One Breaker is Enough: Hidden Topology Attacks on Power Grids

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

Abstract—A coordinated cyber-attack on grid meter readings and breaker statuses can lead to incorrect state estimation that can subsequently destabilize the grid. This paper studies cyber-attacks by an adversary that changes breaker statuses on transmission lines to affect the estimation of the grid topology. The adversary, however, is incapable of changing the value of any meter data and can only block recorded measurements on certain lines from being transmitted to the control center. The proposed framework, with limited resource requirements as compared to standard data attacks, thus extends the scope of cyber-attacks to grids secure from meter corruption. We discuss necessary and sufficient conditions for feasible attacks using a novel graph-coloring based analysis and show that an optimal attack requires breaker status change at only ONE transmission line. The potency of our attack regime is demonstrated through simulations on IEEE test cases.

I. INTRODUCTION

Real time operation of the power grid and computation of electricity prices [1] require accurate estimation of its structure and critical state variables. Remote Terminal Units (RTUs) transmit measurements collected from different grid components to the central control center for state estimation and subsequent use in analyzing grid stability. The collected measurements can be broadly classified into two kinds: meter readings and breaker statuses. The breaker statuses on transmission lines help create the current operational topology of the grid. The meter readings, comprised of line flow and bus power injection measurements, are then used to estimate the state variables over the estimated topology. In a practical setting, the collected measurements suffer from noise, that get added at source or during communication to the control center. The affect of such noise is minimized through placement of redundant/additional meters and use of suitable bad-data detection and correction techniques at the estimator [2].

Cyber-attacks on the power grid refer to corruption of measurements (meter readings and breaker statuses) by an adversary, aimed at changing the state estimation output, without getting detected by the estimator’s checks. The viability of such attacks has in fact been demonstrated through controlled experiments like the Aurora attack in Department of Energy’s Idaho Laboratory and GPS spoofing attack on phasor measurement units (PMUs) [3]. Past literature on cyber-attacks have generally looked at adversaries that change meter data (and not breaker statuses) to affect state estimation. Such data attacks involving injection of malicious data into meters were first analyzed in [4]. Using a DC power flow model for state estimation, the authors of [4] provide an attack design using projection matrices. Following this, several approaches have been discussed to study hidden attacks under different operating conditions. These include mixed integer programming [5], heuristic based detection [6], sparse recovery using $l_0-l_1$ relaxation [7], graph-cut based construction for systems with phasor measurement units (PMUs) [8] among others. The possible economic ill-affects of such hidden data attacks on power markets are presented in [9]. In a recent paper [10], the authors investigates hidden attacks under the more general and potent regime of topology data (breaker statuses) and meter data corruption. All of these cited work on ‘data alone’ or ‘topology and data’ attacks, however, require changing floating point meter measurements in real time. The practicality of this is questionable as significant resources are required to synchronize the changes at multiple meters.

In this paper, we focus on hidden attacks that primarily operate through changes in breaker statuses. Here the adversary changes the statuses of a few operational breakers from 1 (closed) to 0 (open), as well as jams (blocks the communication) of flow measurements on a subset of transmission lines in the grid. However, the adversary does not modify any meter reading to an arbitrary value. We term these attacks as ‘breaker-jammer’ attacks. Note that breaker statuses, unlike meter readings, are binary in nature and fluctuate with lower frequency. They are thus easier to change, even by adversaries with limited resources. Jamming measurements, through jammers or by destruction of communication apparatus, is technologically less intensive than corrupting meter measurements. In fact, jamming does not raise a major alarm as measurement loss due random communication drops occurs under normal circumstances. The ‘breaker-jammer’ attack model was introduced by the authors for grids with a specific meter configuration requiring sufficient line flow measurements in [11]. This work generalizes the framework to any grid with line flow and injection meters and uses a novel graph-coloring analysis to determine the optimal hidden attack. Our graph coloring based analysis is in principal similar to [12] which studies standard data attacks as a graph partitioning problem. However, the similarly ends there as our attack model does not use corruption of meter readings. Instead breaker status changes and line flow jams provide a different set of necessary and sufficient conditions for feasible attacks. The surprising revelation of our analysis is that under normal operating conditions, a single breaker status change (with the
necessary flow measurement jamming) is sufficient to create an undetectable attack. In fact, we show that if a hidden attack can be constructed by changing the status of a set of breakers, then a hidden attack using only one break status change exists as well. This is significant as the adversary can focus on jamming the necessary flow measurements, after selecting a breaker to attack. Further, our attack design does not depend on the current system state or transmission line parameter values, and has low information requirements.

The rest of this paper is organized as follows. We present the system model used in generalized state estimation and describe the attack model in the next section. The graph coloring approach to determine the necessary and sufficient conditions for a hidden attack and elucidating examples are discussed in Section [III]. The design of the optimal hidden attack is discussed in Section [IV] along with simulations on IEEE test cases. Finally, concluding remarks and future directions of work are presented in Section [V].

II. GENERALIZED STATE ESTIMATION IN THE POWER GRID AND TOPOLOGY-BASED ATTACKS

First, we provide a brief description of the notation used. We represent the current operational structural grid of the system by graph $G = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ denotes the set of buses/nodes of size $n_V$ and $\mathcal{E}$ denotes the set of operational edges of size $n_E$. The set of binary breakers statuses for the edges is denoted by the diagonal matrix $D$ of size $n_E \times n_E$. We assume that all lines to be initially operational ($D$ is identity matrix) and ignore any non-operation line for ease of notation. The edge to node incidence matrix is denoted by $M$ of dimension $n_E \times n_V$. Each operational edge $(ab)$ between nodes $a$ and $b$ has a corresponding row $M_{ab}$ in $M$, where $M_{ab} = e_a - e_b$. $e_a$ denotes the standard basis vector in $\mathbb{R}^{n_V}$ with one at the $a$th location. The direction of flow on edge $(ab)$ is taken to be from $a$ to $b$, without any loss of generality. We consider the DC power flow model for state estimation in this paper [13].

The state variables in this model are the bus phase angles, denoted by the $n_V \times 1$ vector $x$. The set of measurements is denoted by the vector $z = \{z_f, z_{inj}\}$. Here line flow measurements are included in $z_f$ and bus injection measurements are included in $z_{inj}$.

State estimation in the power grid relies on the breaker statuses in $D$ for topology estimation and then uses the meter measurements $z$ for estimating the state vector $x$. The relation between $x$ and $z$ in the DC model is given by $z = Hx + e$ where $e$ is the zero mean Gaussian noise vector with covariance matrix $\Sigma_e$. $H$ is the measurement matrix and depends on the grid structure and susceptance of transmission lines. Let the $k_f$th entry in $z$ corresponds to the flow measurement on line $(ab)$. Then $H(k_1, :) = 0$ (the $k_f$th row in $H$) is given by

$$H(k_1, :) = [0 \ldots B_{ab} \ldots - B_{ba} \ldots] = B_{ab}M_{ab}$$

with the non-zero values at the $a$th and $b$th locations respectively. $B_{ab}$ is the susceptance of the line $(ab)$. On the other hand, if the $k_{inj}$th entry corresponds to an injection measurement at node $a$, we have $z(k_2) = \sum_{r(a) \in \mathcal{E}} B_{ar}(x_a - x_r)$. In matrix form, ignoring measurement noise, we can write equations for received measurements as

$$z_f = TBMx$$

(2)

$$z_{inj} = M_{inj}BMx$$

(3)

$B$ is the diagonal matrix of susceptances of lines in $\mathcal{E}$. We arrange the rows in $M$ such that the top $|z_f|$ rows represent the lines with flow measurements. Matrix $T$, comprising of the top $|z_f|$ rows of a $n_E \times n_E$ identity matrix, selects these measured flows. For ease of notation and analysis in later sections, we pad trailing zeros to vector $z_f$ and make it of length $n_E$. Similarly, we pad trailing all-zero rows to $T$ to make it a diagonal square matrix of dimension $n_E$.

$M_{inj}$ on the other hand consists of the columns of $M$ that correspond to the nodes with injection measurements. The optimal state vector estimate $\hat{x}$ is given by minimizing the residual $\sum_{r} \|z - H\hat{x}\|_2$. If the minimum residual does not satisfy a tolerance threshold, bad-data detection flags turn on and data correction is done by the estimator. The overall scheme of topology and state estimation processes followed by bad-data detection and correction is called generalized state estimation (GSE) [13] as illustrated in Figure 1.

**Attack Model:** We assume that the adversary is agnostic and has no information on the current system state $x$ or line susceptance matrix $B$. For attack, the adversary changes the breaker statuses on some lines. The new breaker status matrix, after attack, is denoted by $D - D_a$ where diagonal matrix $D_a$ has a value of 1 for attacked breakers. Similarly, the available flow measurements after jamming are represented by $T - T_a$, with diagonal matrix $T_a$ having a value of 1 corresponding to jammed flows. Let the new state vector estimated after the breaker-jammer attack be denoted by $x + c$, where $c \neq 0$ denotes the change. Note that if the flow measurement on a line is not jammed, its value remains the same following the attack. Using (2), we have

$$(T - T_a)BMx = (T - T_a)BM(x + c)$$

$$\Rightarrow (T - T_a)BMc = 0$$

(4)

It follows immediately that if the breaker status on the $r$th line with flow measurement is changed ($D_a(r, r) = 1$), to avoid detection, its flow measurement needs to be jammed as well ($T_a(r, r) = T(r, r) = 1$). Thus,

$$D_a(T - T_a) = 0$$

(5)

Consider the injection measurements ($z_{inj}$) now, which are not changed during the attack. The breaker attack leads to removal
of lines marked as open from Equation (3), resulting in the following modification.

\[ z_{inj} = M'_{inj}BMx = M'_{inj}(D - D_a)BM(x + c) \]
\[ \Rightarrow M'_{inj}D_aBMx = M'_{inj}(D - D_a)BMc \]
\[ (\text{Using (4))} = M'_{inj}(I - D_a)(I - T + T_a)BMc \]
\[ (\text{Using (5))} = M'_{inj}(I - (D_a + T - T_a))BMc \] (6)

Equation (6) thus states that after the ‘breaker-jammer’ attack, for each injection measurement, the sum of original flows contributed by lines with attacked breakers (left side) needs to be accommodated by changes in estimated flows on lines (connected to the same bus) whose breakers are intact but actual flow measurements are not received (right side). Finally, for unique state estimation following the adversarial attack (with one bus considered reference bus with phase angle \( \theta \)) we need

\[ \text{rank} \left[ \begin{pmatrix} (T - T_a) \\ M'_{inj}(D - D_a) \end{pmatrix} BM \right] = n_y - 1 \quad (7) \]

The necessary conditions for a successful ‘breaker-jammer’ attack that results in a change in estimated state vector consists of equations (4), (5), (6), and (7). In the next section, we describe a graph coloring based analysis of the necessary and sufficient conditions and use it to discuss design of optimal attacks of our regime.

III. ATTACK ANALYSIS BASED ON GRAPH COLORING

For our graph coloring based analysis, we use the following coloring scheme: for any change \( c \) in the estimated state vector, neighboring buses with same value in \( c \) are given same color. Using this, we now discuss a permissible graph coloring corresponding to the requirements of a feasible attack discussed in the previous section. Equation (4) states that if the flow on line \((ab)\) between buses \( a \) and \( b \) is not jammed, \( c(a) = c(b) \) (same color in our scheme). Thus, a set of buses connected through lines with available flow measurements (not jammed) has the same color.

This implies that the grid buses, following a feasible attack, can be divided into groups, each group having a distinct color. The lines between buses of different groups do not carry any flow measurement or are jammed by the adversary. A test example is illustrated in Figure 2. Observe the buses with injection measurements, that are not corrupted by the adversary. For an interior bus \( d \), (all neighboring nodes have the same color as itself), the right side of Equation (4) equates to zero. The left side becomes equal to zero, under normal operating conditions, if breakers on lines connected to bus \( d \) are not attacked. Thus, we have a feasible graph coloring has lines with attacked breakers connected to boundary buses. A boundary bus is one that has neighboring buses of colors distinct from itself.

Now, consider the injection meter installed on any boundary bus. Such buses can exist in two configurations: a) connected to lines with attacked breaker (see bus \( a \) in Figure 2) or b) connected to only lines with correct breaker statuses (node \( b \) if line \((bf)\) did not have a breaker attack). In either case, using (5), we have: each injection measurement placed at a boundary bus provides one constraint relating the values of \( c \) for neighboring differently colored buses.

For further analysis, we now use the coloring constraints highlighted in bold above to construct a reduced grid graph \( \mathcal{G} \) from \( \mathcal{G} \) as follows:

1. In each colored group, club boundary buses without injection measurements with all interior buses into one ‘supernode’ of that color. Make boundary buses with injection measurements into supernodes with the same color. Connect supernodes of same color with artificial lines of zero susceptance.
2. For each line with intact breaker between two buses of different colors, create a line of same impedance between their corresponding ‘supernodes’. Remove supernodes connected only to other supernodes of same color.
3. Make injection measurements on supernodes equal to the sum of original flows on lines with attacked breakers connected to them (positive for inflow, negative for outflow). If no incident line has attacked breaker, make the injection equal to 0.

Figure 3 illustrates the reduced graph construction for the example in Figure 2. Note that in the reduced graph \( \mathcal{G} \), original lines between buses of same color are removed. The included lines exist between buses of different colors and have jammed or unavailable flow measurements. Similarly, injection measurement relation (6) at interior nodes are trivially satisfied.
by c and are ignored. The reduced system, thus, only includes constraints from equation (6) as shown below:

$$\sum_{b,(ab) \in \hat{E}} \hat{B}_{ab}(\hat{c}_a - \hat{c}_b) = \hat{z}_{inj}(a) \quad (8)$$

Here, a and b are supernodes of different colors. The numeric value for the color of supernode a is given by \(\hat{c}_a\) (not the \(a^{th}\) entry in \(\hat{c}\)). \(\hat{B}\) and \(\hat{E}\) are the susceptance matrix and edge set corresponding to the reduced graph \(\hat{G}\). \(\hat{z}_{inj}(a)\) denotes the injection measurement on supernode a with value given by Step 3 in the reduced graph construction. Note that equation (6) for the injection measurements involves rows of the susceptance weighted Laplacian matrix for \(\hat{G}\). A unique solution of \(\hat{c}\) for \(\hat{G}\) in turn provides a uniquely estimated \(c\) in \(G\) after the adversarial attack. We now look at condition (7), necessary for unique state estimation after a feasible adversarial attack in terms of graph coloring. The reduced graph \(\hat{G}\) greatly simplifies our analysis here.

First, it is clear that each color must have at least one supernode or a neighboring supernode (of different color) with injection measurement. Otherwise, the value of \(\hat{c}\) for that color will not be in any injection constraint. This goes against uniqueness of state estimation. Note that the number of degrees of freedom in \(\hat{c}\) (representing distinct values in \(c\)) is one less than the number of colors as one color denotes the reference phase change of 0. Using \(\hat{G}\), we prove the following result regarding permissible graph coloring for unique estimation.

**Theorem 1.** Following a ‘breaker-jammer’ attack, the number of injection measurements at the boundary buses should be one less than the number of distinct colors in the grid buses.

**Proof:** Let the number of colored groups be \(k\). Then the number of independent entries in \(\hat{c}\) is \(k - 1\) (one entry being 0). The total number of linear constraints involving the numeric values in \(\hat{c}\) is equal to the number of injection measurements at the supernodes in \(\hat{G}\). For unique state estimation, the number of injection measurements should thus be greater than or equal to \(k - 1\). We now show that exactly \(k - 1\) injection measurements are needed to get a solution to state estimation. Consider the reduced graph \(\hat{G}\). For real valued line susceptances and for cases where the supernodes having injection measurements do not form a closed ring with no additional branches (see Figure 3), the rank of \(k - 1\) rows is \(k - 1\) and we have unique state estimation. If the reduced graph \(\hat{G}\) contains a closed ring of supernodes with injection measurements, then the measurements will represent the entire susceptance weighted graph Laplacian of the ring, that is rank deficient. However, the real valued entries in \((\hat{z})_{inj}\) that exist on the right side of (8) and are derived from flows on lines with attacked breakers, will not cancel out under normal operating conditions. Further, the adversary designing the attack is unaware of the current system state and will be able to determine if they do. Hence the \(k - 1\) injections measurements constraints will be linearly independent (the adversary will expect this under normal operations). This gives an unique \(\hat{c}\) and \(c\) for a \(k\) distinct colored grid graph.

To summarize, the highlighted statements and Theorem 1 provide the necessary and sufficient conditions for a feasible ‘breaker-jammer attack’ under our graph-coloring scheme. In the next section, we show that the graph coloring approach proves a surprising result that simplifies the design of an optimal attack.

**IV. Design of Optimal Attack**

We call an ‘breaker-jammer’ feasible attack optimal if it requires minimum number of breaker status changes (considering the fact that doing so is significantly more resource draining than measurement jamming). If multiple attacks are possible using the minimum number of breaker changes, we select as optimal the attack that requires the least number of flow measurement jams. Using the reduced graph \(\hat{G}\), we present the following result for the minimum number of breaker changes needed for a feasible attack under normal operating conditions (non-zero real-valued bus susceptances and line flows that are distinct for different grid elements).

**Theorem 2.** If a feasible attack can be designed with \(k\) breaker status changes, then a feasible attack exists such that all but one breaker statuses are changed back to their original operational state (1), while keeping their line flow measurements jammed.

**Proof:** Construct the reduced graph \(\hat{G}\) with its colored supernodes for the feasible attack with \(k\) breakers and necessary flow measurement jams. Let the number of colors in state estimation change \(c\) be \(r + 1\). The length of \(\hat{c}\) is then \(r + 1\). By Theorem 1, there are \(r\) injection measurements at the supernodes that provide constraint equations listed in (8). If we revert the breaker status of an attacked line back to 1 while keeping its flow measurement jammed, the only change in any constraint equation (8) involving that line will be that the injection measurement on the incident node (entry in \(\hat{z}\)) will become 0. Since all but one breakers are brought back to the operational state, at least one injection measurement in \(\hat{z}\) will still remain non-zero and the \(r\) constraint equations will still have linear independence. Thus, state estimation will result in a different but non-zero \(\hat{c}\), leading to a feasible attack.

For example, consider the case in Figure 2 where two breaker statuses are attacked. If the breaker status on line (bf) is changed back to 1 while keeping the flow measurement jammed, the new reduced graph that will be derived is given in Figure 4. As mentioned in Theorem 2, the coloring scheme is still feasible and a non-zero change in state estimation results.

This is a very significant result and simplifies the search for an optimal attack greatly. Since one breaker change is sufficient, the adversary can select each line in turn (\(n_e\) iterations), attack its breaker (change the corresponding entry in diagonal \(D_a\) to 1) and determine the flow measurements that need to be jammed (given by diagonal \(T_a\)) to conduct a feasible attack. The breaker change that requires the minimum number of measurement jams (or maximally sparse \(T_a\)) will then give the optimal attack. The selection of jammed measurements,
after fixing $D_\alpha$, is formulated as:

$$\begin{align*}
\text{minimize} & \quad ||T_a||_0 \\
\text{subject to} & \quad T_a \text{ is diagonal, } T_a \text{ satisfies (4), (5), (6)}
\end{align*}$$

This is simplified in formulation (10) where the jammed measurements (with 1 on diagonal of $T_a$) are given by the non-zero entries in $TM_c$. $l_0 - l_1$ relaxation can be used to approximately solve (10). Since the adversary has no access to the actual state vector $x$, a random non-zero $x$ is used. Similarly, unavailable line susceptance $B$ are replaced with distinct real values. These replacements, under normal conditions, do not affect the optimal solution as they preserve the linear independence of injection constraints given in (8).

$$\begin{align*}
\text{minimize} & \quad ||TM_c||_0 \\
\text{subject to} & \quad M'_{inj}D_\alpha BMx = M'_{inj}BM_c
\end{align*}$$

The rank constraint (7) is not included in the optimization framework and can be checked manually after determining $T_a$, for consistency.

**Experiments:** We simulate our attack model on IEEE 14, 30 and 57 bus test systems [14] and present averaged findings in Figure 5. For each test system considered, we place flow measurements on all lines and injection measurements on a fraction of buses, selected randomly. To design a feasible attack involving a line, we change its breaker status and solve Problem (10) to jam flows measurements to prevent detection. This is repeated for each line to determine the optimal attack. In Figure 5 note that the average number of flow measurements jammed increases with the number of injection measurements. This happens due to an increase in the number of injection constraints that require more measurement jams.

V. CONCLUSION

In this paper, we study topology based cyber-attacks on power grids where an adversary changes the breaker statuses of operational lines and marks them as open. The adversary also jams flow measurements on certain lines to prevent detection at the state estimator. The attack framework is novel as it does not involve any injection of corrupted data into meters or knowledge of system parameters and current system state. Using lesser information and resource overhead than traditional data attacks, our attack regime explores attacks on systems where all meter data are protected from external manipulation. We discuss necessary and sufficient conditions for the existence of feasible attacks through a new graph-coloring approach. The most important result arising from our analysis is that optimal topology based attacks exist that require a single breaker status change. Finally, we discuss an optimization framework to select flow measurements that are jammed to prevent detection of the optimal attack. Its efficacy is presented through simulations on IEEE test cases. Designing protection schemes for our attack model is the focus of our current work.

**REFERENCES**

[1] A. L. Ott, “Experience with PJM market operation, system design, and implementation”, *IEEE Trans. Power Syst.*, vol. 18, no. 2, 2003.
[2] A. Abur and A. G. Expósito, “Power System State Estimation: Theory and Implementation”, New York: Marcel Dekker, 2004.
[3] D. Shepard, T. Humphreys, and A. Fansler, “Evaluation of the vulnerability of Phasor Measurement Units to GPS spoofing”, *International Conference on Critical Infrastructure Protection*, 2012.
[4] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids”, *Proc. ACM Conf. Comput. Commun. Security*, 2009.
[5] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, “Network-aware mitigation of data integrity attack on power system state estimation”, *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, 2012.
[6] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, “Limiting false data attacks on power system state estimation”, *Proc. Conf. Inf. Sci. Syst.*, 2010.
[7] T. Kim and V. Poor, “Strategic Protection Against Data Injection Attacks on Power Grids”, *IEEE Trans. Smart Grid*, vol. 2, no. 2, 2011.
[8] D. Dekar, R. Baldick, and S. Vishwanath, “Data Attack on Strategic Buses in the Power Grid: Design and Protection”, *IEEE PES General Meeting*, 2014.
[9] L. Xie, Y. Mo, and B. Sinopoli, “False data injection attacks in electricity markets”, *Proc. IEEE SmartGridComm*, 2010.
[10] J. Kim and L. Tong, “On topology attack of a smart grid: Undetectable attacks and countermeasures”, *IEEE J. Select. Areas Commun.*, vol. 31, no. 7, 2013.
[11] D. Dekar, R. Baldick, and S. Vishwanath, “Attacking Power Grids with Secure Meters: The Case for using Breakers and Jammers”, *IEEE Infocom CCSES Workshop*, 2014.
[12] A. Gianii, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla, “Smart grid data integrity attacks”, *IEEE Trans. on Smart Grid*, vol. 4, no. 3, 2013.
[13] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*, CRC, 2000.
[14] “Power system test archive”, [http://www.ee.washington.edu/research/pstca](http://www.ee.washington.edu/research/pstca)