Securing Power Distribution Grid Against Power Botnet Attacks

Lizhi Wang, Lynn Pepin, Yan Li, Fei Miao, Amir Herzberg, and Peng Zhang

***Abstract*— A botnet is a collection of internet-facing devices that are compromised and controlled by a malicious hacker. In this paper, we propose an attack utilising a botnet of high-wattage internet-facing devices, which we call a power botnet. Power botnet attacks can decrease the reliability of power supply, damage the power quality and even cause catastrophic consequences in power distribution grid. To study the effects on power distribution systems, we simulate three different types of power botnet attacks using OpenDSS, and show the change of OLTC lifespans under attacks. We then use deep learning methods to detect these attacks. We show successful detection for two of these attacks and a low detection rate for the third attack. To the best of our knowledge, this is the first paper to consider power botnet attacks, and leverage deep learning methods to detect these attacks on power distribution grids. Future work such as detection schemes for more complicated power botnet attacks will be developed based on the results of this work.**

***Index Terms*— Cyber Security, Power Botnet, Load altering attack, Machine Learning, Attack Detection**

I. INTRODUCTION

The modern cyber-physical system is a target for a wide variety of attacks. Earlier generations of attacks target the inner mechanisms of the power grid, such as SCADA, by traditional information technology attacks, using methods such as phishing or denial-of-service to compromise the grid. But the increasing presence of high-wattage Internet of Things (IoT) devices represents a new attack surface for the power grid. These IoT devices can be controlled remotely via the Internet [1], and are notoriously vulnerable to cyberattacks [2]. A single device controlled by an attacker is known as a **bot**. When such a device is capable of demanding high load, we call it a **power bot**, and we call the collection of such power bots a **power botnet**. A hacker controlling a large enough power botnet can create a specially crafted load in the grid, damaging the stability of the grid or accelerating the degradation of the components inside of it.

High wattage devices such as air conditioners and water heaters are connected to the power-grid and are not part of the infrastructure of the power company. Wi-Fi and Bluetooth-based information communication technologies are being deployed increasingly to create "smart devices". Researchers have found vulnerabilities in many IoT devices that could allow attackers to remotely control them, usually via the Internet. As stated, this is called a bot, and a large number of internet-facing high-wattage bots controlled by a single hacker is known as a power botnet. Attackers can utilize a power botnet to exert a coordinated load change in power grid, performing attacks such as by synchronously switching on or off quantities of high wattage devices or changing the set-points synchronously. We call such an attack a *power botnet attack*. Prior work has used the names "dynamic load-altering attack" (DLAA) [3], "coordinated load changing attack" [2], and "Manipulation of Demand via IoT attack" (MadIoT) [1].

Moving towards defense against this class of attacks, it can be beneficial to be able to detect them. For example, detecting attacks could empower an Advanced Distribution Management System (ADMS) to better protect the grid stability. This is a new class of attacks, and the objectives and methods vary. Unlike traditional botnets [4], which usually perform short-term denial-of-service attacks by brute force, a power botnet attack can be successful with subtle and small influences, making it hard to detect.

To the best of the authors' knowledge, no prior literature exists demonstrating a detection mechanism for power botnet attacks. Reference [5] introduced the cyberattacks on sub-station and overloading of the system through compromised digital relays. Reference [3] examined the attacks on load management system by compromising direct load control command signals, demand side management price signals, or cloud computation load distribution, without considering the attack on IoT devices. Reference [1] introduced the concept of IoT botnet attack on power system, it mainly focused on the impact on transmission system, and no detection method was mentioned. But power botnet attacks, including compromising critical loads and changing the settings of protection relays, can jeopardise the normal operation of the power distribution grid and can be highly dangerous.

Detecting an attack is the *binary classification* [6] problem from machine learning literature. Given a recent history of sensor data from the power grid, we want to identify whether or not it is under attack. We simulate an strategic attack against the power distribution grid, using OpenDSS [7] and the IEEE 123-bus test case. We use the results of these attacks to perform binary classification of the network state, as attack or not-attack

The rest of this paper is organised as follows: In Section II, we introduce and describe power botnet attacks on the power grid. In Section III, we introduce the case study and describe the results of such attacks. In Section IV, we introduce a machine learning approach to detect simulated power botnet

This work was supported in part by the National Science Foundation under Grant ECCS-1831811, in part by Eversource Energy, and in part by the Office of the Provost, University of Connecticut.

L. Wang, Y. Li and P. Zhang are with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269, USA.

L. Pepin, F. Miao and A. Herzberg are with the Department of Computer Science and Engineering, University of Connecticut, Storrs, CT 06269, USA.

978-1-7281-1981-6/19/$31.00 ©2019 IEEE

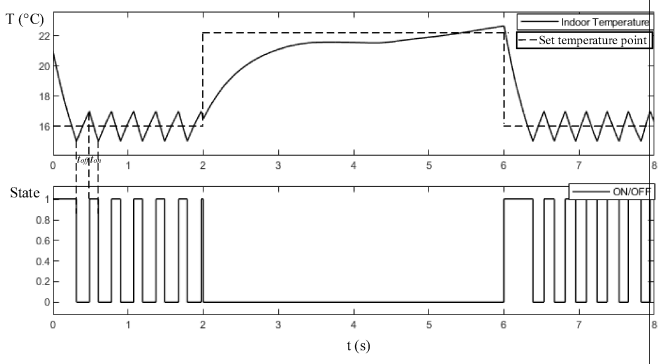

Fig. 1: : Thermal behaviour and power consumption status of air conditioners

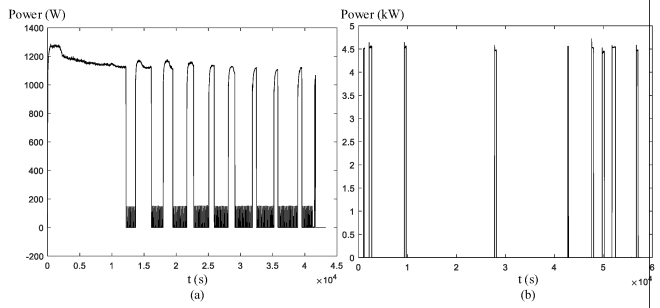

Fig. 2: (a) Power consumption of air conditioner (b) Power consumption of water heater

attacks. In Section V, we discuss the detection result and the simulation of power botnet attack and in section VI we conclude this paper.

II. Attacking Grid Using Power Botnet

A. Introduction to Smart High-wattage appliances

Fig. 1 shows the connection between thermal behaviour and power consumption status of air conditioners. This characteristic provides adversary great chance to manipulate its power consumption by changing the setpoint. Fig. 2 (a) and (b) depict the power consumption of air conditioner and water heater in one day separately. The research data is obtained by monitoring the power consumption of widely used air conditioner LG LW1212ER and water heater E52-50R-045DV [8]. An adversary that has compromised such devices can increase their energy usage by turning them on or increasing their setpoint, thus altering the load of the power grid.

B. Power botnet attack model

The attack surfaces for IoT devices are large. Even if a given IoT device is secure and cannot be directly accessed, an attacker could control it by proxy of any other trusted devices, such as the owner's mobile phone, tablet, or digital home assistant such as Google Home [1]. As is shown in Fig.3, we study power botnet attack on the power distribution system that manipulates power demand side. We assume that an adversary attacks an area by: (i) turning on and off the high wattage smart devices periodically and synchronously,

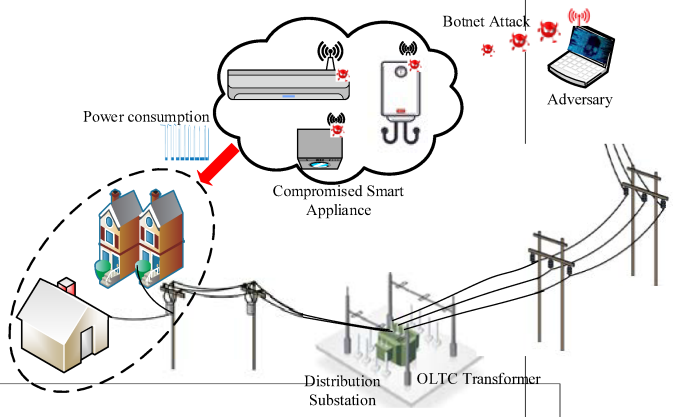

Fig. 3: The power botnet attack model

and (ii) changing the setpoint of air conditioner in the same pattern.

Define the smart device manipulated by adversary as D_a. The load demand changing ΔP of power distribution grid is the mapping of high wattage devices $D_a \rightarrow \Delta P$. $P(t)$, the total load demand is composed of two parts, can be represented as $P(t) = f(P_0, \Delta P)$, where P_0 is the normal load profile. The status of D_a can be represented in (1), where $A \in \{A_{on}, A_{off}\}$ is the control signal by adversary and $T(n)$ is the time sequence defining when the attack happens, in which $T(n)$ can represent how long one attack happens.

$$D_a = \{A,\ T\} \tag{1}$$

where $T = T(n), n = 1, 2, 3, ...$ Then the power load demand can be changed into

$$P(n) = \begin{cases} P_0, & \text{given } A = A_{off}, T(n) = 0 \\ P_0 + \Delta P, & \text{given } A = A_{on}, T(n) = 1 \end{cases} \tag{2}$$

We consider the power botnet attack that a large scale of air conditioners and water heaters are on or off synchronously; and the set points of these high wattage devices are changed. We assume that the attacker performs any of the following three types of power botnet attacks:

1) D_a^I: Attacker's goal is to wear down power system equipment as fast as possible. The time sequence $T(n)$ of attack D_a^I can be $\{010101...\}$. The first attack D_a^I using the alternating pattern to manipulate large scale of high wattage devices synchronously. When the adversary use the method of changing setpoint or turning on devices directly at time $T(0)$, the high wattage devices can be manipulated remotely and synchronously.
2) D_a^{II}: Attackers can remotely change the status of high wattage devices periodically. In each period, the attack can cause long term effect on power system. The time sequence $T(n)$ of attack D_a^{II} can be $\{00001111000000111....\}$, where the lengths of each continuous sequence has a random length.
3) D_a^{III}: Smart attackers using some strategic attack can cause damage to power system with being detected.

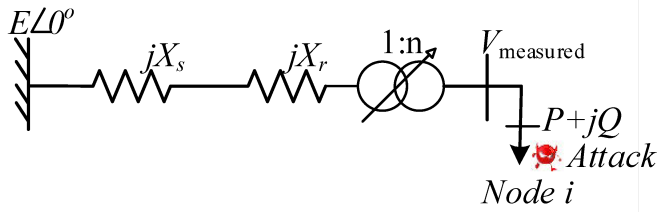

Fig. 4: One line diagram of a feeder with OLTC transformer

The randomly attack can be an useful method to deceive DMS. Attackers can leave no clue for DMS to detect such attack. The value of time sequence $T(n)$ of attack D_a^{III} can be randomly distributed as $\{0010010000100110001.....\}$, where each timeslot individually has a 20% chance to be chosen for attack.

C. Power distribution grid performance under power botnet attack

The power distribution grid can experience extreme condition or even cascading failure attacked by power botnet. The drastic fluctuation of load demand can effect the power quality and reliability of power distribution system; and increase the cost of system operation due to the physical damage to hardware equipment in the system.

1) Attacks can cause the tap changer of OLTC change frequently, which can wear down rapidly this mechanism by causing changes in consumption that will force the OLTC to be invoked and change taps more quickly than under typical conditions. In Fig.4, automatic voltage regulator measures the secondary voltage of transformer ($V_{measured}$) and compares it with a certain reference voltage (V_{ref}). A deadband (ΔU) is set and during normal operating conditions the following relationship is fulfilled [9]. Therefore when the voltage difference in equation (5) caused by power botnet attack becomes greater than the deadband, the regulator change the tap position of OLTC.

$$V_{ref} - V_{measured} < \Delta U \tag{3}$$

2) Attacks can influence the quality of power. As depicted in Fig.4, the increasing of active power consumption in bus node i makes the line voltage drop increase, which leads to the decreasing of voltage V_i. This attack can increase the frequency of low voltage violation in power distribution system. So the quality of power can be affected by such attack.

III. Case Study

In this section, we illustrate the performance of tap changer of OLTC under power botnet attack using the IEEE 123-bus power test system as shown in Fig.5 and detailed in [10]. Let the red nodes represent locations that attackers can manipulate smart devices at, the green nodes represent locations which have no devices under control by an attacker, and the orange nodes represent OLTC regulating transformers. For any point in time, the state vector of d-dimension $\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \mathbf{x}_i, ..., \boldsymbol{x}_{123} \in \mathbb{R}^{d_i}$ is composed of

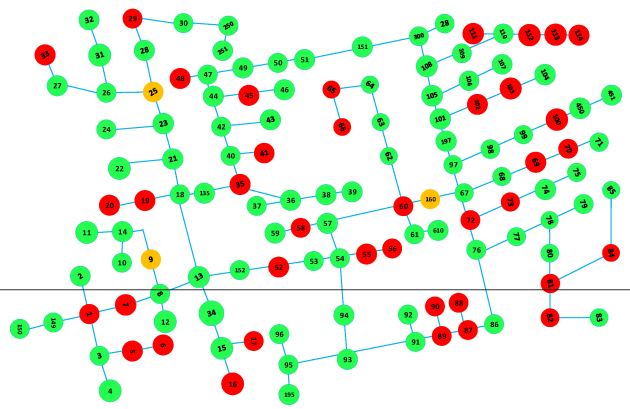

Fig. 5: IEEE 123-bus power test system

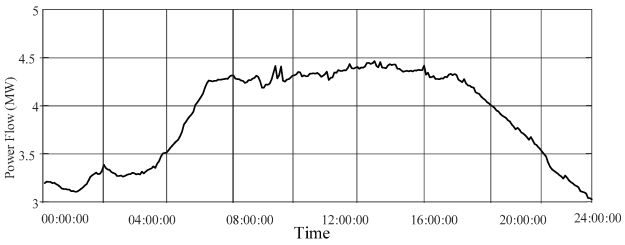

Fig. 6: One example normal load shape

the voltage, current, power, and respective at each of the individual busses, and d_i is the dimension of the state vector f each node. The state of the grid is given by the collection of all such state vectors.

OpenDSS is used to simulate this attack and to record the state vectors of each bus by performing power flow analysis. Fig.6 depicts the normal load profiles of the IEEE 123-bus. The spot load is adopted to evaluate the scale of botnets in each bus. Take bus 42 as an example, the spot load of bus 42 is 20 kW. We assume the average consumption of each house is 5 kW; each house has two air conditioners and one water heater. Then we can evaluate the performance of the system with different percentage of D_a.

Changing the power consumption frequently can be a fast way to wear down OLTC. We assume that the attacker use attack D_a^{II} to manipulate power botnet every 5 minutes. Fig.7 showed the lifespan of OLTC at node 160 with different percentage of devices attacked. Assuming maximal numbers of changes of OLTC is 5000 and all the devices at the red nodes can be attacked. The average life span of OLTC can be 13.70 years under normal condition. However, it takes only 0.52 years when attackers manipulate 80% of all the devices at red nodes to wear down the OLTC.

IV. Deep Learning Method for Power Botnet Attack Detection

We design a binary classification detection scheme for attacks described in the previous section. Three types of attacks were performed and recorded using OpenDSS simulations.

The first type of attacks D_a^I were performed in discrete 1-minute time steps, with 1440 time steps recorded in total.

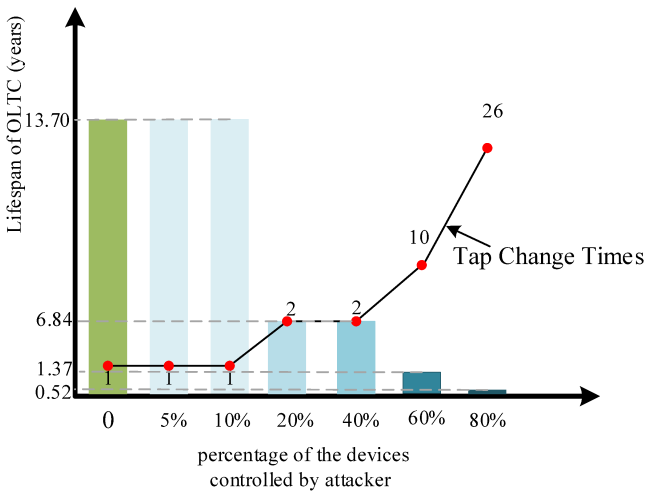

Fig. 7: Lifespan of OLTC under different attack condition

Attacks were performed in an off-on-off-on pattern. The second type of attacks D_a^{II} were performed using long, alternating, continuous off-on-off-on patterns, with 10000 time steps in total. The third type of attacks D_a^{III} involved random attacks, where each time step independently had a 20% chance for an attack to occur within it. For all three attacks cases, the goal is to develop a detector to classify the grid power grid as 'attack' or 'no attack', given the state of the grid as recorded at each of the three OLTC nodes.

The rest of this section is organized as follows: The first subsection describes data-preprocessing and the input data for the neural network. The following subsections describe neural network model structures and the results for detecting attack D_a^I, D_a^{II}, and D_a^{III}, respectively.

A. Data Preprocessing for Neural Network

a) Time features: Any information encoding the time or times slot was removed from the data. This is to prevent the neural network from learning the time slots that attacks occur.

b) Data Normalization: Each node in the power system has a vector that describes the state of the node at each time slot. Each individual value is normalized across its entire history using $[0, 1]$ normalization, except for angles and per-unit tap values. Angles are converted from degrees to an encoded pair $(\sin x, \cos x)$. Per-unit tap values are left unnormalized.

c) The state vector $\boldsymbol{x}_t$: For a given time slot integer $t \in [1, 1440]$, the state of the network is described by the vector $\mathbf{x}_t \in \mathbb{R}^d$, where $d = 54$. This state vector representing the grid is the concatenation of each state vector for each OLTC node in the power grid. The output of the neural network (the target classification value), $\mathbf{y}_t$, is a vector with value $\langle 1, 0 \rangle$ if there is no attack, or $\langle 0, 1 \rangle$ if there is an attack.

d) The state history $X_{(a,b)}$: We use the **sliding window** technique to generate a short-term *state history* X_t of the power grid up to time t. For $a, b \in [1, 1440]$, let $X_{(a,b)}$ be the matrix $[\mathbf{x}_a, \mathbf{x}_{a+1}, ..., \mathbf{x}_b]$, representing a *history* of state vectors. We say $T = b - a + 1$ is the *window size*. That is, T is the amount of state vectors we consider for designing the detector neural network. Given this definition for T, we say $X_t = X_{(t-T,t)}$.

	Predicted no attack	Predicted attack
No attack	4354 ± 17.6	21.50 ± 3.17
Attack	$28.30 + \pm 6.60$	592.9 ± 17.0

TABLE I: Prediction confusion matrix for attack D_a^{II}

e) Data split: The data is processed to create a dataset STATE $= [X_{T-1}, X_T, X_{T+1}, ..., X_T]$, and TARGET $= [\mathbf{y}_{T-1}, \mathbf{y}_T, \mathbf{y}_{T+1}, ..., \mathbf{y}_T]$. The datasets STATE and TARGET are then identically shuffled. The training split is 50%, and the validation split is 10%. In this way, there is a 50-45-5 testing-training-validation split.

B. Detecting Attack D_a^I

Attack D_a^I was the one which used the simple alternating pattern (i.e. 010101...)

a) Model Architecture and Training Parameters: The network used is a linear classifier, implemented as a neural network using the Keras [11] deep learning library. A window size of $T = 12$ was chosen, so the input is a 54×12 matrix. This matrix is flattened and then fully connected to a 2-unit output layer with softmax activation.

b) Experiment and results: The model was then trained over 40 epochs with a batch size of 32, using the binary crossentropy loss function and the Adam optimizer. By default, Keras initializes the Adam optimizer with learning rate 0.001, beta 1 of 0.9, beta 2 of 0.999. [12] 20 trials were performed, and for each of these 10 trials, the model had 100% accuracy, properly classifying every tested value.

C. Detecting Attack D_a^{II}

Attack D_a^{II} was the one which used the alternating pattern with longer sequences with random lengths (i.e. 00000001111111000...). For this generated dataset, out of 10,000 time slots, 1249 were attacks.

a) Model Architecture and Training Parameters: The same network is used to detect attack D_a^{II} as for D_a^I, with the same loss function and optimizer, but a window size of $T = 6$ is chosen. The network is the same linear classifier as the network detecting attack D_a^I, except taking a 54×6 matrix for input.

b) Experiment and results: The model was fit over 20 epochs, and the experiment was repeated 10 times. A confusion matrix is given in Table I. Note: The test set had 4997 samples. Per Table I, on average, there were 4354 true negatives, 592.9 true positives, and a combined total of 49.8 false reports. This corresponds to an average classification accuracy of 99.0%.

D. Detecting Attack D_a^{III}

Attack D_a^I was the one which used 10,000 timeslots, where each timeslot was given an individual 20% chance for the grid to be under attack. This means a simple strategy

	Predicted no attack	Predicted attack
No attack	4020 ± 16.8	$.9000 \pm 0.30$
Attack	$973.5 + \pm 16.8$	$.1000 \pm 0.30$

TABLE II: Prediction confusion matrix for attack D_a^{III}

of assuming no attack should give a minimum accuracy of 80%.

a) Model Architecture and Training Parameters: The model chosen is a multilayer perceptron [6], implemented in Keras, using the same loss and optimisation functions as before. Window size of $T = 12$ was chosen, making the input a 54×12 matrix. The network is described as follows:

1) **Flatten** layer, converts input into vector of size 648.
2) **FC** (fully connected, or *dense*) layer with 256 units, 50% dropout, ReLU activation. [6]
3) **FC** layer, 64 units, 50% dropout, ReLU activation.
4) **FC** layer, 256 units, 50% dropout, ReLU activation.
5) **FC** output layer, 2 units, softmax activation.

b) Experiment and results: The model was fit over 20 epochs, and the experiment was repeated 10 times. Per Table II, on average, there were very few positive predictions, false or true. There were 4020 true negative predictions, and roughly 973.5 false negative predictions. This matrix corresponds to the minimum prediction accuracy of 80%.

V. Discussion

For the two types of attacks D_a^I and D_a^{II}, the neural network detector provides high detection accuracy. This means the power botnet attack detection is feasible under certain scenarios. A linear model and a small amount of input data (sourced from only three OLTC nodes) is used for detection in this work, and in the future we aim to develop more complicated detection mechanisms to match more complicated and subtle power botnet attacks based on this first step trial.

For attack type D_a^{III}, we had prediction accuracy of 80%, corresponding to the minimum accuracy the model should achieve. Note that the neural network detection model for attack D_a^{III} is nonlinear with multiple layers, but is still unable to learn a relationship between the network state and the attack. This may result from the unstructured attack pattern, more analysis about the attack type and detection schemes will be a direction of future work.

Three types of power botnet attacks, the impacts of these attacks and the corresponding detection schemes are shown in this work. In the future, we will analyse more properties and impacts of these attacks, in order to develop better and more robust detection schemes including unsupervised learning techniques. To aid the ADMS in responding to these attacks, attack mitigation and resilient management techniques will also be pursued based on the detection schemes. In anticipation of real-world applicability, future experiment will also involve hardware test-beds.

VI. Conclusions

Power botnet attacks on power distribution grid and the preliminary detection method designed based on deep learning were first introduced in this paper. The effects of power botnet attacks on power distribution grid were analyzed by calculating the tap change of OLTC transformers. Our simulation results show that manipulating different percentages of power botnets can wear down OLTC in different times. More advanced deep learning methods can be leveraged to detect power botnet attacks, and the detection results can be used for defending power distribution system against power botnet attack in the future.

References

[1] S. Soltan, P. Mittal, and H. V. Poor, "Blackiot: Iot botnet of high wattage devices can disrupt the power grid," 2018.

[2] 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 *Proceedings of the 33rd Annual Computer Security Applications Conference*. ACM, 2017, pp. 303–314.

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

[4] W. Yong, S. H. Tefera, and Y. K. Beshah, "Understanding botnet: From mathematical modelling to integrated detection and mitigation framework," in *2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*. IEEE, 2012, pp. 63–70.

[5] Z. Yang, C.-W. Ten, and A. Ginter, "Extended enumeration of hypothesized substations outages incorporating overload implication," *IEEE Transactions on Smart Grid*, 2017.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[7] D. Montenegro, M. Hernandez, and G. A. Ramos, "Real time opendss framework for distribution systems simulation and analysis," in *2012 Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA)*. IEEE, 2012, pp. 1–5.

[8] V. T. ARI, "Research data," http://www.ari.vt.edu/research-data/, accessed October 4, 2018.

[9] D. E. Mawarni, M. V. M. Ali, P. Nguyen, W. Kling, and M. Jerele, "A case study of using oltc to mitigate overvoltage in a rural european low voltage network," in *Power Engineering Conference (UPEC), 2015 50th International Universities*. IEEE, 2015, pp. 1–5.

[10] F. E. Postigo Marcos, C. Mateo Domingo, T. Gómez San Román, B. Palmintier, B.-M. Hodge, V. Krishnan, F. de Cuadra García, and B. Mather, "A review of power distribution test feeders in the united states and the need for synthetic representative networks," *Energies*, vol. 10, no. 11, p. 1896, 2017.

[11] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[12] ——, "Keras; optimizers; adam," https://keras.io/optimizers/#adam/, 2015.