Hidden Attacks on Power Grid: Optimal Attack Strategies and Mitigation

Deepjyoti Deka, *Student Member, IEEE,* Ross Baldick, *Fellow, IEEE,* and Sriram Vishwanath, *Senior Member, IEEE*

***Abstract*—Real time operation of the power grid and synchronism of its different elements require accurate estimation of its state variables. Errors in state estimation will lead to sub-optimal Optimal Power Flow (OPF) solutions and subsequent increase in the price of electricity in the market or, potentially overload and create line outages. This paper studies hidden data attacks on power systems by an adversary trying to manipulate state estimators. The adversary gains control of a few meters, and is able to introduce spurious measurements in them. The paper presents a polynomial time algorithm using min-cut calculations to determine the minimum number of measurements an adversary needs to manipulate in order to perform a hidden attack. Greedy techniques are presented to aid the system operator in identifying critical measurements for protection to prevent such hidden data attacks. Secure PMU placement against data attacks is also discussed and an algorithm for placing PMUs for this purpose is developed. The performances of the proposed algorithms are shown through simulations on IEEE test cases.**

I. Introduction

Stable and secure power grid operation relies on accurate monitoring of the state of its different components, including line currents and bus voltages. The state vector is estimated in a power grid by using a variety of measurement units in different buses and lines operating in the grid. These measurements are also used to facilitate operation of the electricity market through locational marginal pricing [3]. Given their importance in the operation of the grid, error free delivery of data from the distributed meters to the central controller for state estimation is a critical need of every power grid. The effect of incorrect data collection had received attention in the 2003 North-East blackout where incorrect telemetry due to an inoperative state estimator was listed as one of its principal causes [1]. Today Supervisory Control and Data Acquisition (SCADA) systems help relay the measurements to the state estimator of the grid for use in stability analysis and OPF solvers. However the presence of distributed meters spread across the entire geographical area covered by the power grid makes the grid vulnerable to cyber-attacks aimed at introducing malicious measurements. In fact, it has been reported that cyber-hacking had previously compromised the U.S. electric grid [2] and can lead to sub-optimal electricity prices [10].

We consider here a scenario where an adversary in the power grid gains control of some meters in the grid and inject malicious data which can lead to incorrect state estimation. In reality, measurements collected have random noise due to measurement errors, but state estimators are able to overcome those through the use of statistical methods like maximum likelihood criterion and weighted least-square criterion [4]. A coordinated attack on multiple meters by an adversary can evade detection by the standard mechanism present at the estimator and go unobservable and not raise any alarm at the state estimator.

Reference [5] studies this problem of hidden attacks on the power grid which are not detected by tests on the residual of the measurements and shows that a few measurements are enough for the adversary to produce a hidden attack. The authors of [5] also show that the set of protected measurements needed to prevent a hidden attack is of the same size as the number of state variables in the grid. Full protection from any hidden attack is thus very expensive. There have been multiple efforts aimed at studying the construction of malicious attack vectors for the adversary. In [8], the authors discuss the optimal attack-vector construction for the constrained adversary using l_0 and l_1 recovery methods. A constrained adversary is governed by its objective to manipulate the measurements of the minimum number of meters to produce the desired errors in estimation. Reference [6] provides an approach for the creation of the optimal attack vector based on mixed integer linear programming. Such design approaches are NP-hard in general and hence require relaxations of the problem statement and provide approximate solutions at best. In addition, previous work such as that in [9] requires certain assumptions on states of the system which may not hold in general.

In this paper, we consider an adversary with constrained resources. Following the attack model in [8], we define the objective of the adversary as identifying the minimum number of meters that may be manipulated in order to create a hidden attack vector using those meters. We use graph-theoretic ideas like min-cut calculations in determining this optimal attack vector. Unlike previous work, we show that our solution does not require any assumption on the structure of the grid or any relaxation of the problem statement. The complexity of the algorithm for attack vector construction is shown to be polynomial in the number of nodes (buses) and edges (lines) in the power grid. Given the size of large power grids, polynomial running time of the algorithm justifies its significance when compared with NP-hard and brute force methods used in the existing literature.

In addition, the algorithm for attack vector construction does not depend on the exact values of the measurement matrix used in state estimation, relying primarily on the

Deepjyoti Deka, Ross Baldick and Sriram Vishwanath are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail:deepjyotideka@utexas.edu; baldick@ece.utexas.edu; sriram@ece.utexas.edu)

adjacency matrix of the network representing the power grid. This is significant, since the adjacency matrix of grids is often already known or can be approximated by an adversary from publicly available information. Using the novel graph theoretic framework discussed in this paper that requires only the adjacency matrix of the graph representing the power grid, the adversary can, thus, generate the optimal attack vector for malicious hidden attacks on the grid using a polynomial time algorithm.

We demonstrate that the power grid is significantly vulnerable to hidden attacks from even constrained adversaries with limited information (adjacency matrix in our case). We are unaware of any existing work providing optimal solution to this hidden attack problem in polynomial time. The framework developed is easily extended to formulate attack vectors in cases when certain measurements in the system are already protected or the state vector is partially known, i.e., some of the state variables are protected. We also provide algorithms to select, in a greedy fashion, critical measurements for protection to prevent a hidden attack given the prior knowledge of the adversary's resources (maximum number of measurements it can corrupt). Present power grids have additional meters called phasor measurement units (PMUs) installed on a few buses [14],[15]. Placement of a PMU at a given bus in the power grid provides measurement of the voltage phasor at that bus as well as the current flows of all lines incident on that bus. We discuss PMU placement in power grids in the context of hidden data attacks in detail in a separate section. The problems discussed in this context include designing an optimal attack vector for a grid equipped with PMUs, and the selection of locations for placement of PMUs in the grid in order to provide increased protection to the grid.

The main results of this paper are as follows:

- We provide a graph theoretic formulation for the problem of constructing an adversarial optimal attack vector with the minimum number of non-zero values that results in a hidden attack on the grid. Further, we present a algorithm that results in an exact solution to the problem and prove its optimality and polynomial-time complexity.
- We discuss different variations of the adversarial attack problem, given the knowledge of certain prevailing protected measurements and state variables within the grid. We extend this discussion to power systems with PMUs installed at a few buses.
- We present greedy algorithms to select additional protected measurements and locations for placement of PMUs in the grid, in order to hinder an adversarial hidden attack on the grid.
- We study the performance of the algorithms through simulations on IEEE test bus systems and compare them with other algorithms in literature, as well as with brute force techniques.

The rest of this paper is organized as follows. The next section presents a description of the system model used in estimating the state variables in a power grid and formulation of the adversarial attack problem. The novel algorithm to determine the optimal solution to this problem, which includes selection of the measurements to attack and generating the attack vector, is discussed in Section III. Construction of the attack vector in the presence of protected measurements and state variables in the system is discussed in Section IV. Algorithms to select existing measurements in the system for protection to prevent hidden attacks on the grid are provided in Section V. Power grids with installed PMUs are discussed in Section VI wherein we design attack vectors for systems with PMUs and also analyze the placement of additional PMUs against hidden attacks for such systems. Simulations of the proposed algorithms on test IEEE bus systems and comparisons with other algorithms and brute force methods are reported in Section VII. Finally, concluding remarks and future directions of work are presented in Section VIII.

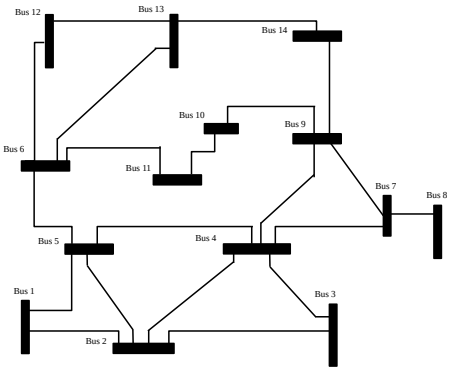

Fig. 1. IEEE 14-bus test system [11]

II. Estimation in the Power Grid and Attack Models

We represent the power grid using an undirected graph (V, E), where V represents the set of buses and E represents the set of transmission lines connecting those buses. There are two kinds of measurements in the power grid in this model: flow measurements and voltage phasor measurements. Meter on a line in E measure the power flow through that line while a meter on a bus in V measure the voltage phasor at that bus. In addition, a PMU can collect both these kinds of measurements (bus voltage and current flows in all lines connected to that bus). This model is sufficiently general and can include cases where some bus voltages or line measurements are measured multiple times for redundancy. Figure 1 shows the graph representation of the IEEE 14 bus test system. It can be found at [11].

We consider the DC power flow model in a power grid

$$z = Hx + e \tag{1}$$

where $z \in \mathbb{R}^m$ is the vector of measurements. $x \in \mathbb{R}^n$ is the state vector and consists of the voltage phase angles at the buses in the grid. The voltage magnitudes are taken as unity in the DC model. H is the measurement matrix which relates the measurements with the state vector and e is the zero mean Gaussian noise vector associated with the measurements. In general, $m > n$ which implies that

there are more measurements than state variables to help provide redundant measurements to the state estimator. Here, H depends on the topology of the network, the location of the meters as well as on the parameters of transmission lines (like resistance and susceptance). The DC power flow on a line $(i,j) \in E$ between buses i and j is given by $B_{ij}(x_i - x_j)$, where B_{ij} is the magnitude of susceptance of the line (i,j). If the k^{th} measurement corresponds to this flow in line (i,j), then z_k is given by:

$$z_k = H_k x = B_{ij}x_i - B_{ij}x_j. \tag{2}$$

The associated row in the measurement matrix (H_k) is a sparse vector with two non-zero values, B_{ij} at the i^{th} position and $-B_{ij}$ at the j^{th} position.

$$H_k = [0..0 \quad B_{ij} \quad 0..0 \quad -B_{ij} \quad 0..0] \tag{3}$$

If z_m measures the voltage phase angle at bus i (the voltage magnitude is considered unity in the DC power flow model), the corresponding row H_m in the measurement matrix is given by:

$$z_m = H_m x = x_i \tag{4}$$

$$\Rightarrow \quad H_m = [0..0 \quad 1 \quad 0..0] \tag{5}$$

Here H_m is a sparse vector with one in the i^{th} position and zero everywhere else. The measurement matrix is, thus, very sparse with a maximum of 2 non-zero values per row. In order to enable correct state estimation, the measurement matrix must have full column rank of n. The state estimator uses the measurements and outputs the estimated state vector $\hat{x}$ by minimizing the residual $\|z - H\hat{x}\|_2$. In normal operation, the magnitude of the minimum residual is smaller than an established threshold which monitors the correctness of the estimate.

Given such an estimator, the adversary corrupts the measurement vector z by adding an attack vector a to generate the new measurements of the form $\hat{z} = z + a$. It is fairly straightforward ([5]) that, if a satisfies

$$a = Hc \tag{6}$$

for some $c \in \mathbb{R}^n$, then the residual calculated by the estimator remains the same as before. Therefore the estimator remains incapable of detecting the presence of an attack, outputting an erroneous state vector estimate $\hat{x} + c$.

Attack vector construction refers to the creation of an optimal attack vector a to corrupt the measurements. Reference [5] creates such attack vectors by using a projection matrix P using the measurement matrix H. A major problem of interest here is the case of a constrained adversary with limited resources. Such an adversary attacks the minimum number of measurements to create a successful hidden attack. Here, an attack is considered successful if at least 1 state variable suffers a change in magnitude after the attack. Equivalently, the construction of the attack vector a^* can be characterized as a solution of the following optimization problem :

$$\begin{aligned} &\min_{a} \|a\|_0 \\ &\text{s.t.} \quad a = Hc, \ c \neq \vec{0} \end{aligned} \tag{7}$$

This is similar to the attacker's problem in [8], where the constraint $c \neq \vec{0}$ is replaced by the constraint $\|c\|_\infty \geq \tau$. Both these problem formulations are essentially the same as every solution of Problem (7) can be suitably scaled to result in a solution obtained in [8]. For every non-zero c, $c\tau/\|c\|_\infty$ satisfies the constraint $\|c\|_\infty \geq \tau$.

In the next section, we present our algorithm for designing an optimal attack vector using graph theory. Unlike [8], we do not need the exact matrix H for our solution, but only the locations of 1s in it.

III. Optimal Attack Vector Design

Consider the m times n measurement matrix H as noted in (3) and (5). It is sparse and every row has either 1 (corresponding to phase angle measurement or 2 (corresponding to line flow measurement) non-zero elements. We first augment one extra column h^g to the right of the matrix H to create a m times $(n+1)$ modified measurement matrix $\hat{H}$ such that for every row H_m with a phase angle measurement, the new column h^g has a value of -1 at the m^{th} location.

$$\hat{H}_m = [H_m \mid -1] \tag{8}$$

For a line flow measurement H_k, the corresponding element in the new column is 0.

$$\hat{H}_k = [H_k \mid 0] \tag{9}$$

The state vector c is now augmented to form a new state vector $\hat{c} = \begin{bmatrix} c \\ 0 \end{bmatrix}$ which has 0 as the final element. We now have

$$a = Hc = \hat{H}\hat{c} = [H \mid h^g] \begin{bmatrix} c \\ 0 \end{bmatrix} \tag{10}$$

Equation (10) holds as the last element of $\hat{c}$ is 0. The new formulation $(\hat{H}, \ \hat{c})$ provides a reference phase angle of 0 (represented by the extra element in $\hat{c}$) such that phase angle measurement at a bus becomes equivalent to a line flow measurement between the bus and reference phase angle. Note that every row in $\hat{H}$ has 2 non-zero elements and corresponds to a line flow measurement now. Next, we state and prove a theorem which will enable us to develop the algorithm for attack vector.

Theorem 1. *There exists a non-zero binary $0-1$ vector c_{opt} of size n times 1 for the optimal attack vector a^* given by Problem (7) such that $\|a^*\|_0 = \|Hc_{opt}\|_0$.*

Proof: Consider Problem (7). Let the optimal attack vector be given by $a^* = Hc^*$. If c^* is a $0-1$ vector, take $c_{opt} = c^*$ and the theorem is trivially true. If c is not a $0-1$ vector, construct n times 1 vector c_{opt} such that $c_{opt}(i) = \mathbf{1}(c^*(i) \neq 0)$, $\forall i \in \{1, n\}$. Consider $\|Hc_{opt}\|_0$. For every non-zero phase angle measurement in Hc^*, we have a corresponding non-zero measurement in Hc_{opt}. However, in case of a non-zero value of line flow measurement in Hc^* between two neighboring buses i and j with $c^*(i) \neq 0, c^*(j) \neq 0$, the corresponding value of Hc_{opt} is 0 as $c_{opt}(i) = c_{opt}(j) = 1$. It, thus, follows from the structure of the H matrix that $\|Hc_{opt}\|_0 \leq \|Hc^*\|_0$. Since $a^* = Hc^*$

is the optimal attack vector with minimum number of non-zero entries, we have $\|Hc_{opt}\|_0 = \|Hc^*\|_0$. Thus, Hc_{opt} also gives an optimal attack vector and $\|a^*\|_0 = \|Hc_{opt}\|_0$. ∎

Now, consider the modified measurement matrix $\hat{H}$ and the augmented vector $\hat{c}$. Using Equation (10) and the previous theorem, we conclude that the optimal attack vector is given by $a^* = \hat{H}\hat{c}_{opt}$ where $\hat{c}_{opt}$ is the $0-1$ vector given by $\hat{c}_{opt} = \begin{bmatrix} c_{opt} \\ 0 \end{bmatrix}$.

Minimizing the number of measurements needed by the adversary to inject the optimal attack vector is equivalent to minimizing the number of non-zero flow measurements given by $\hat{H}\hat{c}$. Since we are concerned only with $\|a\|_0$ and $\hat{c}$ is a $0-1$ vector, we observe that the exact values of susceptance present in $\hat{H}$ are not needed to get the optimal attack vector. In fact, the contribution of a flow measurement of $\hat{H}\hat{c}$ in $\|a\|_0$ will remain the same even if the susceptance B_{ij} of every line included in $\hat{H}$ is changed to 1. We, therefore, create an incidence matrix A_H of dimension m x $(n+1)$ from $\hat{H}$ by replacing every positive element in $\hat{H}$ with a 1 in A_H and each negative element in $\hat{H}$ with to a -1 in A_H. Zeroes are left unchanged. The $(i,j)^{th}$ elements of $\hat{H}$ and A_H are related as

$$A_H(i,j) = 1(\hat{H}(i,j) > 0) - 1(\hat{H}(i,j) < 0) \qquad (11)$$

The main result of this paper which gives the minimum attack vector for the optimization Problem (7) is given in the following theorem involving A_H.

Theorem 2. *The cardinality of the optimal attack vector in Problem (7) with measurement matrix H is equal to the min-cut of the undirected graph of $n+1$ nodes and edges defined by the incidence matrix A_H.*

Proof: From the discussion above, it is clear that for a given $0-1$ vector $\hat{c}$, $\|\hat{H}\hat{c}\|_0 = \|A_H\hat{c}\|_0$. Further, using Theorem 1, the optimization Problem (7) can be written as

$$\begin{aligned} &\min_a \|a\|_0 \qquad (12)\\ \text{s.t. } &a = A_H\hat{c},\ \hat{c} \neq \vec{0} \\ &\hat{c} \text{ is a } 0-1 \text{ vector with } (n+1)^{th} \text{ element } 0 \end{aligned}$$

This is the classical min-cut partition problem in graph theory. The minimum value of $\|q\|_0$ is, thus, given by the magnitude of the min-cut of the undirected graph with A_H as the incidence matrix. ∎

Note that multiple measurements of the same line-flow or phase angle will lead to multiple edges between two nodes in the associated graph. Formally, after pre-processing the initial measurement matrix H to generate $\hat{H}$ and A_H, the optimal attack vector and its cardinality is given by Algorithm 1.

The optimal attack vector consists of the edges in the min-cut and produces a non-zero change in the estimate of the state variables at the nodes which are on the opposite side of the min-cut as the reference node. The resulting attack vector is indeed optimal as its cardinality is equal to the min-cut. The min-cut computation is a well-studied problem in graph theory and has a running time polynomial in the number of nodes and edges in the graph [12]. Reference

Algorithm 1 Optimal Attack Vector (a^*) through Min-Cut

Input: Graph G_H with incidence matrix A_H

```
Compute the min-cut of the graph G_H
c ← 1
Choose (n+1)^th node as root
Remove min-cut edges
Do breadth first path traversal from root
if node i is reached then
    c(i) ← 0
end if
a* ← Hc
```

[13] gives a simple algorithm for computing the min-cut in $O(|V|log|V| + |E|)$ time-steps. Here, $|V|$ and $|E|$ represent the number of nodes and edges in the graph considered. The algorithm presented above, to find the optimal attack vector, has the following distinguishing characteristics which separate it from other algorithms in literature:

- It finds the optimal solution of the optimization Problem (7) without using a relaxation
- It is polynomial-time solvable
- It does not require the exact values of the line susceptance in the grid, using instead, the locations of the measurements in the network.

Next, we show how the Algorithm 1 can be used to design the optimal attack vector in the presence of protected measurements and state variables in the system.

IV. Optimal Attack Vector Construction with Protected Measurements and State Variables

Certain measurements in the power grid are protected from cyber-attacks by encryptions or by geographical isolation. This imposes a constraint on the adversary by requiring that the values of the attack vector a corresponding to protected measurements be made 0. Similarly, certain state variables might be protected from adversarial contamination due to the presence of secure channels of collecting their values. Let S_m be the set of protected measurements and S_v be the set of protected state variables. The optimal attack vector a^* here can be written as the solution of the following optimization problem:

$$\begin{aligned} &\min_a \|a\|_0 \qquad (13)\\ \text{s.t. } &a = Hc,\ c \neq \vec{0} \\ &H^{S_m}c = 0,\ c(i) = 0\ \forall i \in S_v \end{aligned}$$

where H^{S_m} represents the rows in the measurement matrix corresponding to the protected measurements. Here, the adversary needs to ensure that the vector c has values 0 for the protected state variables, while the vector a has values 0 for the protected measurements.

Following Problem (7), we consider the modified measurement matrix $\hat{H}$ (given by Equations (3) and (5)) and the augmented state vector $\hat{c} = \begin{bmatrix} c \\ 0 \end{bmatrix}$ by adding the reference node. We first obtain the incident matrix A_H from the modified

measurement matrix $\hat{H}$ as per Equation (11). We denote the graph represented by the incident matrix A_H as G_H. Every measurement in A_H leads to an edge in G_H of unit weight. The additional constraints due to the protected measurements and state variables are incluuded in the graph G_H through the following modification as follows:
1. Create an edge of infinite weight between the buses with protected state variables in S_v and the reference node.
2. Change the weights of edges with protected measurements in S_m to infinity.

The resultant graph generated from G_H after this modification is denoted by G_H^*. We, now, run the steps outlined in Algorithm 1 on G_H^* to obtain the optimal attack vector as described in the previous section. We call this Algorithm 2 for completion.

Algorithm 2 Optimal Attack Vector (a^*) with protected measurements and state variables

Input: Graph G_H, protected state variables S_v and measurements S_m

1: Modify G_H to generate G_H^*
2: Run Algorithm 1 on G_H^*

In the solution of Algorithm 2, the modified edges (with infinite weight) are not included in the attack vector given by the min-cut to keep the value of the min-cut below infinity. This ensures that the modification of G_H to G_H^* satisfies the constraints arising due to protection and gives the optimal solution.

l_1 **Relaxation:** Problem (13) can also be relaxed and approximately solved using a naïve l_1 relaxation by replacing the non-convex l_0 terms with l_1 terms [16]. However, such an approach leads to attack vector solutions with large cardinality which are sub-optimal. To go around that, we use thresholds in the formulation shown below to solve Problem 13:

$$\begin{aligned} &\min_{a} \|a\|_1 \\ \text{s.t. } &a = Hc,\ c \geq \vec{0},\ 1^T c > \theta_1 \\ &H^{S_m} c = 0,\ c(i) = 0\ \forall i \in S_v \end{aligned} \tag{14}$$

The final attack vector is obtained by thresholding the optimal solution $a^*(i) = 1(a^*(i) > \theta_2), \forall\ 1 \leq i \leq m$, where m is the length of the attack vector. Here, θ_1 and θ_2 are thresholds used to decrease the cardinality of the solution attack vector. θ_1 and θ_2 are taken as 1 and 10^{-3} respectively in our simulations in Section VII.

Until now, we have discussed the adversary's strategy for designing attack vectors towards causing hidden data attacks on the grid. We will use this knowledge in the next section to discuss policies that can be adopted by the power grid system operator or controller to restrict the efficacy of hidden attacks.

V. Protection Strategies against Hidden Attacks

Let us consider the system described in Problem (13). Here, there are pre-existing secure measurements (set S_m) and state variables (set S_v) in the grid. The corresponding set of unprotected measurements and unprotected state variables will be termed S_m^c and S_v^c respectively. For complete protection against hidden attacks, it has been shown in [5] that H^{S_m} should have full column rank. In that case, the only c satisfying the constraint $H^{S_m} c = 0$ is the all-zero vector. However, for full column rank of n, the number of protected measurements needs to be greater than n [5], and incurs a great cost. Instead, we look at the problem of augmenting the set of protected measurements S_m with k measurements selected from the unprotected set S_m^c. Protecting additional measurements leads to an increase in the cardinality of the optimal attack vector $\|a^*\|_0$. This increases the number of compromised measurements needed by the adversary for a successful attack. We formulate this problem as follows:

$$\begin{aligned} &\max_{S^* \in S_m^c} \min_{a} \|a\|_0 \\ \text{s.t. } &a = Hc,\ c \neq \vec{0} \\ &H^{S_m} c = 0,\ H^{S^*} c = 0 \\ &c(i) = 0\ \forall\ i \in S_v,\ |S^*| = k \end{aligned} \tag{15}$$

The set of new protections S^* of cardinality k is then used to update the protected set S_m. As mentioned earlier, S_m, S_v, S_m^c and S_v^c represent the sets of protected measurements, protected state variables, unprotected measurements and unprotected state variables in the grid respectively.

Protecting optimal k additional measurements is equivalent to increasing the weights of k edges in the modified graph G_H^* (outlined in Section IV) to infinity to maximally increase the value of the min-cut. This is a NP-hard problem. A brute- force selection of measurements for protection is computationally intensive and impractical given the large number of candidate measurements in the set S_m^c in a real power grid. Hence, we provide here a greedy approach for Problem 15 in Algorithm 3. Here, S_m is updated in k steps. At each step, the best candidate is chosen in a greedy fashion for protection given a^*, the current optimal attack vector. After including a measurement in the protected set S_m, a^* is updated and used for selecting the next candidate measurement for protection.

Step 4 of Algorithm 3 retains only the measurements represented by the current min-cut of G_H^* as candidates for the next update in S_m. It ignores measurements outside the current min-cut as protecting them does not lead to an increase in the size of the min-cut of the updated graph. This step, thus, leads to an reduction in the number of possible candidates in each step from $m - |S_m|$ to $\|a\|_0$ without any loss of performance. The Algorithm is of course sub-optimal compared to a computationally intensive brute force search of the best measurements for protection.

VI. Power Systems with PMUs: Attacks and Protection

In this Section, we extend the ideas developed in the previous sections to power grids with Phasor Measurement Units (PMUs). A PMU located at a bus in the grid measures its voltage phasor as well as the current flows of all lines incident on that bus [7]. Previous work on PMU placement against hidden measurement attacks [8], [9] assume full protection of

Algorithm 3 Greedy Solution for Additional Protection

Input: Graph G_H^*, attack vector a^*, protected set S_m
Output: Updated G_H^*, a^* and S_m

for $i = 1$ **to** k **do**
  $a_{cm} \leftarrow a^*$
  for $j = 1$ **to** m {m: total measurements } **do**
    if $a^*(j) \neq 0$ **then**
      $G_{temp} \leftarrow G_H^*$
      Protect measurement j in G_{temp}
      Compute optimal attack vector a_{temp} for G_{temp}
      if $\|a_{temp}\|_0 \geq \|a_{cm}\|_0$ **then**
        $cm \leftarrow j$ {current best candidate}
        $a_{cm} \leftarrow a_{temp}$ {current optimal attack vector}
      end if
    end if
  end for
  Protect measurement cm and update G_H^*
  $a^* \leftarrow a_{cm}$, $S_m \leftarrow S_m \cup \{cm\}$
end for

the PMU's measurements. Recently, it has been shown that PMUs do not have full security as they rely on civilian GPS signals for real-time signalling that can thus be corrupted by GPS spoofers [17]. Therefore, we consider both protected and unprotected PMU measurements here.

A. Optimal Attack Vector for Grids with PMUs

The bus phase angles and line flow measurements calculated by the unsecured PMUs are equivalent to other measurements in the grid. We assume here that any measurement in an unsecured PMU can be independently corrupted by an adversary. For secure PMUs, we consider all measurements recorded by them as protected and include them in the protected set S_m and the protected bus phase angles in the set of protected state variables S_v. The attack vector is then given by running Algorithm 2. The optimality of the attack vector follows from the discussion for a general grid with protected measurements given in Sections III and IV.

B. Protection against hidden attack by placing secure PMUs

In this case, we consider secure PMUs such that their measurements are protected against any malicious attack. Each PMU placed at a bus thus creates protected measurements of the bus phase angle and incident line flows on that bus. Optimal Placement of secure PMUs to ensure full protection of all state variables against any adversary is equivalent to a set-cover problem and is NP-hard in general. However, approximate and distributed algorithms based on belief propagation have been shown to provide optimal PMU placement for several IEEE test systems [18]. Here, instead of full protection, we look at the problem of placing additional k secure PMUs to maximally hinder a hidden attack by the adversary. We consider existing protected measurements and protected state variables in the grid and denote them by S_m and S_v respectively. It is worth noting that k PMUs might not be sufficient to provide full protection to the entire state vector. Thus, we look at maximizing the cardinality of the optimal attack vector a^* of the adversary instead. As discussed in Section V, this is done by maximizing the min-cut of the modified graph G_H^* associated with the measurement matrix H and protected sets S_m and S_v. We modify Algorithm 3 and provide a greedy algorithm, Algorithm 4, to determine k bus locations, one at a time for placing secure PMUs. This greedy algorithm runs k times and thus has a small complexity compared to a brute force search. The performance of the algorithm on IEEE test cases is reported in the following section.

Algorithm 4 Greedy Solution for k secure PMU placement

Input: Graph G_H^*, attack vector a^*, protected set S_m
Output: Updated G_H^*, a^*

for $i = 1$ **to** k **do**
  $a_{cm} \leftarrow a^*$
  for $j = 1$ **to** n {m: total buses} **do**
    $G_{temp} \leftarrow G_H^*$
    Place PMU at bus j in G_{temp}
    Compute optimal attack vector a_{temp} for G_{temp}
    if $\|a_{temp}\|_0 \geq \|a_{cm}\|_0$ **then**
      $cm \leftarrow j$ {current best candidate bus}
      $a_{cm} \leftarrow a_{temp}$ {current optimal attack vector}
    end if
  end for
  Place PMU at bus cm and update G_H^*
  $a^* \leftarrow a_{cm}$
end for

VII. Simulations on IEEE test systems

In this section, we evaluate the performance of our proposed algorithms by simulating their performance on different IEEE test bus systems, namely 14-bus, 30-bus and 118-bus systems. Data about these test systems can be found at [11]. All simulations are run in Matlab Version 2009a. We start by discussing the performance of Algorithms 1 and 2 that are used for constructing the optimal attack vector (a^*) of the adversary. We take IEEE-14 bus system and place flow measurements in all lines and voltage measurements on random 60% of the buses. Figure 2 shows the increase in the average size of the optimal attack vector with increase in the fraction of randomly protected measurements placed in the system. We see that plots generated by simulations of our algorithms and that obtained through brute force search for the optimal attack vector overlap completely, depicting optimal performance. It can be seen from the same figure that the performance of our algorithms are much better than the output of the l_1 relaxation given in Problem (14). Next, we plot the output of Algorithm 2 and show its improved performance over the output of a l_1 relaxation approach for 30, 57 and 118 IEEE test bus systems under different system conditions in Figure 3. The output of l_1 relaxation is not close to optimal as commendable performance of $l_0 - l_1$ solvers requires the measurement matrix

to satisfy certain necessary conditions [16]. Such conditions are difficult to satisfy for test-systems where the network and the measurement matrices are not random, as in our case.

We now present results on our approach to Problem (15), which is given by Algorithm 3. Here, we select, in a greedy fashion, the k best measurements for protection such that the minimum number of measurements needed for a successful hidden attack increases the most. Figure 4 shows the performance of our greedy Algorithm 3 for different values of k for the IEEE-14 bus system with flow measurements on all lines, voltage measurements on 60% of the buses and $1/6$ of measurements initially protected. We observe that the performance of the greedy algorithm in increasing the size of the optimal attack vector is comparable to a computationally intensive brute-force selection for protecting additional measurements in this case. We also simulate Algorithm 3 for IEEE 30, 57 and 118 bus systems and plot the average improvement in the cardinality of the optimal attack vector with an increase in the value of k in Figure 5. It is important to note that the minimum cardinality of the optimal attack vector a^* does not increase significantly with a small increase in k or fraction of protected measurements in all the different test systems considered. This observation can be explained using the fact that the test-systems are sparse and have several buses with low degree and thus have a low min-cut. Determination of optimal attack vectors in the presence of secure PMUs for the IEEE 30 and 57 bus systems is shown in Figure 6. In either test system, we place line flow measurements in each bus and phase angle measurements in 60% of the buses. We observe an expected increase in the average size of the optimal attack vector on increasing the fraction of buses randomly selected for placement of secure PMUs. Finally, we show the performance of Algorithm 4 in the IEEE 30-bus system. In the base case, we put line flow measurements in each bus of the system, phase angle measurement in random 60% of the buses and protect $1/10$ of the measurements selected randomly. Algorithm 4 is used to place k additional PMUs on buses to increase the cardinality of the optimal attack vector for the system. We observe again that enough secure PMUs need to be placed in the grid to significantly increase the cardinality of the optimal attack vector as high sparsity of the network graph and low degrees of the nodes keep the graph min-cut low.

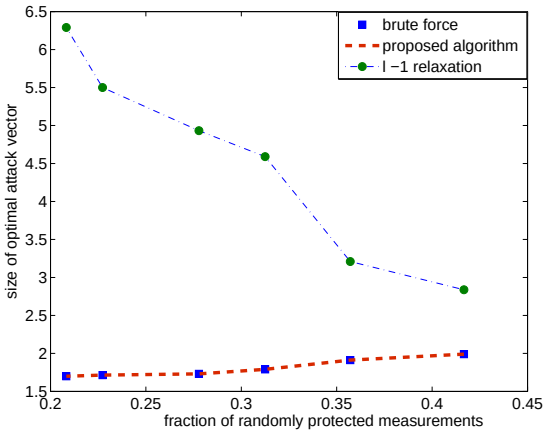

Fig. 2. Optimal hidden attack on IEEE 14-bus system with flow measurements on all lines, voltage measurements on 60% of the buses and fraction of measurements randomly protected

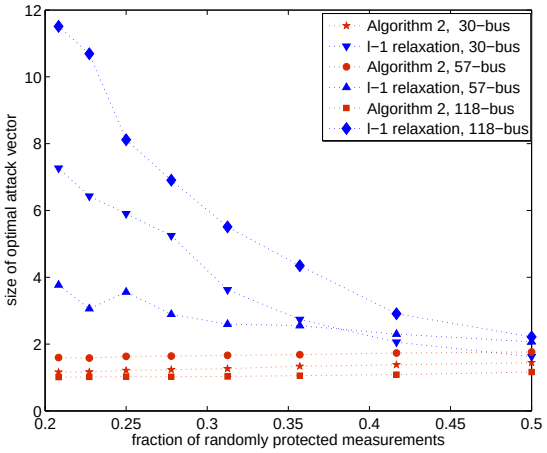

Fig. 3. Optimal hidden attack on IEEE test systems with flow measurements on all lines, voltage measurements on 60% of the buses and fraction of measurements randomly protected

VIII. Conclusion

In this paper, we study an adversarial problem of causing errors in estimation of state variables in a power grid through injection of suitably hidden measurement errors. We formulate this problem in terms of controlling and manipulating a minimum set of measurements in order to affect a successful hidden attack as an l_0 optimization problem. We introduce a novel graph-theoretic approach to designing the optimal attack vector using min-cuts. The proposed algorithm has polynomial time complexity and is shown to result in an optimal output given a configuration of the power grid. We show that our algorithm gives the optimal output even when a fraction of the measurements have existing protection and performs much better than a l_1 relaxation of the problem. From the system operator's perspective, we develop an algorithm to identify measurements in the system that provide additional protection, aimed at preventing and/or reducing the efficacy of hidden attacks by an adversary. Although sub-optimal, the low complexity algorithm can be used to protect measurements to increase the set of measurements that the adversary must control in order to cause a successful hidden attack on the system. Further, we extend the discussion on hidden attacks in the grid to systems with PMUs and discuss design of optimal attack vector for a system with PMUs and placement of additional secure PMUs in the system to prevent such attacks. The advantage of using low complexity algorithms to provide security against hidden attacks is immense for large power grids with several thousand buses and lines. This work can be extended to include other hidden attacks where the adversary is not limited by number of attacked meters but other resources. Another extension includes determining the minimum set of key measurements for protection by the system operator given the knowledge of the adversary's maximum capacity to attack the power grid. This is the focus of our current work.

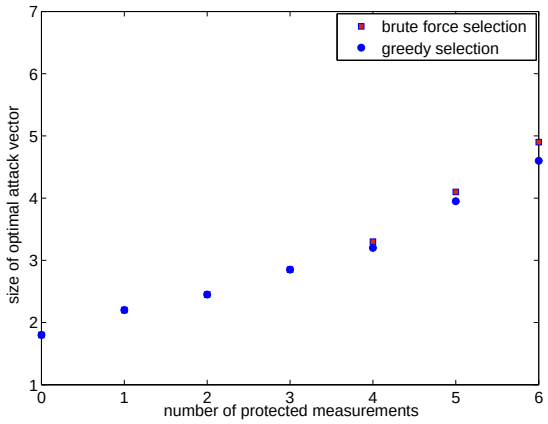

Fig. 4. Protection of additional measurements in IEEE 14-bus system with flow measurements on all lines, voltage measurements on 60% of the buses and $1/6$% of measurements initially protected

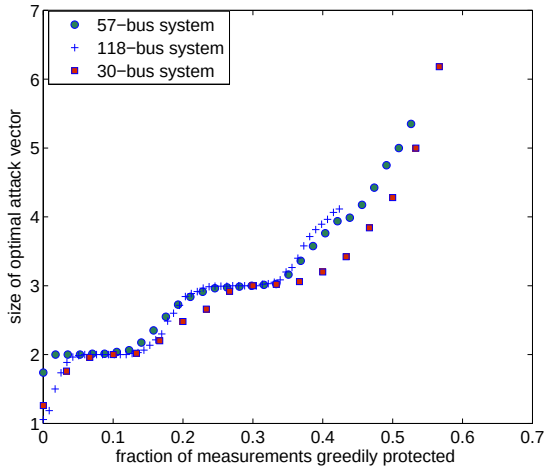

Fig. 5. Greedy protection of additional measurements in IEEE test systems with flow measurements on all lines, voltage measurements on 60% of the buses and $1/6$ of measurements initially protected

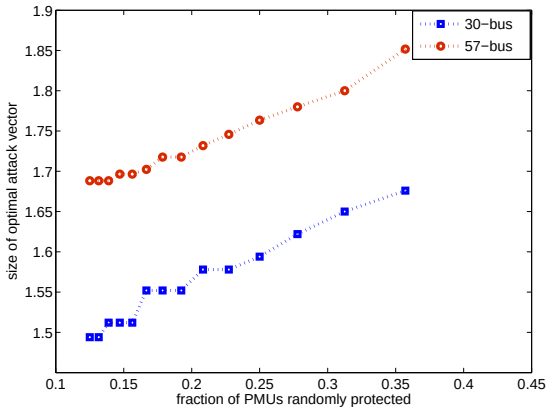

Fig. 6. Optimal hidden attack on IEEE test systems with flow measurements on all lines, voltage measurements on 60% of the buses and PMUs placed randomly on fraction of buses

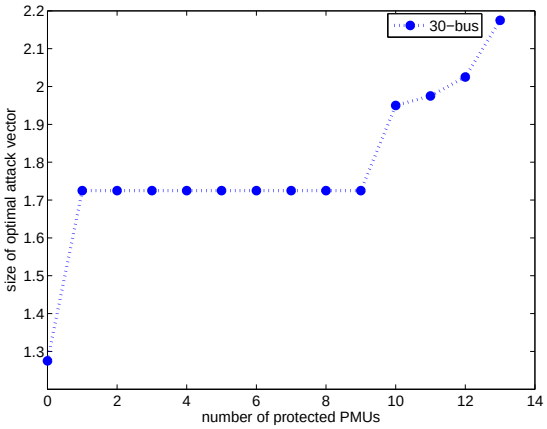

Fig. 7. Greedy placement of additional PMUs in IEEE 30-bus system with flow measurements on all lines, voltage measurements on 60% of the buses and $1/10$ of measurements initially protected

REFERENCES

[1] B. Liscouski and W. Elliot, "Final report on the August 14, 2003 blackout in the United States and Canada: Causes and Recommendations", *A report to US Department of Energy*, 40, 2004.

[2] S. Gorman, "Electricity grid in U.S. penetrated by spies", *Wall St. J.*, 2009.

[3] A. L. Ott, "Experience with PJM market operation, system design, and implementation", *IEEE Trans. Power Syst.*, vol. 18, no. 2, 2003.

[4] A. Abur and A. G. Expósito, "Power System State Estimation: Theory and Implementation", New York: Marcel Dekker, 2004.

[5] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids", *Proc. ACM Conf. Comput. Commun. Security*, 2009.

[6] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attack on power system state estimation", *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, 2012.

[7] R. F. Nuqui and A. G. Phadke, "Phasor measurement unit placement techniques for complete and incomplete observability", *IEEE Trans. Power Del.*, vol. 20, 2005.

[8] T. Kim and V. Poor, "Strategic Protection Against Data Injection Attacks on Power Grids", *IEEE Trans. Smart Grid*, vol. 2, no. 2, 2011.

[9] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Limiting false data attacks on power system state estimation", *Proc. Conf. Inf. Sci. Syst.*, 2010.

[10] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets", *Proc. IEEE SmartGridComm*, 2010.

[11] R. Christie, "Power system test archive", Available: http://www.ee.washington.edu/research/pstca.

[12] L. R. Ford and D. R. Fulkerson, "Maximal flow through a network", *Can. J. Math.*, 1956.

[13] M. Stoer and F. Wagner, "A simple min-cut algorithm", *J. ACM*, 44(4), 1997.

[14] A. G. Phadke, "Synchronized phasor measurements in power systems", *IEEE Comput. Appl. Power*, vol. 6, 1993.

[15] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems", *IEEE Trans. Smart Grid*, vol. 1, 2010.

[16] D. L. Donoho, "Compressed sensing", *IEEE Trans. Inf. Theory*, vol. 52, 2006.

[17] D. P. Shepard, T. E. Humphreys, and A. A. Fansler, "Evaulation of the Vulnerability of Phasor Measurement Units to GPS Spoofing", *International Journal of Critical Infrastructure Protection*, 2012.

[18] D. Deka and S. Vishwanath, "PMU placement and error control using belief propagation", *IEEE SmartGridComm*, 2011.

[19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming", http://stanford.edu/∼boyd/cvx, 2009.