Unidentifiable Attacks in Electric Power Systems

Zhengrui Qin, Qun Li
College of William & Mary, Williamsburg, VA
zhengrui@cs.wm.edu liqun@cs.wm.edu

Mooi-Choo Chuah
Lehigh University, Bethlehem, PA
chuah@cse.lehigh.edu

***Abstract*—The electric power grid is a crucial infrastructure in our society and is always a target of malicious users and attackers. In this paper, we first introduce the concept of unidentifiable attack, in which the control center cannot identify the attack even though it detects its presence. Thus, the control center cannot obtain deterministic state estimates, since there may have several feasible cases and the control center cannot simply favor one over the others. Furthermore, we present algorithms to enumerate all feasible cases under an unidentifiable attack, and propose an optimization strategy from the perspective of the control center to deal with an unidentifiable attack. We briefly evaluate and validate our enumerating algorithms and optimization strategy.**

***Keywords*-Smart Grid, Unidentifiable Attack, State Estimates, False Data Injection, Security, Bad Data Identification.**

Nomenclature

Indices:
- k: feasible case index
- g: generator index
- i, j: bus index

Sets and elements:
- L: set of load buses
- G: set of generator buses
- A: set of all meters
- P: set of protected meters
- D: set of bad meters
- B: set of all buses, $B = L \cup G$
- b_i: bus i
- T: set of transmission lines
- t_{i_j}: transmission line between buses i and j
- t_{i_*}: transmission lines incident to bus i

Constants:
- m: total number of meters, $m = |A|$
- l: total number of feasible cases
- n: total number of buses, $n = |B|$
- r: capacity of the attacker
- C_g: generating cost of generator g
- $C_{shed,i}$: power shedding cost of load bus i
- G_{ij}: conductance between bus i and bus j
- B_{ij}: susceptance between bus i and bus j
- $PG_{g,min}$: min real capacity of generator g
- $PG_{g,max}$: max real capacity of generator g
- $QG_{g,min}$: min reactive capacity of generator g
- $QG_{g,max}$: max reactive capacity of generator g
- PL_{ij}^{min}: min line capacity between bus i and bus j
- PL_{ij}^{max}: max line capacity between bus i and bus j
- $PD_{k,i}$: real demand on bus i in case k
- $QD_{k,i}$: reactive demand on bus i in case k

Variables:
- PG_g: real power generated by generator g
- QG_g: reactive power generated by generator g
- V_i: voltage amplitude of bus i
- θ_i: voltage phase of bus i
- $PS_{k,i}$: real power shedding of bus i in case k
- $QS_{k,i}$: reactive power shedding of bus i in case k
- P_{ij}: real power flow between bus i and j
- Q_{ij}: reactive power flow between bus i and j
- PL_{ij}: power flow between bus i and j
- $D_{shed,k}$: total real power shedding cost for case k

I. Introduction

The electric power grid is a distribution network that connects the electric power generators to customers through transmission lines, and its security and reliability are critical to society. In order to enable its safe and reliable operation, the power grid is monitored continuously by smart meters installed at important locations of the power grid. The meters take various measurements, including real and reactive power injections on buses and real and reactive power flows on transmission lines. Such data is then fed to the control center within the Supervisory Control And Data Acquisition (SCADA) system. Using the collected information, the control center estimates the state variables, which are the voltage amplitudes and phases on buses, and then makes corresponding adjustments to stabilize the power grid.

To obtain reliable state estimates, it is essential for the control center to be fed reliable and accurate meter measurements. However attackers may compromise meter measurements and send malicious data to the control center, thus misleading the control center to make bad decision that may cause severe consequences to the power system. Researchers have developed various techniques to detect bad data measurements [1]–[7], most of which are based on measurement residuals.

However, Liu *et al*. [8] has presented an undetectable false data injection that can defeat all the detection techniques based on measurement residuals. Their results indicate that for medium size power system (e.g. IEEE 30-bus system), the attackers may need to compromise 60 to 75% of all meters before they can succeed in launching an undetectable attack. However, an attacker may either have limited attack resources or only limited access to some meters. Thus, we are interested in exploring if there are other types of attacks that require fewer meters.

In this paper, we focus on unidentifiable attacks, which are different from undetectable attacks discussed in [8]. *In unidentifiable attacks, the control center can detect that there are bad or malicious measurements, but it cannot identify which meters have been compromised*. As a result, the attacker does not need to manipulate as many meters for unidentifiable attacks as when he is launching undetectable attacks. Under

an unidentifiable attack, the control center has no way to simply eliminate some "bad" data and thus get accurate state estimates. However, the control center has to make a decision how much power to generate, no matter good or bad, in response to the attack. We argue that a good decision during such an attack is one that minimizes the total cost which includes generation and penalty cost caused by damages of the attack.

Our main contributions in this paper are as follows:

- We are the first to propose the unidentifiable attack in a smart grid system. We demonstrate the feasibility of this type of attack. An adversary can launch an unidentifiable attack by compromising a smaller number of meters compared with the previously proposed attacks while at the same time confuse the control center on what really happens.
- We propose a heuristic algorithm to enumerate all feasible cases under an unidentifiable attack. The previous classic "bad data detection" algorithms do not work for this attack scenario. Our algorithm is the first to resolve the problems of the previous algorithms. It also significantly reduces the possible solution searching space compared with brute force approach. We show through empirical study that the algorithm can efficiently find all possible attacks.
- Enumerating possible attacks is not equivalent to locating the exact attack. To defend against all possible attack scenarios, we also propose a strategy to minimize the average damage to the system. We formulate the problem as a nonlinear programming problem and solve it through a standard optimization package.
- We model our system in AC mode, which is nonlinear and doubles the number of variables compared to DC mode. The recent security investigations of the smart grid system, such as [8]–[11], are all based on DC mode. Although DC mode can be representative of the power system, AC mode can capture more subtleties and is more complicated and realistic to describe a power system. We believe this is the first piece of work to carefully examine the attacks and solutions in realistic AC mode. The formulation and optimization can be used as a basis for future work.

II. Related Work

To ensure the power system operates correctly, the control center needs to collect measurements to estimate the state variables, and then takes control actions against any contingency. For a system with n buses and m meters, the state estimates are determined through the following model:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e} \tag{1}$$

where $\mathbf{x} = (V_1, ..., V_n, \theta_1, ..., \theta_n)$ is the state variable vector, $\mathbf{z} = (z_1, ..., z_m)$ is the measurement vector, and $\mathbf{e}$ is the measurement error vector.

Bad measurements may exist due to faulty meters, transmission errors or alterations by malicious attackers. Bad measurements can induce the control center to obtain wrong state estimates and result in severe consequences. Researchers have developed lots of approaches on bad data detection and identification since 1970's, such as Identification By Elimination (IBE) [1], [2], Non-Quadratic Criteria (NQC) [3], Hypothesis Testing Identification (HTI) [4], Combinatorial Optimization Identification (COI) [5]–[7]. An early comparative study of the first three approaches can be found in [12]. Besides bad data detection approaches, public-key schemes, such as [13]–[15], can also be implemented to prevent malicious users from manipulating meter measurements.

Liu *et al.* [8] has shown that, given the topology and line impedance of a power system, an attacker can injection malicious data without being detected by the control center. The injected malicious data can introduce arbitrary errors into the state estimates, which could result in huge consequence. In this kind of attacks, the injected malicious data does not change the residual, and thus can circumvent all detectors based on residual checking. In the DC model, to launch an undetectable attack, the attack must manipulate the meter readings from $\mathbf{z}$ to $\mathbf{z+a}$ such that $\mathbf{a} = \mathbf{Hc}$ (in the DC model, Eq (1) is simplified to $\mathbf{z} = \mathbf{Hx} + \mathbf{e}$), where $\mathbf{c}$ is a constant vector to be added to the original state estimates. Since then, the undetectable attack has drawn a lot of attention, such as in [9] where a specific undetectable attack called the load redistribution attack is discussed.

The unidentifiable attack considered in this paper is different from the undetectable attack in that the control center can detect the presence of an attack but cannot identify which meters have been compromised. This is in fact the concept of *nondeducibility* [16] but with an inverse form, in which the attacker maintains the property of nondeducibility. Our unidentifiable attacks aim to confuse the control center to the extent that it does not know what the exact demand scenario is and hence needs to rely on a strategy to deal with such attacks. Compared with undetectable attack, an attacker only needs to manipulate at most half as many meters to launch an unidentifiable attack as those he needs for an undetectable attack. The concept of unidentifiable attacks is hence of great value and more practical, especially for an attacker with limited attack resources. Consequently, this paper complements the research in cyber-physical systems [17]–[22].

III. Unidentifiable Attack

The unidentifiable attack in this paper is a new type of attack in the power system. The formal definition of an unidentifiable attack is as follows.

Unidentifiable Attack: Suppose in a power system with a set of meters A, the attacker compromises a set of meters D, where $D \subset A$. An unidentifiable attack is the attack scenario that satisfies the following two conditions: (1) the control center is able to conclude the presence of bad measurements; (2) the control center cannot deterministically deduce whether

D or D' (or D's) is compromised, where $D' \subset A$, $D' \not\supseteq D$ and $D' \not\subseteq D$.

Remark: From the above definition, it is obvious that it is different from an undetectable attack, which cannot be detected by any means of detection. One would argue that an undetectable attack is a special case of unidentifiable attack, since an undetectable attack is literally unidentifiable. However, we differentiate between these two types of attacks in this work.

To further understand the unidentifiable attack, let us consider an ideal case where the measurements have no error except those which are manipulated by an attacker. Suppose there are $m = m_0 + 2m_1$ measurements which can be divided into three sets, M_0, M_1 and M_2, with cardinalities m_0, m_1 and m_1 respectively. Assume that an attacker has manipulated the set of measurements in M_1. As a whole, the measurements are not consistent, that is, the control center can detect the presence of an attack. Let us further assume that, the measurements $M_0 \cup M_1$ alone are consistent and make the whole system *observable*[1], so are the measurements $M_0 \cup M_2$. In such a scenario, the control center can conclude that either set M_1 or set M_2 are the compromised measurements, even if it knows that there are exactly m_1 compromised measurements. However, the control center has no way to determine the exact set (either M_1 or M_2) that has been compromised. We call such an attack unidentifiable, since the attack on set M_1 confuses the control center to believe that either set M_1 or set M_2 has been compromised. We say this attack has two *feasible cases*, one is $M_0 \cup M_1$ and the other is $M_0 \cup M_2$. Here by a feasible case, we mean a set of meter readings that render the power system observable and hence can produce a set of state variables that is different from the set of state variables produced by any other feasible case.

In the above example, it is easy to understand why the control center is confused, since $|M_1| = |M_2|$. Suppose $m_1 = |M_1| > |M_2| = m_2$ with other assumptions made for the above attack remain unchanged. Will the control center now favor set M_1 as good data over set M_2? It depends. If the control center knows that the largest number of measurements that the attacker can manipulate is smaller than m_1, then it knows that set M_2, instead of M_1, has probably been manipulated. However, if the attacker can manipulate m_1 or more meters, the control center still cannot favor one set over the other.

In the power system, all meters are interactive to some extent. Therefore, changing one meter usually requires changes of many other meters in order to make the changes consistent. From the view of an attack, he intends to change as few meters as possible to generate an unidentifiable attack. Considering also that the meters on generator buses are not easily attacked since the control center usually has direct communication with power plant to verify the meter readings, we in this paper focus on two types of unidentifiable attack, which require relatively less effort of the attacker. One is *load redistribution attack* (Type I), in which the attacker obfuscates the control center whether the power demands on some load buses are redistributed, while the total power demand is unchanged. The other is *load increase attack* (Type II), in which the attacker obfuscates the control center whether the demand on a certain bus is increased, while the demands for other load buses remain the same.

Let us consider two simple examples that illustrate the two types of unidentifiable attack above. To simplify our discussion, we use DC mode in our examples, but we consider AC mode for the rest of the work in this paper. Fig. 1 is a three bus power system. On each bus, there is a power injection meter; on each transmission line, there are two power flow meters, with one at each end of the line. In DC mode, there is no resistance on the transmission line but only susceptance. The susceptance between bus 1 and bus 2 is 280 (we omit the unit, and thereafter); that between bus 2 and bus 3 is 70; and that between bus 1 and bus 3 is 140. Suppose the load on bus 2 is 21, and the load on bus 3 is 35. Before any attack, the meter readings are consistent as shown in Fig. 1.

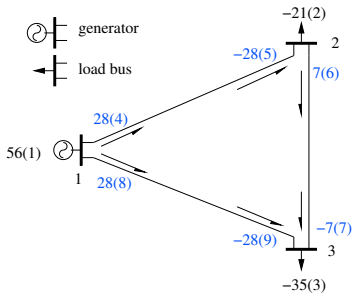

Fig. 1: The meter readings before attack. XX(Y) means that meter Y's reading is XX. A positive value means a power flow comes out of a bus, while a negative value means a power flow goes into a bus.

Now suppose an attacker can manipulate meters {2,3,5,9}. The attacker changes the readings of these four meters to the value shown in Fig. 2. The whole set of data is not consistent, and the control center knows that an attack is present. However, the readings on meters {1,4,6,7,8} are consistent, and they can determine a set of state variables, which corresponds to the load vector $\{bus2, bus3\}$=$\{21, 35\}$. The readings on meters {1,2,3,5,9} are also consistent, while they can determine a different set of state variables, which corresponds to the load vector $\{bus2, bus3\}$=$\{14, 42\}$. Under this scenario, even though the control center knows that four meters have been compromised, it has no way to identify which four have been manipulated. The compromised data can either be meters {2,3,5,9}, or meters {4,6,7,8}. That is, there are two feasible cases, one is meters {1,4,6,7,8} and the other is meters {1,2,3,5,9}. But the control center has no evidence to favor one over the other. In this example, the net effect of the attack is to have the control center guess whether there is a 7 unit load redistribution between bus 2 and bus 3 ({21,35} to {14,42}). To make this load redistribution undetectable, one has to compromise all nine meters except meter 1. However, we only need to compromise 4 meters to make this attack

[1]A set of measurements is said to make the system observable if the states of all the buses can be determined with these measurements. Otherwise, the system is said to be *unobservable* with this set of measurements.

unidentifiable.

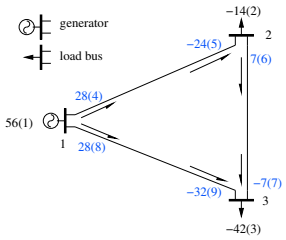

Fig. 2: Attack scenario 1.

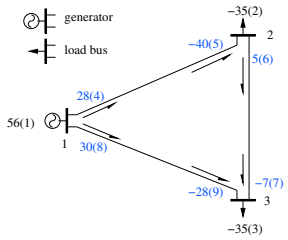

Fig. 3: Attack scenario 2.

Another similar attack is shown in Fig. 3. In this case, meters {2,5,6,8} are compromised. Similarly, the whole set of data is inconsistent, and the control center knows that an attack is present. However, the readings on meters {1,3,4,7,9} are consistent, and they can determine a set of state variables, which corresponds to the load vector $\{bus2, bus3\}$={21, 35}. Similarly, the readings on meters {2,3,5,6,8} are also consistent and produce a different set of state variables, which corresponds to the load vector $\{bus2, bus3\}$={35, 35}. Even though the control center knows that there are four compromised meters, it has no way to identify which four have been manipulated. The compromised data can either be meters {2,5,6,8}, or be meters {1,4,7,9}. That is, there are two feasible cases, one is meters {1,3,4,7,9} and the other is meters {2,3,5,6,8}. In this example, the net effect of the attack is to let the control center guess whether there is a 14 unit load increase on bus 2 (from 21 to 35). Again, only four meters need to be compromised to launch this unidentifiable attack.

In each of the example scenarios described above, there are some meters that have the same readings for the different feasible cases of an unidentifiable attack. With more common readings among different feasible cases, fewer meters need to be compromised to launch an unidentifiable attack.

IV. Optimization Strategy

Under an unidentifiable attack, the control center cannot identify which set of meters is manipulated, even though it knows that some meters are compromised. Suppose that under an unidentifiable attack, the control center finds l feasible attack cases (we will present algorithms to find all feasible cases in section V). To reduce the damage caused by such an unidentifiable attack, the control center has to consider all l feasible attack cases, and tries to find a solution such that the damage to the power system is as small as possible before the set of compromised meters can be identified and eliminated (it is possible that the attack cannot be identified without sending power engineers to conduct a physical check). A good strategy for the control center is to find a power generation solution such that the power system on the average operates at the most economical price without having to favor any particular attack case.

To evaluate whether a power generation solution is good or not, we need to assess the potential damage such a solution yields to each of the feasible attack cases. The damage mainly includes the followings:

- Power shedding on load buses. There is not enough power supply for load buses such that some load buses get less power than their demands which results in the tripping of circuit breakers to shed some loads;
- Overloading of transmission lines. The power flow on a transmission line may go beyond its capacity such that the line trips, possibly resulting in severe consequences, e.g., large area blackout;
- Overpowering on load buses. A load bus may be fed more power than its demand which can result in the power system operating at higher frequency than it can tolerate, tripping certain circuit breakers and causing blackout.

The cost of the whole power system consists of two components: one is related to the cost of power generation, and the other is related to the cost of damages mentioned above. Any good power generation solution should avoid overloading and overpowering scenarios since they both can cause severe consequences. In our proposed strategy to identify good power generation solutions during an unidentifiable attack, we propose to avoid any potential damages caused by overloading and overpowering by including constraints that prevent overloading and overpowering from occurring. Therefore, we only need to include the power generating cost and the penalty of load shedding in our overall cost. Since all l feasible attack cases are unidentifiable and equally possible in the view of the control center, it is reasonable to consider the average damage caused by a power generation solution to all feasible attack cases. We are to find a power generation solution such that the sum of the generating cost and the average damage caused by load shedding is minimized subject to certain constraints that prevent overloading and overpowering from happening. That is,

$$min: \quad \sum_{b_j \in G} C_j PG_j + \frac{1}{l} \sum_{k=1}^{l} D_{shed,k} \tag{2}$$

where $D_{shed,k}$ is defined as follows

$$D_{shed,k} = \sum_{b_i \in L} C_{shed,i} PS_{k,i} \tag{3}$$

The constraints are:

(1) Power shedding constraints:

$$0 \leq PS_{k,i} \leq PD_{k,i}, \forall b_i \in L, 1 \leq k \leq l \tag{4}$$

in which the positive $PS_{k,i}$ guarantees that there is no overpowering.

(2) Power flow and power injection constraints:

$$\sum_{j=1}^{n} V_i V_j (G_{ij} cos(\theta_i - \theta_j) + B_{ij} sin(\theta_i - \theta_j)) - PG_i + PD_{k,i} - PS_{k,i} = 0, 1 \leq i, j \leq n, 1 \leq k \leq l \tag{5}$$

$$\sum_{j=1}^{n} V_i V_j (G_{ij} sin(\theta_i - \theta_j) - B_{ij} cos(\theta_i - \theta_j)) - QG_i + QD_{k,i} - QS_{k,i} = 0, 1 \leq i, j \leq n, 1 \leq k \leq l \tag{6}$$

$$P_{ij} = -V_i^2 G_{ij} + V_i V_j (G_{ij} cos(\theta_i - \theta_j) + B_{ij} sin(\theta_i - \theta_j)), \forall t_{i\_j} \in T \tag{7}$$

$$Q_{ij} = V_i^2 B_{ij} + V_i V_j (G_{ij} sin(\theta_i - \theta_j) - B_{ij} cos(\theta_i - \theta_j)), \forall t_{i\_j} \in T \quad (8)$$

$$PL_{ij} = \sqrt{P_{ij}^2 + Q_{ij}^2} \quad (9)$$

(3) Line transmission capacity constraints:

$$-PL_{ij}^{max} \leq PL_{ij} \leq PL_{ij}^{max}, \forall t_{i\_j} \in T \quad (10)$$

(4) Generator capacity constraints:

$$PG_{g,min} \leq PG_g \leq PG_{g,max}, \forall b_g \in G \quad (11)$$

$$QG_{g,min} \leq QG_g \leq QG_{g,max}, \forall b_g \in G \quad (12)$$

In the above formulation, $PD_{k,i}$ and $QD_{k,i}$, which are determined by the kth feasible attack case, are known. $PS_{k,i}$ and $QS_{k,i}$ are variables. V_i and θ_i, $1 \leq i \leq n$, are auxiliary variables, which connect other variables via Eq(5), Eq(6), Eq(7), and Eq(8). After solving the minimization problem, we will get all the variables, including PG_j, $PS_{k,i}$, V_i and θ_i, $\forall b_j \in G$, $1 \leq k \leq l$, $1 \leq i \leq n$. The control center can then determine the amount of power generation on each generator and the amount of power supply on each load bus, and send these quantities as directives to the corresponding generators and load buses. This is how the control center responds to the unidentifiable attack.

V. Analysis of Unidentifiable Attack

When an unidentifiable attack occurs, the control center first has to detect the presence of an attack. One can use any typical bad data detection scheme proposed by previous work to detect the presence of an attack. Given a power system with n buses and m meters in AC model, as mentioned in Section II, the measurements $\mathbf{z} = (z_1, ..., z_m)$ are functions of the state variables $\mathbf{x} = (V_1, ..., V_n, \theta_1, ..., \theta_n)$. That is, $\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{e}$, where $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), ..., h_m(\mathbf{x}))$, are defined in the following four cases (assume no error, e.g., $\mathbf{e} = \mathbf{0}$):

(i) z is real power injection on bus i:

$$z_i = \sum_{j=1}^{n} V_i V_j (G_{ij} cos(\theta_i - \theta_j) + B_{ij} sin(\theta_i - \theta_j)) \quad (13)$$

(ii) z is reactive power injection on bus i:

$$z_{ii} = \sum_{j=1}^{n} V_i V_j (G_{ij} sin(\theta_i - \theta_j) - B_{ij} cos(\theta_i - \theta_j)) \quad (14)$$

(iii) z is real power flow from bus i to bus j:

$$z_{iii} = -V_i^2 G_{ij} + V_i V_j (G_{ij} cos(\theta_i - \theta_j) + B_{ij} sin(\theta_i - \theta_j)) \quad (15)$$

(iv) z is reactive power flow from bus i to bus j:

$$z_{iv} = V_i^2 B_{ij} + V_i V_j (G_{ij} sin(\theta_i - \theta_j) - B_{ij} cos(\theta_i - \theta_j)) \quad (16)$$

In case of errors or an attack, the detection scheme will compute L_2 norm $||\mathbf{z} - \mathbf{h}(\hat{\mathbf{x}})||_2$, where $\hat{\mathbf{x}}$ is the vector of estimated state variables obtained by a least square estimator. Then the L_2 norm is compared with a predefined threshold τ, and an attack is declared only if $||\mathbf{z} - \mathbf{h}(\hat{\mathbf{x}})||_2 > \tau$.

Next, the control center should discern if the attack is unidentifiable. It can draw such a conclusion by enumerating all feasible cases for the attack. If at least two feasible cases exist, then we can conclude that an unidentifiable attack has occurred. Otherwise, there is no unidentifiable attack.

In the following, we first make some assumptions and formulate the case enumerating problem. Then, we describe algorithms to enumerate all feasible cases, including a brute-force search algorithm and an empirical method that can speed up the brute-force search. Finally, we analyze the complexity and performance of the algorithms.

A. Assumption and problem formulation

First, we assume that the presence of bad measurements does not make the whole system unobservable, which excludes the scenario of undetectable attacks. Due to the limited resources of the attacker, we further assume that the attacker can at most compromise r meters; r is the attack capacity of the attacker, which we assume the control center knows by estimating the effort the attatter can take. Finally we assume that a set of meters, say set P, is protected by the power system operator.

Let set A be the set of all meters and set D be the set of bad meters that the control center deduces. With the above assumptions, the problem of finding all feasible cases under an unidentifiable attack can be formulated as follows:

Enumerate all different sets of D such that:

1) The meters in $A \backslash D$ make the whole power system observable;
2) The meters in $A \backslash D$ are consistent; that is, after solving the state estimation for the power system with meters in $A \backslash D$, the norm of residuals of these meters are zero or less than a predefined threshold τ.
3) The cardinality of D is smaller than or equal to r;
4) $D \cap P = \emptyset$;
5) Different set of D results in different state variables.

B. Enumerating Feasible cases

Given the assumptions and formulation above, our goal is to find all feasible cases that satisfy all the constraints. When r is small, we can use brute-force search to find all feasible cases. However when r is big, the brute-force search becomes very expensive, since its search time grows exponentially with increasing r. However, meters that are compromised in an unidentifiable attack are typically clustered. Thus, if one can identify an attack region where the compromised meters are located, then the search space can be reduced and hence the search process can be sped up.

1) Brute-force Search to Enumerate Feasible Cases: When r is small, we use brute-force search directly. In a brute-force search, every combination that meets all the constraints in the problem formulation is examined. The brute-force search algorithm works as follows.

Alg. 1: *Brute-force Search*
Input: r, the attacker's capability;

Set A, the set of all meters;
Set P, the protected set;
Output: A set F that contains all feasible sets of D.
1: $F = \emptyset$;
2: For i=1, r
3: Check every of $\binom{m}{i}$ bad data combinations, D, except those are supersets of any set in F;
4: If $(A \backslash D) \cap P = \emptyset$, then
5: If $M \backslash D$ pass the residual test, then
6: Put the bad data set D into F;
7: Endif
8: Endif
9: Endfor

In the above algorithm, the brute-force search actually does not check every combination, as shown in line 3. It does not check the combinations that have already been covered by previous combinations that are included in set F. If we have already found a feasible case with a set of meter readings D being declared as bad, then we do not need to check any other sets of meter readings D', where $D' \supset D$.

2) Locate Attack Region Then enumerate Feasible Cases: When the number of meters that an attacker can compromise, r, is large, the brute-force search approach becomes expensive. As we have indicated earlier, the set of compromised meters in an unidentifiable attack are typically located within a clustered region because their readings affect one another. Thus, if we can identify the attack region, then we only need to enumerate all feasible cases which only include meters within the attack region. Such a strategy greatly reduces the search space and hence the search time.

To identify the attack region, we can use existing algorithms based on Identification by Elimination (IBE), such as those discussed in [1], [2]. Though these algorithms cannot exactly identify all the bad data, especially those interacting[2] ones, these algorithms can give us some clues about the attack region.

We propose a three-step scheme for enumerating feasible cases for an unidentifiable attack. In our first step, we use the IBE algorithm since our goal is not to identify all bad data but to roughly locate the attack region. In the IBE algorithm, it first runs the least square estimator and then deletes the measurement with the largest residual, until the norm of the residuals is less than a pre-defined threshold τ. The IBE algorithm works as follows:

Step 1: IBE (Alg. 2)
1: $D = \emptyset$;
2: $A = \{$ all meters $\}$;
3: While the norm of residuals of meters in $A \backslash D \geq \tau$
4: Put the meter with the largest residual in D;
5: Run state estimation with $A \backslash D$;
6: Find the meter that has the largest residual;
7: End

[2]Multiple bad measurements are said interacting if the effect of the interaction can be added up and make a good measurement have the largest residual. Otherwise they are called non-interacting.

After executing *Step 1*, we identify a set of meters, D. We then check where the meters in set D are, and hence can roughly deduce where the attack region is. We define the *attack region*, R, which is a subgraph of the whole power system, using the following algorithm:

Step 2: Attack Region Identification (Alg. 3)
1: $R = \emptyset$;
2: For meter $a \in D$
3: if a is on bus i
4: $R = R \cup b_i \cup t_{i\_*}$;
5: else if a is on line t_{i_j}
6: $R = R \cup b_i \cup b_j \cup t_{i\_j}$;
7: endif
8: endfor

The rationale for the above algorithm is as follows. If a meter is on a bus, its reading (real/reactive) is the summation of all power flows (real/reactive) incident to that bus according to Eq(5) and Eq(6). If a meter is on a branch, its reading is a function of the state variables of the two end buses according to Eq(7) and Eq(8). For interacting bad data, the bad data nearby a good one can make the good one has the largest residual. Thus, if a data is eliminated in *Step 1*, it is either because the data is bad or its nearby data is bad. Therefore, in the attack region, we include both the data with largest residual and its neighbors that affect it directly. That is, if a meter on a bus is declared bad, we include both that bus and all the branches incident to that bus into the attack region; if a meter on a branch is declared bad, we include that line and two end buses into the attack region. Fig. 4 illustrates how the attack region is defined.

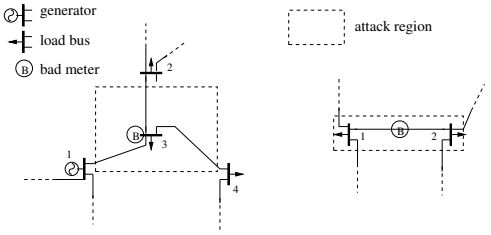

Fig. 4: Attack region illustration. On the left, a meter on bus 3 is declared bad; on the right, a meter on the branch between bus 1 and bus 2 is declared bad.

Therefore, by looking at the region where the bad data are located, we can roughly identify the attack region. However, we cannot guarantee that the attack region defined above includes all the bad meters, which is summarized in the following claim.

Claim: By eliminating the measurement with the largest residual until the remaining ones are consistent, the attack region defined above is not guaranteed to include all bad data. The proof is in the Appendix.

Remark: The example given in our proof in the Appendix is an extreme case. Consider a power system in AC mode,

there are four meters on each transmission line, with two at each end of the line, one for real power flow, the other for reactive power flow; and there are two injection meters on each bus, one for real power and the other for reactive power. To produce the above attack, the attacker has to compromise more than 2/3 of all meters[3]. If the attacker has that much attack resources, he may be able to launch an undetectable attack which may produce larger damages and hence he may not have the incentive to launch an unidentifiable attack.

After obtaining the attack region, we will do a brute-force search in it. However, this brute-force search algorithm is different from Alg. 1, since we need to consider the case that the detected attack region does not include all bad data as proved in the above claim. The algorithm is as follows.

Step 3: Brute-force Search in the attack region (Alg. 4)
Input: R, the detected attack region, with $|R|$ meters in it;
r, the attacker's capability;
Set A, the set of all meters;
Set P, the protected set;
Output: A set F that contains all feasible sets of D.

1: $F = \emptyset$;
2: For i=1, r
3: For every of $\binom{|R|}{i}$ bad data combination, D, except those are supersets of any set in F
4: If $(A \backslash D) \cap P = \emptyset$, then
5: If $M \backslash D$ pass the residual test, then
6: Put D into F;
7: Else
8: Run IBE and update D by including the data with largest residual;
9: Put D in F if $|D| \leq r$;
10: Endif
11: Endif
12: Endfor
13: Endfor

Although the detected attack region may not include all bad data, line 8 in *Step 3* is able to find bad data outside of the detected attack region. The three-step algorithm will find all the feasible cases.

C. Performance Analysis

Given a power system with m measurements and an attacker with capability r, the brute-force search (Alg. 1) is $O(\binom{m}{r})$. This is huge when m and r are both large. Therefore, Alg. 1 only works for either a small power system or an attacker with very limited capability.

When Alg. 1 is not applicable, we should utilize the three-step scheme (Alg. 2-4). Suppose there are $|R|$ meters in the located attack region R, then the complexity is $O(\binom{|R|}{r})$, which is much smaller than that of brute-force search, given that the compromised measurements in an unidentified attack is usually clustered and $|R|$ is much less than m. If the attack region R is not connected, i.e., there are more than one attack regions, we can apply Alg. 2-4 on each connected attack region. We can get the running time by dividing time complexity by CPU capacity.

Our method is better than all existing bad data detection methods in power system under an unidentifiable attack, since they cannot work in case of such attack. In an unidentifiable attack, there are more than one feasible cases. All existing methods can only find one solution, which means that they can at most find one feasible case. The attacker is always able to manipulate a set of measurements such that the set of bad measurements identified by an existing method is different from the set of manipulated measurements. Therefore, none of them can work in the scenario of an unidentifiable attack. In this sense, our method has already greatly eliminated false positive (FP) and false negative (FN) which all existing methods have. However, since our algorithms are heuristic, they may still have FP and FN. If the detected attack region contains all the bad data, there will be neither FP nor FN. Even if the detected attack region does not contain all the bad data, Alg. 4 is able to find some bad data outside of the detected attack region. We believe that these two facts will reduce the rate of FP and FN. As shown in the evaluation next section, there is neither FP nor FN in dealing with the four unidentifiable attacks created with Matpower [23].

VI. Evaluation

In this section, we present the results of several experiments that we conduct. First, we generate four unidentifiable attacks in two bus systems. Second, we locate the attack region and enumerate all possible cases using Alg. 2-4 presented in Section V; at the same time, we show that the IBE method does not correctly identify the set of bad data, especially if the bad data interact with one another. Finally, we evaluate the operating cost of the power system based on the optimization strategy we present in Section IV. Our results show that the optimization strategy we propose for dealing with unidentifiable attacks is a viable solution.

A. Generating unidentifiable attacks

We generate one Type I attack and one Type II attack in each of 14-bus system and 30-bus system, whose topologies are shown in Fig. 5 and Fig. 6 respectively. We generate malicious data using the Matpower tool [23], which is developed to solve power flow problems. Given the topology of a power system, the transmission line characteristics and power loads on buses (load vector), Matpower is able to output the power flow on transmission lines and power injections on buses. We first input one load vector into Matpower and record the first set of meter readings. Then, we feed another load vector into Matpower and record the second set of meter readings. Comparing the two sets of readings, some of them are the same in both sets while others are different. We merge the two sets of readings as follows: for those meter readings that are different in these

[3]For a n bus system with $|T|$ branches, there are $2n + 4|T|$ meters. The attack has to compromise a portion of $\frac{2n+4|T|-(2n-1)}{2n+4|T|} > \frac{2}{3}$, since $|T|$ is usually greater than n.

two cases, we keep some obtained from the first set, and some from the second set. In this way, we can get an unidentifiable attack scenario with two feasible cases.

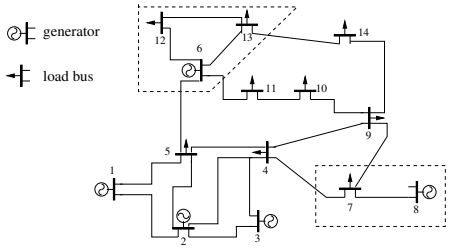

Fig. 5: The topology of 14-bus system in Matpower.

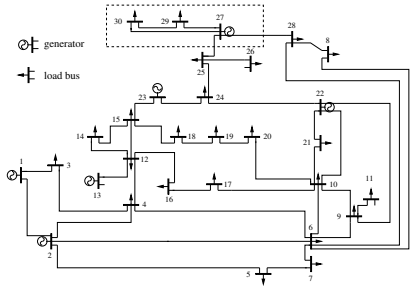

Fig. 6: The topology of 30-bus system in Matpower.

We still consider the two types of attack scenarios illustrated in Section III using the AC mode. For Type I load redistribution attack scenario, we introduce the second feasible case by increasing the load on one bus by a certain amount and decreasing the load on another bus by the same amount. Thus, together with the original case, we obtain an unidentifiable attack with two feasible cases. For Type II load increase attack scenario, we introduce the second feasible case by increasing the load only on one bus by a certain amount; similarly, we get an unidentifiable attack with two feasible cases.

We first generate two unidentifiable attacks in 14-bus system, one for each type. For the Type I attack, the compromised meters are listed in Table I, and the rest meters remain intact and their readings are omitted. The meter readings before the attack is based on the real power load vector $\{bus12, bus13\} = \{6.1, 13.5\}$, and the meter readings after the attack are based on the real power load vector $\{bus12, bus13\} = \{16.1, 3.5\}$; the loads in other buses have the same values as those in the Matpower distribution package. For the Type II attack, the compromised data are listed in Table II. The meters readings before the attack is based on the real power $bus7 = 0$, and the readings after the attack is based on the real power $bus7 = 10$; the loads in other buses have the same values as those in the Matpower distribution package.

TABLE I: Type I attack where seven meters are changed.

Meters	Before attack	After attack
PI on bus13	-13.5	$\mathbf{-3.5}$
PL from bus6 to bus12	8.05	**12.32**
QL from bus6 to bus12	3.31	**4.43**
PL from bus13 to bus6	-17.93	$\mathbf{-14.34}$
QL from bus13 to bus6	-9.99	$\mathbf{-9.33}$
PL from bus12 to bus13	1.87	$\mathbf{-3.96}$
QL from bus13 to bus12	-1.53	$\mathbf{-2.41}$

TABLE II: Type II attack where two meters are changed.

Meters	Before attack	After attack
PL from bus7 to bus8	0	$\mathbf{-10.00}$
PL from bus8 to bus7	0	**10.00**

We also generate two unidentifiable attacks using the 30-bus system in Matpower. Both Type I and Type II attacks are shown in Table III. Columns 3 show the changed meters for the Type I attack scenario. The meters in other region remain unchanged. The meter readings before the attack are based on the real power load vector $\{bus29, bus30\} = \{2.4, 10.6\}$, and the meter readings after the attack are based on the real power load vector $\{bus29\_1, bus30\_2\} = \{12.4, 0.6\}$; the loads in other buses have the same values as those in the Matpower distribution package. Column 4 shows the changed meters for the Type II attack scenario. The meter readings before the attack are obtained when the load of bus 30 is 10.6, the meter readings after the attack are obtained when the load of bus 30 is 20.6; the remaining loads are the same as those in the Matpower distribution package.

TABLE III: Type I and II attacks in 30-bus system. The bold ones are the changed.

Meters	Before attack	After attack (Type I)	After attack (Type II)
PI on bus27	26.91	26.91	**37.63**
QI on bus27	11.39	11.39	**12.74**
PI on bus29	-2.4	$\mathbf{-12.4}$	-2.4
PL from bus27 to bus29	6.17	**9.32**	**10.56**
QL from bus27 to bus29	1.68	1.68	**2.27**
PL from bus29 to bus27	-6.08	$\mathbf{-9.12}$	-6.08
PL from bus27 to bus30	7.12	**3.96**	**13.46**
QL from bus27 to bus30	1.67	1.67	**2.45**
PL from bus30 to bus27	-6.95	$\mathbf{-3.91}$	-6.95
PL from bus29 to bus30	3.68	$\mathbf{-3.28}$	**7.90**
QL from bus29 to bus30	0.61	0.61	**0.88**
PL from bus30 to bus29	-3.65	**3.31**	-3.65

B. Locate the attack region and enumerate feasible cases

For the four attacks listed above, we first use Alg. 2 to get the deleted set D. The deleted set for each attack is listed in Table IV, where where $bsxx\_1/2$ means PI/QI on bus xx respectively, and $brxx\_1/2/3/4$ means the PL/QL on the from-bus and to-bus of branch xx respectively. In 14-bus power system, $br12 = (bus6, bus12)$, $br13 = (bus6, bus13)$, and $br19 = (bus12, bus13)$. In 30-bus power system, $br37 = (bus27, bus29)$, $br38 = (bus27, bus30)$ and $br39 = (bus29, bus30)$. In order to show the effectiveness of IBE, we also list the real compromised set for each attack.

TABLE IV: The deleted sets and compromised set for four attacks.

	Attack	Deleted set	Compromised set
14	Type I	$bs12\_1$, $br19\_3$, $br12\_3$ $br13\_1$, $br19\_2$, $br13\_2$ $br12\_4$	$bs13\_1$, $br12\_1$, $br12\_2$ $br13\_3$, $br13\_4$, $br19\_1$ $br19\_4$
	Tpye II	$bus7\_1$, $bus8\_1$	$br14\_1$, $br14\_3$
30	Type I	$bs30\_1$, $br38\_2$, $br37\_2$ $br38\_4$, $br37\_4$, $br39\_2$ $br39\_4$	$bs29\_1$, $br37\_1$, $br37\_3$ $br38\_1$, $br38\_3$, $br39\_1$ $br39\_3$
	Type II	$bs30\_1$, $br38\_3$, $br37\_3$ $br39\_3$, $br37\_4$, $br39\_4$ $br38\_4$	$bs27\_1$, $bs27\_2$, $br37\_1$ $br37\_2$, $br38\_1$, $br38\_2$ $br39\_2$

As we can see in Table IV, the IBE method cannot identify

the real compromised measurements, and there is even no common element between the deleted set and compromised set. This illustrates that the IBE method cannot identify the interacting bad measurements as mentioned in previous work, such as [3]–[7]. Neither can other bad data detection methods, such as NQC, HIT and COI mentioned in Section II, as explained in Section V-C.

Next we apply Alg. 3 on the deleted set listed in Table IV to get the attack region. Though the deleted sets do not even contain one real compromised measurement, the attack regions obtained from Alg. 3 do contain all the compromised measurements. The four attack regions are shown in the dashed rectangle or trapezoid in Fig. 5 and Fig. 6 (in 30-bus system, the two attack regions are the same).

Finally, we apply Alg. 4 directly to enumerate all feasible cases. For all four unidentifiable attacks, we are able to find out that there are only two feasible cases for each attack, just as same as described in Section VI-A. Furthermore, we can tell the attack type of each attack after obtaining its feasible cases. Let us take the Type II attack in 14-bus system as an example to show the effectiveness of Alg. 4. In the 14-bus system, there are 14 buses and 20 branches. As we assume 4 meters on each branch and 2 meters on each bus, there are 108 meters in total. To calculate the time complexity of the enumerating algorithms in Section V, let us further assume that the attacker can at most compromise 8 meters and there is no protected meter in the power system ($P = \emptyset$). For the brute-force search algorithm, the search space of $\sum_{i=1}^{8} \binom{108}{i}$, is still huge, not to mention all the computations required for state estimation and residual checking. While in the attack region, there are only 16 meters. By localizing the attack region first, the search space is greatly reduced to at most $\sum_{i=1}^{8} \binom{16}{i}$. Actually, the search space is even far smaller than $\sum_{i=1}^{8} \binom{16}{i}$ for two reasons. The first is that we have already found one feasible case via the IBE method. The second is, once a feasible case is found, the brute force search can skip some combinations. For instance, if a solution with 3 bad data has been identified, we do not need to check all bad data combinations which include those 3 bad data.

C. Optimization on the cost

We evaluate our optimization problem using the four unidentifiable attacks we discussed in Section VI-A. For each of the unidentifiable attack, we have already known that there are two feasible cases and what they are. Thus, we only need to feed these feasible cases into the objective function Eq(2) and try to minimize it. We use the free software IPOPT [24] to solve the nonlinear optimization problem. In our analysis, we set the power shedding cost as five times as the cost of the most expensive generator. This setting is reasonable, since the power shedding cost must be higher than the cost of any generator; otherwise, the generator will choose not to satisfy the load demand even it still has available capacity.

1) Type I attack in 14-bus system: In this attack, we change 7 meters as shown in Table I. Under this unidentifiable attack, the control center may either conclude that the power demands of bus 12 and bus 13 are 6.1 and 13.5 (case 1), or they are 16.1 and 3.5 (case 2). These two load vectors are fed together with the constraints into IPOPT to determine the optimal state variables, the voltage and phase on each bus, which can minimize the total cost. In the original Matpower packet, all line capacities are 9900 MVA. In order to examine the impact of line capacities, we adjust the line capacities for the following branches: branch 12, branch 13, and branch 19 to 10 MVA, 25 MVA and 10 MVA respectively. The cost comparison is listed in Table V, in which solution 1 is the optimal solution based on case 1, and solution 2 is the optimal solution based on case 2. "Over-load" means that if the control center gets a solution based on case 1 but it is actually case 2, then some branches will exceed their line capacities. As we can see, our solution is the best, given that the control center cannot favor one case over the other.

TABLE V: The cost comparison for type I attack in 14-bus system.

	If case 1	If case 2	Average
Solution 1	8083	Over-loaded	NA
Solution 2	8594	8594	8594
Our solution	8573	8595	8584

2) Type II attack in 14-bus system: Table II shows type I attack in 14-bus system. The two feasible cases are: the real power demand on bus 7 is either 0 (case 1) or 10 (case 2). In this example, we do not change any line capacity. The cost comparison is listed in Table VI, where "Over-powered" means that if the control center gets a solution based on case 2 but it is actually case 1, then some buses will get more power than their demands. As we can see, our solution is still the best, given that the control center cannot favor one case over the other.

TABLE VI: The cost comparison for type II attack in 14-bus system.

	If case 1	If case 2	Average
Solution 1	8083	10208	9146
Solution 2	Over-powered	8486	NA
Our solution	8087	10081	9084

3) Two attacks in 30-bus system: The evaluation for the two attack in 30-bus system is similar to that in 14-bus system. Here we omit the details but only keep the main results. In the type I attack, the two feasible cases are: the power demands of bus 29 and 30 are 2.4 and 10.6 (case 1), or they are 12.4 and 0.6 (case 2). And we adjust the line capacities for the following branches: branch 37, branch 38, and branch 39 from the original value of 16 MVA to 4 MVA, 8 MVA and 3 MVA respectively. The cost comparison is listed in Table VII. In the type II attack, the two feasible cases are: the real power demand on bus 30 is either 10.6 (case 1) or 20.6 (case 2), and the cost comparison is shown in Table VIII.

Again, we can see that our solutions is the best on average among all the solutions, which shows that our optimization strategy is indeed viable and effective.

TABLE VII: The cost comparison for type I attack in 30-bus system.

	If case 1	If case 2	Average
Solution 1	635.0	Over-loaded	NA
Solution 2	693.1	693.1	693.1
Our solution	680.3	693.7	687.0

TABLE VIII: The cost comparison for type II attack in 30-bus system.

	If case 1	If case 2	Average
Solution 1	581.2	775.0	678.1
Solution 2	Over-powered	623.6	NA
Our solution	581.3	750.7	666.0

VII. Conclusion

In this paper, we introduce the concept of unidentifiable attack in power system, which is a new type of attack never proposed before. In such an attack, the control center cannot obtain a deterministic state estimation, since there may be several possible cases and the control center cannot simply favor one over the others. We then formulate an optimization strategy from the perspective of the control center to deal with an unidentifiable attack such that the average damage caused by the attack can be minimized. Furthermore, we propose a three-step scheme that allows us to find all feasible cases under an unidentifiable attack, in which we locate attack region first and hence significantly reduce the search space when compared to the search space using the brute-force search scheme directly. We evaluate and validate our optimization strategy and enumerating scheme using 14-bus and 30-bus power systems. Our results show that the minimal cost strategy allows the power system to operate at minimal cost irrespective of what the exact attack case is.

ACKNOWLEDGMENT

The authors would like to thank Prof. Bruce McMillin for his valuable suggestions and all the reviewers for their helpful comments. This project was supported in part by US National Science Foundation grants CNS-1117412 and CAREER Award CNS-0747108.

References

[1] F. Schweppe and J. Wildes, "Power system static-state estimation, Part I II & III," *IEEE Trans. on Power Apparatus and Systems*, vol. 89, no. 1, 1970.

[2] E. Handschin, F. Schweppe, J. Kohlas, and A. Fiechter, "Bad data analysis for power system state estimation," *IEEE Trans. on Power Apparatus and Systems*, vol. 94, no. 2, 1975.

[3] H. Merrill and F. Schweppe, "Bad data suppression in power system static state estimation," *IEEE Trans. on Power Apparatus and Systems*, vol. 90, no. 6, 1971.

[4] T. Van Cutsem, M. Ribbens-Pavella, and L. Mili, "Hypothesis testing identification: a new method for bad data analysis in power system state estimation," *IEEE Trans. on Power Apparatus and Systems*, vol. 103, no. 11, 1984.

[5] E. Asada, A. Garcia, and R. Romero, "Identifying multiple interacting bad data in power system state estimation," *IEEE Power Engineering Society General Meeting*, 2005.

[6] S. Gastoni, G. Granelli, and M. Montagna, "Multiple bad data processing by genetic algorithms," in *IEEE Power Tech Conference*, vol. 1, 2003.

[7] A. Monticelli, F. Wu, and M. Yen, "Mutiple Bad Data Identification for State Estimation by Combinatorial Optimization," *IEEE Trans. on Power Delivery*, vol. 1, no. 3, 1986.

[8] Y. Liu, M. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," *Proceedings of ACM CCS*, 2009.

[9] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power system," *IEEE Trans. on Smart Grid*, vol. 2, no. 2, 2011.

[10] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious data attacks on SmartGrid state estimation: Attack strategies and countermeasures," *Proc. of IEEE SmartGridComm*, 2010.

[11] T. Kim and H. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. on Smart Grid*, vol. 2, no. 2, 2011.

[12] L. Mili, M. Ribbens-Pavella, and T. Van Cutsem, "Bad data identification methods in power system state estimation-a comparative study," *IEEE Trans. on Power Apparatus and Systems*, no. 11, 1985.

[13] H. Wang, B. Sheng, and Q. Li, "TelosB implementation of elliptic curve cryptography over primary field," College of William and Mary, Tech. Rep. WM-CS-2005-12, October 2005.

[14] H. Wang and Q. Li, "Efficient implementation of public key cryptosystems on MICAz and TelosB motes," College of William and Mary, Tech. Rep. WM-CS-2006-7, October 2005.

[15] H. Wang, B. Sheng, C. C. Tan, and Q. Li, "WM-ECC: an elliptic curve cryptography suite on sensor motes," College of William and Mary, Tech. Rep. WM-CS-2007-11, 2007.

[16] T. Gamage and B. McMillin, "Nondeducibility-based analysis of cyber-physical systems," *Critical Infrastructure Protection III*, 2009.

[17] H. Wang, C. Tan, and Q. Li, "Snoogle: A search engine for the physical world," *IEEE INFOCOM*, 2008.

[18] S. Ren, Q. Li, H. Wang, X. Chen, and X. Zhang, "Analyzing object detection quality under probabilistic coverage in sensor networks," *Int'l Workshop Quality of Service (IWQoS)*, 2005.

[19] Z. Ling, J. Luo, W. Yu, X. Fu, D. Xuan, and W. Jia, "A new cell counter based attack against Tor," *Proceedings of ACM CCS*, 2009.

[20] D. Xuan, R. Bettati, and W. Zhao, "A gateway-based defense system for distributed DoS attacks in high-speed networks," *Workshop on Information Assurance and Security*, vol. 1, 2001.

[21] M. Ding, F. Liu, A. Thaeler, D. Chen, and X. Cheng, "Fault-tolerant target localization in sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, 2007.

[22] K. Xing, M. Ding, X. Cheng, and S. Rotenstreich, "Safety warning based on highway sensor networks," *IEEE Wireless Communications and Networking Conference*, vol. 4, 2005.

[23] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. on Power Systems*, no. 99, 2011.

[24] A. Wachter and L. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, 2006.

VIII. Appendix

Proof: Suppose the whole system has n buses with m measurements, then the system has $2n-1$ state variables. The adversary has modified $m-2n+1$ measurements, and the remaining $2n-1$ measurements are all critical and can make the system observable (and hence satisfy *Assumption 1*). Now the $2n-1$ measurements can give a deterministic solution for the state variables, but any $2n-2$ measurements of the same set cannot. Now let us select $2n-2$ measurements out of the set of $2n-1$ measurements, and refer to the remaining one measurement as R. The $2n-2$ measurements yield many feasible solutions of state variables for that particular power system. We select one of the feasible solutions, which is different from the one obtained using the set of $2n-1$ measurements. The adversary then modify the $m-2n+1$ measurements based on this selected feasible solution. Obviously, the largest residual will then occur on the meter that measures R. After eliminating R, the rest of the measurements are consistent. *Step 2* only identifies the attack region as a small neighborhood around R and hence does not include all bad data in the set which contains the $m-2n+1$ readings. □