Fundamental Limits of Cyber-Physical Security in Smart Power Grids

Yue Zhao, Andrea Goldsmith, and H. Vincent Poor

***Abstract*— Cyber-physical security of power systems under power injection attacks that alter generation and loads is studied. The system operator employs Phasor Measurement Units (PMUs) for detecting such attacks, while attackers devise attacks that are *unobservable* by such PMU networks. For the NP-hard problem of finding the *sparsest* unobservable attacks, it is shown that the solution has a simple form with probability one, namely, $\min\big(\kappa(\mathcal{G}^{\mathcal{M}}), M\big) + 1$, where $\kappa(\mathcal{G}^{\mathcal{M}})$ is the vertex connectivity of an augmented graph, and M is the number of PMUs. The constructive proof allows one to find the entire set of the sparsest unobservable attacks in polynomial time. Furthermore, the geometric interpretation of unobservable attacks leads to a natural characterization of their potential impacts. With optimized PMU deployment, the sparsest unobservable attacks and their potential impact as functions of the number of PMUs are evaluated numerically for IEEE 30, 57, 118, 300-bus systems and Polish 2383, 2737, 3012-bus systems. It is observed that, as more PMUs are added, the maximum potential impact among all the sparsest unobservable attacks drops quickly until it reaches the minimum sparsity.**

I. Introduction

Modern power networks are increasingly dependent on information technology in order to achieve higher efficiency, flexibility and adaptability. The development of more advanced sensing, communications and control capabilities for power grids enables better situational awareness and smarter control. However, security issues also arise as more complex information systems become prominent targets of cyber-physical attacks: not only can there be data attacks on measurements that disrupt situation awareness [1], but also control signals of many power grid components including generation and loads can be hijacked, directly leading to physical misbehavior of power systems [2]. Therefore, to achieve reliable and secure operation of a smart power grid, it is essential for the system operator to minimize (if not eliminate) the feasibility and impact of such cyber-physical attacks.

Recently, there has been considerable research concerning data injection attacks on sensor measurements, particularly from supervisory control and data acquisition (SCADA) systems. A common and important goal among these works is to pursue the integrity of network *state estimation*, that is, to successfully detect the injected data attack and recover the correct system states. The feasibility of constructing false data injection attacks to pass bad data detection schemes and alter estimated system states was first shown in [1]. There, a natural question arises as to how to find the *sparsest unobservable* data injection attack, as sparsity is used to model the complexity of an attack, as well as the resources needed for an attacker to implement it. However, finding such an *optimal attack* requires solving an NP-hard l_0 minimization problem. While efficiently finding the sparsest unobservable attacks in general remains an open problem, many interesting solutions under special problem settings have been developed (see, e.g. [3] and [4]). Furthermore, as PMUs become increasingly deployed in power systems, network situational awareness for grid operators is significantly improved compared to using legacy SCADA systems only. However, the high installation costs of PMUs still prohibit large-scale deployment. Thus, the problem of how to economically deploy PMUs to best facilitate the state estimator to detect data injection attacks becomes an interesting problem that many studies have addressed (see, e.g. [5] and [6] among others.)

Compared to data attacks that target state estimators, cyber-physical attacks that directly disrupt power network physical processes can have a much faster (and often stronger) impact on power grids. In addition to physical attacks implemented by hacking control signals or directly intruding upon grid components, several types of load altering attacks have been shown to be practically implementable via Internet-based message attacks [2]. Topological attacks are another type of physical attack which have been considered in [7]. Furthermore, there have been studies of dynamic power injection attacks [8], [9].

In this paper, we investigate a general type of cyber-physical attacks in power systems, namely, *power injection attacks* that alter generation and loads in the network. We consider a grid operator that employs PMUs to (partially) monitor the network for detecting power injection attacks. Since power injection attacks disrupt the power system states immediately, the timeliness of PMU measurement feedback is essential. We study the open l_0 minimization problem of finding the sparsest unobservable attacks given any set of PMU locations. We prove that this in general NP-hard problem has a simple solution with probability one, namely, the sparsity of the optimal solution is $\min\big(\kappa(\mathcal{G}^{\mathcal{M}}), M\big) + 1$, where $\kappa(\mathcal{G}^{\mathcal{M}})$ is the vertex connectivity of an augmented graph of the original power network, and M is the number of PMUs. Furthermore, the geometric interpretation of these sparsest unobservable attacks leads to a natural characteri-

This research was supported in part by the DTRA under Grant HDTRA1-08-1-0010, in part by the Air Force Office of Scientific Research under MURI Grant FA9550-09-1-0643, and in part by the Office of Naval Research under Grant N00014-12-1-0767.

Y. Zhao is with the Dept. of Electrical Engineering, Princeton University, Princeton, NJ, 08544 USA, and with the Dept. of Electrical Engineering, Stanford University, Stanford, CA, 94305 USA (e-mail: yuez@princeton.edu).

A. Goldsmith is with the Dept. of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: andrea@stanford.edu).

H. V. Poor is with the Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

zation of their potential impact. Accordingly, among all the sparsest unobservable attacks, an attacker can easily find the one with the greatest potential impact. Finally, for all possible numbers of PMUs with optimized placement, we evaluate the sparsest unobservable attacks in terms of their sparsity and potential impact in IEEE 30, 57, 118, 300-bus and Polish 2383, 2737, 3012-bus systems.

The remainder of the paper is organized as follows. In Section II, models of the power network, power injection attacks and PMUs are established. We then formulate the minimum sparsity problem of unobservable attacks. In Section III, we prove that the minimum sparsity of unobservable attacks can be found in polynomial time with probability one. The potential impact of unobservable attacks is characterized based on a geometric interpretation. In Section IV, numerical evaluation of the sparsest unobservable attacks in IEEE benchmark test cases and Polish power systems are provided. Conclusions are drawn in Section V. Due to space limitations, proof details (except for Theorem 2) are omitted here, and can be found in [10].

II. Problem formulation

A. Power network model

We consider a power network with N buses, and denote the set of buses and the set of transmission lines by $\mathcal{N} = \{1, \dots, N\}$ and $\mathcal{L} = \{1, \dots, L\}$ respectively. For a line $l \in \mathcal{L}$ that connects buses n and m, denote its reactance by x_l as well as x_{nm}, and define its *incidence vector* $\boldsymbol{m}_l$ as follows:

$$\boldsymbol{m}_l(i) = \begin{cases} 1, & \text{if } i = n, \\ -1, & \text{if } i = m, \\ 0, & \text{otherwise.} \end{cases}$$

Based on the power network topology and line reactances, we construct a weighted graph $\mathcal{G} = \{\mathcal{N}, \mathcal{L}, \boldsymbol{w}\}$ where the edge weight $w_l \triangleq \frac{1}{x_l}, \forall l = 1, \dots, L$. We employ the DC power flow model in this paper, and the real power injections $\boldsymbol{P} \in \mathbb{R}^N$ and the voltage phase angles $\boldsymbol{\theta} \in \mathbb{R}^N$ satisfy $\boldsymbol{P} = \boldsymbol{B}\boldsymbol{\theta}$, where $\boldsymbol{B} = \sum_{l=1}^{L} \frac{1}{x_l} \boldsymbol{m}_l \boldsymbol{m}_l^T \in \mathbb{R}^{N \times N}$ is the *Laplacian* of the weighted graph $\mathcal{G}$. Furthermore, the power flow on line l from bus n to bus m equals $P_{nm} = \frac{1}{x_{nm}}(\theta_n - \theta_m)$.

We consider attackers inflicting power injection attacks that alter the generation and loads in the power network. We denote the power injections under normal conditions by $\boldsymbol{P}$, and denote a power injection attack by $\Delta\boldsymbol{P} \in \mathbb{R}^N$. Thus the post-attack power injections are $\boldsymbol{P} + \Delta\boldsymbol{P}$.

B. Sensor model and unobservable attacks

We consider the use of PMUs by the system operator for monitoring the power network in order to detect power injection attacks. With PMUs installed at the buses, we consider the following two different sensor models:

1) A PMU securely measures the voltage phasor of the bus at which it is installed.
2) A PMU securely measures the voltage phasor of the bus at which it is installed, as well as the current phasors on all the lines connected to this bus[1].

We denote the set of buses with PMUs by $\mathcal{M}$ $(\subseteq \mathcal{N})$, and let $M \triangleq |\mathcal{M}|$ be the total number of PMUs, where $|\cdot|$ denotes the cardinality of a set. We say that a power injection attack $\Delta\boldsymbol{P}$ is *unobservable* if it leads to *zero* changes in all the quantities measured by the PMUs. With the first PMU model described above, we have the following definition:

Definition 1 (unobservability condition): An attack $\Delta\boldsymbol{P}$ is unobservable if and only if

$$\exists \Delta\boldsymbol{\theta}, \text{ such that } \Delta\boldsymbol{P} = \boldsymbol{B}\Delta\boldsymbol{\theta} \text{ and } \Delta\boldsymbol{\theta}_{\mathcal{M}} = \boldsymbol{0}, \qquad (1)$$

where $\Delta\boldsymbol{\theta}_{\mathcal{M}}$ denotes the $M \times 1$ sub-vector of $\Delta\boldsymbol{\theta}$ obtained by keeping its M entries whose indices are in $\mathcal{M}$.

With the second PMU model described above, for any bus $n \in \mathcal{N}$, it is immediate to verify that the following three conditions are equivalent:

1) There are no changes of the voltage phasor at n and of the current phasors on all the lines connected to n.
2) There are no changes of the voltage phasor at n and of the power flows on all the lines connected to n.
3) $\forall n' \in N[n]$, there is no change of the voltage phasor at n', where $N[n]$ is the closed neighborhood of n which includes n and its neighboring buses $N(n)$.

Thus, for forming unobservable attacks, the following two situations are equivalent to the attacker:

- The system operator monitors the set of buses $\mathcal{M}$ with the second PMU model;
- The system operator monitors the set of buses $N[\mathcal{M}]$ with the first PMU model,

where $N[\mathcal{M}]$ is the closed neighborhood of $\mathcal{M}$ which includes all the buses in $\mathcal{M}$ and their neighboring buses $N(\mathcal{M})$. Thus, the unobservability condition with the second PMU model is obtained by replacing $\mathcal{M}$ with $N[\mathcal{M}]$ in (1). Without loss of generality (WLOG), we employ the first PMU model in the following analysis, and all the results can be directly translated to the second PMU model.

C. Sparsest unobservable attacks

In forming an unobservable attack, an attacker generally has two objectives: minimize execution complexity and maximize its impact on the grid. Note that these two objectives can be competing interests that are not simultaneously achievable. For an attack vector $\Delta\boldsymbol{P}$, we use its zero norm $\|\Delta\boldsymbol{P}\|_0$ to model its complexity. For minimizing attack complexity, an attacker is interested in finding the sparsest attacks that satisfy the unobservability condition (1):

$$\begin{aligned} \min_{\Delta\boldsymbol{\theta}} \ & \|\Delta\boldsymbol{P}\|_0 \\ s.t. \ & \Delta\boldsymbol{P} = \boldsymbol{B}\boldsymbol{\Delta\theta}, \ \Delta\boldsymbol{\theta}_{\mathcal{M}} = \boldsymbol{0}, \ \Delta\boldsymbol{\theta} \neq \boldsymbol{0}. \end{aligned} \qquad (2)$$

Equivalently, a more compact form of (2) is as follows:

$$(2) \ \Leftrightarrow \ \min_{\Delta\boldsymbol{\theta}_{\mathcal{M}^c} \neq \boldsymbol{0}} \|\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c} \Delta\boldsymbol{\theta}_{\mathcal{M}^c}\|_0, \qquad (3)$$

[1] In practice, the second PMU measurement model is achieved by installing PMUs on all the lines connected to a bus.

Because of the non-convexity of the l_0 norm, problem (3) is in general NP-hard. Instead of applying existing heuristics (e.g., l_1 relaxation), in Section III, we will solve (3) by analyzing the network topology and the structure of the Laplacian matrix $\boldsymbol{B}$. We will then characterize the potential impact associated with unobservable attacks.

III. Solving for the sparsest unobservable attacks

In this section, we study the problem of finding the sparsest unobservable attacks given any set of PMUs $\mathcal{M}$ (cf. (3)). We first show an important role of the *vertex connectivity* of the grid in lower bounding the optimum of (3). We then derive an upper bound of (3). By further exploiting the geometric insights behind the upper bound, we close the gap between the lower and upper bounds, and provide a complete solution to the minimum sparsity problem (3). Finally, we characterize the potential impact of unobservable attacks based on their geometric interpretations.

A. The role of vertex connectivity in lower bounding the sparsity of unobservable attacks

We first make the following definitions:

Definition 2 (Vertex cut): A vertex cut of a connected graph $\mathcal{G}$ is a set of vertices whose removal renders $\mathcal{G}$ disconnected.

Definition 3 (Vertex connectivity): The vertex connectivity of a graph $\mathcal{G}$, denoted by $\kappa(\mathcal{G})$, is the size of the minimum vertex cut of $\mathcal{G}$, i.e., it is the minimum number of vertices that need to be removed to disconnect the remaining graph.

We now state the following theorem.

Theorem 1: For a connected power grid $\mathcal{G} = \{\mathcal{N}, \mathcal{L}, \boldsymbol{w}\}$, assume that the line reactances x_l ($l \in \mathcal{L}$) are independent continuous random variables strictly bounded away from zero from below. $\forall M \leq \kappa(\mathcal{G})$, for *any* set of buses $\mathcal{M} \subset \mathcal{N}$, $|\mathcal{M}| = M$, in order to have $\Delta\boldsymbol{\theta}_{\mathcal{M}} = \boldsymbol{0}$, one of the following must be true with probability one:

- There is no power injection attack at any bus in the grid, i.e., $\Delta\boldsymbol{P} = \boldsymbol{0}$; or
- There must be at least $M+1$ buses with non-zero power injections from an attack, i.e., $\|\Delta\boldsymbol{P}\|_0 \geq M+1$.

Theorem 1 provides a lower bound on the optimum of (3) which holds with probability one, i.e., no matter which set of M buses' phase angles are monitored by PMUs,

- if $M \leq \kappa(\mathcal{G})$, there must be at least $M+1$ power injections in any non-zero unobservable attack;
- if $M > \kappa(\mathcal{G})$, there must be at least $\kappa(\mathcal{G})+1$ power injections in any non-zero unobservable attack.

In sum, the minimum sparsity of unobservable attacks is lower bounded almost surely by $\min(\kappa(\mathcal{G}), M)+1$. We will see in the following that while this lower bound is not always tight, a modification of it will render the optimum of (3).

B. An upper bound on the minimum sparsity of unobservable attacks

To derive an upper bound on the minimum sparsity of unobservable attacks, we exploit the fact that solving (3) is equivalent to finding the sparsest non-zero vector in the *range space* of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$. We have the following theorem:

Theorem 2: For a connected grid, the optimum of (3) is upper bounded by $|N(\mathcal{M}^c)|+1 (\leq M+1), \forall 1 \leq M \leq N-1$.

Proof: By re-indexing the buses, WLOG, i) let $\mathcal{M}^c = \{1, 2, \dots, N-M\}$, and ii) let $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ have the following partition as depicted in Figure 1(a):

1) The top submatrix $\boldsymbol{B}_{\mathcal{M}^c\mathcal{M}^c}$ is an $(N-M)\times(N-M)$ full-rank matrix.
2) The middle submatrix (which will be shown to be $\boldsymbol{B}_{N(\mathcal{M}^c)\mathcal{M}^c}$) consists of all the remaining rows each of which has at least one *non-zero* entry.
3) The bottom submatrix is an *all-zero* matrix.

In particular, from the definition of the Laplacian, the middle submatrix of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ as described above is exactly $\boldsymbol{B}_{N(\mathcal{M}^c)\mathcal{M}^c}$ as its row indices correspond to those buses not in $\mathcal{M}^c$ but connected to at least one bus in $\mathcal{M}^c$. Note that the middle and the bottom sub-matrices can be degenerate. Now, we let

$$\Delta\boldsymbol{\theta}_{\mathcal{M}^c} = \boldsymbol{B}^{-1}_{\mathcal{M}^c\mathcal{M}^c}\boldsymbol{e}_1, \tag{4}$$

where $\boldsymbol{e}_1 \in \mathbb{R}^{(N-M)\times 1}$ is the natural basis $[1, 0, \dots, 0]^T$. Then, we construct an attack vector $\Delta\boldsymbol{P} = \boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}\Delta\boldsymbol{\theta}_{\mathcal{M}^c}$: it has a 1 at its index 1, and some possibly non-zero values at the indices that correspond to $N(\mathcal{M}^c)$, but has *zero values at all other indices.* Thus,

$$\|\Delta\boldsymbol{P}\|_0 \leq |N(\mathcal{M}^c)| + 1 \leq M + 1. \tag{5}$$

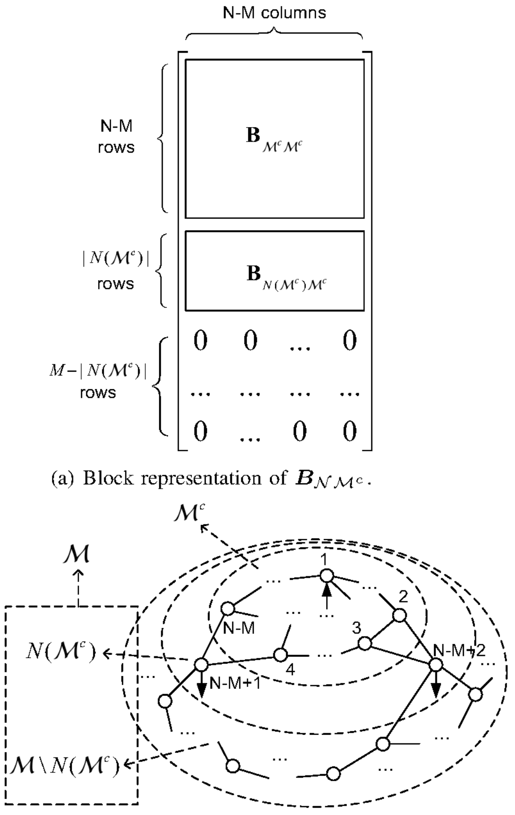
■

From Theorem 1 and 2, we have closed the gap between the lower and upper bounds on the optimum of (3) for the case of $M \leq \kappa(\mathcal{G})$, and proved that the optimum in this case equals $M+1$. To further solve the case of $M > \kappa(\mathcal{G})$, we would like to improve the upper bound $|N(\mathcal{M}^c)|+1$.

We observe that, by selecting a *subset* of the columns of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ to form a new submatrix $\boldsymbol{B}_{\mathcal{N}\mathcal{S}}, \mathcal{S} \subset \mathcal{M}^c$, and partitioning $\boldsymbol{B}_{\mathcal{N}\mathcal{S}}$ into three submatrices in the same way as $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ is partitioned in Figure 1(a), it is possible that the resulting middle submatrix $\boldsymbol{B}_{N(\mathcal{S})\mathcal{S}}$ has *fewer* rows than $\boldsymbol{B}_{N(\mathcal{M}^c)\mathcal{M}^c}$, leading to even sparser unobservable attacks with sparsity $|N(\mathcal{S})|+1$. By further developing this idea based on the geometric insights behind the proof of Theorem 2, we provide a complete solution of (3) next.

C. Closing the gap between lower and upper bounds on the minimum sparsity of unobservable attacks

We start by providing a geometric interpretation of Theorem 2. As shown in Figure 1(a), if $\mathcal{M}\backslash N(\mathcal{M}^c) \neq \emptyset$, all the buses can be partitioned into three subsets $\mathcal{M}^c, N(\mathcal{M}^c)$ and $\mathcal{M}\backslash N(\mathcal{M}^c)$, corresponding to the row indices of the top, middle and bottom submatrices of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ respectively. Moreover, $N(\mathcal{M}^c)$ is a *vertex cut of $\mathcal{G}$ that separates $\mathcal{M}^c$ from $\mathcal{M}\backslash N(\mathcal{M}^c)$*. The $|N(\mathcal{M}^c)|+1$-sparse attack $\Delta\boldsymbol{P}$ (cf. (5)) is formed by injecting power at *one* of the buses among $\mathcal{M}^c$ (i.e., bus 1), and extracting power at and only at the buses in the vertex cut $N(\mathcal{M}^c)$, such that the phase angle changes at $\mathcal{M}$ are all zero.


N-M columns
N-M rows
$\mathbf{B}_{\mathcal{M}^c\mathcal{M}^c}$
$|N(\mathcal{M}^c)|$ rows
$\mathbf{B}_{N(\mathcal{M}^c)\mathcal{M}^c}$
$M-|N(\mathcal{M}^c)|$ rows
0 0 ... 0
... ...
0 ... 0 0
$\mathcal{M}^c$
$\mathcal{M}$
$N(\mathcal{M}^c)$
$\mathcal{M}\backslash N(\mathcal{M}^c)$
1
2
3
4
N-M
N-M+1
N-M+2

(a) Block representation of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$.

(b) A 3-sparse unobservable power injection attack.

Fig. 1. Sparse attacks with the voltage phase angles at the bus $1, 2, \ldots N-M$ affected.

An example that illustrates the above formation of an $|N(\mathcal{M}^c)|+1$-sparse attack is depicted in Figure 1(b). The set $\mathcal{M}^c = \{1, 2, \ldots, N-M\}$ interacts with the set $\mathcal{M}\backslash N(\mathcal{M}^c)$ via only *two* buses — $N(\mathcal{M}^c) = \{N-M+1, N-M+2\}$. As a result, there exists a 3-sparse attack that injects/extracts power at bus 1, $N-M+1$ and $N-M+2$, such that the phase angle changes at $\mathcal{M}$ are all zero.

More generally, for all $\boldsymbol{B}_{\mathcal{N}\mathcal{S}}$, $\mathcal{S} \subseteq \mathcal{M}^c$, the above geometric interpretation applies as well: Given such $\mathcal{S}$, $N(\mathcal{S})$ is a *vertex cut that separates* $\mathcal{S}$ *from* $\mathcal{M}\backslash N(\mathcal{S})$. Therefore, finding the minimum $|N(\mathcal{S})|$ is similar to finding a *minimum vertex cut of a graph, but with a key restriction of* $\mathcal{S} \subseteq \mathcal{M}^c$ that represents the unobservability condition (1). Note that the minimum vertex cut of $\mathcal{G}$ may not be a legitimate $N(\mathcal{S}), \mathcal{S} \subseteq \mathcal{M}^c$ for the following reasons: It can happen that the minimum vertex cut *disconnects* the remaining part of $\mathcal{M}$ such that part of $\mathcal{M}$ is in $\mathcal{S}$, violating $\mathcal{S} \subseteq \mathcal{M}^c$. To ensure the restriction $\mathcal{S} \subseteq \mathcal{M}^c$, we construct a variation of the original grid $\mathcal{G}$:

Definition 4: Given a set of buses $\mathcal{M} \subseteq \mathcal{N}$, $\mathcal{G}^{\mathcal{M}}$ is defined to be the following augmented graph based on $\mathcal{G}$: $\mathcal{G}^{\mathcal{M}}$ has the same set of buses as $\mathcal{G}$, and in addition to all the existing edges of $\mathcal{G}$, an edge is added between every pair of buses in $\mathcal{M}$ with its weight an independent continuous random variable strictly bounded away from zero from below.

We then have the following lemma that ensures the legitimacy of a vertex cut as $N(\mathcal{S})$ for some $\mathcal{S} \subseteq \mathcal{M}^c$:

Lemma 1: For any (including the minimum) vertex cut of the augmented graph $\mathcal{G}^{\mathcal{M}}$, there is no pair of buses in $\mathcal{M}$ that is disconnected by this cut.

From Lemma 1, any vertex cut (corresponding to $N(\mathcal{S})$ for some $\mathcal{S} \subseteq \mathcal{M}^c$) of the augmented graph $\mathcal{G}^{\mathcal{M}}$ partitions $\mathcal{G}^{\mathcal{M}}$ into three subgraphs:

1) $\mathcal{S}(\subseteq \mathcal{M}^c)$ which consists of all the buses that are disconnected from $\mathcal{M}$ after removing the cut set $N(\mathcal{S})$.
2) $N(\mathcal{S})$ which is the vertex cut set itself.
3) $\mathcal{N}\backslash N[\mathcal{S}]$ which contains (not necessarily exclusively) the remaining buses in $\mathcal{M}$ after removing this cut set.

By the same arguments as in the constructive proof of Theorem 2, an $|N(\mathcal{S})|+1$-sparse unobservable attack can be formed by changing the power injections at *one* of the buses in $\mathcal{S}$ as well as at all the buses in the cut set $N(\mathcal{S})$.

Now, note that an attack is unobservable by $\mathcal{M}$ in $\mathcal{G}$ *if and only if* it is unobservable by $\mathcal{M}$ in $\mathcal{G}^{\mathcal{M}}$. We then have the following theorem that closes the gap between the lower and upper bounds of (3):

Theorem 3: For a connected power grid $\mathcal{G} = \{\mathcal{N}, \mathcal{L}, \boldsymbol{w}\}$, assume that the line reactances x_l ($l \in \mathcal{L}$) are independent continuous random variables strictly bounded away from zero from below. Given any $\mathcal{M} \subseteq \mathcal{N}, |\mathcal{M}| = M$ denoting the set of buses with PMUs, the minimum sparsity of unobservable attacks, i.e., the global optimum of (3), equals $\min\big(\kappa(\mathcal{G}^{\mathcal{M}}), M\big) + 1$ with probability one.

Finally, we note that finding the minimum vertex cut of a graph is computationally efficient. Thus, Theorem 3 not only shows that the sparsest attack (3) can be found in polynomial time with probability one, but also crystalizes the geometric interpretation of the optimal solutions. In particular, the constructive solution based on the minimum vertex cut of the augmented graph $\mathcal{G}^{\mathcal{M}}$ is guaranteed to be optimal for solving (3). Next, we will show that this geometric insight leads to a natural characterization of the potential impacts of unobservable attacks.

D. Potential impacts of unobservable attacks

As discussed in the last section, a vertex cut of $\mathcal{G}^{\mathcal{M}}$, denoted by $N(\mathcal{S})$ for some $\mathcal{S} \subseteq \mathcal{M}^c$, partitions $\mathcal{G}^{\mathcal{M}}$ into $\mathcal{S}$, $N(\mathcal{S})$ and $\mathcal{N}\backslash N[\mathcal{S}]$. We now show that, *given any power injection changes in* $\mathcal{S}$, appropriate power injection changes in $N(\mathcal{S})$ exist such that the voltage phase angles in $\mathcal{N}\backslash\mathcal{S}$ ($\supseteq \mathcal{M}$) do not change. We illustrate this for the case of $\mathcal{S} = \mathcal{M}^c$, and generalizations to $\mathcal{S} \subseteq \mathcal{M}^c$ follow immediately.

As in the proof of Theorem 2 (cf. Figure 1(a)), for any power injection changes in $\mathcal{S} = \mathcal{M}^c$, denoted by $\Delta\boldsymbol{P}_{\mathcal{M}^c}$, we let $\Delta\boldsymbol{\theta}_{\mathcal{M}^c} = \boldsymbol{B}^{-1}_{\mathcal{M}^c\mathcal{M}^c}\Delta\boldsymbol{P}_{\mathcal{M}^c}$. Then, we let $\Delta\boldsymbol{P}_{N(\mathcal{M}^c)} = \boldsymbol{B}_{N(\mathcal{M}^c)\mathcal{M}^c}\Delta\boldsymbol{\theta}_{\mathcal{M}^c}$. Finally, we let $\Delta\boldsymbol{P}_{\mathcal{M}\backslash N(\mathcal{M}^c)} = \boldsymbol{0}$ and $\Delta\boldsymbol{\theta}_{\mathcal{M}} = \boldsymbol{0}$. From the block representation of $\boldsymbol{B}_{\mathcal{N}\mathcal{M}^c}$ as in Figure 1(a), the above constructions satisfy

- $\Delta\boldsymbol{P} = \boldsymbol{B}\Delta\boldsymbol{\theta}$, and
- The voltage phase angles in $\mathcal{N}\backslash\mathcal{S} = \mathcal{M}$ do not change.

Consequently, we have the following corollary:

Corollary 1: As long as an attacker takes control of the power injections of all the buses in a vertex cut $N(\mathcal{S})$, it can

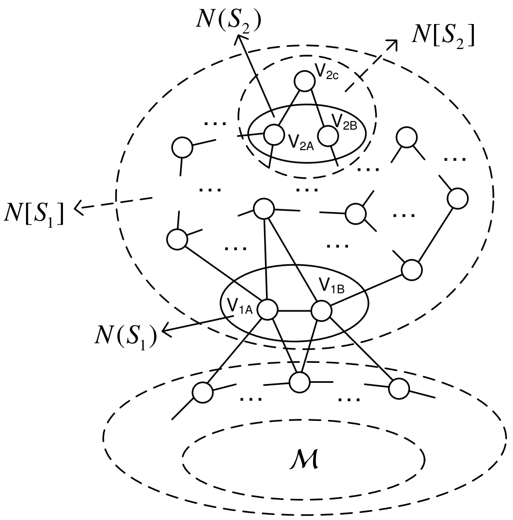


Fig. 2. An illustration of two minimum vertex cuts with the same size but different potential impacts.

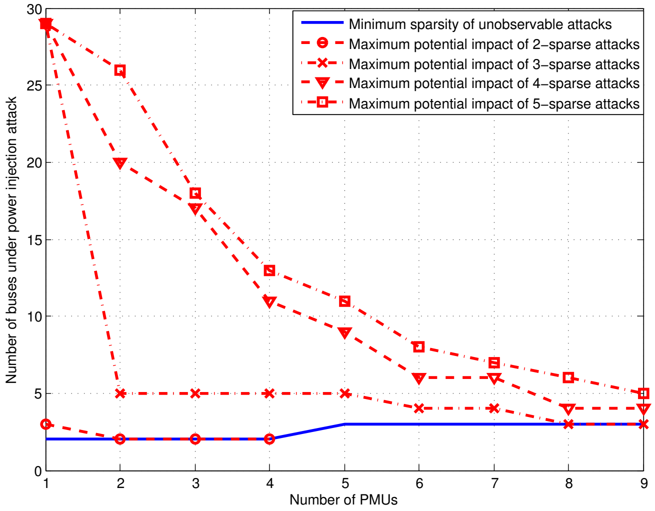


Fig. 3. Minimum sparsity of unobservable attacks and maximum potential impact of 2, 3, 4, 5-sparse attacks as functions of M; IEEE 30-bus system.

always *cancel out* the effects of anything that happens within $\mathcal{S}$ on the measurements by the set of PMUs $\mathcal{M}$ $(\subseteq \mathcal{N}\backslash\mathcal{S})$.

From Corollary 1, by taking control of the buses in a cut $N(\mathcal{S})$, an attacker is able to hide from the system operator a power injection attack with a zero norm as large as

$$|N[\mathcal{S}]| = |N(\mathcal{S})| + |\mathcal{S}| \ (\gg |N(\mathcal{S})| + 1). \tag{6}$$

Accordingly, we define the potential impact of unobservable attacks as follows:

Definition 5: For any vertex cut $N(\mathcal{S})$ for some $\mathcal{S} \subseteq \mathcal{M}^c$, the potential impact of unobservable attacks by controlling power injections at $N(\mathcal{S})$ is $|N[\mathcal{S}]|$.

Employing Definition 5, we can *differentiate* the potential impacts of multiple sparsest unobservable attacks *with the same sparsity*. In particular, multiple minimum vertex cuts can exist for the same augmented graph $\mathcal{G}^{\mathcal{M}}$. Then, each of these cuts leads to a different sparsest unobservable attack of the same size (constructed by controlling the buses in this cut as well as one other bus disconnected from $\mathcal{M}$ by it). However, different cuts may disconnect different portions of the network from $\mathcal{M}$, leading to vastly different potential impacts of unobservable attacks. An illustration is depicted in Figure 2. In this example, two vertex cuts both of size two, $N(\mathcal{S}_1) = \{V_{1A}, V_{1B}\}$ and $N(\mathcal{S}_2) = \{V_{2A}, V_{2B}\}$, are noted as enclosed by solid ovals. Accordingly, both cuts enable 3-sparse unobservable attacks. However, their potential impacts are significantly different. Cut $N(\mathcal{S}_2)$ only disconnects one other bus, namely $\mathcal{S}_2 = \{V_{2C}\}$ from the set of PMUs $\mathcal{M}$, and hence its potential impact equals $|N[\mathcal{S}_2]| = 3$; in comparison, cut $N(\mathcal{S}_1)$ disconnects all the vertices above $N(\mathcal{S}_1)$ from $\mathcal{M}$, and hence its potential impact equals $|N[\mathcal{S}_1]| \gg 3$. With this definition of potential impact, it is then natural for an attacker to *seek the sparsest unobservable attack with the greatest potential impact*.

Finally, we conclude this section by noting the following fact: the minimum sparsity and the potential impacts of unobservable attacks are fully determined with probability one by the *network topology and the locations of the PMUs*.

IV. Numerical evaluation

In this section, we evaluate the sparsest unobservable attacks and their potential impacts when the system operator deploys PMUs at optimized locations. We have seen in Section III that the minimum sparsity and potential impacts of unobservable attacks are determined fully by the network topology and the PMU placement. Unlike network states and network parameters which can vary over short and medium time scales, the transmission network topology (or the set of possible topologies) typically stays the same over long time scales. The above motivates the system operator to optimize the PMU placement according to the network topology. For the best performance in countering power injection attacks, the system operator wants to *raise the minimum sparsity of unobservable attacks, as well as mitigate the maximum potential impact of unobservable attacks*. The geometric interpretations of the sparsest unobservable attacks and their potential impacts allow us to develop an efficient PMU placement algorithm for the system operator to pursue both objectives. The details of the algorithm are omitted here due to space limitations, and can be found in [10].

We evaluate our results in IEEE 30-bus, IEEE 57-bus, IEEE 118-bus, IEEE 300-bus, Polish 2383-bus, Polish 2737-bus, and Polish 3012-bus systems. The evaluation is performed based on the software toolbox MATPOWER [11]. In each of these systems, we generate a set of PMUs based on the developed placement algorithm, with the number of PMUs increasing from one until all attacks become observable. For all numbers of PMUs, the minimum sparsity of unobservable attacks as well as the maximum potential impact among the sparsest unobservable attacks are found.

Specifically, the minimum sparsity of unobservable attacks and the maximum potential impact among these sparsest attacks both as functions of the number of PMUs M are plotted

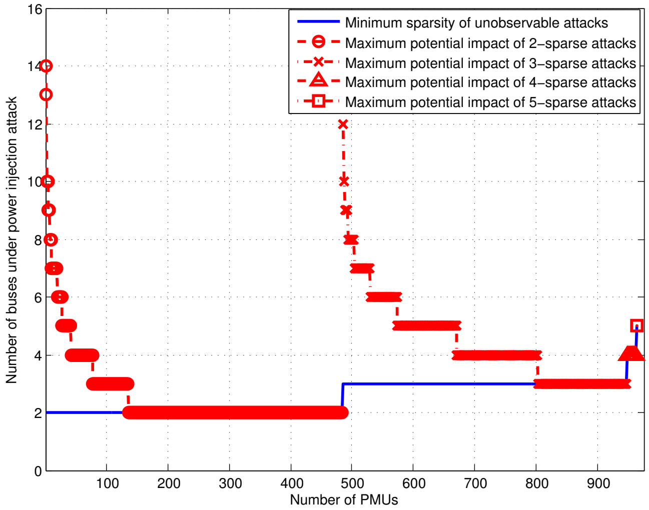


Fig. 4. Minimum sparsity of unobservable attacks and the maximum potential impact of the sparsest attacks as functions of M; Polish 3012-bus system.

for the IEEE 30-bus power system and the Polish 3012-bus system in Figure 3 and 4 respectively. In addition, for the IEEE 30-bus system, the maximum potential impacts among all 2-sparse, 3-sparse, 4-sparse and 5-sparse unobservable attacks for the entire range of M are plotted. (Note that the minimum sparsity of unobservable attacks does not exceed 3 for all M.)

We make the following observations which appear in all seven of the evaluated systems:

- In *all* seven systems, all the attacks become observable with *less than a third* of the buses installed with PMUs (assuming the second PMU model). The average percentage of the number of PMUs needed to have full network observability equals 31.1%. This number resembles a well-known estimate of such percentage to be one third [12].
- The topologies of the tested power systems tend to allow sparse power injection attacks. In other words, the vertex connectivity of these power networks is often small. Furthermore, there are often *many ties* when finding unobservable attacks with the minimum sparsity: this is why even after adding a lot more PMUs into the network, with each addition eliminating the previous sparsest attack, the minimum sparsity can still remain the same.
- While there are many ties of unobservable attacks with the same sparsity, the potential impacts among them can vary significantly. Moreover, as more PMUs are added, the maximum potential impact among all the sparsest unobservable attacks drops quickly until it reaches the minimum sparsity.

V. Conclusion

We have studied cyber-physical attacks that alter power generation and loads in power networks while remaining unobservable under the surveillance of system operators using PMUs. We have provided an explicit solution to the open problem of finding the sparsest unobservable attacks; the minimum sparsity among all unobservable attacks equals $\min\left(\kappa(\mathcal{G}^{\mathcal{M}}), M\right) + 1$. In deriving this minimum sparsity, a lower bound on it based on the vertex connectivity of the network was shown to hold with probability one. We have then provided a constructive upper bound that successfully closes the gap to the lower bound. This constructive upper bound enables us to find all the sparsest unobservable attacks in polynomial time by finding the minimum vertex cuts of an augmented graph $\mathcal{G}^{\mathcal{M}}$. As a result, $\min\left(\kappa(\mathcal{G}^{\mathcal{M}}), M\right) + 1$ is a fundamental limit of this minimum sparsity that is not only explicitly attainable, but also unbeatable by all possible unobservable attacks. We have further shown that the geometric interpretation of the unobservable attacks allows a natural characterization of their potential impacts. With optimized PMU deployment, we have evaluated the sparsest unobservable attacks and their potential impacts in IEEE 30, 57, 118, 300-bus systems and Polish 2383, 2737, 3012-bus systems. Finally, while this work has studied a static system model and power injection attacks, extension to dynamic systems, measurements and power injection attacks remains an interesting future direction, for which we expect that similar insights on fundamental limits will apply.

References

[1] Y. Liu, P. Ning, and M. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, Article 13, May 2011.

[2] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet-based load altering attacks against smart power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 667–674, December 2011.

[3] K. C. Sou, H. Sandberg, and K. H. Johansson, "On the exact solution to a smart grid cyber-security analysis problem," *eprint arXiv:1201.5019v2*.

[4] J. Hendrickx, K. H. Johansson, R. Jungers, H. Sandberg, and K. C. Sou, "An exact solution to the power networks security index problem and its generalized min cut formulation," *eprint arXiv:1204.6174*.

[5] A. Giani, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla, "Smart grid data integrity attacks: characterizations and countermeasures," *Proc. of IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 232–237, October 2011.

[6] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 326–333, June 2011.

[7] J. Weimer, S. Kar, and K. H. Johansson, "Distributed detection and isolation of topology attacks in power networks," *Proc. of the 1st Int. Conf. on High Confidence Networked Systems*, pp. 65–72, July 2012.

[8] F. Pasqualetti, F. Dorfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, to appear.

[9] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *eprint arXiv:1205.5073*.

[10] Y. Zhao, A. Goldsmith, and H. V. Poor, "Fundamental limits of cyber-physical security in smart power grids," *IEEE Transactions on Automatic Control, Special Issue on Control of Cyber-Physical Systems*, 2013, under review.

[11] 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, Feb. 2011.

[12] T. Baldwin, L. Mili, J. M.B. Boisen, and R. Adapa, "Power system observability with minimal phasor measurement placement," *IEEE Transactions on Power Systems*, vol. 8, no. 2, pp. 707–715, May 1993.