Phasor Measurement Unit Selection for Unobservable Electric Power Data Integrity Attack Detection

Annarita Giani, Russell Bent, and Feng Pan
DSA-4, Los Alamos National Laboratory

Abstract

Electric power system operators make critical decisions based on remote measurements. If those measurements are compromised, it is possible to make decisions that could lead to critical consequences. Of particular concern are unobservable attacks where compromised measurements are not flagged as erroneous by bad data detection algorithms. The use of secure measurement devices, such as PMUs, is one way to recognize such attacks. In this paper we present an algorithm based on integer programming for optimal placement of PMUs to detect unobservable electric power SCADA data integrity attacks. Alternatively, this algorithm can be used to identify minimal sets of existing PMUs whose data is needed to detect unobservable bad data attacks. We show that the algorithm is efficient on practical examples drawn from the power engineering literature.

1 Introduction

In the modern world, societies and economies are increasingly dependent on the services (electric power, natural gas, water, etc.), that infrastructure systems provide. These systems are highly complex and are governed by highly non-linear physics relationships. This complexity makes such systems very difficult to control and operate. Despite this complexity, considerable progress was made in recent years to improve the processes that are used to operate infrastructure systems. The smart grid initiatives are an example of such advances [1, 2]. These control processes are highly dependent on accurate system state data that is remotely measured and transmitted to control systems via advanced supervisory control and data acquisition (SCADA) systems. Transmission of these data represents a point of vulnerability of the system to cyber attackers.

In this paper we focus on data integrity attacks to SCADA systems for electric power. Currently, SCADA devices in power systems measure system states such as power injections at substations (buses), power flows at lines and transformers, voltage values (magnitudes), etc. Historically, such data is measured and transmitted with the expectation that there is noise and error in the measurements and that it does not provide enough information (for example, voltage phase angles) to completely characterize the state of a system. As a result, the power engineering community has developed sophisticated techniques to estimate the state of unobserved portions of the system and filter bad data [3]. These techniques are robust to random failures and expected measurement errors in power systems. However, increasingly, there is concern that it is possible to introduce errors into the data in a coordinated fashion that is undetectable by bad data filters [4]. When such error is introduced by a malicious source (such as a cyber attacker), this error is referred to as a *data integrity attack*. When attacked data is provably able to bypass bad data filters, such attacks are referred to as *unobservable* data integrity attacks [4]. In general, unobservable attacks require the compromise of large numbers of sensors and much recent work has focused on developing general methods to identify worst case sets of sensors based on constraints on how many sensors may be comprised [4, 5, 6, 7]. While such models are important for assessing the vulnerability of a system, they suffer computationally as the models tend to be very difficult to solve.

A subset of unobservable data integrity attacks are those that only require a small number of compromised sensors. It can be argued that such attacks are the most realistic as an attacker has limited resources (e.g., time and information) to plan an attack. This makes attacks that consist of a small number of sensors desirable. These types of attacks are referred to as k-sparse attacks, where k is the number of sensors that are compromised [4, 5]. Recent papers have shown that identifying all possible 3, 4, and 5 sparse attacks requires polynomial time [5], eliminating the computational challenges associated with more general models. More importantly, perhaps, [5] identifies the types of redundant measurements that are required to make unobservable k-sparse attacks detectable. One important measurement (there are other possible measurements that could be used, such as frequency, line flows, etc.) for detecting k-sparse attacks is voltage phase angles. Voltage phase angles are typically estimated from other measurements. Since Phasor Measurement Units (PMUs) directly provide these measurements [8], they are candidate devices for detecting unobservable attacks. In this paper we develop optimization models for optimally placing PMUs to cover undetectable attacks. Alternatively, in the case where PMU deployment is ubiquitous, we show that the optimization models are used to identify the smallest set of PMUs that are needed for use in detecting attacks.

Literature Review The PMU placement problem is generally an NP-Complete problem, and, as discussed in [5], our placement problem is no different. However, there is limited work on optimizing the placement of PMUs to combat k-sparse attacks. Reference [5] optimizes the placement of PMUs using a polynomial time algorithm that is guaranteed to find a sufficient number of PMUs, but is not guaranteed to find the optimal solution. In this paper, we describe a model that is guaranteed to find the optimal solution. In the worst case, this algorithm requires exponential computational time, but is efficient in practice when tested on a wide range of problems.

More generally, there are a number of interesting papers that address similar PMU placement problems. These papers include determining the optimal placement of PMUs in order to improve system observability [9, 10, 11, 12, 13]. There is also work that maximizes the amount of mutual information between PMU measurements and the power system states [14]. Multi-objective criteria (observability, cost, importance and security) are considered in [15]. In references [14, 16, 17], the PMU placement problem is posed in terms of improving state estimation. Interested readers are invited to use reference [18] to see a comprehensive review of different types of PMU allocation problems.

The main contributions of this paper are as follows:

- We developed a mixed integer program for determining the minimal number of PMUs required to defend against an arbitrary set of unobservable attacks.
- The models for placing or selecting PMUs to detect k-sparse attacks are based on PMU capabilities. This paper includes a comparison on the relative merits of each capability, in terms of how many PMUs are required to detect attacks.
- We show that the models are tractable in practice through empirical studies on examples drawn from the power engineering literature.

The remainder of this paper is organized as follows: Section 2 summarizes the previous work on smart grid unobservable attacks. Section 3 introduces how PMUs are used for countermeasures against such attacks. Section 4 describes the mixed integer programming models used for optimal PMU placement or selection. Section 5 discusses experimental results based on simulation on IEEE test cases. The paper ends with a section on conclusions.

2 Unobservable Smart Grid Data Integrity Attacks

For completeness, we first summarize the main results of [5]. Electric power systems are potentially vulnerable to a large number of unobservable data integrity attacks. Data integrity attacks are defined as attacks that modify data that is measured at remote locations (i.e., meters and sensors) either at sensing or during its transmission to other locations (i.e., control centers). Data integrity attacks that are *consistent* with power flow physics and uncompromised data are called *unobservable* [4]. Unobservable attacks require coordination–compromised meter readings must be carefully orchestrated to fall on a low dimensional manifold in order for the attack to be unobservable. Since the attacks are not observable, it is possible for such attacks to cause significant errors in applications like state estimation. References [4, 5] provide a formal definition of unobservable attacks. As compromising a large number of sensors is a difficult task, we focus on attacks that compromise a modest number of meters as discussed in [5]. It describes efficient algorithms to find all unobservable attacks involving the compromise of exactly two power injection meters and an arbitrary number of power meters on lines. In this paper, we use this approach to enumerate sparse attacks. We then optimize PMU resources based on this set of attacks.

One of the interesting attributes of unobservable attacks is that they partition a power network into observable islands. These are disjoint subsets of buses which share the same perceived change of state under attack [5]. Figure 3 shows an example of how an attack partitions a network into islands. Conceptually, this means phase angles shift by the same quantity within each island. If there is a PMU in an island, it can detect when a shift is due to the normal behavior of the system or is only a perceived shift due to an attack. Thus, PMUs render the attack observable. As will be discussed later, during an attack, at most one island will exhibit no shift. From [5] we have the following definitions.

DEFINITION 1: An *attack* $\mathcal{A} = (\mathbb{S},\ a)$ is a set of meters $\mathbb{S}$, and an attack vector $0 \neq a \in \mathbb{R}^{m+n+1}$ where $n+1$ is the number of buses and m is the number of measurements. The nonzero components of a correspond to the *compromised* meters in $\mathbb{S}$, i.e. $k \in \mathbb{S} \iff a_k \neq 0$. Under the attack $\mathcal{A}$, the meter readings are changed by the attacker from their uncompromised values y to the compromised values $y + a$. We abuse language and say that a line is compromised when the meter on that line is compromised. The *sparsity* of the attack $\mathcal{A} = |\mathbb{S}| =$ the number of compromised meters.

□

DEFINITION 2: Consider a power system with the linear DC power flow model $y = Hx$. Let x^o denote the current system state, and $y^o = Hx^o$ denote the uncompromised measurements. An attack $\mathcal{A}$ is called *unobservable* at operating point x^o with respect to the model if there exists some system state consistent with the compromised observations, i.e.

$$\exists \ x^a \ : \ y^o + a = H(x^o + x^a)$$

□

REMARK 1: x^a is the (unique) *perceived state perturbation* associated with attack $\mathcal{A}$. It is the fictitious change of system state necessary to produce the compromised meter readings $y^o + a$. As the model is linear, $\mathcal{A} = (\mathbb{S}, a)$ is unobservable if and only if $a \neq 0, a = Hx^a$ is solvable, and $\mathbb{S}$ indexes the nonzero elements of a. Unobservability of $\mathcal{A}$ under the model does not depend on the current system state x^o.

□

The following is the main theorem of [5] that defines unobservable attacks in terms of the rows in H corresponding to the measurements under attacks.

THEOREM 1: Consider the DC power flow model. Consider an unobservable attack $\mathcal{A} = (\mathbb{S}, a)$. Construct the matrices K and L from H by deleting the rows in $\mathbb{S}$ and by retaining the rows in $\mathbb{S}$ respectively. Then

(a) rank $(K) \leq n - 1$

(b) the attack vector a is a nonzero vector in the subspace:

$$\mathcal{T} = \{a \in \mathbb{R}^{m+n+1} : a = Hx, \ Kx = 0\}$$

□

Is there a typo in the theorem? It says we are deleting and keeping the rows of S? RB

Another crucial concept that was introduced in [5] is the concept of *observable islands*. Observable islands are disjoint subsets of buses that share the same perceived change of state [voltage phase] under the attack. More precisely:

DEFINITION 3: Let $\mathcal{A} = (\mathbb{S},\ a)$ be an unobservable attack, and let x^a be its associated *perceived* change of system state. Partition the set of buses $\mathbb{V}$ into the disjoint union

$$\mathbb{V} = \mathbb{V}_1 \cup \cdots \cup \mathbb{V}_s, \quad \mathbb{V}_i \cap \mathbb{V}_j = \phi \quad \text{for } i \neq j$$

defined by the equivalence classes

$$v_1,\ v_2 \in \mathbb{V}_i \iff x^a(v_1) = x^a(v_2)$$

The sets $\{\mathbb{V}_i\}_{i=1}^{s}$ are called the *observable islands* associated with the attack $\mathcal{A}$.

□

The following results connect observable islands with unobservable attacks. Consider the unobservable attack $\mathcal{A} = (\mathbb{S},\ a)$.

(a) Every compromised line in $\mathbb{S}$ connects a pair of distinct observable islands.

(b) Every line that connects distinct observable islands is either unmetered or compromised.

(c) No lines contained within an observable island are compromised.

Reference [5] also offers a graph theoretical characterization of 3, 4, 5 sparse attacks.

THEOREM 2: Assume all lines are metered. An irreducible attack $(\mathbb{S}, a)$ is 3-sparse if and only if

(a) $\mathbb{S}$ consists of two adjacent injection buses b_1, b_2 and the line ℓ connecting these buses

(b) The connecting line ℓ is a bridge of the power system graph $\mathcal{G}$.

An irreducible attack $(\mathbb{S}, a)$ is 4-sparse if and only if

(a) $\mathbb{S}$ consists of two injection buses b_1, b_2 and two lines ℓ_1, ℓ_2.

(b) The injection buses b_1, b_2 are connected by the lines ℓ_1, ℓ_2 via an intermediate bus b_o.

(c) The connecting lines ℓ_1, ℓ_2 are bridges of the power system graph $\mathcal{G}$.

□

Figures 1 and 2 show how 3,4,5 sparse attacks look like.

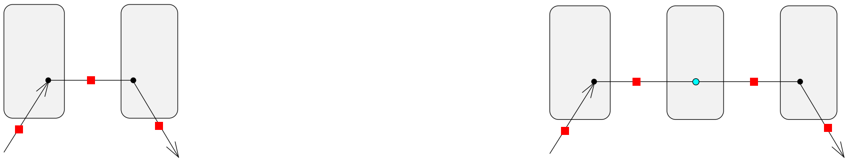

Figure 1: Canonical forms: 3-sparse (left) & 4-sparse (right) irreducible attacks.

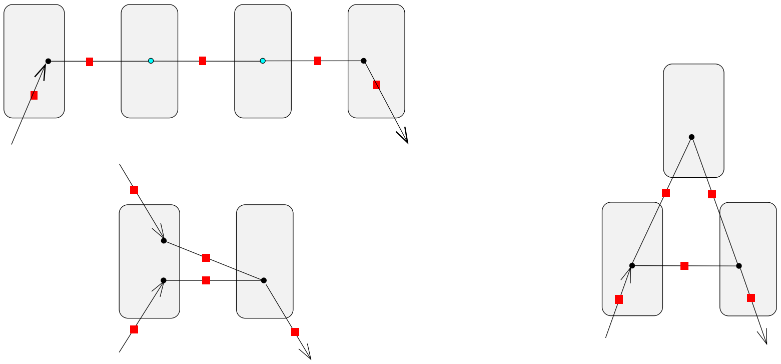

Figure 2: Three canonical forms for 5-sparse irreducible attacks.

Is there a something for 5 sparse attacks lilke you have for 3 and 4 sparse attacks? RB

3 PMU Allocation/Selection for Attack Detection

In power systems, the basic physics of power flows force power to flow from high voltage phase angles to low voltage phase angles. The amount of flow is proportional to the phase angle difference between the source and the sink. Irrespective of flows, the differences in phase angles is important because large deviations cause system instability. While state estimation estimates phase angles on slow time scales (5 minutes or more), instability can occur on much more rapid time scales. Thus, one of the original motivations for developing and deploying PMUs was to directly measure phase angles on fast times scales to provide early warning of a system drifting towards instability.

In this paper, we take advantage of the PMU's capability to directly measure phase angles and line flows to counter unobservable attacks. While there are vulnerability concerns related to PMUs [19, 20], in this paper we assume they are secure [21] as they are engineered for security, use the modern NASPInet infrastructure, and are more secure than older SCADA controls. There are at least two ways a PMU can counter an unobservable attack. First, since PMUs measure line flows, if a PMU is placed at a bus that is connected to a line that crosses island boundaries, it can implicitly measure phase angles in both islands to detect a shift [22].

Second, as discussed earlier, an attack partitions a power network into observable islands (e.g. Figure 3). A PMU placed at bus k offers direct measurement of the voltage phase x_k at that bus which is instead estimated with traditional SCADA devices. Theorem 16 in [5], restated here, states that two PMUs placed in two distinct observable islands are sufficient to thwart the attack.

THEOREM 2: Consider an arbitrary collection of unobservable attacks $\mathbb{A} = \{\mathcal{A}_1, \cdots, \mathcal{A}_p\}$. Let

$$\mathbb{V}_1^k \quad \mathbb{V}_2^k \quad \cdots \quad \mathbb{V}_{s_k}^k$$

denote the observable islands associated with attack $\mathcal{A}_k$. All attacks in $\mathbb{A}$ can be made observable by placing PMUs at buses $\mathbb{B}$

$\iff \quad \forall\, k,\ \exists\, i_1 \neq i_2 :\ \mathbb{V}_{i_1}^k \cap \mathbb{B} \neq \phi,\ \mathbb{V}_{i_2}^k \cap \mathbb{B} \neq \phi$

i.e. every attack has two distinct islands which contain PMUs.

□

Reference [5] has shown that, when using this method, $p+1$ PMUs are sufficient to thwart a collection of p unobservable attacks. In this paper we show that significantly fewer are required.

Given the PMU's capabilities and limited PMU resources, there are three distinct steps required to quantitatively prioritize countermeasure investments. *First*, we must identify all sets of devices that can be used to conduct unobservable attacks. Here we use [5] to enumerate (in polynomial time) all possible data integrity attacks consisting of 3,4,5 sensors. *Second*, we must assess the risk to the system from the attacks. Here, we focus on consequence using the techniques of [23] that calculate economic consequence. It is important to note that our approach is general and can use any method for calculating consequence, including system damage–we use economics for the purpose of demonstrating the approach. *Third*, we must identify the minimum set of PMUs to cover all the attacks, or with limited resources, the placement of PMUs to cover the more critical attacks to minimize consequence. Under PMU deployment scenarios, this approach gives a security criteria for their allocation. If PMU deployment is ubiquitous, but data processing resources are limited, then this approach prioritizes which PMU data streams are used.

4 Optimization Models

In this section we develop mixed integer programming models for determining the optimal PMU placement for detecting k-sparse attacks. The first set of models minimizes the cost of undetected attacks when PMU resources are limited (Section 4.1). The second set of models

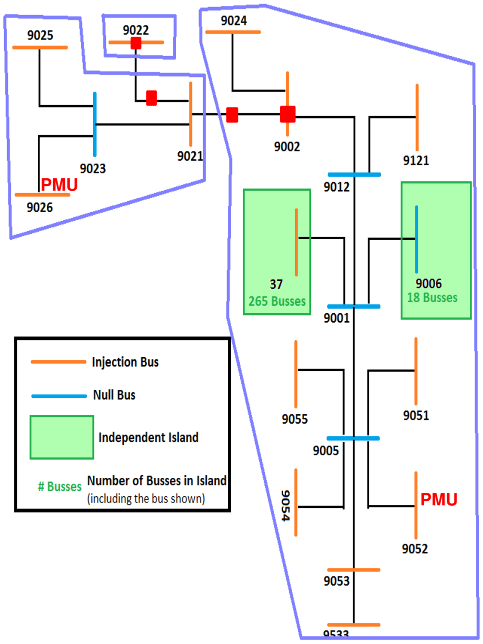

Figure 3: An example of a four sparse attack on the IEEE 300 electric power system. For space constraints, the green rectangles group 265 and 18 buses, respectively. The grid is divided into three islands by the attack (red squares). Two PMUs in two different islands are sufficient as a countermeasure to the attack (for example, buses 9026 and 9052).

minimizes the total number of PMUs required to detect all attacks (Section 4.2).

4.1 Minimizing Cost

Formally, for an attack $a \in A$, the notation η_a is used to denote the cost or consequence of the attack if it is undetected. The cost quantifies the penalty for allowing that attack to go undetected. In this paper, the work of [23] is used to compute η_a. The set $\bar{A}_a$ is then the set of buses adjacent to lines included in the attack. We then introduce the variable x_a that is set to 0 when attack a is detectable and 1 otherwise. Similarly, for a bus $n \in N$, the variable p_n is used to determine if a PMU is placed at n (it is set to 1 if there is a PMU at n, and 0 otherwise). Next, the term I_a is used to denote the set of islands for a. I_a consists of sets of buses such that $\bigcup_{i \in I_a} V(i) = N$ and $\bigcap_{i \in I_a} V(i) = \emptyset$, where $V(i)$ is the set of buses in island i. The variable y_a^i is then used to denote whether or not island i of attack a has a PMU (it is set to 1 if the island has a PMU, and 0 otherwise).

In order to assess the relative value of PMU capabilities for unobservable attack detection, we isolate each of the capabilities in our optimization models.

4.1.1 Phase Angle Shift Detection (PASD)

In our first optimization model, we assume that the only capability for detecting an attack is through the direct detection of phase angle shift. This is achieved by placing a PMU at buses in at least two different islands defined by the attack. In this model, the goal is to place PMUs at buses such that sum of the consequence of undetected attacks is minimized. The model is formally presented here:

$$\text{minimize} \sum_{a \in A} \eta_a x_a \tag{1}$$

$$\text{subject to} \sum_{n \in N} p_n \leq P \tag{2}$$

$$(1 - x_a) \leq \frac{1}{2} \sum_{i \in I_a} y_a^i \qquad \forall a \in A \tag{3}$$

$$y_a^i \leq \sum_{n \in I_a} p_n \qquad \forall a \in A, \forall i \in I_a \tag{4}$$

$$y, p, x \in \{0, 1\} \tag{5}$$

Equation (1) describes the objective function for minimizing the cost of the attacks that are undetected. Equation (2) states the constraint that the number of PMUs is limited by P. Equation (3) states the constraints that assure that an attack is detected when PMUs are placed in two different islands. In the case that only one island contains a PMU, the attack is undetected because of Equation (5) constrains solutions to $\{0, 1\}$. Equation (4) is a constraint that associates an island with the PMUs at buses in that island. The right hand side counts how many buses in the island have PMUs. If there is at least 1, then y_a^i may be set to 1.

4.1.2 Flow Based Shift Detection (FBSD)

In our second optimization model, we use the capabilities of a PMU to measure line flow [22]. The PMUs directly measure the phase angle at its bus and then indirectly measures the phase angle at the connected bus using the line flow measurement, detecting phase angle shifts at either side of the line. The model is formally presented in here:

$$\begin{aligned} \text{minimize} \; & \sum_{a \in A} \eta_a x_a && (6) \\ \text{subject to} \; & \sum_{n \in N} p_n \leq P && (7) \\ & (1 - x_a) \leq \sum_{n \in \bar{A}_a} p_n \quad \forall a \in A && (8) \\ & p, x \in \{0, 1\} && (9) \end{aligned}$$

This model is derived from Equations 1-4 with the following differences. First, the island constraint (Equation 4) is dropped. Second, the phase shift detection constraint (Equation 3) is replaced with the constraint in Equation 8. Equation 8 states that an attack is detected if one of the buses adjacent to a line connecting islands has a PMU, thereby using the PMU's capability to measure line flows and associated angle shift on both sides of the line [22].

4.1.3 Phase Angle Shift and Line Flow Detection (PASLFD)

In our third optimization model, we assume that both attack detection capabilities are available. The model is a combination of Equations 1-4 and Equations 6-9:

$$
\begin{aligned}
\text{minimize } & \sum_{a \in A} \eta_a x_a && && (10)\\
\text{subject to } & \sum_{n \in N} p_n \leq P && && (11)\\
& (1 - x_a) \leq \frac{1}{2} \sum_{i \in I_a} y_a^i + \sum_{n \in \bar{A}_a} p_n && \forall a \in A && (12)\\
& y_a^i \leq \sum_{n \in I_a} p_n && \forall a \in A,\ \forall i \in I_a && (13)\\
& y, p, x \in \{0, 1\} && && (14)
\end{aligned}
$$

In this model, Equations 3 and 8 are combined so that an attack is detected if either one is satisfied (Equation 12).

4.2 Minimizing the Number of PMUs

These optimization models are easily modified to identify the minimal set of PMUs needed to detect all the attacks. This is the problem originally posed by [5] and was not solved to global optimality (until now). The models are modified by removing the maximum PMU constraint (i.e., constraint 2), replacing the x variables with the constant 0, and replacing the objective function with

$$
\text{minimize } \sum_{n \in N} p_n \qquad (15)
$$

5 Experimental Results

In this section we discuss our experimental results and demonstrate that our approach is computationally tractable on a wide variety of systems. We first present results where the objective is Equation (15), i.e. the goal is to minimize the total number of PMUs required to cover all attacks. Second, we present results where the objective is Equation (1), i.e. the goal is to minimize the cost of uncovered attacks when the number of PMUs is limited.

5.1 Minimizing PMUs

In this section we compute the minimum number of PMUs that are needed to detect all 3, 4, and 5 sparse data integrity attacks on IEEE test problems. We compare these results with the upper bound provided in Theorem 16 of [5] and show that in practice (for the IEEE networks) it is empirically computationally tractable to calculate the optimal set. The results are shown in Table 1. The first column labels the IEEE electric power network. The second column provides the upper bound on the number of PMUs required that were derived in [5]. The upper bound in [5] is $p+1$ where p is the number of unobservable attacks. The remaining three sets of columns list the minimum number of PMUs for PASD, FBSD, and PASLFD and the CPU seconds required to solve each problem. The problems were solved on a desktop computer with an Intel Xeon 2.67 Ghz processor using Cplex 12.5.

For these problems, it is clear that the phase shift detection capabilities dominate the line flow detection capabilities of PMUs. Only when there is one possible attack in a system, does FBSD provide better results (RTS-79 and IEEE 57). Otherwise, FBSD does not increase the capability of PASD to cover all attacks. Intuitively this means that there is limited overlap of attack detection in FBSD, whereas there is a great deal of overlap in PASD. This allows PASD to overcome the limitations of needing two instead one PMU to cover an attack. These results also improve the upper bound discussed in [5].

5.2 Minimizing Consequence

In this section we present results on IEEE-118, IEEE-300, RTS-96, and the Polish grid (2383 buses, winter peak) included with the Matpower distribution [24] when PMU resources are limited. The approach of [23] is used to compute attack consequence. Note that some attacks

		PASD		FBSD		PASLFD	
IEEE Test Cases	[5]	PMUs	CPU Time	PMUs	CPU Time	PMUs	CPU Time
RTS-79	2	2	< 0.001	1	< 0.001	1	< 0.001
30	5	2	< 0.001	3	0.06	2	0.04
39	13	4	< 0.001	7	< 0.001	4	< 0.001
57	2	2	0.03	1	< 0.001	1	< 0.001
RTS-96	3	2	< 0.001	2	< 0.001	2	0.03
118	9	5	< 0.001	6	0.02	5	0.01
300	142	55	0.1	63	0.01	55	4.4
2383wp (Polish)	576	233	1.12	310	7.67	233	13.3
2736sp (Polish)	194	136	0.34	152	0.3	136	0.35
3012wp (Polish)	588	254	0.49	344	0.6	254	15.33

Table 1: This table presents the optimal number of PMUs required to detect all 3, 4, 5 sparse data integrity attacks using the PASD, FBSD, and PASLFD models. In each case, the CPU time in seconds is reported. The results are compared with the upper bound discussed in [5]. Results are presented on ten IEEE electric power problems, including three large scale problems based on Poland's electric power grid. In all cases, the problems are solvable in less than 1 CPU second and in all cases, less than 1 CPU minute. In terms of solution quality, the PASD (phase angle shift) model requires fewer PMUs to detect all attacks. Furthermore, the capability of a PMU to measure line flows provides limited additional benefits for detecting attacks.

are eliminated by [23] because they involve buses with no injection, rendering them useless for data integrity attacks that modify injection. In the IEEE Reliability Test System, RTS-96 (73 bus system), there are two 3-sparse attacks ([207, 208], [307, 308]) and no others. The consequence cost of both attacks is 45. Table 2 describes the cost of uncovered attacks for different PMU levels when using PASD. No attacks are covered with one PMU and all attacks are covered when 2 are available. For FBSD and PASLFD one attack of cost 45 is covered with one PMU and the rest of the results remain the same. Similarly, there are five 3-sparse attacks, three 4-sparse attacks and no 5 sparse attacks for IEEE 118. Two attacks have non-zero consequence. An attack on buses 85 and 86, and the line between them has cost 12253. An attack on buses 12 and 117, and the line between has cost 23410. Table 3 describes the cost of uncovered attacks for different PMU levels using PASD. No attacks are

Available PMUs	Covered attacks	Cost uncovered attacks	CPU Time
1	0	90	0.04
2	2	0	0.02
3	2	0	0.02

Table 2: IEEE RTS 96 Test Case. Cost of uncovered attacks in case of limited resources.

covered with one PMU and all attacks are covered when 2 PMUs are available. For FBSD and PASLFD one attack of cost 23410 is covered with one PMU and the rest of the results remain the same.

Due to space limitations, for both IEEE 300 and the polish grid case, the number of sparse attacks is too large to list in tables in this manuscript. Overall, there are 32 3-sparse attacks, 47 4-sparse attacks and 26 5-sparse attacks in IEEE 300 (for a total of 105 attacks). In the IEEE 2383 polish grid, winter peak network (IEEE 2383wp) there are 269 3-sparse attacks, 206 4-sparse attacks, and 91 5-sparse attacks (for a total of 566 attacks).

Figure 4 shows the cost of undetected attacks as function of the number of PMUs that can be placed for the IEEE 300 problem. Figure 5 presents results for the IEEE 2383wp test case. The large initial drop in Figure 5 is due to the fact that only a few attacks have major consequences. Most of the attacks have small effect or no effect at all. In both cases, the PASD capability of PMUs dominates FBSD except for when one PMU is available. This result provides additional evidence that PMU deployment based upon the phase shift capabilities of a PMU are better than deployments based upon a PMUs line flow capabilities. However, unlike the minimum PMU results, when PMUs are placed based upon the combined capabilities (PASLFD) there is some marginal benefit. In the plots, there are occasionally points where slightly fewer PMUs are needed to achieve the same level of consequence reduction when PASLFD is used.

Available PMUs	Covered attacks	Cost uncovered attacks	CPU Time
1	0	35663	0.05
2	2	0	0.03
3	2	0	0.03

Table 3: IEEE 118 Test Case. Cost of uncovered attacks in case of limited resources.

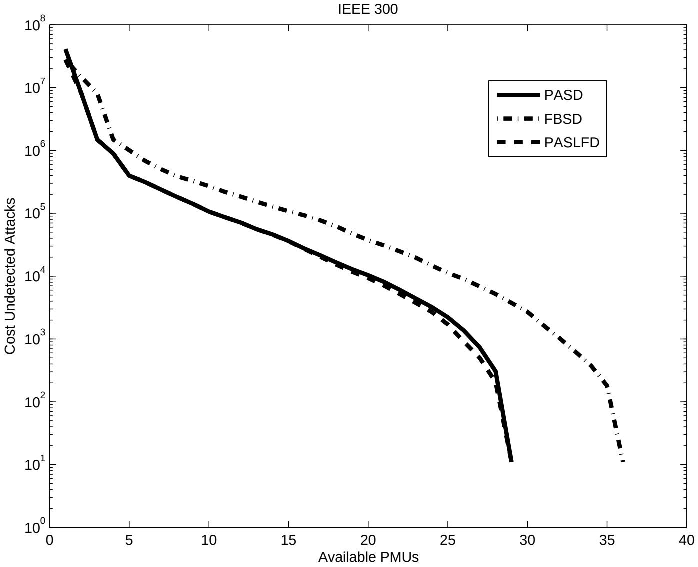

Figure 4: This graph compares the results of PASD, FBSD, and PASLFD on the IEEE 300 problem. The goal is to minimize the cost of uncovered attacks (y axis–log scale) when the number of PMUs is limited (x axis). PMU placement based on phase angle shift capabilities of a PMU require significant fewer PMUs than placements based upon a PMU's line flow shift capabilities. PMU placement based upon both capabilities provide marginal benefit between 15 and 27 PMUs.

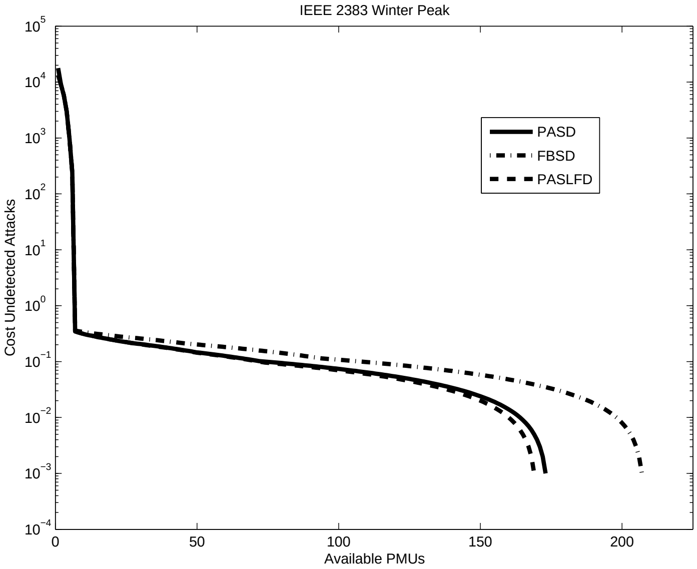

Figure 5: This graph compares the results of PASD, FBSD, and PASLFD on the IEEE Polish 2383 winter peak problem. The goal is to minimize the cost of uncovered attacks (y axis–log scale) when the number of PMUs is limited (x axis). PMU placement based on phase angle shift capabilities of a PMU require significant fewer PMUs than placements based upon a PMU's line flow shift capabilities. Like the IEEE 300 case, PMU placement based upon both capabilities does provide some marginal benefit by reducing the number of PMU's required to meet the same consequence level as PASD or FBSD.

6 Conclusions

Recent years have seen increased interest in understanding the vulnerabilities of electric power grids to cyber attacks. Indeed, recent work by [4, 5, 6, 7] and others has shown that it is possible for an attacker to falsify information sent to the grid operator so that the incorrect information remains consistent with other measurements reported to the operator. In this paper we developed a mixed integer programming model that is used for optimal deployment of PMUs and for optimal selection of existing PMUs to mitigate sparse unobservable attacks. We have shown that despite the NP-Completeness of the problem of minimizing PMU resource utilization for detecting k-sparse attacks, in practice this problem can be solved efficiently for a wide variety of problems.

There are a number of future questions that remain to be answered. First, other types of data integrity attacks need to be considered, including the on/off status of a power lines (either from direct measurements or state estimation [25, 26, 27]), the output of generators, the states of control devices, etc. PMUs could be used to counter such attacks and PMU resource allocation should include mitigating these attacks. Second, it will be important to develop multi-objective PMU placement models that include k-sparse security considerations along with existing considerations such as overall improved state estimation. Finally, it will be interesting to test the approach on larger scale problems (10,000 nodes or more), as this is the size of problems at the interconnection level in the United States. The results on problems with up to 2800 nodes are promising, however, further research is required to confirm that the approach will scale to larger problems.

Acknowledgment

This work was partially supported by the Los Alamos National Laboratory Director's Fellowship *Cyber Security of the Smart Grid* and the Center for Nonlinear Studies (CNLS).

References

[1] H. Farhangi, "The Path of the Smart Grid," *Power and Energy Magazine, IEEE*, vol. 8, no. 1, pp. 18–28, 2010.

[2] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid the new and improved power grid: A survey," *Communications Surveys Tutorials, IEEE*, vol. 14, no. 4, pp. 944–980, 2012.

[3] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, no. 2, pp. 262–282.

[4] Y. Liu, P. Ning, and M. K. Reiter, "False Data Injection Attacks Against State Estimation in Electric Power Grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, pp. 13:1–33, Jun. 2011.

[5] A. Giani, E. Bitar, Garcia, M. M., McQueen, P. Khargonekar, and K. Poolla, "Smart Grid Data Integrity Attacks," *IEEE Transaction on Smart Grids*, vol. 4, no. 3, pp. 1244–1253, 2013.

[6] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious Data Attacks on the Smart Grid," *Smart Grid, IEEE Transactions on*, vol. 2, no. 4, pp. 645–658, 2011.

[7] G. Dan and H. Sandberg, "Stealth Attacks and Protection Schemes for State Estimators in Power Systems," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, 2010, pp. 214–219.

[8] A. Phadke, "Synchronized Phasor Measurements in Power Systems," *Computer Applications in Power, IEEE*, vol. 6, no. 2, pp. 10–15, 1993.

[9] T.-T. Cai and Q. Ai, "Research of PMU Optimal Placement in Power Systems," in *Proceedings of the 5th WSEAS/IASME International Conference on Systems Theory and Scientific Computation*, ser. ISTASC'05. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2005, pp. 38–43.

[10] N. Manousakis, G. Korres, and P. Georgilakis, "Taxonomy of PMU Placement Methodologies," *Power Systems, IEEE Transactions on*, vol. 27, no. 2, pp. 1070–1077, May.

[11] R. H. Shewale, B. krishna Kethineni, U. P. Balaraju, S. Bhil, and P. D. More, “Optimal Placement of Phasor Measurement Unit for Power System Observability by Heuristic Search Method,” *International Journal of Advanced technology and Engineering Research (IJATER)*, vol. 2, no. 2, pp. 128–133, 2012.

[12] R. Sodhi, S. Srivastava, and S. Singh, “Optimal PMU Placement to Ensure System Observability Under Contingencies,” in *Power Energy Society General Meeting, 2009. PES ’09. IEEE*, 2009, pp. 1–6.

[13] F. Aminifar, A. Khodaei, M. Fotuhi-Firuzabad, and M. Shahidehpour, “Contingency-Constrained PMU Placement in Power Networks,” *Power Systems, IEEE Transactions on*, vol. 25, no. 1, pp. 516–523, 2010.

[14] Q. Li, T. Cui, Y. Weng, R. Negi, F. Franchetti, and M. D. Ilic, “An Information-Theoretic Approach to PMU Placement in Electric Power Systems,” *E-print arXiv:1201.2934*, 2012.

[15] O. Linda, A. Giani, M. Manic, and McQueen, “Multi-Criteria Based Staging of Optimal PMU Placement using Fuzzy Weighted Average,” *Proc. of IEEE International Symposium on Industrial Electronics, IEEE ISIE*, 2013.

[16] S. Chakrabarti, E. Kyriakides, and M. Albu, “Uncertainty in Power System State Variables Obtained Through Synchronized Measurements,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 58, no. 8, pp. 2452–2458, 2009.

[17] J. Chen and A. Abur, “Placement of PMUs to Enable Bad Data Detection in State Estimation,” *Power Systems, IEEE Transactions on*, vol. 21, no. 4, pp. 1608–1615, 2006.

[18] W. Yuill, A. Edwards, S. Chowdhury, and S. P. Chowdhury, “Optimal PMU Placement: A Comprehensive Literature Review,” in *Power and Energy Society General Meeting, 2011 IEEE*, July, pp. 1–8.

[19] S. D'Antonio, L. Coppolino, I. A. Elia, and V. Formicola, "Security Issues of a Phasor Data Concentrator for Smart Grid Infrastructure," in *Proceedings of the 13th European Workshop on Dependable Computing*, ser. EWDC '11, 2011, pp. 3–8.

[20] A. A. F. Daniel P. Shepard, Todd E. Humphreys, "Evaluation of the Vulnerability of Phasor Measurement Units to GPS Spoofing Attacks," *International Journal of Critical Infrastructure Protection*, vol. 5, pp. 146–153, 2012.

[21] M. McQueen and A. Giani, "'Known Secure Sensor Measurement' for Critical Infrastructure Systems: Detecting Falsification of System State," *Proc. of the 3rd International Workshop on Software Engineering for Resilient Systems, SERENE*, 2011.

[22] B. Xu and A. Abur, "Observability analysis and measurement placement for systems with pmus," in *Power Systems Conference and Exposition, 2004. IEEE PES*, Oct 2004, pp. 943–946 vol.2.

[23] A. Giani, F. Pan, R. Bent, and K. Poolla, "Phasor Measurement Unit Placement for Unobservable Attack Detection," in *Los Alamos National Laboratory Technical Report*, no. LA-UR-13-24315, 2013.

[24] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, 2011.

[25] H. Singh and F. Alvarado, "Network Topology Determination using Least Absolute Value State Estimation," *IEEE Transactions on Power Systems*, vol. 10, no. 3, pp. 1159–1165, 1995.

[26] D. Singh, J. Pandey, and D. Chauhan, "Topology Identification, Bad Data Processing, and State Estimation Using Fuzzy Pattern Matching," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1370–1379, 2005.

[27] R. Singh, E. Manitsas, B. Pal, and G. Strbac, "A Recursive Bayesian Approach for Identification of Network Configuration Changes in Distribution System State Estimation," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1329–1336, 2010.