# A Game-Theoretic Approach for Power Systems Defense Against Dynamic Cyber-Attacks

Saqib Hasan, *Student Member, IEEE,* Abhishek Dubey, *Senior Member, IEEE,* Gabor Karsai, *Senior Member, IEEE,* and Xenofon Koutsoukos, *Fellow, IEEE*

***Abstract*—Technological advancements in today's electrical grids give rise to new vulnerabilities and increase the potential attack surface for cyber-attacks that can severely affect the resilience of the grid. Cyber-attacks are increasing both in number as well as sophistication and these attacks can be strategically organized in chronological order (dynamic attacks), where they can be instantiated at different time instants. The chronological order of attacks enables us to uncover those attack combinations that can cause severe system damage but remained unexplored due to the non-existent dynamic attack models. Motivated by the idea, we consider a game-theoretic approach to design a new attacker-defender model for power systems. Here, the attacker can strategically identify the chronological order in which the critical substations and their protection assemblies can be attacked in order to maximize the overall system damage. However, the defender can intelligently identify the critical substations to protect such that the system damage can be minimized. We apply the developed algorithms for these models to the IEEE-39 and 57 bus systems based on the attacker/defender budgets. Our results show the effectiveness of these models in improving the system resilience under dynamic attacks.**

***Index Terms*—Cascading failures, Cyber-attack, Dynamic attack, Game theory, Resilience, Smart grid, Static attack.**

## I. Introduction

Recent studies by the National Electric Research Council (NERC) documented that malicious attacks on power grids are much more devastating than the destruction caused by natural calamities [1] and can be instigated through cyber penetration [2] or physical obstruction [3] resulting in large blackouts. Today, power system resilience considering cyber-security has gained significant attention [4] as cyber-attacks are increasing both in number as well as sophistication and are considered as one of the major obstacles towards the reliable system operations [5]–[8]. For instance, due to the technological transformation of the traditional power grids into smart grids, power systems employ a large number of sophisticated and autonomous components such as protection devices, phasor measurement units (PMUs), remote terminal units (RTUs), etc. This increases the potential attack surface by giving rise to new vulnerabilities [9].

The attackers take advantage of such cyber components and gain access to the network by compromising the firewall and can launch catastrophic attacks, compromising system reliability [10] e.g., the recent Ukraine 2015 cyber-attack [11]. What makes the problem worse is the fact that most operators follow the guidelines from NERC [12] requiring only $N-2$ reliability criterion [13], since analysis of higher order contingencies is computationally hard [14], [15], however, a cyber-attack is not limited to only two component failures.

Given such challenges, it is crucial to not only analyze a power system topology for reliability failures but it is also important to analyze the effect of cyber-attacks. In principle this can be approached by considering static attacks, where the devices are affected simultaneously or by dynamically sequenced attacks, which as shown in this paper, can cause significantly higher damage as compared to their static counterparts. Therefore, methods to study dynamic attack are important.

Several frameworks and attack models have been developed to study security vulnerabilities [16]–[26]. A man-in-the middle attack and modeling of cyber-physical switching attacks are presented in [16], [17]. Several data integrity attack studies the effect of manipulating control messages, measurement data in [19], [20]. A special type of false data injection attack, i.e., load redistribution (LR) attack is presented in [21], [22]. The effect of cyber-attack on the voltage stability of support devices is provided in [23]. The work in [24] considers cyber-failures in protection assemblies and provides a platform to obtain new cascading traces. A real-time cyber-physical system testbed that provides mitigation strategies against attacks is discussed in [25]. Additionally, a number of game-theoretic approach based studies exist. For example, an efficient algorithm to solve the defender-attacker-defender problem for system protection is discussed in [27]. In [28], the authors formulate the problem as a minmax non-cooperative game and solved it using genetic algorithm. Moreover, the work in [29] formulates the coordinated attacks on power systems as a bi-level optimization problem. The authors in [30] consider coordinated multi-switch attacks that leads to cascading failures in smart grid. In [31], the authors studied a joint substation-transmission line vulnerability and proposed a component interdependency graph based attack strategy. Based on false data injection attacks, a Markov security game for attacks on automatic generation control is formulated in [32] and a time synchronization based attack is presented in [33]. Further, in [34] the effect of false data injection attacks against state estimation in power grids are studied. Finally, the work in [35], [36] studies the temporal features of attacks in power systems.

However, there are several limitations in these approaches. The frameworks in [16], [17], [25] do not consider a system-wide identification of critical components to compromise. Attack models and strategies referenced in [18]–[23], [27]–[34] focus on simultaneous attacks on different aspects of the system such as opening of circuit breakers, false data injection attacks in monitoring components, etc. However, none of these

approaches consider cyber-attacks from the perspective of time domain, which is a vital facet in cascading failures since the progression of such failures takes at least minutes [37] or at times hours [38]. An attacker can easily and realistically sequence these attacks in a stealthy manner such that the attack mimics the trace of a normal cascading failure that could easily misguide the system operators. Moreover, considering strategically timed cyber-attacks reveal new system vulnerabilities which can not be found using previous approaches and their identification can enhance the overall power system resilience. Further, the attack model in [36] is based on the constructed sequential attack graph (SAG) which can be computationally infeasible for large power networks and most of them do not provide any defense model.

In this paper, we consider a game-theoretic approach to design attacker-defender cyber-attack and defense models for power systems to identify the worst-case dynamic attack. This work proposes a much simpler approach which does not require the construction of complex SAG as required by [36]. Further, we do not choose attacks based on node degree or load which enables us to explore a wider attack area. The specific contributions are:

- A formal dynamic attack model is described, where the attacking cost of any substation and their components is uniform. In this model, the attacker can strategically identify the critical substations and its components to attack at different time instants in order to maximize the system damage constrained by the attacker's budget.
- A formal dynamic defense model is described, where the protection cost of any substation is uniform. In this model, given a defense budget, a defender can strategically identify the most critical substations to prioritize and protect so as to minimize the overall system damage.
- Two efficient polynomial-time algorithms are introduced to identify both the worst-case dynamic attack and a defense strategy which minimizes overall system damage.

Our results (shown using IEEE 39 and 57 bus examples) demonstrate that the approach captures the worst-case dynamic attacks on the power system networks and effectively uses the dynamic defense model to minimize the overall system damage. It also proves the effectiveness and efficiency of our algorithms. Moreover, the attack algorithm is able to maximize the system damage for both static and random attacks.

The remainder of this paper is organized as follows. The system model along with a motivating example is discussed in Section II. Section III and IV give a detailed formal description of the static attack and defense models. The dynamic attack and defense models along with their algorithms are formally presented in Section V and VI. Results are discussed in Section VII followed by the conclusions in Section VIII.

## II. System Model and Motivating Example

We consider a power system $\mathcal{G}_{\mathcal{P}}$, where $U$ is a set of buses, $G$ is a set of generators, $R$ is a set of transmission lines, $L$ is a set of loads, and $P$ is a set of protection assemblies. The power system is divided into substations. Each substation has its own monitoring and control units referred to as RTUs. Let $S = \{S^i\}_{i=1}^m$ be the set of substations. Each substation consists of a set of protection assemblies from $P$. We define $F(S^i)$ as a function that returns the set of protection assemblies in a substation $S^i$. Clearly, the union of all the protection assemblies in every substation represents the set of $P$ in the power network, that is, $\bigcup_{i=1}^{m} F(S^i) = P$. The symbols used have been summarized in Table I. Table II describe the main subroutines referred later in the algorithm sections.

TABLE I: List of Commonly Used Symbols

| Symbol | Description |
|---|---|
| General Symbols | |
| $\mathcal{G}_{\mathcal{P}}$ | power system model |
| $S$ | set of substations |
| $P$ | set of protection assemblies in a power system |
| $S^i$ | $i^{th}$ substation in $S$ |
| $F(S^i)$ | function that returns the set of protection assemblies in substation $S^i$ |
| $B_S$ | substation attack budget |
| $B_P$ | protection assemblies attack budget |
| $B_D$ | substation defense budget |
| $L$ | set of loads in $\mathcal{G}_{\mathcal{P}}$ |
| $L_j$ | $j^{th}$ load in $L$ |
| $I_j$ | current flowing through $j^{th}$ load in $L$ |
| $L_T$ | total power system load |
| Static Attack and Defense Model | |
| $S^{'}$ | set of substations selected from $S$ for static attack |
| $P^{'}$ | set of protection assemblies selected from $P^{'}$ for static attack |
| $A_{P'}$ | static attack on substations $S^{'}$ and protection assemblies $P^{'}$ |
| $D_P$ | defense strategy for static attacks |
| $D_S$ | set of protected substations |
| Dynamic Attack and Defense Model | |
| $k$ | time instant in $\{1, \ldots, T\}$ |
| $S^{'}(k)$ | set of substations selected from $S$ for dynamic attack |
| $P^{'}(k)$ | set of protection assemblies selected from $P$ for dynamic attack |
| $A_{P'}(k)$ | dynamic attack on substations $S^{'}(k)$ and protection assemblies $P^{'}(k)$ |
| $x(k)$ | state of the system at $k^{th}$ time instant |
| $H(k)$ | attack history of the system $\mathcal{G}_{\mathcal{P}}$ |
| $G(H(k))$ | function representing the power system state under the presence of attack history $H(k)$ at time step $k$ |
| $g(H(k))$ | function representing nominal system state with no attack history |
| $D_S$ | set of protected substations |

Let us consider an IEEE-14 bus system [39] to demonstrate the concept of static and dynamic attack. The system is divided into substations shown by blue dotted rectangles labeled as $S^n$, where $n \in \mathbb{N}$. The protection assemblies within the substations are labeled as $PAn$. The transmission lines labeled as 'Rn_m' can be isolated by manipulating the protection assemblies associated with it. Now consider the static attack scenario where the protection assemblies associated with the transmission lines 'R6_13' and 'R7_8' are manipulated to isolate them from the power network simultaneously. This led to removal of lines 'R9_14', 'R6_12', 'R9_10', 'R12_13' and loads 'L 5, L9, L4, and L7' from the power network due to subsequent system overloading. Now, in case of dynamic attack, only transmission line 'R6_13' is isolated initially which causes

TABLE II: List of Methods

| Method Name | Use |
|---|---|
| `Gen_Contin`$(S_{info}, \hat{P}_a)$ | Returns the set of contingencies based on the protection assemblies in $S_{info}$, and $\hat{P}_a$ |
| `Simulate_Model`$(\mathcal{G}_\mathcal{P})$ | Simulates the nominal state of the power system model $\mathcal{G}_\mathcal{P}$ |
| `Isolate_Branches`$(\mathcal{G}_\mathcal{P}, p)$ | Removes branch(es) from the power system model $\mathcal{G}_\mathcal{P}$ associated with the attacked protection assemblies $p$ |
| `Simulate_Contin`$(\mathcal{G}_\mathcal{P}, p, k)$ | Simulates the power system model $\mathcal{G}_\mathcal{P}$ with branch(es) removal at specific time instants $k$ |
| `Get_Branches`$(\mathcal{G}_\mathcal{P}, p)$ | Returns the overloaded branches in the power system model $\mathcal{G}_\mathcal{P}$ post attack |
| `Get_Loads`$(\mathcal{G}_\mathcal{P}, p, k)$ | Returns the load names $l$ that are disconnected in the power system model $\mathcal{G}_\mathcal{P}$ post attack |
| `Get_Damage`$(\mathcal{G}_\mathcal{P}, l)$ | Returns the overall damage in the power system model $\mathcal{G}_\mathcal{P}$ post attack |
| `Obtain_Subs`$(S_{info}, p)$ | Returns the substation(s) corresponding to the attacked protection assemblies $p$ in the power system model $\mathcal{G}_\mathcal{P}$ |

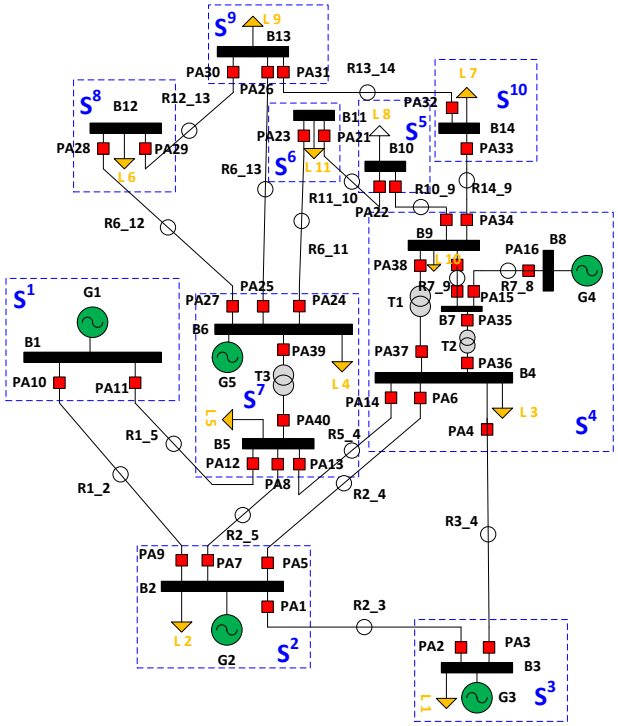



Fig. 1: IEEE-14 Bus System [39]

a cascading failure in lines 'R12_13', 'R9_14' and 'R6_12' due to overloading of these lines. The transmission lines are isolated and at this time another attack is executed, i.e., transmission line 'R7_8' is isolated which results in the outage of lines 'R10_11', and 'R9_10' in the subsequent cascading stage. Post dynamic attack, the system lost a total of five loads namely 'L 5, L8, L9, L4, and L 7' as opposed to 'L 5, L9, L4, and L7' in the static attack scenario. This is obviously a higher damage as compared to the static attack considering the same attacks are executed with a difference in the attack execution time and provides the motivation to the problem.

Before, we dive into the details of dynamic attack and defense models, it is important to understand the problem from the static attack perspective. Therefore, we will first explain the static attack and defense models in detail to give the reader a better understanding. Results from Static attack and defense models are demonstrated in [40] and we build our dynamic attack and defense models on top of it.

## III. Static Attack Model

In this section, we first formulate the static attack model that aims to maximize the load loss in the power system network. Then, we provide an efficient algorithm to identify the worst-case static attack.

### A. Worst-Case Static Attack

The objective of a malevolent attacker is to maximize the load loss and destabilize the power network. To achieve this, first the attacker may gain access to a subset of substations $S^{'} \subseteq S$ where the attacker is resource bounded, i.e., the attacker can compromise at most $B_S$ substations. Now, the adversary can identify the protection assemblies $P^{'} \subseteq F(S^{'})$ to manipulate in order to isolate the transmission lines from the power network where the protection assemblies belong to the selected substations $S^{'}$. The attacker is again resource bounded and can attack at most $B_P$ protection assemblies. Note that, a naive attacker may select a large $B_P$ and probably attack all the protection assemblies within the compromised substations, whereas, a strategic attacker may favor a small $B_P$ as it would enable the attacker to remain undetected for a considerably longer period of time that could provide the attacker with an opportunity to potentially cause higher system damage.

Additionally, note that transmission lines are rated to carry a maximum amount of power and are isolated from the rest of the system in case of limit violations. This action often results in cascading failures causing severe load loss. Manipulating all the protection assemblies of a substation to disconnect power lines may reduce the overall system load. Hence, this may not lead to severe cascading failures causing higher load loss. Next, the attack on a set of substations $S^{'}$ and protection assemblies $P^{'}$ is denoted by $A_{P^{'}}$.

Let $L_j$ denote the $j^{th}$ load in the power network $\mathcal{G}_\mathcal{P}$. The current flowing through each load $L_j$ is given by $I_j$, where $j = 1$ to $n$. Now, we compute the damage function as below:

$$J(A_{P^{'}}) = \frac{\sum_{j=1}^{n} L_j}{L_T} \times 100, \quad \forall I_j = 0 \tag{1}$$

where $L_T$ represents the total system load. The problem is formally defined below.

*Problem 1 (Worst-Case Static Attack):* Given a power system network $\mathcal{G}_\mathcal{P}$, a substation budget $B_S$, and a protection assembly budget $B_P$, find a worst-case static attack $A_{P^{'}}$ that maximizes the damage in the power system network. Formally,

$$\underset{S^{'}}{\operatorname{argmax}} \max_{P^{'} \subseteq F(S^{'})} J(A_{P^{'}}) \tag{2}$$

$$\begin{aligned} |S^{'}| &\leq B_S \\ \forall S^{'}, S^{''} \in S : S^{'} \cap S^{''} &= \emptyset \end{aligned} \tag{3}$$

$$|P^{'}| \leq B_P$$
$$\forall P^{'}, P^{''} \in P : P^{'} \cap P^{''} = \emptyset \tag{4}$$

$$B_S \leq B_P \tag{5}$$

### B. Algorithm for Finding Worst-Case Static Attack

This section describes the algorithm for finding the worst-case static attack in detail.

*1)* `Get_WSA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info})$*:* Algorithm 1 is based on iteratively identifying attacks that maximize the system damage depending upon the budget constraints, i.e., $B_S$ and $B_P$. Here, the algorithm intelligently selects the protection assemblies to manipulate one-by-one that maximizes power system damage and maps it back to their respective substations. This approach reduces the overall run time of the algorithm. It takes as inputs the power system model $\mathcal{G}_\mathcal{P}$, protection assemblies budget $B_P$, and power system substation configuration information $S_{info}$. Further, it identifies the worst-case static attack by identifying a set of critical substations to compromise $S^{'}$, the protection assemblies to manipulate $P^{'}$ and the damage caused by the attack $L_w$.

**Algorithm 1** Algorithm for Finding Worst-Case Static Attack: `Get_WSA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info})$

**Input:** $\mathcal{G}_\mathcal{P}, B_P, S_{info}$
**Initialize:** $L_w \leftarrow 0, P^{'} \leftarrow \emptyset, S^{'} \leftarrow \emptyset, L_g \leftarrow 0$
$P_t \leftarrow F(s)$
$\hat{P}, L_P \leftarrow$ `Get_Static_Attack`$(\mathcal{G}_\mathcal{P}, P_t)$
$L_w \leftarrow L_P, P^{'} \leftarrow \hat{P}$
**for** $k = 2, \ldots, B_P$ **do**
  $\hat{P}_t \leftarrow$ `Get_Contin`$(S_{info}, \hat{P})$
  $\hat{P}, L_P \leftarrow$ `Get_Static_Attack`$(\mathcal{G}_\mathcal{P}, \hat{P}_t)$
  **if** $L_P > L_w$ **then**
    $L_w \leftarrow L_P$, $P^{'} \leftarrow \hat{P}$
  **end if**
  **if** $(L_g - L_w) \leq \varepsilon$ **then**
    **break**
  **else**
    $L_g \leftarrow L_w$
  **end if**
**end for**
$S^{'} \leftarrow$ `Obtain_subs`$(S_{info}, P^{'})$
**return** $S^{'}, P^{'}, L_w$

As a first step, the algorithm identifies the maximum damage causing protection assemblies that can be manipulated from the entire set of protection assemblies using the method `Get_Static_Attack`$(\mathcal{G}_\mathcal{P}, P_t)$. The set of all protection assemblies can be obtained by using the function $F(S)$. Further, for every following iteration, the algorithm identifies the new set of attackable protection assemblies. For instance, let $S_{info}$ be the set that represents the information about the substations and its protection assemblies. If an attacker has attacked a protection assembly $\hat{P}$ from the set of substations $S_{info}$ then in the next iteration, `Get_Contin`$(S_{info}, \hat{P})$ uses the $\hat{P}$ to return a new attackable set of protection assemblies such that the attacker can choose only one new protection assembly from the total number of protection assemblies $P$ in $S_{info}$ that has not been previously attacked. Similarly, in each iteration the algorithm selects the protection assemblies $P^{'}$ to manipulate from the attackable set of protection assemblies that are part of the selected $S^{'}$ in order to isolate transmission lines from the power network. Here, the function `Get_Static_Attack`$(\mathcal{G}_\mathcal{P}, \hat{P}_t)$ identifies the protection assemblies that cause maximum damage and updates the solution if the damage $L_P$ caused by the selected protection assemblies is greater than the worst-case static damage $L_w$, where $\hat{P}_t$ represents the set of protection assemblies that are available for the attack. The function `Get_Static_Attack`$(\mathcal{G}_\mathcal{P}, \hat{P}_t)$ is similar to Algorithm 4, however, it does not consider the time vector for scheduling attacks. The algorithm terminates if no further improvement in system damage is observed. At the end, the substations $S^{'}$ that should be compromised in order to maximize system damage corresponding to the attacked protection assemblies are identified through direct mapping using the method `Obtain_subs`$(S_{info}, P^{'})$. The worst-case running time of Algorithm 1 is non-exponential and is given by $O(|P| \times |B_P|)$.

## IV. Static Defense Model

In this section, first we provide the formulation of the defender model to improve the power system resilience by minimizing the damage/load loss. Then, we provide an efficient algorithm for identifying the critical substations to be protected in order to minimize the system damage considering the static attack model. Here, based on the substations and their components i.e. protection assemblies targeted by the attack, a set of critical substations to be protected is identified.

### A. Defender's Problem

The primary goal of a defender is to improve the power system resilience by protecting the critical substations in order to minimize the possible load loss when an attack is launched. To achieve this, the defender can protect a subset of substations $D_S$ from the total number of substations $S$, i.e., $D_S \subseteq S$. The defender is resource bounded and it can prioritize and protect up to $B_D$ substations due to financial budget constraints because it is impossible to protect and upgrade all the substations simultaneously. Also, a strategic attacker would aim at maximizing the system damage by attacking the most critical substations. Hence, this model can provide important insight upon which substations can be prioritized for the upgrade and protected first against the malicious adversarial attack. The problem is formally described below.

*Problem 2 (Defender's Problem):* Given a power system network $\mathcal{G}_\mathcal{P}$, a defense budget $B_D$, a substation budget $B_S$, a protection assembly budget $B_P$, find a defense strategy $D_P$ to minimize the system load loss. Formally,

$$\underset{D_S}{\operatorname{argmin}} \max_{S^{'} \subseteq S \setminus D_S} \max_{P^{'} \subseteq F(S^{'})} J(A_{P^{'}}) \tag{6}$$

$$|D_S| \leq B_D \tag{7}$$

$$|S^{'}| \leq B_S$$
$$\forall S^{'}, S^{''} \in S : S^{'} \cap S^{''} = \emptyset \tag{8}$$

$$|P^{'}| \leq B_P$$
$$\forall P^{'}, P^{''} \in P : P^{'} \cap P^{''} = \emptyset \tag{9}$$

$$B_S \leq B_P \tag{10}$$

### B. Algorithm for Finding the Critical Substations to Protect

Algorithm 2 starts with an empty set and strategically identifies the critical substations to protect one-by-one such that when an attacker launches an attack the overall system damage can be minimized. The algorithm takes the same inputs as algorithm 1 with the defense budget $B_D$ as an additional input. It then identifies the critical substations $D_S$ to prioritize and protect to minimize the system damage when a static attack is launched.

**Algorithm 2** Algorithm to Find Critical Substations to Protect: $\texttt{Get\_Static\_Defense}(\mathcal{G}_\mathcal{P}, B_P, B_D, S_{info})$

**Input:** $\mathcal{G}_\mathcal{P}, B_P, B_D, S_{info}$
**Initialize:** $D_s^{'} \leftarrow \emptyset, D_S \leftarrow \emptyset, D_S^t \leftarrow \emptyset, \hat{L}_w \leftarrow 100, L_{Prev} \leftarrow 100, L_H \leftarrow \emptyset$
$S_t^{'}, P^{'}, L_w \leftarrow \texttt{Get\_WSA}(\mathcal{G}_\mathcal{P}, B_P, S_{info})$
$L_H \cup L_{Prev}$
**for** $i = 1, \ldots, B_D$ **do**
  $\hat{L}_w \leftarrow 100, flag \leftarrow 0$
  **if** $D_S^t \neq \emptyset$ **then**
    $S_t^{'}, L_{Prev} \leftarrow \texttt{Get\_WSA1}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, D_S^t, \emptyset)$
    $L_H \cup L_{Prev}$
  **end if**
  **for all** $s \in S_t^{'}$ **do**
    $S_s^{'}, P_s^{'}, L_s^{'} \leftarrow \texttt{Get\_WSA2}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, D_S^t, s)$
    **if** $L_s^{'} < \hat{L}_w$ **then**
      $\hat{L}_w \leftarrow L_s^{'}, D_s^{'} \leftarrow s, flag \leftarrow 1$
    **end if**
  **end for**
  $D_S \leftarrow D_S \cup D_s^{'}, D_S^t \leftarrow D_S^t \cup D_s^{'}$
  **if** $\hat{L}_w > min(L_H)$ AND $flag == 1$ **then**
    $D_S \leftarrow D_S \setminus D_s^{'}$
  **else**
    $D_S \leftarrow D_S^t$
  **end if**
**end for**
**return** $D_S$

First, the worst-case static attack is identified using $\texttt{Get\_WSA}(\mathcal{G}_\mathcal{P}, B_P, S_{info})$ which is explained in Algorithm 1. Next, for the first iteration when there are no critical substations in $D_S^t$ to protect, we use the critical substations $S_t^{'}$ identified from the worst-case attack to identify the first substation to protect. $D_S^t$ represents the intermediate solution set for substations to be protected in order to obtain a better solution. We iteratively protect each substation in $S_t^{'}$ and evaluate the overall system damage post static attack using $\texttt{Get\_WSA2}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, D_S^t, s)$. The computed system damage in each iteration is used to select the substation to protect, i.e., $D_S \leftarrow D_S \cup D_s^{'}$, $D_S^t \leftarrow D_S^t \cup D_s^{'}$, where $D_s^{'}$ is the substation that is to be protected and is obtained in the $i^{th}$ iteration. Note that, the function $\texttt{Get\_WSA2}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, D_S^t, s)$ is same as Algorithm 1, however, here the worst-case static attack is computed by eliminating the protected substations $D_S^t$ and the substation $s$ from the attackable list of substations, i.e., $S \setminus (D_S^t \cup s)$. If the computed damage $L_s^{'}$ is smaller than the maximum damage $\hat{L}_w$, the solution is updated. Additionally, in each iteration, if the protected substations set $D_S^t$ is non-empty then a new set of critical substations are identified using worst-case static attack function, i.e., $\texttt{Get\_WSA1}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, D_S^t, \emptyset)$. This function is also same as Algorithm 1, however, the protected substations $D_S^t$ are removed from the attackable list of substations while executing the worst-case static attack on the power system model $\mathcal{G}_\mathcal{P}$. It ensures that once the substations are protected, the attacker can only launch the static attack on the remaining substations depending on the attack budget. The obtained attack can further be utilized to identify the substation to protect considering the defense budget constraints. In the algorithm $L_H$ keeps a track of all the previous load losses obtained after protecting the substations in $D_S^t$ and updates the final solution $D_S$ depending upon the comparison of the obtained damage with the previous system damages. This ensures a better protection mechanism that provides an effective solution. The worst-case running time of Algorithm 2 is non-exponential and is given by $O(|S| \times |B_D| \times |P| \times |B_P|)$.

## V. Dynamic Attack Model

In this section, we first formulate the dynamic attack model then we provide an efficient algorithm for identifying the worst-case dynamic attack that maximizes the system damage.

### A. Worst-Case Dynamic Attack

The objective of the malicious attacker is to destabilize the power system by maximizing the load loss. In order to achieve this, first the attacker can gain access to a subset of substations $S^{'}(k) \subseteq S$ at different time instants $k$, where $k \in \{1, \ldots, T\}$. The attacker is resource bounded and can compromise up to $B_S$ substations. Next, the adversary can identify the protection assemblies $P^{'}(k) \subseteq F(S^{'}(k))$ to manipulate within the selected substations in order to disconnect transmission lines from the power system network at different time instants $k$. Here, the attacker is again resource bounded, i.e., it can manipulate at most $B_P$ protection assemblies. Finally, the dynamic attack on a set of substations $S^{'}$ and protection assemblies $P^{'}$ at time step $k$ is denoted by $A_{P^{'}}(k)$. Now, we compute the dynamic attack damage function as below:

$$J(A_{P^{'}}(k), x(k)) = \frac{\sum_{j=1}^{n} L_j(k)}{L_T} \times 100, \quad \forall I_j(k) = 0 \quad (11)$$

where $k \in \{1, \ldots, T\}$, $x(k)$, and $A_{P^{'}}(k)$ represents the time step, system state, and the attack at time step $k$ respectively. The problem is formally defined below.

*Problem 3 (Worst-Case Dynamic Attack):* Given a power system network $\mathcal{G}_\mathcal{P}$, a substation budget $B_S$, and a protection assembly budget $B_P$, find a worst-case dynamic attack $A_{P^{'}}(k)$ that maximizes the system damage. Formally,

$$\underset{\{S^{'}(k)\}_{k=1}^{T}}{\operatorname{argmax}} \; \max_{(\{P^{'}(k) \subseteq F(S^{'}(k))\}_{k=1}^{T})} \sum_{k=1}^{T} J(A_{P^{'}}(k), x(k)) \quad (12)$$

$$x(k) = \begin{cases} G(H(k)), & \text{if} \quad H(k) = \{A_{P^{'}}(i)\}_{i=1}^{k-1} \\ g(H(k)), & \text{if} \quad H(k) = \emptyset \end{cases} \quad (13)$$

$$\sum_{k=1}^{T} |S^{'}(k)| \leq B_S$$
$$\forall k, k^{'} \in \{1, \ldots\ldots, T\} : S^{'}(k) \cap S^{'}(k^{'}) = \emptyset, k \neq k^{'} \quad (14)$$

$$\sum_{k=1}^{T} |P^{'}(k)| \leq B_P \tag{15}$$

$$\forall k, k' \in \{1, ....., T\} : P^{'}(k) \cap P^{'}(k') = \emptyset, k \neq k'$$

$$B_S \leq B_P \tag{16}$$

where, $x(k)$ represents the state of the system at time step $k$ and $H(k)$ represents the attack history of the system.

### *B. Algorithm for Finding Worst-Case Dynamic Attack*

This section describes the algorithm for finding the worst-case dynamic attack in detail.

*1)* $\texttt{Get\_WDA}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k)$*:* Algorithm 3 is based on iteratively identifying the attacks at specific instants in time depending upon the budget constraints, i.e., $B_S$ and $B_P$. Here, $S_{info}$ denotes power system substation configuration, $a_k$ denotes the possible attack time vector, $L_w^d$ represents the worst-case dynamic damage and $a_k^d$ represents the identified time vector at which the attack needs to be executed.

First, we use $\texttt{Get\_WSA}(\mathcal{G}_\mathcal{P}, S_{info}, B_P)$ to identify the worst-case static attack described as Algorithm 1 in Section III. Here, we identify the maximum damage causing attack that provides the substations to compromise $S^{'}$, and the protection assemblies $P^{'}$ within the substations to manipulate in order to isolate the transmission lines from the power network assuming the attacks take place at the same time. The set of $P^{'}$ is iteratively used to generate a new set of contingencies $C$ using $\texttt{Gen\_Contin}(P^{'}, P^d)$. The contingencies $C$ are used by $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a_{temp}^d, a_k)$ (explained as Algorithm 4) which returns the maximum damage $L^*$ causing attack consisting of substations and associated protection assemblies $P^*$ and the attack time vector $a^*$. In each iteration one attack is intelligently identified along with its time instant vector $a_{temp}^d$ and added to the solution. Note that during the contingency generation process, $P^*$ is utilized in such a way that the search space remain much smaller than the exhaustive search but still effective. In each iteration, if the maximum damage $L^*$ obtained from $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a_{temp}^d, a_k)$ is larger than the worst-case dynamic damage $L_w^d$ then the solution is updated. At the end, the method $\texttt{Obtain\_subs}(S^{'}, P^{'}(k))$ is used to obtain the direct mapping of the substations to be attacked. This is possible because the corresponding protection assemblies belong to the respective substations. This process reduces algorithm run time and provides effective solution.

*2)* $\texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a_{temp}^d, a_k)$*:* Given a set of contingencies, Algorithm 4 identifies the protection assemblies one-by-one and the best sequence in which the attack can be executed to maximize the power system damage. Here, $a_{temp}^d$ represents the attack time vector of set of contingencies in $C$. Note that, the attack vector $a_{temp}^d$ of any contingency $C(i,j)$ represents the time instants of the previously attacked protection assemblies in $C(i,j)$. Since protection assemblies are identified one-by-one and added to the solution, the maximum damage causing protection assembly that needs to be identified in any iteration will have an empty time instant ([]) in $C(i,j)$ before the algorithm is executed. Further, for any iteration in Algorithm 3, Algorithm 4 computes the maximum damage causing attack identifying the set of protection assemblies $P^*$ to manipulate within the identified substations $S^{'}$, the system damage $L^*$ caused by the attack, and the time instants $a^*$ at which the attacks need to be executed.

**Algorithm 3** Algorithm for Finding Worst-Case Dynamic Attack: $\texttt{Get\_WDA}(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k)$

**Input:** $\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k$
**Initialize:** $L_w^d \leftarrow 0, P^{'}(k) \leftarrow \emptyset, S^{'}(k) \leftarrow \emptyset, a_k^d \leftarrow \emptyset, a_k^{'} \leftarrow 0$
$S^{'}, P^{'}, L_w \leftarrow \texttt{Get\_WSA}(\mathcal{G}_\mathcal{P}, S_{info}, B_P)$
$S^{'}(k) \leftarrow S^{'}, P^{'}(k) \leftarrow P^{'}, L_w^d \leftarrow L_w$
**for all** $p \in P^{'}$ **do**
  $a_k^d \leftarrow a_k^d \cup a_k^{'}$
**end for**
**for all** $p \in P^{'}$ **do**
  $P^d \leftarrow \emptyset, a^d \leftarrow a_k^{'}, a_{temp}^d \leftarrow a^d$
  $P^d \leftarrow P^d \cup p$
  **for** $i = 1, \ldots, (|P^{'}|)$ **do**
    $C \leftarrow \texttt{Gen\_Contin}(P^{'}, P^d)$
    $P^*, L^*, a^* \leftarrow \texttt{Get\_Dynamic\_Attack}(\mathcal{G}_\mathcal{P}, C, a_{temp}^d, a_k)$
    $P^d \leftarrow P^*, a_{temp}^d \leftarrow a^*$
    **if** $L^* \geq L_w^d$ **then**
      $L_w^d \leftarrow L^*, P^{'}(k) \leftarrow P^*, a_k^d \leftarrow a^*$
    **end if**
  **end for**
**end for**
$S^{'}(k) \leftarrow \texttt{Obtain\_subs}(S^{'}, P^{'}(k))$
**return** $S^{'}(k), P^{'}(k), L_w^d, a_k^d$

For each contingency, the algorithm first simulates the power system model in its nominal state, i.e., without any attack. Then, depending upon a contingency $C(i,j)$ and the attack vector $a_{temp}^d$, all the transmission lines associated with $C(i,j)$ are removed from the power network for which the time instants are '0', i.e., initial attack. The power system $\mathcal{G}_\mathcal{P}$ is then simulated with the initial attack and is further evaluated for the secondary effects of this attack, i.e., additional system overloads. If there are any overloaded transmission lines they are identified and removed from the power network. Additionally, if there are any other attacks in $C(i,j)$ that are available to be executed using the attack vector $a_{temp}^d$ at any other time instants they are also identified and executed. Next, the algorithm uses the time instant vector $a_k$ to manipulate the protection assembly with empty time instant to isolate the associated transmission line such that it maximizes the system damage. The power system model is then simulated with the contingencies $(P_{C(i,j)} \cup P_{a_k})$ and its associated attack vector $(a_{C(i,j)} \cup a_k^*)$. Next, the amount of system damage caused by the attack is computed for every contingency set in $C$. If the computed load $L_C$ is larger than the maximum damage $L^*$ in any iteration, the solution is updated. Note that, after evaluating each contingency set in $C$, the power system model is set back to its nominal state.

## VI. Dynamic Defense Model

In this section, first we provide the formulation of the defender model then we provide an efficient algorithm for identifying the critical substations to be protected to minimize the system damage.

**Algorithm 4** Algorithm for Finding Dynamic Attack: `Get_Dynamic_Attack`$(\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k)$

**Input:** $\mathcal{G}_\mathcal{P}, C, a^d_{temp}, a_k$
**Initialize:** $L^* \leftarrow 0, P^* \leftarrow \emptyset, a^* \leftarrow \emptyset, a^*_k \leftarrow \emptyset, P_{a_k} \leftarrow \emptyset$
**for** $i = 1, \ldots, |C|$ **do**
  `Simulate_Model`$(\mathcal{G}_\mathcal{P})$
  **for** $k = 1, \ldots, |a_k|$ **do**
    $P_{C(i,j)} \leftarrow \emptyset, k_c \leftarrow 0, a_{C(i,j)} \leftarrow \emptyset$
    **for** $j = 1, \ldots, |a^d_{temp}|$ **do**
      **if** $a^d_{temp}(j) = 0$ **then**
        `Isolate_Branches`$(\mathcal{G}_\mathcal{P}, C(i,j))$
        $a_{C(i,j)} \leftarrow a_{C(i,j)} \cup a^d_{temp}(j)$
        $P_{C(i,j)} \leftarrow P_{C(i,j)} \cup C(i,j)$
      **end if**
    **end for**
    `Simulate_Contin`$(\mathcal{G}_\mathcal{P}, P_{C(i,j)}, a_{C(i,j)})$
    $e \leftarrow 1$
    **while** $e = 1$ **do**
      $e \leftarrow 0, k_c \leftarrow k_c + 1$
      $c \leftarrow$ `Get_Branches`$(\mathcal{G}_\mathcal{P}, C(i,j))$
      **if** $|c| \neq 0$ **then**
        **for** $y = 1, \ldots, |c|$ **do**
          `Isolate_Branches`$(\mathcal{G}_\mathcal{P}, c(y))$
        **end for**
        $e \leftarrow 1$
      **end if**
      **for** $j = 1, \ldots, |a^d_{temp}|$ **do**
        **if** $k_c = a^d_{temp}(j)$ **then**
          `Isolate_Branches`$(\mathcal{G}_\mathcal{P}, C(i,j))$
          $P_{C(i,j)} \leftarrow P_{C(i,j)} \cup C(i,j)$
          $a_{C(i,j)} \leftarrow a_{C(i,j)} \cup a^d_{temp}(j)$
        **end if**
      **end for**
      **if** $k_c = a_k(k)$ **then**
        `Isolate_Branches`$(\mathcal{G}_\mathcal{P}, C(i, |C(i)| - 1))$
        $P_{a_k} \leftarrow C(i, |C(i)| - 1), a^*_k \leftarrow k_c$
      **end if**
      `Simulate_Contin`$(\mathcal{G}_\mathcal{P}, P_{C(i,j)} \cup P_{a_k}, a_{C(i,j)} \cup a^*_k)$
      $L_l \leftarrow$ `Get_Loads`$(\mathcal{G}_\mathcal{P}, P_{C(i,j)} \cup P_{a_k}, a_{C(i,j)} \cup a^*_k)$
      $L_C \leftarrow$ `Get_Damage`$(\mathcal{G}_\mathcal{P}, L_l)$
    **end while**
    **if** $L_C > L^*$ **then**
      $L^* \leftarrow L_C, P_t \leftarrow P_{C(i,j)}, P_i \leftarrow P_{a_k}$
      $a^{'}_C \leftarrow a^*_k, a_C \leftarrow a_{C(i,j)}$
    **end if**
    `Simulate_Model`$(\mathcal{G}_\mathcal{P})$
  **end for**
**end for**
$P^* \leftarrow P_t, a^* \leftarrow a_C$
$P^* \leftarrow P^* \cup P_i, a^* \leftarrow a^* \cup a^{'}_C$
**return** $P^*, L^*, a^*$

### A. Defender's Problem

The objective of the defender is to improve the power system resilience by minimizing the load loss possible. In order to achieve this, defender can protect a subset of substations $D_S$ from the total number of substations $S$ in the power system network, i.e., $D_S \subseteq S$. Due to financial budget constraints, the defender is resource bounded and can prioritize and protect at most $B_D$ substations. The problem is formally defined below.

*Problem 4 (Defender's Problem):* Given a power system network $\mathcal{G}_\mathcal{P}$, a defense budget $B_D$, a substation budget $B_S$, a protection assembly budget $B_P$, find a defense strategy $D^{'}_P$ to minimize the damage/load loss when an attacker launches a dynamic attack at different time instants $k$. Formally,

$$\underset{D_S}{\operatorname{argmin}} \max_{\{(S'(k) \subseteq S \setminus D_S)\, (P'(k) \subseteq F(S'(k)))\}_{k=1}^{T}} \sum_{k=1}^{T} J(A_{P'}(k), x(k)) \tag{17}$$

**Algorithm 5** Algorithm for Finding Critical Substations to Protect: `Get_Dynamic_Defense`$(\mathcal{G}_\mathcal{P}, B_P, B_D, S_{info}, a_k)$

**Input:** $\mathcal{G}_\mathcal{P}, B_P, B_D, S_{info}, a_k$
**Initialize:** $D^{'}_s \leftarrow \emptyset, D_S \leftarrow \emptyset, \hat{L}_w \leftarrow 100, L_{Prev} \leftarrow 100, L_H \leftarrow \emptyset$
$S^{'}_t(k), P^{'}(k), L^d_w, a^d_k \leftarrow$ `Get_WDA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k)$
$L_H \cup L_{Prev}$
**for** $i = 1, \ldots, B_D$ **do**
  $\hat{L}_w \leftarrow 100, flag \leftarrow 0$
  **if** $D^t_S \neq \emptyset$ **then**
    $S^{'}_t(k), L_{Prev} \leftarrow$ `Get_WDA1`$(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k, D^t_S, \emptyset)$
    $L_H \cup L_{Prev}$
  **end if**
  **for all** $s \in S^{'}_t(k)$ **do**
    $S^{'}_s(k), P^{'}_s(k), L^{'}_s \leftarrow$ `Get_WDA2`$(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k, D^t_S, s)$
    **if** $L^{'}_s < \hat{L}_w$ **then**
      $\hat{L}_w \leftarrow L^{'}_s, D^{'}_s \leftarrow s, flag \leftarrow 1$
    **end if**
  **end for**
  $D_S \leftarrow D_S \cup D^{'}_s, D^t_S \leftarrow D^t_S \cup D^{'}_s$
  **if** $\hat{L}_w > min(L_H)$ AND $flag == 1$ **then**
    $D_S \leftarrow D_S \setminus D^{'}_s$
  **else**
    $D_S \leftarrow D^t_S$
  **end if**
**end for**
**return** $D_S$

$$x(k) = \begin{cases} G(H(k)), & \text{if} \quad H(k) = \{A_{P'}(i)\}_{i=1}^{k-1} \\ g(H(k)), & \text{if} \quad H(k) = \emptyset \end{cases} \tag{18}$$

$$|D_S| \leq B_D \tag{19}$$

$$\sum_{k=1}^{T} |S^{'}(k)| \leq B_S \tag{20}$$
$$\forall k, k' \in \{1, ....., T\} : S^{'}(k) \cap S^{'}(k') = \emptyset, k \neq k'$$

$$\sum_{k=1}^{T} |P^{'}(k)| \leq B_P \tag{21}$$
$$\forall k, k' \in \{1, ....., T\} : P^{'}(k) \cap P^{'}(k') = \emptyset, k \neq k'$$

$$B_S \leq B_P \tag{22}$$

where, $x(k)$ represents the state of the system at time step $k$ and $H(k)$ represents the attack history of the system.

### B. Algorithm for Finding the Critical Substations to Protect

Algorithm 5 starts with an empty set and intelligently identifies the critical substations to protect one-by-one such that when an attack is launched the overall system damage can be minimized. The algorithm takes the same inputs as Algorithm 3 with the defense budget $B_D$ as an additional input and identifies the critical substations $D_S$ to protect.

First, the worst-case dynamic attack is identified by using `Get_WDA`$(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k)$ which is described as Algorithm 3. Next, if there are no critical substations in $D_S$, We use the critical substations $S^{'}_t(k)$ identified from the worst-case dynamic attack to identify the first substation to protect. We iteratively protect each substation in $S^{'}_t(k)$ and evaluate the overall system damage post dynamic attack using `Get_WDA2`$(\mathcal{G}_\mathcal{P}, B_P, S_{info}, a_k, D^t_S, s)$. The computed system damage in each iteration is used to select the substation to protect, i.e., $D_S \leftarrow D_S \cup D^{'}_s$. A track of intermediate solution $D^t_S \leftarrow D^t_S \cup D^{'}_s$ is kept in order to obtain a better solution.

Note that, the function $\texttt{Get\_WDA2}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, s)$ is same as Algorithm 3, however, here the worst-case dynamic attack is computed by eliminating the protected substations $D_S^t$ and the substation $s$ from the attackable list of substations, i.e., $S \setminus (D_S^t \cup s)$. If the computed damage $L_s^{'}$ is smaller than the maximum damage $\hat{L}_w$, the solution is updated.

Additionally, in each iteration, if the protected substations set $D_S^t$ is non-empty then a new set of critical substations are identified using the worst-case dynamic attack function, i.e., $\texttt{Get\_WDA1}(\mathcal{G}_{\mathcal{P}}, B_P, S_{info}, a_k, D_S^t, \emptyset)$. This function is also same as Algorithm 3, however, the protected substations $D_S^t$ are removed from the attackable list of substations while executing the worst-case dynamic attack on the power system model $\mathcal{G}_{\mathcal{P}}$. It ensures that once the substations are protected, the attacker can only launch the dynamic attack on the remaining substations based on the attack budget. The obtained attack can further be utilized to identify the substation to protect considering the defense budget constraints. In the algorithm $L_H$ keeps a track of all the previous load losses obtained after protecting the substations in $D_S^t$ and updates the final solution $D_S$ depending upon the comparison of the obtained damage with the previous system damages. This ensures a better protection mechanism that provides an effective solution.

## VII. Evaluation

We considered two standard IEEE systems, the 39 bus and 57 bus systems to evaluate our approach. We used a modified version of the steady state simulator discussed in [15] to perform the analysis. First, we discuss how randomly chosen attacks can be optimized using our dynamic attack model. Next, we show the optimization of the worst-case static attacks using the dynamic attack model. Then, we present the dynamic defense results that show the reduction in the overall system damage/load loss. Finally, we evaluate the performance of our algorithm's execution time for the dynamic attack and defense algorithms in comparison with the naive exhaustive search algorithm.

### A. Optimizing Random Attacks

Figure 2 shows the optimization of the random attacks using the dynamic attack model discussed in Section V. Here, depending upon the attack budget (up to 6), we randomly picked the components to attack from the power system model. Then, we used these attacks as inputs to our dynamic attack algorithm to obtain a strategic sequence in which the attacks can be executed so as to maximize the system damage. We performed our evaluation on the IEEE 39, 57 bus system and the results are shown in Figure 2. The x-axis represents the attack budget whereas the y-axis represents the system damage, i.e., load loss. Red, green color markers represent the random and strategic dynamic attacks respectively.

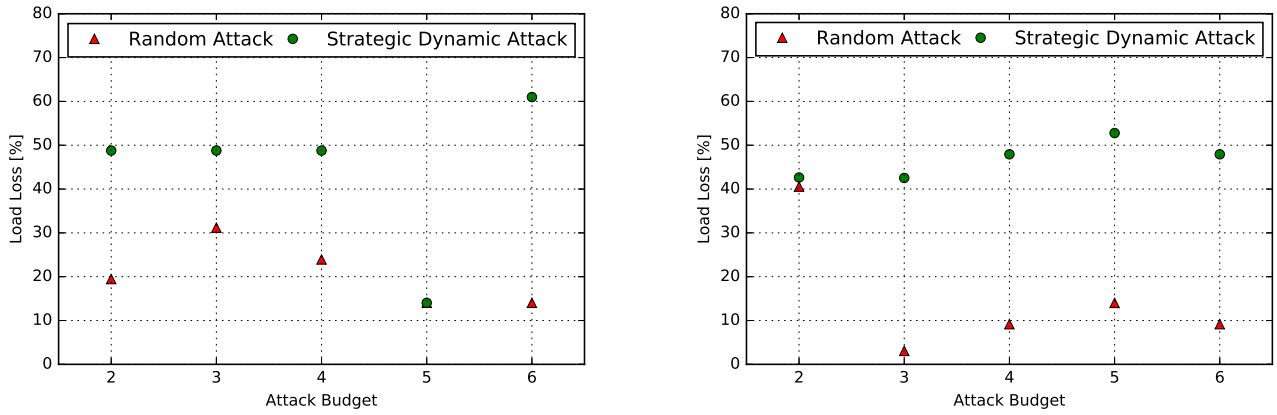
For both the standard IEEE systems, we can clearly see from Figures 2a and 2b that our dynamic attack algorithms described in this paper are able to strategically identify the specific instants (or sequences) at which different attacks can be executed and maximize the system damage for a randomly identified set of components to attack. From Figure 2a, for an attack budget of 6 in IEEE-39 bus system the random attack caused a load loss of 14.03%, however, the same attack when executed at different instants in time, i.e., dynamic attack resulted in a total load loss of 60.99%. The dynamic attack on the same components caused a 334.71% higher load loss than the static attack. For the same attack budget in the IEEE-57 bus system the random attack caused a load loss of 9.16%, whereas, the dynamic attack resulted in a load loss of 47.93% as shown in Figure 2b. This dynamic attack load loss is 423.25% higher than the random attack.

### B. Optimizing Static Attacks

We perform the analysis on the same IEEE systems. First, we identified the worst-case static attack and then we use it to identify the worst-case dynamic attack in order to further maximize the system damage. Figure 3 shows the results for the optimization of the worst-case static attack using our dynamic attack model and algorithm. The x-axis represents the attack budget, whereas, the y-axis represents the system damage. Red, green colored markers represent the worst-case static and dynamic attacks respectively. Here, we consider an attack budget of up to 6 components/substations.

From Figures 3a and 3b it is clear that the dynamic attack causes higher damage with different attack budgets. As shown in Figure 3a, for an attack budget of 2 in IEEE-39 bus system the worst-case static attack caused a load loss of 84.27%, however, the optimized worst-case dynamic attack resulted in a load loss of 96.60%. Here, the dynamic attack on the same components caused a 14.63% higher load loss. Similarly, for the IEEE-59 bus system in Figure 3b, the worst-case static attack caused a load loss of 50.70%, whereas, the optimized worst-case dynamic attack resulted in a load loss of 54.15% for an attack budget of 3. The dynamic attack caused a higher load loss by 6.80%. Note that, the worst-case static attacks are already identified as the attacks that cause maximum damage, however our dynamic attack algorithms are still able to optimize them for obtaining even higher system damage if there is a possibility for optimization. The dynamic attack algorithm results from Figure 3 clearly show that the dynamic attacks on the same components that are identified from the static attack scenario when scheduled and executed strategically resulted in a higher system damage. Note that, in Figure 3, the static and dynamic attack load loss becomes equal for some attack budgets because there is no additional load loss possible within the system. Also note that, for some attack budgets the difference in the load loss between the static attack and the dynamic attack can remain very small because the additional loads that gets disconnected during the dynamic attack maybe smaller in magnitude as compared to the total load loss. However, if the additional load loss is larger in magnitude, then this difference can be significantly larger as shown by attack budget 2, 3 in Figures 3a and 3b respectively.

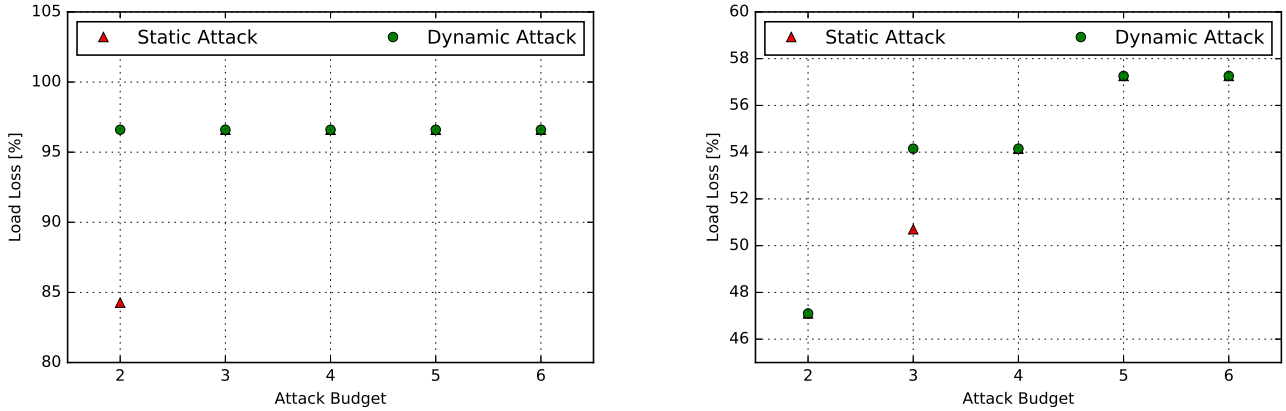
We have shown the exact cascade progression for one of the static and dynamic attack scenarios in Table III that can easily answer the question of 'how dynamic attacks can have higher impact?'. For both the attack scenarios, we consider the same substations and its components to attack, but the only




(a) IEEE-39 bus system

(b) IEEE-57 bus system

Fig. 2: Random Attacks Vs Dynamic Attacks: Load loss as a function of various attack budgets for different standard IEEE systems.




(a) IEEE-39 bus system

(b) IEEE-57 bus system

Fig. 3: Static Attacks Vs Dynamic Attacks: Load loss as a function of various attack budgets for different standard IEEE systems.

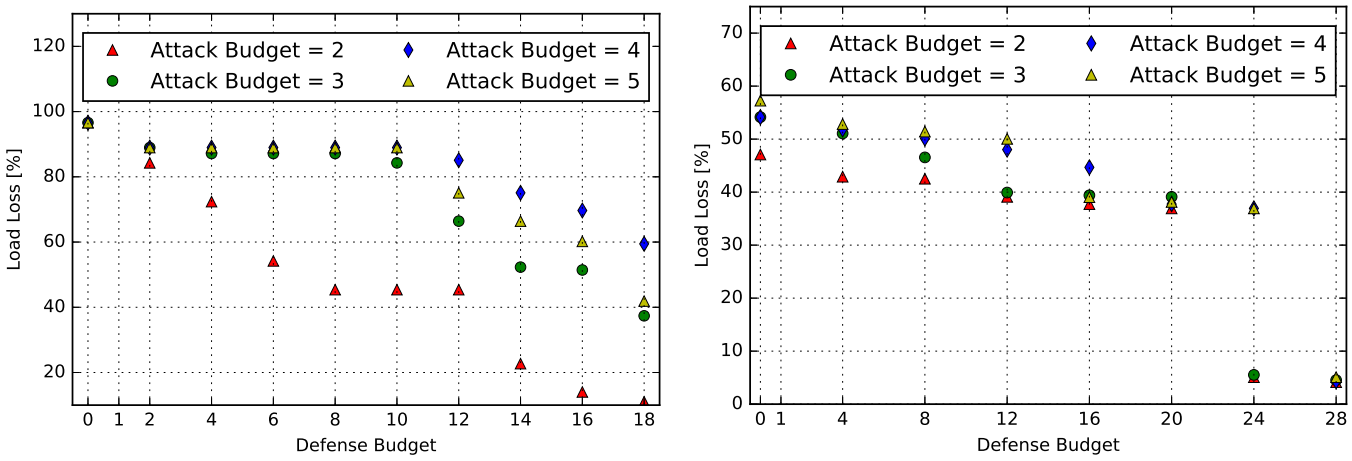
difference is the attack time. For the static attack scenario with an attack budget of 2, Table III shows that both the attacks are launched at the same time [0, 0] ([0, 0] indicates simultaneous attack or static attack). As a result of the static attack the transmission lines associated with the attacked protection assemblies are isolated. This resulted in a sequence of cascading failures as shown by the 'Stage 1 Outages' through 'Stage 4 Outages' in Table III and the total system load loss was observed to be 84.27%.

Now, we consider the same substations and protection assemblies for the dynamic attack scenario. Here, the initial attack takes place at time instant 0 that initiated a cascading event causing subsequent failures (Stage 1 Outages in Table III). At time instant 1, another attack was launched that further weakened the system causing Outages through Stage 2 to Stage 5 resulting in a significant damage to the system. The overall system load loss was observed to be 96.60% (Stage 2 and Stage 5 Outages in Table III) which is considerably higher than the static attack. Note that the specific time at which these attacks can be executed are computed using the algorithms described in Section V.

### *C. Minimizing System Damage Using Dynamic Defense*

We evaluate our defense model and algorithm using the standard IEEE-39 and 57 bus systems. Figure 4 shows the load losses in the power system at different attack budgets when a dynamic attack is launched after the critical substations are intelligently identified and protected depending upon the defense budget. In each figure, the x-axis represents the defense budget and the y-axis represent the total system damage. Red, green, blue, and yellow colored markers represents the attack budgets 2, 3, 4 and 5 respectively. The respective color markers at the defense budget 0 represent the total system damage without any defense.

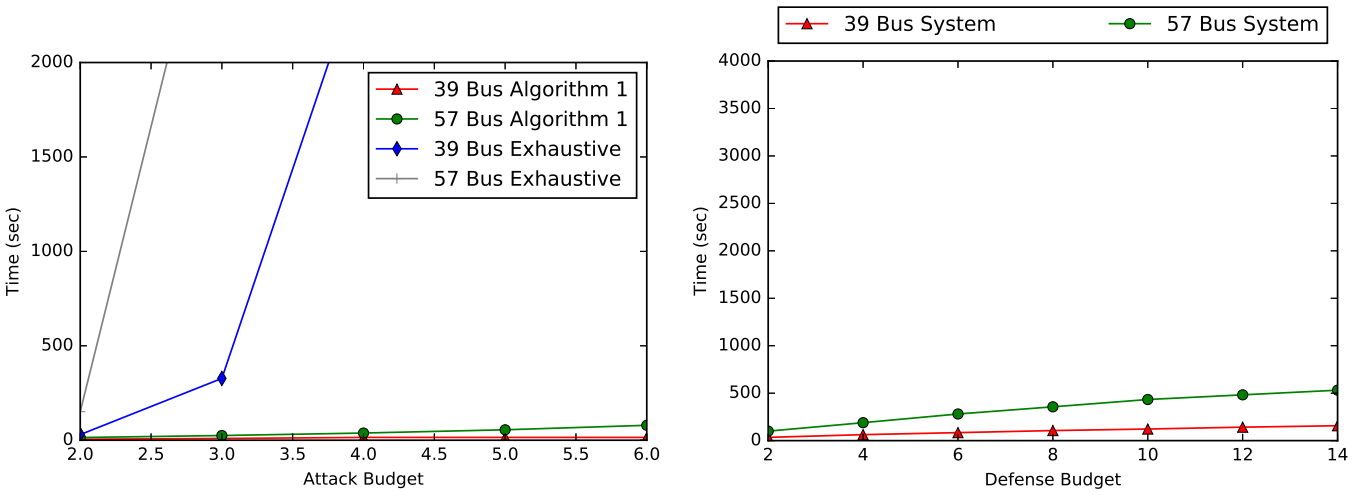
From Figure 4, we can clearly see that by intelligently selecting and protecting the critical substations of the power network, the system damage can be significantly reduced for IEEE-39 bus system (Figure 4a) and 57 bus system (Figure 4b) when a dynamic attack is launched. In Figure 4a, for an attack and a defense budget of 2, the load loss is reduced from 96.60% to 84.27%, that is, a total of 12.76% reduction in load loss. Moreover, for the same attack budget and a




(a) IEEE-39 bus system

(b) IEEE-57 bus system

Fig. 4: Dynamic Defense: Load loss as a function of various defense budgets for different standard IEEE systems.




(a) Execution time of attack analysis

(b) Execution time of defense analysis

Fig. 5: Analysis execution time for attack and defense for different standard IEEE systems

defense budget of 18, a total of 88.58% reduction in load loss is observed. For other attack budgets, as the defense budget increases we can see significant improvement in the reduction of total system load loss.

### *D. Performance of the Dynamic Attack and Defense Algorithms*

Here, we compare the execution time of our dynamic attack and defense algorithms with the naive exhaustive search algorithms. We use the same standard IEEE systems to perform our analysis. Figure 5 shows the dynamic attack and defense execution time with respect to the exhaustive search. In each figure, the x-axis represents either the attack budget or the defense budget and the y-axis represents the time taken by the algorithm to identify the attack or defense. The details of the markers are shown in the legend box of Figure 5.

From Figure 5a, we can clearly see that the time taken to identify the dynamic attack for IEEE-39, 57 bus system increases very slightly with increase in the attack budget. However, the time taken to identify the attack using the exhaustive search algorithm is observed to be exponential even at smaller attack budgets. The exhaustive search execution time in Figure 5a represents the time taken to identify the maximum damage causing static attack. Moreover, the exhaustive search execution time for identifying the maximum damage causing dynamic attack will be much larger than the time taken to identify the static attack. Similarly, it is clear from Figure 5b that the time taken to identify the defense increases slowly with the increase in the defense budget. We know that dynamic defense via exhaustive analysis will take much longer than the exhaustive attack since it will have to first identify the attack and then identify the defense. Hence, if we compare only against the attack time, it still shows that the developed approach is much faster than the exhaustive search. Therefore, as demonstrated in Figure 5, our algorithms prove to be far more efficient than the naive exhaustive search.

TABLE III: Scenario representing the maximization of system damage using dynamic attack model

| Static Attack | | Dynamic Attack | |
|---|---|---|---|
| Initial Attack | Attack time vector: [0, 0]<br>Substations compromised: [$S^{13}$, $S^{24}$]<br>Protection assemblies attacked: [PA10, PA16]<br>Transmission lines Isolated due to the attacked protection assemblies: [‘$R16\_19$’, ‘$R2\_3$’]<br>Load loss: ‘$0\%$’ | Initial Attack | Attack time vector: [0]<br>Substations compromised: [$S^{24}$]<br>Protection assemblies attacked: [PA16]<br>Transmission lines Isolated due to the attacked protection assemblies: [‘$R2\_3$’]<br>Load loss: ‘$0\%$’ |
| Stage 1 Outages | Isolation of transmission lines due to the secondary effect of the outages from the initial attack: [‘$R2\_25, R25\_26, R18\_17, R27\_26$’], Load loss: ‘$0\%$’ | Stage 1 Outages | Isolation of transmission lines due to the secondary effect of the outages from the initial attack: [‘$R2\_25, R18\_17$’], Load loss: ‘$0\%$’ |
| Stage 2 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 1: [‘$R6\_5, R14\_15, R14\_13, R10\_13, R26\_28, R21\_22$]<br>Load loss: ‘$35.48\%$’ | Additional Attack | Attack time vector: [1]<br>Substations compromised: [$S^{13}$]<br>Protection assemblies attacked: [PA10]<br>Transmission lines Isolated due to the attacked protection assemblies: [‘$R16\_19$’]<br>Load loss: ‘$0\%$’ |
| Stage 3 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 2: [‘$R8\_7, R6\_7, R10\_11, R6\_11$]<br>Load loss: ‘$64.80\%$’ | Stage 2 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 1: [‘$R6\_5$’], Load loss: ‘$0\%$’ |
| Stage 4 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘$R9\_39, R8\_9$’], Load loss: ‘$84.27\%$’ | Stage 3 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 2: [‘$R8\_7, R6\_7, R4\_14, R14\_13, R10\_13$’], Load loss: ‘$7.25\%$’ |
| | | Stage 4 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘$R9\_39, R8\_9, R21\_22, R24\_23$’], Load loss: ‘$56.42\%$’ |
| | | Stage 5 Outages | Isolation of transmission lines due to the secondary effect of the outages from the stage 3: [‘$R25\_26, R17\_27, R27\_26, , R16\_17, R26\_28, R28\_29, R26\_29, R16\_21$’], Load loss: ‘$96.60\%$’ |

## VIII. Conclusions and Future Work

We described the static and dynamic cyber-attack and defense models for electrical power systems using game-theoretic approach. From the attacker's perspective, we provide an efficient and effective algorithm that is able to strategically identify the dynamic attacks that maximizes the system damage by considering both random attacks as well as worst-case static attacks. We also provide an efficient algorithm from defenders perspective that identifies the critical substations to protect in order to minimize the overall system damage. Our results shows that, under financial budget constraints, intelligently selecting the substations to prioritize and protect can significantly improve the power system resilience. In addition, these algorithms are efficient and perform significantly better than the exhaustive search even with the complex dynamic attack and defense models. As part of the future work, the attacker-defender models can be easily extended to consider randomness, i.e., a success probability can be associated with an attack and a defense that can give us more insight to improve the power system resilience under probabilistic scenarios. Further, under unknown circumstances where the defender has no idea whether an attacker follows a static attack model or a dynamic attack model, a defense strategy that could improve the overall power system resilience irrespective of the attack model can be an interesting direction to explore.

## Acknowledgment

This work is funded in part by the NSF under the award number CNS-1329803 and the NSF FORCES project under the award number CNS-1238959. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF or FORCES. The authors would like to thank Amin Ghafouri and Charles Hartsell for their help and discussions related to the work presented here.

## References

[1] N. R. Council *et al.*, *Terrorism and the electric power delivery system*. National Academies Press, 2012.

[2] P. J. Hawrylak, M. Haney, M. Papa, and J. Hale, "Using hybrid attack graphs to model cyber-physical attacks in the smart grid," in *Resilient Control Systems (ISRCS), 2012 5th International Symposium on*. IEEE, 2012, pp. 161–164.

[3] J. David, "Double threat: Us grid vulnerable on two fronts," *CNBC. Retrieved from http://www. cnbc. com/id/101306145*, 2014.

[4] C. W. Draffin Jr, "Cybersecurity white paper," 2016.

[5] "A cyberattack on the u.s. power grid," https://www.cfr.org/report/cyberattack-us-power-grid/, Council on Foreign Relations, 2017.

[6] D. B. Rawat and C. Bajracharya, "Cyber security for smart grid systems: Status, challenges and perspectives," in *SoutheastCon 2015*. IEEE, 2015, pp. 1–6.

[7] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Computer Networks*, vol. 57, no. 5, pp. 1344–1371, 2013.

[8] S. Gorman, "Electricity grid in us penetrated by spies," *The Wall Street Journal*, vol. 8, 2009.

[9] Q. T. Review, "Enabling modernization of the electric power system," *U.S. Department of Energy*, 2015.

[10] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, 2012.

[11] T. Pultarova, "Cyber security-Ukraine grid hack is wake-up call for network operators [news briefing]," *Engineering & Technology*, 2016.

[12] S. NERC, "Top-004–2: Transmission operations," *North American Electric Reliability Corporation*, 2007.

[13] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis & Design, SI Version*. Cengage Learning, 2012.

[14] M. J. Eppstein and P. D. Hines, "A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1698–1705, 2012.

[15] S. Hasan, A. Ghafouri, A. Dubey, G. Karsai, and X. Koutsoukos, "Heuristics-based approach for identifying critical n-k contingencies in power systems," in *Resilience Week (RWS), 2017*. IEEE, 2017, pp. 191–197.

[16] S. Liu, S. Mashayekh, D. Kundur, T. Zourntos, and K. Butler-Purry, "A framework for modeling cyber-physical switching attacks in smart grid," *IEEE Transactions on Emerging Topics in Computing*, 2013.

[17] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, E. G. Im, Z. Yao, B. Pranggono, and H. Wang, "Man-in-the-middle attack test-bed investigating cyber-security vulnerabilities in smart grid scada systems," 2012.

[18] H. Lin, H. Alemzadeh, D. Chen, Z. Kalbarczyk, and R. K. Iyer, "Safety-critical cyber-physical attacks: Analysis, detection, and mitigation," in *Proceedings of the Symposium and Bootcamp on the Science of Security*. ACM, 2016, pp. 82–89.

[19] S. Sridhar and G. Manimaran, "Data integrity attack and its impacts on voltage control loop in power grid," in *PES General Meeting, 2011*.

[20] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Transactions on Industrial Informatics*, 2015.

[21] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731–1738, 2012.

[22] V. Turau and C. Weyer, "Cascading failures caused by node overloading in complex networks," in *Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG), IEEE 2016*, pp. 1–6.

[23] B. Chen, S. Mashayekh, K. L. Butler-Purry, and D. Kundur, "Impact of cyber attacks on transient stability of smart grids with voltage support devices," in *Power and Energy Society General Meeting (PES), 2013*.

[24] S. Hasan, A. Chhokra, A. Dubey, N. Mahadevan, G. Karsai, R. Jain, and S. Lukic, "A simulation testbed for cascade analysis," in *Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017 IEEE*. IEEE, 2017, pp. 1–5.

[25] S. Poudel, Z. Ni, and N. Malla, "Real-time cyber physical system testbed for power system security and control," *International Journal of Electrical Power & Energy Systems*, vol. 90, pp. 124–133, 2017.

[26] J. Yan, Y. Zhu, H. He, and Y. Sun, "Multi-contingency cascading analysis of smart grid based on self-organizing map," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 4, pp. 646–656, 2013.

[27] W. Yuan, L. Zhao, and B. Zeng, "Optimal power grid protection through a defender–attacker–defender model," *Reliability Engineering & System Safety*, vol. 121, pp. 83–89, 2014.

[28] K. Hausken and G. Levitin, "Minmax defense strategy for complex multi-state systems," *Reliability Engineering & System Safety*, vol. 94, no. 2, pp. 577–587, 2009.

[29] Y. Xiang, L. Wang, and N. Liu, "Coordinated attacks on electric power systems in a cyber-physical environment," *Electric Power Systems Research*, vol. 149, pp. 156–168, 2017.

[30] S. Liu, B. Chen, T. Zourntos, D. Kundur, and K. Butler-Purry, "A coordinated multi-switch attack for cascading failures in smart grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1183–1195, 2014.

[31] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He, "Joint substation-transmission line vulnerability assessment against the smart grid," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 1010–1024, 2015.

[32] Y. W. Law, T. Alpcan, and M. Palaniswami, "Security games for risk minimization in automatic generation control," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 223–232, 2015.

[33] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 87–98, 2013.

[34] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[35] J. Yan, Y. Zhu, H. He, and Y. Sun, "Revealing temporal features of attacks against smart grid," in *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*. IEEE, 2013, pp. 1–6.

[36] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He, "Resilience analysis of power grids under the sequential attack," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2340–2354, 2014.

[37] B. Liscouski and W. Elliot, "Final report on the august 14, 2003 blackout in the united states and canada: Causes and recommendations," *A report to US Department of Energy*, vol. 40, no. 4, 2004.

[38] H. Pidd, "India blackouts leave 700 million without power," *The guardian*, vol. 31, 2012.

[39] http://icseg.iti.illinois.edu/power-cases/, ICSEG.

[40] S. Hasan, A. Ghafouri, A. Dubey, G. Karsai, and X. Koutsoukos, "Vulnerability analysis of power systems based on cyber-attack and defense models."