A Novel Distributed and Stealthy Attack on Active Distribution Networks and a Mitigation Strategy

Pirathayini Srikantha, *Member, IEEE*, Jingyuan Liu, *Student Member, IEEE*, and Jagath Samarabandu, *Senior Member, IEEE*

***Abstract*—Rapid advances in smart devices tremendously facilitate our day-to-day lives. However, these can be exploited remotely via existing cyber vulnerabilities to cause disruption at the physical infrastructure level. In this paper, we discover a novel distributed and stealthy attack that uses malicious actuation of a large number of small-scale loads residing within a distribution network (DN). This attack is capable of cumulatively violating the underlying operational system limits, leading to widespread and prolonged disruptions. A key element of this attack is the efficient use of attack resources, planned via Stackelberg games. To mitigate this type of an attack, we propose a countermeasure strategy which adaptively suppresses adverse effects of the attack when detected in a timely manner. The effectiveness of the proposed mitigation strategy is demonstrated via theoretical convergence studies, practical evaluations and comparisons with the state-of-the-art using realistic load flow and DN infrastructure models.**

I. Introduction

WITH the advent of the internet of things (IoT), a large number of traditionally passive devices are now equipped with communication capabilities and the ability to make intelligent decisions. In the electricity sector, recent advances in the metering infrastructure (e.g. smart meters) and loads (e.g. smart appliances) are revolutionizing power distribution, management and consumption. However, these devices are also associated with well-documented vulnerabilities that can be exploited by an adversary to inflict costly and irreversible system-wide disruptions.

As such, a significant flaw in the Zigbee wireless communication protocol has been identified recently [1]. This protocol is widely used to connect not only consumer-centric cyber-enabled loads but also smart meters and data concentrators that collect, process and transmit power consumption data to the electric power utility (EPU) [2]. The potency of this flaw was highlighted when a drone was used to apply an “over-the-air” adversarial firmware update on smart light bulbs to remotely actuate a set of strobing light patterns [3]. Standard security mechanisms implemented in these smart bulbs were bypassed via side-channel attacks. This work has been successful in exhibiting how a vulnerability in the cyber-domain can be exploited to instigate physically observable disruptions by using readily available low-cost commercial equipment. This case study has significant implications for the electric grid - especially at the distribution network (DN) level.

The DN is a low-voltage system that delivers power to consumers. The proliferation of smart appliances in the residential sector has introduced many points of vulnerability in the DN as these utilize communication protocols such as Zigbee which are associated with well-documented flaws. As smart loads are typically connected to the Internet, an adversary will be able to remotely compromise these in large numbers [4]. These threats are not limited only to consumer-centric devices. Infrastructure management devices such as the smart meters and data concentrators also communicate via these protocols [2]. The effect of actuating a large number of small-scale appliances specifically targeting physical system limits can be devastating as protection devices (e.g. fuses, circuit breakers, etc.) can be triggered and result in widespread service outages [5]. Due to the cumulative nature of the attack, the EPU will have great difficulty isolating misbehaving devices individually. Hence, securing the DN from cyber-attacks using only cryptographic techniques no longer suffices. Active defence mechanisms consisting of power injecting elements that adapt to current conditions of the system are necessary to mitigate the physical effects of such cyber-physical attacks.

In this paper, we show how an adversary may exploit cyber vulnerabilities of power devices to mount a new type of stealthy distributed attack on a power grid at the DN level by targeting system limits. We also propose a defence mechanism to effectively mitigate the adverse effects of such an attack. The attack scheme involves strategically actuating a large number smart appliances that have been compromised in order to overwhelm the physical limits. The effective use of attack resources is planned by applying Stackelberg games. The proposed defence mechanism counteracts this by deploying active power injection resources such as storage devices and/or distributed generation located within the DN. Existing literature pertaining to securing DN can be categorized into three general classes: resilience against communication channel attacks; mitigation of the effects of data tampering; and maintaining system limits. In the first class, emphasis has been placed on detecting cyber intrusion and securing smart grid communication protocols from attacks such as jamming and denial-of-service (DoS) [6]–[8]. In the second category, methods by which the state information generated in the DN can be tampered via false data injection attacks have been explored and recent advances allow these attacks to be detected in real-time [9], [10]. In the final class of proposals, resilience has been incorporated into the core of algorithms that coordinate power injecting and consuming elements residing within the DN to maintain system limits [11], [12]. Attack construction

P. Srikantha, J. Liu and J. Samarabandu are with the Electrical and Computer Engineering department at Western University. Emails: psrikan@uwo.ca, jliu2325@uwo.ca, and jagath@uwo.ca

techniques compromising consumer loads and IoT devices are presented in references [4], [13]. However, to the best of the authors' knowledge, no attack strategies have been proposed in the literature for actuating a large number of small-scale compromised sources in a stealthy manner to cause targeted system-wide disruptions in the DN.

Thus, the main contributions of this paper are four-fold: 1) We identify how Stackelberg games can be utilized by the attacker to target specific DN buses for maximum disruption by evaluating the capabilities of existing defence mechanisms deployed by the EPU; 2) We show how an adversary may use population games to stealthily coordinate a large number of compromised smart loads via periodic broadcasts of cost signals and iterative revisions of individual load actuation; 3) We propose real-time and adaptive defence mechanisms based on data aggregation to mitigate the adverse effects of adversarial attacks; and 4) The effectiveness of the proposed attack-mitigation scheme is demonstrated via practical simulations conducted on IEEE 33-bus, IEEE 69-bus and Brazilian 136-bus DN systems along with theoretical convergence results and comparisons to state-of-the-art coordination strategies. The remainder of this paper is organized as follows. Sec.II introduces the system model utilized in this paper for the construction of the attack-mitigation scheme. Sec. III and IV present the attack and mitigation schemes proposed in this paper. Sec. V presents the results of applying these strategies to practical systems and the paper is concluded in Sec. VI.

II. System Model

In this section, the attack-mitigation model, notations and steady-state DN system operational limits are presented.

A. Attack-Mitigation Model

The DN considered in this paper is composed of power injecting components managed by the EPU and a large number of active loads equipped with communication capabilities (e.g. smart appliances) that are susceptible to cyber-attacks. As such, following are the assumptions made in our attack-mitigation model (also adopted in references [10], [11]):

1) Attacker will be able to observe power injections/absorptions across the DN;
2) Attacker can compromise smart loads via the communication channel;
3) Attacker can securely communicate with all compromised devices;
4) Attacker has no information about the operating state of EPU resources after the attack has commenced;
5) EPU will be able to detect the onset of the attack;
6) EPU will detect corrupt state measurements; and
7) EPU can securely communicate with all of its resources.

Assumptions 1 and 2 are valid as it is possible to exploit existing vulnerabilities in software systems and communication protocols to gain access to cyber-enabled devices such as data concentrators and smart appliances [14]. Moreover, as smart loads are typically connected to the commonly accessible Internet, it will be possible for the adversary to breach into these devices from a remote location [15]. Assumption 3 is feasible as the attackers will have complete control over the attacked loads via the cyber channels. Assumption 4 is realistic as once the onset of an attack has been identified, the EPU will either switch to a higher level of security for system monitoring regardless of the latencies and overheads incurred or completely isolate monitoring equipments from the public domain [2]. Assumption 5 is supported by the widespread integration of the advanced metering infrastructure (AMI) in the DN that produces comprehensive real-time power consumption measurements [2]. Excessive stress and congestions on the DN will be evident from these measurements. Assumption 6 stems from recent advances in the detection of false data injections into state measurements [10]. The final assumption is practical as the EPU will utilize secure channels to communicate with its ancillary devices [14]. Furthermore, Assumptions 4-7 indicate a best case scenario and violating these will only make it easier for the attack to succeed.

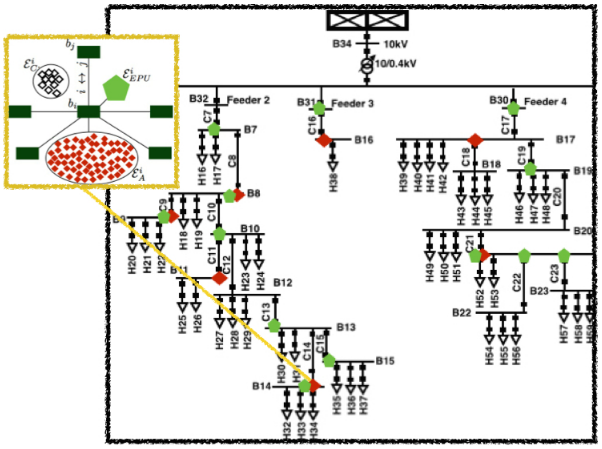


Figure 1: Active DN (Pillai et al [16]) consisting of defence (green) and attacked (red) resources.

B. Notations

The primary function of the DN is to deliver power transported by the transmission network (from large-scale synchronous plants) to end-user appliances. The DN is composed of n buses $\mathcal{B} = \{b_1 \dots b_n\}$ and l lines $\mathcal{L} = \{i \leftrightarrow j\}$ connecting buses $b_i, b_j \in \mathcal{B}$ where b_i is closer to the feeder bus than b_j. Each bus b_i consists of active cyber-physical elements (e.g. storage, smart appliances, electric vehicles and distributed generation) that are directly controlled by the consumers ($\mathcal{E}_C^i$), EPU ($\mathcal{E}_{EPU}^i$) or the adversary ($\mathcal{E}_A^i$). In Fig. 1, an illustration of a low-voltage Danish DN is presented [16]. It is possible for a bus to consist of either, both or none of these components. The adversary will employ resources in $\mathcal{E}_A = \{\mathcal{E}_A^1, \dots, \mathcal{E}_A^n\}$ to violate steady-state limits of DN lines and buses by maximizing the deviations of system states from nominal operating values and the EPU will coordinate elements in $\mathcal{E}_{EPU} = \{\mathcal{E}_{EPU}^1, \dots, \mathcal{E}_{EPU}^n\}$ to minimize these deviations. Thus, the attacker utilizes *evasion* to maximize system disruption while the EPU behaves like a *pursuer* to mitigate the ensuing adverse effects.

C. Steady-state System Limits

Steady-state limits in the active DN are typically associated with 1) power flow in DN lines and 2) bus voltages. The maximum apparent power flow limit imposed on each DN line $i \leftrightarrow j$ is $\bar{s}_{i,j}$ and this depends on the physical attributes of the line such as reactance and conductance. As the DN is a low-voltage system and the DN lines are associated with low reactance to resistance ratio, excessive power flow in the lines caused by significant power consumption by loads will not only violate power flow limits but also lead to steep voltage drops in buses that can result in under-voltage conditions. Thus, voltage V_i on bus b_i is subject to upper and lower limits (i.e. $\underline{V}_i \leq |V_i| \leq \bar{V}_i$). Typically, for safe operation, bus voltages in the DN are required to operate within the $\pm 5\%$ interval of the nominal voltages. In the per unit (p.u.) scale, this translates to $\underline{V}_i = 0.95$ and $\bar{V}_i = 1.05$ where the nominal voltage is $V_i = 1.0$. If any of the afore-mentioned limits are violated, disruptive and costly damages will be incurred by DN equipment and loads.

III. Design of Stealthy Attacks on DN

The goal of the attacker is to efficiently coordinate compromised active elements belonging to the set $\mathcal{E}_A$ to maximize the attack impact on the DN. The incremental nature of the attack will render isolation of compromised nodes difficult for the EPU. The attack is therefore stealthy. To achieve this, the attacker will first assimilate how its limited resources can be utilized effectively to maximize the adverse effects on the system via solution concepts in Stackelberg games (as outlined in Sec. III-A). Through this process, the adversary will identify specific DN buses to target in the attack. After breaching into active cyber-enabled loads residing within these buses via the communication channel, the attacker will design distributed actuation strategies that will be executed by these compromised loads for stealthily saturating DN line limits over-time. The attacker will strive to breach into a large number of small-scale power consuming elements such as smart appliances and electric vehicles residing within the target buses and leverage on the notion of anonymization in population game theory (PGT) as detailed in Sec. III-B to prevent premature isolation of the attack.

A. Attack Planning via Stackelberg Games

The adversary will aim to maximize the disruption in the DN by strategically deploying its resources. First, the state measurement signals generated in the DN will be observed by the attacker over a period of time to glean insights into the characteristics (e.g. availability) of defence mechanisms deployed by the EPU (based on Assumption 1). This information will then be leveraged to systematically target specific buses that have the greatest potential to succumb to adversarial stress. For this, theoretical constructs from Stackelberg games are applied. Although Stackelberg games have been used in the literature in the context of energy management [17] and vulnerability analysis [18], this is the first time these are being effectively utilized for attack planning.

Stackelberg games model natural competition in a system consisting of a dominant *leader* and a *follower* [19]. In this class of games, the leader will always make the first move. The follower will observe the action taken by the leader and then react using best-response strategy whereby it will aim to minimize its losses with respect to the leader's decision. This leader-follower paradigm fits well into the target identification process in the active DN as the EPU is analogous to the leader and the adversary is akin to the follower. The EPU is primarily involved in the design of the active DN and it will deploy active defence and protection mechanisms that are naturally subject to limitations (e.g. economical, capacity constraints, etc.). The adversary silently observes the behaviour of the system via state measurement signals and deduces the attributes of the defence mechanisms deployed by the EPU. These insights are then used to target specific buses in the active DN for maximal disruption. Opposite roles will be studied in future work involving planning studies where the EPU will deploy defence mechanisms based on empirically observed behaviour of adversaries. As such, the optimization problem $\mathcal{P}_T$ is formulated as a typical *follower* problem listed in the work of Yin et al [19] for which parameters are computed based on the attacker's observations.

$$\mathcal{P}_T : \max_{0 \leq a_i \leq 1} \sum_{i=1}^{n} \Big[a_i \big(c_i U_a^c(b_i) + (1 - c_i) U_a^u(b_i) \big) - f_i(a_i) \Big]$$

Optimization variables $a = \{a_1 \ldots a_n\}$ represent the probabilities of attacking buses $\mathcal{B}$. c_i is probability of the EPU activating its resources in $b_i \in \mathcal{B}$. $U_a^c(b_i)$ is the utility for the attacker if the target bus b_i is covered (i.e. defended) by the EPU. $U_a^u(b_i)$ is the utility for the attacker if the target bus is uncovered (i.e. not defended) by the EPU. $f_i(a_i)$ is the cost incurred by the attacker for deploying its resources for attacking b_i. As the utility for the attacker if the target is covered is lower than otherwise, $U_a^u(b_i) - U_a^c(b_i) > 0$. We design these utility functions as follows:

$$U_a^c(b_i) = (\min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i\big] - \tilde{d}_i) * (\tilde{p}_{i,j} - \tilde{p}_i) \tag{1}$$

$$U_a^u(b_i) = (\min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i) - \underline{d}_i\big] * (\tilde{p}_{i,j} - \tilde{p}_i) \tag{2}$$

$$c_i = P^i_{d_i \geq \tilde{d}_i} \tag{3}$$

$p_{i,j}^{lim}$ is the upper limit of real power flow on line $i \leftrightarrow j$, $\tilde{p}_{i,j}$ is the average power flow on the line, $\check{p}_i$ is the adversarial real power injection at bus i that will force the local bus voltage magnitude $|V_i|$ to drop to the minimum allowable bus voltage $\underline{V_i}$ when no other actuation takes place in the DN, $\mathcal{E}_A^i$ is the available real power injection capacity for the attacker at b_i, $\tilde{d}_i$ is the average power injection capacity of EPU defence resources at b_i while $\underline{d}_i$ is the average minimum power injection capacity of the EPU and $\tilde{p}_i$ is the average power injection of b_i. As the adversary intends to instigate widespread outage in the system, the utilities are designed to account for the minimum adversarial power injection that will either saturate the local line $i \leftrightarrow j$, or cause a local bus voltage drop that reduces the voltage limit margin to 0, or completely exhaust locally available adversarial power injection capability. This is captured by the minimum term in both U_a^c and U_a^u. As $\tilde{d}_i > \underline{d}_i$, the relation $U_a^u(b_i) - U_a^c(b_i) > 0$

holds. The probability c_i of the EPU covering b_i is defined to be the proportion of time EPU capacity is greater than the average capacity of these resources. The adversary can utilize local measurements to compute $\tilde{p}_{i,j}$, $\tilde{d}_i$ and $P^i_{d_i \geq \tilde{d}_i}$. The computation of $\check{p}_i$ is more involved and this will be discussed next. We show that $\check{p}_i$ can be computed just based on the impedances of lines in $\mathcal{P}_i$ which is composed of all the lines that form a path from b_i to the feeder node b_0 as follows:

$$\check{p}_i = (|V_i|^2 - \underline{V}_i)/(\sum_{(g,l)\in\mathcal{P}_i} 2r_{g,l}) \tag{4}$$

where $|V_i|^2$ is the magnitude of voltage at b_i prior to the instigation of local attack and $r_{g,l}$ is the resistance of line $g \leftrightarrow l$. To derive this relation, we first note that due to power balance requirements, the voltage drop from the feeder b_0 which has a voltage of 1 p.u. to bus b_i is the following [20]:

$$1 - |V_i|^2 = \sum_{(g,l)\in\mathcal{P}_i} \Big[2r_{g,l}\tilde{p}_{g,l} - (r_{g,l}^2 + x_{g,l}^2)|\tilde{I}_{g,l}|^2 \Big] \tag{5}$$

where $x_{g,l}$ is the reactance of the line $g \leftrightarrow l$ and $\tilde{I}_{g,l}$ is the average current flow on the line prior to adversarial actuation. The second term in the summation which is quadratic represents power losses and this is typically eliminated as the power flow on the line is much larger than these losses. This holds in our system model as the attacker strives to saturate power links. The adversarial power draw $\check{p}_i$ will increase the real power flow from $\tilde{p}_{g,l}$ to $\tilde{p}_{g,l} + \check{p}$. The ensuing change in bus voltage can be computed by subtracting Eq. 5 prior to the adversarial draw with Eq. 5 attained after the adversarial draw to obtain Eq. 4. $\check{p}_i$ can be computed with no knowledge of power flow or bus voltages in other parts of the DN.

Closed-form solution exists for $\mathcal{P}_T$ if the cost function $f_i(a_i)$ is convex and twice continuously differentiable. This is typically the case as greater the probability of attack, the more resources will be required for the attack and thus greater will be the cost incurred by the attacker and this can be modelled as a convex function. Furthermore, $a_i \in [0, 1]$ is a continuous interval. To derive the solution a_i, the first-order condition for optimality is applied where the gradient of the objective function in $\mathcal{P}_T$ is set to 0:

$$a_i = [f_i'^{-1}\Big((c_i U_a^c(b_i) + (1 - c_i)U_a^u(b_i)\Big))]_0^1 \tag{6}$$

where the notation $[.]_0^1$ implies that the minimum and maximum values a_i can take are 0 and 1. This value is normalized to obtain $a_i^{nom} = a_i / \max_{i\in 1\ldots n}(a_i)$. The attack probability vector $a = \{a_1^{nom} \ldots a_n^{nom}\}$ is then used by the adversary to decide on which buses will be targeted for compromise. This way, the attacker will not need to use an exorbitant amount of resources to attack all DN buses but rather focus on the weaknesses in the EPU's defence mechanism to perpetrate a deeply resonating attack.

B. Attack Actuation via Population Games

After identifying the target buses, the adversary will proceed to compromising as many cyber-enabled loads possible (based on Assumption 2) located within each one of these buses in order to obtain direct actuation access. While an unsophisticated attacker may force all of these devices to draw the maximum amount of power possible, ensuing sudden surges in power flow will trip passive protection devices such as fuses/circuit breakers. This will isolate the problematic components from the rest of the DN [5], and thus minimize the effect on the overall system. However, we will show that a smart attacker will be able to deploy their attack that progresses without any impulsive surges and avoid immediate isolation.

Theoretical constructs from population game theory (PGT) are employed to design stealthy actuation of compromised devices within each target bus. PGT models interactions between a large number of anonymous agents equipped with identical discrete strategy set $y = \{y_1 \ldots y_m\}$. The key feature of the PGT framework is that the impact of individual actions executed by the agents is *incremental*. The general state of the system is inferred via aggregate measures $x = \{x_1 \ldots x_m\}$. x_i represents the fraction of agents in the population using strategy $y_i \in y$. Each agent will switch its strategy from y_i to y_j as dictated by the switching probability $\rho_{i,j}(F(x))$ and this is a function of the cost F of the current state x of the system. If F satisfies specific conditions (discussed later), then the convergence of these random cost-based revisions to a unique and optimal solution x^* where $F(x^*) = 0$ is guaranteed.

The PGT paradigm supports many requirements entailed in the construction of stealthy attacks in the DN. As the adversary aims to utilize a large number of small-scale appliances to execute the attack in the target bus b_i, the compromised appliances are analogous to agents in PGT and form the population $\mathcal{P}_i$ where $|\mathcal{P}_i| = q_i$. PGT allows for incremental and anonymous changes in power consumption of these loads and this will render the detection and isolation of the compromised devices difficult. One specific deviation of the DN attack setup from PGT is due to the different power consumption ratings and settings associated with each appliance. Thus, all appliances will not be equipped with the same strategy set. In order to overcome this issue, if an appliance does not support a particular level of power consumption $y_a \in y$, it will select a feasible setting that is the closest to y_a and this is not cause for concern as the impact of a single appliance in the grand scale is very minor due to $\lim_{q_i\to\infty} \frac{\delta_a^j}{q_i {x^i}^T y} = 0$ where δ_a^j is the deviation of agent j from strategy y_a).

In the DN attack setup, each bus b_i will be coordinated by the adversary independently from one another. Hence, large number of compromised loads at each bus will form separate populations. This will allow the attacker to focus on weaknesses present at specific buses while also eliminating interdependencies. Thus, all notations introduced from this point on for the attack construction will be variables and parameters maintained separately at each bus and index i representing bus b_i will not be included unless necessary. The system state x at target bus b_i is a vector representing the proportion of agents in $\mathcal{P}_i$ utilizing specific strategies in y and belongs to the simplex $\triangle = \{x | \sum_{j=1\ldots|y|} x_j = 1, x_j \geq 0 \ \forall \ j = 1 \ldots |y|\}$. $\mathcal{P}_A^i$ is solved by the attacker to compute actuation in bus b_i.

$$\mathcal{P}_A^i : \min_{x\in\triangle} \sum_{k=1}^{|y|} \frac{1}{2} q_i x_k^2 y_k$$
$$\text{s.t.} \sum_{k=1}^{|y|} q_i y_k x_k = \min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i)\big]$$

It is evident from this formulation that the attacker aims to minimize the usage of strategies with higher magnitudes in order to prevent detection by the EPU via the quadratic objective. The coupling constraint dictates the overall power consumption in the target bus b_i to either saturate the local line, result in a voltage drop close to acceptable limits or exhaust available adversarial resources. This will result in DN congestions that can lead to widespread outages as shown in Sec. V. Although solving $\mathcal{P}_A^i$ to obtain the optimal solution x^* is straightforward, it is difficult to assign specific actuation strategies to individual compromised agents in a central manner to achieve this aggregate state as each appliance will be subject to different local constraints and settings. Hence, the cost $F(x) = \{F_1(x) \dots F_{|y|}(x)\}$ is used as a general broadcast signal by the attacker to coordinate the compromised loads in a distributed manner (Assumption 3). In particular, F is:

$$F_k(x_k) = q_i x_k y_k + \nu^* q_i y_k \;\forall\; k \in \{1 \dots |y|\} \tag{7}$$

where ν^* is the optimal Lagrangian multiplier associated with the coupling constraint in $\mathcal{P}_A^i$. F is in fact the gradient of the Lagrangian $\mathcal{L}_A^i(x,\nu)$ constructed for $\mathcal{P}_A^i$:

$$\mathcal{L}_A^i(x,\nu) = \sum_{k=1}^{|y|} \frac{1}{2} q_i x_k^2 y_k + \nu(\sum_{k=1}^{|y|} q_i y_k x_k - \min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i)\big]) \tag{8}$$

The optimal Lagrangian dual variable ν^* results from solving $\max_\nu \mathcal{L}_A^i(x^*,\nu)$ where x^* is obtained by solving $\mathcal{P}_A^i$. $F(x)$ is broadcast to all agents residing within bus b_i periodically every τ seconds. These signals are used by the adversary's agents for local strategy revisions. The simplex condition on x is intrinsically satisfied due to the distributed strategy selection by compromised loads. When the coupling constraint is feasible, these distributed revisions will converge to the optimal solution x^* as $(x^* - x)^T F(x^*) < 0 \;\forall\; x \in \triangle$ holds [21], [22]. x^* is the global solution as $\mathcal{P}_A^i$ is strictly convex.

$F(x)$ is utilized by individual agents to revise local actuation strategies at random time instances based on the switching probability (referred to as *projection* revision [22]):

$$\rho_{j,k}(F^i(x)) = \Big[\frac{F_j(x) - F_k(x)}{q_i x_j}\Big]_+ \tag{9}$$

These revisions will result in the state dynamic: $\dot{x}_i = \sum_{j\in y} x_j \rho_{j,i}(F(x)) - x_i \sum_{j\in y} \rho_{i,j}(F(x))$ composed of the rate at which agents switch into strategy y_i and the rate at which agents switch out of strategy y_i. Stochastic effects are averaged out due to the strong law of large numbers. As such, projection revision specifically results in the state dynamic: $\dot{x}_i = \frac{1}{n}\sum_{y_j\in y} F_j(x) - F_i(x)$. This is in fact negative of the projection of the cost function $F(x)$ onto the simplex $\triangle$. As $F(x)$ is the gradient of $\mathcal{L}_A^i(x,\nu^*)$, this implies that the system will gravitate towards lower potential at each revision. Moreover, according to the Lyapunov theory, this dynamic results in exponentially fast convergence to the optimal solution x^* as the Lyapunov function is in fact $\mathcal{L}_A^i(x,\nu^*)$ [22].

At optimality, the attacker would have successfully saturated all links connected to bus b_i and this will lead to excessive voltage drops and equipment failure. These revisions will be carried out in each one of the DN buses selected as targets by the adversary which will result in the widespread reduction of the DN system topology as illustrated in Sec. V. Algorithm 1 presents a summary of our proposal for stealthy attack construction in the DN.

Algorithm 1 Attack Construction

```
1) Target Identification
  for $b_i \in \mathcal{B}$ do                                ▷ Compute average parameters
      $U_a^c(b_i) \leftarrow (\min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i)\big] - \tilde{d}_i) * (p_{j,i}^{lim} - \bar{p}_{j,i})$
      $U_a^u(b_i) \leftarrow (\min\big[p_{i,j}^{lim} - \tilde{p}_{i,j}, \check{p}_i, \mathcal{E}_A^i)\big] - \underline{d}_i) * (p_{j,i}^{lim} - \bar{p}_{j,i})$
      $c_i \leftarrow P^i_{d_i \geq \bar{d}_i}$
  end for
  $a \leftarrow$Solve $\mathcal{P}_T$, $a^{nom} \leftarrow a / \max_{i\in 1\dots n} a$
  for $a_i \in a$ do                                          ▷ Identify target buses for attack
      if rand(1) $< a_i^{nom}$ then
          $\mathcal{A} \leftarrow b_i$
      end if
  end for
2) Compromised Load Actuation Construction at $b_i \in \mathcal{A}$
  Central Aggregation:
  for every $\tau$ seconds do
      $F_j(x_j) \leftarrow q_i x_j y_j + \nu^* q_i y_j \;\forall\; j \in \{1 \dots |y|\}$
      Broadcast $F(x)$ to all adversary agents in $b_i$
  end for
  Individual Agents:
  Initialize strategies to $y_1$, $t_{rev} \leftarrow rand(T)$
  if $t > t_{rev}$ and $rand(1) < \rho_{i,j}(F(x))$ then
      Switch to feasible strategy closest to $y_j$
      $t_{rev} \leftarrow t_{rev} + rand(T)$
  end if
```

IV. EPU Countermeasure via Dual Updates

Upon detecting anomalous power consumption in the DN from aggregate measurements (Assumption 5), the EPU will react by employing all available resources in $\mathcal{E}_{EPU}$ to minimize system state deviations from nominal values. Defence mechanisms such as storage devices and distributed generation are becoming financially viable due to incentive mechanisms from governments and advances in technologies [5]. The EPU will utilize every power injecting element $\mathcal{E}_{EPU}^i$ located within each bus $b_i \in \mathcal{B}$ to offset abnormal power consumption across the entire DN as it is not possible to single out compromised devices effectively due to the stealthy nature of the proposed attack strategy. An iterative approach based on monotone operators will be employed by the EPU for the systematic and cost-effective actuation of these devices. PGT is not applicable here as the EPU resources are limited and are highly varying.

A. Countermeasure Formulation

The EPU will attempt to restore balance in net power demands of the DN by optimally activating all of its secure power injecting components using current state measurement signals (Assumptions 6 and 7). As it is possible for the adversary to cumulatively draw large amounts of power that is much greater than the power injecting capacity locally available within that bus, the EPU will leverage on the surplus capacity of resources located in other buses to offset this imbalance. Thus, the EPU countermeasure problem $\mathcal{P}_{EPU}$ is formulated for the entire DN (unlike the adversary which targets specific buses).

$$\mathcal{P}_{EPU}: \min_{p_g} \sum_{i=1}^{|\mathcal{B}|} C_i(p_g^i)$$
$$\sum_{i=1}^{|\mathcal{B}|} p_g^i = \sum_{i=1}^{|\mathcal{B}|} p_a^i,\ \underline{p}_g^i \leq p_g^i \leq \bar{p}_g^i\ \forall\ b_i \in \mathcal{B}$$

p_g^i is the power injected by the EPU resource $\mathcal{E}_{EPU}^i$ located in b_i and p_a^i is the overall adversarial power draw in b_i (EPU can detect this due to Assumptions 5 and 6 and past trends in power demands). $C_i(p_g^i)$ is the cost of dispatching these resources and is set to be $(p_g^i)^2/(1-|V_i|^2)$ which is strictly convex. This cost allows for greater power injection in buses that are attacked (i.e. bus voltage magnitudes are closer to lower limits). The coupling constraint dictates that the overall power injected by the EPU must match the overall adversarial power consumption in the system. The inequalities impose upper $\bar{p}_g^i$ and lower $\underline{p}_g^i$ bounds associated with the generation capacity of $\mathcal{E}_{EPU}^i$. The EPU will not be directly solving $\mathcal{P}_{EPU}$ as it will not be aware of local conditions that affect the cost functions and generation limits of each bus. This process is distributed whereby every EPU resource will respond to general signals broadcast by the EPU based on local constraints and costs.

B. Distributed EPU Coordination Strategy

EPU signals and resource actuation are computed using the Karush Kuhn Tucker (KKT) conditions which are necessary and sufficient for optimality in convex optimization problems [21]. The first condition is based on *primal feasibility* which indicates that the optimal solution ${p_g^i}^*$ must satisfy the constraints listed in the primal problem $\mathcal{P}_{EPU}$. The second condition is based on *dual feasibility* ${\lambda_1^i}^* \geq 0$, ${\lambda_2^i}^* \geq 0$ and $\nu^* \in \mathbb{R}\ \forall\ b_i \in \mathcal{B}$ which is associated with the dual problem of $\mathcal{P}_{EPU}$. The optimal dual variables λ_1^* and λ_2^* are associated with the inequality constraints pertaining to the generation capacities and ν^* is associated with the coupling constraints of $\mathcal{P}_{EPU}$. The third condition is derived from the *complementary slackness* property of dual variables and primal constraints: ${\lambda_1^i}^*(\underline{p}_g^i - {p_g^i}^*) = 0$ and ${\lambda_2^i}^*({p_g^i}^* - \bar{p}_g^i) = 0\ \forall\ b_i \in \mathcal{B}$. These indicate that if the optimal solution does not activate the inequality constraints (i.e. there is some slack in the constraint and solution), then the λ terms are 0. Otherwise, when $\underline{p}_g^i - {p_g^i}^* = 0$ or ${p_g^i}^* - \bar{p}_g^i = 0$, these dual variables will not necessarily be 0. The final condition is based on *stationarity*: $\frac{\partial L}{\partial x_i} = C_i'({p_g^i}^*) + \nu^* - {\lambda_1^i}^* + {\lambda_2^i}^* = 0$. This is attained from the first-order optimality condition which dictates that the gradient of the Lagrangian dual function with respect to the primal variable is 0 for x^*, λ_1^*, λ_2^* and ν^*.

These KKT conditions are used to eliminate λ_i variables and derive the following dependencies between ν and p_g^i.

$$\text{If } \nu^* \geq -C_i'(\underline{p}_g^i) \text{ then } {p_g^i}^* = \underline{p}_g^i \tag{10}$$
$$\text{If } -C_i'(\bar{p}_g^i) \leq \nu^* \leq -C_i'(\underline{p}_g^i) \text{ then } {p_g^i}^* = {C_i'}^{-1}(-\nu^*) \tag{11}$$
$$\text{If } \nu^* \leq -C_i'(\bar{p}_g^i) \text{ then } {p_g^i}^* = \bar{p}_g^i \tag{12}$$

Based on this, power injection by each source is:

$$p_g^i(\nu) = [{C_i'}^{-1}(-\nu)]_{\underline{p}_g^i}^{\bar{p}_g^i} \tag{13}$$

The overall power injection in the system is a summation of individual actuation of all devices in $\mathcal{E}_{EPU}$ and this must be equal to the overall adversarial power consumption $\sum_{i=1}^{|\mathcal{B}|} p_a^i$. The difference between the overall EPU power injection and adversarial consumption is $d(\nu)$:

$$d(\nu) = \sum_{i=1}^{|\mathcal{E}_{EPU}|} [{C_i'}^{-1}(-\nu)]_{\underline{p}_g^i}^{\bar{p}_g^i} - \sum_{i=1}^{|\mathcal{B}|} p_a^i \tag{14}$$

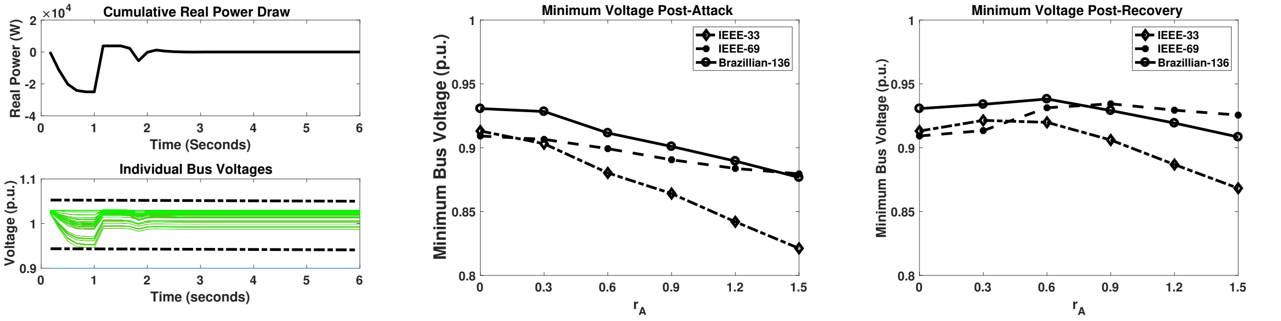
As $d(\nu)$ is a monotonically non-decreasing function with respect to $-\nu$, the solution ν^* where $d(\nu^*) = 0$ is a unique solution [21]. The EPU will identify ν^* via binary search. For this, an initial guess of ν is periodically broadcast to all $\mathcal{E}_{EPU}$ every τ seconds. This signal is used for local actuation by the EPU resources. This is repeated until $d(\nu) \in \pm\epsilon$ as detailed in Algorithm 2. The interval $[\nu_{min}, \nu_{max}]$ for ν is initialized using the power injection capacity ratings of $\mathcal{E}_{EPU}$ which are available to the EPU in advance and worst possible costs incurred in bus b_i based on current DN conditions. The convergence rate of this algorithm is $\mathcal{O}(\log(\frac{\nu_{max}-\nu_{min}}{\epsilon}))$.

Algorithm 2 Attack Construction

```
1) Signal Computation by EPU:
$\nu_{min} \leftarrow \max(-C_i'(\underline{p}_g^i))$, $\nu_{max} \leftarrow \min(-C_i'(\bar{p}_g^i))$
$\nu \leftarrow \nu_{max}$
for every $\tau$ seconds when $|d(\nu)| > \epsilon$ do
    Broadcast $\nu$
    if $d(\nu) > 0$ then
        $\nu_{min} \leftarrow (\nu_{max} + \nu_{min})/2$,
    else
        $\nu_{max} \leftarrow (\nu_{max} + \nu_{min})/2$
    end if
    $\nu \leftarrow (\nu_{max} + \nu_{min})/2$
end for
2) Distributed Actuation by EPU Resources:
for every $\tau$ received do
    $p_g^i \leftarrow [{C_i'}^{-1}(\eta_1)]_{\underline{p}_g^i}^{\bar{p}_g^i}$
end for
```

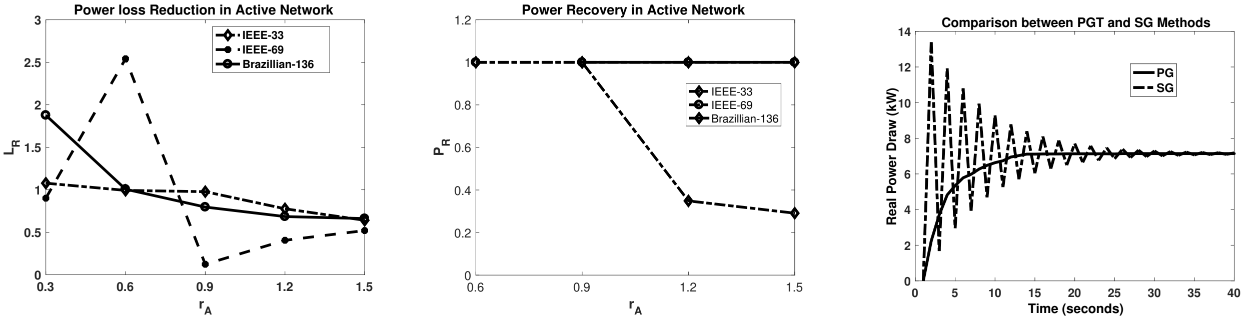
V. Results

In this section, we verify the performance of the proposed attack and mitigation strategies via simulations conducted using realistic DNs such as IEEE 33-bus, IEEE 69-bus and Brazilian 136-bus systems. We explore the impact on bus voltages, power losses, and DN loads when the proposed attack-mitigation strategies are applied. These present interesting insights on how cyber attacks perpetrated on consumer appliances can induce deeply resonating system failures and how these adverse consequences can be averted with active self-healing defence mechanisms. Then, we present a comparative study.



(a) State evolution in 33-bus DN. (b) Minimum bus voltage post-attack. (c) Minimum bus voltage post-recovery.

Figure 2: Impact of application of attack and countermeasures.



(a) Branch power losses. (b) Load recovery. (c) Comparison of PGT vs SG

Figure 3: Impact of EPU Resources and Comparative Study.

A. Evolution of System States with DN Attack-Mitigation

In Fig. 2a, we present the evolution of system states when an adversary compromises buses in a 33-bus DN based on the computed attack probabilities and coordinates the corresponding consumer loads via the proposed population game theoretic method. The attack takes place over one second and thereafter the EPU dispatches its resources to protect the system. In the first sub-plot, the cumulative real power consumed by compromised loads evolves exponentially fast as expected for the state dynamic induced by the projected strategy revision technique. The impact of this abnormal distributed power draw on bus voltage magnitude is illustrated in the second subplot. The bus voltage magnitudes of all buses in the system are plotted here. It is clear that the onset of the attack has resulted in the dip of bus voltage magnitudes perilously close to the lower limits. If the attack had not been intercepted by the EPU, this continued trend will have resulted in the violation of bus voltage limits and the onset of outages as illustrated in the next section. At the 1 second mark, the EPU deploys all available defence mechanisms to avert threshold violations via the proposed countermeasure involving dual updates. It is clear that the EPU is able to rapidly compensate adversarial power draw and increase the safety margins of bus voltage magnitudes.

B. Minimum Bus Voltage Magnitude

Next, we analyze the impact of adversarial resources on minimum bus voltage magnitudes in 33-bus, 69-bus and 136-bus DNs. First, we examine the case where only the attack takes place in Fig. 2b, then in Fig. 2c, we analyze the impact on the system in the post recovery period when EPU's defence mechanisms are deployed. The x-axis represents the proportion r_A of the adversarial resources with respect to the EPU resources. As expected post-attack, when the adversarial capacity increases, the impact on the bus voltage magnitudes is pronounced and exhibits a downwards trend in Fig. 2b. The 33-bus system is most affected due to the unbalanced topology of the DN buses [26]. When the EPU resources are deployed, in post-recovery it is clear that the voltage profile has improved significantly. However, when the adversarial resources have more than half the capacity of the EPU, the minimum bus voltage magnitudes in 33-bus, 69-bus and 136-bus systems reflect a downward trend. This is expected as the limited capacity of the EPU will not be able to compensate for the significant power draw enacted by the adversary across the DN system. Thus, it is imperative to deploy adequate defence mechanisms to offset adversarial attacks like these.

C. Branch Power Losses

Excessive adversarial power draw can result in power losses in the DN lines. In Fig. 3a, the average power loss reduction ratio (post recovery versus post attack) L_R in DN lines achieved post recovery for various values of r_A is examined. When the adversarial resources are scarce (i.e. r_A is close to 0), the power loss improvement resulting from the EPU resources allows for better performance than the original state of the DN (i.e. when $L_R > 1$). This is due to EPU resources having greater capacity than compromised loads at various buses. Hence, greater reduction of power losses in lines/branches can be attributed to reverse power flow. Thus, active defence mechanisms not only avert adversarial attacks but can also be utilized for improving the voltage profile of the DN.

D. Load Recovery

The main goal of the adversarial attack is to trigger protection mechanisms that will result in widespread power outages in the DN. When system limits (e.g. bus voltage magnitudes and apparent power flow limits) are violated, outages will ensue. In Fig. 3b, the proportion of loads that have been recovered post deployment of EPU resources is studied. When adversarial resources are lower than EPU resources, load recovery is 100% (i.e. r_A is low). However, when the adversarial resources exceed the EPU resources, some loads are severed in the 33-bus system while this is not observed in the 69-bus and 136-bus systems. This can be attributed to the deeply linear arrangement of buses in the 33-bus system. When the adversary is successful imposing undue stress on specific buses, then all descendent buses will be isolated or affected. This will result in the severing of a large number of loads from the network. This is not the case with the 69-bus and 136-bus system which has more balanced tree topologies. Thus, this key insight also highlights the importance of designing a DN with a balanced network structure.

E. Comparison

Both the attack and mitigation strategies involve the coordination of either adversarial or defence resources. With defence resources, convergence is highly rapid as illustrated in Fig. 2a. However, adversarial coordination involves a large number of consumer-centric loads with incremental impact on the system and this is not the case with the EPU resources. Thus, the performance of the coordination mechanism with respect to the number of active elements is an important consideration. For an adversary to *stealthily* exact an attack, the cumulative power draw should not result in irregular oscillations which will not only trip protection devices such as a circuit breakers but also alert the EPU to abnormal activities in the system. There exist a large range of coordination algorithms in the literature such as consensus algorithms [23], [24] based on decentralized mechanisms which are associated with linear convergence rates that are not tractable in large systems. Other algorithms such as that based on sub-gradient (SG) [25] method is utilized to compute coordinating signals which are then utilized by individual agents to refine local strategies. We show in Fig. 3c that the SG method results in significant ringing in comparison to the proposed PG strategy. These oscillations can prematurely trigger protection devices and this is not the goal of the adversary. The proposed PG strategy allows for the rapid coordination of a large number of loads with no oscillations or other disruptive artifacts.

VI. Conclusions

Existing vulnerabilities in the cyber-domain can be leveraged by an adversary to perpetrate sophisticated attacks on the DN system. An adversary can effectively coordinate a large number of small-scale consumer-centric devices in a stealthy and seamless manner. We have also shown that if the onset of an attack can be detected in a timely manner, the EPU can deploy its resources in a distributed manner to offset the adverse effects of the attack on the system. As future work, we intend to study how the EPU can detect stealthy cyber-physical attacks in a timely manner. Moreover, we will also study how the DN can be designed to be robust to stealthy attacks and also how defence mechanisms can be planned to maximally avert failures while incurring minimal cost.

References

[1] J. Markoff, “Hackers Have New Entries With the Internet of Things,” *The New York Times*, Nov. 3, 2016.

[2] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati and G. P. Hancke, “Smart Grid Technologies: Comm. Tech. and Standards,” *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529-539, 2011.

[3] E. Ronen, A. Shamir, A.O. Weingarten, and C. Oflynn, “IoT Goes Nuclear: Creating a ZigBee Chain Reaction,” *IEEE Symposium on Security and Privacy*, 2017.

[4] Y. Dvorkin and S. Garg, “IoT-enabled distributed cyber-attacks on transmission and distribution grids,” *North Am. Pow. Sym.*, 2017.

[5] J. D. Glover, T. J. Overbye, and M. S. Sarma, *Power system analysis & design*. Australia: Cengage Learning, 2017.

[6] U. K. Premaratne, J. Samarabandu, T. S. Sidhu, R. Beresh, and J. Tan, “An Intrusion Detection System for IEC61850 Automated Substations”, *IEEE Trans. on Power Delivery*, vol. 25, no. 4, pp. 2376-2383, 2010.

[7] K. Gai, M. Qiu, Z. Ming, H. Zhao, and L. Qiu, “Spoofing Jamming Attack Strategy Using Optimal Power Distributions in Wireless Smart Grid Networks,” *IEEE TSG*, vol. 8, no. 5, pp. 2431-2439, 2017

[8] P. Srikantha and D. Kundur, “Denial of Service Attacks and Mitigation for Stability in Cyber-enabled Power Grid,” *IEEE PES & Innovative Smart Grid Technologies Conf.*, 2015.

[9] Y. Liu, P. Ning and M. K. Reiter, “False data injection attacks against state estimation in electric power grids”, *ACM Conf. on Comp. and Comm. Sec.*, pp. 21-32, 2009.

[10] Y. He, G. J. Mendis and J. Wei, “Real-Time Detection of False Data Injection Attacks in Smart Grid: A Deep Learning-Based Intelligent Mechanism,” *IEEE Trans. on Sm. Grid*, vol. 8, no. 5, pp. 21-32, 2017.

[11] Y. Liu, H. Xin, Z. Qu, and D. Gan, “An Attack-Resilient Cooperative Control Strategy of Multiple Distributed Generators in Distribution Networks,” *IEEE Trans. Sm. Grid*, vol. 7, no. 6, pp. 2923-2932, 2016.

[12] P. Srikantha and D. Kundur, “Resilient Distributed Real-Time Demand Response via Population Games,” *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2532-2543, 2017.

[13] A. H. Mohsenian-Rad and A. Leon-Garcia, “Distributed Internet-Based Load Altering Attacks Against Smart Power Grids,” *IEEE Trans. on Smart Grid*, vol. 2, no. 4, pp. 667-674, Dec. 2011.

[14] R. Rodrigo, et al. “Securing the Internet of Things.” *Computer*, vol. 44, no. 9, pp. 51-58, 2011.

[15] A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi,“Internet of Things for Smart Cities,” *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32, 2014.

[16] J. Pillai, P. Thogersen, J. Moller, and B. Bak-Jensen. “Integration of Electric Vehicles in Low Voltage Danish Distribution Grids,” *IEEE PES GM*, pp. 1-6, 2012.

[17] J. Chen and Q. Zhu, “A Stackelberg Game Approach for Two-Level Distributed Energy Management in Smart Grids,” IEEE Trans. on Smart Grid, pp. 1-1, 2017.

[18] D. Shelar and S. Amin, “Analyzing Vulnerability of Electricity Distribution Networks to DER Disruptions,” ACC, 2015.

[19] Z. Yin, D. Korzhyk, C. Kiekintveld, V. Conitzer, and M. Tambe, “Stackelberg versus Nash in Security Games: Interchangeability, Equivalence, and Uniqueness,” *Journal of Artificial Intelligence Research*, vol. 41, pp. 297-327, Jun. 2011.

[20] S. Low, “Convex relaxation of optimal power flow part I: Formulations and equivalence”, *IEEE Trans. on Cont. of Net. Sys.*, vol. 1, no. 1, pp.15-27, 2014.

[21] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge Univ. Pr., 2011.

[22] W. H. Sandholm, *Economic Learning and Social Evolution: Population Games and Evolutionary Dynamics*. MIT Press, 2011.

[23] J. Rivera, P. Wolfrum, S. Hirche, C. Goebel, and H.-A. Jacobsen, "Alternating Direction Method of Multipliers for decentralized electric vehicle charging control," *IEEE Conf. on Dec. and Cont.*, 2013.

[24] C. Chen, J. Wang, and S. Kishore, "A Distributed Direct Load Control Approach for Large-Scale Residential Demand Response," *IEEE Trans. on Pow. Sys.*, vol. 29, no. 5, pp. 2219-2228, 2014.

[25] J. Joo and M.D. Ilic., "Multi-layered Optimization of Demand Resources using Lagrange Dual Decomposition". *IEEE Trans. on Smart Grid*, vol. 4, no. 4, pp. 2081-2088, 2013.

[26] M. E. Baran and F. F. Wu, "Optimal capacitor placement of radial distribution systems", *IEEE Trans. on Pow. Delivery*, vol. 4, no. 2, pp. 725-734, 1989.

Pirathayini Srikantha is currently an Assistant Professor in the Department of Electrical and Computer Engineering at Western University. She received her B.A.Sc. degree in Systems Design Engineering from the University of Waterloo in 2009 and her M.A.Sc. degree in Electrical and Computer Engineering from the same institute in 2013. She obtained her Ph.D. degree from The Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto in 2017. She is a certified Professional Engineer (P.Eng.) in Ontario. Her main research interests are in the areas of large-scale optimization and distributed control for enabling adaptive, sustainable and resilient power grid operations. Her work has been published in premier smart grid journal and conference venues. Her research efforts have received recognitions that include the best paper award (IEEE Smart Grid Communications) and runner-up best poster award (ACM Women in Computing). She is also actively involved in professional and social activities. She has served as the Workshop Chair, Session Chair and Technical Program Committee member in IEEE conferences. She is a reviewer in numerous IEEE transactions journals.

Jingyuan Liu received his B. Sc. degree in aerospace engineering from the University of Illinois-Urbana Champaign, Champaign, IL, USA in 2015, and MSc. degree in aerospace engineering from the Purdue University, West Lafayette, Indiana, USA, in 2017. He is currently pursuing the Ph.D. degree from the Department of Electrical and Computer Engineering, Western University, London, ON, Canada. His current research interests include optimization and control of power systems and smart grid security, with a particular focus on the application of game theoretic and convex optimization techniques to distribution network reconfiguration and optimal power flow problems.

Jagath Samarabandu received the BSc (Eng) degree in electronics and telecommunication with first class honours from the University of Moratuwa, Sri Lanka, in 1982, and the MS and PhD degrees in electrical engineering from the State University of New York at Buffalo in 1990 and 1994, respectively. He was awarded the Fulbright Scholarship in 1987. He has been with the Department of Electrical and Computer Engineering at the University of Western Ontario since 2000. His research interests include intelligent systems, network security and image recognition.