A Trilevel Model for Segmentation of the Power Transmission Grid Cyber Network

Bryan Arguello,Emma S. Johnson, and Jared L. Gearhart

Abstract—Network segmentation of a power grid's communication system is one way to make the grid more resilient to cyber attacks. We develop a novel trilevel programming model to optimally segment a grid communication system, taking into account the actions of an information technology (IT) administrator, attacker, and grid operator. The IT administrator is given an allowance to segment existing networks, and the attacker is given a fixed budget to attack the segmented communication system in an attempt to inflict damage on the grid. Finally, the grid operator is allowed to redispatch the grid after the attack in order to minimize damage. The resulting problem is a trilevel interdiction problem, which we solve by leveraging current research in bilevel branch and bound. We demonstrate the benefits of optimal network segmentation through case studies on the 9-bus WSCC system and the 30-bus IEEE system. These examples illustrate that network segmentation can significantly reduce the threat posed by a cyber attacker with perfect knowledge of the grid.

Index Terms—Power system protection, Cyberspace, Networked control systems, Optimization, Multilevel systems

Nomenclature

Sets

$\mathcal{E}$	Security enclaves
$\mathcal{B}$	Balancing authority entities
$\mathcal{C}$	Control center entities
$\mathcal{S}$	Substation entities
$\mathcal{T}$	Tiers of entities in the communication network: $\mathcal{T} = \{\mathcal{B}, \mathcal{C}, \mathcal{S}\}$
$\mathcal{Z}$	Pairs of adjacent tiers in communication network: $\mathcal{Z} = \{(\mathcal{B}, \mathcal{C}), (\mathcal{C}, \mathcal{S})\}$. Generic members of $\mathcal{Z}$ are always denoted (A, B) where A is the tier above and B is the tier below.
$\mathcal{E}_0(T)$	Existing enclaves in tier $T \in \mathcal{T}$
$\mathcal{E}_1(T)$	New enclaves in tier $T \in \mathcal{T}$
$\mathcal{E}(T)$	Enclaves in tier $T \in \mathcal{T}$: $\mathcal{E}(T) = \mathcal{E}_0(T) \cup \mathcal{E}_1(T)$
T_n	Entities in tier T which communicate with entity n one tier above
$\mathcal{K}$	Transmission lines

$\mathcal{G}$	Generators
$\mathcal{G}_s$	Generators at substation s
$\mathcal{L}$	Loads
$\mathcal{L}_s$	Loads at substation s
$\mathcal{R}$	Relays
$\mathcal{R}_s$	Relays at substation s
$\mathcal{R}_d$	Relays which control load d
$\mathcal{R}_k$	Relays which control line k
$\mathcal{R}_g$	Relays which control generator g

Parameters

$Q_{n,e}$	Binary indicating if existing enclave e is in entity n
U	Maximum number of enclaves attacker can penetrate
Θ_k	Transformer shift angle on line k
B_k	Line charging susceptance of line k
D_d	Demand at load d
$s(d)$	Substation served by load d
$s(g)$	Substation served by generator g
$\overline{P}_g$	Maximum real power output of generator g
$\overline{F}_k$	Thermal limit of line k
$o(k)$	Origin bus of line k
$d(k)$	Destination bus of line k

Binary Variables

IT Operator Decisions:

$x_{e,r}$	Indicates whether or not enclave $e \in \mathcal{E}(\mathcal{S})$ communicates with relay r
$y_{e,f}$	Indicates whether or not enclave e communicates with enclave f
$q_{n,e}$	Indicates whether or not enclave e is in entity n
$t_{e,n}$	Indicates whether or not enclave e communicates with entity n (in the tier below)

Attacker decisions:

z_e	Indicates whether or not the attacker infiltrates enclave e
δ_r	Indicates whether or not relay r is compromised
v_k	Indicates whether or not line k is operational
u_d	Indicates whether or not the load d is online
w_g	Indicates whether or not generator g is operational

Continuous Variables

Grid operator decisions:

θ_s	Voltage angle at substation s
p_g	Real power output of generator g
f_k	Real power flow on line k
l_d	Real power load shed at load d
L	Real power total load shed

I. INTRODUCTION

THE electric grid, like many infrastructure systems, historically relied on air-gapped and specialized control networks for security. For convenience and cost savings, grid/utility owners are increasingly integrating traditional information technologies (IT) into their control systems. This shift has increased the attack surface for electric grids and has made them more vulnerable to cyber attacks. This was illustrated in the December 2015 cyber attack against the Ukrainian power grid [1]. According to [2], nation-state actors in particular pose a significant threat to cyber-physical systems and also possess the ability to gather the necessary information in advance of attacks. Criminal organizations, terrorists, hackers, and hacktivists may also pose a threat to these systems [3].

Network segmentation is one strategy that has been proposed to improve cyber network security [4]. Dividing networks into small segments and restricting communications between segments can limit the scope of attacks and the attackers' ability to pivot within a network. Several governmental and regulatory agencies have provided guidance that motivates segmentation and offer strategies for implementing it [5], [6], [7], [8]. However, as noted in [9], this guidance tends to provide general design principles and architectures, but not specific recommendations or tools for cyber network designers.

Designing segmentation strategies for cyber-physical systems presents a particularly difficult challenge as it requires consideration of two complex interconnected systems. In an IT setting, segmentation decisions can be made based on business functions and the expected consequences of disruptions to these functions (e.g. interruption to payment systems, loss of customer data). In a grid setting, the cyber and physical layers are connected and distributed over wide areas; coordinated attacks create the potential for consequences that are greater than the sum of their parts. A key recommendation in [3] is the need to "evaluate the potential risk of a coordinated cyberattack on geographically distributed targets."

In this paper, we explore the use of optimization to identify strategies for segmenting cyber networks. We use an explicit representation of the cyber and grid network layers to capture the interconnected nature of these systems and the resulting impacts of attacks. We also employ trilevel optimization techniques to capture the adaptive nature of and interactions between cyber defenders, attackers, and infrastructure operators. Finally, given the potential capabilities of the attacker, we assume that the attacker has full knowledge of the system when crafting attacks.

A. Literature Review

Several researchers have previously considered applying optimization to cyber security problems that are related to network segmentation. In [10], a heuristic optimization algorithm is used to place the minimum number of intrusion detection systems (IDS) on the nodes of a cyber network while ensuring that no nodes are more than n hops from the nearest detection system. This heuristic approach is also applied by [11] to place trust systems on the smart grid cyber network.

In [12], trust system placement is combined with network segmentation. A mixed-integer linear program (MILP) is used to partition a power system supervisory control and data acquisition (SCADA) network into a collection of enclaves such that a trust node is placed on at least one end of all arcs connecting different enclaves. This ensures that all communications between segments are scanned. The authors also develop a heuristic algorithm for partitioning and placing trust nodes on larger networks.

In [9], simulated annealing is used to identify segmentation strategies for a computer network in an air operations center that supports 40 military missions. An emulation test bed is used to simulate the performance of each segmentation strategy against a "dumb" worm attack that propagates in a manner similar to a disease in a Susceptible-Exposed-Infectious-Recovered (SEIR) model. The quality of each segmentation strategy is based on mission delay and the availability of devices on the network. In [13], this approach is extended to an algorithm that co-evolves the parameters available to the attacker and defender in order to optimize both of their strategies. The attacker still employs a "dumb" worm attack and cannot control how it propagates, but they tune the intensity of the attack to balance the number of components compromised against the probability of detection.

In [14], a heuristic method that uses minimum spanning trees is used to improve performance and scalability for the trust node placement problem on smart grid SCADA systems. Variants of this model are considered in [15] and [16]. These variants account for node centrality and communication latency when deciding where to place trust nodes. In [17], this approach is further expanded to consider both link coverage (the number of links adjacent to a trust node) and path tolerance (the longest path not passing through a trust node) when making placement decisions.

In all of the previously mentioned references that focus on cyber-physical security, only the cyber network is considered, while effects on the underlying physical system it controls are not modeled. There are several examples where segmentation decisions are made based on the underlying physical system managed by the cyber network. In [18], an emulation test bed is used to compare two segmentation strategies developed by subject matter experts (SME's) for SCADA systems in a chemical plant. They demonstrate that in process-flow operations, security and safety gains can be made by separating SCADA devices for in-flow and out-flow values. This approach is extended in [19] to include an SME-informed heuristic approach that creates segmentation strategies to separate control of in-flows and out-flows. In [20], network segmentation is used to prevent domino effects in storage tank facilities (e.g. chained explosions of nearby tanks) by an attacker who can attack at most one segment. Belief networks are used to estimate the probability of chained reactions and a graph centrality measure is used to find the segmentation strategy. These works use testbeds to study the effects of network segmentation on the underlying physical system. However, they do not use optimization to make their

network segmentation decisions.

Prior research has used bilevel programming to consider the effects on the power grid from an intentional attack, though these model only the physical grid, not the cyber communication system. For example, [21] and [22] formulate and solve a bilevel program that models an attacker with perfect knowledge of the grid and the response of a grid operator. The attacker directly de-energizes power grid components such as transmission lines and generators. The grid operator then responds through generator redispatch and load shed to minimize unmet demand. This attacker-defender model is adapted for a cyber attacker in [23], though still without modeling the cyber network explicitly.

Finally, several authors have employed trilevel programming to preemptively protect the power grid against malicious attacks in [24], [25], [26], and [27]. These works focus on hardening power grid components such as transmission lines, generators, and buses against optimal physical attacks on the grid. However, they do not consider the grid's communication system in modeling these attacks.

B. Contributions and Paper Organization

In this work, we propose segmenting the grid's communication network optimally by explicitly modeling both the cyber network and the physical network. Specifically, we use trilevel optimization to model the interconnected decisions made by an IT administrator segmenting a grid's communication network, a cyber attacker maliciously trying to damage the grid by attacking the segmented network, and a grid operator mitigating damage from the attack. The main contributions of this paper are:

- The formulation of a trilevel programming model to identify provably optimal segmentation strategies for a power transmission cyber-physical system which reduce the severity of the worst attack available to a malicious cyber attacker.
- The use of an infrastructure model (i.e., a DC optimal power flow (DCOPF) grid model) to inform network segmentation decisions.
- A description of how to reformulate the trilevel model as a bilevel program, which can be solved using algorithms from the literature (e.g. bilevel branch-and-bound).
- A case study showing network segmentation results for both the 9-bus WSCC and 30-bus IEEE test cases, where segmentation reduces the severity of the worst-case attack by 31% and 56%, respectively.

In the remainder of the paper, Sections II and III give a description of the model. Section IV details our solution methodology. Section V presents our case studies on the 9-bus and 30-bus IEEE systems. We conclude in Sections VI and VII with ideas for future work and conclusions.

II. Cyber-Physical Network Segmentation Model

We first describe the trilevel network segmentation model at a high level. We give the mathematical formulation in the next section. When we talk about the cyber-physical system, we draw a distinction between *entities*, the bodies which control the grid, *enclaves*, the networks that make up an entity's IT infrastructure, and *physical components* such as the lines, generators, and loads on the physical grid. While our model is more general, in this paper we consider three types of entities: balancing authorities, control centers, and substations.

A. Modeling the Communication Network

For a description of how a power transmission grid communication system network is structured, we refer the reader to [28]. We simplify the communication network to a 3-tier forest where each substation enclave, control center enclave, and balancing authority enclave is represented as a node, and parent-child relationships represent that the parent enclave communicates with the child enclave (i.e., data flows between the parent and child enclave). Every substation enclave is the child of exactly one control center enclave. In turn, every control center enclave is the child of exactly one balancing authority enclave. We show an example of the forest representation of the 9-bus WSCC test system communication network in Fig. 1a. In this example, every entity has only one enclave. The balancing authority enclave communicates with two different control center enclaves. The first of these communicates with three substation enclaves, and the second with six substation enclaves.

The physical power grid is connected to its communication system through relays, which are drawn as the leaves of the tree in Fig. 1a. Individual relays are controlled by exactly one substation enclave and can be used by an attacker to de-energize transmission lines, generators, and loads. In Fig. 1c, each shaded block of grid components is colored to indicate which control center enclave from Fig. 1a ultimately controls it.

B. Modeling Network Segmentation

In the outermost problem of the trilevel model, we model the decisions of an IT administrator segmenting the network in order to minimize the load shed from a worst-case cyber attack. We model network segmentation by allowing the IT administrator to partition any security enclave into two or more security enclaves while still respecting the underlying control structure among the entities. By only subdividing existing assignments, we aim to generate new segmentation strategies that are compatible with the original topology. In the graph representation, segmentation corresponds to expanding the graph by dividing nodes. If a node is divided, the new subnetwork must be assigned to the same entity as the original node. For example, if a substation enclave is divided, both new enclaves are subnetworks of the same substation entity. Edges are added to the expanded graph according to the following rules:

1) A new enclave may only communicate with (i.e., be a child of) a parent enclave if the entities containing the new enclave and the parent enclave communicated in the original graph. As an example, a substation enclave may only communicate with a control center enclave if the corresponding substation and control center originally communicated before segmentation.

(a) The 9-bus system communication network before segmentation: Arcs indicate that the parent component communicates with the child component. The balancing authority, control center, and substation enclaves comprise the communication system, and its control over the relays is shown as the leaves of the tree. (Note that in this example, we include fewer relays than in the 9-bus system used in the case study in Section V-B.)

(b) The 9-bus system communication network after segmentation. The balancing authority enclave has been subdivided into two enclaves, each of which communicates with one of the control centers. The contol centers also have two enclaves each, as does Substation 4.

(c) The 9-bus physical network before segmentation: Grid components are shaded based on what control-center enclave in the communication network controls them, and the dotted line represents what balancing authority enclave controls the components it contains. Before segmentation, there is only one balancing authority enclave. Buses 1, 2, and 3 are all controlled by relays controlled by Substations 1, 2, or 3, and hence controlled by Control Center 1. The rest of the network is controlled by Control Center 2.

(d) The 9-bus physical network after segmentation: Now the network is divided between two balancing authority enclaves, and each of the two sets of components controlled by Control Centers 1 and 2 are now divided between two different enclaves in their control center. Dotted angled lines inside of the shaded regions indicate grid components which are on different relays. For example, two of the lines adjacent to Bus 4 are on a relay controlled by the green (cross-hatched) enclave of Control Center 2, and the other line is controlled by the red (striped) enclave.
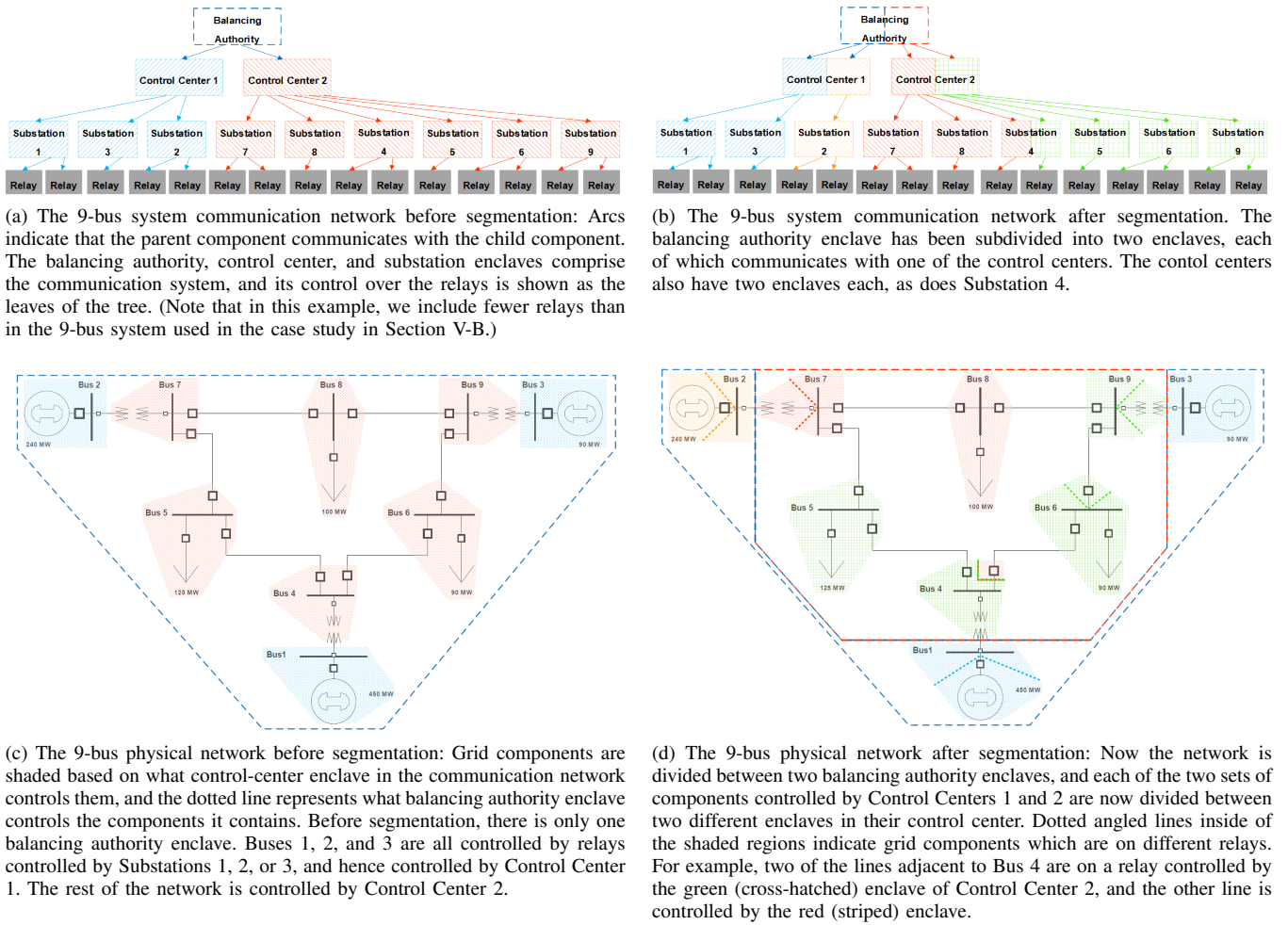
Fig. 1. The 9-bus system and a graph representation of its communication network, both before and after segmentation. The communication network is shown before segmentation in (a) and after segmentation in (b). The physical grid is shown before segmentation in (c) and after segmentation in (d). Colors are used to map which physical components are controlled by which control-center enclaves in the communication network.

2) After partitioning a substation enclave, that substation's relays must be reassigned so that each relay is controlled by exactly one of the new substation enclaves. The result of this partitioning is that a substation's grid connections are separated by the substation's enclaves, making that substation less vulnerable overall.

3) Every substation enclave must communicate with at least one relay.

Continuing our example, see Fig. 1b for a possible segmentation of the 9-bus network. We see that, after network segmentation, the balancing authority has two enclaves, one for each control center. Both control centers gain a new enclave. For each control center, the substations it originally controlled are divided between its two new enclaves. Finally, Substation 4 is segmented into two enclaves, each communicating with a separate enclave within Control Center 2. Note that each of the two relays controlled by Substation 4 is assigned to a different enclave. The corresponding diagram of the physical grid is shown in Fig. 1d.

C. Attacker-Defender Model

In the inner two problems of the trilevel model, we model a cyber attacker who finds the highest-load-shed attack on the segmented network given that the grid operator can redispatch after their attack in order to minimize load shed. We assume that attacks begin by infiltrating balancing authority enclaves, then progress by gaining access to the control center enclaves followed by substation enclaves, until the attack reaches the relays, where it can produce an effect on the grid. At each step, the attacker can never infiltrate a node in the graph without first infiltrating its parent. We assume the attacker has an upper limit on the number of enclaves that can be compromised, and every time the attacker gains access to a security enclave, exactly one unit of this budget is consumed. In other words, one unit of budget is used for every node in the graph the attacker reaches. When the attacker gains access to a substation enclave, we assume that the attacker will use all relays controlled by that substation enclave to de-energize every accessible grid component. To gain some intuition about the benefits of network segmentation, note for example that originally in Fig. 1, an attacker would need to access three enclaves to gain control of the two relays controlled by Substation 4. After segmentation, the attacker would need to access five enclaves to gain control of the same relays.

Finally, given an attack, the grid operator solves a DCOPF

on the remaining components to minimize total load shed.

III. TRILEVEL FORMULATION

We now give a mathematical formulation of the problem described in Section II. Because many constraints involving (parent, child) entity pairs hold at multiple levels – both (balancing authority, control center) and (control center, substation) – we use the set $\mathcal{T}$ to denote the set of all three communication network entities and the set $\mathcal{Z} = \{(\mathcal{B},\mathcal{C}),(\mathcal{C},\mathcal{S})\}$ to denote adjacent tiers of entities. The trilevel model is:

$$\min_{(x,y,q,t)\in\mathcal{D}} \quad \max_{(\delta,z,u,v,w)\in\mathcal{A}(x,y)} \quad \min_{(\theta,f,p,l)\in\mathcal{O}(u,v,w)} \sum_{d\in\mathcal{L}} l_d, \quad (1)$$

where $\mathcal{D}$ is the feasible region of the network designer, $\mathcal{A}(x,y)$ is the feasible region of the attacker, given the design decisions x and y, and $\mathcal{O}(u,v,w)$ is the DCOPF feasible region given the attack defined by u, v, and w.

We model $\mathcal{D}$ with the following constraints:

$$\sum_{r\in\mathcal{R}} x_{e,r} \geq 1, \qquad \forall e \in \mathcal{E}(\mathcal{S}) \tag{2}$$

$$\sum_{e\in\mathcal{E}(\mathcal{S})} x_{e,r} = 1, \qquad \forall r \in \mathcal{R} \tag{3}$$

$$q_{s,e} \leq \sum_{r\in\mathcal{R}_s} x_{e,r}, \qquad \forall s \in \mathcal{S}, e \in \mathcal{E}_1(\mathcal{S}) \tag{4}$$

$$q_{s,e} \geq x_{e,r}, \qquad \forall s \in \mathcal{S}, r \in \mathcal{R}_s, e \in \mathcal{E}_1(\mathcal{S}) \tag{5}$$

$$Q_{s,e} \leq \sum_{r\in\mathcal{R}_s} x_{e,r}, \qquad \forall s \in \mathcal{S}, e \in \mathcal{E}_0(\mathcal{S}) \tag{6}$$

$$Q_{s,e} \geq x_{e,r}, \qquad \forall s \in \mathcal{S}, r \in \mathcal{R}_s, e \in \mathcal{E}_0(\mathcal{S}) \tag{7}$$

$$\sum_{e\in\mathcal{E}(A)} y_{e,f} = 1, \qquad \forall (A,B) \in \mathcal{Z}, f \in \mathcal{E}(B) \tag{8}$$

$$\sum_{n\in T} q_{n,e} = 1, \qquad \forall T \in \mathcal{T}, e \in \mathcal{E}_1(T) \tag{9}$$

$$t_{e,n} \leq \sum_{f\in\mathcal{E}_0(B)} y_{e,f} Q_{n,f} + \sum_{f\in\mathcal{E}_1(B)} y_{e,f} q_{n,f},$$
$$\forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A), n \in B \tag{10}$$

$$t_{e,n} \geq y_{e,f} Q_{n,f}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A),$$
$$n \in B, f \in \mathcal{E}_0(B) \tag{11}$$

$$t_{e,n} \geq y_{e,f} q_{n,f}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A),$$
$$n \in B, f \in \mathcal{E}_1(B) \tag{12}$$

$$t_{e,n} \leq Q_{m,e}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}_0(A),$$
$$m \in A, n \in B_m \tag{13}$$

$$t_{e,n} \leq q_{m,e}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}_1(A),$$
$$m \in A, n \in B_m \tag{14}$$

$$y_{e,f} \in \{0,1\}, \forall (e,f) \in (\mathcal{E}(\mathcal{C})\times\mathcal{E}(\mathcal{S}))\cup(\mathcal{E}(\mathcal{B})\times\mathcal{E}(\mathcal{C}))$$
$$\tag{15}$$

$$x_{e,r} \in \{0,1\}, \qquad e \in \mathcal{E}(\mathcal{S}), \forall r \in \mathcal{R} \tag{16}$$

$$q_{n,e} \in \{0,1\}, \qquad \forall (n,e) \in \cup_{T\in\mathcal{T}}(T\times\mathcal{E}_1(T)) \tag{17}$$

$$t_{e,n} \in \{0,1\}, \quad \forall (e,n) \in (\mathcal{E}(\mathcal{C})\times\mathcal{S})\cup(\mathcal{E}(\mathcal{B})\times\mathcal{C}) \tag{18}$$

In (2), we require that every substation enclave controls at least one relay, and in (3) we enforce that every relay is controlled by exactly one substation enclave. Referring back to

the graph in Fig. 1, the $x_{r,e}$ variables represent whether or not there is an arc from enclave e in the substation tier to relay r. Constraints (4)-(7) enforce that an enclave e is assigned to a substation if and only if some relay controlled by that substation is controlled via e.

In (8), we require that, between any two adjacent tiers, every enclave in the tier below is controlled by exactly one enclave in the tier above. Since $y_{e,f}$ determines whether or not there is an arc from enclave e to enclave f, these constraints enforce that substation enclaves and control center enclaves have exactly one parent.

Constraints (9) guarantee that, for each of the tiers of the communication network, every new enclave is assigned to exactly one entity (substation, control center, or balancing authority depending on the tier). Constraints (10)-(12) enforce that enclave e communicates with entity n in the layer below e if and only if entity n has an enclave f that communicates with e. Returning to the graphs in Fig. 1, $t_{e,n}$ is an indicator of whether enclave e is a parent of some enclave assigned to entity n. Constraints (13)-(14) enforce that if enclave e communicates with entity n in the network layer below it, then e must be assigned to an entity m which controls n. Note that B_m is an instance of the set T_n with $T = B$ and $n = m$. Finally, (15)-(18) give variable domains.

Note that (10) and (12) contain bilinear terms. Since all variables in these products are binaries, they can be reformulated linearly with the introduction of one new binary variable per bilinear term. That is, we introduce $\beta_{e,f,n} \in \{0,1\}$ for all $(A,B) \in \mathcal{Z}, e \in \mathcal{E}(A), f \in \mathcal{E}_1(B), n \in B$ and the constraints

$$\beta_{e,f,n} \leq y_{e,f}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A),$$
$$f \in \mathcal{E}_1(B), n \in B \tag{19}$$

$$\beta_{e,f,n} \leq q_{n,f}, \qquad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A),$$
$$f \in \mathcal{E}_1(B), n \in B \tag{20}$$

$$\beta_{e,f,n} \geq y_{e,f} + q_{n,f} - 1, \quad \forall (A,B) \in \mathcal{Z}, e \in \mathcal{E}(A),$$
$$f \in \mathcal{E}_1(B), n \in B \tag{21}$$

and replace the $y_{e,f} q_{n,f}$ terms in (10) and (12) with $\beta_{e,f,n}$

For fixed network design decisions x and y, $\mathcal{A}(x,y)$ is defined by:

$$\sum_{e\in\mathcal{E}} z_e \leq U \tag{22}$$

$$z_f \leq \sum_{e\in\mathcal{E}(A)} y_{e,f} z_e, \qquad \forall (A,B) \in \mathcal{Z}, \forall f \in \mathcal{E}(B) \tag{23}$$

$$\delta_r = \sum_{e\in\mathcal{E}(\mathcal{S})} x_{e,r} z_e, \qquad \forall r \in \mathcal{R} \tag{24}$$

$$v_k \leq (1-\delta_r), \qquad \forall k \in \mathcal{K}, r \in \mathcal{R}_k \tag{25}$$

$$v_k \geq \sum_{r\in\mathcal{R}_k} (1-\delta_r) - |\mathcal{R}_k| + 1, \qquad \forall k \in \mathcal{K} \tag{26}$$

$$w_g \leq (1-\delta_r), \qquad \forall g \in \mathcal{G}, r \in \mathcal{R}_g \tag{27}$$

$$w_g \geq \sum_{r\in\mathcal{R}_g} (1-\delta_r) - |\mathcal{R}_g| + 1, \qquad \forall g \in \mathcal{G} \tag{28}$$

$$u_d \leq (1-\delta_r), \qquad \forall d \in \mathcal{L}, r \in \mathcal{R}_d \tag{29}$$

$$u_d \geq \sum_{r\in\mathcal{R}_d} (1-\delta_r) - |\mathcal{R}_d| + 1, \qquad \forall d \in \mathcal{L} \tag{30}$$

Constraint (22) enforces the attacker's budget. Constraint (23) requires that the attacker can only access enclaves which are controlled by already-accessed enclaves. That is, with respect to the graphs in Fig. 1, the attack must begin at a balancing authority enclave, and no child can be accessed if its parent was not. Constraint (24) enforces that a relay is compromised when the enclave that controls it has been accessed by the attacker. Constraints (23) and (24) contain bilinear terms that can be linearized using the same technique that linearizes (10) and (12).

Constraints (25)-(26) enforce that the line k is opened if and only if some relay that controls it is compromised. Constraints (27)-(30) do the same for generators and loads.

Finally, given the attack decision, we define $\mathcal{O}(u, v, w)$, the operator's DCOPF feasible region:

$$\sum_{k \in \{k' | d(k')=s\}} f_k - \sum_{k \in \{k' | o(k')=s\}} f_k$$
$$+ \sum_{g \in \mathcal{G}_s} p_g = \sum_{d \in \mathcal{L}_s} (D_d - l_d) \qquad \forall s \in \mathcal{S} \qquad (31)$$

$$f_k = B_k v_k(\theta_{o(k)} - \theta_{d(k)} - \Theta_k) \qquad \forall k \in \mathcal{K} \qquad (32)$$

$$-\overline{F}_k \le f_k \le \overline{F}_k \qquad \forall k \in \mathcal{K} \qquad (33)$$

$$0 \le p_g \le w_g \overline{P}_g \qquad \forall g \in \mathcal{G} \qquad (34)$$

$$(1 - u_d)D_d \le l_d \le D_d \qquad \forall d \in \mathcal{L} \qquad (35)$$

$$-\pi \le \theta_s \le \pi \qquad \forall s \in \mathcal{S} \qquad (36)$$

Equation (31) enforces flow balance at each substation, while the line power flow approximation is enforced in (32). We linearize (32) by replacing it with:

$$f_k \le B_k(\theta_{o(k)} - \theta_{d(k)} - \Theta_k) + \qquad \forall k \in \mathcal{K} \qquad (37)$$
$$B_k(2\pi + \Theta_k)(1 - v_k)$$

$$f_k \ge B_k(\theta_{o(k)} - \theta_{d(k)} - \Theta_k) \qquad \forall k \in \mathcal{K} \qquad (38)$$
$$- B_k(2\pi + \Theta_k)(1 - v_k)$$

$$-\overline{F}_k v_k \le f_k \le \overline{F}_k v_k \qquad \forall k \in \mathcal{K} \qquad (39)$$

Constraints (33) and (34) enforce thermal limits and maximum generation capacity, forcing flow or generation to 0 if the component is compromised. Note that we can drop (33) after adding (39). Also note that in (34), we assume the minimum generation capacity for all generators is 0. While this is not accurate, it is necessary in order to make (31)-(36) feasible for all possible values of u, v, and w, the importance of which we describe in Section IV. In (35), we upper bound load shed by the total load and require that we shed all compromised loads. Bounds on the phase angles are enforced in (36).

IV. Solution Methodology

Because the grid operator's problem is linear and has an objective opposite the attacker's objective, we can take its dual and linearize the bilinearities that appear in the dual objective function in order to reformulate the inner two levels as one mixed integer linear program (MILP). For details on this procedure, see [22]. Note that the linearization requires upper bounds on the dual variables. The authors of [29] show that verifying the correctness of these dual bounds is as hard as solving the original bilevel problem. Thus,

we adopt the heuristic commonly used in prior literature, which is to assume that the duals are bounded above by the power capacity of the component corresponding to their index [21]. After this reformulation, where we take the dual of the linear program which minimizes $\sum_{d \in \mathcal{L}} l_d$ over the set $\{(\theta, f, p, l) : (31), (34)\text{-}(36), (37)\text{-}(39)\}$ and combine the resulting maximization problem with the attacker problem, we have a bilevel problem with integer leader and mixed integer follower:

$$\min_{(x,y,q,t,\beta) \in \mathcal{D}} \max_{\boldsymbol{X} \in \mathcal{A}_D(x,y)} L, \qquad (40)$$

where

$$\boldsymbol{X} = (\delta, z, u, v, w, f, p, d, \theta, \mu, \xi^+, \xi^-,$$
$$\lambda^+, \lambda^-, \gamma, \alpha^+, \alpha^-, \beta^+, \beta^-)$$

and $\mathcal{A}_D(x, y)$ is defined by:

$$(22)\text{-}(30)$$

$$\sum_{s \in \mathcal{S}} \left[\mu_s \sum_{d \in \mathcal{L}_s} D_d - \pi(\beta_s^+ + \beta_s^-) \right]$$
$$+ \sum_{d \in \mathcal{L}} D_d[\alpha_d^+(1 - u_d) - \alpha_d^-]$$
$$- \sum_{k \in \mathcal{K}} [B_k(\Theta_k + (2\pi + \Theta_k)(1 - v_k))(\xi_k^+ + \xi_k^-)$$
$$+ \overline{F}_k v_k(\lambda_k^+ + \lambda_k^-)] - \sum_{g \in \mathcal{G}} \overline{P}_g \gamma_g w_g = L \qquad (41)$$

$$\mu_{d(k)} - \mu_{o(k)} + \xi_k^+ - \xi_k^- + \lambda_k^+ - \lambda_k^- = 0, \quad \forall k \in \mathcal{K} \quad (42)$$

$$\mu_{s(g)} - \gamma_g \le 0, \quad \forall g \in \mathcal{G} \qquad (43)$$

$$\mu_{s(d)} + \alpha_d^+ - \alpha_d^- \le 1, \quad \forall d \in \mathcal{L} \qquad (44)$$

$$\beta_s^+ - \beta_s^- + \sum_{k \in \{k' | o(k')=s\}} B_k(\xi_k^- - \xi_k^+)$$
$$+ \sum_{k \in \{k' | d(k')=s\}} B_k(\xi_k^+ - \xi_k^-) = 0, \qquad \forall s \in \mathcal{S} \quad (45)$$

$$\xi^+, \xi^-, \lambda^+, \lambda^-, \gamma, \alpha^+, \alpha^-, \beta^+, \beta^- \ge 0 \qquad (46)$$

where μ is the dual of the balance constraint (31), ξ^- and ξ^+ are the duals of (37) and (38) respectively, λ^+ and λ^- are the duals of (33), γ is the dual of the upper bound in (34), α^+ and α^- are the duals of (35), and β^+ and β^- are the duals of (36). Constraint (41) calculates the dual objective value. Constraints (42), (43), (44), and (45) are the dual constraints corresponding to f, p, l, and θ, respectively. Constraint (41) includes many bilinear terms, all the product of a non-negative dual variable and a binary variable. These can be reformulated with the addition of auxiliary continuous variables. For example, to reformulate $\alpha_d^+ u_d$, we introduce $\overline{\alpha}_d^+$ in place of the product in constraint (41) and add the constraints

$$\alpha_d^+ - D_d(1 - u_d) \le \overline{\alpha}_d^+ \le \alpha_d^+ + D_d(1 - u_d) \qquad (47)$$

$$0 \le \overline{\alpha}_d^+ \le D_d u_d, \qquad (48)$$

where we are assuming that α_d^+ is bounded above by the total demand at load d.

Note that for every outer-problem solution $(x, y, q, t, \beta) \in \mathcal{D}$, problem (40) has a finite objective value. That is,

$\mathcal{A}_D(x, y) \neq \emptyset$. This is because, regardless of the network segmentation and the attack, it is always possible to prevent infeasibility by shedding load or turning off generators. We refer to this property as *relatively complete recourse.* Note that it is because the generator dispatch lower bounds are 0 that we have this property.

Problem (40) is a mixed-integer bilevel model, which we solve using the bilevel branch and bound algorithm from [30]. In [31], [32], and [30], callbacks within the IBM CPLEX solver [33] are used to create a bilevel branch-and-cut solver. The authors make their software available for academic use at [34]. To easily generate the required MPS file for the bilevel branch and bound solver, we formulated our model with Pyomo [35] and [36].

A. An Equivalent Solution Methodology

We found that, when solving (40) with the solver from [34], the solver could not close the gap within a time limit of 72 hours. We suspect this may be a symptom of numerical issues in the cuts. To overcome this, we solved a slightly modified problem, in which the variable β and the constraints it appears in are moved to the inner problem of (40). Formally, the formulation is as follows:

$$\min_{(x,y,q,t) \in \mathcal{D}'} \max_{(\boldsymbol{X}, \beta) \in \mathcal{A}'_D(x,y,q)} L \qquad (49)$$

where $\mathcal{D}' = \{(x, y, q, t) : (2)\text{-}(9), (11), (13)\text{-}(18)\}$ and

$$\mathcal{A}'_D(x, y, q, t) = \{(\boldsymbol{X}, \beta) : (19)\text{-}(30), (41)\text{-}(46)\}.$$

Note that, in (49), we have lost the relatively complete recourse property: Fixing the outer level problem's choice of y and q, it is possible to choose β in order to make one or more of (19)-(21) infeasible. However, because the solver from [34] uses the high-point relaxation of (49) to get a lower bound, it assumes the outer problem will not choose to make the inner problem infeasible. That is, we are actually solving the problem:

$$\min_{(x,y,q,t) \in \mathcal{P}} \max_{(\boldsymbol{X}, \beta) \in \mathcal{A}'_D(x,y,q)} L, \qquad (50)$$

where

$$\mathcal{P} = \{(x, y, q, t) \in \mathcal{D}' : \exists (\boldsymbol{X}, \beta) \text{ s.t. } (\boldsymbol{X}, \beta) \in \mathcal{A}'_D(x, y, q)\}$$

Essentially, there is an implicit constraint added to the outer problem that its solution is in the projection of the inner problem's feasible region onto the outer problem variables. We therefore show that (50) is equivalent to (40).

We first show that the optimal objective value of (50) is a lower bound to that of (40). Note first that $\mathcal{P}$ equals the projection of $\mathcal{D}$ onto the (x, y, q, t) variables. This follows from the relatively complete recourse property of (40). We can rewrite $\mathcal{P}$ as

$$\mathcal{P} = \{(x, y, q, t) \in \mathcal{D}' : \exists \beta \text{ s.t. } (19)\text{-}(21)\}$$

Next, note that the projection of $\mathcal{A}'_D(x, y, q)$ onto the space of the $\boldsymbol{X}$ variables is a restriction of $\mathcal{A}_D(x, y)$ (since constraints (19)-(21) have been moved to the inner problem). Since, in projected space, we have an equivalent outer problem region and a restricted inner problem region, we have a lower bound.

Next, we show that this bound is achieved. To see this, consider a solution $(x, y, q, t, \beta, \boldsymbol{X})$ to (50). Then, by our argument above, $(x, y, q, t, \beta) \in \mathcal{D}$, and by the definition of $\mathcal{A}'_D(x, y, q)$, $\boldsymbol{X} \in \mathcal{A}_D(x, y)$. Thus, this solution is feasible in (40). This completes the argument: When using the branch and bound solver from [34] to solve (49), we are solving (40) exactly.

V. RESULTS

We present a case study on the 9-bus WSCC and 30-bus IEEE systems [37] using the methodology given in Section IV-A. Note that, as we have formulated the network segmentation problem, the network designer is forced to use their entire budget, regardless of if every enclave in the budget is necessary to reduce that load shed the attacker can cause. While it would be possible to reformulate the problem to avoid this, it would be necessary to use the network designer's objective to motivate them to be frugal with their budget. This is no longer an interdiction problem, and is in essence a multi-objective variant. For simplicity, we leave this as future work, and instead suggest solving the proposed model for a variety of network designer budgets in order to find a minimal one which achieves a satisfactory reduction in load shed. We will show an example of this process for the 9-bus system.
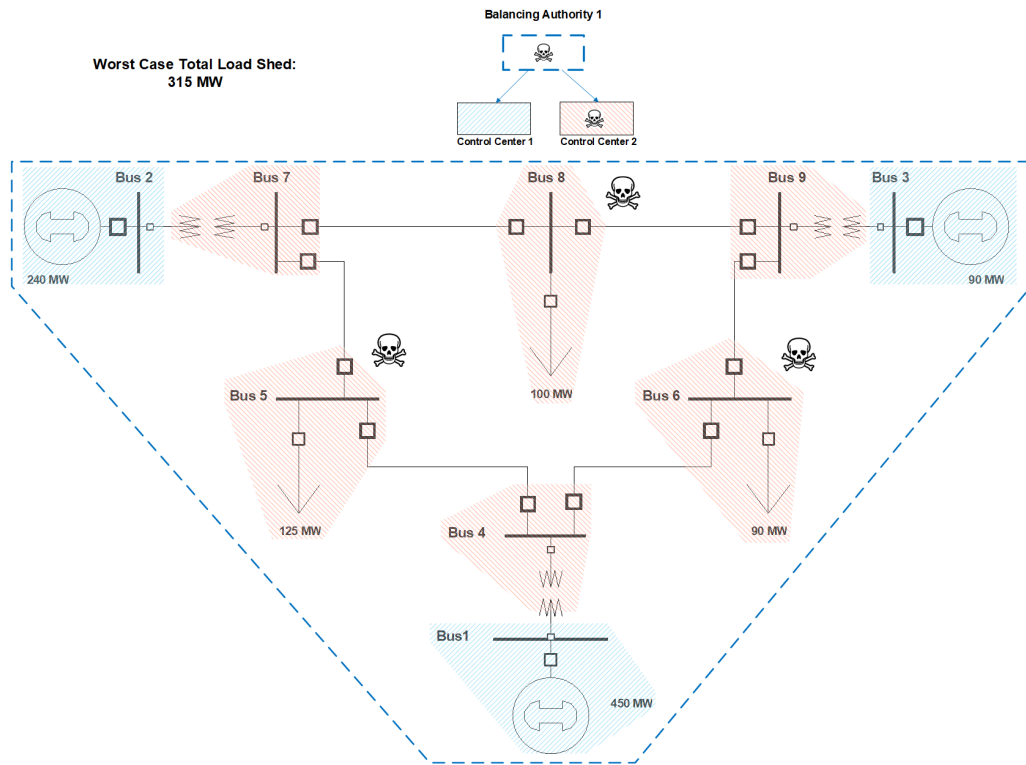
A. Hardware and Software Specification

We ran our experiments on a Linux server with two Intel® Xeon® Silver 4210 CPU @ 2.20GHz processors and 256 GB RAM. For the CPLEX-based solver from [34], we used CPLEX 12.9.0 with two threads.
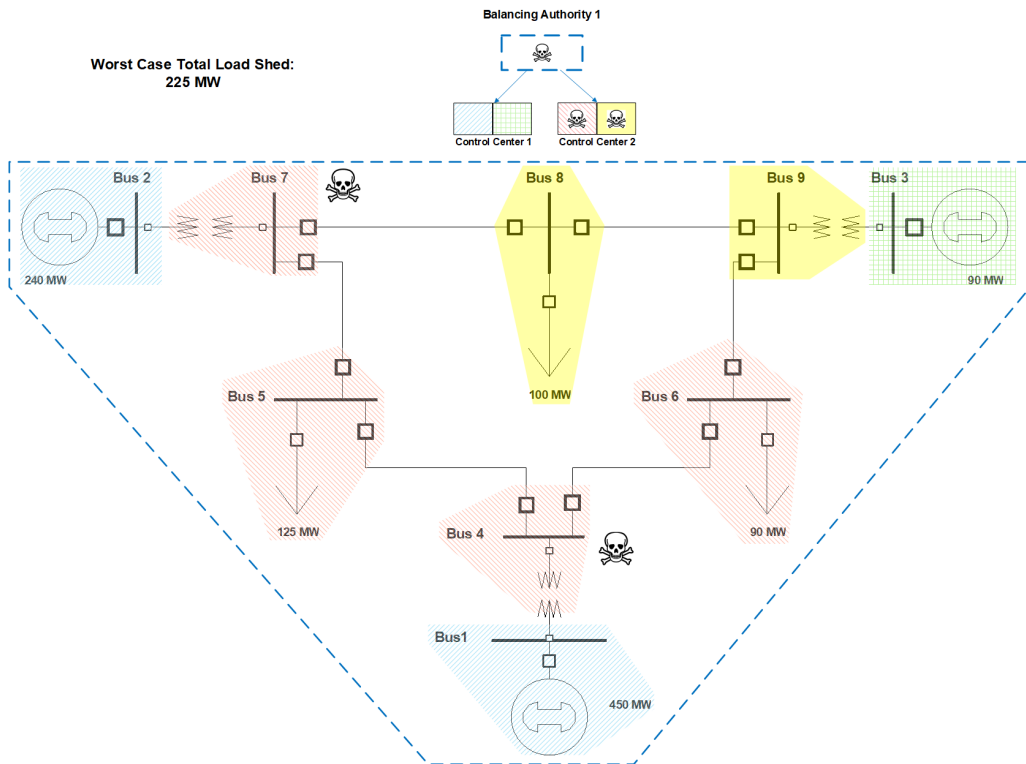
B. 9-bus WSCC System

To create intuitive results that are easy to validate, we first use the 3-generator 9-bus WSCC test system [37]. We use a simple communication network consisting of a single balancing authority and two control centers. The three substations with generators are assigned to one control center and the remaining six substations are assigned to the other control center. There is exactly one security enclave at each substation and each grid component at that substation is controlled through a relay hosted by its substation's security enclave. See Fig. 2a for a depiction of this cyber-physical system before segmentation. Note that this starting topology does not correspond exactly to the communication network we showed in Fig. 1. In Fig. 1, we reduced the number of relays in the system for the sake of readability.

We do two case studies on this network: The first defends against an attacker with a budget of 5 and the second defends against an attacker with a budget of 8. We begin with the case where the budget is 5. Before network segmentation, the attacker can cause a complete blackout by pivoting from the balancing authority, through the second control center, to Substations 5, 6, and 8. These substations contain all of the system's load of 315 MW. See Fig. 2a for a visualization of this attack. Note that the attacker has multiple optimal solutions: For example, they could alternatively pivot from the

(a) The 9-bus system before segmentation. The worst-case attack with budget 5 is marked with a skull-and-crossbones on each enclave and component which is compromised. The attacker infiltrates the Balancing Authority enclave, the Control Center 2 enclave, and the three substation enclaves whose relays control the loads and lines at Buses 5, 6, and 8.

(b) The 9-bus system after segmentation, where the network designer was allowed a budget of two control center enclaves and the attacker had a budget of five enclaves. After segmentation, the attacker spends three units of budgets to infiltrate the balancing authority enclave and both the Control Center 2 enclaves. They only have two units of budget remaining, so they infiltrate buses 4 and 7, which islands the two largest generators, resulting in 225 MW of load shed. Note that the attack on the second yellow Control Center 2 enclave does not cause any load shed; it only allows the attacker to use the full attack budget.

Fig. 2. The 9-bus system before and after segmentation in figures (a) and (b) respectively, when the attacker has a budget of 5 enclaves.

TABLE I
PARTIAL PARAMTER SWEEP OF NETWORK DESIGNER BUDGETS FOR THE
9-BUS SYSTEM WITH ATTACK BUDGET 5.

No. of Substation Enclaves	No. of Control Center Enclaves	No. of Balancing Authority Enclaves	Worst-Case Attack Load Shed (MW)
0	1	0	315
1	0	0	315
1	1	1	315
0	2	0	225
0	0	2	190

balancing authority to Control Center 1, and then compromise Substations 1, 2, and 3. This allows them to shut down all the generators, also resulting in 315 MW of load shed.

In Table I, we present results on a partial parameter sweep of network designer budgets for this problem. We see that two enclaves at either the balancing authority or control center level are necessary to prevent any load shed. Furthermore, two enclaves at the balancing authority level prevent 35 MW more load shed than two enclaves at the control center level.

Figure 2b displays the optimal network segmentation for a budget of two control center enclaves. The model segments both the control center networks into two enclaves, and reallocates the substations so that it is no longer possible to compromise all three generators or all three loads. After network segmentation, one worst-case attack infiltrates both enclaves in Control Center 2. The attacker can then only access two substations. They choose Substations 4 and 7 to cause a load shed of 225 MW, reducing the load shed possible in the unsegmented network by 28.6%. Note that the attack on the solid yellow right enclave of Control Center 2 is unnecessary. However, due to the segmentation, there is nothing productive for the attacker to do with this last unit of budget.

In the second case study, we decide an optimal network segmentation to protect against an attacker with a budget of 8 enclaves. Clearly, since the attacker with budget 5 was able to shed all the load (see Figure 2a), in the unsegmented case with budget 8, the attacker sheds 315 MW of load also. Using a process similar to what we showed in Table I, we select a network designer budget of two balancing authority enclaves and four control center enclaves. The resulting segmentation is visualized in Fig. 3. The balancing authority is divided into three enclaves, each of which communicates with two control center enclaves. No load bus or generator bus is on the same control center enclave, so, to remove all the loads or all the generators, the attacker would need to spend three units of budget at the substation level and three at the control center level. However, this leaves them only two at the balancing authority level. Because each balancing authority enclave is mapped to one Control Center 1 enclave and one Control Center 2 enclave, they would need nine or more units of budget to attack enough enclaves to cause a complete blackout. Thus, they cannot shed all the load, and 90 MW (the smallest load) is served. Again, network segmentation reduced the load shed by 28.6%. Note in the attack shown in Fig. 3, the attacker wastes two units of budget: Similarly to the budget-5 case, the attacks on the blue balancing authority enclave and the purple control center enclaves are unproductive, but they cannot be
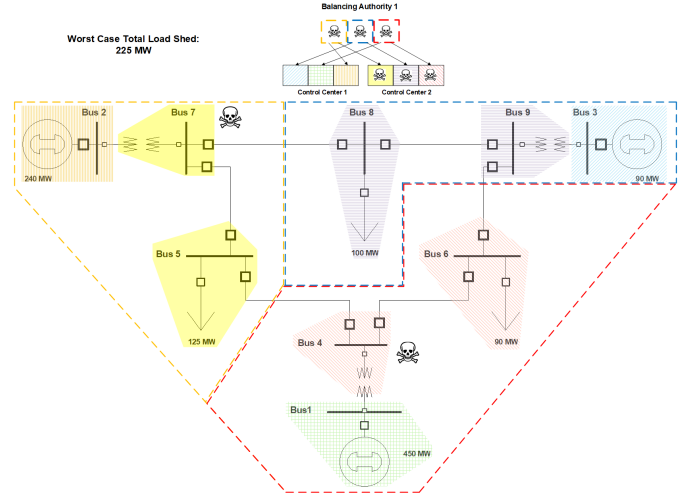

Fig. 3. The 9-bus system after segmentation with a network designer budget of 2 balancing authority enclaves and 4 control center enclaves and an attack budget of 8 enclaves.

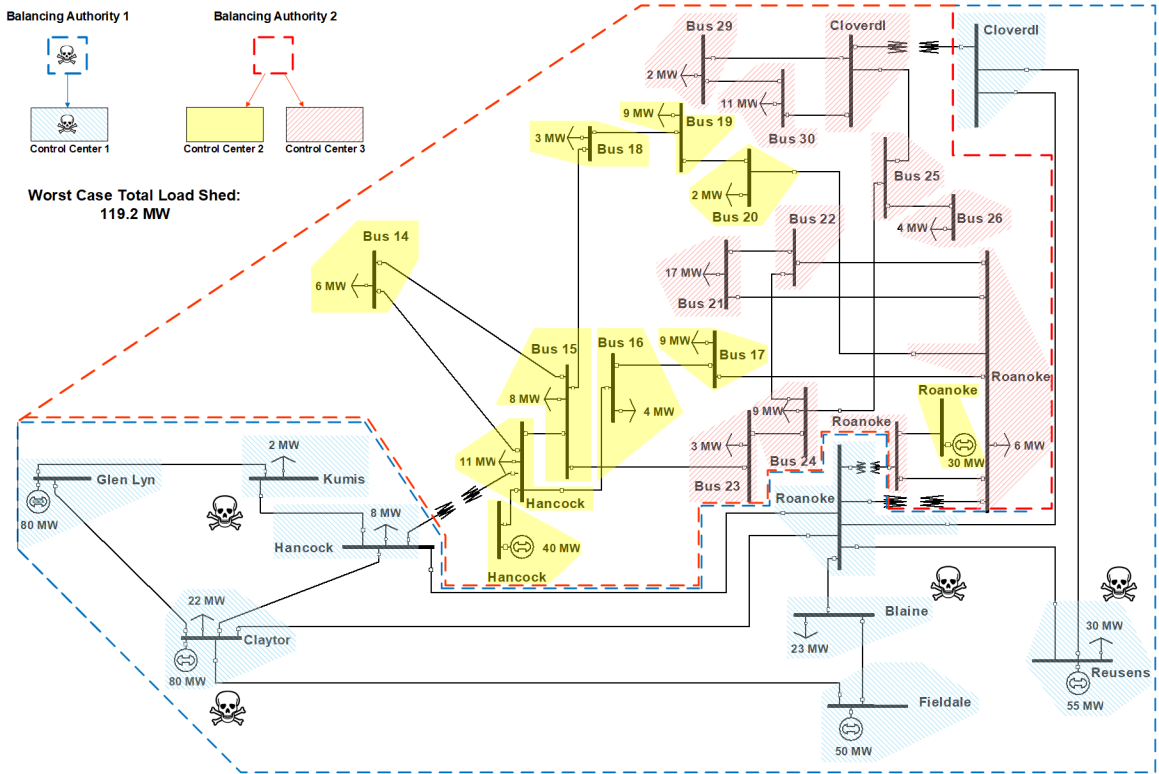reassigned so that the attack increases the load shed.

Both of these examples illustrate that optimal network segmentation isolates the most critical infrastructure in a way that drastically reduces the damage caused by the worst-case attack.
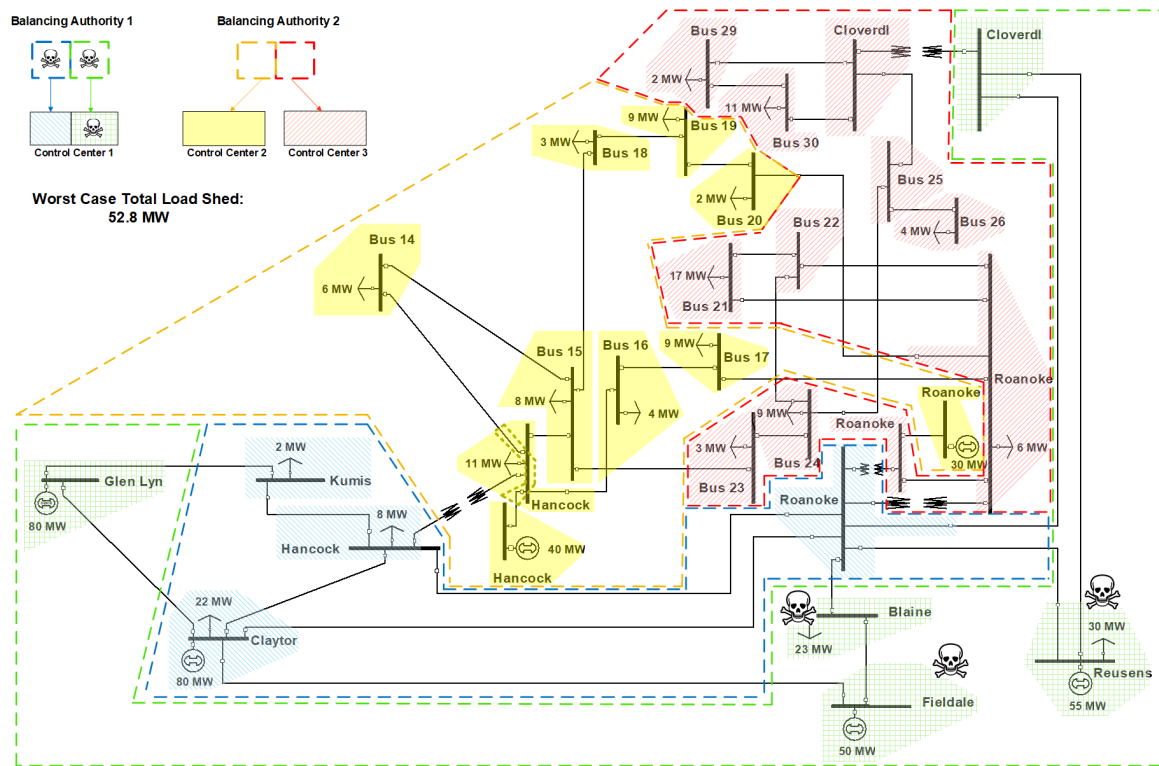
C. 30-bus System

The 30-bus IEEE system provides a larger example with a more complex topology. For this system, we use the voltage levels to create a fictitious communication network. All high-voltage grid components communicate with a single control center and balancing authority. The lower-voltage grid components are split between two control centers. Both these control centers communicate with a second balancing authority. This network is depicted in Fig. 4a. Balancing Authority 1 and Control Center 1 control the high-voltage grid components, and the remaining two control centers and Balancing Authority 2 control the low-voltage components.

In this case study, we consider an attack budget of six enclaves. We again search the results for a range of network designer budgets, trying to find a minimal one with satisfactory reduction of load shed. We present results where we allow two extra balancing authority enclaves, one extra control center enclave, and one extra substation enclave. Our model segments both balancing authorities into two security enclaves each and divides Control Center 1 into two enclaves. Because substations controlled by Control Center 1 have the largest loads and generation, the model assigns substations to the two enclaves in Control Center 1 so that the largest loads communicate with only one control center enclave. Control Centers 2 and 3 are each assigned to their own enclave in Balancing Authority 2. Finally, the substation enclave is used at the Hancock bus in Control Center 2 with 11 MW of load: The relays controlling the lines to Bus 14 and to the Hancock bus with the 40 MW generator are on their own enclave.

Before network segmentation, the attacker can cause a load shed of 119.2 MW, or 63.1% of the total load, by pivoting

(a) The 30-bus system before segmentation: The network is controlled by two balancing authorities and 3 control centers, each with only one security enclave. The worst-case attack with budget 6 compromises Balancing Authority 1, Control Center 1, and four substation enclaves controlling buses Blaine, Claytor, Kumis, and Reusens.

(b) The 30-bus system after segmentation: Balancing Authority 1 and Control Center 1 have two enclaves each, meaning the attacker would require 4 units of budget to gain access to the same substations. In lieu of this, they attack only the green control center enclave, compromising Blaine, Fieldale, and Reusens. While it is possible for them to also attack Glen Lyn or Cloverdl, the resulting load shed is not higher.

Fig. 4. The 30-bus system before and after segmentation in figures (a) and (b) respectively, with the worst-case attack with budget 6 marked in both versions.

from the first balancing authority, through its control center, to substations Kumis, Claytor, Blaine, and Reusens. See Fig. 4a. After network segmentation, the worst-case attack uses two units of budget to give the attacker access to the substations controlled by the green control center enclave. This allows them to attack substations Blaine, Fieldale, and Reusens, causing a total load shed of 52.8 MW, or 27.9% of the total demand. This means that network segmentation reduced the worst-case load shed by 55.7%. See Fig 4b for a depiction. Note that the attack on the blue balancing authority enclave is wasted. In fact, the attack on Fieldale is also not necessary: All of the load shed stems from directly compromising the loads at Blaine and Reusens, and the generators at Glen Lyn and Fieldale are not necessary for serving the remaining load.

This case study illustrates an intuitive network segmentation strategy. The largest loads are connected to substations Blaine and Reusens, so an attacker will expend all resources to infiltrate those substations. However, this network segmentation strategy ensures that the attacker will not be able to shed any further load. While it may seem intuitive to put Blaine and Reusens on separate control center enclaves, note that the attacker still has enough budget to access both, expending four units of budget to attack both Balancing Authority 1 enclaves and both Control Center 1 enclaves, then using the remaining two units to attack Blaine and Reusens. The segmentation therefore limits the attacker's ability to compromise additional enclaves.

Finally, we note that the segmentation of Balancing Authority 2 and of the Hancock substation are indeed necessary in this example. Without segmenting Balancing Authority 2, there is an attack on 6 enclaves which sheds 104.7 MW of load by compromising the Balancing Authority 2 enclave, the enclaves in Control Centers 1 and 2, the Roanoke substation with the 6 MW load, and the Cloverdl substation under Control Center 3. By segmenting Balancing Authority 2, this attack is prevented since it takes an additional unit of budget to gain access to both control centers. With only this segmentation and not the extra substation enclave at Hancock, there still exists an attack which shed 53 MW of load, targeting the yellow Balancing Authority 2 enclave, Control Center 2, the Hancock substation with the 11 MW load, Bus 15, Bus 17, and Bus 20. Including the new Hancock substation enclave, the worst attack on the low-voltage side of the network is reduced to 51.7 MW, compromising the red Balancing Authority 2 enclave, Control Center 3, the Roanoke substation with the 6 MW load, Bus 23, and the accessible Cloverdl substation.

This case study shows that even moderate budgets of additional enclaves can greatly decrease the possible load shed achievable by a cyber attacker, in this case reducing it by more than half.

VI. Future Work

The network segmentation model we develop can be improved and extended in several ways. More details from the communication network presented in [28] can be included to result in a more realistic network. The attacker model can be made more realistic by adding more complex movement throughout the network, using enclave access fees based on real data, and adding context to the attacker's budget. For example, an IT administrator could perform analysis to determine the maximum number of enclaves that could likely be comprised before the IDS systems on the network would generate an alert, and use that as the basis for making segmentation decisions. We use the simplest grid operator model available. Future work can focus on using a higher-fidelity power flow model and allowing grid operator transmission switching. In addition, solution strategies that would allow solving the model on realistically-sized networks would be of interest.

VII. Conclusion

Segmentation is a cyber defense strategy that has been proposed for improving network security. However, specific strategies for optimally implementing these recommendations, while accounting for the underlying system being protected, are lacking. To address this, we develop a trilevel cyber-physical power-transmission-system network-segmentation model that is the first of its kind. It models an IT administrator who must decide how to preemptively segment the power transmission system communication network in preparation for an attacker with perfect information of the whole system. We use an attacker model where a malicious actor must move from balancing authorities through control centers to substations in order to disable grid components through relays. The attacker anticipates how the grid operator will use a DCOPF to redispatch the generators after the attack is executed. We reformulate our model to a mixed-integer bilevel problem and use bilevel branch-and-cut to solve it. Finally, we present network segmentation results on the 9-bus WSCC and 30-bus IEEE test systems, demonstrating the benefit of optimal network segmentation in terms of reducing the worst-case load shed the attacker can achieve.

VIII. Acknowledgments

References

[1] Cybersecurity and Infrastcture Security Agency, "Cyber-attack against ukrainian critical infrastructure ICS Alert - IR-ALERT-H-16-056-01," 2016, https://us-cert.cisa.gov/ics/alerts/IR-ALERT-H-16-056-01, Last accessed on 2020-08-15.

[2] Daniel Coates, "Statement for the record worldwide threat assessment of the US intelligence community," 2019, congressional Testimonies 2019.

[3] Government Accountability Office, "Critical infrastructrion protection-actions needed to address significant cybersecurity risks facing the electric grid," 2019, GAO-19-332.

[4] National Security Agency, Central Security Service, "Segregate networks and functions, Fact Sheet MIT-012FS-2013," 2016, https://apps.nsa.gov/iaarchive/library/ia-guidance/security-tips/segregate-networks-and-functions.cfm, Last accessed on 2020-08-15.

[5] US Department of Homeland Security, Industrial Control Systems Cyber Emergency Response Team, "Recommended practice: Improving industrial control system cybersecurity with defense-in-depth strategies," 2016.

[6] V. Y. Pillitteri and T. L. Brewer, "Guidelines for smart grid cybersecurity," 2014, NIST Interagency/Internal Report (NISTIR) - 7628 Rev 1.

[7] K. Stouffer, V. Pillitteri, M. Suzanne Lightman, and A. A. Hahn, "Guide to industrial control systems (ICS) security," 2016, NIST Special Publication 800-82 Revision 2.

[8] A. C. S. Centre, "Implementing network segmentation and segregation," 2020, australian Signals Directorate.

[9] N. Wagner, C. . Şahin, M. Winterrose, J. Riordan, J. Pena, D. Hanson, and W. W. Streilein, "Towards automated cyber decision support: A case study on network segmentation for security," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2016, pp. 1–10.

[10] S. B. Jeong, Y. W. Choi, and S. Kim, "An effective placement of detection systems for distributed attack detection in large scale networks," in *Information Security Applications*, C. H. Lim and M. Yung, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 204–210.

[11] Y. Zhang, L. Wang, and W. Sun, "Trust system design optimization in smart grid network infrastructure," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 184–195, 2013.

[12] J. M. C. Gonzalez, K. M. Hopkinson, G. H. Greve, M. D. Compton, J. Wilhelm, S. H. Kurkowski, and R. W. Thomas, "Optimization of trust system placement for power grid security and compartmentalization," *IEEE Tran. Power Syst.*, vol. 26, no. 2, pp. 550–563, 2011.

[13] E. Hemberg, J. R. Zipkin, R. W. Skowyra, N. Wagner, and U.-M. O'Reilly, "Adversarial co-evolution of attack and defense in a segmented computer network environment," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1648–1655.

[14] M. M. Hasan and H. T. Mouftah, "Optimal trust system placement in smart grid SCADA networks," *IEEE Access*, vol. 4, pp. 2907–2919, 2016.

[15] ——, "A study of resource-constrained cyber security planning for smart grid networks," in *2016 IEEE Electrical Power and Energy Conference (EPEC)*, 2016, pp. 1–6.

[16] ——, "Latency-aware segmentation and trust system placement in smart grid SCADA networks," in *2016 IEEE 21st International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, 2016, pp. 37–42.

[17] ——, "Optimization of trust node assignment for securing routes in smart grid SCADA networks," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1505–1513, 2019.

[18] B. Genge and C. Siaterlis, "An experimental study on the impact of network segmentation to the resilience of physical processes," in *NETWORKING 2012*, R. Bestak, L. Kencl, L. E. Li, J. Widmer, and H. Yin, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 121–134.

[19] ——, "Physical process resilience-aware network design for SCADA systems," *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 142 – 157, 2014, 40th-year commemorative issue.

[20] R. Arief, N. Khakzad, and W. Pieters, "Mitigating cyberattack related domino effects in process plants via ICS segmentation," *Journal of Information Security and Applications*, vol. 51, p. 102450, 2020.

[21] J. Salmeron, K. Wood, and R. Baldick, "Worst-case interdiction analysis of large-scale electric power grids," *IEEE Tran. Power Syst.*, vol. 24, no. 1, pp. 96–104, 2009.

[22] A. L. Motto, J. M. Arroyo, and F. D. Galiana, "A mixed-integer LP procedure for the analysis of electric grid security under disruptive threat," *IEEE Tran. Power Syst.*, vol. 20, no. 3, pp. 1357–1365, 2005.

[23] A. Castillo, B. Arguello, G. Cruz, and L. Swiler, "Cyber-physical emulation and optimization of worst-case cyber attacks on the power grid," in *2019 Resilience Week (RWS)*, vol. 1, 2019, pp. 14–18.

[24] A. Delgadillo, J. M. Arroyo, and N. Alguacil, "Analysis of electric grid interdiction with line switching," *IEEE Tran. Power Syst.*, vol. 25, no. 2, pp. 633–641, 2009.

[25] X. Wu and A. J. Conejo, "An efficient tri-level optimization model for electric grid defense planning," *IEEE Tran. Power Syst.*, vol. 32, no. 4, pp. 2984–2994, 2017.

[26] N. Alguacil, A. Delgadillo, and J. M. Arroyo, "A trilevel programming approach for electric grid defense planning," *Computers & Operations Research*, vol. 41, pp. 282–290, 2014.

[27] W. Yuan, L. Zhao, and B. Zeng, "Optimal power grid protection through a defender–attacker–defender model," *Reliability Engineering & System Safety*, vol. 121, pp. 83 – 89, 2014.

[28] N. Gaudet, A. Sahu, A. E. Goulart, E. Rogers, and K. Davis, "Firewall configuration and path analysis for smartgrid networks," in *2020 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)*. IEEE, 2020, pp. 1–6.

[29] T. Kleinert, M. Labbé, F. Plein, and M. Schmidt, "Technical note—there's no free lunch: On the hardness of choosing a correct big-M in bilevel optimization," *Operations Research*, vol. 68, no. 6, pp. 1716–1721, 2020.

[30] M. Fischetti, I. Ljubić, M. Monaci, and M. Sinnl, "On the use of intersection cuts for bilevel optimization," *Mathematical Programming*, vol. 172, no. 1-2, pp. 77–103, 2018.

[31] ——, "Intersection cuts for Bilevel optimization," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9682, pp. 77–88, 2016.

[32] ——, "A New General-Purpose Algorithm for Mixed-Integer Bilevel Linear Programs," *Operations Research*, vol. 65, no. 6, pp. 1615–1637, 2017.

[33] [Online]. Available: https://www.ibm.com/analytics/cplex-optimizer

[34] [Online]. Available: https://msinnl.github.io/pages/bilevel.html

[35] M. L. Bynum, G. A. Hackebeil, W. E. Hart, C. D. Laird, B. L. Nicholson, J. D. Siirola, J.-P. Watson, and D. L. Woodruff, *Pyomo–optimization modeling in python*, 3rd ed. Springer International Publishing, 2021, vol. 67.

[36] W. E. Hart, J.-P. Watson, and D. L. Woodruff, "Pyomo: modeling and solving mathematical programs in python," *Mathematical Programming Computation*, vol. 3, no. 3, pp. 219–260, 2011.

[37] [Online]. Available: https://icseg.iti.illinois.edu/power-cases