Optimal Data Attacks on Power Grids: Leveraging Detection & Measurement Jamming

Deepjyoti Deka, Ross Baldick, and Sriram Vishwanath
Department of Electrical & Computer Engineering, The University of Texas at Austin
Email: deepjyotideka@utexas.edu, baldick@ece.utexas.edu, sriram@ece.utexas.edu

***Abstract*—Meter measurements in the power grid are susceptible to manipulation by adversaries that can lead to errors in state estimation. This paper presents a general framework to study attacks on state estimation by adversaries capable of injecting bad-data into measurements and further, of jamming their reception. Through these two techniques, a novel 'detectable jamming' attack is designed that changes the state estimation despite failing bad-data detection checks. Compared to commonly studied 'hidden' data attacks, these attacks have lower costs and a wider feasible operating region. It is shown that the entire domain of jamming costs can be divided into two regions, with distinct graph-cut based formulations for the design of the optimal attack. The most significant insight arising from this result is that the adversarial capability to jam measurements changes the optimal 'detectable jamming' attack design only if the jamming cost is less than half the cost of bad-data injection. A polynomial time approximate algorithm for attack vector construction is developed and its efficacy in attack design is demonstrated through simulations on IEEE test systems.**

***Index Terms*—PMU, phasor measurement, jamming, data injection, detectable attack, graph-cut**

I. Introduction

As power grids around the world move towards smarter devices and distributed control, it has led to large scale placement of metering, including PMUs [1], for real-time data collection. This can have several positive implications for the grid, notably monitoring of the grid state for improved reliability and optimal electricity prices. However, 'smart' meters and associated communication infrastructure are vulnerable to adversarial attacks by rogue agents and online viruses. Examples of these attacks include GPS spoofing attack on PMUs [2], 'Dragonfly' virus [3], Aurora test attack [4] among others. Such data attacks can lead to incorrect estimation of the grid state and result in large scale blackouts. The extreme consequences of adversarial attacks and counter strategies has attracted significant interest from the research community. [5] first introduced the problem of undetectable data attacks that bypass standard bad-data tests present in the state estimator. The optimal attack vector of compromised measurements is constructed in [5] using projection matrices. Subsequent work has looked at the problem of constructing the optimal attack under different grid conditions and adversarial objectives. Attack construction that requires minimum number of measurement corruptions are presented in [6] using $l_0 - l_1$ relaxation. Reference [7] analyzed a system with phasor measurements and used mixed integer linear programming to create the optimal attack. For systems with PMUs and line flow measurements, [8], [9] discusses graph-cut based attacks on specific buses on the grid and protection strategies against them. Similarly, other protection schemes have been discussed in the literature, including but not limited to heuristic protection schemes [10], greedy schemes [6], [9] among others.

It is worth noting that most research on power grid cybersecurity has focussed on designing 'hidden' attack vectors that completely evade bad-data detection tests at the state estimator. However, the authors of [11] showed that this is not necessary and introduced data 'framing' attacks that require changing the values at half of the measurements in the attack vector while damaging the other half. The attack here is initially detected by the estimator but becomes successful after the damaged measurements are removed. In [12], a generalization of this called a 'detectable' attack model was presented for systems where a subset of the measurements are secure and hence incorruptible. The authors in [12] showed that by focussing on the bad-data identifier in the state estimator, the cardinality of the optimal 'detectable' attack on average can be reduced by greater than 50% of the cardinality of 'hidden' attacks ($(50\%$ in the worst case)). More importantly, the 'detectable' attack framework in [12] produces feasible attacks in operating regimes where no 'hidden' attacks are possible.

In this work, we consider the 'detectable' attack framework in [12] but with one major modification - in addition to modifying insecure measurements (bad-data injection) as described in previous work, the adversary here is capable of jamming measurement communication to the state estimator. Note that measurement jamming can be conducted using commercial jammers, Denial of Service attack [13] or even by physical damage to the meter or communication channel. Compared to bad-data injection that requires measurements to be changed by precise real values, measurement jamming is in fact less resource-intensive. The overarching goal of this work is thus to *study the impact of adding measurement jamming to the adversary's arsenal on the design of the optimal data attacks.* We show that the entire range of values for measurement jamming costs can be divided into two intervals with different graph-cut based optimal attack formulations. Specifically, we prove that measurement jamming significantly alters the optimal attack strategy only when the jamming cost is less than half the cost of data-injection. We provide recursive min-cut based algorithms to design the optimal attack over the entire range of jamming cost values and show the cost improvement derived from measurement jamming through simulations on

978-1-4673-8289-2/15/$31.00 ©2015 IEEE

IEEE test cases [14]. By discussing the scope of measurement jamming as an adversarial strategy, our work thus provides a potent and realistic generalization of current data attack frameworks that undermines measures of grid resilience based on invulnerability to 'hidden' attacks.

The rest of this paper is organized as follows. The next section presents a description of the system models used in state estimation, bad-data detection and identification. The novel adversarial attack model with jamming is introduced in Section III along with conditions necessary for attack feasibility. Section IV analyzes how the cost of jamming affects the attack strategy and grid resilience, and presents a graph theoretic formulation for the optimal attack design. Our algorithm to design an optimal attack vector is presented in Section V. Simulations of the proposed algorithm for the range of jamming and bad-data injection costs on IEEE bus systems and comparisons with existing work are shown in Section VI. Finally, concluding remarks and future directions of work are presented in Section VII.

II. State Estimation and Bad-Data Detection in Power Grids

We denote the power grid by a set V of buses (nodes) connected by a set E of transmission lines (directed edges).

Measurement Model: We use the DC power flow model [15], which assumes unit voltage magnitudes at all nodes and inductive lines. It is given by:

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

Here $z \in \mathbb{R}^m$ is the m length vector of measurements consisting of a) flow measurements on lines, and b) voltage phasor measurements on buses. $x \in \mathbb{R}^n$ denotes the state vector of length $|V|$ that comprises of the phase angles at all buses in the grid. H is the measurement matrix and e is a zero mean Gaussian noise vector with known covariance Σ. Let the k_1^{th} and k_2^{th} entries in z represent the power flow on line (i,j) (from nodes i to j) and the voltage angle at node i respectively. Then, $z(k_1) = B_{ij}(x(i) - x(j)),\ z(k_2) = x(i)$. Here B_{ij} is the susceptance of line (i,j). The corresponding rows in H thus have the following structure:

$$H(k_1,:) = [0..0 \quad B_{ij} \ 0..0 \ -B_{ij} \ 0..0] \tag{2}$$

$$H(k_2,:) = [0..0 \quad 1 \quad 0..0 \ \ 0 \ \ 0..0] \tag{3}$$

We assume full column rank of H for unique state estimation. Further, without loss of generality, we introduce a $(n+1)^{th}$ reference bus with phase angle 0 in our system and represent it by augmenting 0 to the state vector x. Using (3) any bus angle measurement can be considered equivalent to a flow on a hypothetical line of unit susceptance between the bus and the reference bus. Thus, we add an extra column h^g corresponding to the reference bus in matrix H to get $z = Hx = [H|h^g]\begin{bmatrix} x \\ 0 \end{bmatrix}$. Here $h^g(k) = -1$ if $z(k)$ measures a phase angle and 0 otherwise. Abusing notation, we use x and H to denoted the augmented state vector and measurement matrices respectively. Observe that in the augmented system, all measurements now correspond to flow measurements.

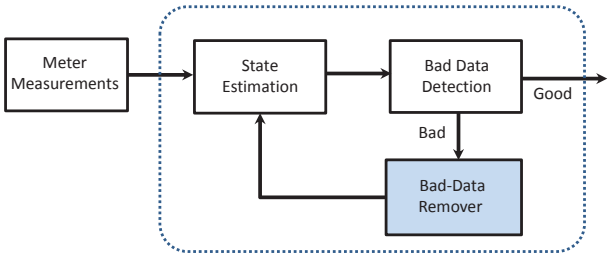

Fig. 1. State Estimator for a power system [16], [15]

State Estimator: We consider a least-square state estimator in the grid as shown in Figure 1 [16], [15]. The state vector estimate x^* for measurement vector z is generated by minimizing the weighted residual $J(x,z) = \|\Sigma^{-.5}(z - Hx)\|_2$ over variable x. Following estimation, a threshold (λ) based bad-data detector determines the presence of erroneous measurements by the following test:

$$\begin{aligned} \|\Sigma^{-.5}(z - Hx^*)\|_2 \leq \lambda \ \ &\text{accept} \quad x^* \\ > \lambda \ \ &\text{detect bad-data} \end{aligned} \tag{4}$$

If the test detects bad-data, the estimator tries to identify the erroneous measurements as described below, following which the state estimate is recomputed.

Bad-data Removal: The measurement residual vector r for measurement z and estimated x^* is given by [16], [15]:

$$r = z - Hx^* = [I - H(H^T\Sigma^{-1}H)^{-1}H^T\Sigma^{-1}]z \tag{5}$$

with variance R_r. Assuming that each measurement is independently affected by natural bad data, the optimal removal strategy involves removing the least number of measurements from z to satisfy the threshold condition in Eq. (5) while preserving system observability. For multiple bad-data entries, this optimal strategy is known to be a non-convex problem [16], [12] and iterative schemes may be used.

We assume in the remainder of this paper that the measurement data z, in the absence of any adversarial manipulation, is capable of producing the correct state estimate x^* by passing the bad-data detection test in (5).

A. Attack Models

Let a denote the adversarial attack vector added to measurement vector z to generate the compromised measurement vector $z+a$. Traditional attack models focus on bypassing the bad-data detector by ensuring that the measurement residual in (4) remains unchanged. Mathematically, this requires $a = Hc$ for some non-zero $c \in \mathbb{R}^n$ as $\|\Sigma^{-.5}(z - Hx^*)\|_2 = \|\Sigma^{-.5}(z + a - H(x^* + c))\|_2$. This is termed a **'Hidden' Attack** that produces an erroneous state vector $x^* + c$ [5]. Next we describe 'detectable' data attacks [12] that are the focus of this paper.

'Detectable' Data Attack: From the bad-data removal scheme described earlier, it is clear that an attack vector $a \neq 0$ will change the state estimate if removal of some

$k < \|a\|_0$ measurements (distinct from the attack vector) satisfies (4). For any $Hc \neq 0$, consider an adversarial strategy that excludes (or does not corrupt) less than 50% of the non-zero entries in Hc from the attack vector d. Note that the attack gets detected but the non-zero terms in $(Hc - d)$ instead of d are identified as bad-data and removed. This happens as $\|d\|_0 > \|Hc - d\|_0$. After removal, the remaining measurements that include the non-zero terms in d satisfy the threshold leading to a 'detectable' attack. In the next section, we formulate the design of the optimal 'detectable' data attack and use it to analyze changes that arise due to the adversarial capability to jam measurements.

III. 'Detectable' Attack with Measurement Jamming

In a general setting, some measurements in the grid may be incorruptible due to geographical isolation or encryption. We denote this set of measurements secure from adversarial corruption by S. Note that measurements in S suffer from noise induced bad-data. The remaining insecure measurements open to adversarial manipulation belong to set S^c. The minimum cost 'detectable' attack is given by the non-zero terms in the optimal vector d^* for the following optimization problem [12]:

$$\min_{d\in\{0,1\}^m, c\in\mathbb{R}^{n+1}} \|d\|_0 \tag{P-1}$$
$$\text{s.t. } a = Hc, c \neq \mathbf{0}, c(n+1) = 0, d(i) = 0 \ \forall i \in S$$
$$\|d\|_0 > \|a\|_0/2 \ \text{ (for feasibility)} \tag{6}$$
$$rank(DH) = n, \ diag(D) = \mathbf{1} - (\mathbf{1} - d) * a_0 \tag{7}$$

Here, $a*b$ refers to the element-wise multiplication between vector a and b, while a_0 denotes the sparsity pattern in vector a. Condition (6) ensures that the estimator removes measurement entries corresponding to non-zero terms in $(\mathbf{1}-d)*a$ as bad-data, instead of the data injected in $d*a$. D is a diagonal matrix whose diagonal entries are 0 for removed data and 1 otherwise. The attack passes the bad-data detection test as it lies in the column space of DH, the measurements matrix after bad-data removal. It is worth restating that as each row in H (augmented) corresponds to a flow measurement, H is equivalent to a susceptance weighted incidence matrix of a graph G_H with $n+1$ nodes and edges given by rows in H. Due to this structure of H, it can be shown that the optimal attack strategy for Problem P-1 doesn't change if H is replaced by the un-weighted incidence matrix A_H of graph G_H ($A_H(i,j) = 1(\hat{H}(i,j) \geq 0) - 1(\hat{H}(i,j) \leq 0)$) [8], [9], [12] and c is restricted to be a $0-1$ vector. Further, the optimal attack for Problem P-1 has the following graph-theoretic formulation.

Theorem 1 ([12, Theorem 2])**.** *Let C^* denote the minimum cardinality cut in G_H with a minority of secure cut-edges ($|C^* \cap S| < |C^*|/2$). An optimal 'detectable' attack for Problem P-1 is given by any $\lfloor 1+|C^*|/2 \rfloor$ cut-edges in $C^* \cap S^c$ (insecure cut edges).*

The proof can be found at [12]. Observe that if d is restricted to an all-1 vector, Problem P-1 gives the minimum cardinality cut that does not include any secure edge in S^c, which is the optimal 'hidden' attack [8], [9].

'Detectable Jamming' Attack: We now analyze an adversary with the capacity to jam insecure measurements in addition to injecting bad-data to manipulate measurements. Secure measurements are considered to be unjammable as well. Let p_J and p_I be the cost associated with jamming and bad-data injection into an insecure measurement respectively. We assume that $0 \leq p_I \leq p_I$ to be the range of values for p_J as jamming is less resource intensive than bad-data injection. Consider a cut C in graph G_H. Let n_S^C and $n_{S^c}^C$ denote the number of secure and insecure edges in cut C with $n_{S^c}^C > n_S^C$ as shown in Fig. 2. By Theorem 1, 'detectable' attack requires injection into k^C ($k^C > |C|/2$) insecure edges in C at a cost of $p_I k^C$. Instead, consider a different strategy where the adversary jams k_J^C insecure measurements. As jammed measurements are not received by the control center, the cut-size effectively reduces to $|C| - k_J^C$. If the remaining $n_{S^c}^C - k_J^C$ insecure cut edges are greater in number than the n_S^C secure edges, the adversary can still attack k_I^C ($> \frac{|C|-k_J^C}{2}$) measurements and generate a feasible attack as depicted in Fig. 2. The cost of this new attack is $p_I k_I^C + p_J k_J^C$. We term it a 'detectable jamming' attack to distinguish it from the original 'detectable' attack without jamming. The design of the optimal 'detectable jamming' attack is formulated as follows:

$$\min_{d_J, d_I\in\{0,1\}^m, c\in\mathbb{R}^{n+1}} p_J\|d_J\|_0 + p_I\|d_I\|_0 \tag{P-2}$$
$$\text{s.t. } a = A_H c, c \neq \mathbf{0}, c(n+1) = 0$$
$$d_J + d_I \in \{0,1\}^m \tag{8}$$
$$d_J(i) = d_I(i) = 0 \ \forall i \in S \tag{9}$$
$$\|d_I\|_0 > (\|a\|_0 - \|d_J\|_0)/2 \ \text{ (for feasibility)} \tag{10}$$
$$rank(DA_H) = n, diag(D) = \mathbf{1} - (\mathbf{1} - d_J - d_I) * a_0 \tag{11}$$

The non-zero values in d_J and d_I give the measurements to jam and injection bad-data respectively in the optimal attack. Note that in Problem P-2, we replace H with incidence matrix A_H and make c a $0-1$ vector as discussed for Problem P-1. Condition (8) ensures that data injection and jamming cannot occur at the same measurement. The remaining conditions arise from incorruptibility of secure measurements (9), feasibility of 'detectable' attack (10) and full system observability after bad-data removal (11). From the discussion preceding Problem P-2, it is clear that the optimal 'detectable jamming' attack has a graph-cut based construction as stated below.

Lemma 1. *Let C denote a cut in G_H with ($n_{S^c}^C > |C|/2$) insecure cut-edges. A feasible 'detectable jamming' attack is given by jamming ($k_J^C \geq 0$) and injecting data into ($\lfloor 1 + \frac{|C^*|-k_J^C}{2} \rfloor > 0$) insecure cut-edges at a cost of $p_J k_J^C + p_I \lfloor 1 + \frac{|C^*|-k_J^C}{2} \rfloor$. The optimal attack is given by minimizing the attack cost over variable k_J^C (jammed edges) for all feasible cuts C.*

It is noteworthy that if $k_J^C = 0$ in Lemma 1, we obtain the optimal 'detectable' attack (no jamming) as a feasible 'de-

tectable jamming' attack. This leads to the following important properties.

Corollary 1.
- *The set of system configurations with feasible 'detectable jamming' attacks is identical to that of 'detectable' attacks and is a superset of that of hidden attacks.*
- *The cost of the optimal 'detectable jamming' attack is never greater than the cost of optimal 'detectable' attack and never greater than* $.5 + 1/|C_h^*|$ *times the cost of optimal 'hidden' attack,* $|C_h *|$ *being the cardinality of optimal 'hidden' attack in the system.*

The first property arises as the feasibility requirement for 'detectable jamming' and 'detectable' attacks are the same and less strict than that for 'hidden' attacks. The second property follows from the fact that every 'detectable' attack is a feasible 'detectable jamming' attack, while injecting bad-data into $1 + \lfloor |C_h^*| \rfloor/2$ measurements of every 'hidden' attack constitutes a feasible 'detectable' attack. The simulation results in Section VI demonstrate that the average impact of 'detectable jamming' attack is indeed more substantial than these worst-case bounds. In the next section, we discuss how the jamming cost p_J affects the design of the optimal attack vector of our regime.

IV. Effect of Jamming Cost on Attack Construction

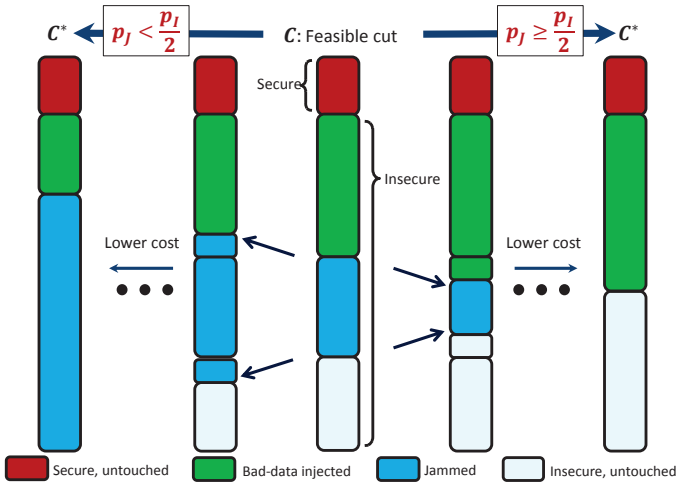

Fig. 2. Effect of jamming cost p_J and bad-data injection cost p_I on the minimum cost attack C^* derived from a feasible cut C. When $p_J < p_I/2$, attack cost is reduced by replacing one bad-data injection with jamming two measurements as shown on the left of C. For $p_J \geq p_I/2$, attack cost is reduced by replacing two jammed measurements by one measurement with bad-data injection while leaving the other untouched as shown on the right of C.

As mentioned earlier, we consider the jamming cost p_J to lie in the interval $[0, p_I]$ where p_I is the bad-data injection cost. Consider a feasible cut C with $n_{S^c}^C$ insecure edges and n_S^C secure edges in G_H as shown in Fig. 2. By Theorem 1, a feasible 'detectable jamming' attack comprises of selecting $(k_J^C \geq 0)$ and $(k_I^C = \lfloor 1 + \frac{|C| - k_J^C}{2} \rfloor > 0)$ insecure edges for jamming and bad-data injection respectively, at a overall cost of p^C given by:

$$p^C = (p_J - \frac{p_I}{2})k_J^C + p_I \frac{|C| + 2 - (|C| - k_J^C) \mod 2}{2} \tag{12}$$

where $\mod$ denotes the remainder operator. We divide the range of p_J into two intervals: A $(0 \leq p_J < p_I/2)$ and B $(p_I/2 \leq p_J \leq p_I)$. Note that in interval A, the cost p^C is a decreasing function of k_J^C. Therefore, the minimum cost attack for feasible cut C is obtained by jamming the maximum permissible $(n_{S^c}^C - n_S^C - 1)$ insecure edges. The remaining $(n_S^C + 1)$ insecure edges are injected with bad-data to maintain feasibility.. The attack cost is given by

$$p^C = (p_I - p_J)n_S^C + p_J n_{S^c}^C + (p_I - p_J) \tag{13}$$

Ignoring constant $(p_I - p_J)$, this equals C's cut-weight if secure and insecure edges are given weights of $(p_I - p_J)$ and p_J respectively. Thus, if $p_J < p_I/2$, the optimal 'detectable jamming' cut corresponds to the feasible cut C^* with lowest cut-weight, where secure and insecure edges have weights of $(p_I - p_J)$ and p_J respectively. Next consider interval B $(p_I/2 \leq p_J \leq p_I)$. In Eq. (12), if k_J^C is reduced by 2, the $\mod 2$ term remains unchanged and cost p^C decreases. Hence the optimal attack for cut C corresponds to either $k_J^C = 0$ or $k_J^C = 1$, otherwise the attack cost can be reduced further. By directly checking the $\mod 2$ term's contribution, we note that the optimal attack for C is given by $(k_J^C = 0, k_I^C = (1 + |C|)/2)$ for odd $|C|$, and $(k_J^C = 1, k_I^C = |C|/2)$ for even $|C|$. In either case, the attack cost can be expressed as follows:

$$p^C = p_J(1 - |C| \mod 2) + p_I \lfloor (1 + |C|)/2 \rfloor \tag{14}$$

As this is an increasing function of the cut-size $|C|$, the optimal 'detectable jamming' attack in interval B corresponds to the feasible cut C^* with lowest cut-size. Fig. 2 clearly demonstrates the creation of an optimal attack for a feasible cut C given the relative values of p_I and p_J. We summarize this discussion by presenting our main theorem for optimal 'detectable jamming' attack construction.

Theorem 2. *The cost optimal 'detectable jamming' attack for measurement graph* G_H *with jamming cost* p_J *and bad-data injection cost* p_I *is given by:*
A)$p_J < p_I/2$*: Give weights of* $p_I - p_J$ *and* p_J *to secure and insecure edges respectively in* G_H*. Find the minimum weight feasible cut* C^* *with* $n_S^{C^*}$ *secure edges. Of the insecure measurements, use* $(n_S^{C^*} + 1)$ *for bad-data injection and jam the rest.*
B)$p_J \geq p_I/2$*: Find the minimum cardinality feasible cut* C^* *in unweighted* G_H*. Use* $\lfloor (1 + |C^*|)/2 \rfloor$ *insecure measurements for bad-data injection and jam* $(1 - |C^*| \mod 2)$ *measurements.*

The following deductions follow immediately from Theorem 1 and Theorem 2,

Corollary 2.
- *For $p_J \geq p_I/2$, the optimal 'detectable jamming' and 'detectable' (without jamming) attacks correspond to the same cut C^*, that has minimum cardinality among all feasible cuts.*
- *For $p_J = 0$, the optimal 'detectable jamming' attack corresponds to the cut C^*, that has the minimum number of secure edges among all feasible cuts.*

In the next Section, we present our algorithm to construct the optimal attack described in Theorem 2 and Corollary 2.

V. Algorithm For Attack Construction

To confirm the existence of a feasible attack, we need to identify a cut with a majority of insecure edges in the graph. Theorem 3 in [12] proves that this is equivalent to the 'ration-cut' problem, a known NP-hard problem. Thus, the design of the optimal 'detectable jamming' attack is hard in general as well. We now provide an approximate algorithm (Algorithm 1) for attack vector construction. For $p_J < p_I/2$, we create weighted graph G_H with secure (insecure) edges having weight $p_I - p_J$ (p_J). For $p_J \geq p_I/2$, we consider unweighted G_H. Using Theorem 2, the optimal attack, in either case, is generated using the minimum weighted feasible cut in G_H.

Working: Algorithm 1 proceeds by computing the minimum weight cut C in G_H (Step 1) and checks if it is a feasible cut (Step 3). If C is infeasible, one secure edge is selected randomly in C and its edge-weight is increased by β (Step 4). We consider two cases, one where β is taken as the secure edge-weight (finite) and the other where it is taken as ∞. Following the increase, the algorithm recomputes the minimum weight cut and checks for feasibility. This process is iterated until a feasible cut is obtained (construct the attack vector) or the cut-weight reaches a threshold $\gamma < \infty$ (declare no solution).

Algorithm 1 'Detectable Jamming' Attack Construction

Input: β, γ, G_H with secure (S) and insecure (S^c) edges weighted based on p_J, p_I

1: Compute min-weight cut C in G_H
2: $w_C \leftarrow$ weight of C
3: **while** ($w_C < \gamma, 2|C \bigcap S| \geq |C|$) **do**
4: Randomly pick edge $i \in C \bigcap S$ and increase its weight by β
5: Compute min-weight cut C in G_H
6: $w_C \leftarrow$ weight of C
7: **end while**
8: **if** $2|C \bigcap S| < |C|$ **then**
9: Construct attack vector using Theorem 2
10: **else**
11: Declare no solution
12: **end if**

Note that for $\beta = \infty$, in the worst case, there are $|S|$ min-cut computations (one for each secure edge) of complexity $O(|V||E| + |V|^2 \log |V|)$ [17] giving Algorithm 1 a computational complexity of $O(|S||V||E| + |S||V|^2 \log |V|)$. However, as the algorithm is approximate, it might not return a solution in every case. In the next section, we show simulation results on designing optimal attacks by Algorithm 1 in IEEE test systems. We also demonstrate the capacity of 'detectable jamming' attacks in overcoming high placement of secure measurements in the systems considered.

VI. Results on IEEE test systems

We discuss the performance of Algorithm 1 in designing 'detectable jamming' attacks by simulations on IEEE 14-bus and 57-bus test systems [14]. In each simulation run, we put flow measurements on all lines in the considered test system and phase angle measurements on 60% of the system buses, selected randomly. Over multiple simulations, we vary the fraction of secure measurements and record the trends in average cost of constructing 'detectable jamming' attack. We consider difference values over the range of jamming cost p_J $(0, p_I/4, 3p_I/4)$, and different values of parameter β (finite and ∞) in Algorithm 1. The trend in average optimal cost of 'detectable jamming' attacks for the 14 bus system is presented in Fig. 3 for configurations that allow feasible 'hidden' attacks. To demonstrate the efficacy of our approach, we also plot average costs of 'hidden' and 'detectable' (no jamming) attacks. Note that the average 'detectable jamming' attack cost in Fig. 3 is significantly below the upper bound (Corollary 1). The average costs are observed to eventually decrease with increasing secure measurements in the system. This trend is due to increasing number of system configurations resilient to 'hidden' attacks, that are ignored while computing the average attack costs. Further, it is apparent that changing the value of β does not affect the performance of Algorithm 1 much. Similarly, Fig. 4 includes the reduction in average attack cost due to jamming for the 57 bus system, with finite β. Finally, Fig. 5 plots the increase in number of resilient operating regimes with an increase in the number of secure measurements in the system. It can be seen that compared to 'hidden' attacks, 'detectable' and 'detectable jamming' attacks pose a much greater threat to the grid vulnerability as the number of secure operating regimes in the latter hardly increases with increasing number of secure measurements. To conclude, the simulations prove the dual adversarial benefits of 'detectable jamming' attacks: lowering of attack cost and increased insensitivity to presence of secure measurements.

VII. Conclusion

We introduce a new data attack framework on power grids termed 'detectable jamming' attacks, where an adversary uses measurement jamming in addition to changing meter readings (bad-data injection) to create a change in state estimation despite violating the bad-data detection test. This is ensured by leading the state estimator to incorrectly label uncorrupted correct data as bad-data. The worst-case attack cost of 'detectable jamming' attacks is approximately half of the optimal 'hidden' attack cost, while the capability to overcome secure measurements is more pronounced. We show that the design

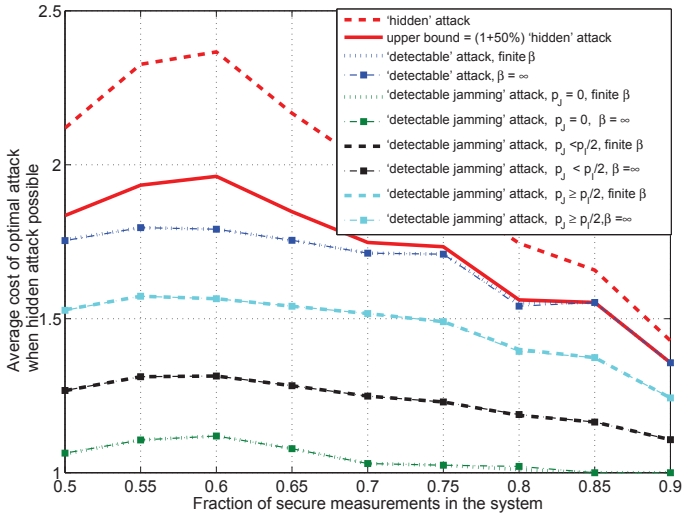

Fig. 3. Average cost of optimal attacks ('hidden', 'detectable' and 'detectable jamming') for different values of β (weight of secure edge and ∞) by Algorithm 1 on the IEEE 14 bus test system with protection on a fraction of measurements selected randomly. The bad-data injection cost (p_I) is taken as 1. The jamming costs (p_J) considered are $0, 1/4(< p_I/2), 3/4(> p_I/2)$. Only configurations with feasible 'hidden' attacks are considered to compute the average costs.

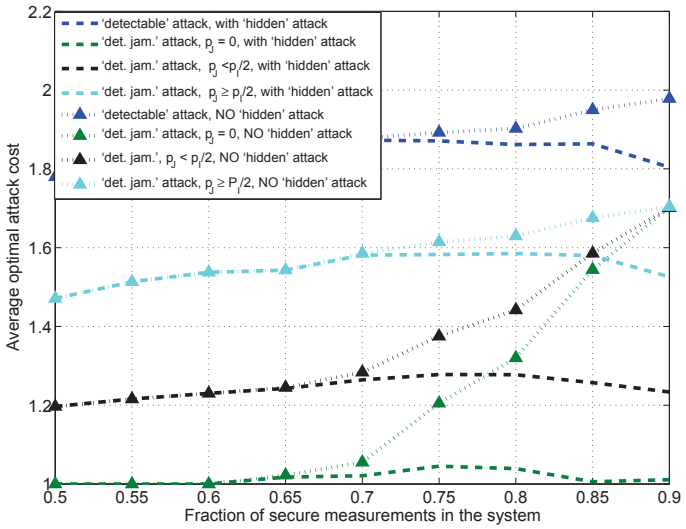

Fig. 4. Average cost of optimal attacks ('detectable' and 'detectable jamming') by Algorithm 1 (with finite β) on the IEEE 57 bus test system with protection on a fraction of measurements selected randomly. The bad-data injection cost (p_I) is taken as 1. The jamming costs (p_J) considered are $0, 1/4(< p_I/2), 3/4(> p_I/2)$.

of the minimum cost attack of this regime is equivalent to a constrained graph cut problem that takes two different forms, dependent on the relative values of jamming and data injection costs. Further we show that in comparison to 'detectable' (no jamming) attacks, our jamming reliant framework significantly alters the optimal attack design if the jamming cost is less than half the cost of bad-data injection. We present an iterative min-cut based approximate algorithm with polynomial complexity to determine the optimal cut. We demonstrate the adversarial benefits of our proposed 'detectable jamming' framework through simulations of our algorithm on IEEE test cases for different jamming costs and system conditions. Designing

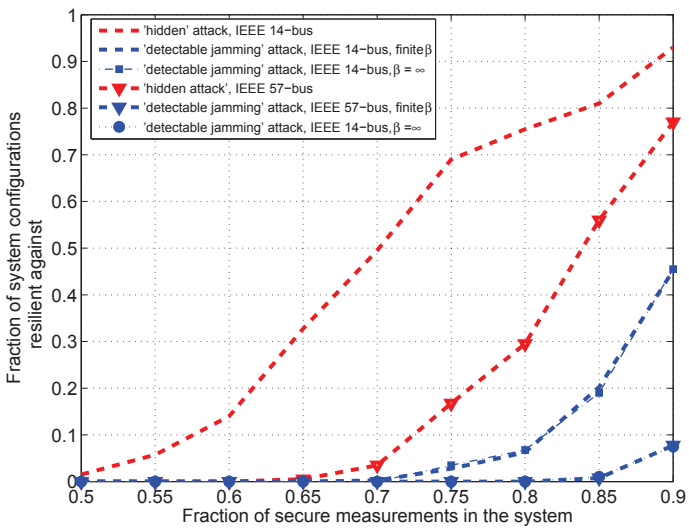

Fig. 5. Average fraction of configurations with no feasible 'hidden' and 'detectable jamming' attacks given by Algorithm 1 in IEEE 14 and 57 bus test systems with increasing fraction of secure measurements.

optimal security measures against this attack regime is the object of our current research in this domain.

REFERENCES

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

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

[3] http://www.nytimes.com/2014/07/01/technology/energy-sector-faces-attacks-from-hackers-in-russia.html

[4] J. Meserve, "Staged cyber attack reveals vulnerability in power grid", CNN, 2007. Available: http://www.cnn.com/2007/US/09/26/power.at.risk/index.html.

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

[7] 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.

[8] D. Deka, R. Baldick, and S. Vishwanath, "Optimal Hidden SCADA Attacks on Power Grid: A Graph Theoretic Approach", *ICNC*, 2014.

[9] D. Deka, R. Baldick, and S. Vishwanath, "Data Attack on Strategic Buses in the Power Grid: Design and Protection, *IEEE PES General Meeting*, 2014.

[10] 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.

[11] J. Kim, L. Tong, and R. J. Thomas, "Data Framing Attack on State Estimation with Unknown Network Parameters", *Asilomar Conference on Signals, Syst., and Computers*, 2013.

[12] D. Deka, R. Baldick, and S. Vishwanath, "Data Attacks on the Power Grid DESPITE Detection", *IEEE PES Innovative Smart Grid Technologies*, 2015. (available at http://arxiv.org/abs/1505.01881)

[13] L. Shichao, L. P. Xiaoping, and S. E. Abdulmotaleb, "Denial-ofservice (dos) attacks on load frequency control in smart grids", *IEEE PES Innovative Smart Grid Technologies*, 2013.

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

[15] A. Abur and A. G. Exposito, "Power System State Estimation: Theory and Implementation", *CRC*, 2000.

[16] A. Monticelli, "State estimation in electric power systems: a generalized approach", *Kluwer Academic Publishers*, 1999.

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