Smart Grid Vulnerability and Defense Analysis Under Limited Attack Budget

Tu N. Nguyen, *Senior Member, IEEE,* Bing-Hong Liu, *Member, IEEE,* Nam P. Nguyen, *Member, IEEE,*
Vijay K. Shah, *Member, IEEE,* and Jung-Te Chou

***Abstract*—Most of today's infrastructure systems can be efficiently operated, thanks to the intelligent power supply of smart grids. However, smart grids are highly vulnerable to malicious attacks, mainly due to the interplay between the components (e.g., substations and transmission lines). The failure of one or more critical components may trigger the sequential failure of other components, resulting in a cascading failure and the eventual breakdown of the whole system. Thus, in order to protect a smart grid system, two important challenges arises for an operator: (i) *to identify and protect the most critical components in the system.*; if attackers find the the most critical components of the system, the whole system is at risk, and (ii) *to design an efficient system recovery strategy*; if the system already suffers from attacks, a well-designed strategy can greatly help minimize the scale of damage. There are efforts in the existing literature to design different ranking methods for node or link identifications. However, they simply select the highest degree nodes or common links as the critical ones and these prior solutions neither consider the interdependency between components in the system nor the role of system users, which is one of the critical factors that impacts the smart grid vulnerability assessment. [This motivates us to study a more general and practical problem in the context of smart grid vulnerability assessment] [THIS MOTIVATION IS NOT IN LIEU WITH THE PAPER TITLE.], referred to as *Maximum-Impact through Critical-Line with Limited Budget* (MICLLB) problem. We propose an efficient algorithm, called *Greedy Based Partition Algorithm* (GBPA) to solve the MICLLB problem. In addition, we also design an algorithm, namely *Homogeneous-Equality Based Defense Algorithm* (HEBDA) to help reduce the damages in case the system are suffering from the attacks. Through rigorous theoretical analysis and experimentation, we demonstrate that our proposed methods perform well within reasonable bounds of computational complexity.**

***Index Terms*—Smart power grids, cascading failure, attack, defense.**

I. Introduction

SMART grids provide electric power to almost all critical systems, and are thus, considered one of the most important infrastructures in the world. However, smart grids are extremely vulnerable to cascading failures [1], [2], [3], [4], [5], [6] wherein initial failures of one or more components may trigger the sequential failure of other components. This vulnerability may result in a devastating consequence if the initial component failures are critical ones. In the smart grid, each component (node) has load and capacity that may leverage a cascading effect, resulting in a great impact on other components, even started with a single failure of a single node [7], [8]. After a node fails, its load is redistributed to its neighbors because of the balance changes of flows. Likewise, it leads to a global redistribution of loads over the entire system. This can trigger a cascade of overloading failures due to a limited capacity on each node [9].

Leveraged by advances in electrical and networking technologies, the smart grids architecture has been widely studied for enhancing the performance of power grids in order to deliver electricity efficiently [10], [11], [12]. A common smart grid uses supervisory control and data acquisition (SCADA) systems to manage the electricity distribution [13]. The operation center of smart grids is responsible for making decisions in response to outage incidents, which is based on information collected from the status of system components such as stations and consumers [14], [15]. Moreover, the operation center is able to provide consumers with the *real-time* expense/price of the consumers' usage [16]. Therefore, consumers can adjust the time-of-use plans in order to reduce their expenses. This communication is one of advantages of the smart grid compared with the traditional power grid. However, if the operation center receive incorrect data in regards to consumers' usage, it possibly make wrong decisions in power distribution plan, resulting in potential failures.

Many existing works have proposed to mislead the state estimation process that aim to tend the operation center making wrong decisions. This kind of malicious attacks is called false data injection attack [17]. The authors in [18] formulate an attack to smart grids based on false data injection and aim to maximize the total operation costs of the system. They focus on characterizing the load buses and the line measurements in order to interrupt the dispatch of electricity. Moreover, in [19], the authors also try to interfere the dispatch of electricity by preventing a complete topology information sent to the operation center. These existing works aim to inject false data to the smart grids to increase the smart grid's operation cost. The smart grid attack using false data injection has attracted a lot of attention and been studied in various perspectives. In [20], they propose a framework to maximize the damage if a smart grid uses forged topology information. Additionally, in [21], the authors design a successful false data injection attack using limited topology information. They show the possibility to attack a smart grid with a limited attack resources.

T. N. Nguyen is with the Department of Computer Science, Purdue University Fort Wayne, Fort Wayne, IN 46805, U.S.A. (e-mail: nguyent@pfw.edu.) B.-H. Liu and J.-T. Chou are with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 80778, Taiwan. (e-mail: bhliu@nkust.edu.tw and qq3025566@gmail.com.) N. P. Nguyen is with the Department of Computer and Information Sciences, Towson University, MD, U.S.A. (e-mail: npnguyen@towson.edu.) V. K. Shah is with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061, U.S.A. (e-mail: vijays@vt.edu.)

Moreover, in [22], the authors propose an undetectable attack that also uses false data injection to modify the real price of electricity market. Furthermore, in [23], they aim to discover the impact and emphasize the importance for defenders to develop some strategies against this kind of attack. Although, various models of smart grid attack are proposed to study the vulnerability [18], [19], [20], [21], [22], [23], these works fail to address the problem of how to destroy the system or make a large blackout, instead they mainly present different ranking methods that use the false data injection. to increase operation cost of the system.

In recent times, there has been a surge in the attempts to study the smart grid vulnerabilities under the targeted attack that focus on characterizing critical components [24], [25]. The authors in [24] rely on the maximum flow to identify the lines with maximum load. They consider these lines as the most critical components that their failures have a great impact on the entire system. There is research work [25] that designs an attack by manipulating the price rate of users to trigger the cascading failure on the smart grids. They also consider some particular lines that cause the most overloaded transmissions as the critical components. Although there are many algorithms proposed for the problem of attacking smart grids, they only aim to destroy the nodes or transmission lines, and there is no research that studies the smart grid vulnerability in practical scenarios such as quantifying the total impact of attacks on users (consumers). Intuitively, the satisfaction of users for daily electricity use is one of major factors impacting on the way we understand the smart grid vulnerability and design appropriate defense strategy. This motivates us to come up with an idea of classifying users and components based on their importance. The aim of this work is to study a more general and practical problem in terms of smart grid vulnerability assessment and system defense analysis under the limited attack budget, referred to as the Maximum-Impact through Critical-Line with Limited Budget (MICLLB) problem.

While investigating the MICLLB problem, we identified many new challenges when compared to previous works. We summarize these main challenges as follows.

- **C1:** To understand the vulnerability of smart grids, we have to not only understand the property of each individual component, but also need to characterize the interdependency between components because the failure of any component can lead to unpredictable subsequent failures. In previous works, researchers fail to address the problem of interdependency between components nor consider the role of users that is one of critical factors impacting on the vulnerability of smart grids. It is clearly much harder and requires a complex technique to understand deeply the vulnerability of smart grids.
- **C2:** To achieve the maximum total impact of attacks on users, it requires an efficient attack algorithm in the smart grid. In previous works, they introduce some efficient algorithms for cascading failure attacks because they consider the attack budget is unlimited. As always, these assumptions are unrealistic since the scale of attacks really depends on available attack resources. In practice, sometimes the attack budget is insufficient, and a cascading failure does not always work. Thus, how to design a full approach for realistic attacks and defenses in smart grids is a challenge.
- **C3:** The third challenge is how to theoretically analyze the framework to design an actual attack and defense algorithm. Since the investigated problem (MICLLB) is NP-complete (we discuss the difficulty of the problem in detail in Section IV-A), there is no way to get exactly the subsequent failures of the whole system, and it is also difficult to determine relevant components for the order of attacks. The question of guarantee for maximum total impact of attacks on users becomes harder and harder. Hence, the performance of the desired algorithm not only depends on an comprehensive evaluation but also requires a specific technique and mechanism for the smart grid vulnerability analysis.

[Looks like there are two focuses of this work - impact on users and critical components. This needs to come clearly in the paper. Currently, it is little all over the place.] [Will come back to revise this point in detail]

To address these challenges, we first propose a novel attack algorithm, referred to as the Greedy Based Partition Algorithm (GBPA), to maximize the total impact on users. We then design an appropriate defense algorithm, namely Homogeneous-Equality Based Defense Algorithm (HEBDA) to protect critical components that helps reduce the impact of the GBPA attack on users. Comprehensively, we implement our proposed methods through simulations and analyses. We summarize the main contributions of this paper as follows.

- We study the problem of designing an efficient attack method to maximize the total impact on users in smart grids, referred to as the Maximum-Impact through Critical-Line with Limited Budget (MICLLB) problem.
- We characterize the system's topology and range all the components based on their critication using our proposed metric. Based on the critication ranking ontained we establish the ordered set of attacks ($\mathcal{ATK}$) for the MICLLB problem. We propose a new attack algorithm using the obtained $\mathcal{ATK}$, referred to as the Greedy Based Partition Algorithm (GBPA). Then, an appropriate defense algorithm, called Homogeneous-Equality Based Defense Algorithm (HEBDA) is proposed to secure the critical components.
- Theoretical analyses of GBPA and HEBDA are provided to support their performance. We also conduct intensive simulations to compare GBPA and HEBDA with related methods. The results show that the GBPA and HEBDA provide better performance in terms of the total impact on consumers than the existing solutions for the MICLLB problem.

Organization: The paper is organized as follows. Section I discusses related works on smart grid vulnerability problem. In section II, we describe the smart grid model, the DC power flow approximation, the interplay of transmission line failure, the attack model, and the problem definition. We introduce related formulation, metric, and attack algorithm in section

III-A. In section III-B, we present an appropriate defense algorithm. Section IV provides the time complexity analysis and the difficulty of the investigated problem. In Section V, we evaluate the performance of the GBPA and HEBDA. Finally, we make key concluding remarks in Section VI.

II. Preliminaries

A. Smart Grid Network Model

[A figure would be nice.] We abstract a general smart grid as a system architecture that is composed of generators, transformer stations, and consumers. A generator is responsible for supplying power to the system. We assume that a transformer station is an intermediate station that basically helps transfer power and they are with no power generation as well as no power consumption. Consumers consist of a set of consumption sectors such as houses, industries, and data centers. Hence, a smart grid can be modeled by a weighed directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$, where every node $u \in \mathcal{V}$ belongs to either the set of generators G, transformer stations T or consumers C, i.e., node set $\mathcal{V} = G \cup T \cup C$. In addition, $\mathcal{E}$ represents the set of power transmission lines, wherein each line $(u, v) \in \mathcal{E}$ connects two nodes u and v. The rest of parameters in the graph $\mathcal{G}$ is defined as follows:

- $f : \mathcal{E} \rightarrow \mathbb{Z}$ is a line power flow function,
- $\delta : \mathcal{E} \rightarrow \mathbb{Z}_0^+$ is a line maximum capacity function,
- $\tau : \mathcal{E} \rightarrow \mathbb{Z}_0^+$ is a line cost function, and
- $\omega : \mathcal{V} \rightarrow \mathbb{Z}_0^+$ is a node weighting function.

Each edge (u, v) is also associated with a weight $\epsilon(u, v)$ representing the operation parameter of the system. The higher $\epsilon(u, v)$ is, the more power flow is distributed from u to v. In this paper, the terms *link* and *line* are sometimes used interchangeably. Likewise, we also use the terms *node* and *vertex* interchangeably.

In a smart grid, each generator $u \in G$ has power generation output p_u and each consumer v has a demand d_v. The power transmission line (u, v) has power flow $f(u, v)$ and maximum capacity $\delta(u, v)$. Normally, maximum capacity is less than its physical or thermal limit. The power flow on each transmission line is therefore always maintained below the line's maximum capacity. If the power flow on a link rises over its maximum capacity, it results in out-of-service, mainly due to thermal limit. A consumer is usually connected to generators by several lines. When a transmission line is out of service, the total demand at its terminal may be shifted to other transmission lines. The workload increases quickly and become a large burden to the rest of lines, resulting in a cascading failure in the system.

B. DC Power Flow Approximation

Recently, a great number of attempts has been made towards the modeling of the behavior of the power grids. It is worth mentioning that there exists the equations, called alternative current (AC) power flow, that can estimate the flow values for each component in the network [26]. The AC power flow equations are non-convex and non-linear and illustrate behavior of the power grids for both active and reactive powers [25], [27], [28]. While the active power flow estimation in the AC model is extremely hard because it depends on the complex voltage at nodes, direct current (DC) load flow analysis appears as the best solution to estimate line flows across the grid with a simplification and linearisation of AC power flow equations. The active power flow $f(u, v)$ within (u, v) is defined as follows:

$$f(u, v) = \frac{|I_u|\,|I_v|}{x_{uv}} \sin(\theta_{uv}), \tag{1}$$

where $|I_u|$ and $|I_v|$ represent the voltage amplitude at node u and v, respectively; θ_{uv} represents the voltage phase difference between node u and node v, and x_{uv} is the impedance of (u, v). In order to reduce the running time and making the linearization feasible in solving the power flow problem, most of previous works [29] rely on the linearization or DC power flow approximation by using the assumptions as follows:

- Voltage phase differences are small, i.e., $\sin(\theta_{uv}) \approx \theta_{uv}$.
- Line resistance is ignored that $x_{uv} = r_{uv}$, in which r_{uv} is the reactance of (u, v).
- Reactive power balance equations are neglected.

In the previous work [30], the authors clarify that above assumptions result in the DC load flow's performance reaching ten times faster compared to AC load flow's performance. Thus, we adopt the DC power flow approximation to estimate the flow values in the network for our paper. We summarize the equations and cases that reflect characteristics of the DC power flow approximation model in power grids.

$$\mathcal{O}(u) = \sum_{v \in \mathcal{N}_u^+} f(u, v), \quad \forall u, v \in \mathcal{V} \tag{2}$$

$$\mathcal{I}(u) = \sum_{v \in \mathcal{N}_u^-} f(u, v), \quad \forall u, v \in \mathcal{V}, \tag{3}$$

where $\mathcal{N}_u^+$ and $\mathcal{N}_u^-$ are the sets of outgoing neighbors and incoming neighbors of u, respectively; likewise, $\mathcal{O}(u)$ and $\mathcal{I}(u)$ represent the total of outgoing power flow and total of incoming power flow of node u, respectively. These equation lead the following recurrence node u with $\mathcal{O}(u)$ and $\mathcal{I}(u)$.

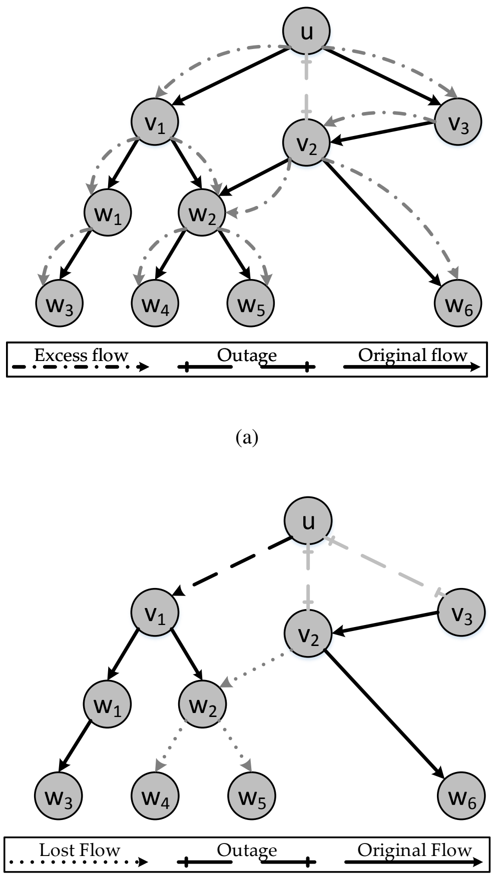
Case I: $\mathcal{O}(u) - \mathcal{I}(u) > 0 \quad \forall u \in \mathcal{V}$. In this case, the total of outgoing power flow of node u is greater than its total of incoming power flow. The power output p_u at node u is not negative and thus node u is a generator $(u \in G)$.

Case II: $\mathcal{O}(u) - \mathcal{I}(u) < 0 \quad \forall u \in \mathcal{V}$. In this case, the total of outgoing power flow of node u is smaller than its total of incoming power flow. The power output p_u at node u is negative and thus node u is a consumer $(u \in C)$.

Case III: $\mathcal{O}(u) - \mathcal{I}(u) = 0 \quad \forall u \in \mathcal{V}$. In this case, the total of outgoing power flow of node u is equal to its total incoming power flow, the power output p_u at node u is equal to zero, thus, the role of node u is a transformer station $(u \in T)$.

C. The Interplay of Transmission Line Failure

A power grid is designed such that generators, consumers, transformers, and lines work below under their capacity. A cascading failure effect can occur in different denominators.



(a)

(b)

Fig. 1: The power flow redistribution procedure. (a) A cascading redistribution. (b) The system changes after redistribution process.

In this paper, we consider one of the common denominators, wherein the failure of certain lines can lead to the failure of other lines due to the interdependency with previous failed line/lines. A transmission line (u, v) fails, if the power flow $f(u, v)$ through the line increases beyond its maximum capacity $\delta(u, v)$. Once this happens, its power flow $f(u, v)$ is redistributed to u's other outgoing neighbors. Each line (u, w), $\forall w \in \mathcal{N}_u^+ \setminus \{v\}$ will receive an additional power flow $f(u, w)^+$ which is proportional to weight $\epsilon(u, w)$ of the line (u, w) from u to w, thus, defined as follows:

$$f(u,w)^+ = f(u,v) \times \frac{\epsilon\,(u,w)}{\sum_{t \in N_u^+ \setminus \{v\}} \epsilon\,(u,t)}, \tag{4}$$

Due to the power flow redistribution, the total of incoming power flow of node w is elevated to $\mathcal{I}(w) + f(u, w)^+$. When the total of incoming power flow of a node increases, the total of outgoing power flow of the node increases as well. The total of outgoing power flow of terminal node decides the power flow of lines that connect the node to its outgoing neighbors. When the power flow of some lines exceed their maximum capacity, hence, fail in the next time step. Meanwhile, due to the failure of line (u, v), the total of incoming power flow of node v is deducted an amount $f(u, v)$, resulting in decreasing total of incoming power flow of nodes in v's outgoing neighbors (N_v^+). Then, we have the total of incoming power flow of $z \in \mathcal{N}_v^+$ is $\mathcal{I}(z) - f(u, v)$.

Take fig. 1(a) for example, link (u, v_2) fails, its load is redistributed to u's other neighbors v_1 and v_3. The line (u, v_1) will receive an additional power flow $f(u, v_1)^+ = f(u, v_2) \times {}^{\epsilon(u,v_1)}/_{\epsilon(u,v_1)+\epsilon(u,v_3)}$. Likewise, we have $f(u, v_3)^+ = f(u, v_2) \times {}^{\epsilon(u,v_3)}/_{\epsilon(u,v_1)+\epsilon(u,v_3)}$. Meanwhile, due to the line (u, v_2) fails, the total of incoming power flow of node v_2 is deducted an amount $f(u, v_2)$. However, it also receives an additional power flow from (v_3, v_2) due to an increment of the power flow of the link (u, v_3). Then, the total of incoming power flow of node v_2 is $\mathcal{I}(v_2) - f(u, v_2) + f(u, v_3)^+$. The process will be propagated to other nodes in time steps.

D. Attack model

The occurrence of initial transmission lines (links)'s failure can be triggered by false data injection attacks wherein attackers inject some false data into the measurement system. Consequently, the circuit breaker will be activated automatically by operation center to protect other transmission lines. The failures are propagated from initial failed links to other links in time steps. The redistribution process will stop at the end when there are no more failed nodes. To assess the vulnerability of a smart grid system more practically under the cascading failure potential, we consider the following factors:

- Power grid design in fact indicates that it is a huge system and there are situations wherein the effort of attackers is insufficient to take all users down or to make a large power system blackout initially due to lack of resources (limited budget). **[Not clear to me - Attacker takes down the transmission lines, not the users..]****[will come back to address this]** In this paper, we consider the attack budget factor, denoted by $\mathcal{B}_{ATK}$. Likewise, each transmission line (u, v) has a robustness factor $\tau(u, v)$ and a higher value $\tau(u, v)$ reflects the robustness of the line (u, v). From an attacker perspective, this is also the cost that they need to spend to take down the line.
- Also, in reality, there are particular regions or users (consumers) that need to be power supplied consistently. Therefore, users must be weighted by their importance. Higher weights are given to more important targets. **[What targets?]** Let $\omega(u)$ denote the weight of an user u.

The idea behind our proposed attack algorithm in this paper is to maximize damages to smart grids that are expressed through the total impact on users (consumers). These requirements make the problem practical, but also difficult. This motivates us to study a more general and practical problem in terms of attack and defense.

E. Problem definition

In this paper, while given a smart grid represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$ wherein each line (u, v) has a power flow (load) $f(u, v)$, a capacity $\delta(u, v)$, a cost $\tau(u, v)$, each node u has a weight $\omega(u)$, and an attack budget

$\mathcal{B}_{ATK}$, the investigated problem is to find an ordered set of links $\mathcal{ATK}$ such that this ordered failures of links in $\mathcal{ATK}$ maximizes the total impact on consumers (we define and discuss the total impact on users in detail in section III-A), referred to as the Maximum-Impact through Critical-Line with Limited Budget (MICLLB) problem. The MICLLB problem is formally illustrated as follows:

INSTANCE: Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$, an attack budget $\mathcal{B}_{ATK}$, and a constant $k \in \mathbb{Z}^{+}$.

QUESTION: Does there exist an ordered set of lines (links) $\mathcal{ATK}$, that is, a sequence of lines in $\mathcal{ATK}$, for attacks, such that the total impact on consumers/users is not less than k?

Theorem 1: The MICLLB problem is NP-complete.

Details of the difficulty of the MICLLB problem are discussed in detail in Section IV-A.

III. Strategic Algorithm

In this section, we first introduce an attack algorithm, namely Greedy Based Partition Algorithm (GBPA) for solving MICLLB problem in section III-A. Then, an appropriate algorithm for defense, namely Homogeneous-Equality Based Defense Algorithm (HEBDA) is proposed in section III-B.

A. Greedy Based Partition Algorithm (GBPA)

To evaluate the vulnerability of a smart grid system, most of previous works focus on characterizing the higher load nodes or lines to trigger a cascading failure effect that may cause further failures. However, these approaches may miss out on a critical point that smart grid is not only vulnerable to cascading failures, but also to the unbalanced power distribution. Thus, We comprehensively consider the power grid that is vulnerable to both cascading failure attacks and normal attacks.

*1) **Cascading failure attack**:* A cascading failure that is triggered (*decided or emphasized*) by the right initial failures, is not like the failures occurred at high degree nodes or high load links that may lead to a large failure of its neighbors but does not expose a cascading effect. Therefore, we pay more attentions on determining critical lines and raking their critiation as follows.

- Non-critical link: Let $\mathcal{UC}$ denote the set of non-critical links. A non-critical link $(u, v) \in \mathcal{UC}$ is a link whose failure does not affect any users. This happens due to the instability of the system or previously attacks' impact, or the link is not directly (or indirectly) connected to any users.
- Ranking component's critiation: There is a way to make a large blackout that is *cutting-off* all outgoing power flow from generators. However, recall that there is a limited budget to attacks, so that determining the right target lines decides the scale of destruction from attacker perspective. Essentially, the redistribution power is a process that power flow is redirected by stations (nodes), we will focus on determining the most critical outgoing links between many outgoing links of many nodes.

Considering the interdependency between nodes and links, and between links, we need to determine the highly vulnerable nodes, in order to determine the most critical links. Recall that a node u that is a consumer ($\mathcal{O}(u) - \mathcal{I}(u) < 0$), is weighted by its importance $\omega(u)$. It is obvious that providing good services to these important consumers are the first priority of the provider since they are critical and potential customers. Let $\beta(u)$ denote the satisfactory factor of the consumer u. The unsatisfactory factor is estimated based on the total supplied power in the incoming links and the demand of the consumer.

$$\beta\left(u\right) = \frac{d_u - \left|\mathcal{I}\left(u\right) - \mathcal{O}\left(u\right)\right|}{d_u} \tag{5}$$

Equation 6 indicates that the higher $\beta(u)$ is, the less satisfactory of the consumer u will be. Intuitively, those nodes with a higher weight and a higher unsatisfactory factor will potentially be targets for attackers. Thus, we introduce a metric representing for the *impact on consumer* ($\forall u \in C$), denoted by $\gamma(u)$.

$$\gamma\left(u\right) = \omega\left(u\right) \times \beta\left(u\right) \tag{6}$$

However, from the attacker's point of view, selecting a target to attack based on only the *impact on consumer* (γ) may require a higher cost, which may rise a much higher total spending costs but increase little total impact. To avoid this scenario, we define a new metric for every consumer $u \in C$, referred to as the *average impact value per incremental cost* of u, denoted by $\mathcal{R}(u)$. The $\mathcal{R}(u)$ is defined as follows:

$$\mathcal{R}(u) = \frac{\gamma\left(u\right)}{{}^{\mathcal{O}(u)}/_{\lambda(u)}}, \tag{7}$$

where $\lambda(u)$ is the total costs needed to trigger a cascading failure at node u. Obviously, in worst case, all outgoing links of u need to be taken down:

$$\lambda\left(u\right) = \sum_{v \in N_u^{+}} \tau\left(u, v\right) \tag{8}$$

However, in fact, we need to consider the co-impact of initial attacked links to trigger a large failure. Thus, the failure of a few outgoing links of node u is sufficient to take down all of its outgoing links. This happens if the total loads of earlier attacked link/links is higher than the capacity of the remaining links.

$$\delta\left(u, v\right) \leq \sum_{t \in N_u^{+} \setminus \{v\}} f\left(u, t\right) \tag{9}$$

Take fig. 1(b) for example. In this example, we have $\delta\left(u, v_1\right) \leq f\left(u, v_2\right) + f\left(u, v_3\right)$. Then, the failures of (u, v_2) and (u, v_3) cause the overload at (u, v_1), resulting in the failure of (u, v_1). This implies that there is no need to attack all outgoing links to isolate a node. If there exists the co-impact that can trigger a large failure to completely isolate a node u, then the total costs $\lambda(u)$ can be reduced as below:

$$\lambda\left(u\right) = \sum_{v \in N_u^{+}} \tau\left(u, v\right) - \sum \tau\left(u, t\right), \tag{10}$$

$\forall v, t \in N_u^{+} / \sum f\left(u, v\right) > \sum \delta\left(u, t\right)$ and $(u, t) \notin \mathcal{UC}$. Thus, the the *average impact value per incremental cost* of nodes is evaluated as shown in procedure 1 (EVOC).

Procedure 1 EVOC ($\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$)

```
for each vertex $u \in \mathcal{V}$ do
    $\lambda(u) \leftarrow 0$
    $total \leftarrow 0$
    for each vertex $v \in \mathcal{N}_u^+$ do
        if $(u, v) \notin \mathcal{UC}$ then
            $total \leftarrow \lambda(u) + \tau(u, v)$
        end if
    end for
    $min \leftarrow total$; $target \leftarrow null$
    if $\sum f(u, v) > \sum \delta(u, t)$ $(\forall v, t \in N_u^+)$ then
        for each $(u, t) \notin \mathcal{UC}$ do
            if $total - \tau(u, t) < min$ then
                $min \leftarrow total - \tau(u, t)$
                $target \leftarrow t$
            end if
        end for
    end if
    $\lambda(u) \leftarrow min$
    $\mathcal{R}(u) \leftarrow \frac{\gamma(u)}{\mathcal{O}(u)/\lambda(u)}$
end for
```

After evaluating the the *average impact value per incremental cost* of all nodes, we can determine the target links for a cascading failure attack. Afterward, a new graph with remaining nodes and links (after each batch of the cascading failure) is updated as well.

*2) **Unbalanced load distribution attack**:* The *average impact value per incremental cost* of nodes can help determine the target nodes and then links to attack with a cascading failure effect, however, it is not practical if attack budget is not sufficient to rise up a cascading failure effect. As mentioned, in case the budget is sufficient, to isolate a node we can attack a certain number of its outgoing links to trigger a cascading failure effect and make a larger blackout. However, in case the attacking resources and budget are insufficient, a tricky problem behind the interplay of links is exposed. Let's consider a power grid as shown in fig. 1(a). The failure of (u, v_2) triggers the power flow redistribution happened at u. It seems that (u, v_2)'s failure impacts on the users w_4, w_5, and w_5, but it is not. Because (u, v_2)'s power flow is redistributed to v_1 and v_3, and keeps serving w_4, w_5, and w_5. Then its failure impact is negligible. We call this phenomenon as the reverted flow.

To handle this situation, we consider the second attack case. In case of lack of attack budget, we first classify sub-regions in the power grid into *satisfied region* and *unsatisfied region*. Let $\mathcal{T}(v)$ be the Depth-First-Search (DFS) tree rooted at v. Also, let $\mathcal{D}_{\mathcal{T}(v)}$ and $\mathcal{S}_{\mathcal{T}(v)}$ denote the total demands of nodes and the total power supplies to all nodes in $\mathcal{T}(v)$, respectively.

- Unsatisfied region: is the region with lack of power supplies ($\mathcal{D}_{\mathcal{T}(v)} \geq \mathcal{S}_{\mathcal{T}(v)}$)
- Satisfied region: is the region without lack of power supplies ($\mathcal{D}_{\mathcal{T}(v)} \leq \mathcal{S}_{\mathcal{T}(v)}$)

The idea behind the classification is to direct the attacks to *unsatisfied regions* that originally suffers from the outage. Then, the effect of the early outage is strengthened by the later attacks. The details of proposal algorithm is in Algorithm 1.

Algorithm 1 GBPA ($\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega), \mathcal{R}, \mathcal{B}_{ATK}$)

```
Call: EVOC ($\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$) ▷ Compute $\mathcal{R}(u)$ for every node $u \in \mathcal{V}$
Let $\mathcal{C}$ be the set of nodes in descending critication ranking ordering
Let $\Psi$ be the total costs has been used to attack.
Let $\mathcal{ATK}$ be the ordered set of links
Let $\mathcal{T}$ be total weights of failed nodes
$\mathcal{ATK} \leftarrow \emptyset$; $\Psi \leftarrow 0$; $\mathcal{T} \leftarrow 0$
while $\Psi \leq \mathcal{B}_{ATK}$ do
    for each vertex $u \in \mathcal{C}$ do
        if $\lambda(u) \leq \mathcal{B}_{ATK} - \Psi$ then                ▷ Enough budget
            for each vertex $v \in \mathcal{N}_u^+$ do
                if $(u, v) \notin \mathcal{UC}$ and $\sum f(u, v) > \sum \delta(u, t)$ $(\forall v, t \in N_u^+)$ then
                    $\mathcal{ATK} \leftarrow \mathcal{ATK} \bigcup (u, v)$ )
                    $\mathcal{T} \leftarrow \mathcal{T} + \omega(u)$
                    Update a new graph and remove all failed nodes in $\mathcal{G}$ with the failure of node $u$
                end if
            end for
        end if
    end for
    for each vertex $u \in \mathcal{C}$ do
        if $\lambda(u) \geq \mathcal{B}_{ATK} - \Psi$ && $\exists$ $v \in \mathcal{N}_u^+ / (u, v) \in \mathcal{UC}$ || $\mathcal{D}_{\mathcal{T}(v)} \geq \mathcal{S}_{\mathcal{T}(v)}$ then      ▷ Lack budget
            for each vertex $v \in \mathcal{N}_u^+$ do
                if $\tau(u) \leq \mathcal{B}_{ATK} - \Psi$ \$\$ $\mathcal{D}_{\mathcal{T}(v)} \geq \mathcal{S}_{\mathcal{T}(v)}$ then
                    $\mathcal{ATK} \leftarrow \mathcal{ATK} \bigcup (u, v)$ )
                    $\mathcal{T} \leftarrow \mathcal{T} + \omega(u)$
                    Update a new graph and remove all failed nodes in $\mathcal{G}$ with the failure of node $u$
                end if
            end for
        end if
    end for
end while
return $\mathcal{ATK}$, $\mathcal{T}$
```

B. Homogeneous-Equality Based Defense Algorithm(HEBDA)

In Section III-A, the GBPA is proposed to find an ordered set of links $\mathcal{ATK}$ such that the impact (total weights) on users from failures of link in $\mathcal{ATK}$ is maximized. To protect the system, in this section we design a defense algorithm to help determine the links that need to be secure against GBPA, referred to as Homogeneous-Equality Based Defense Algorithm (HEBDA). Because there is a limited attacking budget $\mathcal{B}_{ATK}$, we also consider a limited defending budget, namely $\mathcal{B}_{DEF}$. To use the budget efficiently, we introduce a metric, referred to as the homogeneous equality.

$$\mathcal{HE} = \left\lceil \frac{\mathcal{B}_{ATK} + \sum\limits_{(u,v) \in \mathcal{ATK}} \tau(u, v)}{|\mathcal{ATK}|} \right\rceil \tag{11}$$

The idea behind the HEBDA is to secure critical links in $\mathcal{ATK}$ with more costs. Since the links' cost is elevated, it is more robust and becomes harder to be taken down under attacks. However, due to the limited defending budget, priority is given to the most critical links to be secure. The details of HEBDA are shown in Algorithm 2.

Algorithm 2 HEBDA ($\mathcal{G}, \mathcal{ATK}, \mathcal{B}_{ATK}, \mathcal{B}_{DEF}$)

```
Evaluate $\mathcal{HE}$

$\mathcal{HE} \leftarrow \left\lceil \frac{\mathcal{B}_{ATK} + \sum_{(u,v)\in\mathcal{ATK}} \tau(u,v)}{|\mathcal{ATK}|} \right\rceil$

if $\mathcal{HE} \geq \mathcal{B}_{ATK}$ then
    $\mathcal{HE} = \mathcal{B}_{ATK} + 1$
end if
for each link $(u,v) \in \mathcal{ATK}$ do
    if $\tau(u,v) \leq \mathcal{HE}$ then
        if $\mathcal{HE} - \tau(u,v) \leq \mathcal{B}_{DEF}$ then
            $\tau(u,v) \leftarrow \mathcal{HE}$
            $\mathcal{B}_{DEF} \leftarrow \mathcal{B}_{DEF} - (\mathcal{HE} - \tau(u,v))$
        end if
    end if
end for
while $\mathcal{B}_{DEF}$ do
    Let $(u,v)$ be the link with the lowest cost.
    $\tau(u,v) \leftarrow \tau(u,v) + 1$
    $\mathcal{B}_{DEF} \leftarrow \mathcal{B}_{DEF} - 1$
end while
```

IV. Analysis

A. Time Complexity of Strategic Algorithms

Theorem 2: The time complexity of the GBPA is bounded in $O(nm + nlogn + n^3)$, where n, and m denote, respectively, the number of nodes, and the number of links.

Proof: Suppose there are at most n nodes and m links in the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau, \omega)$, where $|\mathcal{V}| = n$ and $\mathcal{E} = m$. In the EVOC procedure, it takes at most $O(m)$ to compute the critication ranking for each node, and therefore it takes $O(nm)$ to compute the critication ranking for n nodes in $\mathcal{V}$. In the first two steps of the GBPA, it requires $O(nlogn)$ to sort (heap sort) the set $\mathcal{C}$ in descending critication ranking ordering. In the while loop of the GBPA, it is clear that at most n nodes are considered. In each iteration for the *enough budget case*, it requires at most $O(n^2)$ to find a suitable node to attack, and in each iteration for the *lack budget case*, it also requires at most $O(n^2)$ to find a suitable node to attack. The while loop therefore requires at most $O(n^3)$. The time complexity of the GBPA is bounded in $O(nm + nlogn + n^3)$, which completes the proof. ■

Theorem 3: The time complexity of the HEBDA is bounded in $O(m)$, where m denotes the number of links.

Proof: Because at most m links are selected in $\mathcal{ATK}$, the for loop requires at most m iterations. In addition, because the next while loop requires at most $O(m)$ to find the link with the lowest cost, we have the time complexity of the HEBDA is bounded in $O(m)$, which thus completes the proof. ■

B. The Hardness of the MICLLB Problem

In this subsection, we show the hardness of the MICLLB problem by first showing that a sub-problem of the MICLLB problem, referred to as the Reduced Maximum-Impact through Critical-Line with Limited Budget (RMICLLB) problem, is NP-hard. The RMICLLB problem is the MICLLB problem with $\omega(u) = 1$ ($\forall u \in \mathcal{V}$); that is, each node has an equal weight to others. The problem is illustrated as follows:

INSTANCE: Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, f, \delta, \tau)$, and an attack budget $\mathcal{B}_{ATK}$, and a constant $k \in \mathbb{Z}^+$.

QUESTION: Does there exist an ordered set of lines (links) $\mathcal{ATK}$, that is, a sequence of lines in $\mathcal{ATK}$, for attacks, such that the total number of failed nodes is not less than k?

Lemma 1 is used to show that the RMICLLB problem is NP-hard. We first use the set cover problem [31] to show the hardness of the RMICLLB problem. Then, the hardness of the MICLLB problem is provided in Theorem 1. The set cover problem is illustrated as follows:

INSTANCE: Given a set X, and a collection C of X, where $|X|$ denotes the size of X.

QUESTION: Does C have an exact cover for X, that is, a subset $C' \subseteq C$ such that x is contained in exactly one member of C' for each $x \in X$?

For example, let $X = \{1, 2, 3, 4, 5, 6\}$ and $C = \{C_1, C_2, C_3\}$, where $C_1 = \{1, 2, 4, 6\}$, $C_2 = \{2, 5, 6\}$, and $C_3 = \{3, 5\}$. In this example, C has an exact cover $C' = \{C_1, C_3\}$ for X.

Lemma 1: The RMICLLB problem is NP-hard.

Proof: In the set cover problem, each member $x \in X$ can be treated as a node $t \in \mathcal{V}$; each member $c \in C$ can be treated as an attacked link (u, v); a node $t \in \mathcal{V}$ fails due to the failure of the link (u, v) if $x \in X$ is included in $c \in C$; and k is set to $|X|$. Because each $x \in X$ is contained in exactly one member of C', and a node $t \in \mathcal{V}$ is failed by the cascading failure caused by one attack, the set cover problem is also an RMICLLB problem. We have that the set cover problem is a sub-problem of the RMICLLB problem. Because the set cover problem is NP-hard [31], the RMICLLB problem is also NP-hard. This completes the proof. ■

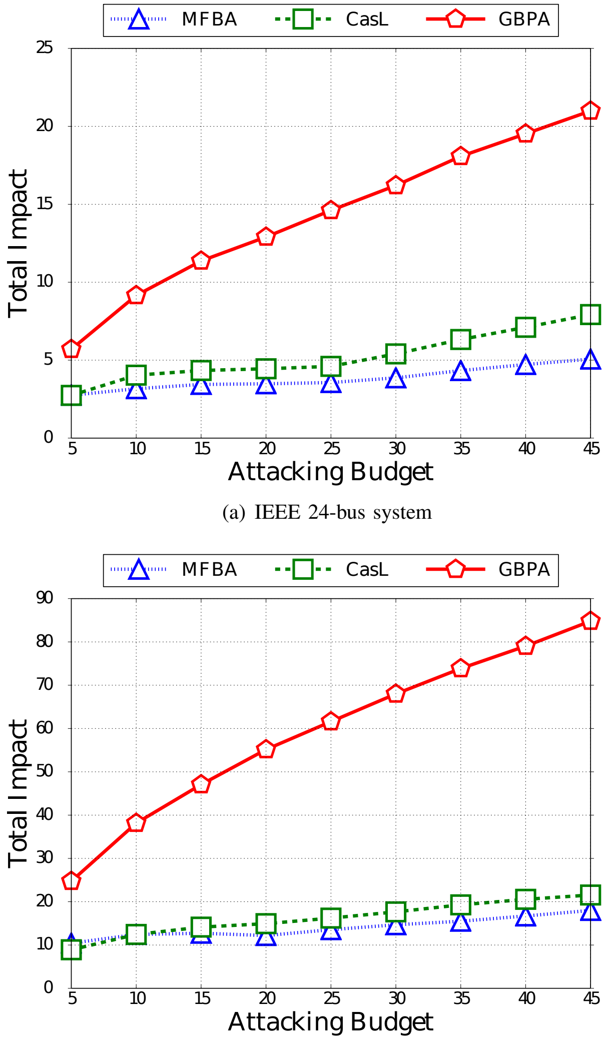
Theorem 1: The MICLLB problem is NP-complete.

Proof: By Lemma 1, because the RMICLLB problem, which is a sub-problem of the MICLLB problem, is NP-hard, the MICLLB problem is NP-hard. In addition, it is clear that the MICLLB problem belongs to the NP class. We therefore have that the MICLLB problem is NP-complete. ■

V. Performance Evaluation

In this section, we conduct simulations to evaluate the performance of the proposed algorithms performed using C++ and PYPOWER [32]. In the simulations, two benchmark power systems are used to generate the power grid topologies: IEEE 24-bus system [33] with 24 vertices and 38 edges and IEEE 118-bus system [34] with 118 vertices and 177 edges. The cost of links are randomly selected within the interval [1,μ] ($\mu \in \mathbb{N}$). We set up two scenarios to test the effect of the cascading failures caused by attack algorithm GBPA as well as to test the performance of the defense algorithm, HEBDA.

a) Attacking scenario: We simulate the GBPA and two other attack strategies, the Maximum-Flow Based Algorithm (MFBA) [24] and the Cascade Potential Ranking Lines Algorithm (CasL) [25], and then compare them in terms of the total weights of failed nodes. In the MFBA, a virtual sink vertex and a virtual source vertex are added into the network and connected with each sink vertex (i.e. the vertex whose out-neighborhood is empty) and each source vertex (i.e. the



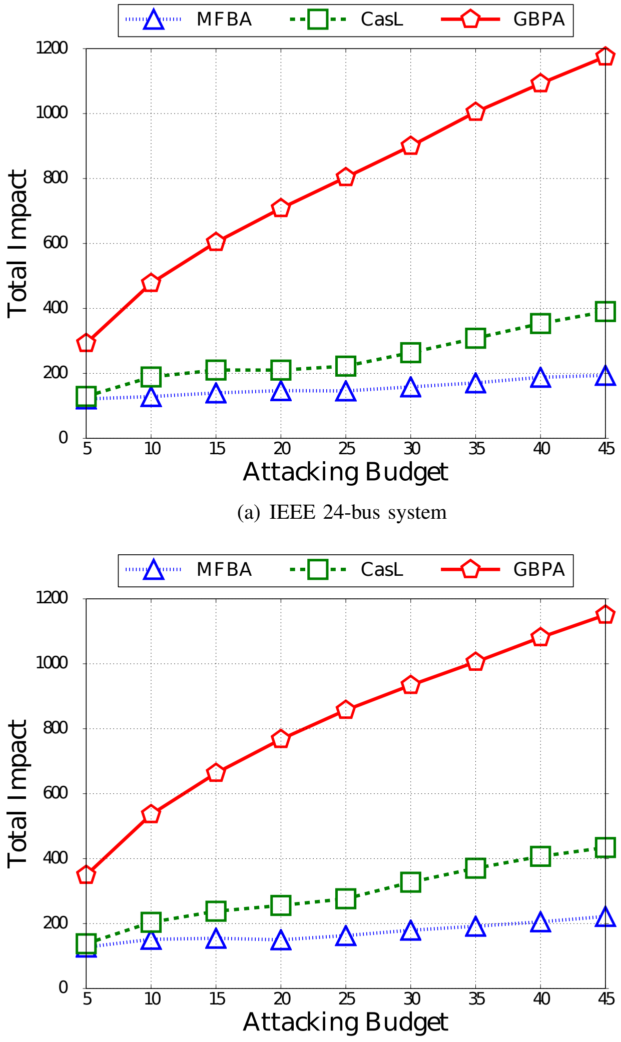
(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 2: Total impacts achieved by the MFBA, the CasL, and the GBPA with $\mathcal{B}_{ATK}$ ranging from 5 to 45, where $\mu = 25$ and users' importance are randomly selected within the interval [1,5] in different power systems.

vertex whose in-neighborhood is empty). After that, an index for vulnerability assessment, referred to as the centrality index, is calculated as the ratio of current flow by the maximum flow. The centrality index is considered as the link vulnerability index in the MFBA. The links with higher centrality are used more frequently, and thus, they are more vulnerable. In the CasL, a similar index is used for vulnerability assessment, namely the cascading potential, that is counted by the numbers of failed links caused by the initial failure. A link with a higher cascading potential is more vulnerable since it tends to a higher number of failures.

Fig. 2(a) and fig. 2(b) show the total weights of failed users under different attacking budgets with $\mu = 25$ in different power systems, where the users' importance are selected randomly within the interval [1,5]. It is obvious that the GBPA demonstrates a better performance than the MFBA and the CasL. This is because in the GBPA, each node is evaluated by the critication ranking that help select the right target links and manage the attacking resources well to attack efficiently.Although other algorithms aim to increase the number of failed links, their strategies fail to select and attack the right components, which results in a lower impact on further failures.



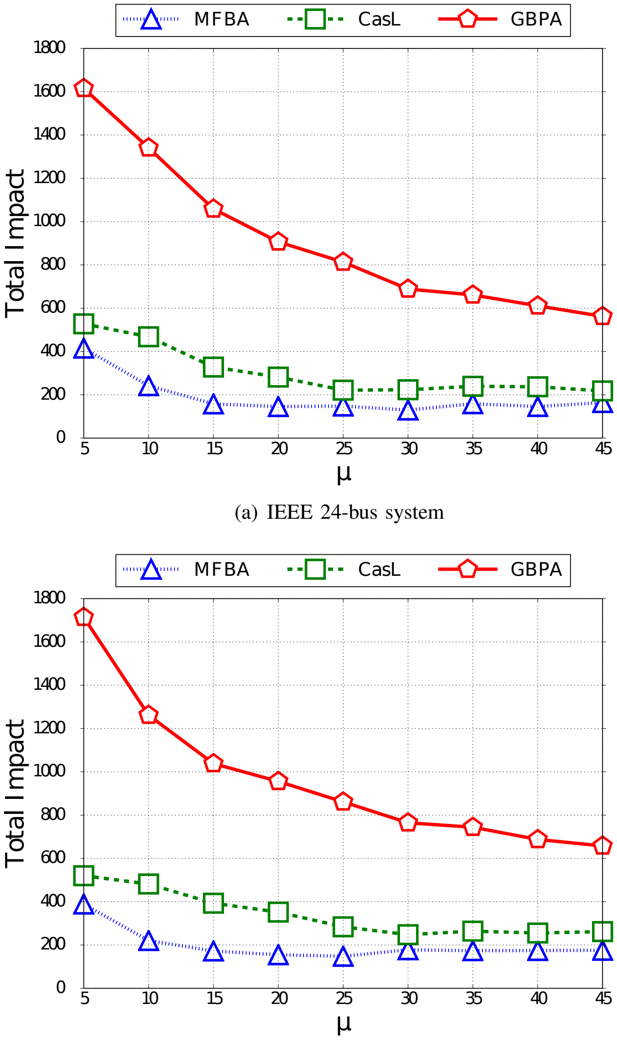
(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 3: Total impacts achieved by the MFBA, the CasL, and the GBPA with $\mathcal{B}_{ATK}$ ranging from 5 to 45 where $\mu = 25$ and users' importance are selected as its demand in different power systems.

Fig. 3(a) and fig. 3(b) show the simulation results concerning the total weights of failed users with the users' importance are set as their demand. It is clear that the more attacking budgets, the more weights of failed users caused by three algorithms. Obviously, the more available attacking resources, the more links can be taken down, and thus, the more impacts on users. Similar to the results in fig. 2, the GBPA provides a better performance than the MFBA and the CasL.

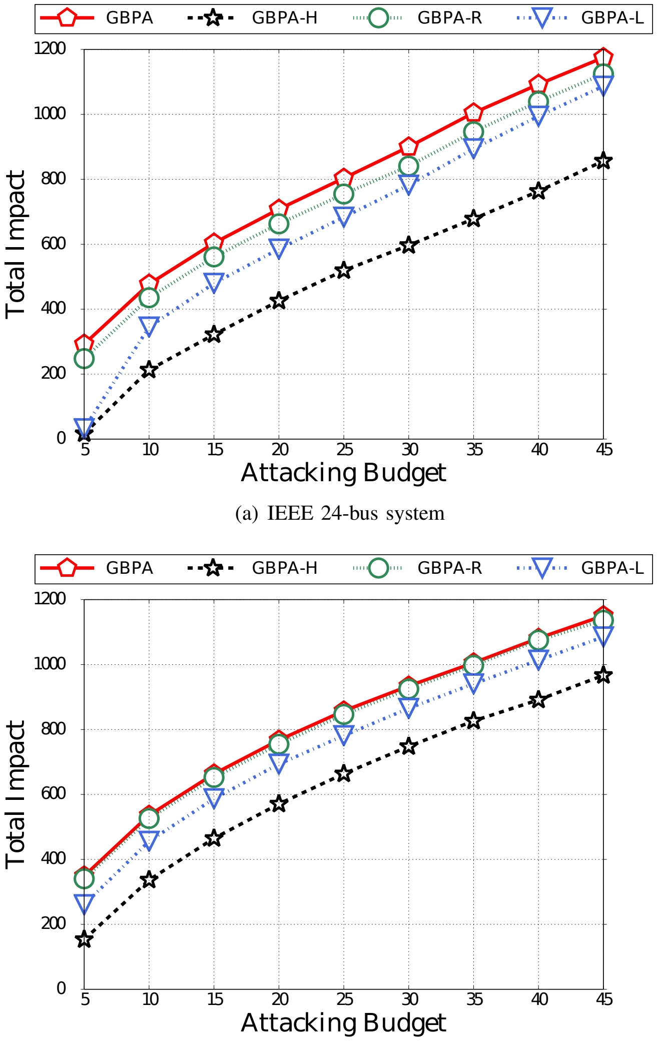
From fig. 4(a) and fig. 4(b), we observe that the more μ, the more average costs secured on links, and therefore, the system is more robust. That is why the total weights of failed users in all algorithms is low when the μ is increased. However, since



(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 4: Total impacts achieved by the MFBA, the CasL, and the GBPA with μ ranging from 5 to 45 when $\mathcal{B}_{ATK} = 25$ in different power systems.



(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 5: Total impacts achieved by the HEBDA and the baseline algorithms against the GBPA with $\mathcal{B}_{ATK}$ ranging from 5 to 45 where $\mu = 25$ and $\mathcal{B}_{DEF} = 25$ in different power systems.

the μ is varied, the GBPA is still an efficient attack and the performance of GBPA is then better than the others.

b) Defending scenario: Here, the random defense algorithm (RDA) and the lowest robustness defense algorithm (LRDA) are introduced as the baseline algorithms to compare with the HEBDA in terms of the total weights of failed users. The RDA is designed by randomly selecting links to secure more costs. On the other hand, the LRDA is designed to secure the links that have the lowest cost fist. In section V-0a, the GBPA demonstrates with the best performance between three algorithms, the following simulations, the HEBDA, RDA, and LRDA are implemented in response to the GBPA, namely the GBPA-H, GBPA-R, and GBPA-L, respectively.

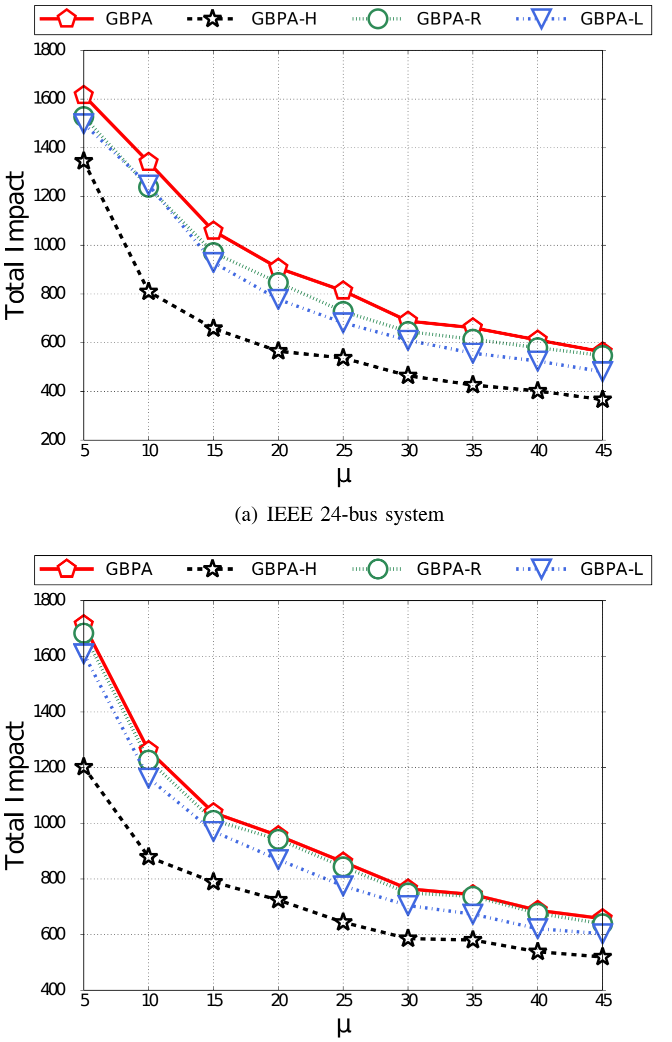
Fig. 5(a) and fig. 5(b) show the simulation results in terms of the total weights of failed users under different $\mathcal{B}_{ATK}$ with $\mu = 25$ and $\mathcal{B}_{DEF} = 25$. It is clear that the GBPA-H has a better performance on reducing the total damage compared to the baseline algorithms.

Fig. 6(a) and fig. 6(b) show the comparison of the total weights of failed users between the GBPA-H, GBPA-R, GBPA-L, and GBPA with the μ is varied, $\mathcal{B}_{ATK} = 25$ and $\mathcal{B}_{DEF} = 25$. Obviously, the GBPA-H provides a better performance than the others in terms of the reduction of total damage caused by the GBPA.

Fig. 7(a) and fig. 7(b) show similar results as in fig. 6. That is, the GBPA-H outperforms the GBPA-R, and the GBPA-L, and the total weights of failed users caused by the GBPA decreases with increasing the defending budget.

VI. Conclusion

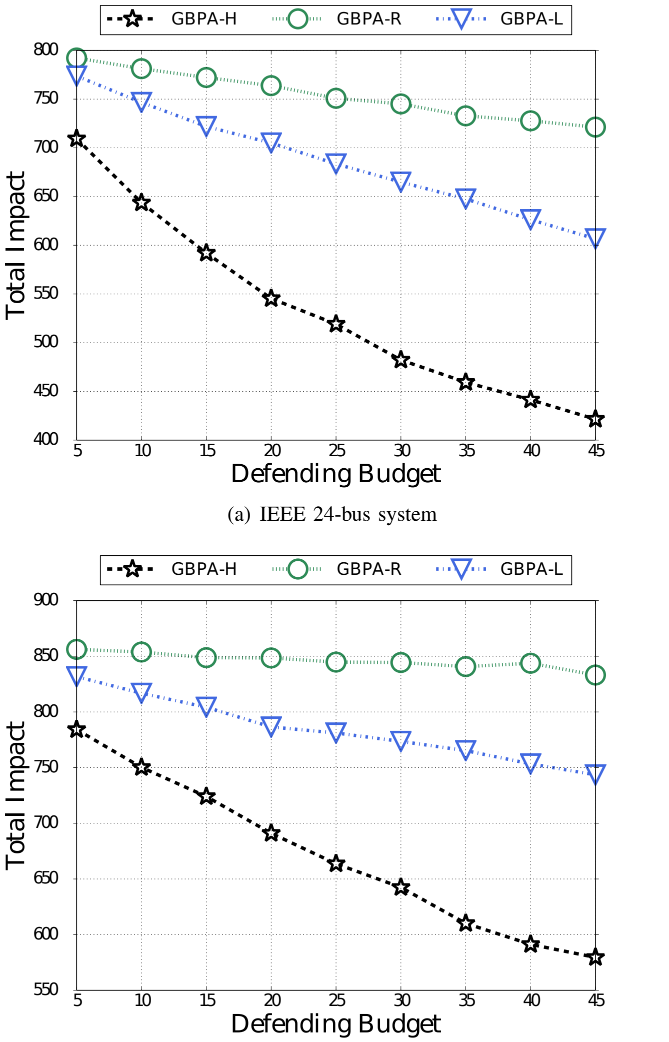
In this paper, we have investigated the problem of designing an efficient attack strategy that maximizes the total impact on users/consumers in smart grids, termed, Maximum-Impact through Critical-Line with Limited Budget (MICLLB) problem. We then introduced a new metric to range the critical components and used critication ranking to establish the ordered set of attacks ($\mathcal{ATK}$) for the MICLLB problem.



(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 6: Total impacts achieved by the HEBDA and the baseline algorithms against the GBPA with μ ranging from 5 to 45 where $\mathcal{B}_{ATK} = 25$ and $\mathcal{B}_{DEF} = 25$ in different power systems.



(a) IEEE 24-bus system

(b) IEEE 118-bus system

Fig. 7: Total impacts achieved by the HEBDA and the baseline algorithms against the GBPA with $\mathcal{B}_{DEF}$ ranging from 5 to 45 where $\mu = 25$ and $\mathcal{B}_{ATK} = 25$ in different power systems.

We proved that MICLLB problem is NP-complete, and subsequently, proposed an efficient attack algorithm (GBPA) based on the obtained $\mathcal{ATK}$ that maximizes the total impact of attacks on users. In addition, an appropriate defense algorithm (HEBDA) has been designed to secure the critical components against the GBPA. Extensive simulation results demonstrate that both GBPA and HEBDA better performance with a higher attack impact on users than the most recently related attack methods.

Acknowledgment

This work was supported by the Purdue Research Foundation under the 2020 Summer Faculty Grant.

References

[1] Vaiman, Bell, Chen, Chowdhury, Dobson, Hines, Papic, Miller, and Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Transactions on Power Systems*, vol. 27, pp. 631–641, 2012.

[2] G. Zhang, Z. Li, B. Zhang, and W. A. Halang, "Understanding the cascading failures in indian power grids with complex networks theory," *Physica A: Statistical Mechanics and its Applications*, vol. 392, pp. 3273–3280, 2013.

[3] Y. Zhu, J. Yan, Y. Sun, and H. He, "Revealing cascading failure vulnerability in power grids using risk-graph," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, pp. 3274–3284, 2014.

[4] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with dc power flow model and transient stability analysis," *IEEE Transactions on Power Systems*, vol. 30, pp. 285–297, 2015.

[5] P. Dey, R. Mehra, F. Kazi, S. Wagh, and N. M. Singh, "Impact of topology on the propagation of cascading failure in power grid," *IEEE Transactions on Smart Grid*, vol. 7, pp. 1970–1978, 2016.

[6] P. D. H. Hines, I. Dobson, and P. Rezaei, "Cascading power outages propagate locally in an influence graph that is not the actual grid topology," *IEEE Transactions on Power Systems*, vol. 32, pp. 958–967, 2017.

[7] A. G. Phadke and J. S. Thorp, "Expose hidden failures to prevent cascading outages [in power systems]," *IEEE Computer Applications in Power*, vol. 9, no. 3, pp. 20–23, July 1996.

[8] R. Leelaruji and V. Knazkins, "Modeling adequacy for cascading failure analysis," in *2008 Australasian Universities Power Engineering Conference*, Dec 2008, pp. 1–6.

[9] J. Seo, S. Mishra, X. Li, and M. T. Thai, "Catastrophic cascading failures in power networks," *Theoretical Computer Science*, vol. 607, pp. 306–

319, 2015.

[10] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power and Energy Magazine*, vol. 3, pp. 34–41, 2005.

[11] H. Farhangi, "The path of the smart grid," *IEEE Power and Energy Magazine*, vol. 8, pp. 18–28, 2010.

[12] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid—the new and improved power grid: A survey," *IEEE Communications Surveys Tutorials*, vol. 14, pp. 944–980, 2012.

[13] T. Roth and B. McMillin, "Physical attestation in the smart grid for distributed state verification," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2017.

[14] G. Locke and P. D. Gallagher, "Nist framework and roadmap for smart grid interoperability standards, release 1.0," *National Institute of Standards and Technology*, vol. 33, 2010.

[15] Y. Zhang, L. Wang, Y. Xiang, and C. W. Ten, "Power system reliability evaluation with scada cybersecurity considerations," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1707–1721, 2015.

[16] F. Li, Y. Wei, and S. Adhikari, "Improving an unjustified common practice in ex post lmp calculation," *IEEE Transactions on Power Systems*, vol. 25, pp. 1195–1197, 2010.

[17] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, pp. 1630–1638, 2017.

[18] 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, pp. 1731–1738, 2012.

[19] X. Liu and Z. Li, "Local load redistribution attacks in power systems with incomplete network information," *IEEE Transactions on Smart Grid*, vol. 5, pp. 1665–1676, 2014.

[20] D. H. Choi and L. Xie, "Economic impact assessment of topology data attacks with virtual bids," *IEEE Transactions on Smart Grid*, pp. 1–9, 2017.

[21] X. Liu, Z. Bao, D. Lu, and Z. Li, "Modeling of local false data injection attacks with reduced network information," *IEEE Transactions on Smart Grid*, vol. 6, pp. 1686–1696, 2015.

[22] M. Esmalifalak, H. Nguyen, R. Zheng, L. Xie, L. Song, and Z. Han, "A stealthy attack against electricity market using independent component analysis," *IEEE Systems Journal*, pp. 1–11, 2017.

[23] X. Liu, Z. Li, X. Liu, and Z. Li, "Masking transmission line outages via false data injection attacks," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 1592–1602, 2016.

[24] J. Fang, C. Su, Z. Chen, h. sun, and P. Lund, "Power system structural vulnerability assessment based on an improved maximum flow approach," *IEEE Transactions on Smart Grid*, pp. 1–1, 2017.

[25] S. Mishra, X. Li, T. Pan, A. Kuhnle, M. T. Thai, and J. Seo, "Price modification attack and protection scheme in smart grid," *IEEE Transactions on Smart Grid*, pp. 1–12, 2017.

[26] D. V. Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. L. Kling, "Usefulness of dc power flow for active power flow analysis with flow controlling devices," in *The 8th IEE International Conference on AC and DC Power Transmission*, 2006.

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

[28] Y. Koç, M. Warnier, R. E. Kooij, and F. M. Brazier, "An entropy-based metric to quantify the robustness of power grids against cascading failures," *Safety science*, vol. 59, pp. 126–134, 2013.

[29] Z. Bao, Y. Cao, G. Wang, and L. Ding, "Analysis of cascading failure in electric grid based on power flow entropy," *Physics Letters A*, vol. 373, pp. 3032–3040, 2009.

[30] C. Coffrin, P. Van Hentenryck, and R. Bent, "Approximating line losses and apparent power in ac power flow linearizations," in *2012 IEEE Power and Energy Society General Meeting*, July 2012, pp. 1–8.

[31] N. Alon, B. Awerbuch, and Y. Azar, "The online set cover problem," in *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, ser. STOC '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 100–105. [Online]. Available: https://doi.org/10.1145/780542.780558

[32] F. Milano, "A python-based software tool for power system analysis," in *IEEE Power Energy Society General Meeting*, 2013, pp. 1–5.

[33] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh, "The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee," *IEEE Transactions on Power Systems*, vol. 14, pp. 1010–1020, 1999.

[34] I. Pena, C. Brancucci, and B. M. Hodge, "An extended ieee 118-bus test system with high renewable penetration," *IEEE Transactions on Power Systems*, pp. 1–1, 2017.