Detecting False Data Injection Attacks on DC State Estimation

Authors: Rakesh B. Bobba, Katherine M. Rogers. Qiyan Wang, Himanshu Khurana, Klara Nahtstedt and Thomas J. Overbye

Presenter: Yang Yu

Submitted in Partial Fulfillment of the Course Requirements for
ECEN 689: Cyber Security of the Smart Grid
Instructor: Dr. Deepa Kundur

Overview

- Introduction
- Background Knowledge
- Motivation and Approach
- Assessment and Future Works

1.Introduction

- State estimation is an important power system application to estimate the state of the power transmission system
- Traditionally, it was assume the techniques used to detect and identify the bad measurements can also detect malicious attacks
- However, recent work by Liu and others [2] shows that a well organized attack could inject false data without being detected.

1.Introduction

- Two approaches to protect state estimation:

  1) design robust control algorithms that can detect or tolerate malicious data modification

  2) protect the sensor measurements and other data from being manipulated

1.Introduction

- Main contribution of the authors:

  1) Showing that protecting a set of basic measurements is necessary and sufficient to detect the previous undetectable false data injection attacks on DC state estimation

  2) Determining how to find these measurements.

2. Background knowledge

- m — The number of measurements
- n — The number of state variables
- **H** — m × n Jacobian matrix representing the topology
- **x** — n × 1 vector of state variables
- **z** — m × 1 vector of measurements
- **e** — m × 1 vector of measurements errors
- $\hat{\mathbf{x}}$ — n × 1 vector of estimated state variables
- **W** — m × m diagonal matrix
- τ — Threshold for the L2 norm based detection of bad measurements
- $\mathbf{z}_a$ — m × 1 measurement vector with bad measurements
- **a** — m × 1 attack vector
- **c** — n × 1 vector of estimation errors introduced due to **a**
- **M** — The set of sensor or measurements indices
- **V** — The set of state variable indices
- $I_{\bar{m}}$ — The set of indices of protected sensor measurements
- $I_{\bar{v}}$ — The set of indices of independently verified state variables
- I_m — The set of indices of potentially manipulated measurements
- I_v — The set of indices of potentially manipulated state variables
- **p** — The number of protected measurements
- **q** — The number of independently verified state variables

2.Background Knowledge

- Common formulation of the state estimation problem when using a DC power flow model :

$$z = Hx + e$$

where **x** = $(x_1, x_2, \ldots, x_n)^T$ represents the true states of the system, **z** = $(z_1, z_2, \ldots, z_n)^T$ represents the sensor measurements, H is an m by n Jacobian matrix, representing the topology and **e** = $(e_1, e_2, \ldots, e_n)^T$ is the random errors in the measurement.

- Assuming **e** is zero mean and Gaussian, the state estimation is given by:

$$\hat{x} = (H^T W H)^{-1} H^T W z$$

Where **W** is a diagonal matrix whose elements are the measurement weights.

2.Background Knowledge

- Bad Measurements Detection:

$$\left\| z - H\hat{x} \right\| > \tau \qquad \text{bad data is present}$$

$$\left\| z - H\hat{x} \right\| < \tau \qquad \text{bad data is not present}$$

- False Data Injection Attacks:

  If attackers could find a attack vector **a** = **Hc**, then the resulting manipulated measurement $\mathbf{z}_a = \mathbf{z} + \mathbf{a}$ will able to pass the bad measurement detection algorithm

2.Background Knowledge

- Proof:

$$\left\| z_a - H\hat{x}_{bad} \right\| = \left\| z + a - H(\hat{x} + c) \right\|$$

$$= \left\| z - H\hat{x} + (a - Hc) \right\| = \left\| z - H\hat{x} \right\|$$

when **a** = **Hc**

As we can see, the L_2 norm remains the same as normal situation, so that the false data injection attacks bypass the detection.

3. Motivation and Approach

- An obvious approach to defend such attacks is to protect all the sensor measurements. However, this is not always feasible. So the author aim to identify

  1) a set of sensors
  2) a set of state variables

  such that, when the measurements from the sensors in the chosen set are protect and the values of state variables from the chosen set can be verified independently, then an attacker cannot find any attack vectors

  Furthermore, the author also identify the smallest of such sets.

3. Motivation and Approach

- Adversary Model:

  1) the adversary has access to the topology matrix **H** which is determined by the power network topology and line impedances

  2) the adversary has the capability to manipulate sensors measurements, either by compromising the sensors or the communication between the sensors and the control center

  3) the attacker is restricted to compromising only the specific sensors denoted by the set I_m ,which is because the other measurements are protected by the grid operator.

  4) assume the grid operator can independently verify the values of a few chosen state variables, denoted by $I_{\bar{v}}$,and the attacker, in order to avoid detection, is constrained not to inject false data into those variables

3. Motivation and Approach

- Let $I_{\bar{m}}$ denote the set of indices of measurements that are protected by the grid operator. Let $I_{\bar{v}}$ denote the set of indices of states variables that the attackers cannot inject false data into. Under the previous assumptions,

  A successful false data injection attack have satisfied the following three conditions:

  1) $\mathbf{a} = \mathbf{Hc}$

  2) $\mathbf{a}_i = 0$ for $i \in I_{\bar{m}}$

  3) $\mathbf{c}_j = 0$ for $j \in I_{\bar{v}}$

  where $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)^T$ is the attack vector and $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)^T$ represents the estimation error introduced by the attack vector **a**.

3. Motivation and Approach

- As a result, the operation need to identify smallest set of sensors, $I_{\bar{m}}$, and smallest set of state variables, $I_{\bar{v}}$,such that an attacker cannot find an attack vector that satisfies the above three conditions.

- Two approaches to identifying optimal $I_{\bar{m}}$ and $I_{\bar{v}}$:

  1) Brute-Force Search
  2) Protecting Basic Measurements

3. Motivation and Approach

- Approach I : Brute-Force Search:

  Let p be the number of the protected measurements, q be the number of independently verified state variables. The grid operator can pick at random a fixed q out of n state variable and p out of m sensors, and check if any attack vectors that satisfy the above three conditions.

  Let H = (h_1, h_2, . . . , h_n), where h_i denotes the ith column vectors of H. Set H_s = (h_{j1} , h_{j2} , . . . , h_{jn-q}) and c_s = (c_{j1} , c_{j2} , . . . , c_{jn-q})^T where $j_i \notin I_{\bar{v}}$ for $1 \le i \le n-q$. Let $P_s = H_s(H_s^T H_s)^{-1} H_s^T$ and $B_s = P_s - I$. Also let B_s = (b_1, b_2, . . . , b_m), where b_i denote the column vectors of B_s. Set B'_s = (b_{i1} , b_{i2} , . . . , b_{im-p}) and a_0 = (a_{i1} , a_{i2} , . . . , a_{im-p})^T , where $i_r \notin I_{\bar{m}}$ for $1 \le r \le m-p$

$$a = Hc \Leftrightarrow B_s a = 0 \Leftrightarrow B'_s a' = 0$$

3. Motivation and Approach

- Since B'_s is a m by m-p matrix, the problem will become if the select measurements and variables ensure that the rank of B'_s is m-p. If rank of B'_s is m-p, then the false data injection attack will be detected.

- The brute-force approach have to search through $C_m^p \times C_n^q$ combinations for a given p and q, where $1 \le p \le m$ and $1 \le q \le n$

- Reduce combinations: $1 \le q \le \frac{n}{10}$, $1 \le p \le \frac{m}{2}$

- Although we could reduce the combinations, brute-force approach still takes lots of time, a simulation on the IEEE 14-bus system could not be done in two days.

3. Motivation and Approach

Number of Protected Sensors	Number of Verifiable State Variables	Number of Defensible Configurations	Percentage of Defensible Configurations
7	0	0	0
8	0	329245	14%
9	0	1991771	35%
6	1	0	0
7	1	18954135	75%
6	2	12288444	62%

Table1: Number of protected sensors and verifiable state variables needed to detect false data injection attacks for IEEE 9-bus system [1]

3. Motivation and Approach

- Approach II: Protecting Basic Measurements
- A set of basic measurements in state estimation is a minimum set of measurements which is sufficient to ensure observability
- **Theorem**: When there are no verifiable state variables, it is necessary and sufficient to protect a set of basic measurements in order to be able to detect false data injection attacks.
- **Corollary**: If there are q verifiable state variables it is necessary and sufficient to protect a set of basic measurements corresponding to the remaining n-q state variables in order to be able to detect false data injection attacks.

3. Motivation and Approach

- From the theorem and the corollary, the author concludes that the minimum number of protected or verifiable quantities is equal to n, i.e. the number of state variables.

- Determining the protected set:

  Other than the brute-force approach, the author present a more efficient approach to find the basic measurements.

  First a LU decomposition is performed on H

$$\widetilde{H} = P \cdot H = L_{AA} \cdot U_b$$

  where P is a row permutation matrix, which maps the rows of H to $\overline{H}$

3. Motivation and Approach

- The new basis is given by

$$L'_{AA} = \begin{bmatrix} I_n \\ R \end{bmatrix}$$

where I_n is the n by n identity matrix. Rows of I_n are correspond to the n basic measurements, and rows of R are correspond to redundant measurements.

We use P to map the first n measurements in the new basis back to the original measurements. This gives us one set of basic measurements.

Other basic measurements sets could be derived after this.

3. Motivation and Approach

- L'_{AA} could tell which measurements we may switch out to obtain a new set of basic measurements.
- Example:

$$L'_{AA} = \begin{array}{c} p_2 \\ p_3 \\ p_{24} \\ p_{12} \\ p_{34} \\ p_{23} \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & -0.5 & 0.5 & 0 \end{pmatrix}$$

The basic measurements are (p_2, p_3, p_{24}, p_{12}), but p_3 and p_{24} can be replaced by either p_{34} or p_{23}. Then we could have other basic measurements such as (p_2, p_{34}, p_{23}, p_{12}).

The key is the rows switched out the redundant measurement set must be linearly independent.

3. Motivation and Approach

- In summary, there are many choices of sets of basic measurements could be protected. The selection process may proceed as follows:

  1) Determine a satisfactory set of basic measurements using the above method

  2) Decide which state variables will be made verifiable, then add them to the candidate set and optionally remove an equal number of protected measurements.

4.Assessment and Future Works

- The authors successfully shows that protecting a set of basic measurements is necessary and sufficient to detect the previous undetectable false data injection attacks on DC state estimation and provide several efficient approaches to find these measurements. Some proof is not given because they could be found in the reference papers.

- The paper enables future researchers to consider:

  1) Topology changes: topology may changes either due to the normal line outages or malicious attacks.

  2) Generic False Data Injection Attacks: even the attacker could not find a attack vector such a = Hc, he may still able to find a such that

$$\left\| z - H\hat{x} + (a - Hc) \right\| \le \tau$$

Reference

- [1] Rakesh B. Bobba, Katherine M. Rogers, Qiyan Wang, Himanshu Khurana, Klara Nahrstedt and Thomas J. Overbye, “Detecting False Data Injection Attacks on DC State Estimation”, in *Preprints of the First Workshop on Secure Control Systems, CPSWEEK 2010, 2010.*

- [2] Y. Liu, M. K. Reiter, and P. Ning, “False data injection attacks against state estimation in electric power grids,” in CCS ’09: Proceedings of the 16th ACM conference on Computer and communications security. New York, NY, USA: ACM, 2009, pp. 21–32.

- [3] J. Chen and A. Abur, “Placement of pmus to enable bad data detection in state estimation,” Power Systems, IEEE Transactions on, vol. 21, no. 4, pp. 1608–1615, Nov. 2006.

- [4] J. London, L. Alberto, and N. Bretas, “Analysis of measurement-set qualitative characteristics for state-estimation purposes,” Generation, Transmission Distribution, IET, vol. 1, no. 1, pp. 39 –45, January 2007.