MaDIoT 2.0: Modern High-Wattage IoT Botnet Attacks and Defenses

Tohid Shekari, *Georgia Institute of Technology;* Alvaro A. Cardenas, *University of California, Santa Cruz;* Raheem Beyah, *Georgia Institute of Technology*

https://www.usenix.org/conference/usenixsecurity22/presentation/shekari

This paper is included in the Proceedings of the 31st USENIX Security Symposium.

August 10–12, 2022 • Boston, MA, USA

978-1-939133-31-1

Open access to the Proceedings of the 31st USENIX Security Symposium is sponsored by USENIX.

MaDIoT 2.0: Modern High-Wattage IoT Botnet Attacks and Defenses

Tohid Shekari
School of Electrical and Computer Engineering
Georgia Institute of Technology
t.shekari@gatech.edu

Alvaro A. Cardenas
Computer Science and Engineering
University of California, Santa Cruz
alacarde@ucsc.edu

Raheem Beyah
School of Electrical and Computer Engineering
Georgia Institute of Technology
rbeyah@coe.gatech.edu

Abstract

The widespread availability of vulnerable IoT devices has resulted in IoT botnets. A particularly concerning IoT botnet can be built around high-wattage IoT devices such as EV chargers because, in large numbers, they can abruptly change the electricity consumption in the power grid. These attacks are called Manipulation of Demand via IoT (MaDIoT) attacks. Previous research has shown that the existing power grid protection mechanisms prevent any large-scale negative consequences to the grid from MaDIoT attacks. In this paper, we analyze this assumption and show that an intelligent attacker with extra knowledge about the power grid and its state, can launch more sophisticated attacks. Rather than attacking all locations at random times, our adversary uses an instability metric that lets the attacker know the specific time and geographical location to activate the high-wattage bots. We call these new attacks MaDIoT 2.0.

1 Introduction

Most attacks on power systems (e.g., Ukrainian power grid blackout, Dragonfly 2.0, and Aurora) target the central control systems of the electrical grid [1–4]. While targeting control systems is a direct attack, adversaries can also attack power systems indirectly through the consumer side. Given the proliferation of IoT vulnerabilities and the growing availability of high-wattage devices with Internet connectivity, security researchers are starting to analyze Manipulation of Demand via IoT (MaDIoT) attacks [5–7].

In a MaDIoT attack, the adversary uses a botnet consisting of high-wattage IoT devices to change the power system's load abruptly; these attacks might cause frequency instabilities, line failures, and increased operating costs [5]. A followup-work by Huang et al. [6] argued that a missing piece in Soltan's analysis [5] was a model of the protection mechanisms already in place in the power grid to prevent problems caused by natural events (e.g., sudden changes in the generation/load imbalance). They then showed how these protections (e.g., under-frequency load shedding or the time delay before disconnecting an overloaded transmission line) would significantly reduce the impact of MaDIoT attacks. In particular, Huang et al. argue that the embedded protection systems in the power grid will prevent widespread blackouts from adversaries launching MaDIoT attacks. While the negative impact of MaDIoT attacks on operations might not be as severe as previously thought, a recent work by Shekari et al. showed that MaDIoT attacks can be used to affect the electricity market (fraudsters can predict the market and future changes in the load because of the botnet they control) [7].

In this paper, we revisit the problem of the impact of MaDIoT attacks on the security of the power grid. So far, the original argument is that these attacks are dangerous [5], and follow-up work showed that they were not as severe as initially thought [6]. A missing analysis in these previous works is that they considered MaDIoT attacks as an all-or-nothing effort (e.g., turning on all bots simultaneously or turning them all off) [5, 6]. This spreads the attack throughout the power system equally and randomly. In this paper, we show that sometimes "less" is better. In particular, we show that by carefully turning on devices in specific geographical locations, we can target the system more methodically. In particular, we propose MaDIoT 2.0; our new attack looks at voltage stability indices and then targets the geographical areas where the system is more vulnerable from the stability perspective. Our contributions include the following:

- We propose and analyze a new MaDIoT 2.0 attack.
- We show that MaDIoT 2.0 has a significantly better success rate compared to the previous attacks (i.e., [5, 6]) while requiring a fewer number of compromised IoT devices, which makes it more feasible in practical situations.
- We conduct numerical studies to investigate the effectiveness of MaDIoT 2.0 with real-world data obtained by crawling the websites of independent system operators (ISOs) and the Bloomberg Terminal.

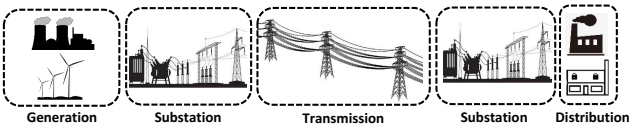

Figure 1: A typical power grid and its different sectors.

- All our models and simulations are released as open-source software so that other researchers can reproduce our results [8]. We are also sharing high-order models for standalone devices (e.g., generators, transmission lines, and loads), protective relays, and the controllers of standalone devices.

- We discuss short-term and long-term countermeasures to minimize the damaging consequences and severity of MaDIoT 2.0.

The rest of this paper is organized as follows. Section 2 explains the basic structure of the bulk power system and its primary control and protection schemes. The intuition behind MaDIoT 2.0 attacks and the threat model is discussed in Section 3. The detailed formulation and mechanism of the attack model are in Section 4. Section 5 evaluates the performance of the proposed approach with real-world case studies and shows its better performance over the previous works. Countermeasures are presented in Section 6. Related work is summarized in Section 7. We conclude the paper and discuss possible future work in Section 8. The feasibility of the attack and the data of the benchmark systems are in the appendix.

2 Background

A typical power grid is divided into different sectors illustrated in Figure 1 [9]. Generation, transmission, and distribution sectors are connected through substations [10]. Each substation includes high voltage equipment such as power transformers (to change the voltage level of the circuits) and circuit breakers (switches for connecting/disconnecting lines), and also control and protection devices such as protective relays (to detect faults), voltage and current measurement sensors, and remote terminal units (RTU) to communicate with the control center via the SCADA system [11].

Most of the energy is produced by power plants in the generation sector. The voltage at generators is originally medium voltage (e.g., 13.8 kV). This generated power is then stepped up to a higher voltage level (e.g., 500 kV) to be transmitted over long distances. This voltage level change is performed to reduce energy losses in transmission lines (higher voltage levels imply smaller currents, which lead to lower transmission losses). Eventually, the electricity is stepped back down to a medium voltage level by distribution substations near end users. The distribution sector generally feeds the consumers within a limited geographical area with medium voltage [9].

2.1 Control of Power Systems

The total demand for the power grid is continuously changing. To preserve system stability and avoid any large-scale blackouts, the output power of generators must match the demand in real-time [12]. These variations change the load of the transmission lines; some of them might even work while being overloaded, depending on the grid operating point [6]. Therefore, to relieve overloaded transmission lines, the configuration of the system is modified (by switching the transmission lines) so that the energy will be transmitted to the end-users via different routes. These strategies are a part of the power system control mechanism. Power system control is defined as the set of local and wide-area algorithms which help the system operator maintain grid stability. The main goal of operators is to ensure that the grid is continuously delivering electric energy to consumers at the nominal voltage and frequency with an acceptable error (i.e., 120 ± 20 V and 60 ± 0.5 Hz in the US) while keeping the generation-demand balance. The most important power system control schemes are: i) primary frequency control (governor), ii) automatic voltage control (AVR), and iii) automatic generation control (AGC). These controllers have either local or wide-area mechanisms.

Electric power has two main elements: i) active and ii) reactive power. Active power is the part that flows from the generator to the load, and reactive power is the part that is continuously flowing from source to load and is returning back to the source. The existence of reactive power makes the transfer of active power possible in electrical circuits [12].

The primary frequency controller is locally installed in each generator. It changes (increases/decreases) the active output power of the generator in response to any frequency change in the system, which is a sign of load-generation imbalance in the grid. As a rule of thumb, whenever the generation exceeds the demand in the grid, the system frequency becomes greater than the nominal value, and whenever the demand exceeds the generation, it drops below the nominal frequency [11].

The AVR is similarly installed in each generator with the goal of maintaining the voltage level of the generator within allowable ranges; it achieves this by changing the reactive power output of the generator [12].

Finally, AGC is a wide-area controller that changes the output power of the system generators to recover the frequency to its nominal value if the primary frequency controllers are not able to fully recover the system frequency change to the allowable range [12]. Wide-area controllers gather and analyze the data from the entire grid and make decisions and issue commands to multiple components throughout the system. These controllers use a private network for receiving data and sending commands; this network is called the SCADA system.

2.2 Protection of Power Systems

When a severe fault or incident occurs in the system (e.g., short circuit fault in a transmission line, sudden outage of a big power plant, etc.) and the physical damage to the grid components or a widespread blackout is inevitable, power system protection schemes will intervene to isolate the faulty area while keeping as much of the transmission network still in operation [11]. This means that there may be localized outages that can be easily repaired, and these are caused to prevent the interconnected bulk system from going down.

These protection schemes can be categorized into local and wide-area methods. The local protection schemes usually detect and isolate the faulty component in the system to prevent damage to the equipment and prevent the fault from spreading to the entire grid. On the other hand, wide-area protection schemes gather and process data from different parts of the grid through the SCADA system to detect any faults and react to them accordingly. From the technical perspective, wide-area protection schemes employ more sophisticated data analysis methods and are able to detect and resolve more complex faults in the system [11]. Table 1 lists the most common protection schemes used in the bulk power system.

These protection schemes play an important role in keeping the grid stable following severe natural events or accidents, and as shown by Huang et al. [6], they can even protect the grid from basic MaDIoT attacks. However, basic MaDIoT attacks and natural events have not taken an adversarial look at protection mechanisms. In the next Section, we discuss what types of attacks can bypass and even use the existing protection mechanisms against the grid.

3 Threat Model

3.1 Intuition Behind MaDIoT 2.0

The main objective of the MaDIoT attack is to manipulate the consuming power of compromised high-wattage IoT devices (turning them on/off at the same time) to cause a blackout in the target power grid. To understand the limitations of a MaDIoT attack, we first need to understand the intuition behind it. We simulate the classical benchmark IEEE 39-node (also known as the model of the New England power grid) test system to test different attack scenarios and determine which attack scenarios are more practical to implement to cause a blackout in the grid. We considered the load profile of one week per season (the modeling details are in Section 5 the appendix and in our public repository [8]. The central assumption of a MaDIoT attack is that the attacker has access to a high-wattage IoT botnet whose bots are evenly distributed in all of the system nodes. This means that adversaries can alter the system demand at different nodes at their will by turning on the bots simultaneously to take down the entire power grid.

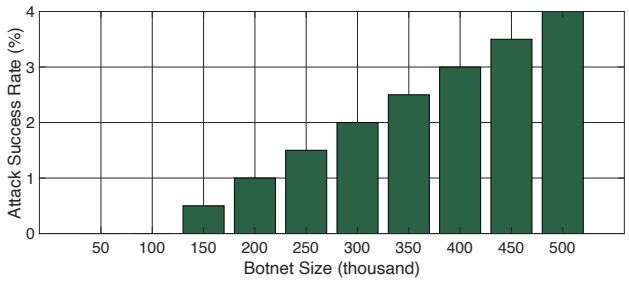

Figure 2: Success rates of MaDIoT 1.0 attacks vs. size of the high-wattage IoT botnet.

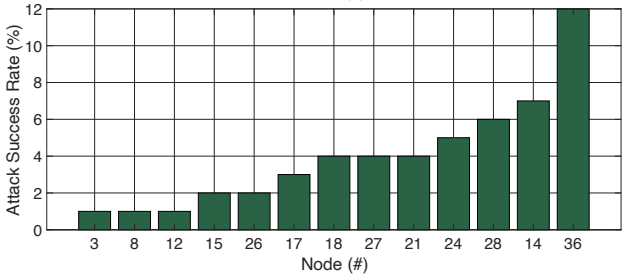

Figure 3: Success rates when attacking a single node (IEEE 39 node benchmark). Only the nodes with the highest success rates are in the figure.

In MaDIoT, the adversary turns on all of the compromised devices evenly distributed in the system nodes to give a shock to the power grid and cause a widespread blackout [5, 6]. Figure 2 shows the success rate of MaDIoT 1.0 for different botnet sizes. As we can see, the results are consistent with [6], and the system control and protection schemes are able to recover the stability of the power grid following the attacks in most of the simulated cases, and hence, the attack's success rate is very low. The low success rate can even lead to the early detection of the attack due to so many unsuccessful tries (consumers might report malfunctioning devices, or the electric utility can spot these anomalies).

Given that in historical power grid blackouts, the system collapse occurs due to a severe supply-demand imbalance in limited geographical areas of the grid, implementing the MaDIoT attack from a single node seems more reasonable than the previous attack scenario. Based on this, we simulated MaDIoT attacks by turning on the compromised IoT devices located in a single node of the power grid. Figure 3 illustrates the results of the new attack scenario. Note that only the nodes with the highest success rates are shown in the figure. We can see in Figure 2 that the success rate of the attacks in some nodes is larger than that of the previous scenario (although most nodes have a lower success rate). Our main question is whether we can further increase the attack's success rate or not.

The MaDIoT attack seems to have a significantly higher success rate if we launch it from several nodes that are close together (neighboring nodes). By performing the brute force simulations (simulating all the possible attacks from three nodes, n(3, 39) = 9,139), we noticed that attacking the grid

Table 1: The Most Common Protection Schemes Used in Power Grids

Name of the Protection Scheme	Local or Wide-Area	Aim
Distance [13]	Local	Short circuit detection in transmission lines
Overcurrent [14]	Local	Overload detection in transmission lines
Overvoltage Load Shedding (OVLS) [11]	Local/Wide-Area	Overvoltage detection in grid nodes
Undervoltage Load Shedding (UVLS) [15]	Local/Wide-Area	Undervoltage detection in grid node
Under-Frequency Load Shedding (UFLS) [16, 17]	Local/Wide-Area	Underfrequency detection in grid nodes
Over-Frequency Generation Rejection (OFGR) [18]	Local/Wide-Area	Overfrequency detection in generation nodes
Differential [19, 20]	Local	Fault detection in power transformers and transmission lines
Out-of-Step [21]	Local	Out of synchronous detection in power generators
Loss-of-Excitation [22]	Local	Excitation system fault detection in power generators

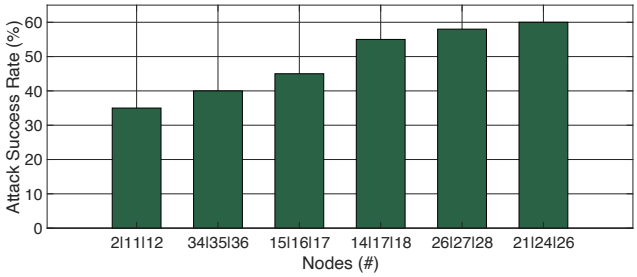

Figure 4: Success rate of an IoT botnet attack (IEEE 39 node benchmark) when launched from three nodes. Only the 3-nodes with the highest success rates are in the figure.

from three nodes reveals the best results. Figure 4 illustrates the performance of the attack when launched from three different locations in the power grid. The figure lists the highest success rates from all the possible simulated cases. We can see that the success rate of the new attack scenario is considerably higher than that of the previous ones. There are several technical reasons why the last strategy causes the most effective attacks. As the first reason, when a severe supply-demand imbalance occurs in a limited geographical area of the grid, the generators detect this energy shortage and try to increase their generated power to compensate for the deficit. However, the distant generators cannot deliver their share to the faulty area as long transmission lines have considerable power losses. Accordingly, this persistent power shortage remains in the faulty area and causes a cascading outage throughout the grid. In such situations, the central system protection scheme is under-frequency load shedding (UFLS). UFLS tends to evenly drop a portion of the grid's load to bring back the supply-demand balance. However, because it decreases the system loads evenly throughout the system, it cannot solve the problem, and a widespread blackout inevitably happens. We will show how a modification to the existing UFLS scheme can reduce the attack's success rate in Section 6.

Since the power grid demand and its operating status change over time, the best nodes to launch the MaDIoT attack change accordingly. Therefore, it is important to periodically (e.g., every 15 minutes) find the best nodes for the MaDIoT attack to increase the attack's success rate. In addition, given a power grid structure and its operating condition, there is a question on how an adversary calculates the best nodes of the system to attack in its current state. Our detailed analysis shows a strong correlation between the best nodes to attack and their voltage stability margins. The adversary can compute the voltage stability margins of different nodes in the system and choose the ones close to instability to launch the attack from those nodes. We explain the detailed hierarchical approach of our attack in the following Section.

One might ask here if we can choose these three nodes randomly to cause a blackout in the targeted grid. Technically speaking, the attack works well if we choose the top nodes (from the voltage stability perspective) that are close to the instability. For example, in the New England test system, there are n(3, 39) = 9,139 possible attack combinations, and only the top 5-6 choices return reasonable success rates after the implementation of the attack. The probability of choosing these nodes randomly is only 0.055%, which is very unlikely. In addition, due to a large number of unsuccessful attacks, the chance of consumers or operators detecting the high-wattage IoT botnet attack becomes high. The probability of choosing three good noes to attack gets smaller with larger power grids. For example, in a system with 118 nodes, this probability is 0.0019%.

3.2 Overview of MaDIoT 2.0

As in [5–7], we assume that vulnerable high-wattage IoT devices have been already compromised and are part of a botnet that can be directly controlled by the attacker–we discuss this assumption in Section 9. To launch more sophisticated attacks, the adversary needs to obtain the graph of the targeted grid through reconnaissance, phishing, or available automatic tools. This is a one-time analysis for each power grid and can be done offline. A detailed explanation of how an attacker can obtain the grid graph is in Section 4.1.

With this information, a MaDIoT 2.0 attack has two main online stages: in the first stage, the attacker obtains basic information about the targeted power grid, and in the second phase, the attacker analyzes the data to find the right time and

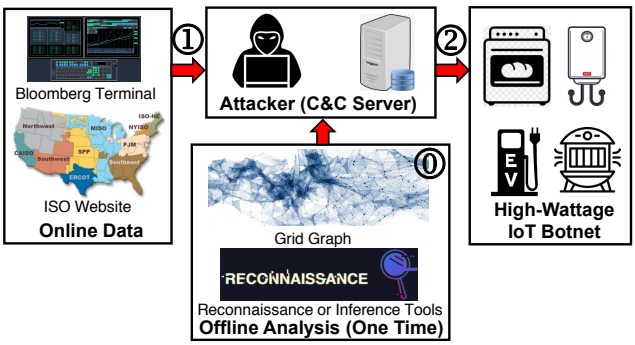

Figure 5: Overview of MaDIoT 2.0 attacks. ⓪ Preliminary offline analysis: the attacker obtains the grid graph through reconnaissance and inference tools, ① Data acquisition stage: the attacker collects the real-time operation information about the current operating point of the grid by crawling the online websites of ISOs and Bloomberg Terminal, ② System analysis stage: the attacker analyzes the raw data and determines the weakest points of the grid from the stability perspective. Finally, the attacker assesses the feasibility of the attack with the available high-wattage IoT botnet and implements it if it is feasible.

place to launch the attack. Figure 5 shows the threat model and overview of the new MaDIoT 2.0 attack. We now describe these steps in more detail.

3.2.1 Data Acquisition Stage

In this stage, the attacker's goal is to obtain real-time information on the state of the power grid. With the emergence of deregulated electricity markets in recent years, real-time data is publicly available through the ISO websites or stock trading tools like Bloomberg Terminal [23]. Three examples of online data sharing in ISOs websites are given in [24–26]. This data is openly available so traders and market players can monitor the changes in the underlying system and quickly adapt their bids in the market for maximum profit.

By crawling ISO websites and the Bloomberg Terminal, an attacker can learn the power production and consumption of the system in different nodes[1] and estimates the stability margin of the grid in various geographical regions. In particular, the attacker can obtain each node's integrated power generation/consumption, which is explicitly shared via the methods outlined above.

3.2.2 System Analysis Stage

After learning the state of each node, the next step for the attacker is to identify the most vulnerable nodes based on voltage stability analysis. The motivation to conduct this study is that one of the most common root causes of big blackouts is the voltage instability following severe technical incidents such as big imbalances between energy generation and consumption (e.g., caused by the outage of critical power plants due to technical faults) [12, 16, 17].

A classical way to identify vulnerable nodes is through voltage stability indices. There are multiple voltage stability indices in the literature, and they can help the attacker rank the nodes and determine the most vulnerable ones in real-time. We will give a detailed explanation of these indices along with their performance in Sections 4 and 5.

Once the attacker ranks nodes according to their stability margin, he needs to evaluate the likelihood that an attack on the top nodes (e.g., the top 3 most vulnerable nodes) will bring the power system down. The adversary sends the attack command to the relevant bots located in critical nodes if the attack is feasible. According to the numerical results presented in Section 5, the attack's success[2] rate can be as high as 91%. If the attack is not feasible, the attacker will wait until the beginning of the following scanning cycle (e.g., five minutes mainly depending on the refresh rate of the public information) and start again from the first stage. As we will discuss in Section 5, MaDIoT 2.0 attacks have two advantages: i) the attack is executed only when there is a good chance for success (so the existence of the botnet is not discovered because of failed attacks), and ii) a successful attack to bring down the power grid requires less IoT bots than in previous work because it only targets the weakest nodes of the system.

4 MaDIoT 2.0

4.1 Preliminary Offline Analysis

Before we describe the online tasks of MaDIoT 2.0, we discuss a preliminary offline step. To launch successful attacks, the attacker needs to gather some basic information about the architecture of the power grid. This architecture does not change over long periods of time (years), and we only need to obtain them once for every target grid. Although the power companies and ISOs do not explicitly share this type of information, the attackers can acquire them with indirect methods such as phishing or social engineering [27]. In addition to this, researchers have shown that there is enough openly available information to infer in great detail the topology and configuration of power grids [28].

Because all of the transmission lines and substations in the bulk power grid are outdoors, they can be identified with online mapping services such as Google Maps [29]. The attacker can easily follow transmission lines from the power plants to the distribution substations and obtain the graph of the entire system. The size and shape of the tower reveal the voltage level of the respective transmission line and the attacker can estimate the technical parameters of the line by

[1] The terms nodes and buses are interchangeably used in power systems. Each node represents a relatively wide geographical area (e.g., a metropolitan city such as Atlanta or a big power plant).

[2] Success is defined as a complete blackout in the entire grid, and failure is defined as the recovery of the grid from the attack

multiplying the length of the line to the per-unit values of the relevant tower. Although this process might take some time if the attacker does it manually, the processing time here is not important because it is a one-time analysis for every system. In addition, the attacker can develop computer vision algorithms to automatically (or semi-automatically) generate the graph of a given system using Google Maps satellite pictures [30–32]. Reference [31] illustrates how the graph of the Chilean power grid can be inferred by an open-source tool.

4.2 Data Acquisition Stage

As we briefly explained in Section 3.2, once the attacker has some basic information of the topology of the grid, the next step is to start monitoring the state of the system and find vulnerable nodes in real-time. As explained before, this data (power generation and consumption in each node) of the power grid is released on the website of ISOs (e.g., NY-ISO and CAISO) and is updated every 5 minutes [24, 25]. In addition to this, this data can also be accessed through an advanced stock trading tool called Bloomberg Terminal [23]. To extract the operation data, the attacker uses a crawler. To verify this, we collected all the system operation data of the California and New York power grids from January 2020 to January 2021.

4.3 System Analysis Stage

Once the attacker has information about the state of the system, the next step is to determine the weakest nodes of the grid from the voltage stability perspective. Note that since the power generation and consumption in different nodes of the system change constantly, the weakest points of the target grid will change accordingly. In particular, we want to exploit voltage instability as changing voltages in multiple nodes is easier than attempting to change the frequency of the grid. Frequency is a global variable in power grids while voltage is a local variable in each of the system nodes. Changing a node's voltage can be done by slightly changing the demand at that node while changing the system frequency requires the demand change in the entire system.

We can create voltage instabilities when the load increases in nodes where the *voltage stability margin* is at its critical point [15]. Therefore to find vulnerable nodes, we need to compute the voltage stability margin. While finding the exact stability margin is computationally expensive and cannot be solved in real-time, the power grid community has developed a set of voltage stability indices that approximately rank the system nodes based on their voltage stability margins. We now introduce two candidate options for estimating this quantity.

4.3.1 Voltage Magnitude of Nodes (Index 1)

During the normal operation of the power grid, the operators want to keep the voltage magnitude of system nodes constant, and only allow deviations between 0.95 and 1.05 per unit (p.u.). Undervoltage protection schemes use the voltage magnitude of the system nodes as an indication of their stability margin [15]. In these schemes, a lower voltage magnitude implies a lower stability margin in the node, and hence, during emergency conditions when the voltage magnitude is too low, the protection scheme drops a portion of the loads that are fed through the critical nodes to keep the system stability and recover the voltage value to its nominal range.

This index is very easy to compute (we only need the value of the voltage at each node), and therefore our first candidate to identify the most vulnerable nodes is the voltage magnitude of nodes. We will show later (Figure 10.c2) that if the attacker increases the system demand in several critical nodes where the voltage magnitude is in the minimum range, the resulting effect then further drops voltage magnitudes below the normal range and causes cascading outages in the power grid.

4.3.2 Modal Analysis (Index 2)

One of the most efficient methods to calculate the voltage stability margin of the system nodes is the modal analysis based on the Jacobian matrix [33–35]. To calculate this index for different nodes, the relationship between the system states and the active and reactive power of system nodes can be written as:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} J_{P\theta} & J_{PV} \\ J_{Q\theta} & J_{QV} \end{bmatrix} \begin{bmatrix} \Delta \theta \\ \Delta V \end{bmatrix}, \tag{1}$$

where ΔP, ΔQ, $\Delta\theta$, and ΔV are the vectors representing active and reactive power changes and voltage magnitude and angle change in the system nodes. The elements of the Jacobian matrix (i.e., $J_{P\theta}$, J_{PV}, $J_{Q\theta}$, and J_{QV}) are calculated based on the results of the load flow analysis performed in the previous stage (the data acquisition stage helps us perform this analysis easily). Since the reactive power and voltage changes have strong relationships with each other, it is reasonable to assume $\Delta P = 0$ [33]. Therefore, we can write:

$$\Delta Q = [J_{QV} - J_{Q\theta} J_{P\theta}^{-1} J_{PV}] \Delta V = J_R \Delta V. \tag{2}$$

Reversing the matrix above equation, one can rewrite it as:

$$\Delta V = J_R^{-1} \Delta Q. \tag{3}$$

In order to simplify the calculations, a decomposition can be used to calculate J_R^{-1} as:

$$J_R^{-1} = E_R \xi^{-1} E_L. \tag{4}$$

Hence, (3) can be transformed into:

$$\Delta V = E_R \xi^{-1} E_L \Delta Q = \sum_l \frac{E_{R,l} E_{L,l}}{\lambda_l} \Delta Q, \quad (5)$$

where E_R and E_L are the right and left eigenvector matrices of J_R and ξ is the diagonal eigenvalue matrix of J_R. Also, $E_{R,l}$ and $E_{L,l}$ denote the lth column and lth row of E_R and E_L and λ_l stands for lth eigenvalue of J_R. All things considered, the V-Q sensitivity of node k can can be computed through:

$$VQS_k = \frac{\partial V_k}{\partial Q_k} = \sum_l \frac{\mu_{kl} \eta_{lk}}{\lambda_l}, \quad (6)$$

in which μ_{kl} and η_{lk} are the kth element of $E_{R,l}$ and $E_{L,l}$. The negative value of VQS_k implies the voltage instability in the node k in the grid. During the normal operation of the power system, VQS_k will be positive for all of the system nodes which means that the system is stable from the voltage stability perspective; however, we can use this index to rank the system nodes and see which one of them are more prone to the instability point. The lower value for the VQS_k index implies that the kth node of the system is closer to the voltage instability and a small load shock can make it unstable.

Index 2 is more computationally intensive (i.e., $O(n^5 \times (\log n)^2)$ time complexity) for larger systems because a large matrix (the dimensions are increasing linearly with the total number of the grid nodes, i.e., $n \times n$) is getting inverted for its calculation. Index 1 on the other hand, is only looking at ΔV in equation (1) (which is a subset of Index 2) and is very easy to compute (i.e., $O(1)$ time complexity) regardless of the grid size. So, in short, index 2 performs better because it is a more accurate model of the system, yet it requires more effort by the attacker. For very large grids where the calculation of index 2 is not practical, index 1 can be used which has some errors but reveals a satisfactory success rate for the attack. In smaller grids, we can use index 2 because of its higher attack success rates.

4.3.3 Checking the Feasibility of an Attack

After the attackers find the most vulnerable nodes in the grid, they need to evaluate the feasibility of the attack before its implementation. To do so, the adversary can use modal analysis to determine whether the grid will be unstable following the implementation of load alteration attack in the weakest nodes of the system. In the first step, the attackers update the power consumption in the weakest nodes of the power grid based on the available potential of the high-wattage botnet. Then, they compute AC load flow analysis and recalculate $J_{P\theta}$, J_{PV}, $J_{Q\theta}$, and J_{QV} based on the new updates. Finally, the voltage stability index of different nodes can be obtained through (6). If following the attack implementation, the index VQS of weak nodes becomes negative, this implies that the grid will be unstable; otherwise, the available botnet is not strong enough to take down the power grid and the attack scenario is not feasible in the current operating point. If the attack does not succeed with the current state, the attacker needs to wait for at least five minutes so that the operating point of the grid and power consumption in different nodes change. Then, the adversaries will repeat the previous analysis until they find an attack with a high success likelihood. The accuracy of this evaluation is numerically evaluated in Section 5. We will show that by using a suitable voltage stability index, the success rate of the attack can be as high as 90%.

5 Experiments and Discussion

Due to the irreparable economic and social damages caused by the real-world implementation of MaDIoT 2.0, we use simulation results to show the performance of the proposed attack (as was done in previous work [5, 6]). In this line, instead of adopting simplified models and simulations, we leveraged an advanced, commercial power system simulation software called DIgSILENT PowerFactory [36].

5.1 Test Case and Component Modeling

To evaluate the performance of the proposed attack, we use a standard test power grid, called the New England power system—this is also known as the IEEE 39-bus test system. The New England power system includes 39 buses (nodes), 32 transmission lines, 24 power transformers, and 10 generators. The total base load of the system is 6097.1 MW (active load) and 1408.9 MVAr (reactive load) [37]. This test case is greater than the New York and California ISOs together, and hence, is a reasonable representation of power grids in various countries. We will also use a smaller test case (IEEE 9-bus test system) which has been used in the previous relevant works [5, 6]. The IEEE 9-bus test system has 6 transmission lines, 2 power transformers, 3 generators with a total active and reactive power generation of 350 MW and 244 MVAr. Further explanation about the used test cases can be found in Section 10 (Appendix).

To capture the system dynamics and minimize the simulation errors, we used the best recommendation from power system standards [12]. An eighth-order model is used for representing the dynamics of generators. In this model, the mechanical part of the generator is formulated by a second-order state-space equation and the electrical part by a sixth-order system [12]. Furthermore, the IEEE-type DC1A excitation system for modeling the AVRs of generators and an appropriate governor model are employed in our simulations [12].

We also adopt a combinational load model in our simulations, where the static and dynamic parts of the composite model are represented by a polynomial model (i.e., a mixture of power constant, current constant, and impedance constant loads) and a third-order induction motor, respectively [16, 17]. This model of the system allows us to study the dynamic

behavior of the power grid in response to MaDIoT 2.0 attacks with minimal errors.

We model the system for 24 hours in each season. We chose 24 hours because the daily load curve of each power grid is similar and we can reasonably extrapolate the analyzed results to longer periods of time such as one year [24, 25]. Previous works have studied the power grid behavior during the non-strategic MaDIoT botnet attacks only in one snapshot (e.g., 5 minutes) [5, 6]. Therefore, previous results cannot be reliably extrapolated to longer time periods.

Table 2 summarizes the list of new models and simulation contributions of the current work over the recent related works. Our work presents the most comprehensive and high-fidelity modeling of real-world power systems when compared to previous works. Our improved models are highly important because of three main reasons:

1. Ignoring the detailed modeling of the system controllers (i.e., governor, AVR, AGC) in the time domain simulations can lead to considerably erroneous results compared with practical situations because these controllers contribute to the system recovery when a severe incident occurs in the grid [12].

2. When a big disturbance happens in the grid, system protection schemes (i.e., distance, overcurrent, OVLS, UVLS, UFLS, OFGR, differential, out-of-step, and loss-of-excitation) seek to locate and isolate the faulty area to limit the damaging consequences of the widespread outages and area of the blackout. Hence, overlooking these schemes in the simulations will lead to erroneous results [38].

3. The detailed modeling of the system components (i.e., generators, static, and dynamic loads) play an important role in the power system dynamic studies. During unstable conditions, the dynamic behavior of the loads and system generators makes the situation worse and pushes the grid towards a more unstable point where recovery is hard [16].

This comprehensive system modeling approach can be used in future studies as the benchmark.

Table 2: Modeling and Simulation Contributions of the Current Work Over the Recent Related Works

Element	Our Work	Soltan et. al. [5]	Huang et. al. [6]
Governor	✓	✗	✓
AVR	✓	✗	✗
AGC	✓	✗	✗
Distance	✓	✗	✗
Overcurrent	✓	✓	✓
OVLS	✓	✗	✗
UVLS	✓	✗	✓
UFLS	✓	✓	✓
OFGR	✓	✗	✗
Differential	✓	✗	✗
Out-of-Step	✓	✗	✗
Loss-of-Excitation	✓	✗	✗
Static Load	✓	✓	✓
Dynamic Load	✓	✗	✗
Generator Model	✓	✗	✗

5.2 Evaluation of MaDIoT 2.0

Grid Simulation and Botnet Features We first consider the New England power grid for 24 hours. We assume that the attacker obtains the power consumption of the simulated power grid every 5 minutes. Therefore, there are 288-time intervals in which the available high-wattage IoT botnet can be used to take down the power grid. We chose a 24-hour grid simulation because of the daily load patterns of the grid [39]. For our initial evaluation, we assume that the attacker has access to a botnet with 150,000 bots, each of which can consume 3 kW of electrical power ($\sim$ 3,850 bots per node since the New England grid model has 39 nodes). In each node, the attacker needs to control $\sim$0.2% of the total real power of the grid to have a consistent attack with high success rate. According to our calculations, there are roughly 52,000 residential high-wattage IoT devices in each of the grid nodes in the New England power grid (more than 2 million devices in the entire grid). Assuming that the attacker cannot compromise more than one bot in each residential home, he would need to compromise 3,850 homes (from 52,000 total homes within each node) in each of the grid nodes to have a very high success rate in his attack.

We use each of the voltage stability indices as a binary predictor (if the algorithm gives a positive, it also gives us the suitable nodes to launch the attack) at each simulation time period. For the current time, the predictor can determine whether to launch the attack or not. With the high accurate simulation of the power grid, we evaluate the performance of predictions. We have the following basic definitions used in calculating the evaluation metrics.

- **True Positive:** When the predictor gives a positive (i.e., it recommends an attack) and the attack causes a blackout.

- **True Negative:** When the predictor gives a negative (i.e., it recommends no attack) and a blackout in the system is not possible even with a brute-force attack.

- **False Positive:** When the predictor gives a positive (i.e., it recommends an attack) but the attack's implementation does not cause a blackout.

- **False Negative:** When the predictor gives a negative (i.e., it recommends no attack) but we find a "blackout"

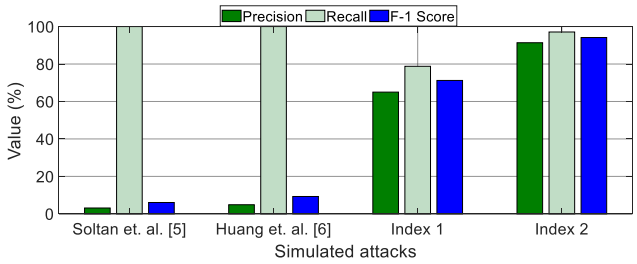

Figure 6: The performance of the proposed attack methods along with the approaches developed in the most recent works obtained from the New England power grid [5, 6].

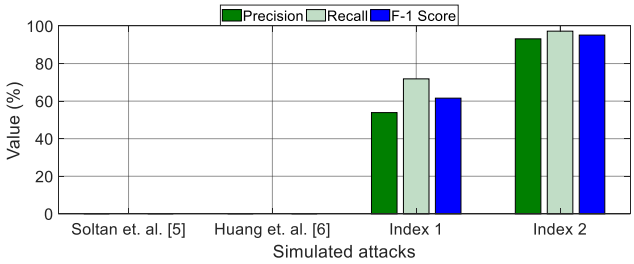

Figure 7: The performance of the proposed attack methods along with the approaches developed in the most recent works obtained from the IEEE 9-bus test system [5, 6].

in the system by using a brute-force search of all possible attacks (i.e., there is a successful MaDIoT attack, but the predictor fails to see it).

Attack Performance in Test Grids Figure 6 illustrates the precision, recall, and the F-1 score of the proposed attack methods along with the approaches developed in the most recent works obtained from the New England power grid [5, 6]. Here, the precision or recall alone are not suitable to judge the performance of the studied attack methods. Since the methods proposed in [5, 6] have zero false negatives, their recall is 100%. However, they have a high false-positive rate, and hence, low performance. Accordingly, we use the F-1 score to judge the overall performance of the simulated attack scenarios because it represents a combination of both precision and recall. As can be seen in Figure 6, the proposed MaDIoT 2.0 attacks outperform the previous methods with a relatively large margin. As previously reported by Huang et al. [6], we confirm that in most of the cases the conventional protection schemes are able to control the disturbance caused by ***random attacks*** and keep the system stability following the IoT botnet attack. However, MaDIoT 2.0 attacks can bypass these protections and create a cascading failure of the bulk power system. The reason behind this observation is that we target the weakest nodes of the grid while the proposed attack scenarios in [5, 6] are uniformly spread over the grid. Simulating the IEEE 9-bus test system reveals similar results with 0%, 0%, 61%, and 93% F-1 scores that are associated with Soltan et. al. [5], Huang et. al. [6], Index 1 and Index 2, respectively (see Figure 7).

Performance of Indices 1&2 The other interesting observation is that Index 2 for the voltage stability margin indicator performs much better than Index 1 for launching MaDIoT 2.0 attacks. Based on this observation, it is apparent that Index 2 is a better indicator for determining the voltage stability margin in system nodes. To validate this claim, we solved the exact model of the system equations in all of the 288-time intervals and obtained the error-free ranking of the system nodes in terms of their voltage stability margins. Figure 8 depicts the normalized accuracy of Index 1 and Index 2 compared with the ground truth which was acquired through performing the computationally intensive calculations (that cannot be performed in real-time) on the New England power grid. According to this figure, Index 2 represents the ranking of the grid nodes based on their voltage stability margins with lower errors and this justifies the results we observed in Figure 6. It should be noted that the exact equations of the entire grid took longer than 24 hours to calculate (for all 288 time intervals) and this is why we cannot use the exact model in the real-time attack mechanism. On the other hand, the calculation of the proposed indices in each of the time intervals took less than 5 sec (24 minutes for all 288-time intervals) and this makes the MaDIoT 2.0 attack feasible in practice. Again, simulation of the IEEE 9-bus system returns the same similar pattern where the normalized accuracy of Index 2 is much higher than that of Index 1. This is aligned with the aforementioned performance metrics discussed in the previous paragraph.

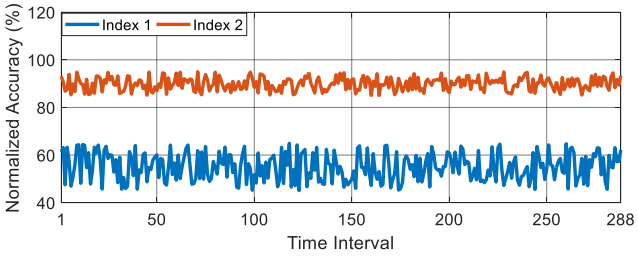

Figure 8: The accuracy of the proposed Index 1 and Index 2 compared with the exact calculations (ground truth).

Effect of Botnet Size on Success Rates We also considered how the size of the botnet changed the effectiveness of all attacks. Figure 9 shows the F-1 score of different attack methods versus the number of compromised high-wattage IoT devices that the attacker can control in the new England power grid. As can be seen, not only the proposed attack methods have a higher F-1 score compared to the previous attack mechanisms in the literature, but also they require smaller botnets to cause a large-scale blackout of the bulk power system. In addition to this, as the size of the available botnet increases, the success rate of the proposed attacks increases as well. However, after a certain point (150,000 bots) the increase in the attack success rate saturates and it does not respond to an increase in the size of the available botnet. In the IEEE 9-bus test system, the optimal success rate for the MaDIoT 2.0 attack is achieved when the size of the botnet is 5,500 bots. The reason for this observation and its difference with the

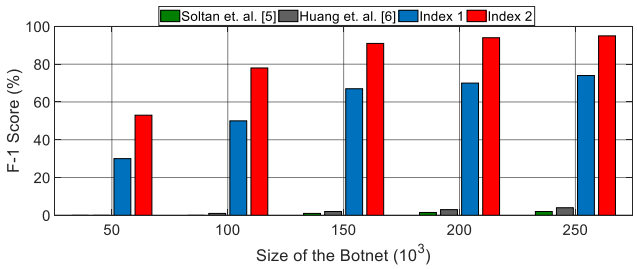

Figure 9: The success rates of different IoT botnet attacks versus the size of the available high-wattage IoT botnet in the New England power grid.

New England power grid is that the IEEE 9-bus test system is much smaller than the New England power grid. Based on this, one can reasonably infer that for smaller power grids, we will need smaller botnets to achieve the highest attack success rate.

Effect of the Attack on Node Voltages and System Frequency It is interesting to see how the proposed attack causes a system collapse. To do so, we illustrate the voltage profile of the system nodes and the grid frequency during different attacks in the New England power grid. Figure 10 shows the system frequency and the voltage profile of the New England power grid during one of the simulated time intervals. In this Figure, ⓪, ①, ②, and ③ denote the attack time, protection and control involvement time, system recovery to the normal condition, and wide-area blackout (voltage collapse), respectively. In this figure, (a1), (b1), and (c1) illustrate the system frequency over time (the frequency is unique and equal in every power grid). Also, (a2), (b2), and (c2) depict the system voltage magnitude in different nodes over time (each node has its voltage magnitude, and they are not equal in the entire grid). As we can see, the system controllers and the protection schemes are able to recover the grid to the normal operating state following the attack scenarios proposed in [5, 6]. However, since our attack targeted the weak nodes of the system, power grid controllers and protection schemes are unable to handle the emergency conditions following the attack, and a system-wide voltage collapse becomes inevitable. More specifically, the voltage magnitude of the system nodes collapses after a few seconds of starting the MaDIoT 2.0 attacks. The blackout is confirmed almost ten seconds after the attack.

Effect of Botnet Time Delay on the Attack's Performance An important factor that has not been studied in the previous works is the network delay of the bots' activation during the manipulation of demand attacks. Practically speaking, the attacker cannot simultaneously activate/deactivate the bots in the botnet because of the latency and randomness in the underlying communication network. To model the network latency associated with different bots, we considered a normal probability distribution function (pdf) for the network latency. Considering the standard deviation of 100 msec, Figure 11 depicts the successful rate of different attack methods versus various values for the mean of the normal pdf in the New England power grid. As can be seen, the latency of the communication network does not have significant effects on the proposed and previous attack methods. The MaDIoT 2.0 attacks still have a very high chance to cause a system-wide blackout in the studied power system. Simulating this delay in the IEEE 9-bus test system resulted in a similar pattern where no notable change was observed during the increase of the botnet delay.

Effect of Available Bots in System Nodes Although MaDIoT 2.0 attacks have excellent performance in causing system-wide blackouts in power grids, they have their own limitations. While MaDIoT 2.0 requires fewer bots, it also requires the attacker to have a presence in all nodes. This is a direct factor in the botnet operator's success, i.e., if the operator does not have enough bots in a location that is a "weak point", the adversary might not be able to launch a successful attack sometimes. It should be noted that the weak points of the power grid change as the loading of different nodes change around the clock. Accordingly, even if the adversary has compromised few bots in certain nodes, he still should be able to cause a blackout in the target grid at certain times. To verify these explanations, we did an experiment in the New England power grid. Figure 12 shows the performance of the MaDIoT 2.0 attacks with different coverage of nodes that consist of high-wattage bots. As expected, although the attack performance declines following the decrease in the number of nodes having high-wattage bots, the success of MaDIoT 2.0 attacks are still higher than MaDIoT 1.0. In practice, there are many high-wattage IoT devices that can be compromised in each grid node. Smart HVAC systems and EVs are only two classes of high-wattage devices that could be compromised and leveraged by the attacker as a part of his botnet. There are currently more than 30 million smart HVAC systems in the north America region [40]. Also, the number of EVs on U.S. roads is projected to reach 18.7 million in 2030, up from 1 million at the end of 2018. This is about 7% of the 259 million vehicles (cars and light trucks) expected to be on U.S. roads in 2030 [41].

Maximum Waiting Time for the Attacker The maximum amount of time the attacker will have to wait so that a feasible attack scenario occurs depends on the operating point of the grid and its general stability margin. Modern grids are often operated near to their stability limits to use the maximum capacity of the grid components and to postpone expensive grid expansion plannings. For this reason, a typical power grid such as the New England test system forces the attacker to wait roughly 3 hours. While impractical, (due to the high operational cost of non-optimal grid operation), operating a grid with a higher stability margin would increase the time the attacker would have to wait to launch a successful attack.

Effect of Line Parameter Errors The other important aspect of the MaDIoT 2.0 attacks is their performance sensitivity to estimation errors by the attacker. To analyze this in detail,

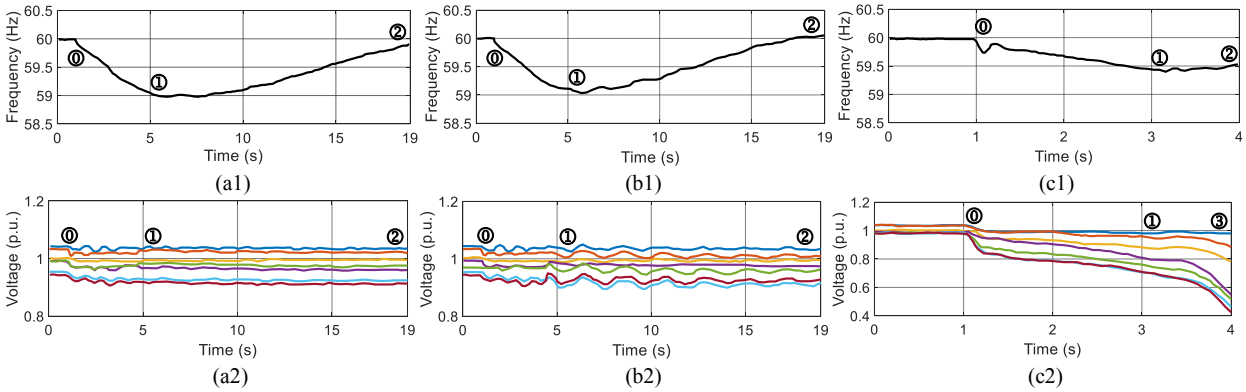

Figure 10: The system frequency and voltage profile of the New England power grid during one of the simulated time intervals following the implementation of different attacks: a) Soltan et. al. [5], b) Huang et. al. [6], and c) Proposed approach using Index 2. Note: ⓪, ①, ②, and ③ represent the attack time, protection and control involvement time, system recovery to the normal condition, and wide-area blackout (voltage collapse), respectively.

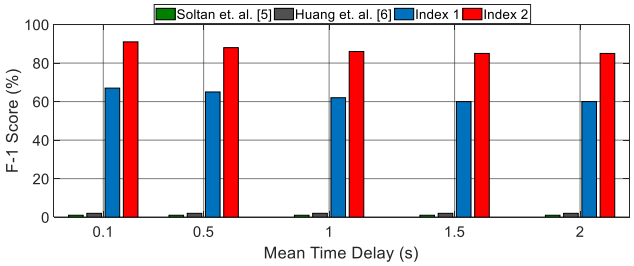

Figure 11: The performance of the different IoT botnet attacks in the New England power grid versus the mean value of the normal pdf used for modeling the network latency.

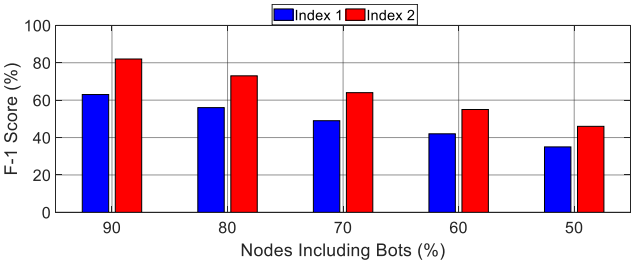

Figure 12: The performance of the MaDIoT 2.0 attacks versus the percent of the nodes including high-wattage IoT bots.

we performed two different experiments in the New England power grid. In the first experiment, we assumed that the line parameters of the grid transmission lines which are obtained through offline analysis are not 100% accurate. Figure 13 depicts the performance of the MaDIoT 2.0 attacks versus the mean error in the New England power grid line parameters. According to this figure, the attack performance decreases as the error becomes bigger. However, this performance reduction is not that significant and the MaDIoT 2.0 attacks are still relatively effective in the presence of reasonable estimation accuracy in the transmission line parameters. Since the attack impact is lower with imperfect information, there is value in adding deception. In fact, the defender can create

Figure 13: The performance of the MaDIoT 2.0 attacks versus the mean error in the New England power grid line parameters.

fake diagrams and substations in documents about the system to create fake nodes; however, they also need to be consistent with the physical-world information the attacker can get (e.g., build a phony substation that can be seen in Google Maps), which can be challenging in practice.

Effect of Erroneous Grid Operation Data In our final experiment, we considered the effectiveness of our data-driven countermeasure, so we assumed that the data associated with the power generation/consumption in different nodes is publicly released with some errors. Figure 14 shows the performance of the MaDIoT 2.0 attacks versus the mean error in the New England power grid nodes' power generation/consumption. As it was expected, the performance of the MaDIoT 2.0 attacks drastically declines with the increased error in the grid nodes' power generation/consumption. This performance reduction is more severe than that of the previous experiment shown in Figure 13. It is worth mentioning that Figure 14 implies that the first data-driven countermeasure (limiting the online data sharing) cannot be alone used for eliminating the risk of MaDIoT 2.0 attacks. However, it definitely reduces (∼26% performance reduction for index 2 in the presence of 5% error) the attacker chance of launching a successful attack to cause blackout in the entire grid.

As a final point, according to recent studies, the emerg-

Figure 14: The performance of the MaDIoT 2.0 attacks versus the mean error in the New England power grid nodes' power generation/consumption.

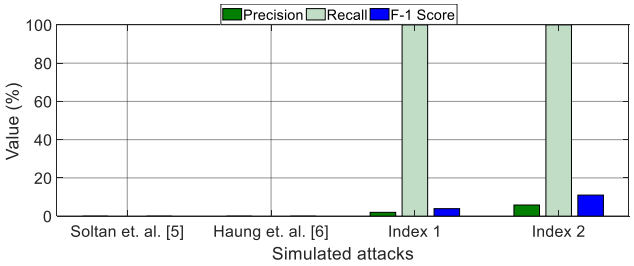

Figure 15: The performance of different manipulation of demand attacks after the modification of UFLS and UVLS schemes in the New England power grid.

ing trends such as the proliferation of new technologies in power grids such as rooftop solar panels and electric vehicles (EVs) reduce the stability margin of the power grids [12], and hence, increase the chance of successful MaDIoT 2.0 attacks in practical cases. Therefore, developing novel countermeasures against MaDIoT 2.0 is needed in future studies.

6 Countermeasures

The most effective way to prevent MaDIoT 2.0 attacks is to update the protection schemes of the power grid so that they can recover system stability following any unpredictable shocks caused by similar attacks. In our experiments, we saw that UFLS and UVLS are the two main protection schemes involved during MaDIoT 2.0 attacks.

Conventional UFLS and UVLS schemes drop a predetermined amount of the power grid loads when the system frequency and voltage drop severely following a technical event such as the outage of a power plant or a heavy transmission line. However, system operators do not consider the situations such as MaDIoT 2.0 attacks when they are configuring them. The current protection schemes drop the system loads that are evenly distributed in the entire grid. However, as shown in Section 5.2, this strategy is unable to protect the grid and recover it following the proposed IoT botnet attack. While one possible way to fix this issue is to drop the loads where the attack was launched, it is hard to detect and identify the location of the MaDIoT 2.0 attack in the grid. The reason is that it is hard to distinguish between a natural load change or technical event and a MaDIoT 2.0 attack from the system monitoring perspective.

One of the effective indicators which could be leveraged to detect the region of the manipulation of demand attacks is to use the voltage falling rate of grid nodes [42]. We observed that during the IoT botnet attacks, the voltage falling rate in the nodes that are close to the attacked nodes is much higher than that of the far nodes. Therefore, we revised the setting of the existing protection schemes so that they will first drop loads of the system in the nodes where the voltage falling rate is bigger than the other nodes. This adaptive protection scheme will shed some loads in the area of the attack and will help the system recover from the attack.

Figure 15 illustrates the performance of different manipulation of demand attacks after the modification of UFLS and UVLS schemes with the explained logic in the New England power grid. As we can see, the F-1 score of the studied attack mechanisms significantly drops with this modification. The high recall occurs because we have zero false negatives in the test cases. Figure 16 depicts the system frequency and voltage profile following the implementation of a MaDIoT 2.0 attack considering the modified adaptive protection scheme in the New England power grid. We can see that the system was going to become unstable following the attack; however, the modified adaptive protection scheme is able to identify the region of the attack and drop the loads accordingly. This eventually helps the grid to fully recover from the attack and prevent a system-wide blackout.

This is just an improvement to existing protection schemes against IoT botnet attacks. However, we also need to check that our modified protection scheme also works against natural incidents and compare it with the existing protection schemes. We simulated the New England power grid applying a set of natural events and observed the performance of the modified UFLS scheme. According to our observations, since the modified UFLS algorithm drops the similar amount of loads to the conventional one (but from different locations), it is able to recover the system stability following natural events. The reason is that for natural random events it does not matter where we shed loads, it only matters what the amount of the load shed is. However, further in-depth analysis is required to verify the effectiveness of the modified protection scheme and also to identify any possible weaknesses against natural events.

7 Related Work

Power system cybersecurity has been widely studied in the past few years [1, 2, 4, 5, 43–62]. Attacks on power systems can be classified into three main groups based on the ultimate goal of the attacker: i) attacks targeting the power grid communication infrastructure, ii) attacks targeting the power grid standalone components, and iii) attacks targeting the power

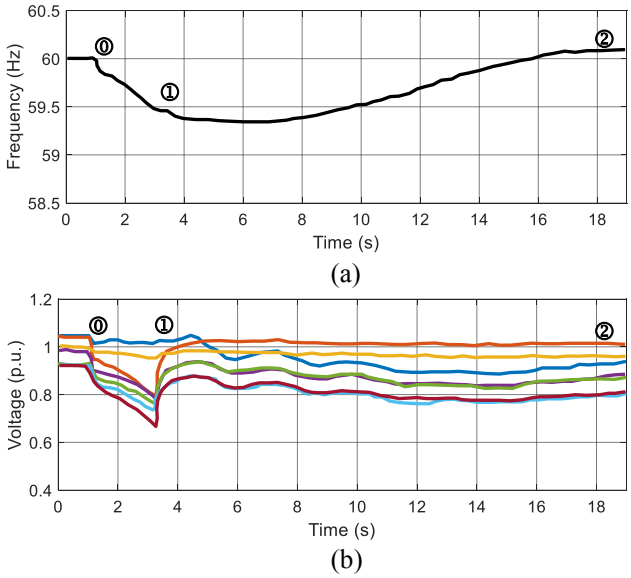

Figure 16: The system frequency and voltage profile of the New England power grid during one of the time intervals following the implementation of MaDIoT 2.0 attack considering the modified adaptive protection scheme. Note: ⓪, ①, and ② represent the attack time, protection and control involvement time, and system recovery to the normal condition, respectively.

grid through consumers (load altering attacks).

In the last class of attacks, the adversaries try to indirectly affect the normal operation of the system by attacking pricing signals or consumer devices [5, 58–63]. These indirect attacks were first introduced by Mohsenian-Rad and Leon-Garcia [58]. In this work, the attackers compromise the load control signals associated with big industrial loads and data centers. Big data centers can be one of the primary targets of load-altering attacks [59]. Barreto et al. [63] considered how attacks could manipulate the demand-response market to change abruptly the load of the grid with devices that respond automatically to energy prices.

The interest in load-altering Botnet attacks started in 2017 and 2018 [5, 61, 62]. Dvorkin et al. [61] proposed an optimization-based approach that requires a complete knowledge of the power grid (topology of the grid, detailed parameters of the transmission lines/generators, and real-time regional generation/demand). To overcome this challenge, Dabrowski et al. proposed a new method to increase the total system demand through remotely activating CPUs, GPUs, hard disks, screen brightness, and printers to cause frequency instability in the European power grid [62]. Although the new approach did not require as much detailed information about the power grid, the number of compromised devices needed for a successful attack is quite high because the devices do not consume a lot of power. Soltan et al. proposed the use of high-wattage IoT devices to launch various types of attacks (frequency instability, power line cascade tripping, and black start restoration interruption) on a power grid to cause blackouts in the entire system [5]. This novel attack, called Manipulation of demand via IoT (MadIoT), was further analyzed by Huang et. al. [6] and it was shown that the introduced attack is not as effective as it was illustrated in [5]. The new analysis in [6] revealed that while MaDIoT attacks could have negative effects on the power grid operation, it is hard to cause a widespread system collapse. According to these analysis, the existence of conventional protection schemes in the grid can effectively protect the system against random MadIoT attacks. However, these protection schemes were mainly designed to help the power grid withstand against credible contingencies and there was no consideration of the manipulation of demand attacks in their configurations [16, 38].

8 Conclusion and Possible Directions

In this paper, we introduced MaDIoT 2.0: a hierarchical two-stage attack mechanism that leverages the potential of high-wattage IoT botnets to attack the power grid and cause a widespread blackout in the entire system. The performance of the developed attack methods is evaluated using extensive simulations and the results showed the superiority of MaDIoT 2.0 over the previously studied attack mechanisms. More specifically, the success rates of the new IoT botnet attack were 91% and 67% for voltage stability Index 1 and Index 2, respectively. In addition, MaDIoT 2.0 requires a smaller number of bots involved in the attacks, since it targets the weakest nodes of the system in the current operating state. Finally, we discussed and showed the effectiveness of proposed countermeasures to mitigate or reduce the damaging consequences of the studied attacks.

In closing, we recommend the following next directions:

- System operators should reconsider the current unnecessary online data sharing mechanisms and policies. Access to historical and real-time system data can be easily leveraged for malicious purposes.
- Further research is required to develop additional MadIoT attacks and effective protection schemes to help the power grid withstand emerging high-wattage botnet attacks.

Acknowledgements

We would like to sincerely appreciate the comments from NERC cybersecurity team which greatly improved the manuscript quality. We also thank the anonymous reviewers who provided us with constructive feedback, our shepherd Ryan Gerdes, and Bing Huang and Ross Baldick who reviewed an early version of this paper. This work is supported by NSF grants CNS-1929580, CNS-1929406, CNS-1929410, and CNS-1931573 and AFRL/DHS grant FA8750-19-2-0010.

References

[1] K. Zetter. (July 2018) Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid. [Online]. Available: https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/

[2] J. Kennedy, "Dragonfly: Western energy sector targeted by sophisticated attack group." [Online]. Available: https://www.symantec.com/blogs/threat-intelligence/dragonfly-energy-sector-cyber-attacks

[3] R. M. Lee, M. J. Assante, and T. Conway, "ICS defense use case: Analysis of the cyber attack on the Ukrainian power grid," *Electricity Information Sharing and Analysis Center, SANS ICS*, 2016.

[4] M. Zeller, "Myth or reality – does the Aurora vulnerability pose a risk to my generator?" in *64th Ann. Conf. for Protective Relay Engineers*, 2011, pp. 130–136.

[5] S. Soltan, P. Mittal, and H. V. Poor, "BlackIoT: IoT botnet of high wattage devices can disrupt the power grid," in *27th USENIX Security Symp.*, 2018, pp. 15–32.

[6] B. Huang, A. A. Cardenas, and R. Baldick, "Not everything is dark and gloomy: Power grid protections against IoT demand attacks," in *28th USENIX Security Symp.*, 2019, pp. 1115–1132.

[7] T. Shekari, C. Irvene, A. A. Cardenas, and R. Beyah, "MaMIoT: Manipulation of energy market leveraging high wattage IoT botnets," in *Proc. of the ACM SIGSAC Conf. on Computer and Commun. Security (CCS)*. ACM, 2021, pp. 1338–1356.

[8] "MaDIoT 2.0." [Online]. Available: https://github.com/MaDIoT20/MaDIoT-2.0.git

[9] T. Shekari, C. Bayens, M. Cohen, L. Graber, and R. Beyah, "RFDIDS: Radio frequency-based distributed intrusion detection system for the power grid." in *Network and Distributed System Security (NDSS) Symp.*, 2019, pp. 1–15.

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

[11] K.-P. Brand, V. Lohmann, and W. Wimmer, *Substation automation handbook*. Utility Automation Consulting Lohmann Bremgarten, Switzerland, 2003.

[12] M. Eremia and M. Shahidehpour, *Handbook of electrical power system dynamics: modeling, stability, and control*. John Wiley & Sons, 2013, vol. 92.

[13] G. Ziegler, *Numerical distance protection: principles and applications*. John Wiley & Sons, 2011.

[14] F. Sui, A. Rezaei-Zare, M. Kostic, and P. Sharma, "A method to assess GIC impact on zero sequence overcurrent protection of transmission lines," in *IEEE Power & Energy Society General Meet*, 2013, pp. 1–5.

[15] M. Begovic, D. Fulton, M. Gonzalez, J. Goossens, E. Guro, R. Haas, C. Henville, G. Manchur, G. Michel, R. Pastore *et al.*, "Summary of" system protection and voltage stability"," *IEEE Trans. Power Del.*, vol. 10, no. 2, pp. 631–638, 1995.

[16] T. Shekari, F. Aminifar, and M. Sanaye-Pasand, "An analytical adaptive load shedding scheme against severe combinational disturbances," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4135–4143, 2015.

[17] T. Shekari, A. Gholami, F. Aminifar, and M. Sanaye-Pasand, "An adaptive wide-area load shedding scheme incorporating power system real-time limitations," *IEEE Syst. J.*, vol. 12, no. 1, pp. 759–767, March 2018.

[18] Z. Song, Y. Lin, C. Liu, Z. Ma, and L. Ding, "Review on over-frequency generator tripping for frequency stability control," in *IEEE PES Asia-Pacific Power and Energy Engineering Conf. (APPEEC)*, 2016, pp. 2240–2243.

[19] A. Guzman, Z. Zocholl, G. Benmouyal, and H. J. Altuve, "A current-based solution for transformer differential protection. i. problem statement," *IEEE Trans. Power Del.*, vol. 16, no. 4, pp. 485–491, 2001.

[20] S. Dambhare, S. Soman, and M. Chandorkar, "Adaptive current differential protection schemes for transmission-line protection," *IEEE Trans. Power Del.*, vol. 24, no. 4, pp. 1832–1841, 2009.

[21] D. A. Tziouvaras and D. Hou, "Out-of-step protection fundamentals and advancements," in *57th Ann. Conf. for Protective Relay Engineers*, 2004, pp. 282–307.

[22] J. Berdy, "Loss of excitation protection for modern synchronous generators," *IEEE Trans. Power Apparatus and Syst.*, vol. 94, no. 5, pp. 1457–1463, 1975.

[23] "Bloomberg Terminal." [Online]. Available: https://en.wikipedia.org/wiki/Bloomberg_Terminal

[24] New York Independent System Operator, "Load Data." [Online]. Available: https://www.nyiso.com/load-data

[25] California Independent System Operator, "Reliability Requirements." [Online]. Available: http://www.caiso.com/planning/Pages/ReliabilityRequirements/Default.aspx#Historical

[26] Pennsylvania and New Jersey Independent System Operator, "Energy Market." [Online]. Available: https://www.pjm.com/markets-and-operations/energy.aspx

[27] J. Wilson, "Phishing Attacks: Why Energy Companies and Utilities Are Getting Zapped." [Online]. Available: https://www.agari.com/email-security-blog/phishing-attacks-why-energy-companies-and-utilities-are-getting-zapped/

[28] A. Keliris, C. Konstantinou, M. Sazos, and M. Maniatakos, "Open source intelligence for energy sector cyberattacks," in *Critical Infrastructure Security and Resilience*. Springer, 2019, pp. 261–281.

[29] "Google Maps," 2020. [Online]. Available: https://www.google.com/maps/

[30] C. Arderne, C. Zorn, C. Nicolas, and E. Koks, "Predictive mapping of the global power system using open data," *Scientific data*, vol. 7, no. 1, pp. 1–12, 2020.

[31] H. Kim, D. Olave-Rojas, E. Álvarez-Miranda, and S.-W. Son, "In-depth data on the network structure and hourly activity of the central chilean power grid," *Scientific data*, vol. 5, no. 1, pp. 1–10, 2018.

[32] W. Medjroubi, U. P. Müller, M. Scharf, C. Matke, and D. Kleinhans, "Open data in power grid modelling: new approaches towards transparent grid models," *Energy Reports*, vol. 3, pp. 14–21, 2017.

[33] J. Tang, J. Liu, F. Ponci, and A. Monti, "Adaptive load shedding based on combined frequency and voltage stability assessment using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 2035–2047, 2013.

[34] B. Gao, G. Morison, and P. Kundur, "Voltage stability evaluation using modal analysis," *IEEE trans. Power Syst.*, vol. 7, no. 4, pp. 1529–1542, 1992.

[35] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-hill New York, 1994, vol. 7.

[36] "DIgSILENT PowerFactory," 2020. [Online]. Available: https://www.digsilent.de/en/

[37] "IEEE 39-Bus System," 2020. [Online]. Available: https://icseg.iti.illinois.edu/ieee-39-bus-system/

[38] P. M. Anderson, *Power system protection*. Wiley, 1998.

[39] (December 2021) EIA's Hourly Electric Grid Monitor provides timely data about electricity usage patterns. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=43295

[40] (December 2017) Homes with Samrt Thermostats in North America. [Online]. Available: https://www.statista.com/statistics/625868/homes-with-smart-thermostats-in-north-america/

[41] (December 2017) EEI Celebrates 1 Million Electric Vehicles on U.S. Roads. [Online]. Available: https://www.eei.org/resourcesandmedia/newsroom/Pages/Press%20Releases/EEI%20Celebrates%201%20Million%20Electric%20Vehicles%20on%20U-S-%20Roads.aspx#:~:text=EEI%2FIEI%20report%3A-,The%20number%20of%20EVs%20on%20U.S.%20roads%20is%20projected%20to,on%20U.S.%20roads%20in%202030

[42] M. Abedini, M. Sanaye-Pasand, and S. Azizi, "Adaptive load shedding scheme to preserve the power system stability following large disturbances," *IET Gen., Transm. & Dist.*, vol. 8, no. 12, pp. 2124–2133, 2014.

[43] X. Qin, K. Mai, N. Ortiz, K. Koneru, and A. A. Cardenas, "Cybersecurity and resilience for the power grid," *Resilient Control Architectures and Power Systems*, 2021.

[44] C. Barreto, J. Giraldo, A. A. Cardenas, E. Mojica-Nava, and N. Quijano, "Control systems for the power grid and their resiliency to attacks," *IEEE Security & Privacy*, vol. 12, no. 6, pp. 15–23, 2014.

[45] S. Sridhar, "Cyber risk modeling and attack-resilient control for power grid," *Ph.D. dissertation, Dept. Elect. Comput. Eng., Iowa State Univ., Ames, IA, USA*, 2015.

[46] K. G. Boroojeni, M. H. Amini, and S. S. Iyengar, *Smart grids: security and privacy issues*. Springer, 2017.

[47] A. Gholami, T. Shekari, M. H. Amirioun, F. Aminifar, M. H. Amini, and A. Sargolzaei, "Toward a consensus on the definition and taxonomy of power system resilience," *IEEE Access*, vol. 6, pp. 32 035–32 053, 2018.

[48] J. Kim, C. Lee, H. Shim, Y. Eun, and J. H. Seo, "Detection of sensor attack and resilient state estimation for uniformly observable nonlinear systems having redundant sensors," *IEEE Trans. Autom. Control*, vol. 64, no. 3, pp. 1162–1169, 2019.

[49] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inform. and Syst. Security*, vol. 14, no. 1, pp. 1–33, 2011.

[50] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[51] C. Lee, H. Shim, and Y. Eun, "Secure and robust state estimation under sensor attacks, measurement noises, and process disturbances: Observer-based combinatorial approach," in *European Control Conf. (ECC)*, 2015, pp. 1872–1877.

[52] Y. Shoukry, P. Nuzzo, N. Bezzo, A. L. Sangiovanni-Vincentelli, S. A. Seshia, and P. Tabuada, "Secure state reconstruction in differentially flat systems under sensor attacks using satisfiability modulo theory solving," in *54th IEEE Conf. on Decision and Control (CDC)*, 2015, pp. 3804–3809.

[53] J. Meserve, "Mouse click could plunge city into darkness, experts say." [Online]. Available: http://www.cnn.com/2007/US/09/27/power.at.risk/index.html

[54] "FOIA response documents." [Online]. Available: http://s3.documentcloud.org/documents/1212530/14f00304-documents.pdf

[55] "Dragonfly: Western energy sector targeted by sophisticated attack group." [Online]. Available: https://www.symantec.com/blogs/threat-intelligence/dragonfly-energy-sector-cyber-attacks

[56] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: Implications for false data injection attacks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3317–3318, 2017.

[57] L. Garcia, F. Brasser, M. H. Cintuglu, A.-R. Sadeghi, O. A. Mohammed, and S. A. Zonouz, "Hey, my malware knows physics! attacking PLCs with physical model aware rootkit." in *Network and Distributed System Security (NDSS) Symp.*, 2017, pp. 1–15.

[58] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet-based load altering attacks against smart power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 667–674, 2011.

[59] Z. Xu, H. Wang, Z. Xu, and X. Wang, "Power attack: An increasing threat to data centers," in *Network and Distributed System Security (NDSS) Symp.*, 2014, pp. 1–15.

[60] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad, "Dynamic load altering attacks against power system stability: Attack models and protection schemes," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2862–2872, 2016.

[61] Y. Dvorkin and S. Garg, "IoT-enabled distributed cyber-attacks on transmission and distribution grids," in *North American Power Symp. (NAPS)*, 2017, pp. 1–6.

[62] A. Dabrowski, J. Ullrich, and E. R. Weippl, "Grid shock: Coordinated load-changing attacks on power grids: The non-smart power grid is vulnerable to cyber attacks as well," in *Proc. of the 33rd Ann. Computer Security Applications Conf. (ACSAC)*, 2017, pp. 303–314.

[63] C. Barreto, A. A. Cárdenas, N. Quijano, and E. Mojica-Nava, "CPS: Market analysis of attacks against demand response in the smart grid," in *Proceedings of the 30th Annual Computer Security Applications Conference*, 2014, pp. 136–145.

[64] M. Antonakakis *et al.*, "Understanding the Mirai botnet," in *26th USENIX Security Symp.*, 2017, pp. 1093–1110.

[65] S. Edwards and I. Profetis, "Hajime: Analysis of a decentralized internet worm for IoT devices," *Rapidity Networks*, vol. 16, 2016.

[66] P. Pierluigi, "LuaBot is the first Linux DDoS botnet written in Lua Language." [Online]. Available: https://securityaffairs.co/wordpress/51155/malware/linux-luabot.html

[67] G. Dan, "BrickerBot, the permanent denial-of-service botnet, is back with a vengeance." [Online]. Available: https://arstechnica.com/information-technology/2017/04/brickerbot-the-permanent-denial-of-service-botnet-is-back-with-a-vengeance/

[68] H. Jiang, Y. Liu, and J. N. Matthews, "IP geolocation estimation using neural networks with stable landmarks," in *IEEE Conf. Comp. Communications Workshops*, 2016, pp. 170–175.

[69] [Online]. Available: https://geo.ipify.org/blog/ip-geolocation-accuracy-how-reliable-is-the-technology

[70] C. Cimpanu. (2016) You Can Now Rent a Mirai Botnet of 400,000 Bots. [Online]. Available: https://www.bleepingcomputer.com/news/security/you-can-now-rent-a-mirai-botnet-of-400-000-bots/

[71] D. Goodin. (2018) New IoT botnet offers DDoSes of once-unimaginable sizes for $20. [Online]. Available: https://arstechnica.com/information-technology/2018/02/for-sale-ddoses-guaranteed-to-take-down-gaming-servers-just-20/

[72] R. Joven and E. Ananin. (2018) DDoS-for-Hire Service Powered by Bushido Botnet. [Online]. Available: https://www.fortinet.com/blog/threat-research/ddos-for-hire-service-powered-by-bushido-botnet-.html

[73] D. Goodin. (December 2017) 100,000-strong botnet built on router 0-day could strike at any time. [Online]. Available: https://arstechnica.com/information-technology/2017/12/100000-strong-botnet-built-on-router-0-day-could-strike-at-any-time/

[74] S. Boulevard. (2018) Here's how anyone with $20 can hire an IoT botnet to blast out a week-long DDoS attack. [Online]. Available: https://securityboulevard.com/2018/08/heres-how-anyone-with-20-can-hire-an-iot-botnet-to-blast-out-a-week-long-ddos-attack/

[75] imperva. (2019) Booters, Stressers and DDoSers. [Online]. Available: https://www.imperva.com/learn/application-security/booters-stressers-ddosers/

[76] Electricity Consumers Resource Council (ELCON), "The Economic Impacts of the August 2003 Blackout ." [Online]. Available: https://elcon.org/wp-content/uploads/Economic20Impacts20of20August20200320Blackout1.pdf

[77] US Energy Information Administration, "2018 Average Monthly Bill- Residential," 2020. [Online]. Available: https://www.eia.gov/electricity/sales_revenue_price/pdf/table5_a.pdf

[78] F. D. Zahlay and K. R. Rao, "Neuro-prony and taguchi's methodology-based adaptive autoreclosure scheme for electric transmission systems," *IEEE Trans. Power Del.*, vol. 27, no. 2, pp. 575–582, 2012.

9 Attack Feasibility

9.1 From the IoT Botnet Perspective

One of the main questions about manipulation of demand attacks is how feasible it is to gain control of a botnet of high-wattage devices within a limited geographical area. Historically, the number of IoT botnets in recent years has increased dramatically, with famous IoT botnets including Mirai, LuaBot, Hajime, and BrickerBot [64–67]. Similar to previous work [5, 6], we assume that the attacker has access to a high-wattage IoT botnet.

The key aspect of our attack is that the adversary needs to activate the attack only in specific geographical areas (such as a city). The differentiation between the location of the compromised bots can be initially done by classifying and using their IP addresses. There are various research efforts providing IP location with median errors of just 3.4km [68], and free and commercial IP geolocation databases claiming to locate cities (the nodes we are interested in for our study) with an accuracy of over 85% [69].

Compared to previous work [5, 6], MaDIoT 2.0 attack requires fewer IoT bots to be successful. The reason is that MaDIoT 2.0 attacks target the most vulnerable nodes of the grid instead of launching attacks spread over all nodes. As shown in Section 5, an adversary with 150,000 bots can effectively take down a typical power grid with MaDIoT 2.0 attacks. Considering IoT botnets such as Mirai had over six hundred thousand compromised devices [64], the future existence of a high-wattage IoT botnet with 150,000 bots is not unlikely.

In practice, there are many high-wattage IoT devices that can be compromised in each grid node. Smart HVAC systems and EVs are only two classes of high-wattage devices that could be compromised and leveraged by the attacker as a part of his botnet. There are currently more than 30 million smart HVAC systems in the north America region [40]. Also, the number of EVs on U.S. roads is projected to reach 18.7 million in 2030, up from 1 million at the end of 2018. This is about 7% of the 259 million vehicles (cars and light trucks) expected to be on U.S. roads in 2030 [41].

Entrepreneurial attackers can compromise high-wattage devices and then offer them for rent. This practice is common in current botnets [70–75]. The available botnet rental services provide clients with the capability to launch a limited or unlimited (for premium users) number of attacks per day with a guaranteed minimum duration from minutes to hours. Since MaDIoT 2.0 attacks take less than a minute, all of the currently available botnets satisfy this time requirement. The cost of renting a typical IoT botnet is negligible compared with the cost of a typical blackout. For example, Anderson Economic Group (AEG) estimates the likely total cost of the 2003 northeast US blackout to be between $4.5 and $8.2 billion [76].

9.2 From the End User's Perspective

In order to make the attack repeatable, the adversary should try to keep it as stealthy as possible. From the end user's perspective, we discuss i) the effect of the attack on the billing statement of the homeowners, and ii) the possibility of attack detection and prevention in each home/building. The financial effect of the MaDIoT 2.0 attack on each of the homeowners depends on the duty cycle of the attack as well as the total power consumption at home. As mentioned earlier, MaDIoT 2.0 attacks use the compromised high-wattage devices for less than one minute. Therefore, even if the adversary launches this attack multiple times each month, its effect in the household is minimal. According to the energy information administration (EIA), the average electricity consumption of Americans is approximately 914 kWh per month. Tennessee has the highest electricity consumption at 1,282 kWh per residential customer, and Hawaii has the lowest at 517.75 kWh per residential customer [77]. Assuming that each of the high-wattage IoT bots consumes 3 kW of electricity and considering the duration of multiple typical attacks in each month (30 minutes), each compromised home will pay 0.11%-0.28% additional payment for electricity bills, which we believe is unnoticeable. To answer the second question, the possibility of the MaDIoT 2.0 attack is detected by home or device owners is almost negligible because the duration of the attack is very short (e.g., 10 seconds) to raise any suspicions. In addition to this, even if the home owner notices the unwanted activation/deactivation of the compromised high-wattage devices, he would likely think that it is happening because of a software bug in the device and a simple restart would resolve the issue but it is already too late and the blackout has already occurred (the attack usually takes less 30 seconds to cause a blackout in the entire system). Note that in the worst case, individual bot detection and losing few bots would not thwart

the entire attack.

10 Basic Data of Test Cases

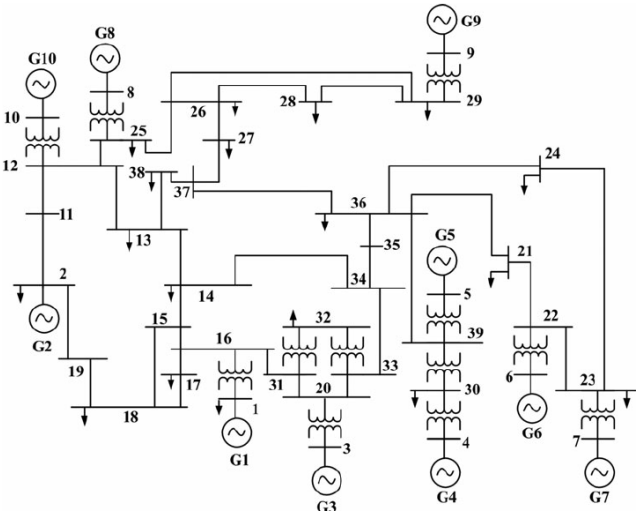

Figure 17: The single-line diagram of the New England power grid (IEEE 39-bus test system) [37].

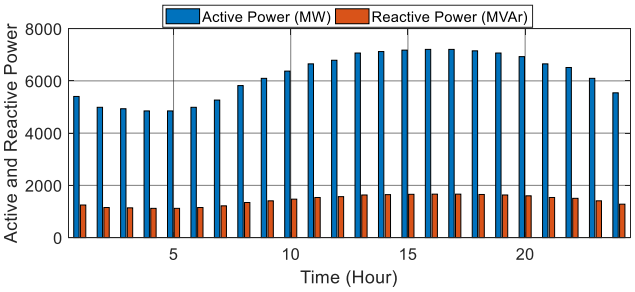

Figure 18: The active and reactive powers of the New England power grid during the evaluation period.

In this Section, the basic data for simulating the test systems were given, which might be used in future studies as benchmark cases. The single-line diagram of the New England power grid is shown in Figure 17 [37]. The total base load of the system is 6097.1 MW (active load)and 1408.9 MVAr (reactive load). To expand the base load, we used a daily load profile in the grid meaning that at every minute, the load of the grid nodes changes similar to the real-world cases. The daily load profile we consider is illustrated in Figure 18 for the New England power grid. The dynamic and static parameters of the system generators are outlined in Table 3.

The grid graph or the single-line diagram of the IEEE 9-bus test system is illustrated in Figure 19 [78]. For the IEEE 9-bus system, we used a relatively similar load curve shown in Figure 20. Also, the dynamic and static parameters of the system generators in this system are outlined in Table 4.

Table 3: Generator Parameters for the New England Power Grid

G	H	x'_d	x'_q	x_d	x_q	T'_{do}	T'_{qo}	x_l
G1	500	0.006	0.008	0.02	0.019	7.0	0.7	0.003
G2	30.3	0.0697	0.170	0.295	0.282	6.56	1.5	0.035
G3	35.8	0.0531	0.0876	0.2495	0.237	5.7	1.5	0.0304
G4	28.6	0.0436	0.166	0.262	0.258	5.69	1.5	0.0295
G5	26.0	0.132	0.166	0.67	0.62	5.4	0.44	0.054
G6	34.8	0.05	0.0814	0.254	0.241	7.3	0.4	0.0224
G7	26.4	0.049	0.186	0.295	0.292	5.66	1.5	0.0322
G8	24.3	0.057	0.0911	0.290	0.280	6.7	0.41	0.028
G9	34.5	0.057	0.0587	0.2106	0.205	4.79	1.96	0.0298
G10	42.0	0.031	0.008	0.1	0.069	10.2	0.0	0.0125

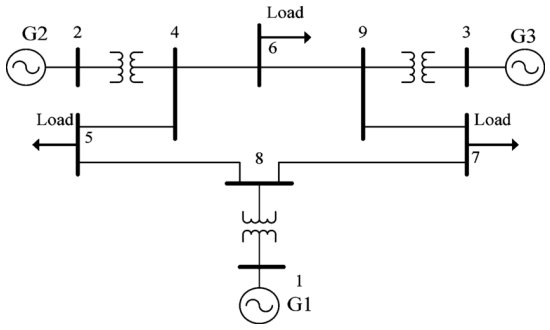

Figure 19: The single-line diagram of the IEEE 9-bus test system [78].

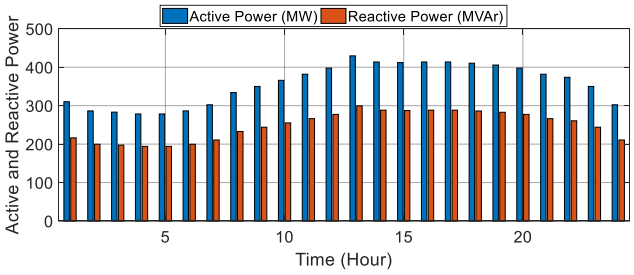

Figure 20: The active and reactive powers of the IEEE 9-bus test system during the evaluation period.

Table 4: Generator Parameters for the IEEE 9-Node System

G	H	x'_d	x'_q	x_d	x_q	T'_{do}	T'_{qo}	x_l
G1	0	0.0608	0.0969	0.146	0.0969	8.96	0	0.0032
G2	0	0.1198	0.1969	0.8958	0.8645	6.0	0.535	0.026
G3	0	0.1813	0.25	1.3125	1.2578	5.89	0.6	0.036