Algorithms for Power Grid State Estimation after Cyber-Physical Attacks *

Saleh Soltan
Electrical Engineering
Columbia University
New York, NY
saleh@ee.columbia.edu

Gil Zussman
Electrical Engineering
Columbia University
New York, NY
gil@ee.columbia.edu

ABSTRACT

We present methods for estimating the state of the power grid following a cyber-physical attack. We assume that an adversary attacks an area by: (i) disconnecting some lines within that area (*failed lines*), and (ii) obstructing the information from within the area to reach the control center. Given the phase angles of the nodes outside the attacked area under either the DC or AC power flow models (before and after the attack), the provided methods can estimate the phase angles of the nodes and detect the failed lines inside the attacked area. The novelty of our approach is the transformation of the line failures detection problem, which is combinatorial in nature, to a convex optimization problem. As a result, our methods can detect any number of line failures in a running time that is independent of the number of failures and is solely dependent on the size of the attacked area.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Reliability, availability, and serviceability; G.2.2 [**Discrete Mathematics**]: Graph Theory—*Graph algorithms, Network problems*

Keywords

Power Grids; Cyber Attacks; Physical Attacks; Information Recovery; Graph Theory; Algorithms

1. INTRODUCTION

Power grids are vulnerable to cyber-physical attacks. A cyber-physical attack may sabotage the information flow to the control center and cause physical failure by remotely disconnecting some of the lines. The recent cyber-physical attack on the Ukranian grid revealed the devastating consequence of such an attack on power grids [2].

We assume that an adversary attacks an area by: (i) disconnecting some lines within that area (*failed lines*), and (ii) obstructing the information from within the area to reach the control center. As a result of an attack, some lines get disconnected, and the phase angles and the status of the lines within the *attacked area* $H = (V_H, E_H)$ become unavailable (Fig. 1). Our objective is to recover the phase angles and detect the disconnected lines using the information available outside of the attacked area.

A line failure in the power grid results in changes to flows and phase angles throughout the grid. We use this property and show that it is possible to estimate the state in the attacked area using the information available outside of the area. We *develop methods for retrieving information from the attacked area* by applying matrix analysis and graph theoretical tools to the matrix representation of the DC equations. In particular, we demonstrate that under the DC power flow model, solving a convex optimization problem can exactly recover the phase angles and detect any number of line failures when the attacked area has certain topological properties. The novelty of our approach is the transformation of the line failures detection problem, which is combinatorial problem, to a convex optimization problem.

We also study the same problem when the phase angle measurements are noisy. We show that by relaxing some of the constraints, the same optimization problem can be used for the information recovery in the noisy scenarios. We numerically evaluate the performance of this method and show that if the Signal to Noise Ratio (SNR) is high enough, it can successfully recover the phase angles and detect the line failures inside the attacked area.

Finally, we adapt a similar idea to show that the same convex optimization method can be used to estimate the phase angles and detect the line failures accurately when the phase angles are given under the AC power flow model. We evaluate the performance of the method in the IEEE 118-bus system, and show that it estimates the phase angles of the buses with less that 1% error, and can detect the line failures with 80% accuracy for all single and double line failure scenarios.

The considered problem is very similar to the problem of line failure detection using phase angle measurements [5,10–12]. Up to two line failures detection, under the DC power flow model, is studied [10,11]. Since the provided methods in [10,11] are brute force search methods that need to search the entire failure space, the running time of these methods grows exponentially as the number of failures increases. Hence, these methods cannot be generalized to detect higher order failures. To the best of our knowledge, the methods provided in this work, are the only methods that can detect line failures in a polynomial running time independently of the number of line failures.

*This extended abstract provides a short summary of the papers that appeared in [7–9]. This work was supported in part by DARPA RADICS under contract #FA-8750-16-C-0054, funding from the U.S. DOE OE as part of the DOE Grid Modernization Initiative, and DTRA grant HDTRA1-13-1-0021.

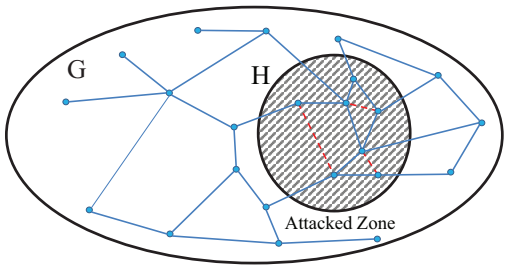


Figure 1: G is the power grid graph and H is a subgraph of G that represents the attacked area. An adversary attacks an area by disconnecting some of its lines (red dashed lines) and disallowing the information from within the area to reach the control center.

2. MODEL AND DEFINITIONS

2.1 DC Power Flow Model

To develop the recovery methods, we use the DC power flow model, which is widely used as an approximation for the non-linear AC power flow model [3, 4]. We represent the power grid by a connected undirected graph $G = (V, E)$ where $V = \{1, 2, \ldots, n\}$ and $E = \{e_1, \ldots, e_m\}$ are the set of nodes and edges corresponding to the buses and transmission lines, respectively. Each edge e_i is a set of two nodes $e_i = \{u, v\}$. p_v is the active power *supply* ($p_v > 0$) or *demand* ($p_v < 0$) at node $v \in V$ (for a *neutral node* $p_v = 0$). We assume *pure reactive* lines, implying that each edge $\{u, v\} \in E$ is characterized by its *reactance* $r_{uv} = r_{vu}$.

Given the power supply/demand vector $\vec{p} \in \mathbb{R}^{|V|\times 1}$ and the reactance values, a *power flow* is a solution $\mathbf{P} \in \mathbb{R}^{|V|\times|V|}$ and $\vec{\theta} \in \mathbb{R}^{|V|\times 1}$ of:

$$\sum_{v\in N(u)} p_{uv} = p_u, \ \forall\ u \in V \tag{1}$$

$$\theta_u - \theta_v - r_{uv}p_{uv} = 0, \ \forall\ \{u, v\} \in E \tag{2}$$

where $N(u)$ is the set of neighbors of node u, p_{uv} is the power flow from node u to node v, and θ_u is the phase angle of node u. When the total supply equals the total demand in each connected component of G, (1)-(2) has a unique solution.[1] Eq.(1)-(2) are equivalent to the following matrix equation:

$$\mathbf{A}\vec{\theta} = \vec{p} \tag{3}$$

where $\mathbf{A} \in \mathbb{R}^{|V|\times|V|}$ is the *admittance matrix* of G,[2] defined as follows:

$$a_{uv} = \begin{cases} 0 & \text{if } u \neq v \text{ and } \{u, v\} \notin E, \\ -1/r_{uv} & \text{if } u \neq v \text{ and } \{u, v\} \in E, \\ -\sum_{w\in N(u)} a_{uw} & \text{if } u = v. \end{cases}$$

The (node-edge) *incidence matrix* of G is another useful matrix that we use in detecting line failures. It is denoted by $\mathbf{D} \in \{-1, 0, 1\}^{|V|\times|E|}$ and defined as follows (for an arbitrary orientation of the edges),

$$d_{ij} = \begin{cases} 0 & \text{if } e_j \text{ is not incident to node } i, \\ 1 & \text{if } e_j \text{ is coming out of node } i, \\ -1 & \text{if } e_j \text{ is going into node } i. \end{cases}$$

[1]The uniqueness is in the values of p_{uv}s rather than θ_us (shifting all θ_us by equal amounts does not violate (2)).

[2]The admittance matrix $\mathbf{A}$ can also be considered as the *weighted Laplacian matrix* of the graph.

2.2 Attack Model

We assume that an adversary attacks an area by: (i) disconnecting some lines within that area (*failed lines*), and (ii) obstructing the information (e.g., status of the lines and phase angle measurements) from within the area to reach the control center. We assume that disconnecting lines within the area does not make G disconnected and the supply/demand values do change after the attack. The methods provided here can be used for general attack scenarios including the change in the supply/demand values and separation of the grid into islands, if the control center is aware of the changes in the supply/demand values.

Fig. 1 shows an example of an attack on the area represented by $H = (V_H, E_H)$. We denote the set of failed lines in area H by $F \subseteq E_H$. Upon failure, the failed lines are removed from the graph and the flows redistribute according to the DC power flows (in Section 5, we assume that the power flows redistribute according the AC power flows). Our objective is to estimate the phase angles and detect the failed lines inside the attacked area using the changes in the phase angles of the nodes outside the area.

Notation. We denote the complement of the area H by $\bar{H} = G\backslash H$. $\mathbf{D}_H \in \{-1, 0, 1\}^{|V_H|\times|E_H|}$ is the submatrix of $\mathbf{D}$ with rows from V_H and columns from E_H. $\vec{\theta}_H$ and $\vec{\theta}_{\bar{H}}$ are the vectors of phase angles of the nodes in H and $\bar{H}$, respectively. If X, Y are two subgraphs of G, $\mathbf{A}_{X|Y}$ denotes the submatrix of $\mathbf{A}$ with rows from V_X and columns from V_Y. We use the prime symbol ($'$) to denote the values after an attack. For a column vector $\vec{y}$, $\|\vec{y}\|_1 := \sum_{i=1}^{n} |y_i|$ is its l_1-norm, $\|\vec{y}\|_2 := (\sum_{i=1}^{n} y_i^2)^{1/2}$ is its l_2-norm, and $\text{supp}(\vec{y}) := \{i | y_i \neq 0\}$ is the index set of its nonzero elements.

3. STATE ESTIMATION

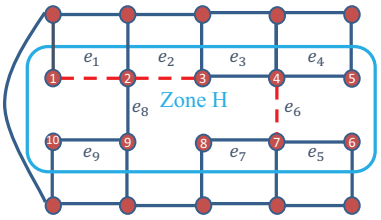
We can formulate the state estimation problem after a cyber-physical attack as follows: Given the attacked area H, $\vec{\theta}$, and $\vec{\theta}'_{\bar{H}}$, the objective is to estimate $\vec{\theta}'_H$ and detect F. We proved in [7] that if the attacked area H is known, then there always exists feasible vectors $\vec{x} \in \mathbb{R}^{|E_H|}$ and $\vec{y} \in \mathbb{R}^{|V_H|}$ satisfying the conditions of the following optimization problem such that $\text{supp}(\vec{x}) = F$ and $\vec{y} = \vec{\theta}'_H$:

$$\begin{aligned} &\min_{\vec{x},\vec{y}} \|\vec{x}\|_1 \text{ s.t.} \\ &\mathbf{A}_{H|H}(\vec{\theta}_H - \vec{y}) + \mathbf{A}_{H|\bar{H}}(\vec{\theta}_{\bar{H}} - \vec{\theta}'_{\bar{H}}) = \mathbf{D}_H\vec{x} \\ &\mathbf{A}_{\bar{H}|H}(\vec{\theta}_H - \vec{y}) + \mathbf{A}_{\bar{H}|\bar{H}}(\vec{\theta}_{\bar{H}} - \vec{\theta}'_{\bar{H}}) = 0. \end{aligned} \tag{4}$$

The novelty of this method is that it provides a convex relaxation for the line failures detection problem which is combinatorial in nature. Notice that the optimization problem (4) can be solved efficiently using Linear Program (LP). We proved in [7] that under various conditions on H and the set of line failures F, the solution to (4) is unique. Therefore the relaxation is exact and the set of line failures can be detected by solving (4). In particular, it is proved in [7] that if H is acyclic and there is a matching between the nodes in H and $\bar{H}$ that covers H, the solution to (4) is unique for any set of line failures.[3]

In the following sections, we numerically show that by relaxing the equalities in (4), a similar approach can be used

[3]We proved in [7] that the solution to (4) is unique under less restricted conditions. However, here we only focus on the simplest case.



SNR = 50dB

Actual $\vec{\theta}'_H$	Recovered $\vec{\theta}'_H$
0.1062	0.1030
0.0882	0.0872
0.0042	0.0075
0.0062	0.0036
0.0342	0.0334
-0.0758	-0.0752
-0.0971	-0.0937
-0.0925	-0.0919
0.0442	0.0441
0.0322	0.0321

$\vec{x}$	-0.0074	-0.0068	0	0	0	-0.0857	0	0	0
supp($\vec{x}$)	1	1	0	0	0	1	0	0	0

Figure 2: An example of an attack and recovered information in the presence of the measurement noise for SNR= 50dB. Red dashed lines show the attacked lines.

to detect line failures in the attacked area with the same properties, when the phase angle measurements are noisy and also when the phase angles are obtained based on the AC power flow model.

4. STATE ESTIMATION IN THE PRESENCE OF MEASUREMENT NOISE

In this section, we discuss the problem of information recovery after an attack in the presence of a measurement noise and uncertainty [8]. We follow [6] and model the measurement noise by changing (3) to $\mathbf{A}(\vec{\theta}-\vec{e}) = \vec{p}$ where $\vec{e} \in \mathbb{R}^{|V|\times 1}$ is a Gaussian measurement noise with a diagonal covariance matrix Σ. Following [12], $\vec{e}$ can also account for the perturbations in $\vec{p}$ after failures. It is obvious that in the presence of noise, the optimization problem (4) has no feasible solution. However, since the l_1-norm is relatively robust against noise, one possible approach to generalize the optimization problem (4) to the noisy case is to relax the conditions as follows:

$$\min_{\vec{x},\vec{y}} \|\vec{x}\|_1 \text{ s.t.}$$
$$\|\mathbf{D}_H\vec{x} - \mathbf{A}_{H|H}(\vec{\theta}_H - \vec{y}) - \mathbf{A}_{H|\bar{H}}(\vec{\theta}_{\bar{H}} - \vec{\theta}'_{\bar{H}})\|_2 < \epsilon_1 \quad (5)$$
$$\|\mathbf{A}_{\bar{H}|H}(\vec{\theta}_H - \vec{y}) + \mathbf{A}_{\bar{H}|\bar{H}}(\vec{\theta}_{\bar{H}} - \vec{\theta}'_{\bar{H}})\|_2 < \epsilon_2.$$

It is easy to see that the optimization problem (5) is a *second-order cone program* that can be solved using gradient decent methods. The values of ϵ_1 and ϵ_2 can be estimated based on the noise level. After solving (5), the line failures can then be detected as before by $F = \{e_i | i \in \text{supp}(\vec{x})\}$.

We show by simulations that solving the optimization problem (5) can correctly recover the phase angles and detect the failed lines in the presence of the measurement noise depending on the SNR level.We consider the graph G and the attacked area H as shown in Fig. 2 (it is easy to see that H is acyclic and there is a matching between the nodes in H and $\bar{H}$ that covers nodes in H). Notice that the graph in Fig. 2 can be part of a much bigger graph. However, only the local information is needed to recover the information inside the attacked area.

Fig. 2 shows an attack scenario with high SNR value and the information recovered by solving (5). As can be seen, the attacked lines can be detected accurately in this case by computing supp($\vec{x}$).

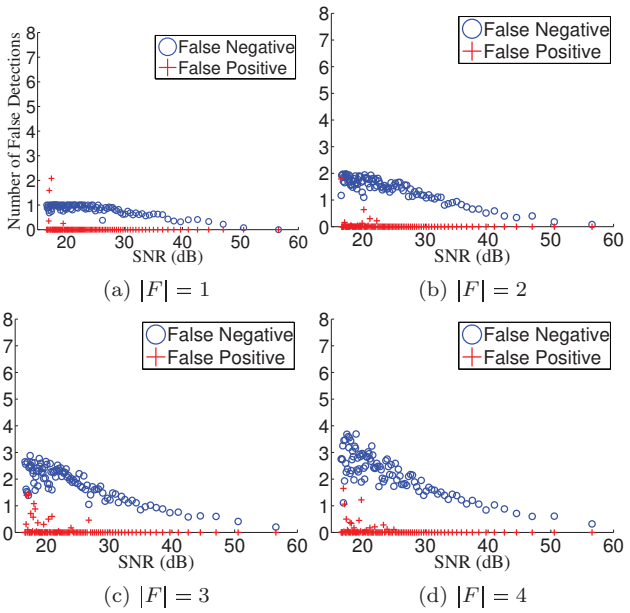


Figure 3: The average number of false negatives and positives in detecting line failures by solving (5) in the presence of the measurement noise versus the SNR. Each data point is the average over 100 trials. (a)-(d) Show this relationship for different number of line failures ($|F|$). Fig. 2 provides the detailed information for a point in (c).

Fig. 3 shows the average number of false negatives and positives in detecting line failures by solving (5) versus the SNR level for different numbers of line failures. As can be seen, for any number of line failures, when the SNR is above a certain level (e.g., 40 dB) the solution to (5) can detect the line failures with acceptable accuracy (less than one false negative and zero false positives on average).

5. STATE ESTIMATION UNDER THE AC POWER FLOW MODEL

The optimization (5) can also be used to detect line failures when the phase angles are given under the AC power flow model [9]. We consider the standard AC power flow equations (for the details of the equations see [3]). The only difference here compared to the noisy scenario is that ϵ_1, ϵ_2 cannot be estimated easily. To overcome this issue, the idea is to change ϵ_i from s_i to t_i and compute the solution to (5) for each setup. If $\mathcal{F}$ is an array that contains all the detected line failures for each setup, then the appearance frequency of each line in $\mathcal{F}$ gives a rough probability that the line is failed. P_F denotes the appearance frequency table of the lines in $\mathcal{F}$. Moreover, the computed vector $\vec{y}$ in each iteration is an estimate of the phase angles inside the attacked area. By computing the mean of all the estimated phase angle vectors in each iteration, one can improve this estimation.

We use the IEEE 118-bus benchmark system as the test network [1] and consider the attacked zone H within this network as shown in Fig. 4. It is easy to see that H is acyclic. Although the rest of nodes in the IEEE 118-bus system are not displayed, there is also a matching between the nodes in H and $\bar{H}$ that covers H.

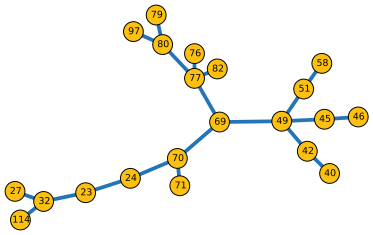


Figure 4: Topology of the attacked area H within the 118-bus system with 21 nodes and 22 lines used in the simulations in Section 5.

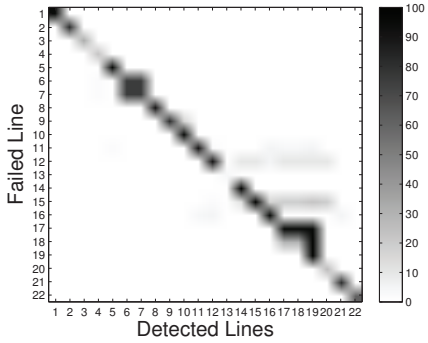


Figure 5: Detected line failures after all single line failures in area H within the IEEE 118-bus system. The color intensity of each (i, j) square shows the number of times line j is detected as failed when only line i is actually failed.

Fig. 5 shows the heat map of the detected line failures after all possible single line failures in H. As can be seen, in most of the cases, the correct line is detected as the most probable failed line. If we use the appearance frequency table P_F and a threshold value t to detect the most likely failed lines, for $t = 0.5$, for almost 80% of the cases there are no false negatives or false positives. For $t = 0.2$, for almost 95% of the cases, there are no false negatives while for almost 80% of the cases there are no false positives either. Moreover, for all single line failure scenarios, the error in the estimated phase angles is below 1%.

We also consider all double line failure scenarios in H. As in the single line failures, the phase angle estimation is very accurate. For all double line failure scenarios, the error in the estimated phase angles is below 2%. Since there are many double line failure cases, we cannot show the failed lines detection results as a matrix heatmap. However, we can show the number of false negatives and positives if we use the appearance frequency table P_F and a threshold value t to detect the most likely failed lines. As can be seen in Fig. 6, for $t = 0.2$ for more than 80% of the cases there is no false negative. Moreover, for more than 80% of the cases there is less than a single false positive line detection.

6. CONCLUSION

We provided convex optimization based methods to estimate the state of the grid following a cyber-physical attack under both the DC and AC power flow models. We demonstrated that these methods can exactly recover the phase angles and detect the line failures under the DC power flow model. However, when the measurements are noisy or the phase angles are given under the AC power flow model, we numerically showed that these methods can still recover the information with low error.

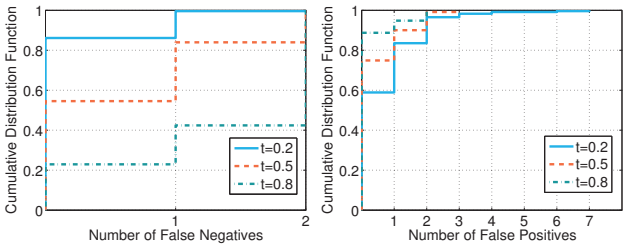


Figure 6: The CDF of the number of false negatives and positives in detecting double line failures in H within the 118-bus system using the threshold value t.

We believe that the presented methods can accurately estimate the state of the grid for less constrained attacked areas as well. Moreover, these method can also be used in different context such as false data detection. Exploring these directions is part of our future work.

7. REFERENCES

[1] IEEE benchmark systems. Available at `http://www.ee.washington.edu/research/pstca/`.

[2] Analysis of the cyber attack on the Ukrainian power grid, Mar. 2016. `http://www.nerc.com/pa/CI/ESISAC/Documents/E-ISAC_SANS_Ukraine_DUC_18Mar2016.pdf`.

[3] A. R. Bergen and V. Vittal. *Power Systems Analysis*. Prentice-Hall, 1999.

[4] D. Bienstock. *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*. SIAM, 2016.

[5] M. Garcia, T. Catanach, S. Vander Wiel, R. Bent, and E. Lawrence. Line outage localization using phasor measurement data in transient state. *IEEE Trans. Power Syst.*, 31(4):3019–3027, 2016.

[6] J. Kim and L. Tong. On topology attack of a smart grid: undetectable attacks and countermeasures. *IEEE J. Sel. Areas Commun*, 31(7):1294–1305, 2013.

[7] S. Soltan, M. Yannakakis, and G. Zussman. Joint cyber and physical attacks on power grids: Graph theoretical approaches for information recovery. In *Proc. ACM SIGMETRICS'15*, June 2015.

[8] S. Soltan, M. Yannakakis, and G. Zussman. Power grid state estimation following a joint cyber and physical attack. *To appear in IEEE Trans. Control Netw. Syst.*, 2017.

[9] S. Soltan and G. Zussman. Power grid state estimation after a cyber-physical attack under the AC power flow model. In *Proc. IEEE PES-GM'17*, 2017.

[10] J. E. Tate and T. J. Overbye. Line outage detection using phasor angle measurements. *IEEE Trans. Power Syst.*, 23(4):1644–1652, 2008.

[11] J. E. Tate and T. J. Overbye. Double line outage detection using phasor angle measurements. In *Proc. IEEE PES-GM'09*, July 2009.

[12] H. Zhu and G. B. Giannakis. Sparse overcomplete representations for efficient identification of power line outages. *IEEE Trans. Power Syst.*, 27(4):2215–2224, 2012.