploy those devices across a power grid – so that a comprehensive coverage can be provided at a relatively low cost – remains a challenge. In this paper, we present a sensitivity study of a PMU-based fault detection method using three different distance metrics. This study can serve as a guideline for efficient PMU deployment. To illustrate the effectiveness of this approach, we have derived an alternative PMU placement plan for a power grid. Experimental results show that our PMU placement reduces the required PMU deployment by more than 80% as compared to the original placement, yet still provides similar level of accuracy in fault detection.

1 INTRODUCTION

Phasor measurement units (PMUs), or synchrophasors, are devices that are deployed in power systems to measure phase angles and magnitudes of the electrical waves in real time, for monitoring the health of the power grid. A significant amount of work has been done in analyzing real-time PMU data for detecting faults (Jiang et al., 2000) (Liang et al., 2014), oscillations (Liu and Venkatasubramanian, 2008), as well as tracking fault locations (Chang et al., 2008). However, these approaches assume a comprehensive coverage of PMUs on the power grid. Due to the installation cost, instrumenting every bus with PMUs is not always practical. Therefore, it is critical to efficiently deploy a limited number of PMUs so that comprehensive coverage in terms of fault detection can be provided. Instead of developing fault detection methods based on a known topology of PMUs on a power grid, we take a different approach in which we use an existing fault detection algorithm as a guideline to derive more efficient PMU placement plans.

As the first step toward better deployment of PMUs, it is essential to quantitively analyze how the distance between a fault and the PMU(s) used to detect that fault impact the detection process itself. To this end, we have carried out a sensitivity study of distances in PMU-based fault detection. Specifically, we use a PMU-based fault detection method that we previously developed as a baseline, and investigate the accuracy of this method with respect to the distance between the fault location and the PMUs being utilized. Here, three distance metrics are studied: topological distance, logical distance, and electrical distance. Topological distance is derived from the system schema of the grid, i.e., number of hops between two sites. Logical distance is the Pearson correlation coefficient which is derived from two PMU data streams. Electrical distance is derived from the Ybus of the power grid, representing another way to elucidate the electrical structure of a power grid (Hines et al., 2010).

The results of the sensitivity study indicate that there is a potential to accurately detect faults even when no PMU is placed on a bus immediately adjacent to the fault. Rather, fault detection remains accurate within a small neighborhood near the fault, and then accuracy falls off as distance from the fault increases. This relationship creates an opportunity for efficient PMU deployment. Specifically, we use the results of the sensitivity study as a guideline and develop a PMU placement algorithm which derives deployment solutions based on distance constraints. We verified our algorithm using PMU data collected from a power grid; experimental results show that our PMU
placement plan uses 80% fewer PMUs compared to
the original placement; however, it provides similar
level of accuracy in detecting faults.

The remainder of the paper is organized as fol-
lows. Section 2 reviews related work in both PMU-
based fault detection and PMU placement in a smart
grid. Section 3 introduces the background of this
work, including the fault detection method we use
in this study, as well as a description of the dataset
we use. The experimental results from the sensitivity
study we carried out on three different distances are
presented in Section 4. In Section 5, we develop a
new PMU deployment algorithm which derives possi-
ble PMU placement solutions based on the results
from the sensitivity study. Section 6 concludes the
paper and proposes future directions for this work.

2 RELATED WORK

With the growing popularity of using phasor mea-
surement units (PMUs) to monitor power systems and
enhance their reliability, there is increasing interest
in analyzing real-time PMU data to detect and loca-
tate faults in the power grid. A significant amount
of work has been done to detect or monitor certain
conditions of a power grid by leveraging informa-
tion extracted from PMU data. Jiang et al. propose
an online approach for fault detection and localiza-
tion using SDFT (smart DFT) (Jiang et al., 2000).
Liu et al. use Frequency Domain Decomposition for
detecting oscillations (Liu and Venkatasubramanian,
2008). Kazemi et al. propose a multivariable re-
gression model to track fault locations using PMU
data (Chang et al., 2008). A more comprehensive sur-
vey can be found in (Glavic and Van Cutsem, 2011).
These approaches assume a comprehensive PMU de-
ployment across the smart grid. Most recently, with
the emergence of big data analytics, a variety of ma-
chine learning techniques have been applied to an-
alyze PMU data in power grid systems, including
classification (Alsafasfeh, 2010), clustering (Antoine
and Maun, 2012), artificial neural networks (Mishra
et al., 2008), Support Vector Machines (Gomez et al.,
2011), and regression trees (Zheng et al., 2013).

Along with the work of monitoring the power grid
using PMUs, the challenge of optimizing PMU place-
ment has also attracted much attention. This is be-
cause deploying PMUs is expensive and a per-bus
coverage of PMU deployment is not always practi-
cal (Mili et al., 1990). Traditional approaches formu-
late PMU deployment as an optimization problem, in
which the power grid is modeled as a graph, and the
objective is to deploy PMUs at a minimum number of
nodes so that the state of the whole power grid is ob-
servable (Anderson and Chakrabortty, 2012a) (An-
derson and Chakrabortty, 2012b) (Brueni, 1993) (Hay-
nes et al., 2002). This problem has been proven to be
NP-complete. It has also been proven that no more
than 1/3 of the nodes in a connected graph of at least
3 nodes are requirement to be equipped with PMUs
in order to provide coverage for the whole power
grid (Brueni and Heath, 2005).

Besides these graph-theoretic approaches, simu-
lation based methods have also been used in devel-
oping optimal placement for PMUs. For instance,
in (Liu et al., 2012), a generic algorithm is proposed to
find optimal deployment for heterogeneous measure-
ment devices, including both PMUs and Smart Meter-
ings systems. The results are tested using simulation.
Similar simulation approaches are used in (Zhu et al.,
2009) to evaluate a PMU placement method which
aims for improving the accuracy of state estimation
of the grid. In (Lien et al., 2006), a concept of fault-
location observability is proposed. PMUs are placed
on buses based on the one-bus spaced deployment
strategy, and then the results are tested using simu-
lation. In (Pegoraro et al., 2012), generatation from
renewable sources are considered in the deployment
of PMU and smart metering.

In this paper we propose a novel approach to de-
rive an anytime-optimal PMU placement plan guided
by a sensitivity study of a PMU-based fault detection
method. Our work differs from the previous work in
the following aspects. First, instead of proposing a
theoretical deployment algorithm, we take a practical
approach by developing a fault detection algorithm
first, and then carrying out a sensitivity study which
serves as the guideline for PMU deployment. Sec-
ond, instead of using simulation, we used real PMU
data collected over one-year period on a smart grid
in Pacific NorthWest region of the United States of Amer-
ica.

3 BACKGROUND

In order to derive efficient PMU deployment plans
for a smart grid, we first develop a simple fault de-
tection method, and verify this method using a year
worth of real PMU data collected from a smart grid in
the Pacific NorthWest of the United States of Amer-
ica. The dataset we used in this study, as well as the
fault detection method are presented in the following
sections.
3.1 Dataset

The dataset we use in this research is from Bonneville Power Administration, the first utility agency that implements a comprehensive adoption of synchrophasors in their wide-area monitoring system. The smart grid is located in the Pacific Northwest area in the United States, and it contains both 500KV and 230KV buses. In this grid, there are 31 sites which are equipped with PMUs to measure voltage, current, and frequency data. The dataset we use in this paper is collected from October 17, 2012 to September 16, 2013. During this time period, there are 107 documented faults, including single-line-to-ground faults, line-to-line faults, and three-phase faults.

3.2 Fault Detection Method

The fault detection method used in this study was developed based on a theoretical analysis on the characteristics of faults, as well as the BPA datasets (Liang et al., 2014). The algorithm is a threshold based decision tree, which classifies faults into the three fault types – single-line-to-ground, line-to-line, and three-phase – using the voltage sag values on all three phases. This fault detection method classifies faults based on a set of pre-defined threshold values on voltage sags, which are calculated by surveying the dataset for the voltage sag values during the same type of faults. The voltage sag values are in p.u., normalized with respect to the voltage level at the steady state of that phase. The steady state voltage magnitude is calculated as follows: 1) use a sliding window to scan a period before the fault occurs; 2) calculate the median of voltage magnitude within the window; 3) if the fluctuation of the voltage magnitude in the window is within a small range, we consider the median of voltage steady state; otherwise, we keep searching by moving the sliding window forward, until we find a steady state.

It has been shown that the accuracy of this method is more than 96% (Liang et al., 2014). Further, it is worth noting that we develop this simple fault detection technique to serve as a baseline for our sensitivity study, but the approach presented in this paper is not limited to this fault detection method. Any other approaches for fault detection could also be used for this purpose.

4 SENSITIVITY STUDY

When a fault occurs on a power grid, the signature is typically visible at nearby locations although the signature is typically reduced in magnitude. This fact makes it possible to optimize the PMU placement by removing PMUs which provide redundant coverage. To fully understand the impact of distances on smart grid fault detection, we have carried out a sensitivity study on the fault detection method described in Section 3. Specifically, we have investigated the accuracy of the method when fault signatures are observed by PMUs at different locations on the power grid.

4.1 Distance Metrics

A smart grid consists of a large number of interconnected sites. In order to efficiently deploy PMUs across the grid, we must first analyze the impact of various types of faults on the entire grid. Figure 1 shows an example grid which includes 5 sites. Suppose at certain point in time, a single-line-to-ground fault (SLG) occurs at Site 1, the impact of this fault is usually observable from other locations of the grid. As shown in Figure 1, this impact can be captured by our fault detection method $FD(t)$ which is executed on various sites. Here in this example, the fault type (SLG) is successfully detected at sites 1, 2, and 3. However, as indicated in Figure 1, Site 4 and 5 have failed to detect the fault, simply because they are further away. In this specific example, if a PMU is deployed on any of the sites 1, 2, or 3, this particular single-line-to-ground fault can be detected.

As the first step toward efficient deployment of PMUs across a power grid, we have carried out a sensitivity study of our fault detection method with respect to the distances to the faulted locations (sites) using real PMU data gathered from BPA's smart grid. Note that the distance between two sites can be repre-
sent  in different ways. In our work, we have investigated three different distance metrics: topological distance, logical distance, and electrical distance.

Topological distance, or hop distance, is a distance metric for estimating geographical distances. To calculate the topological distance between two sites, we represent the grid as a graph with interconnected nodes. Each node is a site with a functioning PMU and edges between nodes are transmission lines that are also monitored by one or more PMUs. We then use Dijkstra’s algorithm (Dijkstra, 1959) to derive the shortest path between any two sites, and construct a distance matrix for the grid. Note that the topological distance matrix is static for a given smart grid, because it is derived from the topology of the grid.

Logical distance is a dynamic distance metric representing the linear correlation between two data streams. In this work, we use Pearson correlation coefficient (PCC) as the metric for logical distance. The PCC of two data streams $X(x_1, x_2, ..., x_n)$ and $Y(y_1, y_2, ..., y_n)$ can be calculated as follows:

$$PCC = \frac{\sum_{i=1}^{n} (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{Y})^2}}$$

(1)

The value of PCC ranges from $-1$ to $1$, representing the linear relationship between the two PMU data streams: $0$ indicates no linear relationship, $1$ and $-1$ indicate linear relationships, in positive or negative direction.

Electrical distance is an electrical cohesiveness metric, as proposed in (Cotilla-Sanchez et al., 2012). The electrical distance between buses $i$ and $j$ is obtained from the quadrant of the ac power flow Jacobian that measures the incremental change in voltage phase angle difference between $i$ and $j$ for an incremental active power transaction between $i$ and $j$. This power flow Jacobian is itself computed from a combination of the $Y_{bus}$ (nodal admittance matrix), and generation and load information. In this particular set of experiments we build a ‘nominal’ power flow Jacobian by assuming that power injections are small increments, whereby the Jacobian is basically inherited from the $Y_{bus}$ structure (this is analogous to a ‘flat start’ before solving the power flow problem).

The above three metrics represent distances between sites in different ways. However, they are also related. For example, the topological distance, i.e.,

The above three metrics represent distances between sites in different ways. However, they are also related. For example, the topological distance, i.e.,

hop distance, has been shown to be a good indicator for electrical distance. Figure 2 shows a box-plot which depicts the relationship between these two types of distances derived from our dataset. As shown in Figure 2, the electrical distance increases when the topological distance increases. Note that a topological distance of $-1$ indicates that the two sites are not connected by a path that is monitored by PMUs. Those sites generally have higher electrical distances too, as shown in the figure.

4.2 Fault Detection Sensitivity

With the three distance matrices being calculated, we can then analyze the accuracy of our fault detection algorithm presented in Section 3 across the whole smart grid, with respect to the distance to the faulted site.

4.2.1 Topological Distance

For each documented fault in our dataset, we first execute our fault detection algorithm on every PMU site, using the data being recorded during the time of the fault, then we compare the results with the ground truth (the recorded fault type), to determine whether the fault is correctly detected on each site. We then associate each result with the topological distance between the site where the data is collected and the faulted site. After all the faults have been analyzed, we calculate the accuracy of our fault detection method on a certain topological distance.

Figure 3 shows the histogram of the accuracy of the fault detection method at various distances away from the location of the fault. There is a clear correlation between the accuracy of the fault detection method and the topological distance from the fault location to the PMU where the data is collected. Specifically, if the PMU is located within 2 hops from
the fault location, the accuracy of our fault detection method is above 80% for all 107 recorded faults in the dataset. The accuracy decreases as topological distance increases.

4.2.2 Logical Distance

We represent logical distance between two sites using Pearson correlation coefficient (PCC) as shown in Equation 1. This distance metric models the similarity of two data streams within a certain period of time, therefore it is a dynamic metric which changes over time. In our study, we calculate logical distance using a 15-second time window preceding a fault, and analyze the accuracy of the fault detection method with respect to this distance metric. Specifically, for each fault, we first calculate the logical distance between the fault location and any other PMU site using the 15-second time window before the fault, then we execute the fault detection method using the data collected at each of the non-fault site, and associate the accuracy with the logical distance.

The results of the sensitivity analysis with respect to the logical distance are shown in Figure 4. In general, the fault detection methods have higher accuracy when it is executed on a PMU site which has higher correlation with the fault location.

4.2.3 Electrical Distance

Similarly, we have carried out a sensitivity study on electrical distance, which is a static distance metric derived from the schema of the smart grid, as described in Section 4.1. The results are shown in Figure 5. As expected, the accuracy of our fault detection method decreases when the electrical distance to the fault location increases. However, the accuracy is close to 100% when the electrical distance is within 0.010.

To further investigate the electrical distance, we have calculated the number of added signals when we gradually increase the electrical distance boundary. Figure 6 shows the results. Green indicates the signals on which the fault detection method can accurately detect faults, while blue bars indicate the total set of signals for within a specific electrical distance from the fault’s source. When the electrical distance is within 0.010, the fault detection method is accurate for most of the signals. When the boundary increases, more signals result in incorrect detection results.

The sensitivity study presented in this section indicates that it is not necessary to deploy PMUs at every site in a smart grid. This is simply because most of the faults can be detected from a site which is within a certain distance from the fault location.

5 PMU DEPLOYMENT

Based on the sensitivity study, we can derive an improved PMU deployment plan for the smart grid.
Here we use the topological distance as an example to illustrate this approach.

The sensitivity study suggests that the accuracy of our fault detection method is reasonably good when the PMU is located within two hops from the fault. Based on this observation, we have developed a PMU placement algorithm which deploys PMUs in the way that each site in the smart grid is less than or equal to \( n \) hops away from a PMU. Algorithm 1 shows how a PMU placement plan is derived.

The algorithm randomly picks a starting site to place a PMU, then traverses the grid while iteratively placing PMUs on the sites as needed. Depending on the starting site, and choices made during the course of execution, the algorithm may return different solutions. Figure 7 shows an example placement solution which uses only 6 PMUs (highlighted in gray in the figure), as opposed to 31 in the original smart grid schema. The placement is derived with a topological distance constraint of 3, which means that any site on the smart grid is less than or equal to 3 hops away from a site with a PMU.

To evaluate the PMU placement plan shown in Figure 7, we have simulated a smart grid with these 6 PMUs, and execute our fault detection method using data from these 6 PMUs. For the 107 recorded faults, the accuracy of the fault detection on the 6-PMU smart grid is 93.9%, only 2.1% less than the accuracy of the original smart grid with 31 PMU-equipped sites (96%). These results illustrate that it is possible to accurately detect faults using significantly less PMUs, creating the opportunities of cost savings for PMU deployment.

Since our PMU placement approach shown in Algorithm 1 can generate a number of different solutions for one distance constraint, to illustrate the accuracy of the new PMU placement plans generated by this algorithm, we randomly picked some placement solutions generated by our PMU deployment algorithm using various of distance constraints, and calculated their accuracy. Figure 8 shows a comparison of the accuracy of these solutions with respect to their distance constraints. The number of PMUs of each solution is also noted in the figure. As shown in the figure, it is possible to achieve more than 90% of accuracy in fault detection with only 6 PMUs. As expected, when the distance constraint increases, the accuracy decreases, so does the number of required PMUs.

Based on a recent cost analysis report by Department of Energy (Department Of Energy, 2014), the average cost per PMU ranges from $40k to $180k, which includes cost for procurement, installation and commis-

---

ALGORITHM 1: PMU Placement

```text
randomly pick a site \( S_0 \);
place PMU at \( S_0 \);
\( S_0.distance = 0 \);
Queue.push(\( S_0 \));

while \( Queue \) is not empty do
    nextSite = Queue.pop();
currentDistance = nextSite.distance;
    if \( nextSite.distance \leq 2 \) then
        Visited.append(nextSite);
        for neighbor in nextSite’s neighborset
            if neighbor is not visited
                neighbor.distance = nextSite.distance + 1;
                Queue.push(neighbor);
        end
    else
        place PMU at nextSite;
        nextSite.distance = 0;
        Queue.push(nextSite);
    end
end

if Queue is empty && not all sites are visited then
    randomly pick an unvisited site;
    Place PMU at newSite;
    newSite.distance = 0;
    Queue.push(newSite);
end
```

---

2 The actual site names are not displayed in Figure 2 due to security reasons. Figure 2 shows 36 sites because we included 5 sites which do not have PMUs installed in the original grid, in order to maintain the original topology. Note that the new PMU deployment plan we generated only places PMUs on the sites that are PMU-equipped in the original power grid.
sioning. Therefore, our work can potentially result in significant cost savings for a smart grid.

![Figure 7: An Example PMU Placement Plan (PMU sites are highlighted in gray)](image)

### 6 CONCLUSION

Synchrophasors are widely used in power grids to enhance situation awareness, and robustness of power delivery. A significant amount of work has been done in fault detection using PMU data. However, these approaches assume a predefined PMU deployment scheme across the smart grid. Since deploying PMUs is costly, it is not necessarily practical or scalable to equip every bus with a PMU. A more cost-effective solution for PMU deployment is needed.

While traditional approaches usually simulate the power grid and derive efficient PMU deployment plans, in this paper, we take a novel approach to investigate ways to efficiently deploy PMUs across a power grid based on the accuracy of detecting faults. The goal of this work is to deploy the fewest PMUs while still providing comprehensive coverage in terms of fault detection. Specifically, we first developed a fault detection method based on voltage sags of three phases, and then carried out a sensitivity study on the accuracy of this method, with respect to the distance from the fault location. To this end, we have investigated three different types of distances, namely topological distance, logical distance, and electrical distance. The sensitivity study shows that our fault detection method can achieve high accuracy when it is executed using data collected within certain distance to the fault location. This creates opportunities to detect faults using less PMUs. We then developed a PMU deployment algorithm which derives valid solutions for PMU deployment based on a pre-defined distance constraint. The evaluation results show that our new PMU deployment plan can achieve high accuracy with less than one-fifth of PMUs originally deployed in the smart grid.

Our work is ongoing in several directions. First, we will develop new fault detection approaches using machine learning techniques, such as classification and clustering. Second, we will create visualization tools based on our study, for enhancing the real-time situation awareness of a smart grid. Third, we will investigate the possibilities to repurpose our fault detection techniques for solving other problems on a smart grid, such as data cleansing, and cyber security challenges.

### ACKNOWLEDGMENT

The generous support from Bonneville Power Administration and OregonBest through the NW Energy XP Award is gratefully acknowledged. The authors also would like to thank Bonneville Power Adminis-


