Budget-Constrained Reinforcement of SCADA for Cascade Mitigation

Vajiheh Farhadi, Sai Gopal Vennelaganti, Ting He, Nilanjan Ray Chaudhuri, and Thomas La Porta
Pennsylvania State University, University Park, PA, USA. E-mail: {vuf8,suv66,tzh58,nuc88,tfl12}@psu.edu.

***Abstract*—We study the impact of coupling between the communication and the power networks as it affects a SCADA-based preventive control system. Today power grids use power lines to carry control information between components in the grid and a control center using power line carrier communication (PLCC). Thus a failure in the power grid will cause a failure in the control network and may reduce the capability of preventive control that in turn increases the risk of cascading failures. We pose the problem of allocating a limited number of non-PLCC communication links (e.g., microwave links) that are immune to failures in the power grid to maximize our controllability over the grid under power system failures, so as to maximize the total demand served at the end of cascade. By formulating the problem as a nonlinear integer programming problem, we establish its hardness and identify a generic heuristic that can find an approximate solution within controllable time. We further develop a domain-specific heuristic that utilizes both graph-theoretic and power system information to achieve similar performance as the generic heuristic at a much lower computational complexity. Our evaluations based on a $2,383$-bus Polish system demonstrate that only a few non-PLCC links, when placed correctly, can substantially improve the robustness of the grid as measured by the total demand served at the end of cascade.**

***Index Terms*—DC-QSS, cascading failure, blackouts, preventive control, SCADA, PLCC**

I. Introduction

Cascading failures in power grids have led to wide-spread socio-economic disruptions [1]. Therefore, it is of high importance to improve our understanding of this phenomenon and develop defense mechanisms.

In this paper, we focus on the interaction between the power grid and the control network during cascading failures. The control network monitors elements of the power grid, and issues command to some of these elements when failures occur to mitigate a cascade. Any degradation that limits the ability of the control network to either monitor or control elements in the power grid will increase the risk of a larger cascade of failures. Ideally, the control network, consisting of sensors/actuators, communication links, routers, and a controller, should be deployed independently of the power grid, with battery backup for all its components. However, this is an expensive solution, especially for deploying dedicated communication links to connect the controller to all the elements in the power grid.

For this reason, utility companies have used power lines to carry control information using the technique of *power line carrier communication (PLCC)* [2]–[6]. PLCC enables communication between substations using low [7], [8], medium [8], [9], or high voltage [10], [11] power lines. PLCC links are much less expensive than non-PLCC (i.e., dedicated) links [12]–[14], but will fail when power lines fail if no backup communication medium is used, thereby degrading connectivity of the control network. In this paper, we examine the problem of reinforcing a fully PLCC-based communication network with the addition of a limited number of reliable more expensive dedicated non-PLCC communication links.

Traditionally, PLCC was used for one-way communication to monitor, control, and tele-protect the power grid [15]. In the 1980s and early 1990s, after studies regarding the implementation of the Supervisory Control and Data Acquisition (SCADA) in both Europe and the United States, bi-directional communication by PLCC was invented [12]. The strong growth of PLCC data services and its economic benefit has made it a promising component of information infrastructure [6], [16]–[18]. The mixture of PLCC and other technologies extends the utilization of PLCC for power generation and control [19]. While it may seem that the advances of non-PLCC communications to fiber-optic technology [20], [21] eliminate the need for PLCC, PLCC links remain an essential component of modern power grids [7], such as the Siemens PowerLink PLCC system [13] and General Electric T&D Power Utilities [22], due to their cost savings.

In this work, we show that a carefully designed communication network containing a few strategically placed non-PLCC links together with PLCC links can provide almost the same protection against cascading failure as a dedicated communication network with 100% non-PLCC links at a fraction of the cost.

A. Summary of Contributions

In this paper, we consider cascading failures in a coupled system of a power grid with a SCADA-based communication network that is geographically co-located with the power grid, through which a Control Center (CC) collects data from sensors and dispatches preventive control commands to generators and loads. Our contributions are:

1) We propose to combine the cost efficiency of PLCC links and the reliability of non-PLCC links in designing the communication network [23], posed as an optimization of maximizing the total power demand served after cascade by

Financial support from NSF Grant Award ECCS 1836827 is gratefully acknowledged.

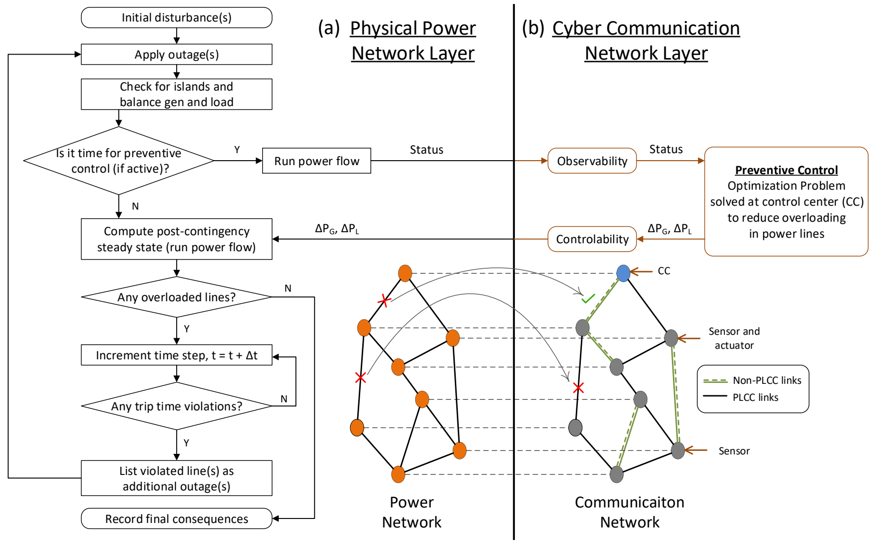


Figure 1. Flow chart highlighting the modeling of the coupled cascading failure in power and communication networks.

selecting a limited number of non-PLCC links, leaving the rest as PLCC.

2) As the demand served after cascade is not an explicit function of the decision variables, we propose a proxy objective function capturing the controllability of nodes in the grid, weighted by their importance in the system topology and the contribution of generation/load. We formulate the underlying optimization as a nonlinear programming (NLP) problem, which is proved to be NP-hard. We then apply a generic heuristic and develop a domain-specific heuristic that explicitly maximizes the controllability of the generators and the loads most likely to be needed in preventive control.

3) We evaluate the proposed algorithms on a $2,383$-bus Polish power system. The results show that (i) the algorithms designed to maximize the proxy objective function can effectively increase the demand served after cascade, (ii) a small number of properly placed non-PLCC links together with (unreliable) PLCC links can achieve almost the same performance as a perfectly reliable communication network, and (iii) while performing similarly as a properly configured generic heuristic, the proposed domain-specific heuristic is significantly faster.

Roadmap: Section II provides background information about cascading failures and preventive control. Section III formulates our problem of budget-constrained design of communication links and presents our proposed algorithms. Section IV evaluates our algorithms against benchmarks. Finally, Section V concludes the paper.

II. Background and Motivation

A. Modeling Coupled Cascading Failure

Broadly, there are three types of cascading failure models reported in the literature - DC-quasi-steady-state (QSS), AC-QSS, and dynamic models. Both AC-QSS models [24], [25], which are based on the AC power flow, and dynamic models [26], [27], have challenges in simulating cascade propagation, especially for large-scale networks.

DC-QSS models [1], [28]–[30], which neglect resistive losses and assume a uniform voltage profile, are computationally efficient and thus suitable for statistical analysis of large-scale systems. Therefore we use DC-QSS models of cascades in this paper.

Although there is a significant body of literature in the area of cascading failures in power grid [1], [24]–[28], very few papers [29], [30] consider the coupled cascading failure of the power grid and the associated communication network. Both [29] and [30] use the DC-QSS model for the power grid.

In [29], the authors propose a model of a smart power grid coupled to a communication network and show that increased power-communication coupling decreases vulnerability, in contrast to the percolation model. However, the robustness can be enhanced and the failure propagation constrained by interconnecting networks that have different modes of failure between the communication and power networks.

In [30], real-world scenarios are considered in which the locations of failures in a coupled power-communication network might be unknown or only partially known. While this work has detailed models for the power gird, the models of the communication network are oversimplified which is the focus of our work.

Figure 1 illustrates the model of cascading failure in a coupled power and communication network. After the impact of the initial outages, the power grid may be segmented into islands. The generation and load in each island are balanced. In order to achieve the balance, either generation is curtailed or the load is reduced uniformly across all generation or load nodes, respectively. If an island does not have any generators, then a complete blackout is assumed in the island. The overloaded branches, which result from the updated line flows, are tripped according to the tripping delays of overcurrent relays, meaning that for a particular line to trip, it must remain overloaded for the duration of its trip time. Once more lines are tripped, the power grid may be segmented into further islands, where load and generation need to be re-balanced, and the entire process is repeated until there are no potential line trips, i.e, no overloaded lines in the network. At this stage, the cascade propagation comes to an end.

As described earlier, we consider a geographically collocated SCADA-based communication network that connects the CC to sensors and actuators. This is a common assumption in work on power-communication overlays [29]. The connections between the CC and these sensors/actuators can be affected by failures in the power grid if PLCC communication is used under the assumption that no backup communication is deployed. Therefore, the cascading failure propagates in a coupled manner through both power and communication networks.

B. Modeling Preventive Control

A preventive controller at the CC is introduced, as illustrated in Fig. 1, which is assumed to have updated information of all the elements in the power grid, e.g., line flows, breaker status, and power output/consumption of generators and load centers. Based on this information, the controller tries to alleviate line overloading by issuing control commands. Both the sensing information and the control commands are communicated via a communication network.

The objective of the preventive controller is to stop cascade propagation by reducing the overloading in lines, which is defined as $\mathbf{L}_{over}$. This objective is achieved by solving the following optimization problem [29]:

$$\begin{aligned} \min_{\Delta P_G, \Delta P_L} \; & -\mathbf{1}^\intercal \mathbf{\Delta P}_L + \lambda^\intercal \mathbf{L}_{over} \\ \text{s.t.} & \mathbf{\Delta P}_G - \mathbf{\Delta P}_L = \mathbf{B}\mathbf{\Delta}\theta, \qquad \Delta\theta_i = 0, \forall \in \Omega_{ref} \\ & \Delta L_{ij} = \frac{\Delta\theta_i - \Delta\theta_j}{x_{ij}}, \qquad \forall\, i,j \in \mathcal{M} \\ & |\mathbf{L}_{\mathcal{M}} + \Delta \mathbf{L}| \le \mathbf{L}_{max} + \mathbf{L}_{over},\ \mathbf{L}_{over} \ge \mathbf{0} \\ & -(\mathbf{P}_G)_{\mathcal{M}} \le (\mathbf{\Delta P}_G)_{\mathcal{M}} \le \mathbf{0} \\ & -(\mathbf{P}_L)_{\mathcal{M}} \le (\mathbf{\Delta P}_L)_{\mathcal{M}} \le \mathbf{0} \\ & (\mathbf{\Delta P}_G)_{\overline{\mathcal{M}}} = \mathbf{0}, (\mathbf{\Delta P}_L)_{\overline{\mathcal{M}}} = \mathbf{0} \end{aligned} \tag{1}$$

Here, the elements of vectors $\mathbf{P}_G$ and $\mathbf{P}_L$ represent the generation and load power at each bus, respectively, and Δ represents the change in these quantities. The objective function (1) minimizes the total amount of load shedding ($-\mathbf{1}^\intercal \mathbf{\Delta P}_L$) and the weighted sum of all overloads such that no further controls on generators or loads can improve the objective. λ is the uniform weight vector. Similarly, a change in the phase angle at the ith bus is defined as $\Delta\theta_i$. The matrix $\mathbf{B}$ is the admittance matrix of the network. The variables $\mathbf{L}$ and $\mathbf{L}_{max}$ represent the actual power flow and its allowable maximum value, respectively. The subscript $\mathcal{M}$ implies that the corresponding quantities belong to the measurable set, and the subscript $\overline{\mathcal{M}}$ implies the opposite. To get more details of the optimization problem (1), please refer to [29]. As presented in Fig. 1, line status and branch flows $\mathbf{L}_{\mathcal{M}}$ are taken as inputs to the optimization problem, and the solution gives load shedding and generation reduction values as outputs.

A communication network with 100% non-PLCC links is cascade-free. In this case, $\overline{\mathcal{M}} = \emptyset$ and the preventive control produces the best possible performance. However, in the presence of PLCC-based links the coupled cascade propagation is affected in a complex manner because power line failure impacts the preventive control. The larger the number of such link failures, the larger the set $\overline{\mathcal{M}}$. This implies that more sensors become unobservable and more actuators become uncontrollable. This limits the effectiveness of cascade prevention, which in turn exacerbates the loss of controllability in a closed-loop fashion.

In practice, preventive control algorithms will run at regular intervals, and there will be a delay in line tripping. To evaluate these effects, we run preventive control optimization following the first outage for every 30 s, which is the shortest possible time for SCADA-based preventive control systems, and take the tripping delay into account.

C. Motivating Example

To understand the potential value of preventive control in mitigating cascading failures, we evaluate the performance of a network with preventive control. Considering the Polish network during winter $1999-2000$ [31] as a test system, which includes $2,383$ buses and $2,896$ branches, we randomly fail 5% of the buses in the power grid and calculate the served power after cascade propagation in scenarios (i) having a network of 100%-PLCC links versus (ii) no communication network. Results in Table I show superior performance in presence of a communication network.

We use a specific example in Fig. 2 to illustrate why a suitably designed communication network can help preventive control to mitigate cascade. The red links in Fig. (2a) are the failures initiating the cascade, which prompt the overloading and tripping of the black links in Figs. (2b-c), respectively. In Fig. (2b), there is no communication network. In Fig. (2c), the CC communicates with and controls the two red generators.

Without control (Fig. (2b)), the cascade causes the loss of 134 lines with the residual power 13653.30 MW. The active power injections of the left (g_1) and the right (g_2) red generators after the cascade are 96.90 MW and 643.0 MW, respectively. With control of only the two nodes g_1 and g_2 (Fig. (2c)), e.g., by connecting them to the CC through non-PLCC links, the cascade only causes the loss of 10 lines with the residual power 24460.20 MW. This is achieved by changing the active power injections at generators g_1 and g_2 to 56.34MW and 645.13MW, respectively, which prevents line overload while satisfying energy conservation.

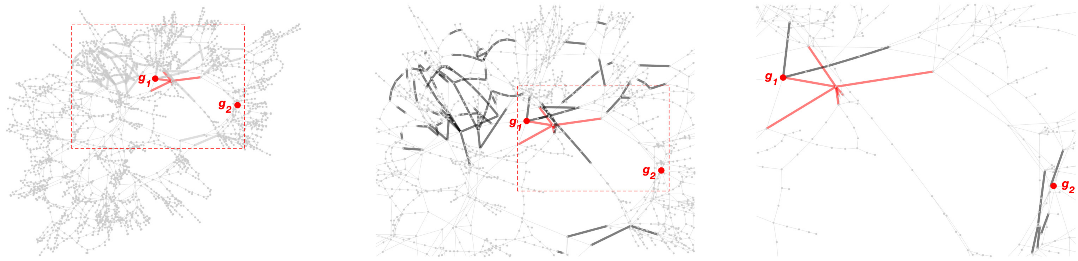
This example not only demonstrates the benefit of preventive control, but also shows that much of the benefit can be achieved by controlling a small number of critical nodes.

III. Budget-Constrained Reinforcement of Communication Network

We develop algorithms to design the communication network as a mix of PLCC and non-PLCC links under a budget

Table I
THE EFFECT OF PREVENTIVE CONTROL ON CASCADE IN (I) 100%-PLCC LINKS AND (II) NO COMMUNICATION NETWORK.

Scenario	no. of cascade steps	no. of tripped lines	residual power
(i)	5	308	22472.14 MW
(ii)	20	504	7790.8 MW



(a) (b) (c)

Figure 2. The effect of preventive control on cascade propagation. (a) initial failures in the Polish grid (red links), (b) cascading failures (black links) without control (as a zoom-in of the rectangle in (a)), (c) cascading failures (black links) with control of the two red nodes (as a zoom-in of the rectangle in (b)).

constraint to facilitate the mitigation of cascading failures through preventive control. To rule out other impacts, we assume that all the communication nodes (including sensors, actuators, and relays) have battery backups and are hence immune to the cascade and cannot be affected by failures in the power grid.

A. Problem Formulation

We formulate the problem as an optimization of non-PLCC link placement under a budget constraint. As PLCC links are much less expensive than non-PLCC links [12]–[14], we start with a baseline where all the communication links are PLCC, and then turn a selected subset of links into non-PLCC links to improve the robustness against cascades. To capture resource limitations, we impose a budget B on the number of non-PLCC links, but our solution can be easily extended to incorporate heterogeneous costs of non-PLCC links.

Ideally, we want to maximize the total demand served when the cascade stops. This objective function, however, faces the challenge that it is not an explicit function of the placement of non-PLCC links. To address this challenge, we propose to use a proxy objective function as follows.

We model the power grid as an undirected graph $\mathcal{G}(\mathcal{N}, \mathcal{L})$ with no self-loops or multiple edges. Let P^i, $\forall i \in \mathcal{N}$ denote the set of all possible paths between the CC and node i. Modeling each path $m \in P^i$ as a set of links it traverses, the total reliability of the network [32], measured by the expected number of nodes connected to the CC after (initial) failure, is defined as (assuming nodes do not fail):

$$\sum_{i\in\mathcal{N}}[1-\prod_{m\in P^i}(1-\prod_{l\in m}\rho_l)], \tag{2}$$

where ρ_l denotes the reliability (i.e., complement of failure probability) of link l. Let $R_l \in \{0,1\}$ indicate whether link l is a non-PLCC link, and p_1/p_0 denote the reliability of non-PLCC/PLCC link, respectively (assuming $p_1 > p_0$). Then $\rho_l = R_l p_1 + (1-R_l)p_0$. We formulate the optimization of non-PLCC links as follows:

$$\max \sum_{i\in\mathcal{N}} \gamma_i (1-\prod_{m\in P^i}(1-\prod_{l\in m\cap\mathcal{L}}(R_l p_1+(1-R_l)p_0)) \tag{3a}$$

$$\text{s.t.} \sum_{l\in\mathcal{L}} R_l \le B, \tag{3b}$$

$$R_l \in \{0,1\}, \qquad \forall l \in \mathcal{L}. \tag{3c}$$

In words, (3) aims at selecting up to B non-PLCC links to maximize the controllability after (initial) failure, measured by the expected total weight of all the nodes remaining connected to the CC. Intuitively, the more nodes the CC is connected to, the better it can observe/control the grid, and hence the better it can mitigate the cascading of failures. However, not all the nodes are equally important, and hence we use the weight $\gamma_i \ge 0$ to reflect the importance of observing and controlling node i.

Remark: Intuitively, γ_i should reflect both the topological importance (e.g., centrality) and the service importance (e.g., power injection) of node i. We find that defining γ_i as "the betweenness centrality (BC) of node i" $\times$ "the real power injected at node i" yields the best performance (see Fig. 4), where BC of a node is the frequency that it appears on the shortest paths between all pair of nodes in the graph [33].

B. Complexity Analysis

We prove that (3) is NP-hard by a reduction from the Steiner tree problem.

Theorem 1. *Problem (3) is NP-hard.*

Proof. The (graph) Steiner tree problem takes as input an undirected graph $\mathcal{G}_0$ with non-negative edge weights and a subset of vertices called *terminals*, and seeks to find a tree that is a subgraph of $\mathcal{G}_0$ with minimum weight to connect all the terminals. The decision problem associated with the Steiner tree problem is "whether exists a solution to the Steiner tree problem with integer edge weights, such that the total weight of the Steiner tree is no greater than a given natural number k". This problem is one of Karp's 21 NP-complete problems [34].

Using the above problem, we will show that the decision problem of "whether there is a feasible solution to (3) that connects the CC to all the non-zero-weight nodes by non-PLCC links" is NP-hard. The NP-hardness of this decision problem implies the NP-hardness of a special case of the optimization problem (3) for $p_1 = 1 > p_0$, as otherwise we can solve the optimization problem and compare the achieved objective value with $\sum_{i\in\mathcal{N}}\gamma_i$. As each node i with $\gamma_i > 0$ contributes $\le \gamma_i$ to the objective value, with "=" achieved only if it is connected to the CC via a perfectly reliable path (consisting of only non-PLCC links), it is easy to see that the optimal objective value equals $\sum_{i\in\mathcal{N}}\gamma_i$ if and only if there is a feasible solution to (3), under which every non-zero-weight node is connected to the CC by a path of non-PLCC links, thus solving the decision problem.

Construction: We construct $\mathcal{G}$ according to $\mathcal{G}_0$, except that each edge in $\mathcal{G}_0$ of weight e (a positive integer) is represented by a tandem of e links in $\mathcal{G}$. The budget B is set to k. One of the terminals is set as the CC, and the other terminals as nodes with non-zero weights. The rest nodes have zero weight.

Claim: The answer to the Steiner tree decision problem is "yes" if and only if the answer to the decision version of the above-constructed instance of (3) is "yes".

Proof of the claim: If the decision problem associated with the Steiner tree gives "yes", i.e., there is a tree $\mathcal{T}_0$ in $\mathcal{G}_0$ with total weight no more than k that connects all the terminals, then the corresponding tree $\mathcal{T}$ in $\mathcal{G}$ will connect the CC with all the non-zero-weight nodes while covering no more than B links. Hence, setting $R_l = 1$ for links in $\mathcal{T}$ and $R_l = 0$ otherwise will connect the CC to all the non-zero-weight nodes by non-PLCC links within budget B. Conversely, if the CC can reach all the non-zero-weight nodes through no more than B non-PLCC links, then $\mathcal{G}$ must contain a tree $\mathcal{T}$ of no more than B links that contains the CC and all the nodes with non-zero weights. Then, the corresponding tree $\mathcal{T}_0$ in $\mathcal{G}_0$ must be a Steiner tree with a total weight no greater than k that connects all the terminals. □

C. Algorithm Design

The NP-hardness of the optimal solution to (3) motivates our search for efficient heuristics.

1) Generic heuristic: We first apply a generic heuristic algorithm that is a *genetic algorithm* [35], that belongs to a non-deterministic class of algorithms that provide suboptimal solutions in controllable time. It works by modifying a population of possible solutions repeatedly such that the population evolves toward an optimal solution. At each step, the genetic algorithm arbitrarily picks solutions from the current population to be parents and produces the children for the next step. Due to the possibly exponential complexity in enumerating all possible paths, we limit P^i for each $i \in \mathcal{N}$ to a set of up to N simple paths [36] from the CC to node i with length $\leq L$, where L and N are design parameters that will be tuned later (see Fig. 5).

2) Domain-specific heuristic: As shown later (Fig. 5), the generic heuristic needs to search a large solution space to achieve reasonable performance, which is computationally expensive. This motivates us to develop the following alternative that uses domain-specific insights. As the ultimate objective is to ensure that the re-dispatch-based preventive control can effectively mitigate the propagation of cascading failure, we will focus solely on maximizing the controllability of the generators and the loads whose re-dispatching is most likely to be needed during preventive control – given the budget constraint. Although this method appears more complicated, we will show that it is much more computationally efficient than the generic heuristic in Section III-C1 while achieving almost the same performance in terms of served loads.

Algorithm 1: Candidate Non-PLCC Link and Candidate Node Selection

input : Power grid data, location of Control Center (CC)
output : Set of candidate non-PLCC links $\mathcal{L}_n$ and candidate control nodes $\mathcal{N}_n$, and subgraphs $\mathcal{V}$

Step 1: Identify subgraphs based on the degree of load and generation nodes and a pre-determined hop count h. G_i/L_i: total generation/load in the ith subgraph;
Step 2: Within the ith subgraph – If $G_i > L_i (L_i > G_i)$, then choose all generator nodes (load nodes) with degree > 1 as candidate control nodes. If $G_i = L_i$, choose all load nodes (degree > 1) in this subgraph as candidate control nodes. Set of such nodes are $\mathcal{N}_n$;
Step 3: Solve the problem of finding the tree of shortest paths on graphs to connect the CC to all candidate control nodes from Step 1. Set of such links are $\mathcal{L}_n$;
return $\mathcal{L}_n$, $\mathcal{N}_n$, $\mathcal{V}$;

Algorithm 2: Non-PLCC Link Selection

input : Budget B, $\mathcal{L}_n$, $\mathcal{N}_n$, $\mathcal{V}$, location of CC, power grid data
output : Set of non-PLCC links $\mathcal{L}_c$

Calculate subgraph weights: Define k_i: no. of nodes of the ith subgraph $\in \mathcal{V}$ connected to the remaining portion of the grid. Calculate weights $w_i = min(L_i, G_i)/k_i$;
Sort subgraphs: Order the subgraphs in descending order of their weights;
$\mathcal{L}_c \leftarrow \emptyset, i \leftarrow 1, \mathcal{G}_c \leftarrow \emptyset$;
while $|\mathcal{L}_c| \leq B$ **do**
  while $i \leq |\mathcal{V}|$ **do**
    if *the ith subgraph does not contain the CC* **then**
      Connect one Gateway node $g \in \mathcal{N}_n$ in the subgraph to one Gateway node $g \in \mathcal{N}_n$ in a subgraph containing the CC using minimum number of non-PLCC links $l \in \mathcal{L}_n$;
      $i \leftarrow i + 1$, $\mathcal{L}_c \leftarrow \mathcal{L}_c \cup \{l\}, \mathcal{G}_c \leftarrow g$;
    else
      $i \leftarrow i + 1$
  return i;
  $i \leftarrow 1$;
  while $i \leq |\mathcal{V}|$ **do**
    if *the ith subgraph contains the CC* **then**
      Calculate betweenness centrality (BC) of nodes within the area w.r.t. the gateway nodes $\in \mathcal{G}_c$ and the CC. Assign non-PLCC links $l \in \mathcal{L}_n$ incident on the node with highest BC;
      $i \leftarrow i + 1$, $\mathcal{L}_c \leftarrow \mathcal{L}_c \cup \{l\}, \mathcal{G}_c \leftarrow g$;
    else
      $i \leftarrow i + 1$
  return i;
  $i \leftarrow 1$;
  while $i \leq |\mathcal{V}|$ **do**
    if *the ith subgraph does not contain the CC* **then**
      Calculate BC of nodes within the area w.r.t. the gateway nodes $\in \mathcal{G}_c$ and the candidate control nodes. Assign non-PLCC links $l \in \mathcal{L}_n$ incident on the node with highest BC;
      $i \leftarrow i + 1$, $\mathcal{L}_c \leftarrow \mathcal{L}_c \cup \{l\}, \mathcal{G}_c \leftarrow g$;
    else
      $i \leftarrow i + 1$
  return i;
return $\mathcal{L}_c$;

To this end, we propose a two-part algorithm, shown in Algorithms 1 and 2. In the first algorithm, we follow the steps described below to identify subgraphs and candidate non-PLCC links. The second algorithm selects the non-PLCC links.

Identify subgraphs (Step 1): Ideally, extensive planning

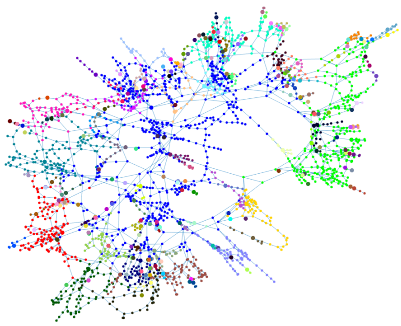


Figure 3. Subgraphs of the system for the domain-specific heuristic, $h = 10$.

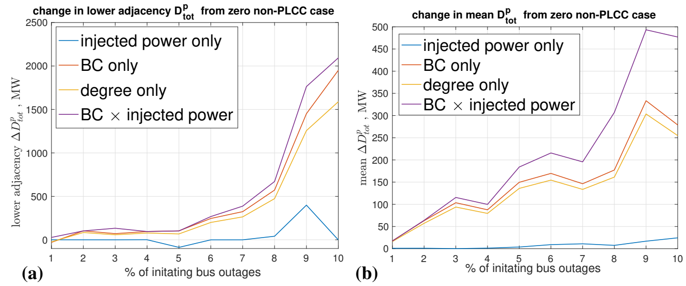


Figure 4. Performance evaluation in terms of change in D^p_{tot} with respect to 100% PLCC case for different node weights (non-PLCC links are placed by the generic heuristic under $L = 200$, $N = 50$, and $B = 17$).

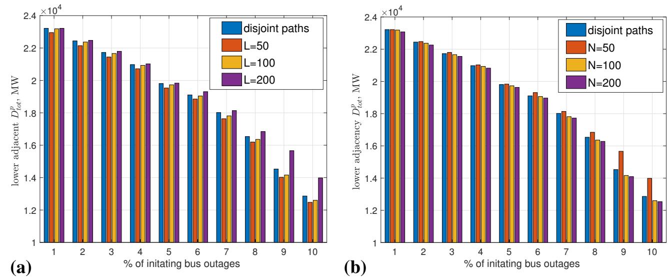


Figure 5. Performance evaluation for the generic heuristic under different design parameters: (a) $N = 50$ & varying L, (b) $L = 200$ & varying N ($B = 17$ in both cases).

studies should give us the information regarding subgraphs that are likely to form during cascading failures. In absence of this information for the system under consideration, we propose a simple heuristic that is motivated by the intuition to partition the power grid by clustering nodes around each high-degree generator/load into a subgraph, which is likely to form an island during cascade. Then, we try to find candidate non-PLCC links connecting the CC to the actuators in each subgraph. Alternative graph clustering algorithms can be found in [37] – exploration of such algorithms are beyond the scope of this paper.

To that end, we first sort the generation and load buses according to their degree, then we build subgraphs around them consecutively. We consider the subgraphs formed by all nodes that can be reached in a pre-determined number of hops from the root node (the radius of the neighborhood). We ignore the nodes that are already included by an existing subgraph. This procedure is repeated until we cover all the nodes. In this context, we use words ‘subgraph’ and ‘area’ (a power grid domain-specific terminology) interchangeably.

Figure 3 shows the graph of the Polish system with subgraphs marked on it with different colors, where roots are specified by increasing the sizes of their markers. Nodes within the same subgraph are at most 10 hops away from the root.

Find candidate non-PLCC links and candidate control nodes (Steps 2, 3): This is a two-step process. Keeping in mind that generation and load nodes need to be connected to the CC for receiving preventive control commands (see, Section II-B), we aim to find a subset of such nodes that are candidates for this. In Step 2, we find the subgraphs where $G_i \neq L_i$ – these areas affect line flows in external areas. Since we always reduce generation and load, we choose generator nodes as candidates when $G_i > L_i$ and vice-versa. The reason behind neglecting nodes with degree 1 is that outage of a line connected to this node will result in disconnection of the generator/load, thereby rendering the non-PLCC communication useless. In Step 3, we solve the problem of finding the tree of shortest paths on graphs to connect the CC to all candidate control nodes.

Algorithm 2 aims to select the non-PLCC links from the candidate control nodes based on graph-theoretic and domain-driven metrics. The steps are described next:

Calculate subgraph weights (Line 1) and sort subgraphs (Line 2): The proposed algorithm prioritizes subgraphs based on weights w_i, which depend on the value of the generation or load, whichever is smaller ($min(L_i, G_i)$), and the connectivity between this subgraph and the rest of the grid (k_i). The logical argument is as follows: **(a)** higher $w_i = min(L_i, G_i)/k_i$ implies that the subgraph has a relatively lower connectivity with other subgraphs and thus a higher probability to form an island during cascade, while **(b)** a subgraph with a higher w_i also contains relatively more generation and load, and is thus more critical to control in the case of islanding to prevent further cascade propagation within the subgraph. We define a gateway node as one of the candidate control nodes connecting the particular subgraph to another gateway node in another subgraph. We use the term ’gateway’ because this particular subgraph joins rest of the system through this node via a non-PLCC link. Thus, we connect a gateway node of a subgraph with high w_i but not the CC, to a gateway node in a subgraph containing the CC. If there are multiple options to do this, then the shortest path is chosen.

Establish non-PLCC links (Lines $3-28$): For establishing non-PLCC links within each area, we use BC measures considering the CC, gateway nodes and candidate control nodes as shown in Algorithm 2.

IV. Performance Evaluation

We evaluate the proposed solutions on the Polish network during winter $1999-2000$ peak condition from Matpower [31]. This system includes $2,383$ buses, $2,896$ branches, and 327 generators. We consider initial bus outages varying from $1\% - 10\%$ of all the buses. For each failure scenario, 500 random sets of node (i.e., bus) outages have been considered. The CC is situated on bus 7 of the Polish network, which is one of the highest degree nodes. We set p_0 and p_1, the reliability of PLCC and non-PLCC links in (3) to 0.99 and 0.9999, respectively [38].

Impact of node weight definition: We start by comparing the performance under different definitions of node weight γ_i.

We compare four different definitions of weights: (i) power injection, (ii) degree, (iii) BC, and (iv) power injection × BC. Under each definition, we solve (3) by the generic heuristic under $L = 200$, $N = 50$ and $B = 17$. Figs. 4(a-b) show the lower adjacency, the smallest data point that is not an outlier in the plot which is 1st quartile - 1.5×inter-quartile range, and the mean of the change in total post-contingency demand served (ΔD^p_{tot}) with respect to the case of 100% PLCC links, where $D^p_{tot} = 24558$ MW before any failure. The results show that representing a node weight by "the BC of the node" × "the real power injected at the node" performs the best, as it considers both the topological and the service importance of the node. We also calculated the reliability of the network as defined in (2) under the non-PLCC links placed under each of these objectives for different definitions of weights. The results follow the same trend as Figs. 4(a-b), but are not shown in the paper. This leads us to conclude that a higher reliable network leads to higher post-contingency served demand after cascade. In the sequel, we will use this definition of weight for the generic heuristic.

Configuration of generic heuristic: We then compare the performance of the generic heuristic under different limits on the length (L) and the number (N) of paths between each node and the CC. For comparison, we also evaluate a variation of this heuristic that requires the paths in each P^i ($i \in \mathcal{N}$) to be disjoint with each other (without limitation on path length). Fig. 5 shows the performance in mitigating cascades in terms of the lower adjacency of the improvement in post-contingency demand served (similar comparison has been observed for the mean and other percentiles). The results indicate that when using the generic heuristic, allowing overlap between paths, and choosing a smaller N and a larger L improve the efficacy in mitigating cascades. The reason lies in the limitation of genetic algorithm where the solution quality may deteriorate with the increase of problem size.

Configuration of domain-specific heuristic: Next, we evaluate the performance of the domain-specific heuristic under different settings of the design parameter h, the maximum hop count in generating subgraphs in Algorithm 1. Fig. 6 shows the impact of the h, evaluated by the lower adjacency and the mean of ΔD^p_{tot} as in Fig. 4. The results suggest that the parameter h significantly affects the performance of the resulting control system as it determines the number and sizes of the generated subgraphs. As it can be seen, the performance starts deteriorating for h that are larger than 10. Choosing a smaller h, doesn't provide a higher percentage of served demands as it excessively increases the number of candidate control nodes/candidate Non-PLCC links. For this particular power grid, we find that setting $h = 10$ leads to better performance, which will be the setting used for this heuristic in the sequel.

Under the above setting, the Polish system is divided into 227 subgraphs by Algorithm 1. Table II summarizes different statistics of these subgraphs, including their sizes (defined as the number of nodes within the particular subgraph) and the variables k_i and w_i used by Algorithm 2 in placing non-PLCC links. Recall that k_i is the number of nodes connecting the ith subgraph to the remaining portion of the grid, and $w_i = min(L_i, G_i)/k_i$ where G_i/L_i is the total generation/load in the ith subgraph.

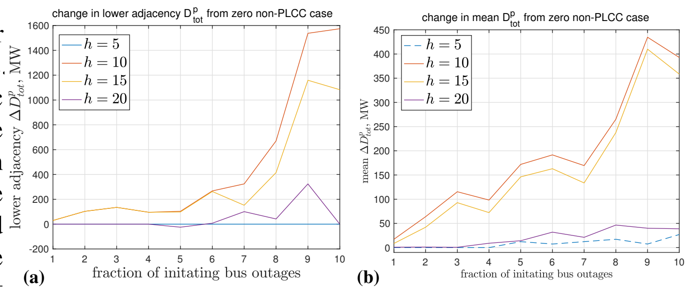


Figure 6. Performance evaluation for the domain-specific heuristic under different design parameters ($B = 17$).

Table II
STATISTICAL MEASURES OF VARIABLES IN THE DOMAIN-SPECIFIC HEURISTIC UNDER $B = 17$

Variable	min	max	mean	median	STD
size	1	596	10.5	1	46.2
k_i	1	139	2.47	1	9.89
w_i	0	83.16	3.816	0	12.26

We have also evaluated a variation of the domain-specific heuristic, where the subgraph weight is defined as $u_i = \alpha w_i + (1-\alpha)v_i$, where $v_i := |G_i - L_i| \times k_i$ and the variable $\alpha = (0, 1]$ determines the relative importance of weights w_i and v_i. The intuitive motivation is that a subgraph with higher v_i has a higher probability to remain connected to other subgraphs and a higher impact due to higher $|G_i - L_i|$ (excess generation or load). The results, which are omitted due to space limitation, indicate that setting $\alpha = 1$ leads to better performance.

Overall comparison: Finally, Fig. 7 compares the performance of all the algorithms in terms of the change in total post-contingency demand served ΔD^p_{tot} with respect to the case of 100% PLCC links ('$B = 0$'). In addition to the proposed heuristics and the baseline of randomly selecting non-PLCC links ('Random'), we consider an intuitive benchmark of allocating non-PLCC links by ranking the nodes in descending order of their BC and selecting all the links incident to each node as non-PLCC links until the budget runs out ('BC'). Different statistical measures of ΔD^p_{tot} are shown.

Results show that while both of the proposed heuristics outperform random selection, the BC-based benchmark performs slightly better than the proposed heuristics when the budget is sufficiently high (e.g., $B = 174$ as in Figs. 7 (a-d)). However, for a lower budget (e.g., $B = 17$ as in Figs. 7 (e-h)), the BC-based benchmark degrades severely, and the proposed heuristics perform much better. This demonstrates the importance of jointly considering the topology information and the power system information in selecting non-PLCC links, instead of solely based on the topology information as in the BC-based method. Moreover, with suitably tuned parameters ($L = 200$, $N = 50$), the generic heuristic can slightly outperform the domain-specific heuristic ($h = 10$), justifying the choice of objective function in (3).

However, we note that the domain-specific heuristic runs significantly faster than the generic heuristic (with an average

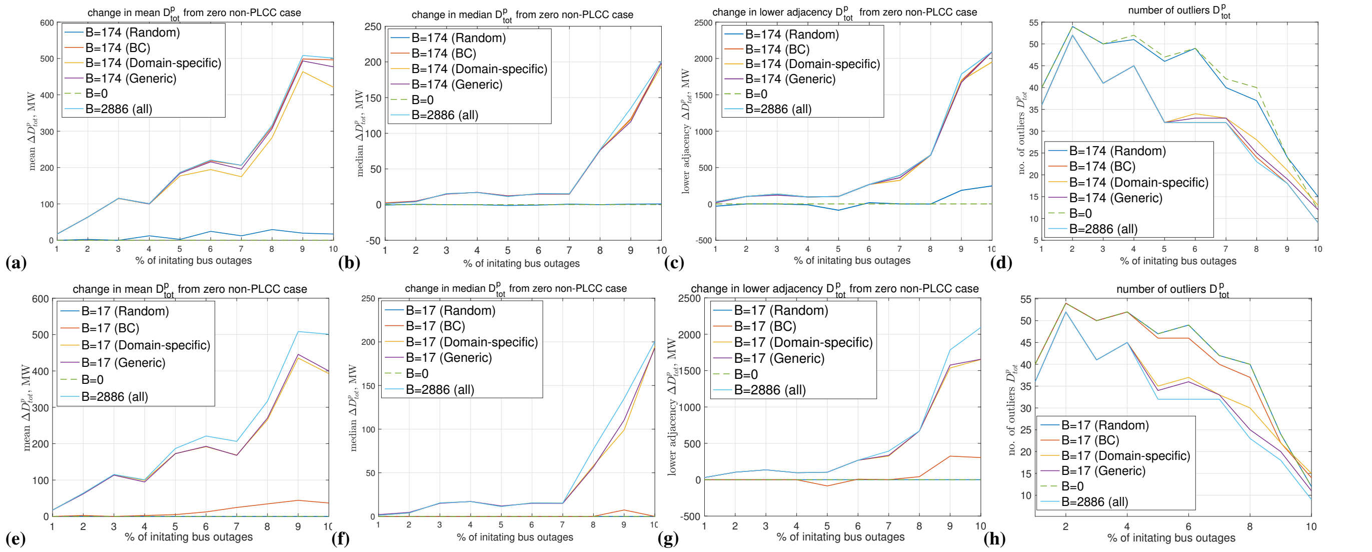


Figure 7. Performance evaluation of change in D_{tot}^p with respect to 100% PLCC case (non-PLCC links are placed by different methods under (a-d) $B = 174$, (e-h) $B = 17$ ($L = 200$ and $N = 50$ in the generic heuristic and $h = 10$ in the domain-specific heuristic).

running time of 32.5 s compared to 2441 s). Furthermore, it can be seen from Fig. 7(f) that using as few as 17 non-PLCC links, the proposed heuristics can serve a median post-contingency demand that is close to the ideal case where all the 2,886 links are non-PLCC. On the other hand, the performance deteriorates significantly with simplistic placement of these non-PLCC links (e.g., 'Random', 'BC'). This result signals the importance of placing non-PLCC links properly when under a budget constraint.

We also see from Fig. 7(d,h) that the comparison in terms of the number of outliers (extremely rare cases) is in line with the comparison in terms of the other statistics: a design leading to statistically higher demand served also has fewer outliers. We note that the upper adjacency, the largest data point that is not an outlier in the plot which is 3rd quartile + 1.5×inter-quartile range, demonstrates insignificant variations among different designs, which are not shown here since they represent outages that are non-critical.

Fig. 8 shows the communication layer of the system under different designs with a budget of 17 and 174, respectively, before imposing any failure. The plots visually describe the different design principles followed by each heuristic: the BC-based heuristic (Fig. 8(c,f)) builds a "backbone" of non-PLCC links, which works well when there is sufficient budget but poorly when the budget is highly limited; in contrast, the proposed heuristics strategically place non-PLCC links at the weakest parts of the network and leverage PLCC links when there is sufficient connectivity.

V. Conclusion

We study the impact of coupling between the communication and the power networks as it affects a SCADA-based preventive control system that leverages power line carrier communication (PLCC) to reduce the cost of deploying the communication network. As the failure of a power transmission line will fail the piggybacked PLCC link, we focus on improving the robustness of such a control system against cascading failures by allocating a limited number of non-PLCC links that are immune to power grid failures, which is formulated as a nonlinear integer programming problem. We establish the NP-hardness of the optimal solution and propose two heuristics that achieve different tradeoffs between the computational efficiency and the efficacy in mitigating cascades. Our evaluations based on a 2,383-bus Polish network demonstrate the promising result that a control system using only a few strategically-placed non-PLCC links and PLCC links elsewhere can achieve almost the same efficacy in mitigating cascading failures as a much more expensive control system that only employs non-PLCC links.

References

[1] S. Pahwa, C. Scoglio, and A. Scala, "Abruptness of cascade failures in power grids," *Scientific reports*, vol. 4, p. 3694, 01 2014.

[2] H. Hadlach, H. Touijer, M. Zahri, M. El alami, and M. Habibi, "Modeling of a smart grid monitoring system using power line communication," in *2017 International Renewable and Sustainable Energy Conference (IRSEC)*, 2017, pp. 1–4.

[3] K. Ali, A. X. Liu, I. Pefkianakis, and K. Kim, "Distributed spectrum sharing for enterprise powerline communication networks," in *2018 IEEE 26th International Conference on Network Protocols (ICNP)*, 2018, pp. 367–377.

[4] A. Majumder and J. Caffery, "Power line communication: An overview," *Potentials, IEEE*, 2004.

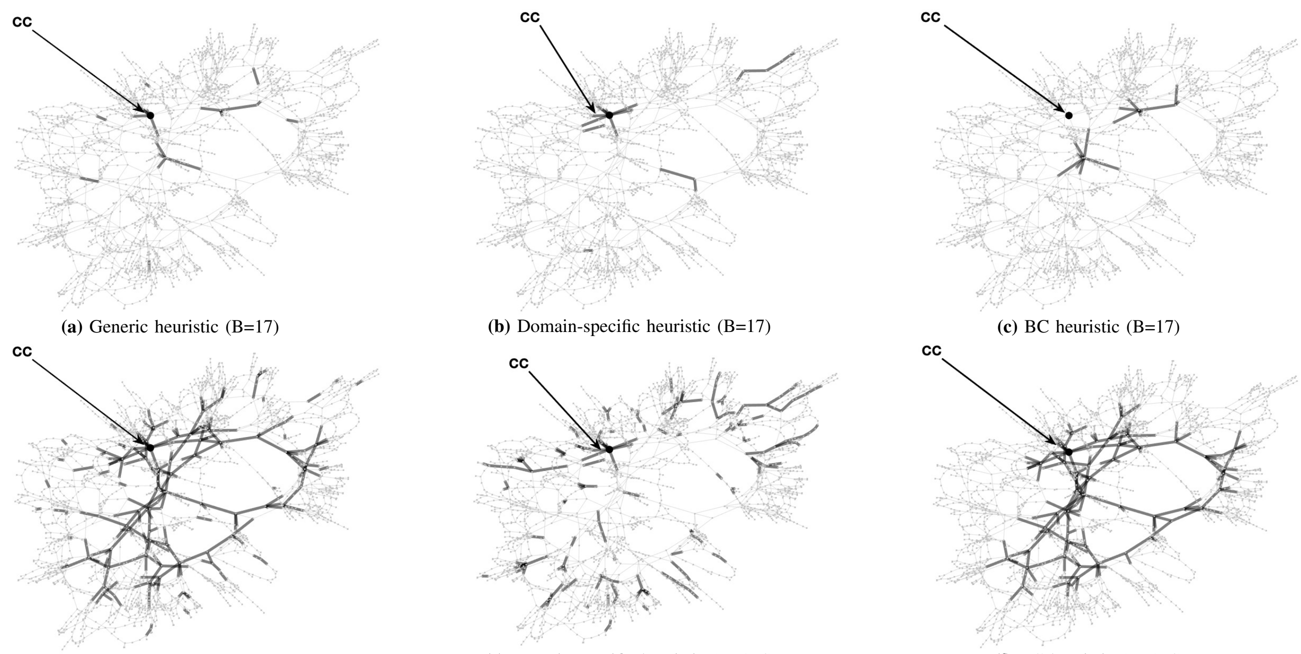
[5] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: The role of power line communications in the smart grid," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 998–1027, 2011.

[6] D. H. H. Akkermans and H. Ottosson, "The report on transmission of data over the electricity power lines." *AKMC Spectrum EnerSearch,*, Europe, June 1998.

[7] R. Alaya and R. Attia, "Narrowband powerline communication measurement and analysis in the low voltage distribution network," in *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2019, pp. 1–6.

[8] A. Mengi, S. Ponzelar, and M. Koch, "The itu-t g.9960 broadband plc communication concept for smartgrid applications," in *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2017, pp. 492–496.

[9] S. Canale, A. Di Giorgio, A. Lanna, A. Mercurio, M. Panfili, and A. Pietrabissa, "Optimal planning and routing in medium voltage powerline communications networks," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 711–719, 2013.



(a) Generic heuristic (B=17) **(b)** Domain-specific heuristic (B=17) **(c)** BC heuristic (B=17)

(d) Generic heuristic (B=174) **(e)** Domain-specific heuristic (B=174) **(f)** BC heuristic (B=174)

Figure 8. Graphs showing the non-PLCC links selected by different heuristics. Dark edges: non-PLCC, light edges: PLCC, CC: control center.

[10] R. Pighi and R. Raheli, "On multicarrier signal transmission for high-voltage power lines," in *International Symposium on Power Line Communications and Its Applications, 2005.*, 2005, pp. 32–36.

[11] D. Hyun and Y. Lee, "A study on the compound communication network over the high voltage power line for distribution automation system," in *2008 International Conference on Information Security and Assurance (isa 2008)*, 2008, pp. 410–414.

[12] K. W. Louie, A. Wang, P. Wilson, and P. Buchanan, "Discussion on power line carrier applications," in *2006 Canadian Conference on Electrical and Computer Engineering*, 2006, pp. 655–658.

[13] "Power network telecommunication powerlink – power line carrier system." [Online]. Available: www.siemens.com

[14] X. Zhang, Y. G. Gao, Y. Guo, and J. He, "Design and implementation of power line carrier intelligent home appliance control system," in *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, 2011, pp. 158–161.

[15] P. A. Brown, "Power line communications—past present and future," *in Proceedings of the International Symposium on Power Line Communicatons and Its Applications (ISPLC '99),*, 1999.

[16] B. S. Sushma, R. Roopesh, S. Gurugopinath, and R. Muralishankar, "Performance characterization of broadband powerline communication for internet-of-things," in *2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, 2019, pp. 146–151.

[17] N. Graf, I. Tsokalo, and R. Lehnert, "Validating broadband plc for smart grid applications with field trials," in *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2017, pp. 497–502.

[18] K. Ali, I. Pefkianakis, A. X. Liu, and Kyu-Han Kim, "Boosting powerline communications for ubiquitous connectivity in enterprises," in *2016 IEEE 24th International Conference on Network Protocols (ICNP)*, 2016, pp. 1–2.

[19] S. Sarkar and P. Kundu, "A proposed method of load scheduling and generation control using gsm and plcc technology," 01 2015, pp. 47 (5.)–47 (5.).

[20] Q. Wang, M. Pipattanasomporn, M. Kuzlu, Y. Tang, Y. Li, and S. Rahman, "Framework for vulnerability assessment of communication systems for electric power grids," *IET Generation, Transmission Distribution*, vol. 10, no. 2, pp. 477–486, 2016.

[21] L. Xu, Q. Guo, T. Yang, and H. Sun, "Robust routing optimization for smart grids considering cyber-physical interdependence," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5620–5629, 2019.

[22] "Telecommunications for power utilities." [Online]. Available: https://www.ge.com/digital/sites/default/files/download-assets/Utilities-Communications-Brochure-GE.pdf

[23] "Supervisory control and data acquisition (scada) systems," *Technical information bulletin 04-1, National Communications Systems*. [Online]. Available: https://www.cedengineering.com/userfiles/SCADA%20Systems.pdf, 2004.

[24] D. Bienstock, "Adaptive online control of cascading blackouts," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–8.

[25] M. Almassalkhi and I. Hiskens, "Model-predictive cascade mitigation in electric power systems with storage and renewables, part i: Theory and implementation," in *2015 IEEE Power Energy Society General Meeting*, 2015, pp. 1–1.

[26] J. Song, E. Cotilla-Sanchez, G. Ghanavati, and P. D. H. Hines, "Dynamic modeling of cascading failure in power systems," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2085–2095, 2016.

[27] S. K. Khaitan, Chuan Fu, and J. McCalley, "Fast parallelized algorithms for on-line extended-term dynamic cascading analysis," in *2009 IEEE/PES Power Systems Conference and Exposition*, 2009, pp. 1–7.

[28] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with dc power flow model and transient stability analysis," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 285–297, 2015.

[29] M. Korkali, J. Veneman, B. Tivnan, J. Bagrow, and P. Hines, "Erratum: Reducing cascading failure risk by increasing infrastructure network interdependence," *Scientific Reports*, vol. 8, p. 46959, 03 2018.

[30] D. Z. Tootaghaj, N. Bartolini, H. Khamfroush, T. He, N. R. Chaudhuri, and T. L. Porta, "Mitigation and recovery from cascading failures in interdependent networks under uncertainty," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 2, pp. 501–514, 2019.

[31] [Online]. Available: matpower.org/docs/ref/matpower5.0/case2383wp.html

[32] C. Report and W. G. . of Study Committee 35, "Requirements and performance of packet switching networks with special reference to telecontrol," August (1991). [Online]. Available: https://cigreindia.org/CIGRE%20Lib/Tech.%20Brochure/

[33] J. Golbeck, "Chapter 21 - analyzing networks," in *Introduction to Social Media Investigation*, J. Golbeck, Ed. Boston: Syngress, 2015, pp. 221 – 235. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780128016565000214

[34] H. Promel and A. Steger, "The steiner tree problem: A tour through graphs algorithms and complexity," 01 2002.

[35] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, pp. 65–85, 1994.

[36] [Online]. Available: https://xlinux.nist.gov/dads//HTML/simplepath.html

[37] S. Schaeffer, "Graph clustering," *Computer Science Review*, vol. 1, pp. 27–64, 08 2007.

[38] A. E. Jahromi and Z. B. Rad, "Optimal topological design of power communication networks using genetic algorithm," *Scientia Iranica*, vol. 20, pp. 945–957, 2013.