2015 American Control Conference
Palmer House Hilton
July 1-3, 2015. Chicago, IL, USA

Analyzing Vulnerability of Electricity Distribution Networks to DER Disruptions

Devendra Shelar and Saurabh Amin

***Abstract*— We formulate a sequential (Stackelberg) game for assessing the vulnerability of radial electricity distribution networks to disruptions in Distributed Energy Resources (DERs). In this model, the attacker disrupts a subset of DER nodes by remotely manipulating the set-points of their inverters. The defender (network operator) responds by controlling the non-compromised DERs and by imposing partial load reduction via direct load control. The attacker's (resp. defender's) objective is to maximize (resp. minimize) the weighted sum of cost due to the loss of voltage regulation and the cost of load control. For the sequential play game where the attacker (resp. defender) is the leader (resp. follower) and under linear power flow equations, we show that the problem reduces to standard bilevel network interdiction problem. Under our assumptions on the attack model, we obtain a structural insight that the attacker's optimal strategy is to compromise the downstream DER nodes as opposed to the upstream ones. We present a small case study to demonstrate the applicability of our model for vulnerability assessment of distribution networks.**

I. Introduction

Over the next two decades, the growth in the percentage of data owing through the electricity grid's communication networks is expected to exceed the growth of electricity flowing through the grid's electric networks [13]. New functionalities such as integration of Distributed Energy Resources (DERs), wide-area monitoring and control, and demand-response mechanisms constitute major thrusts of smart grid initiatives. In particular, the rate of deployment of DERs such as photovoltaic (PV) systems is expected to grow steadily at the distribution side of the electric grid [1], [17]. Large-scale deployment of DERs provide new opportunities to improve the reliability and efficiency of the distribution networks (DNs). DERs can provide power during critical periods, and reduce load on bulk generators during peak loading conditions. The deployment of DERs may result in lower line losses on average, because they are typically located closer to the loads relative to the bulk sources. The inverter-enabled PVs can also provide reactive power compensation that can be used for voltage regulation in contingency situations [7], [17].

However, DERs may also introduce certain undesirable effects, for example, faults due to reverse power flows, safety issues for maintenance crews, and loss of voltage regulation due to their intermittent nature. Importantly, some of these effects can pose significant risks to DNs that are prone to adversarial manipulation. Specifically, the ongoing DER developments require interconnections that are facilitated by information and communications technologies (ICT) [13]. The use of commercial off-the-shelf ICT introduces new security vulnerabilities. These vulnerabilities can be exploited by malicious hackers who want to cause damage to the DN or even outages across multiple DNs.

In this article, we focus on assessing the vulnerability of a DN with DERs (e.g., PV systems) connected to its nodes. We are specifically motivated by a two-stage failure scenario that was recently ranked as critical by the experts in the area of cyber-physical security of DNs [14]. First, the DER systems can be shutdown if a threat agent (attacker) manages to send spoofed supervisory control and data acquisition (SCADA) control commands. Second, the DER SCADA may unintentionally issue control commands that are invalid for the current operating conditions, resulting in disconnection of significant number of DER nodes. Fig. 1 shows a two-step attack that motivates our problem formulation.[1]

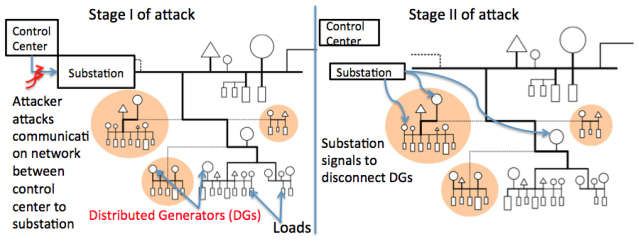


Fig. 1. Illustration of a two step attack motivated by scenarios in [14]. In the first stage, the threat agent (attacker) compromises the communication between the control center (or DER SCADA) and the substation, and introduces incorrect set-points. In the second stage, the substation selectively disconnects a number of DER nodes.

We assume that the attacker's objective is:

(a) To impact the distribution operations by causing loss of voltage regulation (LOVR), and
(b) To induce the defender to impose temporary load control in response to the loss of voltage regulation.

We assume that the attacker is resource constrained, i.e., the attacker can only target a fraction of DER nodes. For this

D. Shelar and S. Amin are with department of Civil and Environmental Engineering (CEE), Massachusetts Institute of Technology, MA, USA. Their work was supported by EPRI grant for "Modeling the Impact of ICT Failures on the Resilience of Electric Distribution Systems" contract ID: 10000621, and the NSF project "CPS: Frontiers: Collaborative Research: Foundations Of Resilient CybEr-physical Systems (FORCES)" (award number: CNS-1239054). Emails: {shelard,amins}@mit.edu

[1]The DN illustration is due to Dr. Alexandra von Meier (UC Berkeley).

attack model, we are interested in determining which DERs are critical under worst-case attacks, and how should the network operator (defender) optimally respond after he observes that a set of DERs are disconnected. Notably, previous work in vulnerability assessment of electricity grids has mainly focused on transmission networks, including the study of cascading failures (see for example, [4],[5], and [15]). The typical model of vulnerability is disruptions in transmission lines and/or generators. However, these approaches do not extend to DNs, especially when the DER nodes are prone to adversarial manipulation. In addition, the specific characteristics of DNs like high resistance-to-reactance (R/X) ratios require consideration of both active and reactive power in the power flow model used for vulnerability assessment.

Our contributions are as follows:

i) In Section II, we model attacker-defender interactions on the radial DNs as a two-stage game where in the first stage the attacker (leader) compromises the control center-substation communications and manipulates the set-points of a subset of DER nodes. In the second stage, the defender (follower) responds by controlling the non-compromised DERs and by imposing partial load shedding via direct load control. Our formulation captures the attacker's (resp. defender's) objective to maximize (resp. minimize) the weighted sum of LOVR and cost due to load control.

ii) In Section III, we show that under the assumption of linear power flows, the aforementioned problem can be solved using standard computational techniques developed for bilinear network interdiction problems (BLNIP). In Section IV, we obtain a structural insight on how the attacker should expend his resources in compromising DER nodes (downstream versus upstream) for the case when the defender does not respond to the attacker's actions. Finally, in Section V, we present a small case study to demonstrate the validity of this insight for the case when the defender optimally responds to the attacker's actions.

II. Problem Formulation

A. Distribution Network Model

We summarize the network model of radial electric distribution systems. We follow the standard modeling approach in the literature [2], [3] and [17]. Consider a tree network of nodes and distribution lines $(\mathcal{N}, \mathcal{E})$. Let $\mathcal{N}_0 \coloneqq \mathcal{N}\backslash\{0\}$ denote the set of all nodes except the substation (labeled as node 0). Let $V_i \in \mathbb{C}$ denote the complex voltage at node i, and $\nu_i = |V_i|^2$ denote the square of the voltage magnitude. Assume that the magnitude of the substation voltage $|V_0|$ is constant. $|V_0|$ is also the nominal voltage at every node. Let $I_{ij} \in \mathbb{C}$ denote the current flowing from node i to node j on line $(i,j) \in \mathcal{E}$. The square of the magnitude of the current is denoted by $\ell_{ij} = |I_{ij}|^2$. A distribution line $(i,j) \in \mathcal{E}$ (also denoted by $i \to j$) has a complex impedance $z_{ij} = r_{ij} + \mathbf{j}x_{ij}$, where r_{ij} and x_{ij} denote the resistance and the reactance of the line (i,j), respectively. Due to the operational requirements of the DN, the voltage magnitudes are constrained as :

$$\underline{\nu}_i \le \nu_i \le \overline{\nu}_i, \tag{1}$$

where $\underline{\nu}_i = |\underline{V}_i|^2$ and $\overline{\nu}_i = |\overline{V}_i|^2$, where $\underline{V}_i$ and $\overline{V}_i$ are the lower and upper bounds for the square of nodal voltage magnitude at node i, respectively.

1) Load Model: Loads can be modeled as constant impedance (**Z**), constant current (**I**) and constant power (**P**) models, or a combination of the three models [8]. Here we only consider constant power loads. Let $sc_i \coloneqq pc_i + \mathbf{j}qc_i$ denote the power consumed by a load at node i, where pc_i and qc_i are the real and reactive components of the power consumed. Let $sc_i^{\mathrm{d}} \coloneqq pc_i^{\mathrm{d}} + \mathbf{j}qc_i^{\mathrm{d}}$ denote the power demanded by a node i, where pc_i^{d} and qc_i^{d} are the real and reactive components of sc_i^{d}. Under our assumptions, $pc_i \le pc_i^{\mathrm{d}}$ and $qc_i \le qc_i^{\mathrm{d}}$. If node i has no load connected to it, we assume that no power is consumed at that node, i.e, $pc_i = qc_i = 0$.

2) PV Model : [2] Let $sg_i \coloneqq pg_i + \mathbf{j}qg_i$ denote the power generated by PV connected to node i, where pg_i and qg_i. Here qg_i models the supply of reactive power by the inverter. This reactive power is bounded by the apparent power capability of the inverter, denoted as $\overline{sg}_i$:

$$-\sqrt{\overline{sg}_i^2 - (pg_i)^2} \le qg_i \le \sqrt{\overline{sg}_i^2 - (pg_i)^2} \tag{2}$$

See [7], [17] for more information of (2). Again, if a node i has no PV connected to it, we assume that no power is generated at that node, i.e, $pg_i = qg_i = 0$. Finally, the setpoints of real and reactive power outputs are denoted by $\widetilde{pg}_i$ and $\widetilde{qg}_i$, respectively. We assume that the control center (defender) can configure the setpoints such that the following constraints hold.

$$0 \le \widetilde{pg}_i \le \overline{sg}_i \tag{3a}$$

$$-\sqrt{\overline{sg}_i^2 - (\widetilde{pg}_i)^2} \le \widetilde{qg}_i \le \sqrt{\overline{sg}_i^2 - (\widetilde{pg}_i)^2} \tag{3b}$$

We will clarify the relationship between the power outputs and the setpoints later.

3) Power Flow Equations: Following [2], the 3-phase balanced power flow equations can be stated as follows:

$$P_{ij} = \sum_{k:j\to k} P_{jk} + r_{ij}\ell_{ij} + pc_j - pg_j \tag{4a}$$

$$Q_{ij} = \sum_{k:j\to k} Q_{jk} + x_{ij}\ell_{ij} + qc_j - qg_j \tag{4b}$$

$$\nu_j = \nu_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) + (r_{ij}^2 + x_{ij}^2)\ell_{ij} \tag{4c}$$

$$\ell_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{\nu_i} \tag{4d}$$

where P_{ij}, Q_{ij} denote the real and reactive power flowing from node i to node j on line $(i,j) \in \mathcal{E}$, respectively, and $S_{ij} = P_{ij} + \mathbf{j}Q_{ij}$ denotes the complex power flowing on

[2]Throughout the paper, we will use the term PV to denote the *complete PV-inverter assembly* attached to a node of DN.

line $(i,j) \in \mathcal{E}$. Equations (4a) and (4b) are the conservation equations for real and reactive power flows. Equation (4c) gives a relationship between the voltage loss and the power flows, and equation (4d) is simply the current-voltage-power relationship.

The following linear power flow (LPF) equations approximate the aforementioned nonlinear equations:

$$P_{ij} = \sum_{k:j\to k} P_{jk} + pc_j - pg_j \tag{5a}$$

$$Q_{ij} = \sum_{k:j\to k} Q_{jk} + qc_j - qg_j \tag{5b}$$

$$\nu_j = \nu_i - 2(r_{ij}P_{ij} + x_{ij}Q_{ij}) \tag{5c}$$

$$\ell_{ij} = \frac{P_{ij}^2 + Q_{ij}^2}{\nu_i}. \tag{5d}$$

Note that, although equation (5d) is a non-linear equation, it can be invoked after all the voltages and power flows are calculated using linear equations (5a)–(5c).

B. Two-Stage (Stackelberg) Game

We now describe a two-stage sequential game where the attacker's (leader's) objective is to compromise PVs and induce LOVR and/or load control. The defender observes the attacker's actions (i.e. he has perfect information about the leader's play). The defender responds by manipulating the settings of one or more loads and by operating the non-compromised PVs, given the constraints imposed by the setpoints. An equivalent representation of this game is the network interdiction model. Classically, such problems are viewed as bilevel optimization problems with binary valued decision variables for the outer (attacker) problem and continuous valued decision variables for the inner (defender) problem.

1) Attacker Model: We assume that the control center can remotely change the set points of PV inverters ($\widetilde{pg}, \widetilde{qg}$). Such a capability is needed so that the control center could ensure safety by preventing voltage spiking due to excess generation in the DN. In our attack model, the adversary intercepts the communication between the control center and substation and introduces incorrect setpoints. Based on these false setpoints, the substation commands the PVs to disrupt and draw VAR from the DN instead of supplying power and VAR to the DN.

We adopt a worst-case approach to evaluate DN vulnerability to such PV disruptions. The real and reactive power setpoints of the PVs that are introduced by the attacker are denoted by $\widetilde{pg}^a$ and $\widetilde{qg}^a$ respectively. We further introduce a vector $\delta \in \{0,1\}^N$, where $N = |\mathcal{N}_0|$. The vector δ denotes the state of the PVs. A PV at node i that is compromised is denoted by $\delta_i = 1$. If the PV node is not compromised, then $\delta_i = 0$. We denote the combined set of attacker strategies by $\psi = (\delta, \widetilde{pg}^a, \widetilde{qg}^a)$. The actual relationship between the desired attacker setpoints $(\widetilde{pg}^a, \widetilde{qg}^a)$ and the actual setpoints of the PVs is presented later (see equation (9)).

It is reasonable to assume that the number of disrupted PVs is bounded by $M \leq N$. This can model the resources available to the attacker and/or the scope of actions that a single substation can take. This resource constraint imposes that the attacker can attack at most M PVs, i.e.,

$$\sum_{i\in\mathcal{N}_0} \delta_i \leq M. \tag{6}$$

The objective of the attacker is to generate an attack vector δ constrained by (6) to violate voltage constraint (1) at one or more nodes, and to induce the defender to respond by satisfying only partial demand at a subset of nodes.

2) Defender Model: The defender model assumes that the defender actions comprise of a tuple $\phi = (\gamma, \widetilde{pg}^d, \widetilde{qg}^d)$. The vector $\gamma \in [0,1]^N$ denotes the load control parameter, i.e. the extent to which the loads are satisfied. $\gamma_i \in [\underline{\gamma}_i, 1]$ denotes the fraction of the demand that is satisfied at load i, where $\underline{\gamma}_i \geq 0$ is the maximum permissible fraction of load control at node i, i.e.,

$$\underline{\gamma}_i \leq \gamma_i \leq 1. \tag{7}$$

For each $i \in \mathcal{N}_0$, the actual real and reactive power drawn by the loads are given as follows:

$$pc_i = \gamma_i pc_i^{\mathrm{d}}, \quad qc_i = \gamma_i qc_i^{\mathrm{d}}. \tag{8}$$

The vectors $\widetilde{pg}^d$ and $\widetilde{qg}^d$ denote the PV setpoints as desired by the defender. The exact relationship between the actual PV setpoints and the desired attacker and defender real and reactive power setpoints of the PVs is given by:

$$\widetilde{pg}_i = \delta_i \widetilde{pg}_i^a + (1-\delta_i)\widetilde{pg}_i^d, \tag{9a}$$

and

$$\widetilde{qg}_i = \delta_i \widetilde{qg}_i^a + (1-\delta_i)\widetilde{qg}_i^d. \tag{9b}$$

We assume that the actual outputs of the PV, i.e., the real power pg_i and the reactive power qg_i attain the respective setpoints $\widetilde{pg}_i$ and $\widetilde{qg}_i$ relatively quickly, in comparison to the total duration of the attacker-defender interaction.

3) Loss Function: We define the cost incurred due to loss of voltage regulation as follows:

$$L_{\mathrm{LOVR}} \coloneqq \max_{i\in\mathcal{N}_0} w_i \left(\underline{\nu}_i - \nu_i\right)_+,$$

where $w_i \in \mathbb{R}_+$ is the relative importance assigned to the voltage constraint (1) at node $i \in \mathcal{N}_0$, and $\max_{i\in\mathcal{N}_0} w_i \left(\underline{\nu}_i - \nu_i\right)_+$ is the maximum (taken over all the non-substation nodes) of the weighted positive parts of the difference between the lower bound $\underline{\nu}_i$ and nodal voltage square ν_i. We define the cost incurred due to load control as follows:

$$L_{\mathrm{LL}} \coloneqq \sum_{i\in\mathcal{N}_0} C_i \left(1-\gamma_i\right) pc_i,$$

where $C_i \in \mathbb{R}_+$ is the cost of shedding unit load at node i.

For an attacker choice ψ and a defender choice γ, the composite loss function can be expressed as follows:

$$L(\psi, \phi) = L_{\mathrm{LOVR}} + L_{\mathrm{LL}}. \tag{10}$$

In the strictly competitive, perfect information sequential game the objective of the attacker (resp. the defender) is to maximize (resp. minimize) the composite loss function:

$$\min_{\psi} \max_{\phi} L(\psi, \phi) \quad \text{s.t.} \quad (3)-(4), (6)-(9). \tag{11}$$

III. Sequential Game with Linear Power Flow

A. Transformation to Bilevel Network Interdiction Problem

The original maximin problem formulation (11) has non-linear power flow equations as constraints. We anticipate that its complete solution will involve the application of techniques from convex programming [7] as well as advanced methods in integer programming [15], e.g. generalized Bender's decomposition. Our goal in this paper is modest. We make a first step in a) gaining intuition about the critical nodes, and b) understanding the tradeoff between the loss in voltage regulation vs. load shedding induced due to load control by the defender. To achieve this goal, we solve the sequential game under linear power flow equations (LPF). Fortunately, we know from the literature [2], [7], that LPF equations may provide reasonable (local) approximations to non-linear power flows in a wide range of conditions.

Before moving forward we need to define certain parameters. Let $K_{ij} \coloneqq \frac{r_{ij}}{x_{ij}}$, and let $\underline{K}$ and $\overline{K}$ denote the minimum and maximum of the K_{ij}s over all the edges $(i,j) \in \mathcal{E}$. Next, we introduce the following choices of attacker set-points for the compromised PV node i:

$$\widetilde{pg}_i^a = 0 \text{ and } \widetilde{qg}_i^a = -\overline{sg}_i, \tag{12}$$

and define the following choices of defender set-points for the non-compromised PV node j:

$$\widetilde{pg}_j^d = \frac{K_{avg}\overline{sg}_j}{\sqrt{K_{avg}^2+1}} \text{ and } \widetilde{qg}_j^d = \frac{\overline{sg}_j}{\sqrt{K_{avg}^2+1}}, \tag{13}$$

where $K_{avg} \coloneqq \frac{\sum_{(i,j)\in\mathcal{E}} r_{ij}}{\sum_{(i,j)\in\mathcal{E}} x_{ij}}$ is the ratio of sum of resistances over all the lines to the sum of reactances over all the lines. Note that $K_{avg} \in [\underline{K}, \overline{K}]$.

Now, consider the following simplified and approximate version of the sequential game (11):

$$[\text{ADLPF}] \quad \max_{\psi} \min_{\phi} L(\psi, \phi) \tag{14}$$

$$\text{s.t.} \quad (5)-(9), (12)-(13), \tag{15}$$

where the NPF equations (4) are replaced by the LPF equations (5). Here we fix the attacker and defender setpoints using equations (12) and (13), respectively, and note that these setpoints define the $\widetilde{pg}, \widetilde{qg}$ in (9).

It is straightforward to reformulate [ADLPF] as follows:

$$[\text{ADLPF}*] \quad \min_{\psi} \max_{\mathbf{y}} \; -t - \sum_{i\in\mathcal{N}_0} (1-\gamma_i)C_i \tag{16}$$

$$s.t. \quad (5)-(9), (12)-(13) \tag{17}$$

$$\nu_i - t \le \underline{\nu}_i \tag{18}$$

where we define $\mathbf{y} \coloneqq (\gamma, P, Q, \nu, pc, qc, pg, qg)$. It can be checked that the above problem can be reformulated as the following BLNIP:

$$\begin{aligned} [\text{ADLP1}] \quad z_1^* &= \min_{\mathbf{x}\in X} z_1(\mathbf{x}), \text{where} \\ z_1(\mathbf{x}) &\equiv \max_{\mathbf{y}} \; \mathbf{c}^T\mathbf{y} \\ &\text{s.t. } A\mathbf{y} \le b \\ &\quad 0 \le \mathbf{y} \le U(\mathbf{1}-\mathbf{x}), \end{aligned}$$

where $\mathbf{y}$ denotes the defender actions, $\mathbf{x} \coloneqq \delta$ denotes the attacker actions, U denotes how the attacker's actions restrict the defender's actions. Under no attack, the defender's actions are constrained as $0 \le y_i \le u_i$, where $u_i = U_{ii}$. We refer the reader to [18] for further details on a known computational approach for solving BLNIP.

Note that in formulating [ADLPF$*$], we need to fix the setpoints $\widetilde{pg}^a$, $\widetilde{qg}^a$, $\widetilde{pg}^d$ and $\widetilde{qg}^d$ apriori because the inner problem of BLNIP is an LP, while the original equation (3) is non-linear.

B. Choices of real and reactive power set-points

In this Section, we introduce three Claims regarding the choice of setpoints made by the attacker (resp. defender) for any compromised (resp. non-compromised) PV node. The formal proofs of these Claims are omitted for the sake of brevity. First, we claim that the attacker setpoints, as defined in (12), are optimal:

Claim 1: The preferred attacker setpoints for the PVs will be such that the real power setpoint is zero, and the reactive power setpoint is negative of the apparent power capability.

Intuitively, the worse case attack will induce as much loss of supply as possible. From (3), we can argue that the maximum loss of supply will occur when the attacker setpoints are chosen according to (12).

Second, the following Claim suggests a range/intervals (upper and lower bounds) from which the defender should choose the setpoints of non-compromised PVs for the purpose of maintaining desirable voltage profile. These intervals provide rough guidelines about defender setpoints for heavy loading conditions that are likely to arise under attack scenarios.

Claim 2: Under linear power flow assumption, it is desirable to select defender set-points for each non-compromised node j such that:

$$\widetilde{pg}_j^d \in \left[\frac{\underline{K}\overline{sg}_j}{\sqrt{\underline{K}^2+1}}, \frac{\overline{K}\overline{sg}_j}{\sqrt{\overline{K}^2+1}}\right], \text{and}$$

$$\widetilde{qg}_j^d \in \left[\frac{\overline{sg}_j}{\sqrt{\overline{K}^2+1}}, \frac{\overline{sg}_j}{\sqrt{\underline{K}^2+1}}\right].$$

Intuitively, when the load in the DN is high, the defender the output of the PVs will lie at the boundary of the semi-circle generated by equation (3). Further, one can show that the

ratio of the coefficients of pg_j and qg_j in the expression of ν_k for any node k, lie in the interval $[\underline{K}, \overline{K}]$. Hence, applying first order conditions to maximize voltages of different nodes, we obtain the respective ranges specified in the Claim 2.

Finally, we make the following claim for homogeneous DNs:

Claim 3: If the R/X ratio is the same, say K, for all the lines, then the defender's preferred setpoints will be $\widetilde{pg}_i^d = \frac{K\overline{sg}_i}{\sqrt{K^2+1}}$ and $\widetilde{qg}_i^d = \frac{\overline{sg}_i}{\sqrt{K^2+1}}$.

From Claim 1 and Claim 3, we obtain the following expressions for the actual setpoints of PV nodes (see (9)):

$$\widetilde{pg}_i = \begin{cases} \frac{K\overline{sg}_i}{\sqrt{K^2+1}} & \text{if } \delta_i = 1 \\ 0 & \text{if } \delta_i = 0, \end{cases}$$

and

$$\widetilde{qg}_i = \begin{cases} \frac{\overline{sg}_i}{\sqrt{K^2+1}} & \text{if } \delta_i = 1 \\ -\overline{sg}_i & \text{if } \delta_i = 0. \end{cases}$$

Finally, thanks to the Claim 3, the loss function $L(\psi, \phi)$ can be written as $L(\delta, \phi)$, as the optimal attacker preferred setpoints are already known.

IV. Optimal Attack Strategy for the Case of No Defender Response

For practical considerations, it is of interest to look at the strategies of the attacker if the defender does not respond in time and/or faces the constraint that load shedding cannot be introduced. Of course, in reality the defender may be able to manipulate other parameters like on-load tap changer, capacitors, etc. Still, the time-scale at which these devices respond may be much larger. Thus, we would like to study the following simple problem where the attacker faces a static defender:

$$\max_{\delta} \; L_{\text{LOVR}}(\delta, \phi)$$
$$\text{s.t.} \quad (3), (5) - (8)$$

We are able to arrive a structural result for this simple problem which we explain next. First, let us introduce the following useful definition:

Definition A: Let $\mathcal{P}_i$ and $\mathcal{P}_j$ denote the paths from the substation node 0 to node i and node j, respectively. The common path resistance between the nodes i and j, denoted by R_{ij}, is the sum of resistances of lines belonging to the intersection of paths $\mathcal{P}_i$ and $\mathcal{P}_j$:

$$R_{ij} := \sum_{(k,l) \in \mathcal{P}_i \cap \mathcal{P}_j} r_{kl}, \text{ for } i, j \in \mathcal{N}_0.$$

The common path reactance between two nodes i and j, denoted by X_{ij}, can be defined similarly:

$$X_{ij} := \sum_{(k,l) \in \mathcal{P}_i \cap \mathcal{P}_j} x_{kl}, \text{ for } i, j \in \mathcal{N}_0.$$

Consider $i, j, k \in \mathcal{N}_0$ where i is a *pivot node* we focus on. We define a pivot node as any node $i \in \mathcal{N}_0$ whose voltage changes we are interested in investigating when other nodes are compromised. The main idea is that by considering each node as a pivot node and iterating over all the nodes we will be able to find the optimal attacker plan. Assume for the sake of simplicity that the PVs at nodes j and k are homogeneous, i.e., $\overline{sg}_j = \overline{sg}_k$. Let $\Delta_j(\nu_i)$ be the change in the square of magnitude of the voltage at node i that will result after the PV at j^{th} node is disrupted. Further, let $J := \{j : \delta_j = 1\}$ denote the subset of PV nodes that may be disrupted. Let $\Delta_j(\nu_i)$ denote the change in the square of voltage magnitude when the PVs belonging to set J are disrupted. The effect of PV disruption at either node j or k on the node i depends upon the locations of node j, k relative to node i. To characterize the locations of j, k, we first define an order over nodes relative to the pivot node i. Let for any given node $i \in \mathcal{N}$, $\mathcal{P}_i$ be the path from the root node to node i. So, $\mathcal{P}_i$ is an ordered set of edges starting from the root node and ending on node i; see Fig. 2. Given a pivot node $i \in \mathcal{N}_0$, we want to order all the nodes in $\mathcal{N}$ as upstream or downstream to each other relative to i. We define the relative ordering $\preceq_i$ as follows:

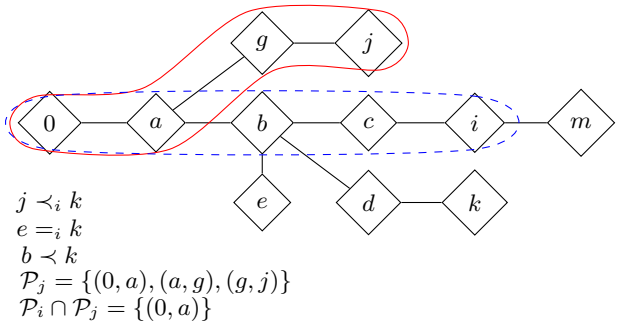


Fig. 2. Precedence description of the nodes for a tree network

Definition B: Nodes $i, j, k \in \mathcal{N}_0$ and i is the pivot node.

$$j \preceq_i k \iff \mathcal{P}_i \cap \mathcal{P}_j \subseteq \mathcal{P}_i \cap \mathcal{P}_k$$

$$j \prec_i k \iff \mathcal{P}_i \cap \mathcal{P}_j \subset \mathcal{P}_i \cap \mathcal{P}_k$$

$$j =_i k \iff \mathcal{P}_i \cap \mathcal{P}_j = \mathcal{P}_i \cap \mathcal{P}_k$$

The following theorem states that, provided the outputs of the PVs at node j and node k are identical, if the node j is upstream to node k relative to the pivot node i, then the PV disruption at node k will have more impact on ν_i than the PV disruption on node k. If, neither j is upstream or downstream of k relative to i, then the individual effects of PV disruptions at j, k on ν_i would be identical. Further, the total effect of PV disruptions of a subset of nodes J will be the sum of individual effects of PV disruption at nodes in J.

Theorem 1: For a tree network, given nodes $i, j, k \in \mathcal{N}_0$ where i is the pivot node and $\preceq_i$ is as defined in B, the effects of PV disruptions on ν_i are as follows:

i) $j \prec_i k$ and $pg_j + \mathbf{j}qg_j = pg_k + \mathbf{j}qg_k$ and $\overline{sg}_j = \overline{sg}_k \implies \Delta_j(\nu_i) < \Delta_k(\nu_i)$

ii) $j =_i k$ and $pg_j + \mathbf{j}qg_j = pg_k + \mathbf{j}qg_k$ and $\overline{sg}_j = \overline{sg}_k \implies \Delta_j(\nu_i) = \Delta_k(\nu_i)$

iii) $\forall J \subseteq \mathcal{N}, \Delta_J(\nu_i) = \sum_{j \in J} \Delta_j(\nu_j)$

Proof: Let Λ_i denote the set of nodes belonging to subtree rooted at node i. Applying the LPF (5a) - (5c) recursively, equation (5c) can be written as

$$\nu_b = \nu_a - 2r_{ab} \sum_{c \in \Lambda_b}^{n} p_c + 2x_{ab} \sum_{c \in \Lambda_b} q_c, \ \forall (a,b) \in \mathcal{E} \quad (20)$$

where p_c, q_c are the net real and reactive power consumed at node c, i.e., $p_c = pc_c - pg_c$ and $q_c = qc_c - qg_c$. Applying equation (20) recursively, we get the following equation.

$$\nu_a = \nu_0 - 2 \left[\sum_{(b,c) \in \mathcal{P}_a}^{i} r_{bc} \sum_{d \in \Lambda_c}^{n} p_d + x_{bc} \sum_{d \in \Lambda_c}^{n} q_d \right]$$

The above equation can be rewritten as

$$\nu_a = \nu_0 - 2 \left[\sum_{d \in \mathcal{N}_0} (p_d R_{ad} + q_d X_{ad}) \right]$$

When $\delta_j = 1$, i.e., the PV j is compromised, the only change that takes place is in the power supplied at node j. Hence,

$$\Delta_j(\nu_i) = 2(pg_j R_{ij} + qg_j X_{ij}).$$

However,

$$\begin{aligned} j \prec_i k \implies & \mathcal{P}_i \cap \mathcal{P}_j \subset \mathcal{P}_i \cap \mathcal{P}_k \\ \implies & R_{ij} < R_{ik} \text{ and } X_{ij} < X_{ik} \end{aligned}$$

$$\begin{aligned} \therefore \Delta_j(\nu_i) &= 2(pg_j R_{ij} + qg_j X_{ij}) \\ &< 2(pg_k R_{ik} + qg_k X_{ik}) = \Delta_k(\nu_i) \end{aligned}$$

$$\begin{aligned} j =_i k \implies & \mathcal{P}_i \cap \mathcal{P}_j = \mathcal{P}_i \cap \mathcal{P}_k \\ \implies & R_{ij} = R_{ik} \text{ and } X_{ij} = X_{ik} \end{aligned}$$

Hence, $\Delta_j(\nu_i) = \Delta_k(\nu_i)$. Due to linearity, the third part of Theorem 1 follows. ■

V. Benchmarking on a Small-sized DN

We present a simple case study to assess the validity of the structural insight obtained from Section IV (and in particular Theorem 1). We also compute the optimal attack plan for the maximin problem [ADLPF$*$] (Section III). Specifically, we implemented the following algorithms:

First, we implemented the standard Bender's Cut (BC) algorithm to solve the bilevel network interdiction problem [ADLPF$*$]. Recall that the inner LP problem uses linear power flow equations. In the following, we use the abbreviation *BC-NPF* to denote the result obtained by evaluating the defender's optimal response under the nonlinear power flow constraints when he faces the optimal attack plan obtained from [ADLPF$*$].

Second, building on Theorem 1, we implemented the following Greedy Algorithm that proceeds iteratively as follows: First, the *GA* assumes that the defender actions are fixed. It considers a pivot node i, and computes the optimal set of PV nodes J_i^* to disrupt such that the impact on ν_i (quantified by $\Delta_{J_i^*}(\nu_i)$) is maximized. The J_i^* is computed by sorting the PV nodes in the decreasing order of the impact which their individual disruption will have on ν_i, and then choosing the top M PVs with largest impact. Second, it selects the pivot node $k \coloneqq \arg\min_{i \in \mathcal{N}_0} \nu_i - \Delta_{J_i^*}(\nu_i)$, that will have the least ν_i after the attack. The *GA* selects J_k^* as the set of PVs that should be compromised by the attacker. For this set of compromised PVs (i.e., attack plan), the optimal defender's actions are computed under the nonlinear power flow constraints. In the next iteration, the new defender actions obtained at the end of previous iteration are assumed to be fixed, and the aforementioned two-step process is repeated. If in any two consecutive iterations, the attack vectors computed are identical, the algorithm outputs the attack vector and the optimal defender actions, and terminates successfully. Otherwise, if the upper limit on the number of iterations is reached, the algorithm terminates unsuccessfully. We use the abbreviation *GA* to denote the results of the greedy algorithm.

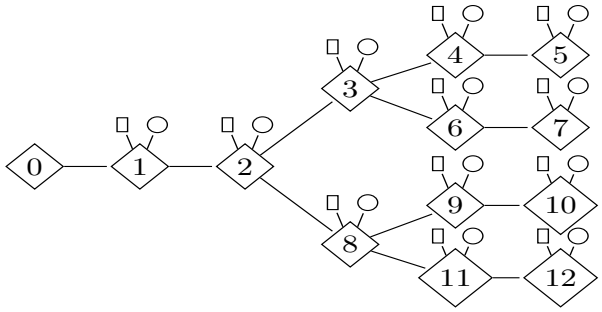


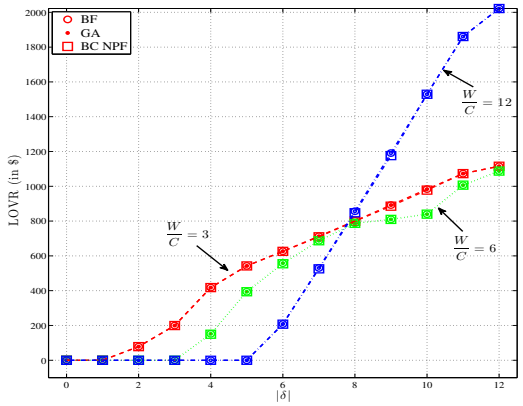
Fig. 3. Tree Network

TABLE I
Parameters of the DN in Figure 3.

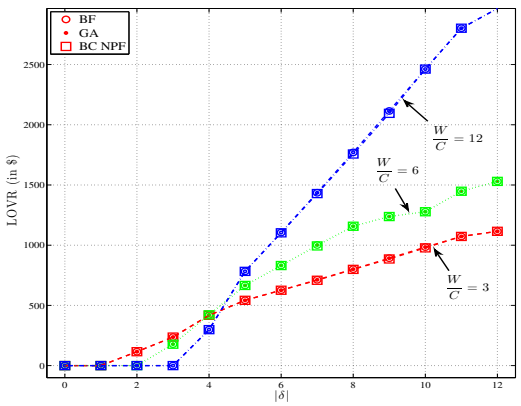
Parameters	Values
$r + \mathbf{j}x$	$(0.33 + 0.38\mathbf{j})\ \Omega/km$
Line Length	$4\ km$
pc_i^{d}	$25\ kW$
$qc_i^{\mathrm{d}} = 0.3pc_i^{\mathrm{d}}$	$7.5\ kvar$
$\overline{pg}_i = 0.7pc_i^{\mathrm{d}}$	$17.5\ kW$
$\overline{sg}_i = 1.1\overline{pg}_i$	$19.25\ kVA$
ν_0	$4\ kV$

Finally, for the sake of benchmarking, we exhaustively enumerate all the possible attack plans for a small-sized network (see Fig. 3), compute the defender's best response by solving the inner minimization problem for each attack plan using NPF, and arrive at the best attack plan. We use the abbreviation *BF* to denote the results from this exhaustive search. We implemented *BC-NPF*, *GA*, and *BF* using the Gurobi Solver in JuMP (*Julia for Mathematical Programming*) [12].

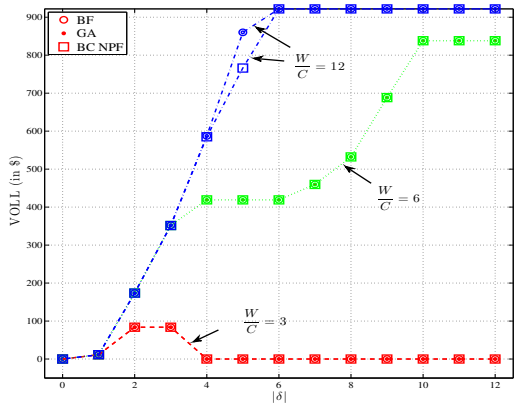
We now list the parameters of the DN considered in this study (Fig. 3). It has one substation node and $N = 12$ loads/PV nodes. We describe all the parameters in *p.u.*



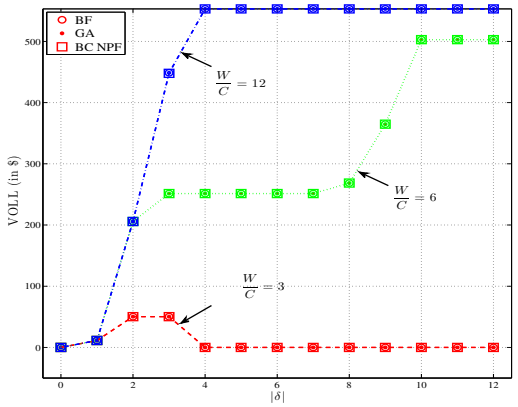
(a) Cost due to LOVR vs $|\delta|$, $\underline{\gamma} = 0.5$



(b) Cost due to LOVR vs $|\delta|$, $\underline{\gamma} = 0.7$



(c) Cost due to LL vs $|\delta|$, $\underline{\gamma} = 0.5$



(d) Cost due to LL vs $|\delta|$, $\underline{\gamma} = 0.7$

Fig. 4. LOVR and LL plots for the 12 Node Distribution Network.

system. In the homogeneous circuit, all the lines, loads and PV-enabled nodes have similar physical properties; they are summarized in Table I.

All the lines have identical impedance per unit length equal to $(0.33 + 0.38)\ \Omega/km$, and each line is $4\ km$ long. Each load node i demands real power equal to $pc_i^{\mathrm{d}} = 25\ kW$, and the reactive power demand is $qc_i^{\mathrm{d}} = 0.3pc_i^{\mathrm{d}}$. This load demand corresponds to power factor of approximately 0.957. Each node has one PV (denoted by diamonds in the Fig. 3). For every PV-enabled node i, the maximum real power output of the PV is $70\ \%$ of the real power demanded by the corresponding load, i.e., $\overline{pg}_i = 0.7pc_i^{\mathrm{d}}$, and the apparent power capability of the inverter is 110% that of the maximum real power output, i.e., $\overline{sg}_i = 1.1\overline{pg}_i$. The nominal phase-to-neutral voltage is taken to be $4\ kV$. The voltage at each node is constrained to be within 0.95 and 1.05 *p.u.* The cost parameter for load control is $C = 7$ \$ per kW, converted to the *p.u.* system.

We use this DN to compare the results from *BC-NPF*, *GA*, and *BF*. These results are presented in Fig. 4. We observe that the performance of three algorithms closely match on the test DN under consideration. In this computational study, we consider scenarios of maximum controllable load percentage $\underline{\gamma} = 50\ \%$ and $\underline{\gamma} = 70\ \%$. For each of these scenarios, we vary the number of PV nodes targeted, M, from 0 to maximum possible value 12. As explained below, we also vary the $\frac{W}{C}$ ratio to capture the effect of different weights to the costs due to LOVR and LL.

In the figure 4a, the $\frac{W}{C} = 3$ ratio roughly corresponds to the minimum $\frac{W}{C}$ ratio for which there is no load control. In this case, the defender the cost-to-benefit ratio of doing load control is so high that the defender chooses to satisfy the loads completely (i.e., no load control). In contrast, the $\frac{W}{C} = 12$ ratio roughly corresponds to the maximum $\frac{W}{C}$ ratio for which the defender tries to invoke maximum admissible load control. The $\frac{W}{C} = 6$ is an intermediate ratio between these two extreme ratios.

From the computational results, we first observe that both the LOVR and LL cost curves start from the origin. This implies that when there is no attack (our nominal starting

condition), the voltages of the nodes are within the lower and upper bounds required for the safe operation. As $|\delta|$ increases, one or both cost curves start increasing. This indicates that as more PVs are attacked, the defender incurs costs due to LOVR, and in addition, he may have to impose load control to better regulate the DN. Let us interpret the physical response of DN after the attack. The apparent power output of the attacked PVs changes to negative value from the positive value in nominal condition. This increases the net load in the network, and causes the voltages to drop below the lower bounds at certain nodes, leading to LOVR. The defender then tries to do load control to reduce the net load in the network, so that the voltage regulation improves.

Perhaps the more interesting observation is that, as the number of compromised PV nodes, M, increases, the cost of LL first increase rapidly and then flattens out. This result can be explained as follows. Depending upon the $\frac{W}{C}$ ratio, there is a subset of downstream loads that are beneficial in terms of the value that defender can obtain by controlling them. That is, if the loads belonging to this subset are controlled, the benefit due to better voltage regulation will outweigh the cost incurred due to load control. In other words, the defender will impose load control on certain downstream loads, because controlling these loads has greater benefit-to-cost ratio. In the contrary, controlling the loads outside this subset, will impose a higher load control cost relative to the cost savings due to improved voltage regulation. Hence, the cost of LL curves increase in the beginning; however, as soon as the load control capability in this subset is exhausted, the defender tries to maintain the demand at other loads. The size of the subset of the beneficial downstream loads depends upon the $\frac{W}{C}$ ratio. The higher the ratio, the more the size of this subset. Hence, the value of $|\delta|$, at which the LL cost curve flattens out increases as the $\frac{W}{C}$ ratio increases.

The cost curve for LOVR also shows interesting behavior as the number of compromised PV nodes increases. This cost is convex increasing for small $|\delta|$, whereas it is concave increasing for large $|\delta|$. This observation can be explained by the fact that the attacker prefers to compromise downstream nodes over upstream ones (which is the main insight from Theorem 1). Initially, the attacker is able to gain a lot by attacking more lucrative downstream nodes. However, as the downstream nodes are eventually exhausted, the attacker has to target the relatively less lucrative upstream nodes. Hence, the marginal benefit of the attacking an additional PV decreases for higher values of $|\delta|$.

VI. Concluding Remarks

In this article, we investigated the problem of assessing the vulnerability of DNs in the face of disruption to DERs such as PV nodes. We modeled the attacker-defender interaction as a bilevel network interdiction problem, where the attacker targets a subset of PVs by changing their setpoints, and the defender responds by controlling the setpoints of non-compromised PVs and by imposing partial load control. We restrict our attention to costs that the defender incurs due to loss in voltage regulation and cost of load control. Other costs such as cost of PV control should be considered in future work. With the help of useful approximations and a case study, we demonstrated the structural result that the attacker targets downstream PVs relative to upstream ones.

Acknowledgements

The authors are grateful for the valuable inputs from Bruno Prestat and Pascal Sitbon.

References

[1] Mesut E. Baran, Hossein Hooshyar, Zhan Shen, and Alex Q. Huang. Accommodating high pv penetration on distribution feeders. *IEEE Trans. Smart Grid*, 3(2):1039–1046, 2012.

[2] H. D. Chiang and Mesut E. Baran. On the existence and uniqueness of load flow solution for radial distribution power networks. *IEEE Transactions on Circuits and Systems*, 37(3):410–416, 1990.

[3] Emiliano Dall'Anese and Georgios B. Giannakis. Sparsity-leveraging reconfiguration of smart distribution systems. *CoRR*, abs/1303.5802, 2013.

[4] A. Delgadillo, J. M. Arroyo, and N. Alguacil. Analysis of Electric Grid Interdiction With Line Switching. *Power Systems, IEEE Transactions on*, 25(2):633–641, May 2010.

[5] A. Delgadillo, J. M. Arroyo, and N. Alguacil. Analysis of Electric Grid Interdiction With Line Switching. *Power Systems, IEEE Transactions on*, 25(2):633–641, May 2010.

[6] Masoud Farivar and Steven H. Low. Branch flow model: Relaxations and convexification. In *CDC*, pages 3672–3679. IEEE, 2012.

[7] Masoud Farivar, Russell Neal, Christopher R. Clarke, and Steven H. Low. Optimal inverter var control in distribution systems with high pv penetration. *CoRR*, abs/1112.5594, 2011.

[8] K. Hatipoglu, I. Fidan, and G. Radman. Investigating effect of voltage changes on static zip load model in a microgrid environment. In *North American Power Symposium (NAPS), 2012*, pages 1–5, Sept 2012.

[9] Ian A. Hiskens. What's smart about the smart grid? In Sachin S. Sapatnekar, editor, *DAC*, pages 937–939. ACM, 2010.

[10] Soumya Kundu and Ian A. Hiskens. Distributed control of reactive power from photovoltaic inverters. In *ISCAS*, pages 249–252, 2013.

[11] J. Lavaei, D. Tse, and Baosen Zhang. Geometry of power flows and optimization in distribution networks. *Power Systems, IEEE Transactions on*, 29(2):572–583, March 2014.

[12] M. Lubin and I. Dunning. Computing in Operations Research using Julia. *ArXiv e-prints*, December 2013.

[13] Ruofei Ma, Hsiao-Hwa Chen, Yu-Ren Huang, and Weixiao Meng. Smart grid communication: Its challenges and opportunities. *IEEE Trans. Smart Grid*, 4(1):36–46, 2013.

[14] Report on Electric Sector Failure Scenarios and Impact Analyses. Electric power research institute. *National Electric Sector Cybersecurity Organization Resource Report*, September 2013.

[15] J. Salmeron, K. Wood, and R. Baldick. Analysis of electric grid security under terrorist threat. *Power Systems, IEEE Transactions on*, 19(2):905–912, May 2004.

[16] Chee-Wooi Ten, Govindarasu Manimaran, and Chen-Ching Liu. Cybersecurity for critical infrastructures: Attack and defense modeling. *Trans. Sys. Man Cyber. Part A*, 40(4):853–865, July 2010.

[17] Konstantin S. Turitsyn, Petr Sulc, Scott Backhaus, and Michael Chertkov. Options for control of reactive power by distributed photovoltaic generators. *Proceedings of the IEEE*, 99(6):1063–1073, 2011.

[18] R.K. Wood. Bilevel network interdiction models: Formulations and solutions. In *Wiley Encyclopedia of Operations Research and Management Science*, 2011.