# Vulnerability Analysis of a Smart Grid with Monitoring and Control System

Maggie Cheng    Mariesa Crow
Missouri University of Science and Technology
Rolla, MO 65401
{chengm,crow}@mst.edu

Robert F. Erbacher
U.S. Army Research Laboratory
Adelphi, MD 20783
Robert.F.Erbacher.civ@mail.mil

## ABSTRACT

Large scale power outage is typically the consequence of cascading failures propagated through a power system. To mitigate failure propagation, wide area monitoring and control system is introduced. However, as the physical system is tightly coupled with the cyber system, new threats are introduced due to possible cyber attacks and failures of the communication system.

This paper presents a new framework for vulnerability analysis. Under this framework, we can identify the vulnerable components and the critical components of a cyber physical system. Distinct from previous work, our model considers the interaction between the different components of the cyber physical system, and models the dynamic evolving process of cascading failures. The impact of a component failure on the system is dynamically changing as the failure propagates. We formulate the vulnerability analysis problem as an optimization problem and present efficient algorithms to solve it. Since instability is the reason of power outage, we use an instability index to measure the negative impact to the system. The results from this optimization problem suggest which components of the system are critical since their failure can most negatively impact the cyber-physical system.

## 1. INTRODUCTION

Cyber-physical system security in smart grid has attracted a lot of research interests in recent years. To take preventive action against potential attacks, it is important to identify the vulnerability as early as possible so that grid operator can enhance the security and robustness of those identified components. In this paper, we present an analytical framework to identify the security holes of a power grid. It is analytical in the sense it identifies the most vulnerable and the most critical components of the system without deliberately probing the system to discover its weaknesses. Previous work on power grid vulnerability analysis is mainly on SCADA system ([2, 3]). Our model is built for the smart grid with wide area monitoring and control, different from SCADA system. Our approach is also different from existing approaches that uses attack graphs or attack trees ([4, 1] etc.) for network vulnerability analysis.

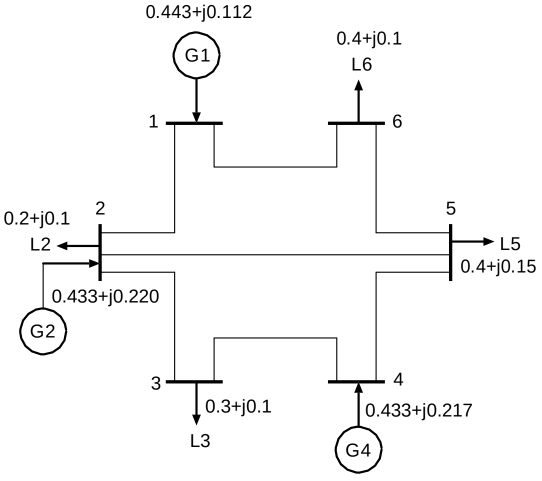

Figure 1: A six-bus three-machine system.

To show the cascading failure propagation in a power grid, we study a simple three-machine system in Figure 1. Suppose instantaneous high load at bus 2 causes power oscillation at bus 2. If transient instability cannot be dampened timely and it may cause power outage in this area. This "fault" may be propagated to bus 3 and the extra load may also cause system instability in that area. If the instability is high and it can continue to increase the power flow in more lines and faults may propagate through the whole power grid and cause large scale power outage. The more the power grid is connected, the more vulnerable it is. An isolated area can only have small scale power outage, but will have it more often; a highly connected power grid will have fewer but larger power outages.

To identify the component that is prone to failure is important since failure of one component may cause chain reaction of more components, and therefore the robustness or reliability of the vulnerable component should be improved. On the other hand, to identify the component whose failure can cause the largest degree of damage to the system is also very important in a security context since the hacker would target at such components to attack. The identification of the vulnerable components and the critical components help the grid operator in long term planning. The operator can reconfigure the generators, power lines or enhance security protection of some components to keep the grid in a reliable and secure state.

The traditional ”cut” approach that is based on the network topology or connectivity alone will not do a satisfactory job in power grid vulnerability analysis. The reason is that power flow dynamics is not considered, and the interaction between different components of the system is not considered. The cascading propagation of failure cannot be captured in such a model. Due to the heterogeneity in load distribution and source capacity, a more sophisticated method is needed than the simple ”cut” approach.

The key idea of this paper that distinguishes itself from previous work is to consider the evolving process of system instability after component failure. System vulnerability is quantified in terms of the cost to the power system, which is related to system instability. When system instability is small, power oscillation caused by load disturbance can be quickly stabilized and no further damage will occur; when system instability is large, power oscillation cannot be dampened timely, so it may cause the tripping of a power line. The tripping of a power line will cause power oscillation in other areas so the chain reaction will continue.

The scope of this paper does not include the countermeasures of potential attacks; it only identifies the vulnerable components and the critical components. It is up to the grid operator to decide what to do with the result of vulnerability analysis. Countermeasure or protection is the next step after vulnerability analysis.

## 2. GRAPH MODEL

We can use a multi-source multi-sink flow network to represent a power grid as follows: source nodes represent generators, sink nodes represent loads, and intermediate nodes represent buses. Directed edges are added from generators to buses and from buses to sinks; undirected edges are added between buses to represent the physical connectivity among them. A directed edge from a generator to a bus or from a bus to a load has no capacity limit, and therefore the edge has capacity set to $c(u, v) = \infty$. Such edges are not subject to failure. An undirected edge between two buses represents the power line with a capacity limit, and therefore we set the edge capacity $c(i, j) = T^*_{ij}$ in both directions. If the power flow $S_{i,j}$ exceeds $T^*_{ij}$, the line will trip off.

If the undirected subgraph induced by the bus nodes is fully connected (i.e., there is a path from every node to every other nodes), then every source node has a connected path to every sink node. When there are multiple sources available, a load may be satisfied by drawing power flows from multiple sources. Intuitively, every sink node would draw power from its nearest source node. If the nearest source cannot satisfy its demand, then the second nearest, and so on. This is because the nearest source has the lowest impedance on the power line so the energy loss is small along the transmission line; It is also because if there is increase in demand, the power supply can ramp up quickly if it is near the sink node so that the power oscillation can be quickly stabilized.

To compute the power flow on power lines when given load condition, we can formulate the problem as a flow network problem. We use the impedance of a wire as the weight of an edge. When computing the power flow, we ignore the capacity constraint since the real power flow does not change its path because of the capacity limit of the power line. The problem can be cast as an optimization problem with an objective of minimizing the total weight, subject to the constraints that flow conservation is satisfied for both active and reactive power, and that the total load demand is satisfied.

## 3. VULNERABILITY ANALYSIS

The objective of vulnerability analysis is to provide a quantitative measure of system instability, and to identify the vulnerable components that are prone to failure and the critical components whose failure can bring a high cost to the power system. A static analysis can be done in one sweep by using the given load and generation configuration, but if the grid operator is actively adjusting his strategy to defend the power grid and the attacker is actively adjusting his strategy to maximize his chance to bring down the system, the defender and the attacker are engaged in a game in which one party aims to maximize a cost function while the other party tries to minimize the same cost function.

### 3.1 Static Analysis

What is considered as a cost to the power system is disturbance to the normal operation. To identify the critical components and the vulnerable components both require quantitative measures of disturbance. We first quantify the degree of disturbance by using a cost function.

Every line $(i, j)$ in the power system has a security limit, $S^*_{ij}$, which is also called the soft limit. The security margin is the distance between the current operating point and the security limit. Ideally, we would like to have a large security margin. There is also a threshold value $T^*_{ij} > S^*_{ij}$, also known as the hard limit. It is desired that power system operates below the security limit. If the power flow $S_{ij}$ exceeds the security limit, the system instability increases and a small increase in load may cause some line flow to go beyond its threshold. If the power flow exceeds $T^*_{ij}$, the breaker on the line will open to protect the line, then the line will trip. After the line has tripped, power flow will redistribute and then another line may exceed its threshold and trip, and the fault may continue to propagate.

In Figure 1, for the given load, the power flows on the lines are all below their security limits. However if the load on bus 2 is increased by $0.2 + j0.1$, the power flow on line $(1, 2)$ will exceeds $S*_{12}$; if the load on bus 3 is increased by $0.2 + j0.1$, the power flows on lines $(2, 3)$ and $(3, 4)$ will both exceed their security limits, leaving no security margins.

Operating the power system with zero security margin is dangerous and potentially it can cause fault propagation. The system instability index is then defined to be:

$$I = \sum_{(i,j)} \max\{S_{ij} - S^*_{ij}, 0\} \qquad (1)$$

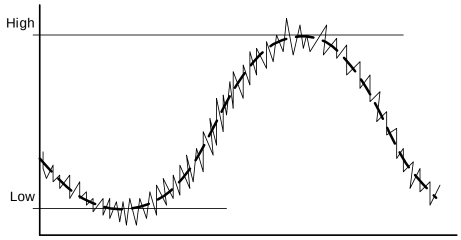

Figure 2: Approximate load distribution in a day.

Remarks:

1. If all power lines operate below their security limits, the instability index is zero; if some line goes beyond the security limit, there is a positive penalty $S_{ij} - S^*_{ij}$ added into the instability index. The power grid operator would try to minimize the system instability index.

2. In equation (1), the maximum value for $S_{ij}$ is $T^*_{ij}$, since if the flow exceeds the threshold, the line will trip so the maximum penalty for line $(i, j)$ is $T^*_{ij} - S^*_{ij}$.

The most vulnerable line is the one that has the smallest security margin: $\arg\min_{(i,j)}(S^*_{ij} - S_{ij})$ if $I = 0$, or the one that exceeds the security limit by a farthest distance: $\arg\max_{(i,j)} S_{ij} - S^*_{ij}$ if $I > 0$. However $S_{ij}$ depends on the load on the system. We will have to confine the load configuration to a few representative cases.

Let $I_B$ denote the system instability index for the base case without disturbance on the load. Let $I_{\Delta P}$ denote the system instability index when $\Delta P$ amount of load is added into the system. Increasing the load by $\Delta P$ at different locations can cause different degrees of disturbance to the system. We use $I_{\Delta P,i}$ to denote the index calculated when the load at bus $i$ is increased by $\Delta P$, then

$$I_{\Delta P} = \max_i I_{\Delta P,i} \quad (2)$$

Also notice that when the load is low, adding a $\Delta P$ amount of load will not necessarily cause system instability to increase; However when the system is already under a high load, adding a $\Delta P$ amount of load will more likely to cause dramatic increase in instability. Therefore it is meaningful to study the system behavior at the peak load. The peak load value can be obtained from historical data. Figure 2 shows that load oscillates between a low value and a high value with some noises. The expected value at the peak load will be used as the base case. The maximum deviation (instead of the standard deviation) from the expected value is used as $\Delta P$ .

The most critical line is the one whose failure will cause a maximum degree of disturbance to the system. After a component failure, the system is more prone to fault when the load increases. To measure the system's sensitivity to increasing load, we define the following metrics, called sensitivity index:

$$S = \frac{I_{\Delta P} - I_B}{\Delta P} \quad (3)$$

We compare S before and after the component failure. The component that causes the maximum increase in $mathcalS$ is identified as the most critical component. Let $S_j$ represent the metrics after the $j^{th}$ component is taken out of service, and S represents the one with the $j^{th}$ component in service. Then the change in the sensitivity index is given as:

$$\Delta S_j = S_j - S \quad (4)$$

and the most critical component is given by

$$j_{critical} = \arg\max_j \Delta S_j \quad (5)$$

## 3.2 Game-Theoretic Analysis

When malicious attacks are possible, the defender will try to allocate his resource to protect the critical components so that the system instability index is low. The attacker, with a limited capacity to launch attacks, would also adjust his strategy accordingly. Initially, the attacker would guess what is the best strategy used by the defender and make the first move. The defender, aiming at minimizing system instability, would allocate his resource to protect the most critical components. The attacker, aiming at causing as much power outage as possible, would adjust which components to attack with the objective of maximizing the system instability. Then the defender chooses which components to protect based on what the attacker has chosen and the attacker adjusts its strategy based on the decision of the opposite side. This iterative process continues until both sides reach an equilibrium or a specific stopping criteria is met. When the iteration stops, the results suggest the critical components to attack since their failure can cause maximum disruption of the system.

### *3.2.1 Vulnerability from Single Attack*

We discuss the game-theoretic model with the assumption that the attacker can attack one component at a time, and the defender can allocate his resource to protect multiple components. In the following, we demonstrate our model by using a simple example: given a power grid with $n$ power lines, the attacker chooses one of the $n$ power lines to attack, the defender allocate resource to the $n$ power lines subject to a total budget constraint.

Probability $p_{ij}$ is the probability of thwarting the attack when the attacker chooses to attack line $j$ while the defender chooses to protect line $i$ with 100% of resource. $p_{jj}$ is given as input based on how difficult it is to protect line $j$. Let $x_i$ be the proportion of resource allocated to protect line $i$, so $\sum_i x_i = 1$. Let $y_j$ be the probability that the attacker will attack line $j$, so $\sum_j y_j = 1$.

Let C1 be the cost to the system when the attack is thwarted so no line is out of service; $C2_j$ be the cost to the system when the defender fails to thwart the attack so line $j$ is taken out of service. We can choose the sensitivity index S and $S_j$ for $C1$ and $C2_j$ respectively.

The objective function is defined as follows:

$$Z = \sum_{j=1} ((1-\sum_i x_i p_{ij})C2_j + \sum_i x_i p_{ij} C1) y_j \quad (6)$$

Obviously $p_{ij} = 0$ when $i \neq j$. Thus $\sum_i x_i p_{ij} = x_j p_{jj}$ is the actual probability of thwarting the attack on line $j$ under the resource allocation. So the objective function becomes

$$Z = \sum_{j=1} ((1 - x_j p_{jj}) C2_j + x_j p_{jj} C1) y_j \qquad (7)$$

If the attacker has a fixed target, i.e., always attack the same line, then $y_j$ is a 0-1 variable, then

$$Z = \max_j \{(1 - x_j p_{jj}) C2_j + x_j p_{jj} C1\} \qquad (8)$$

Otherwise, $y_j$ is a real-valued variable between 0 and 1. In this paper we address the general strategy in which the attacker can attack one component at a time but can change the target to attack for each attempt. The vector $y = y_1, y_2, ..., y_n$ reveals the best strategy to use for the maximum benefit of the attacker: with probability $y_1$, it should attack line 1, with probability $y_2$, it should attack line 2, and so on.

The attacker would like to maximize $Z$ when $x = x_1, x_2, ..., x_n$ is known; the defender would like to minimize $Z$ when $y = y_1, y_2, ..., y_n$ is known. We can formulate a minimization linear program for the defender and a maximization linear program for the attacker. The program can start from the attacker: use an initial guess of x, solve vector y, and then the result is fed into the defender's game. The defender uses y as input, solve the minimization problem to get x, and then the result from the minimization program is fed into the maximization program. When the iterative algorithm terminates, either because the two opponents have reached equilibrium at the objective function $Z$, or a specific criteria has been met, the values in x suggest the best strategy for the defender, and the values in y suggest the best strategy for the attacker.

The iterative approach seems appealing since each party can use the knowledge of the other party's strategy to improve his own strategy. However, when an equilibrium is reached, neither the defender nor the attacker should ever need to know the strategy of the other party o adjust his own. This is because the linear programs solved by the attacker and the defender are duals of each other: the attacker tries to maximize his minimum winning no matter what strategy the defender uses, and the defender tries to minimize the maximum damage to the system no matter what strategy the attacker uses. From the strong duality theory, the optimal solutions of both programs are the same:

$$\max_{\{y:e^T y=1\}} \min_{\{x:e^T x=1\}} Z = \min_{\{x:e^T x=1\}} \max_{\{y:e^T y=1\}} Z \qquad (9)$$

The equilibrium state can be computed without using iterations. The optimal solution for either program can be computed by solving a two-player zero-sum problem.

#### 3.2.2 Vulnerability from Multiple Concurrent Attacks

If the assumption of "attacking one component a time" is relaxed, the attacker can attack any number of components simultaneously, then the problem is no longer linear. If the number of components that can be attacked simultaneously is $k$, and $1 \leq k \leq n$ is a variable, then it takes exponential time just to enumerate the cost incurred by different combination of failures. To make the problem tractable, we can assume a constant number for $k$. For example, when $k = 2$, we would need to run the static analysis to get the cost $C3_{ij}$ when two components $i$ and $j$ both are out of service at the same time. There are $C(n, 2)$ combinations. When $k = 3$, there are $C(n, 3)$ combinations. Using the same optimization framework, we can accurately solve the defender's game and the attacker's game.

## 4. EXTEND TO CONTROL AND MONITORING SYSTEM

The analysis framework can be extended to consider the control and monitoring devices and communication links in a smart grid. The cost function in 3 is based on the system instability index, which is a measurement of how much the line flow exceeds its operational security limit. If a smart grid device other than power lines is considered, we can quantify the cost of its failure or its impact on the power system by using other metrics. For example, if a PMU is attacked, and it happens to be a critical measure of a state variable, then the power system state estimation will not have a solution. But if there is redundant measurement, losing one PMU does not affect the math stability, then the system is still solvable. Solvable and non-solvable can be used as 1 and 0 respectively in the cost function. The same game-theoretic framework can be used to identify the critical components of the cyber system, but it must run separately due to the different definitions of cost functions.

## 5. CONCLUSION

This paper addresses how to identify the vulnerable and the critical components of a smart grid. The proposed method quantifies vulnerability as a cost to the power system, and considers power flow dynamics and interaction of different components in the cost function. The critical components are identified using optimization techniques in a game theory framework. The method is superior to the conventional "cut" approach that only considers the network topology and ignores the roles of the components in the system.

## 6. REFERENCES

[1] P. Maggi, D. Pozza, and R. Sisto. Vulnerability modelling for the analysis of network attacks. In Dependability of Computer Systems, pages 15 –22, june 2008.

[2] C.-W. Ten, C.-C. Liu, and G. Manimaran. Vulnerability assessment of cybersecurity for scada systems using attack trees. In Power Engineering Society General Meeting, pages 1 – 8, 24-28 June 2007.

[3] C.-W. Ten, C.-C. Liu, and G. Manimaran. Vulnerability assessment of cybersecurity for scada systems. IEEE TRANSACTIONS ON POWER SYSTEMS, 23(4):1836–1846, NOVEMBER 2008.

[4] H. Vu, K. Khaw, T. Chen, and F.-C. Kuo. A new approach for network vulnerability analysis. In Local Computer Networks, pages 200 –206, oct. 2008.