Rate Alteration Attacks in Smart Grid

Subhankar Mishra, Xiang Li, Alan Kuhnle, My T. Thai
Dept. of Comp. and Info. Sci. and Eng.
University of Florida, Gainesville, Florida 32611
Email: {mishra, xixiang, kuhnle, mythai}@cise.ufl.edu

Jungtaek Seo
The Attached Institute of ETRI
Korea
Email: seojt@ensec.re.kr

***Abstract*—Smart Grid addresses the problem of existing power grid's increasing complexity, growing demand and requirement for greater reliability, through two-way communication and automated residential load control among others. These features also makes the Smart Grid a target for a number of cyber attacks. In the paper, we study the problem of rate alteration attack (RAA) through fabrication of price messages which induces changes in load profiles of individual users and eventually causes major alteration in the load profile of the entire network. Combining with cascading failure, it ends up with a highly damaging attack. We prove that the problem is NP-Complete and provide its inapproximability. We devise two approaches for the problem, former deals with maximizing failure of lines with the given resource and then extending the effect with cascading failure while the later takes cascading potential into account while choosing the lines to fail. To get more insight into the impact of RAA, we also extend our algorithms to maximize number of node failures. Empirical results on both IEEE Bus data and real network help us evaluate our approaches under various settings of grid parameters.**

I. Introduction

Smart Grid systems along with its advancements such as smart metering, two-way communication capabilities, distributed intelligence and automation of home systems; significantly enhances the efficiency and reliability over the current power grid systems. With increasing use and integration of information technology to deliver the advancements to each and every house, it opens up many possibilities for both the producers and consumers of electricity. However, they also end up in creating opportunities for the attackers by opening up new vulnerabilities in power infrastructures. The basic sectors of the power system, i.e. generation, distribution and control, and consumption are open to a wide range of damaging cyberattacks [1]–[3] and an cyberintrusion [4] attempt may target any sector. Attack on the generation and distribution sectors need much more sophisticated and significant resources as compared to the consumption sector. Therefore, the consumption sector requires a much more attention on the countermeasures of various cyber-attacks.

Among all attacks towards the consumption sector, cyberintrusion attacks that fabricate price signals or messages through the Internet become very crucial due to the following reasons: 1) with the help of automated and distributed software intruding agents, this attack become much easier to launch. Furthermore, because of the load control and automated energy consumption scheduling (ECS) features of the smart grid, these attacks can be very effective. Given the price information and energy consumption needs of the users, ECS units accordingly schedule the timing and amount of energy consumption for each household appliance. Decisions are primarily based on minimizing the cost of energy. As the price information is obtained through the Internet, false price injection can trigger potential load altering attacks exposing the automated residential load control. 2) More importantly, this class of cyberattacks will eventually increase the load at most crucial locations in the grid causing circuit overflow or other malfunctioning that can immediately bring down the grid or cause significant damage to the power transmission and user equipments. This combined with the cascading failure can lead to major blackouts and collapse of the entire system.

Unfortunately, providing a countermeasure for this price rate alteration attack is very challenging. Unlike in case of other cyber attacks such as fabrication of command messages, negligible changes in price messages through the Internet does not come under the radar of Smart Grid security. Also the prices are controlled by the numerous private distributors. This makes the detection of this attack almost impossible. Even with negligible changes in price, the attacker can manipulate the overall distribution of the loads in the Smart Grid. It leads to failure of transmission lines as power flows go over the respective capacities of the network. As this failure cascades through the entire system, it ends up with catastrophic system failures.

Due to the above challenges, in this paper, we first attempt to identify the set of most vulnerable points along with their respective alteration in the rates, which when attacked lead to maximum number of line failures in the system. This will help us in securing Smart Grid against rate alteration attacks, by protecting these critical points. We also take into consideration, the impact of cascading failure after the initial failure of lines. With this we give two perspectives of approaching the problem. First one deals with maximizing failure of lines intially and then extending the effect with cascading failure while the second one takes cascading potential into account while choosing the lines to fail. The node failures which give us a different perspective on the impact of rate alteration attack on the system is also taken into account. Many network parameters are used to evaluate the approaches such as network stress, resistance, alterable rate, maximum resource.

Our contributions are summarized as follows:

- We define a new problem of Rate Alteration Attack (RAA) and prove its hardness and inapproximability of $O(m^{1-\eta})$, m being the number of edges and $\eta > 0$.
- We propose two approaches to the problem, namely

978-1-4799-8381-0/15/$31.00 ©2015 IEEE

MaxL and CasL. (1) MaxL tries to fail as many lines as possible and then calculate the total number of failures using the cascading effect. (2) In CasL we rank the lines on the basis of their cascading potential and then fail those lines one by one till we exhaust the resource.
- We extend the rate alteration attack to see the impact on the number of node failures in the system through MaxN and CasN.
- We experimentally evaluate our proposed algorithms in various settings, from which we infer many insights on the power network behavior to rate alteration attack.

The rest of the paper is orgranized as follows. In Section II the Smart Grid structure and model are described and the problem definition is given. Then, we provide the attacking methods for causing maximum line failures and extend it to maximization of node failures in section IV and section V respectively. Performance evaluation of the proposed algorithms is presented in section VI. The related works are discussed in section VII, which is then followed by conclusion in section VIII.

II. Network Model and Problem Definition

In this section, we describe the network model of the Smart Grid and the associated DC power flow model to understand the power flow dynamics in the power grid.

A. Smart Grid

In this paper, we consider a smart grid represented by a directed graph $G(V, E)$, where each node in V represents either the power generation stations, intermediate power transformer stations, or the consumption sector (houses, industries and data centers etc.) and edges E represent the transmission power lines between nodes. The set of nodes V includes power generation nodes $P \subset V$, consumers $D \subset V$ and other intermediate nodes $O \subset V$ with no power generation or demand.

Each node $i \in P$ has power generation output given by P_i. Every user has a demand D_i and a billing profile B_i. Every user/node i receives the electricity rate r_i from the Internet which we assume to be accessible to the attacker to manipulate or alter. A certain portion of the rate is allowed to change for each user given the baseline constraints set up or hard coded by electricity distribution companies. Hence, we have the maximum rate change (MRC) ρ_i, which represents the maximum change in the rate that the attacker is allowed. The attacker can choose what percentage of the MRC it wants to change denoted by z_i which lies between 0 and 1. Given the automated demand side management, one of the "smart" features of the smart grid, the rate change causes automated increase in the use of the demand for the same household or the company, such as starting up the laundry, more frequent use of the heating/cooling devices, etc. There is a cost associated with the change done by the attacker. Here we consider a linear cost function $c(.)$ for simplicity. The attacker is restricted by the maximum resource A, i.e. the attacker is bound by a given total cost to alter the rates at various positions in the grid.

As explained in previous section about automation of smart grids, and considering the sensitivity of the users to the billing, the total bill is given by $(1 + k_i)B_i$ where B_i represents the user's targeted billing amount and k_i represents the sensitivity of the user towards billing amount. $k_i = 0$ indicates that the user is not willing to pay anymore than the targeted billing amount. For simplification we allow the k_i with a maximum value of 1. The line (i, j) fails when the power flow through the transmission line goes over its capacity.

Every transmission line is a directed edge $(i, j) \in E$ connecting node i with node j, has a maximum capacity u_{ij}. The power flow is given by f_{ij}. When the power through a transmission line goes over the capacity, it breaks due to thermal heating. As we try to balance the total power generation to the total demand of the consumers, the erratic demands can lead to power flow through a transmission line increasing beyond the capacity of the line; leading to the failure of line.

B. Cascading Failure

Cascading failure [8]–[10] is common in power grids when one of the elements fails (completely or partially) and shifts its load to nearby elements in the system. Those nearby elements are then pushed beyond their capacity so they become overloaded and shift their load onto other elements. In comparison to graph-theoretical networks, in the power grid, the power flows have no strict capacity bounds on the lines and are governed by the laws of physics. However, there is a rating threshold marked for each transmission line. When the power flow through a line exceeds this threshold, the line ultimately experiences thermal failure. This outage of lines getting turning off or tripped, alters the network topology and results in a different flow pattern which may cause other line outages. Repetition of this process leads to a cascading failure.

C. DC Power Flow

The common modeling of the behavior of the power grids is done using a system of non-linear, non-convex equations which describe the physics of AC power flows [6]. A reasonably close solution to the system can be produced under normal operating conditions. Then after few more iterations, the Newton-Raphson methods will converge to a useful solution.

However, Newton-Raphson may fail to converge under extreme operating conditions. Even though Newton-Raphson is relatively fast it may be too slow when a large volume of power flow computations is required. For these and other reasons, researchers normally prefer and rely on the linearized or DC power flow approximation. This is solved far more quickly and is proved to be accurate under good operating conditions.

Here, we briefly describe the linearized or DC power flow model. In the linearized approximation, we are given a power grid represented by a directed graph G, where:

- Each node $i \in V$ corresponds to either a power generator (i.e., a supply node), or to a load (i.e., a demand node), or to a node that neither generates nor consumes power. The set of generator nodes are denoted by P.

- If node i corresponds to a generator, then there are values $0 \le P_i^{min} \le P_i^{max}$. If the generator is operated, then its output must be in the range $[P_i^{min}, P_i^{max}]$; if the generator is not operated, then its output is zero. In general, $P_i^{min} > 0$.
- If node i corresponds to a demand, then the "nominal" demand is given by D_i^{nom}. The set of demands or demand nodes is denoted by D.
- The edges E represent power/transmission lines. For each line (i,j), two parameters are given i.e. $x_{ij} > 0$ (the resistance or reactance) and u_{ij} (the capacity).

Now, given a set P of operating generators, the linearized power flow is a solution to the system of constraints is given in the following set of the equations. Here, for each edge (i,j), f_{ij} represents the power flow on the edge (transmission line) (i,j). In the case where $f_{ij} < 0$, power is effectively flowing from j to i. Additionally, the phase angle at node i is given by the variable θ_i. Again, if i is a generator node, then it will have a variable P_i and if i represents a demand node, it will have a variable D_i. Also given a node i, $\delta^+(i)(\delta^-(i))$ is the set of lines oriented out of (into) node i.

The power flow equations are given below:

$$\sum_{(i,j)\in\delta^+(i)} f_{ij} - \sum_{(j,i)\in\delta^-(i)} f_{ji} = \begin{cases} P_i & i \in P \\ -D_i & i \in D \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\theta_i - \theta_j - x_{ij} f_{ij} = 0, \quad \forall (i,j) \quad (2)$$

$$P_i^{min} \le P_i \le P_i^{max}, \quad \forall i \in P \quad (3)$$

$$0 \le D_j \le D_j^{nom}, \quad \forall j \in D \quad (4)$$

In the system $G(V,E)$ given above, constraints (1), (3) and (4) are typical for network flow models, representing flow balance (i.e., net flow leaving a node equals net supply at that node), generator and demand node bounds respectively. Constraint (2) is a commonly used linearization of more complex equations describing power flow physics as explained in the beginning of this section.

Lemma 1. *Let C be given, and suppose G is connected. Then for any choice of nonnegative values P_i (for $i \in C$) and D_i (for $i \in D$) such that*

$$\sum_{i\in C} P_i = \sum_{i \in D} D_i \quad (5)$$

system (2)-(3) has a unique solution in the f_{ij}; thus, the solution is also unique in the $\theta_i - \theta_j$ (over the lines (i,j)).

Lemma 1 concerns the subsystem of $G(V,E)$ consisting of (1) and (2). In particular, the "capacities" u_{ij} play no role in the determination of solutions. When the network is not connected Lemma 1 can be extended by requiring that hold for each component i.e. the total supply and demand within each of the connected component is equal.

D. Problem Definition

We are now ready to formally define our problem as follows.

Definition 1 (Rate alteration attack (Lines) RAA)**.** *Given a Smart Grid system $G(V,E)$, P being the set of power generators, and D being the set of demand nodes, maximum attack resource A, and electricity billing rate r_i, billing constraints $(1+k_i)B_i$ and maximum rate change ρ_i, for each demand node i. Compute an attacking strategy $z = \{z_i\}, i \in D, z_i \in [0,1]$ that alter the rates of those demand nodes, such that the total number of the* ***line*** *failures is maximized.*

III. Complexity

In this section, we first prove the NP-Completeness of RAA. We next study its inapproximability which shows the best approximation ratio that one can do.

Theorem 2. *The rate alteration attack (RAA) is NP-Complete.*

Proof. First, we define the decision version of rate alteration attack problem.

Definition 2 (Decision version of RAA (Lines))**.** *Given a system $\{G(N,P), P, D, A, r_i, k_i, B_i, c_i, \rho_i\}$ RAA asks whether or not there is an attacking strategy $z = \{z_i\}, i \in D, z_i \in [0,1]$ that alter the rates of those demand nodes will result in failure of at least m lines.*

We first prove rate alteration attack problem is in NP. Given the set of demands which rates will be altered in the system, we can verify if number of failed lines is greater than m in polynomial time.

To prove the NP-completeness, we reduce from the maximum coverage problem (MC). The decision is defined as following.

Definition 3 (Decision version of MC)**.** *Given a set $U = \{S_1, S_2, S_3, ..., S_n\}$, the space $E = \cup_{S_i \in U} S_i$ and a number k, MC asks whether or not there exists a subset $S' \in U$ such that $|S'| \le k$ and the number of covered elements$|\cup_{S_i \in S'} S_i|$ is greater than $m(m < |E|)$.*

Here we show how to reduce MC to RAA(see Fig.1). Let (E, U, k) be an instance of MC, and assign one node g as a generator. For each set $s_i \in S$, create a user node u_{s_i}. For each element $e \in E$, create a node n_e, and add edges (g, n_e) to the generator. Now, for each u_{s_i} add edges (n_e, u_{s_i}) iff $e \in S_i$. Set $B_i = |S_i| - 1$, $r_i = 1$, $k_i = 0$. Also, suppose that each node u_{s_i} initially has demand $d_i = |S_i| - 1$, so that the initial flow of edges (n_e, u_{s_i}) will be $1 - \frac{1}{|S_i|}$. Let $\rho_i = \frac{1+\epsilon}{|S_i|+\epsilon}$ and the capacity of lines (n_e, u_{s_i}) to be 1 (all the other lines are assumed not to broken in this attack). So that when we choose to decrease the rate of user node i to its lowest $(r_i - \rho_i)$, the new demand will be $d_i' = |S_i| + \epsilon$. Also, assume that the cost $c_i = 1$ and the budget $A = k$. In addition, let the reactance x_{ij} be equal to 1. It is easy to show that this construction takes polynomial time.

Then we will prove if MC has a solution T, $|T| \le k$ that guarantee a coverage of m elements, RAA has a rate

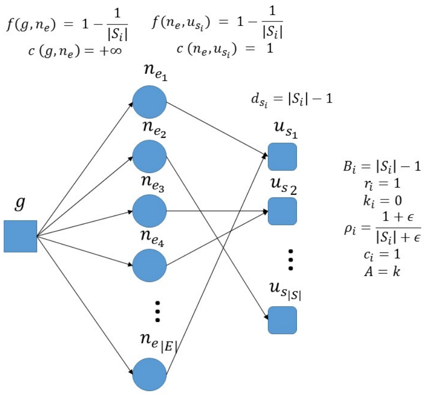


Fig. 1: RAA reduced from MC

alternation strategy to fail m lines with budget k. Also, if RAA has a solution z to fail m lines, there is a solution for MC to cover m elements.

$\Longrightarrow$ Assume MC has a solution $T(|T| \leq k)$ that cover at least m elements. In RAA, if we choose to alter the rates of all the user nodes corresponding to the sets in T to the lowest, i.e. set $z_i = 1, s_i \in T$, then the demand of each selected user nodes will increase to $d'_i = |S_i| + \epsilon$. We keep all the other user nodes unchanged.

So that the flow on edges (n_e, u_{s_i}) will be $1 + \frac{\epsilon}{|S_i|}$ and those edges connected to set nodes in T will be broken. Since there are at least m nodes covered by the sets in T, we are able to fail at least m lines based on this strategy.

$\Longleftarrow$ Now assume RAA has a solution to fail m lines with altering the rates of at most k user nodes. Since all the edges connected to those altered user nodes with $z_i = 1$ must be failed and all the other edges cannot be failed, we know for sure that the m failed lines are all connected to one of the nodes with $z_i = 1$. Thus, if we choose the corresponding sets of those user nodes as a solution of MC, we can cover at least m elements. □

Theorem 3. *There is no $O(m^{1-\eta})$-approximation algorithm for the cascading edge failure problem unles $P = NP$, where m is the number of edges, for any $\eta > 0$.*

Proof. Let an instance $(X, \mathcal{S}, k)$ of the set cover problem be given, with $|X| = g, |\mathcal{S}| = h$. For each $S \in \mathcal{S}$, create generator vertex u_S, demand vertex v_S and transmission line (u_S, v_S) with capacity $1/(|S|+1) - \epsilon$ and reactance r_1. For each $x \in X$, create vertices u_x, v_x and a line (u_x, v_x) with capacity

$$\min_{S: x \in S} 1/(|S|+1) - \epsilon,$$

and reactance r_1. For every set S containing x, add edges $(u_S, u_x), (v_x, v_S)$ with reactance r_1 and capacity 2. Add vertices u_0, v_0 and edges (u_x, u_0), (v_x, v_0) for all $x \in X$ with reactance r_2 and capacity 2. Finally, for $1 \leq i \leq l$, add extra lines (u_0, u_i), (u_i, v_0), each with capacity $k/l - \epsilon$ and reactance r_2. Choose l so that $k/l < 1/g$. It is possible to choose r_1, r_2 such that for each set S, unless all element lines corresponding to elements of S are broken, only a negligible amount of current from u_S to v_S flows over lines of reactance r_2.

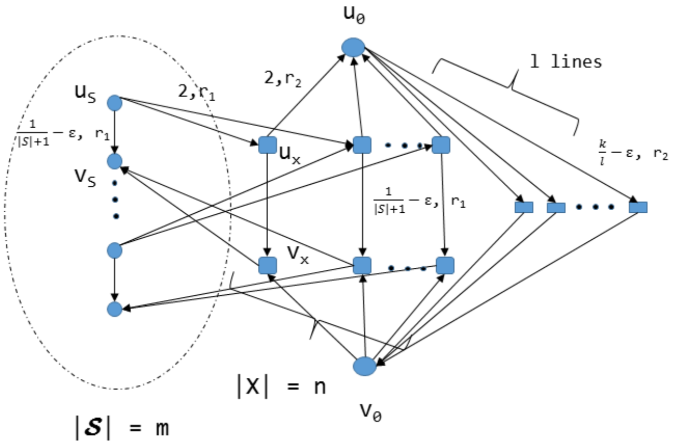

u0
1 lines
2,r2
2,r1
uS
$\frac{1}{|S|+1} - \varepsilon$, r1
vS
ux
$\frac{1}{|S|+1} - \varepsilon$, r1
$\frac{k}{l} - \varepsilon$, r2
vx
|X| = n
v0
|𝓢| = m

Fig. 2: Inapproximability proof using set cover. The labels on the edges represent (capacity,reactance) respectively.

Initially, the demand at each vertex v_S is set to 0, with maximum possible demand equal to 1. We will show that the $2l$ extra lines can be broken with budget k iff there is a set cover of size k.

Suppose there is a set cover SC of size k. For each $S \in SC$, raise the demand of v_S to its maximum value 1. This breaks every element line; hence the extra lines receive flow k/l, which causes them to break. So in total $k+g+2l$ lines break.

Suppose there is no such cover. The most element lines that a feasible solution could break would be $g-1$. So the extra lines have flow at most $k/(l+1) < k/l - \epsilon$. Hence they remain intact. Thus, the largest possible number of lines broken would be at most $k + g - 1$.

Now the total number of lines $m = c_1 g + c_2 h + 2l$. There is a constant K such that

$$K m^{1-\eta} < (g + k + 2l)^{1-\eta}.$$

Suppose

$$l \geq \frac{\left(\frac{g+k}{K}\right)^{1/\eta} - g - k}{2}.$$

Suppose we have an $m^{1-\eta}$-approximation algorithm $\mathcal{A}$. If there is a set cover of size k, $\mathcal{A}$ produces a solution breaking at least

$$\frac{g+k+2l}{m^{1-\eta}} \geq K \frac{g+k+2l}{(g+k+2l)^{1-\eta}} = g + k.$$

edges, while if there is no such cover, it produces a solution of size less than $g + k$. □

IV. Solutions

In this section, we provide solutions to the rate alteration attack problem. Two approaches are considered here and are given as follows:

- The aim here is to initially fail as many lines as possible without considering the cascading failure. Integer programming is formulated for this case. After initial failure of lines, the cascading effect of the same is taken into account. This approached is discussed in section IV-A

- Second approach is explained in section IV-B and consists of ranking the edges according to their cascading potential and then failing the highest ranked with minimum cost. Then the process is repeated until the maximum resource available to the attacker gets exhausted.

A. Maximizing the line failures (MaxL)

We formulate the maximization of number of line failures for the given problem. Let a binary variable y_{ij} indicate the failure of the transmission line $(i, j) \in E$. When $y_{ij} = 1$, the line fails and is 0 otherwise. Our goal is to maximize the total of number of line failures given by $\sum_{e(ij)\in E} y_{ij}$.

The formulation is given as follows:

$$\max \sum_{e(ij)\in E} y_{ij} \tag{6}$$

$$s.t. \quad (1 + k_i)B_i \geq D_i \cdot (r_i - z_i \cdot \rho_i) \qquad \forall i \in D \tag{7}$$

$$\sum_{i\in D} c_i(z_i) \leq A \tag{8}$$

$$y_{ij} < 1 + \frac{f_{ij} - u_{ij}}{u_{ij}} \qquad \forall e(ij) \in E \tag{9}$$

$$Eqs.(1) - (4) \tag{10}$$

$$y_{ij} \in \{0, 1\} \qquad \forall e(ij) \in E \tag{11}$$

$$z_i \in [0, 1] \qquad \forall i \in D \tag{12}$$

where constraint (7), (8), (10) represent the billing calculation resulting in the change in the demand values, constraint on the resource available to the attacker and linearized DC power flow equations. Constraint (9) represents the line breaking scenario, that is, the transmission line breaks when the total power flow through the line exceeds the capacity of the corresponding transmission line.

a) Cascading Effect: Line failures are often associated with the cascading effect. Failure of lines changes the power flows in other lines and results in failure of more lines and often leading to black outs of entire region.

The cascading failure model is described in [7] (extension of model in [8]). At the steady state, G is connected and total supply is equal to the demand. When there is a failure, some edges/lines are removed from the graph G (i.e. gets disconnected). The total supply and total demand are now adjusted within each component by decreasing the demand and supply at loads and generators respectively. Using the power flow equations, the power flow is recalculated. The new flows may exceed the capacity and as a result, the corresponding lines will become overheated. The outages are modeled by moving average of the power flow $\tilde{f}^t_{ij}$: $\tilde{f}^t_{ij} = \alpha f_{ij} + (1 - \alpha)\tilde{f}^{t-1}_{ij}$. The moving average approximates thermal effects, including heating and cooling from prior states to first order [11]. The pseudocode for the cascading failure is given by Algorithm 1.

The outage rule is given as follows:

$$P(\text{(i,j) faults at round } t) = \begin{cases} 1 & \tilde{f}^t_{ij} > (1 + \epsilon)u_{ij} \\ 0 & \tilde{f}^t_{ij} \leq (1 + \epsilon)u_{ij} \end{cases} \tag{13}$$

Algorithm 1: Cascade Failure Template

```
Data: Connected Power Grid Network $G(V, E)$
Result: $S_1$: Lines which failed
$S_2$: Nodes which failed
1 while Network is not stable do
2     Adjust the total demand to the total supply within
      each island.
3     Use equations (1)-(4) to calculate power flows in G.
4     For all lines computer the moving average
      $\tilde{f}^t_{ij} = \alpha f_{ij} + (1 - \alpha)\tilde{f}^t_{ij}$.
5     Remove all lines that have moving average flows
      greater than the capacity ($\tilde{f}^t_{ij} > (1 + \epsilon)u_{ij}$) and add
      to $S_1$.
6     Add the failed nodes to $S_2$.
7     If no more line fails, then network is stable, break
      the loop.
8 Return $S_1, S_2$
```

The adjustment in Step 2 handles the case of islanding, where the line outages create isolated components of the network. A newly created island might have an excess of generation over demand and in such a case we assume that the excess is removed by reducing the output of all generators in that island in equal amounts. The case of excess demand is handled similarly. The process is continued until there are no lines to be failed and there are no overloads. We use the cascading effect of the initial line failures due to rate alterations as formulated in the integer programming above.

B. Cascade Potential Ranking (CasL)

In this section, instead of breaking the lines and then using the cascading template to calculate the total number of lines, we use the cascade algorithm to rank the edges according to their cascading potential. Cascading potential of the edge* is calculated by calculating the numbers of other edges that fail due to load distribution following the edge* failure.

Next, the objective is to fail the highest ranking edge with minimum cost of rate alteration given by $\sum_{i\in D} c_i(z_i)$. The rest of the constraints follow the same reasoning as the maximization of lines failure except constraint (16) which represent the failure of the edge (e^*) with highest cascade potential. The integer programming for calculating the minimum cost for breaking the edge e^* is given by $MCB(e^*)$. The above process is repeated till the attacker runs out of the maximum resource A. Algorithm 2 states the pseudocode for the CasL algorithm.

Minimum cost to break e* [MCB(e*)]

$$\min \sum_{i\in D} c_i(z_i) \tag{14}$$

$$s.t. \quad (1 + k_i)B_i \geq D_i \cdot (r_i - z_i \cdot \rho_i) \qquad \forall i \in D \tag{15}$$

$$f_{ij} > u_{ij} \qquad ij = e^* \tag{16}$$

$$Eqs.(1) - (4) \tag{17}$$

$$z_i \in [0, 1] \qquad \forall i \in D \tag{18}$$

Algorithm 2: Cascade potential ranking lines (CasL) algorithm

```
Data: Connected Power Grid Network $G(V, E)$
Result: Total number of failed lines $L$
Intialize temporary cost $T = 0$
while $T < A$ do
    for each $e \in E$ do
        Remove line $e$
        Cascade_potential(e) = number of failed edges [by Alg. 1]
    Push edges to $E'$ with decreasing order of Cascade_potential()
    for each $e \in E'$ in the order do
        if $MCB(e) + T < A$ then
            e* = e
            break
    $T = T + MCB(e*)$
    $L = L + Cascade\_potential(e*)$
    Remove the failed edges from $E$
Return $L$
```

V. Extension

In this section, we extend our problem definition towards maximization of number of failed nodes instead of maximization of number of failed transmission lines. Maximizing the number of line failures may not always achieve desired results, where maximum lines after failure do not have significant effect on the network and network is still operational with retaining most of the nodes, hence we explore the maximization of the node failures.

Definition 4 (Rate Alteration Attack (Nodes))**.** *Given a Smart Grid system $G(V, E)$, P being the set of power generators, and D being the set of demand nodes, maximum attack resource A, and electricity billing rate r_i, billing constraints $(1+k_i)B_i$ and maximum rate change ρ_i, for each demand node i. Compute an attacking strategy $z = \{z_i\}, i \in D, z_i \in [0, 1]$ that alter the rates of those demand nodes, such that the total number of the* ***node*** *failures is maximized.*

A. Maximizing the node failures (MaxN)

$$\max \sum_{i \in D} w_i \tag{19}$$

$$s.t. \quad (1 + k_i)B_i \geq D_i \cdot (r_i - z_i \cdot \rho_i) \qquad \forall i \in D \tag{20}$$

$$\sum_{i \in D} c_i(z_i) \leq A \tag{21}$$

$$w_i \leq 1 - \frac{\sum_{(j,i) \in \delta^-(i)} f_{ji}}{\sum_{(j,i) \in \delta^-(i)} u_{ji}} \qquad \forall i \in D \tag{22}$$

$$Eqs.(1) - (4) \tag{23}$$

$$w_i \in \{0, 1\} \qquad \forall i \in D \tag{24}$$

$$z_i \in [0, 1] \qquad \forall i \in D \tag{25}$$

We define a binary variable w_i which indicates the failure of the node. Node fails when it has no connectivity with any of the power generators in the network i.e. has no power flows into it. $w_i = 1$ indicates the failure of the node and 0 otherwise. The object is to maximize node failures i.e., maximize $\sum_{i \in D} w_i$.

Constraint 22 calculates the value of the w_i based on the incoming transmission line flow and capacity values. If the total incoming flow is 0, then the node has no connectivity with any power generator in P.

B. Cascade potential ranking (nodes) algorithm

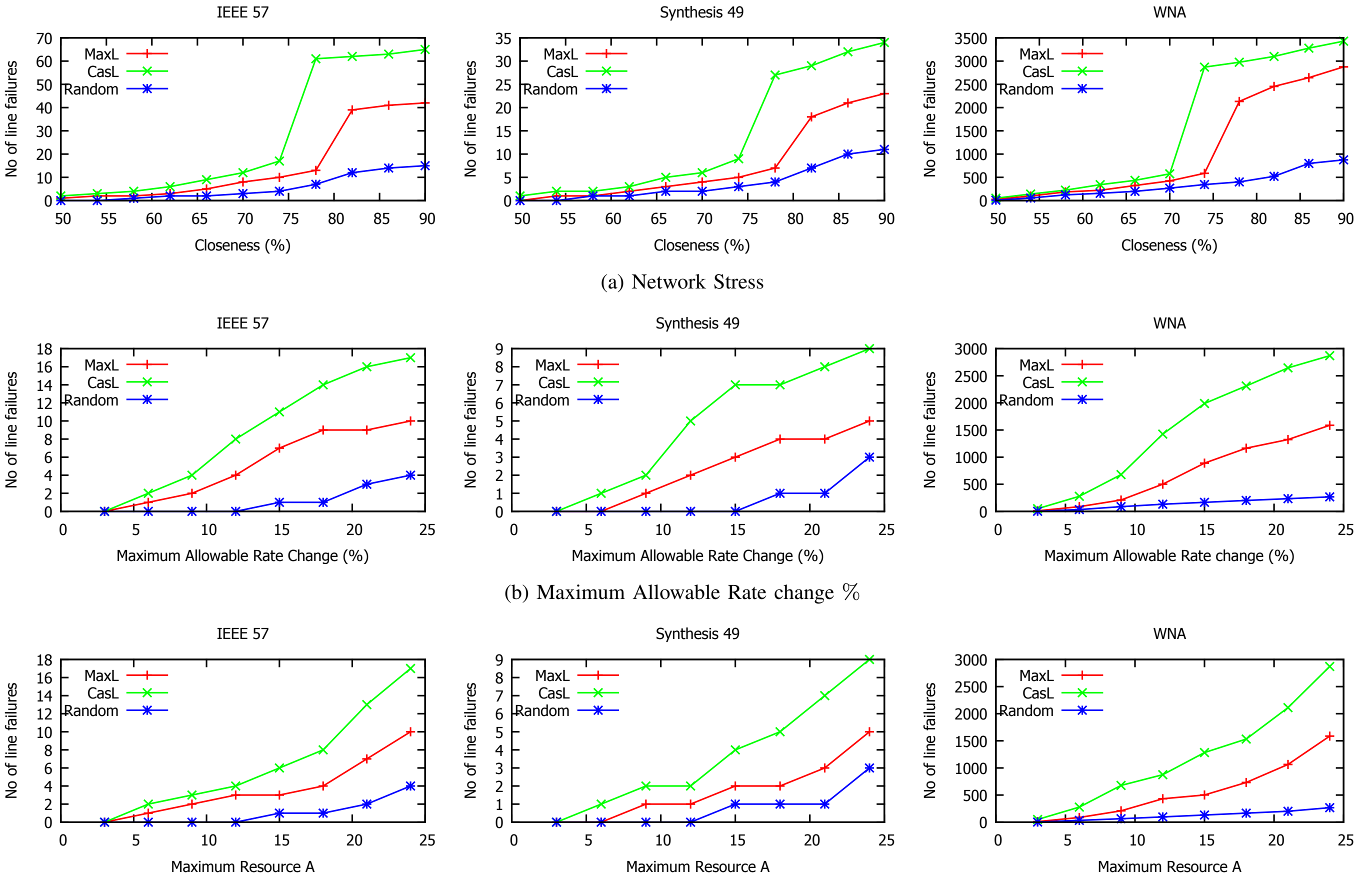
Similar to CasL, the base cascading potential ranking algorithm is used but with the consideration of the importance of failure of nodes. Here, the cascade ranking IP is used to calculate the minimum cost to break the highest cascade potential edge. We continue the similar execution as in case of CasL, until we exhaust the attacker resource A. The Algorithm 3 represents the pseudocode of the this approach and is denoted by CasN.

Algorithm 3: Cascade potential ranking Nodes (CasN) algorithm

```
Data: Connected Power Grid Network $G(V, E)$
Result: Total number of failed nodes $L$
Intialize temporary cost $T = 0$
while $T < A$ do
    for each $e \in E$ do
        Remove line $e$
        Cascade_potential(e) = number of failed nodes [by Alg. 1]
    Push edges to $E'$ with decreasing order of Cascade_potential()
    for each $e \in E'$ in the order do
        if $MCB(e) + T < A$ then
            e* = e
            break
    $T = T + MCB(e*)$
    $L = L + Cascade\_potential(e*)$
    Remove the failed edges from $E$
Return $L$
```

VI. Performance Evaluation

In this section, we evaluate the efficiency of the different algorithms we proposed. In the experiments reported in this section we used a 3.0 GHz Xeon machine with 2 MB L2 cache and 12 GB RAM. All experiments were run using a single core. The LP/IP solver was Cplex [12], with default settings. 50 runs of each cycle was run and averaged for consistency. We use the random algorithm as a baseline to compare our proposed algorithms.



(a) Network Stress

(b) Maximum Allowable Rate change %

(c) Maximum Resource A

Fig. 3: Number of line failures

A. DataSet

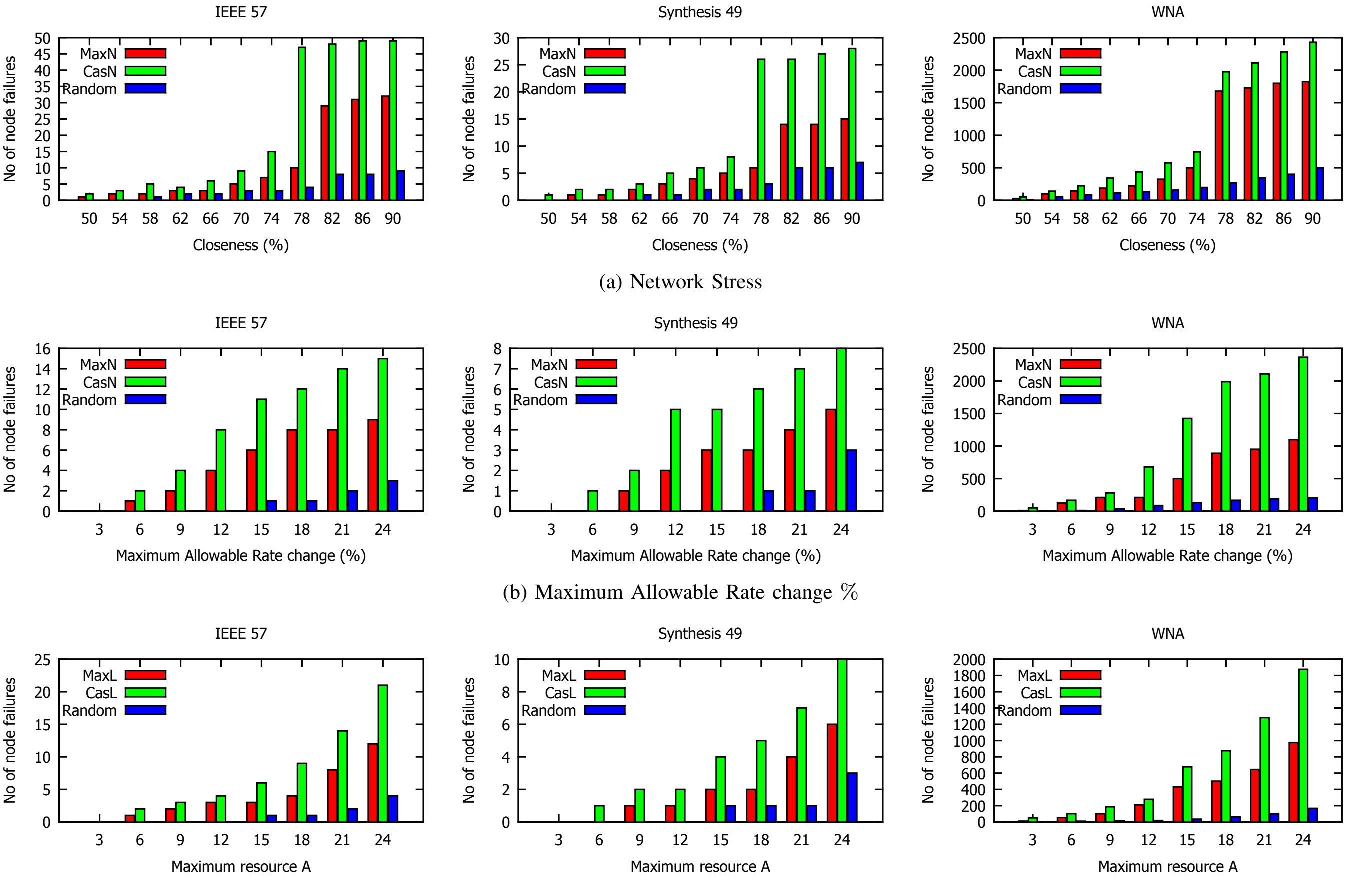
For the experiments, we used the datasets of following types:

1) Two of the IEEE test cases [13]: the 57 bus case (57 nodes, 78 arcs, 4 generators and 38 demand nodes).
2) Artificial example was also created. A square grid network was generated with 49 nodes and 84 arcs. It included 4 generators and 14 demand nodes, rest were intermediate nodes [10].
3) Western North American (WNA) [14]: power grid network with 4941 stations and 6594 transmission lines to run experiments and assign the parameters taking cue from IEEE datasets.

B. Number of line failures

We provide the comparative analysis of the behavior of the different proposed algorithms in three scenarios of datasets. In the Fig. 3a, the algorithms are tested against the network stress. To increase the network stress, the average demands of the users is increased over the whole network. This leads to unique solution in f_{ij}, such that the flows approach the capacity of their corresponding transmission lines. We vary the network stress (closeness) from 50% to 90% to check the response of the three algorithms. We see CasL break the network in a significant way at about $70-80\%$ of the network stress, while MaxL has a bump in the number of transmission line failures when the network is about $77-87\%$ stressed. CasL definitely performs better than MaxL and way better than Random. We also infer that the network stress is an important factor while considering this type of attacks.

Next the reaction of various algorithms while we change the maximum alterable rate change δ_i is evaluated. Note that even if there is a change in the maximum alterable rates, the increase in demand also depends on the billing profile of the user. And hence after a certain variation in maximum alterable rates, there is no impact on the demand of the user. δ_i is varied from 5% to 25% to observe the performance. In this case, we keep the maximum allowable resource constant at 25% to maintain consistency. As argued, the declining effect of the δ_i change, in Fig. (3b) it is evident that as the increase δ_i, the overall change in the number of transmission line failures although increases but at a smaller rate, implying that δ_i change does have diminishing impact because of baseline settings and user sensitivity. Finally we evaluate with the change of maximum resource allocated A to the attackers, which is basically how many houses and how much of rate



(a) Network Stress

(b) Maximum Allowable Rate change %

(c) Maximum Resource A

Fig. 4: Number of node failures

change can be attacked. In Fig. 3c, it can be observed that as A is increased from 5% to 25%, the increase in the failure increases at a faster rate as the attacker has the opportunity to attack more users with a higher degree of rate change with a various combinations of those.

C. Number of node failures

In order to understand the impact of algorithms to the power system in terms of number of node failures, we study the behavior of the different proposed algorithms in three scenarios of datasets and compare them against each other. First the algorithms are evaluated against network stress and then compared with each other by number of node failures caused by attacking scenario. Settings and variation of the stress in the network are calculated and maintained exactly same as in case of line failures. As it can be seen in the Fig. 4a, the CasN algorithm performs better than MaxN and random algorithms in all the three different networks. It can also be observed that the number of node failures is although dependent on the number of line failures but does not equal to the number of transmission failures. The reason being the flow redirection sustains the nodes survival i.e. keeps them connected to the power generators.

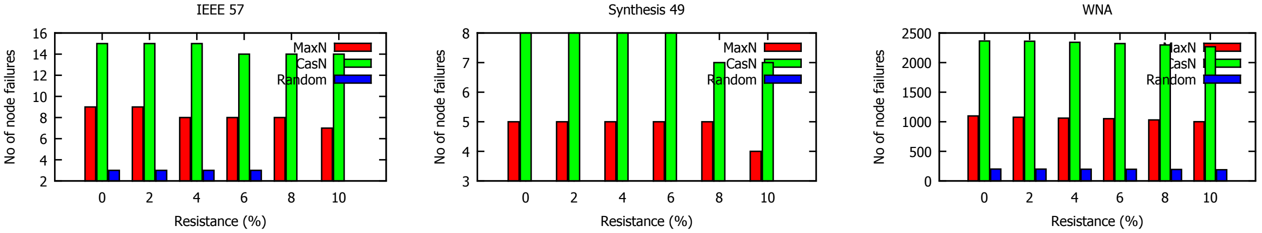
Now, we compare the node failure attacking algorithms while change the maximum alterable rate change (δ_i). The settings and variation is kept in synchronization with the lines failure settings. In Fig. 4b we observe that as δ_i increases , the overall change in the number of node failures although increases but at a smaller rate, implying that δ_i change does have diminishing impact because of baseline settings and user sensitivity, same as in case of the line failures.

Finally we evaluate with the change of maximum resource allocated A to the attackers same as in the case of line failures in Fig. 4c. Rapid increase in the number of node failures as the maximum resource allocated A is increased, also helps us to infer that if we can provide protection to certain portion of the nodes, the effect of this kind can be managed, but that certain portion seems like a unfathomable number for now at around 80%.

D. Attack resistant networks

The performance of the algorithms are compared while changing the resistance of the power grid itself. The power grid is made attack resistant by making some nodes absolutely secured i.e. are assumed to be not vulnerable to any type of cyber attack. Although it is important to note that this might not be a practical scenario, harsh conditions are simulated for the attacker, where the users might be strict organizations, IT companies and highly vigilant users.

The number of the ”strict” or resistant nodes are varied



(a) IEEE 57

(b) Synthesis 49

(c) WNA

Fig. 5: Number of node failures vs Resistance

from 0% to 10% in order to check the response of the various algorithms and are randomly assigned in the network. Each algorithm is ran 100 times, to get a more noise free picture of the reaction. The value of maximum resource A is kept at 24 and and maximum allowable rate change at 24%. As the network becomes more and more resistant, the algorithms go down on their performance. Note that the attacker has no idea which nodes have been protected, hence it just launches the attack, but the effect at the resistant node is null. As we see in the Fig 5, the effect of the resistance of the network although affects the algorithm, the random placements of the resistance without the knowledge of the attacking strategy does not help in preventing most of the damage incurred.

VII. Related Work

A survey and study of Internet based load altering attacks against Smart Grids was presented in [5]. There are three types of cyber-attacks Type-I, Type-II and Type-III aimed at generation, distribution and control, and consumption respectively. Defense mechanisms aimed at protection of Smart Grid consumption sector through protection of command and price signals, protection of smart meters, attack detection and learning of demand patterns, load shedding and load relocation are proposed along with cost-efficient load protection for Type-III attacks. However they do not analyze the effect of rate alteration through Internet and consider the changes in the load measurements without any baseline constraints, as flagging of erratic load increases can be flagged easily beyond those maximum or minimum values.

Load redistribution attacks in Power Systems are modeled by Yuan et. al. [15]. They develop the concept of load redistribution (LR) attacks, a special type of false data injection attacks, and analyze their damage to power system operation in different time steps with different attacking resource limitations. They identify the most damaging LR attack through a max-min attacker-defender model and provide a protection strategy. However, the KKT-based method which is used to identify the most damaging attack from an attacker's perspective is computationally inefficient and also do not include the communication network's role in state estimation of Smart Grids.

VIII. Conclusion

We define a novel problem of rate alteration attacks in the Smart Grid and show the hardness of the problem and inapproximability. Two approaches are studied to exploit the Smart Grid vulnerability to maximize the number of line failures. IP and cascading extension for the first approach and cascading ranking algorithm is provided as the second approach. In addition, we also extend the problem to maximizing the number of node failures. Experimental results on both IEEE bus data and synthetic data along with real data for this new problem, give us insightful knowledge about the reaction of various approaches to different network settings and parameters. For future work, the protection strategy against this newly formulated rate alteration attack can be studied.

IX. Acknowledgement

This work is partially supported by NSF Career Award 0953284, DTRA HDTRA1-14-1-0055 and Korea NSRI.

References

[1] A. R. Metke and R. L. Ekl, *Security technology for smart grid net- works*, IEEE Trans. Smart Grid, vol. 1, no. 1, pp. 99?107, Jun. 2010.

[2] S. M. Amin, *Securing the electricity grid*, Bridge, vol. 40, no. 1, pp.13?20, Mar. 2010.

[3] N. Ye, J. Giordano, and J. Feldman, *Securing the electricity grid*, Commun. ACM, vol. 44, no. 8, pp. 76?82, Aug. 2001.

[4] O. Kosut, L. Jia, R. Thomas, and L. Tong, *Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures*, in Proc. IEEE Int. Conf. Smart Grid Commun., Gaithersburg, MD, Oct. 2010.

[5] Amir-Hamed Mohsenian-Rad, and Alberto Leon-Garcia. *Distributed Internet-Based Load Altering Attacks Against Smart Power Grids*, IEEE Transactions on Smart Grid, Vol. 2, NO. 4, December 2011

[6] A. Bergen and V. Vittal, *Power Systems Analysis*, Prentice-Hall (1999).

[7] Andrey Bernstein, Daniel Bienstock, David Hay, Meric Uzunoglu, and Gil Zussman, *Power Grid Vulnerability to Geographically Correlated Failures Analysis and Control Implications*, IEEE INFOCOM 2014.

[8] J. Chen, J. S. Thorp, and I. Dobson, *Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model,* Int. J. Elec. Power and Ener. Sys., vol. 27, no. 4, pp. 318 326, 2005.

[9] D. Bienstock, *Optimal control of cascading power grid failures*, in PES General Meeting, July 2011.

[10] D. Bienstock and A. Verma, *The N - k problem in power grids: New models, formulations, and numerical experiments*, SIAM J. Optim., vol. 20, no. 5, pp. 2352?2380, 2010.

[11] M. Anghel, K. A. Werley, and A. E. Motter, *Stochastic model for power grid dynamics,* in Proc. HICSS07, Jan. 2007

[12] IBM ILOG CPLEX Optimization Studio. http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud, 2014.

[13] The IEEE reliability test system 1996, IEEE Trans. Power Syst., vol. 14 (1999) 1010 - 1020.

[14] Duncan J Watts and Steven H Strogatz. *Collective dynamics of small-world networks*. nature, 393(6684):440-442, 1998.

[15] Y. Yuan, Z. Li, and K. Ren. *Modeling load redistribution attacks in power systems*. IEEE Trans. Smart Grid, 2(2):382-390, 2011.