Vulnerability Assessment of Social-Smart Grids: An Algorithmic Approach

Lan N. Nguyen, J. David Smith and My T. Thai
Department of Computer & Information Science & Engineering
University of Florida
Gainesville, Florida 32611
Email: {nglan,jdsmith,mythai}@cise.ufl.edu

***Abstract*—Utility providers are gradually deploying social networks as a useful addition to the Smart Grid in order to help engage consumers in energy management and efficient usage. Besides its benefits, is there any negative impact to the Smart Grid? In this paper, we investigate the vulnerability of Smart Grid when integrating into social networks, where attackers utilize misinformation propagation in social network to alter electricity customer's behavior with the goal of causing degradation to power infrastructure. Stand in both perspectives of power facility administrator and adversary, we model the vulnerability assessment of the system under an optimization problem, which enables us to provide theoretical analysis and behavior investigation of the system based on the complexity theory. As solving the problem is challenging, we propose heuristic solutions and show their efficiency on assessing the system's vulnerability in the presence of misinformation attacks. Therefore, we conclude that misinformation attacks must be considered when developing the security model for Socially-enabled Smart Grid technology and planning mitigation techniques.**

I. Introduction

Recent advances in extending existing power grids to Smart Grids and incorporating social networks into the energy ecosystem have revolutionized the landscape of energy management. This opens up a number of opportunities for utility providers to improve energy efficiency by setting prices such that their customers are encouraged to use demanding appliances outside of peak hours. It also allows customers to react to this pricing change by determining a suitable time to use their appliances. Furthermore, with the incorporation of social networks, peak load could be reduced by 6% [1] and overall energy usage could be reduced by roughly 9% [2]. The prospect of taking advantage of social information diffusion to spread energy-saving and efficiency tips has led to a number of recent proposals [3], [4] to integrate data from home Energy Management Systems (EMSs) into social networks.

However, cooperating social network to Smart Grids also opens up a new channel for the attackers to attack the power grid infrastructure. Attackers through social network can use the crowd-shifting strategy [5], [6] as well as misinformation propagation [7], [8], [9] among customers to manipulate their habits of power usage. By misleading a sufficiently large number of users, the attackers are able to create a sudden spike in power load change. Previous works have shown that information, i.e. price rate decreases, can make the new peak in demand up to 60% higher than normal [10], [11]. This significant but unexpected change of loads creates pressure onto the power infrastructure, which, even in the presence of mitigation methods, e.g. load shedding, could cause power components unstable and, in the worst case, lead to a cascading failure. Therefore, in this work, we investigate the disadvantages of integrating social networks into Smart Grid. In particular, we address the following questions: *How dangerous is a misinformation attack on the Social Smart Grid (SSM)? How do we know our power infrastructure is vulnerable to this kind of attack? How will adversaries launch their attacks?*

Assessing the vulnerability of SSM is quite challenging, unfortunately, due to several reasons. First of all, the cascading failure in power grid itself is complicated to model. The integration of social networks brings this complicated behavior to the next level because of their interdependency between social networks and power networks. Failures from one network can be cascaded back and forth between two networks. And finally, it is difficult to capture the behaviors of users reaction to misinformation onto power grids.

In the context of social-power network vulnerability, Pan et al. [12] has studied a problem that given the interdependency between a social network and power network, identify k users in social network to spread misinformation, which would lead to maximum number of failed nodes in power network. Our work extends from that problem in which we consider the interdependence between social network and Smart Grid, an advance extension of power grid with self-protection mechanisms in response to potential failure detection. Furthermore, by formulating the problem of identifying potential attacks in Social-Smart Grid under an optimization problem, we are able to investigate the complex behaviors of the attack based on the complexity theory. Although the problem is proved to be hard to optimize, we propose two heuristic solutions, showing that they can scale up to cripple the large-scale networks. Intuitively, the goal of finding destructive attacks is to assess how robust the Social-Smart Grids is; the more destructive attack we can find, the more vulnerable the system is to this kind of attack. Experimental results has shown the effectiveness of our method, thereby confirming that a fully study on the disadvantages of such social network integration is a must.

Related Work. Previous methods of causing load redistribution have included man-in-the-middle attacks on the Smart Grid communication network, which fabricate electricity price signal to manipulate scheduling settings into causing load spikes [13]. This differs significantly from our work in that the users are receiving false information directly from an authoritative source, rather than propagating second-hand information over a social network. Also, in our problem, instead of fixing number of attacked targets in power network, the misinformation propagation make it more complicate as demand load increases proportionally with number of influenced users in social network. Yuan *et al.* [14] studied optimal attacks on the SCADA control server via manipulating load at individual buses. Further work has included Soltah *et al.* [15], Pan *et al.* [12] and Nguyen *et al.* [16]. Notably, prior work has not considered the role of user behavior in these attacks. Instead the authors assume demand automatically increase when rate reduces. Most of these works assume the absence of self-

protection mechanisms such as load-shedding.

Organization. The rest of the paper is structured as follows. Section II presents how misinformation attack happens in social smart grid and introduces our new optimization model, followed by its inapproximability and non-monotonicity study in section III. Two heuristic solutions to the problem are presented in Section IV. Finally experimental evaluation is detailed in Section V and Section VI concludes our paper.

II. The Misinformation Attack on the Smart Grid

In this section, we describe the models of the social network, Smart Grid and the misinformation attack.

A. Smart Grid

A *Smart Grid* is a communication network built into the power infrastructures to give greater control to both the utility provider and consumer. We model the flow of power in the network with the linearized DC power model, which have been widely used in literature [13], [15], [12]. In the linearized DC model, we are given a power grid represented by a network $G_P = (V_P, E_P)$ with V_P is a set of buses ($|V_P| = n$) and E_P is a set of power line that connecting buses ($|E_P| = m$):

- For each line $e = (i, j) \in E_P$, denote its capacity c_e, its reactance x_e, and its power flow f_e. In some cases, we use f_{ij}, c_{ij}, x_{ij} instead of f_e, c_e, x_e.
- The supply or demand of each bus in V_P on the network is given by the load vector $\beta \in \mathbb{R}^n$, where $\beta_i > 0$ if i is a generator, $\beta_i < 0$ if i is a consumer, otherwise $\beta_i = 0$.
- At each bus $i \in V_P$, the phase angle is given by θ_i. $\theta \in \mathbb{R}^{|V|\times 1}$ is vector of phase angles.

The power flow is given by the system of equations:

$$\mathcal{N} f = \beta \tag{1}$$

$$\mathcal{N}^T \theta - X f = 0 \tag{2}$$

where $\mathcal{N}$ is the incidence matrix of G_P and $X = \text{diag}\{x_e\}$. If $\sum_i \beta_i = 0$, system (1)(2) has a unique solution.

We can rewrite this system as a matrix equation by defining the *admittance matrix* A of the power network:

$$a_{uv} = \begin{cases} 0 & \text{if } u \neq v \text{ and } \{u, v\} \notin E \\ -1/x_{uv} & \text{if } u \neq v \text{ and } \{u, v\} \in E \\ -\sum_{w \in N(u)} a_{uw} & \text{if } u = v \end{cases} \tag{3}$$

Then the matrix equivalent of the power flow system is

$$A\theta = P \tag{4}$$

We denote the Moore-Penrose Pseudo-Inverse matrix (see e.g. [15]) of A as A^+ and use a^+_{uv} to represent its elements.

We next examine a common means by which Smart Grids are protected: load shedding. Load shedding has been shown to be an effective method to preserve the power grid in the presence of cascading line failures [17], [18]. The central idea of load shedding is that when a problem is detected in the power grid, some amount of load will be intentionally "shed" to eliminate the overload in the system, thereby stopping the fault propagation. Zhou Lu et al. [19] introduced two load shedding methods: global and local. Experimental results [19] have shown that local load shedding performs better than global shedding in practical settings. Therefore, in this paper, we focus on local load shedding as our countermeasure when failures are detected. In particular, when a power line $e = (u, v)$ is overloaded ($f_e > c_e$), the loads of node u and v are shed, i.e. $\beta_u = \beta_v = 0$.

Algorithm 1: Cascading Failure

Input : Power Network $G_P = (V_P, E_P)$
while *Network is not stable* **do**
  Balance the supply with the current demand within each connected components of the network.
  Use equations (1) (2) to calculate power flows in E_P.
  If there is edge e such that $f_e > c_e$, shed load connecting to e
  If no load shedding, remove edge e that $f_e > c_e$
  If no line fails and no load shedding, then network is stable, break the loop.
end

Lastly, we describe our model of cascading failure on the power network. Even under load-shedding countermeasures, the failure of one line may cause failures in other lines on the network [19], [20]. Therefore, we model cascading failure as in Alg. 1. To be specific, the local load shedding is implemented when detecting overloaded line. Then, the power flows are recalculated. This process is repeated until the system reaches the state that there is no other nodes to shed (nodes who are incident with overloaded lines but have load 0). Then, we check whether there exists any line $e = (u, v)$ whose $f_e > c_e$, if yes, we disconnect such lines, i.e. $E'_P = E_P \setminus \{e\}$ and recalculation power flow with the new graph $G'_P = (V_P, E'_P)$. All these processes are repeated until there is no more node to shed and no more overloaded lines, the power network is now stable.

B. Social Network

We model the social network as a weighted directed graph $G_S = (V_S, E_S, w)$ with a node set V_S and a directed edge set E_S, where a node $v \in V_S$ represents a user and an edge $(u, v) \in E_S$ exists if v follows u (that is, v can receive information from u). Each (u, v) is associated with a value $w(u, v) \in [0, 1]$ to denote the probability of information propagation. In order to model the information propagation in the social network, we will focus on the widely-applied Independent Cascade (IC) model [21], [22], [23], [24].

In the IC Model, initially no nodes adopt the misinformation. Given a seed set S, the misinformation diffusion proceeds in rounds. In round 0, all nodes $v \in S$ are activated by the misinformation and all other nodes remain inactive. In round $t \geq 1$, all nodes activated at round $t - 1$ will try to activate their neighbors based on the edge weights. A node u activated at round $t-1$ has probability $w(u, v)$ to activate an inactivated neighbor v at time t. u cannot activate any neighbors at any time $t' > t$ and it stays activated till the end. The process stops when no more nodes can be activated.

C. Attack model and Problem definition

In this sub-section, we formulate how misinformation attack happens in social smart grid. Given a social network $G_S = (V_S, E_S)$ and power network $G_P = (V_P, E_P)$, each bus $u_P \in V_P$ is a customer in the power utility service (e.g. a residential house, industrial building) while each node $u_S \in V_S$ is corresponding to a social network user. The interdepency between G_S and G_P is represented as follows:

- Each bus in the power network could be associated to multiple users in social network, which indicates multiple users are allowed to manipulate the load or demand of a bus, e.g. a residential house could have multiple members.
- Each social network user is associated to at most one bus in the power grid. To justify that association, we assume

that even a person can relate to multiple buses (e.g. she owns multiple houses) but at a certain time, she is only capable of manipulating load of at most one bus.
- Each social user u_S, linked with a bus in power network, is also associated with a demand $D(u_S)$. The demand of a power bus in V_P, therefore, is the sum of demand of social users that are linked with it.
- Each social user u_S is also associated with $\mu(u_S)$ - which indicate the change of demand of user u_S when she get influenced by the misinformation, i.e. the demand of u_S becomes $D(u_S) \times \mu(u_S)$ when u_S is influenced.
- For each power bus u_P, when one social user associated with u_P is activated by misinformation, other users associated to u_P are activated as well (i.e. edges connecting those users in G_S have weight 1). This could be justified by the fact that those users tend to have close relationship, e.g. family members. Therefore, the total demand of a power bus is within two values: total demand of associated social users when no one or all are activated by misinformation.

When there is a large amount of users are impacted by misinformation and change their demand simultaneously, the spike of loads pushes pressure on power infrastructure and in the worst case, if overheating happens, the cascading failure occurs as in Alg. 1.

Motivated by this observation, a savvy attacker may exploit the rapid spread of misinformation in the social network to manipulate demand in the power network in order to degrade the power performance. Therefore, studying these attack to have a suitable protection on power systems is a must. A problem is that we have to capture the behavior of power systems and social network when there is misinformation propagation as well as know how attackers will launch their attack. Therefore, we propose an optimization model on which we stand in attacker's perspective to cause maximal damage to power network. We assume the following:

- The objective of the attacker is to damage power system as much as possible. We measure destructiveness of the attack by the number of buses which are deprived of electricity after attack (e.g. get load shed or no connection to power generator because of line disconnection). We call these buses **failed buses** or **failed nodes**.
- Attackers can propagate misinformation from multiple sources, limited by their budget $k \in \mathbb{Z}^+$.

We formally define attacker's objective as follow:

Definition 1. ***M**isinformation **A**ttack on the **S**ocial **S**mart-Grid (MASS): Given smart grid G_P and overlay social network G_S and a budget limit $k \in \mathbb{Z}^+$, find a set of k users in the social network that, when influenced, maximize the number of failed buses in the smart grid.*

In the next sections, we investigate the complexity of MASS problem and propose heuristics to solve the problem.

III. Inapproximability and Non-Monotonicity

Given these models of the social and power networks, we are now ready to analyze the attacker's ability to optimally conduct the MASS attack. Finding the optimal solution for MASS is very important, which could help us devising efficient protection mechanism for the system against the attack. However, it is also a extremely challenging task. First, the behavior of the system in response to such attack is very complicate, which comes from the fact that the objective of MASS is not monotone (Theorem 1). Second, taking theoretical approaches, we show that the optimal objective is inapproximable in polynomial time (assuming $P \neq NP$) within a factor of $O(n^{1-\epsilon})$ (Theorem 2).

Theorem 1 (Non-Monotonicity)**.** *For any $\rho \in (0, \infty)$, there exists a instance of MASS (G_S, G_P) and two set S, T with $S \subset T \subset G_S$, such that the ratio $\frac{\Lambda(S)}{\Lambda(T)} > \rho$ with $\Lambda(S)$ or $\Lambda(T)$ is number of failed nodes in power network G_P if we attack nodes of S or T on social network G_S.*

Proof. For this proof, we need only consider the case where a target set of users S is directly influenced by the attacker. Given the power network $G_P = (V, E)$, there are two generators g_1 and g_2 which can generate infinite supply. We create N user nodes $u_1, u_2, ...u_N$ where the maximum load of each node is $-\beta_{max}(u_i) = 2\epsilon - \delta \ \forall i \in [1, N]$ where $\delta < \epsilon < \frac{1}{2N}$. $\forall i \in [1, N]$, set $c(g_1, u_i) = 2\epsilon$ and $c(g_2, u_i) = \epsilon - \delta$. We then create user node u_0 whose maximum load is $-\beta_{max}(u_0) = 2$ and lines (g_1, u_0) (g_2, u_0) with capacity $c(g_1, u_0) = 1 - \delta$ and $c(g_2, u_0) = 1 + \delta$. Suppose the resistance of each line is 1. At first each user's load is 0.

Consider $S = \{u_0\}$. We raise load of node u_0 to its maximum. By applying (1)-(2), $f(g_1, u_0) = f(g_2, u_0) = 1$ which fails line (g_1, u_0). We continue to recalculate the power flow of each line after (g_1, u_0) fails. This gives $f(g_2, u_0) = 2 > c(g_2, u_0)$, which causes (g_2, u_0) fail and u_0 becomes isolated. Moreover, the power flows $f(g_1, u_i) = f(u_i, g_2) = \frac{1}{N}$ $\forall i \in [1, N]$, which is larger than $c(g_1, u_i) = 2\epsilon < 1/N$ and $c(u_i, g_2) = \epsilon - \delta$. Therefore, all nodes u_i become isolated. Therefore, the total failed buses if we attack S is $\Lambda(S) = N + 1$.

Consider $T = \{u_0, u_1, ...u_N\}$. All user nodes change demand to maximum. By applying (1)-(2), the power flows of edges (g_1, u_0) and (u_0, g_2) are $f(g_1, u_0) = f(g_2, u_0) = 1$, which causes line (g_1, u_0) fail. Meanwhile, $f(g_1, u_i) = f(g_2, u_i) = \epsilon - \delta/2 \ \forall i \in [1, N]$, which is larger than $c(g_2, u_i) = \epsilon - \delta$. Hence, every line (u_i, g_2) is broken. We then recalculate the power flow. Only line (g_2, u_0) fail because $f(g_2, u_0) = 2 > 1 + \delta$. $f(g_1, u_i) = 2\epsilon - \delta < c(g_1, u_i) = 2\epsilon$ $\forall i \in [1, N]$. Hence, only u_0 becomes isolated if we attack T, $\Lambda(T) = 1$.

Therefore, if we choose $N > \rho - 1$ then $\frac{\Lambda(S)}{\Lambda(T)} > \rho$. □

The lack of monotonicity already greatly hinders optimization of MASS. We further show that *even in the absence of load-shedding*, no algorithm can approximate the MASS attack within $O(n^{1-\eta})$.

Theorem 2. *There is no O($n^{1-\eta}$)-approximation algorithm for the MASS attack unless $P = NP$, where n is the number of power demand nodes, for any $0 < \eta < 1$*

Proof. Consider an instance of set cover (U, S, k) where U is the universe, S is collection of subsets of U, and k is the cardinality we are to test. The decision version of set cover, which is NP-complete, asks whether there is a set $C \subset S$ such that $|C| \leq k$ and union of all sets $C_i \in C$ is U. We construct a reduction from MASS to set cover as follows: In the social layer, for each $S_i \in S$ create node v_{S_i} and for each $e_j \in U$, create node v_{e_j}. If $e_j \in S_i$, create edge (v_{S_i}, v_{e_j}) with probability 1, which means if v_{S_i} is activated, v_{e_j} will be affected.

In the power layer, for each $e_j \in U$, create demand node u_{e_j} which is mapped with its corresponding element node in social network. For each u_{e_j} create a generator node g_{e_j} and an edge between them with capacity $c(g_{e_j}, u_{e_j}) = 1 - \epsilon$

and resistance $x(g_{e_j}, u_{e_j}) = 1$. Initially, the load of u_{e_j} is $\beta(u_{e_j}) = 0$. The maximum load $\beta_{max}(u_{e_j}) = 1$. We then create an additional generator node g_0 and edges connecting it to g_{e_j} with resistance r and capacity ∞. Create neutral node n_0 and edges connecting it to each u_{e_j} with resistance r and capacity $1 - \epsilon$. Finally, We create an additional $l > u$ nodes $u_1..u_l$ and connect them to g_0 and n_0 with capacity $u/l - \epsilon$ and resistance r. l will be defined later in the proof. The number of demand node is $n = u + l$. It is possible to choose r large enough such that when $-\beta(u_{e_j}) = \beta(g_{e_j}) = 1$, almost all power flow will cross edge (g_{e_j}, u_{e_j}) and only a negligible amount will use other paths that have lines with capacity r.

Suppose there is a solution to the set cover problem, then every node v_{e_j} in social network would be activated. This lead to every demand node u_{e_j} in power network raising their load to the maximum. We balance supply/demand by letting $\beta(g_{e_j}) = 1$. That makes $f(g_{e_j}, u_{e_j}) > c(g_{e_j}, u_{e_j})$, which cause these lines to fail. We then recalculate the power flow: the in-flow of g_0 would be u. Hence the power flow of each line $f(g_0, u_i) = u/l > c(g_0, u_i)$. Therefore, all demand nodes in power network become failed.

Suppose there is no solution to set cover, which means the maximum number of elements covered is $u - 1$. Similar to previous part, lines corresponding to nodes (g_{e_j}, u_{e_j}) fail where e_j is an element being covered. In this case, the in-flow of g_0 would be at most $u-1$ and there are at least $l+1$ paths to supply power for nodes being covered. Therefore, the power flow $f(g_0, u_i) < \frac{u-1}{l+1} < c(g_0, u_i)$. Moreover, $f(n_0, u_{e_j}) = 1$ where e_j is a node being covered, which causes these lines fail. Therefore, the number of failed node when there is no set cover would be at most $u - 1$.

Let $l > (u-1)n^{1-\eta} - u$ and assume there is $O(n^{1-\eta})$-approximation algorithm $\mathbb{T}$ for MASS problem. If there is set cover solution, the algorithm will find a solution which cause more than $u - 1$ nodes to fail:

$$\Lambda(S_{\mathbb{T}}) \geq \frac{\Lambda(S_{OPT})}{n^{1-\eta}} = \frac{u+l}{n^{1-\eta}} > u - 1$$

where $\Lambda(\cdot)$ is number of failed nodes and $S_{\mathbb{T}}$ is solution of $\mathbb{T}$ and S_{OPT} is optimal solution.

On the other hand, if there is no set cover, the returned solution is at most $u - 1$. Hence, we can use $\mathbb{T}$ to decide the Set Cover problem in polynomial time. □

As the outcome of a MASS attack is both difficult to optimize and predict, we now turn our attention to a pair of reasonable heuristics of the MASS problem to understand how damaging such an attack may be.

IV. Heuristic Attacks

Although we have proposed the self-protection mechanism of smart grid in section II, in reality such systems may not be deployed [13], [15], [12]. Therefore, for each attack solution, we have two versions: one with load-shedding and one without. Interestingly, experiments show that attacks with load shedding in some cases can be more destructive than without protection – a direct result of the non-monotonicity and unpredictability shown in the previous section.

A. *Sequential Node Attack (SENO)*

In this solution, we iteratively identifies which user in social network, who - if being activated - can cause maximum damage to the power system. The impact of a user u_S is estimated by: (1) using forward sampling to identify which users can be influenced by misinformation spreaded by u_S; (2) estimate the impact on power system using Alg 1 when those users increase their demand.

Exact estimation of the impact caused by a user is very expensive, which required us to consider $2^{|E_S|}$ deterministic instances of the social network G_S. Therefore, we reduce the runtime of the estimation by using forward sampling, which similar to the concept of Breadth-First Search as follows: Starting at u_S, we activate a neighbor v_S of u_S with probability $w(u_S, v_S)$. Then with each activated neighbor, we consider activating its not-yet-activated neighbors with corresponding probabilities of their edges. This process is repeated until there is no more activated users. Then, we raise the demand of activated users and recalculate the system's state as in Alg. 1. Each sampling, thus, lets us know how many failed buses that can be caused by activating u_S.

This sampling estimation is repeated in M times and the overall impact of u_S is estimated by taking the mean over all samples. Then the user with highest estimated impact is selected, we then spread the misinformation from that user, observe the misinformation propagation and power system's reaction. After the system is stable, we finds the next user to attack using the same techniques (i.e. sampling, attack and observe) until running out of budget k. The detail of the solution is described in Alg. 2.

Algorithm 2: Sequential Node Attack

Input : $G_S(V_S, E_S), G_P(V_P, E_P), k$ and M
Output: Attack sequence $S = \{u_1, ...u_k\} \subseteq V_S$

$S = \emptyset$
while $|S| < k$ **do**
  for *each* $u_S \in V_S$ **do**
    for M *interations* **do**
      Random sampling the set of activated node $I(u_S)$ by spread misinformation from u_S
      Change demands of nodes in G_P that are associated with nodes in $I(u_S)$
      $\Lambda(u_S) \leftarrow$ number of failed buses in G_S after changing demand (use Alg. 1)
    end
    $\bar{\Lambda}(u_S) \leftarrow$ average number of failed buses causing by u_S in M estimations.
  end
  $v \leftarrow \text{argmax}_{u \in V_S} \bar{\Lambda}(u)$
  $S \leftarrow S \cup \{v\}$
  Spread misinformation from v, observe the activated nodes and change their demand.
  Recalculate power system state using Alg. 1.
end
Return S

The only difference between SENO solution with and without self-protection in Smart Grid is how to estimate the failed buses in each sampling step. In SENO with load shedding, the algorithm finds failed buses as in Alg 1. Meanwhile, the number of failed buses when load shedding is not applied is computed by modifying Alg 1 as follows: when line e is overloaded, e is disconnected and no load is shed.

B. *Sequential Batch Attack Scenario (SEBA)*

In this solution, instead of attacking nodes separately, we targets a batch of critical buses in the power infrastructure who, if increasing demand, could cause enormous damage in the network. The natural question now is *How to identify such set of nodes and how to make sure the budget to cause the associated nodes in social network does not exceed k?.*

Algorithm 3: Sequential Batch Attack

Input : $G_S(V_S, E_S), G_P(V_P, E_P), k$ and M
Output: Attack sequence $S = \{u_1, ...u_k\} \subseteq V_S$

$S = \emptyset$
while $|S| < k$ **do**
  for *each* $u_P \in V_P$ **do**
    $F(u_P) \leftarrow$ the number of failed buses if u_P fails (using Alg. 1)
  end
  $v_P \leftarrow \text{argmax}_{u_P \in V_P} F(u_P)$
  $T \leftarrow$ minimum number of nodes, whose if increasing demand causes failure of v_P. (Equ. 6 if no load shedding or Equ. 9 if there is load shedding)
  $R \leftarrow$ minimal set of social user whose activation causes associated nodes of T influenced (Binary search and WIM)
  $S \leftarrow S \cup R$
  Activate R, observe misinformation propagation and power network reaction until being stable.
end
Return S

With the first question, it is trivial that scanning through all possible set is not a good strategy, which costs us to try $2^{|V_P|}$ sets. We restrict the searching space by observing that there are nodes in power network which are more important than the others, e.g. substation versus residential buses. Failing those nodes can cause failure to other nodes, leading to a huge damage to power system. Therefore, to measure the importance of a power node u_P, we use Alg. 1 to identify number of failed nodes if initially failing u_P. We denote the failed nodes caused by failure of u_P as $F(u_P)$. The attacker then target to fail $v_P = \text{argmax}_{u_P \in V_P} F(u_P)$. To fail the target bus v_P, we devise 2 methods, one is when there is load shedding and one is when there is no load shedding.

If there is no load shedding, two ways to fail v_P are (1) isolating v_P from any generators (in case of demand buses); or (2) disconnecting v_P out of G_P (in case of generators). Suppose L_{v_P} is the minimal set of lines that need to be disconnected to fail v_P.

- With (1), L_{v_P} is found by solving max flow with the pseudo-source s connects to every generator g and the sink is v_P, treating each (s, g) edge as having infinite capacity and each remaining edge with capacity 1.
- With (2), L_{v_P} is simply a set of all edges connecting to v_P. Now, the problem is to identify the minimal set of power buses, whose demand increase will disconnect all edges in L_{v_P}.

We have the following lemma:

Lemma 1. *The change in power flow across edge $e = (i, j) \in E$ when there is load change in bus u is calculated by:*

$$\Delta f_e^u = a_{ij}\Big[\Delta p_u(a_{ju}^+ - a_{iu}^+) + \sum_{v \in V_G} \Delta p_v(a_{iv}^+ - a_{jv}^+)\Big] \quad (5)$$

where Δp_i is the change in load of bus i, Δf_e^u is change in power flow of edge e, V_G is set of generators, and a_{ij}^+ are entries in Moore-Penrose pseudo-inverse of admittance matrix. Note that the load change of u causes the supply change in generators (i.e. Δp_v for all $v \in V_G$) to balance the supply-demand.

Proof. As $\sum_{i \in V_P} \beta_i = 0$ holds, $\Delta p_u + \sum_{v \in V_G} \Delta p_v = 0$ and equations (1)-(2) always have a unique solution. Based on Moore-penrose pseudo-inverse property [15]: if (1)-(2) has a feasible solution, $\hat{\theta} = A^+\beta$ is such a solution. By substituting in the new demand vector, we get $\beta_i' = \beta_i + \Delta p_i$. The remainder follows directly by taking the difference. □

The lemma 1 shows the relation between change in power flow, change in user load, and change in generator supply. By following the rule of demand-supply adjustment given by Soltan et al. [15], we get that $\Delta p_i = (\Delta p_u \times \beta_i)/\Gamma \ \forall i \in V_G$ where $\Gamma = \sum_{v \in V_G} \beta_v$. Then Eqn. (5) becomes:

$$\Delta f_e^u = \Delta p_u \times \left[a_{ij} \times \left(a_{ju}^+ - a_{iu}^+ + \sum_{v \in V_G} \frac{\beta_v(a_{iv}^+ - a_{jv}^+)}{\Gamma}\right)\right]$$

Therefore, the minimal set of power buses (let's call T), whose demand increase can disconnect edges in L_{v_P}, can be identified by solving the following Integer Programming.

$$\min \quad \sum_{u \in V_P} z_u \quad (6)$$

$$\text{s.t.} \quad |f_e + \sum_{u \in V_P} \Delta f_e^u z_u| > c_e \qquad \forall e \in L_{v_P} \quad (7)$$

$$z_u \in \{0, 1\} \qquad \forall u \in V_P \quad (8)$$

where z_u is a variable indicating whether $u \in T$.

If there is a load shedding, v_P fails if at least one line connecting to v_P is overloaded. We call E_{v_P} is set of edges connected to v_P. For each edge e, by using the greedy algorithm and Lemma 1, we can find a set of buses K_e whose load change makes e overloaded. Then, the minimal set of power buses (called T), whose demand increase causes v_P failed, can be obtained by:

$$T = \arg\min_{K_e; e \in N_v} |K_e| \quad (9)$$

With the second question, we directly apply results on the Weighted Influence Maximization (WIM) problem [25], which is defined as: Given a social network G_S, each node $v \in G_S$ is assigned a benefit $b(v)$. Find a set $\mathbb{S} \subset V_S$ of size k such that the total benefit of influenced nodes is maximized.

Therefore, to find a minimal set of users in social network, who - if being activated - can influence associated users of nodes in T, we assign benefit to each nodes u, with $b(u) = 1$ if u is associated to a node in T and 0 otherwise. Let T_S is a set of social users that associate with buses in T. Then, we find a minimum set R of users such that total benefit of activated users by R is at least $|T_S|$ by conducting a binary search on k and solving WIM as a sub-problem. Overall, the SEBA method is described in detail by Alg. 3.

V. Attack Simulation & Experimental Results

In this section, we simulate each attack solution we have described to develop an understanding of the level of risk it poses as well as the complicate behavior of the system in response to each attack. While there are no complete real-world datasets modeling the interdependent social and power networks, we construct reasonable facsimiles by connecting real-world OSN data with real power network topologies. In order to represent our model, we use power systems as the base layer and build the social network on the top of its. Power system datasets are chosen from MATPOWER library [26]. The datasets we use are IEEE 300 buses, Pegase 1354 buses and Polish 3120 bus.

To link the social network with power network, for each demand bus in power network, we randomly link it with a

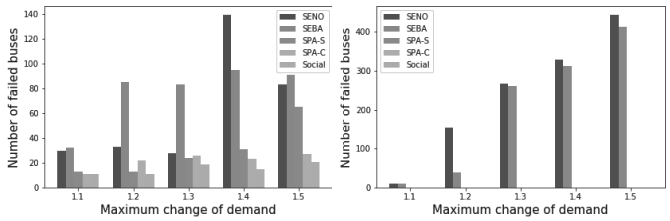
social network and a couple of unlinked neighbors of that node. This linkage is motivated by observation that users who share a same bus tend to have close relationship in social network (e.g. family members tend to be "friends" in Facebook). We assume each user has the same ratio μ of load change when being influenced by the misinformation.

To model information propagation in the social network, we use real-world Facebook network topology from SNAP repository [27], which contains 4039 users. When u and v are associated with the same power bus, the probability u or v influences the other is set to be 1. Otherwise, the probability of user u influences user v, where u and v are friends, is set as $\frac{1}{degree(v)}$ where $degree(v)$ is the in-degree of node v.

We compare with three attack scenarios from [12], which are SPA-C, SPA-S and Social to gain a better understanding of the relative danger these attacks pose. However, we note that these three methods work only when there is no load shedding. The number of estimations on SENO is set to be $M = 1000$. We restrict the runtime to be within 1 hour. Any methods that exceed such limitation with be terminated. This runtime restriction is to simulate the scenario in which the system administrators do not have enough time to react in response to the attack.

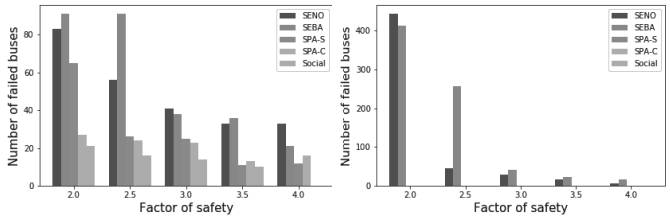
A. Vulnerability of the Social-Smart Grid

In our first experiment, we examine the impact of the change in demand ratio μ on the number of failed buses in the power network. For this test, we set the capacity of each edge e to be $c_e = 2f_e$ where f_e is the stable power flow on e and 2 is the factor of safety α [15], i.e. the factor by which the stable flow of the line must increase in order to fail the line. The number of attacked nodes is 10, 20, and 30 on each of the IEEE 300 bus, Pegase 1354 bus, and Polish 3120 bus topologies, respectively. For a realistic comparison, we test with μ significantly smaller than the factor of safety (from 1.1 to 1.5).



(a) IEEE 300 Bus (b) Polish 3120 Bus

Fig. 1: Impact of each attack with different values of the demand change ratio.



(a) IEEE 300 Bus (b) Polish 3120 Bus

Fig. 2: Impact of each attack with varied values of the factor of safety α.

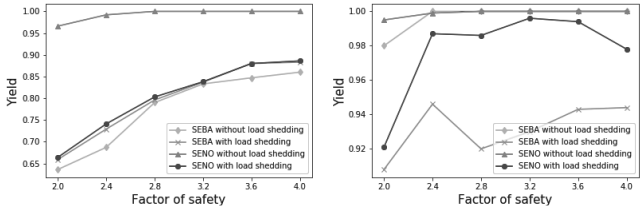
From Fig. 1 and 2 we see that although users' load change is negligible (only 10%), cascading failure and fault propagation in the power network still occurs. Another observation is that: both our proposed methods SENO and SEBA are more efficient than existing methods in term of identifying destructive attacks. Furthermore, they can scale up to run on large-sized power systems such as the Polish 3120-bus dataset, which SPA-S, SPA-C and SOCIAL cannot successfully finish within timing constraint.

Next, we investigate the reduction in efficacy as the factor of safety is increased. We fix the change ratio of demand $\mu = 1.5$ and vary the factor of safety α from 2 to 4. As seen in Fig. 2, even there is a notable drop in performance when we cross to 2μ threshold of 3.0, there still exists node failures with $\alpha = 4.0$, which indicates that we need much more higher α to guarantee the cascading failure does not happen. Therefore, investment in improving the power lines themselves may effectively mitigate the attacks but cause significant expense.

B. Blackout Maximization Under Load Shedding

In this section, we analyze the impact of our two attack scenarios when self-protection via load shedding is applied. Figure 3 shows the yield (ratio between remained load after attack and initial load) as the factor of safety is increased, with μ again fixed at 1.5. In contrast to the cases without load shedding, SEBA outperforms SENO in all cases. To be precise, in experiments on the Polish 3120-bus dataset, the load shedding mechanism seem to be exceptionally effective at protecting power grid under SENO scenario as the power grid loses only 2% of its yield in this case. This can be explained by the fact that SENO attacks each node individually, which is often insufficient to overcome the protection of load-shedding. On the other hand, SEBA attacks in batches, which can overcome the self protection by failing multiple lines simultaneously. Therefore, SEBA should be a more effective method to measure the system vulnerability under attack when there is load shedding.

Under the SEBA attack, the difference in yield between having or not having load shedding is quite small in each case. Notably, the Polish 3120 bus network seems relatively resilient to each attack: the maximum difference in yield between the with- and without-load-shedding is 8%. Oddly, we also see that the SEBA attack in this case is *more effective* with load-shedding. Attackers could take advantage of the fact that load shedding aims to protect *physical infrastructure* rather than keeping high *yield* to maximize destructiveness of their attack, in an other word, deprive people of electricity.



(a) Pegase 1354 Bus (b) Polish 3120 Bus

Fig. 3: Yield with varied values of the factor of safety α

C. Relaxing the Attack Assumptions

We now consider several more difficult attack scenarios. Previously, we had assumed that users who get influenced by misinformation would increase their demand simultaneously.

We now consider the possibility that they instead increase their demand for a period during a discounted interval by assuming that all know about the claimed discount period but that they may not act at the same time. This has the effect of reducing the impact of activating each user. We model this situation as follows: for each influenced user, we randomly select a time interval in which she change her demand. Thus, the overall demand of a bus can receive multiple values rather than only two as stated in Section II. Otherwise, all settings are identical to the above. We call this scenario *No-sim*. In this scenario, we use SEBA as the attack method. Table I shows the result of our attack scenario within this situation. In general:

TABLE I: Percentage of failed buses when relaxing comparing with original results

Method	IEEE 300-bus	Pegase 1354-bus	Polish 3120-bus
No-sim	71.13%	94.91%	103.38%
No-act	45.57%	90.42%	98.79%

- Although people do not react simultaneously, the proposed attack can still cause blackouts.
- The experimental results on the Polish 3120-bus dataset indicate that *more nodes are failed in this scenario* (103.38% of the original number of failed nodes). This highlights the non-monotonicity of the MASS attack shown in our theoretical results.

We next consider the possibility that although a user may be influenced by and propagate the misinformation, *they may not act on it*. We call such scenario *No-act*. This can occur when users share content specifically to inform friends rather than because they intend to make use of it, or by users forgetting to schedule tasks to run during this time. We model this with a probablity p_u that this customer reacts when being activated by misinformation selected uniformly at random for each user. Experiments with $\mathbb{E}[p_u] = 0.5$ are presented in table I. Although on the 300 bus dataset there is a notable drop in performance, in each case a significant number of buses are still failed. On the larger datasets, the level of failure remains remarkably similar to the scenario wherein every activated user is guaranteed to act on the information.

VI. Conclusion

In this paper, we have investigated the negative impacts of integrating a social network into the Smart Grid. Notably, we have shown that a misinformation attack that alters user demand can cause overload in the power network and potentially lead to cascading failure. Thus, finding destructive attacks to the system is of great importance for further protection. Therefore, we model the problem of finding the attack under an optimization problem, whose optimality are very challenging to obtain. Motivated by this observation, we propose two heuristic solutions to the problem, which have shown to be efficient to identify attack strategy with significant damage. Further, this result holds even in the presence of common mitigation techniques (notably: load-shedding). These solutions can be use as efficiently methods to assess the system vulnerability under misinformation attack. Also, this leads us to the conclusion that as Smart Grids are integrated into Social Networks, great care must be taken to preserve the integrity and stability of critical power infrastructure.

VII. Acknowledgements

This work was supported in part by NSF EFRI-1441231, NSF CNS-1814614, and DTRA HDTRA1-14-1-0055.

References

[1] K. C. Chatzidimitriou, K. N. Vavliakis, A. L. Symeonidis, and P. A. Mitkas, "Redefining the market power of small-scale electricity consumers through consumer social networks," in *e-Business Engineering (ICEBE), 2013 IEEE 10th International Conference*. IEEE, 2013.

[2] F. Skopik, "The social smart grid: Dealing with constrained energy resources through social coordination," *Journal of Systems and Software*, vol. 89, pp. 3–18, 2014.

[3] I. G. Ciuciu, R. Meersman, and T. Dillon, "Social network of smart-metered homes and smes for grid-based renewable energy exchange," in *DEST, 2012 6th IEEE International Conference*.

[4] M. Steinheimer, U. Trick, and P. Ruhrig, "Energy communities in smart markets for optimisation of peer-to-peer interconnected smart homes," in *Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2012*. IEEE.

[5] M. LaMonica, "Smart grid needs a dose of social networking," https://www.cnet.com/news/smart-grid-needs-a-dose-of-social-networking/.

[6] R. M. Raafat, N. Chater, and C. Frith, "Herding in humans," *Trends in cognitive sciences*, vol. 13, no. 10, pp. 420–428, 2009.

[7] P. Harper-Slaboszewicz, T. McGregor, and S. Sunderhauf, "Customer view of smart grid—set and forget?" in *Smart Grid*. Elsevier, 2012, pp. 371–395.

[8] C. Y. Ma, D. K. Yau, X. Lou, and N. S. Rao, "Markov game analysis for attack-defense of power networks under possible misinformation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1676–1686, 2013.

[9] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, vol. 7, no. 3, pp. 75–77, 2009.

[10] B. Davito, H. Tai, and R. Uhlaner, "The smart grid and the promise of demand-side management," *McKinsey on Smart Grid*, vol. 3, 2010.

[11] http://www.mpoweruk.com/electricity_demand.htm.

[12] L. N. N. Tianyi Pan, Subhankar Mishra and M. T. Thai, "Threat from being social: Vulnerability analysis of social network coupled smart grid," *IEEE Journal Access*, 2017.

[13] S. Mishra, X. Li, A. Kuhnle, M. T. Thai, and J. Seo, "Rate alteration attacks in smart grid," in *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015, pp. 2353–2361.

[14] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Transactions on Smart Grid*, 2011.

[15] S. Soltan, D. Mazauric, and G. Zussman, "Analysis of failures in power grids," *IEEE Transactions on Control of Network Systems*, 2015.

[16] L. N. Nguyen, J. D. Smith, J. Kang, and M. T. Thai, "Optimal auditing on smart-grid networks," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.

[17] A. Bernstein, D. Bienstock, D. Hay, M. Uzunoglu, and G. Zussman, "Power grid vulnerability to geographically correlated failures—analysis and control implications," in *INFOCOM, 2014 Proceedings IEEE*.

[18] P. Poubeik, P. Kundur, and C. Taylor, "The anatomy of a power grid blackout," *IEEE Power Energy Magazine*, vol. 2006, pp. 22–9, 2006.

[19] X. L. Zhuo Lu, Mingkui Wei, "How they interact? understanding cyber and physical interactions against fault propagation in smart grid," in *Computer Communications (INFOCOM), 2017 IEEE Conference*.

[20] D. Xu and A. A. Girgis, "Optimal load shedding strategy in power systems with distributed generation," in *Power Engineering Society Winter Meeting, 2001. IEEE*, vol. 2. IEEE, 2001, pp. 788–793.

[21] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003, pp. 137–146.

[22] H. Zhang, D. T. Nguyen, H. Zhang, and M. T. Thai, "Least cost influence maximization across multiple social networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 24, no. 2, pp. 929–939, 2016.

[23] D. T. Nguyen, N. P. Nguyen, and M. T. Thai, "Sources of misinformation in online social networks: Who to suspect?" in *Military Communications Conference, 2012-MILCOM 2012*. IEEE, 2012, pp. 1–6.

[24] L. N. Nguyen, K. Zhou, and M. T. Thai, "Influence maximization at community level: A new challenge with non-submodularity," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2019.

[25] H. T. Nguyen, M. T. Thai, and T. N. Dinh, "A billion-scale approximation algorithm for maximizing benefit in viral marketing," *IEEE/ACM Transactions on Networking*, 2017.

[26] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*.

[27] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, Jun. 2014.