Mitigation and Recovery from Cascading Failures in Interdependent Networks under Uncertainty

Diman Zad Tootaghaj, *Student Member, IEEE*, Novella Bartolini, *Senior Member, IEEE*, Hana Khamfroush, Ting He, *Senior Member, IEEE*, Nilanjan Ray Chaudhuri, *Senior Member, IEEE*, Thomas La Porta, *Fellow Member, IEEE*

***Abstract*—The interdependency of multiple networks makes today's infrastructures more vulnerable to failures. Prior works mainly focused on robust network design and recovery strategies after failures, given complete knowledge of failure location. Nevertheless, in real-world scenarios, the location of failures might be unknown or only partially known. In this work, we focus on cascading failures involving the power grid and its communication network with imprecision in failure assessment. We consider a model where functionality of the power grid and its failure assessment rely on the operation of a monitoring system and vice versa. We address ongoing cascading failures with a twofold approach: 1) Once a cascading failure is detected, we limit further propagation by re-dispatching generation and shedding loads, 2) We formulate a recovery plan to maximize the total amount of load served during the recovery intervention. We performed extensive simulations on real network topologies showing the effectiveness of the proposed approach in terms of number of disrupted power lines and total served load.**

***Index Terms*—Interdependent networks; Cascading failures; Power Grids**

I. Introduction

TODAY'S critical infrastructures are highly interdependent. Because of the interdependency between different components, perturbations caused by failures, physical attacks or natural disasters may propagate across different networks. Examples of such coupled critical infrastructures include the food supply and water systems, financial transactions and power grids, transportation systems and food supply, etc. [2, 3]. These critical infrastructures are becoming increasingly correlated and interdependent. Therefore, modeling and understanding the interactions between multiple networks and designing failure resilient infrastructures is crucial for the reliability and availability of many applications and services.

One of the most critical infrastructures in our everyday lives is the power grid. Large-scale blackouts in the power grid due to propagating failures, natural disasters or malicious attacks can severely affect the operation of other interdependent critical infrastructures and cause catastrophic economic and social disruptions. In September 2003, a large cascading blackout, in Italy, led to the shortage of 6400 MW of power, which caused a complete system collapse [4]. A similar event occurred the same year in the Northeast of the United States, leading to over 50 million people losing power for several days. The cascade lasted approximately for four hours, a time sufficient for enabling countermeasures which could have mitigated and limited the blackout propagation. This highlights the necessity of a coherent power control strategy that allows prompt intervention to mitigate or stop the failure propagation. Furthermore, it is crucial to have a strategic recovery plan that ensures effective use of the available resources during the recovery process.

Despite considerable amount of research in the past few decades leading to major improvements in the reliability of communication and power networks, most of the research focused on the recovery of a single network [5–9]. Unlike previous work, we jointly address the following three challenges:

- Providing a realistic model of cascading failures in interdependent networks which takes into account the peculiarity of both the communication network and the power grid, and overcomes the limitations of the simplified epidemic models which cannot represent real infrastructures [10, 11].
- Modeling lack of knowledge in failure localization, by considering the potential uncertainty due to failures in the monitoring systems, which may hamper the use of appropriate countermeasures to prevent a cascading failure or recover network services.
- Progressively restoring the network after a large-scale disruption or a cascading failure in multiple stages, within the limits of recovery resources (time, cost, human personnel).

We make the following contributions to address the above challenges:

- To cope with insufficient knowledge of failure locations, we propose a new Consistent Failure Set (*CFS*) algorithm to determine all failure scenarios, namely the sets of components whose failure is consistent with the observations. *CFS* is used to determine local monitoring interventions that allow the identification of the state of all network components.
- We tackle the problem of mitigating an ongoing cascade (first phase) by formulating a minimum cost flow assign-

We thank, Vittorio Rosato for providing the power grid datasets.
This research is supported by the Defense Threat Reduction Agency under Grant HDTRA1-10-1-0085. The work of Novella Bartolini was supported by NATO under the SPS grant G4936 SONiCS.

D. Z. Tootaghaj, Ting He and T. La Porta are with the Comp. Sci. Dept. in the Pennsylvania State University, PA, USA (emails: {dxz149, tzh58, tlp}@cse.psu.edu), N. Bartolini is with the Comp. Sci. Dept. in Sapienza University, Rome, Italy (email: {bartolini}@di.uniroma1.it), H. Khamfroush is with the Comp. Sci. Dept. in University of Kentucky (email: {khamfroush}@cs.uky.edu), and N. R. Chaudhuri is with the Elec. Eng. Dept. in the Pennsylvania State University, PA, USA (email:{nuc88}@psu.edu). A partial and preliminary version appeared in Proc. IEEE SRDS'17 [1].

ment (*Min-CFA*) as a linear programming optimization problem. *Min-CFA* aims at finding a DC power flow setting that stops the cascading failure with minimum change in power generation and satisfied demand.

- We formulate a progressive recovery problem (second phase) to maximize the satisfied demands (*Max-R*) over multiple consecutive recovery interventions based on the available recovery resources at each stage. We show that *Max-R* is NP-hard, hence we propose two heuristic recovery strategies: 1) *Max-R-Greedy*, as a baseline algorithm, and 2) *Max-R-Backward*, which consider different time spans in scheduling the recovery interventions.
- We perform an extensive experimental evaluation, considering a real network scenario of interdependent power grid and communication network, under different interdependency models. The experiments highlight the efficiency of our cascade prevention algorithm. For example, in a scenario where 60% of the lines of the power grid has failed, the adoption of *Min-CFA* stops the propagation with 54% of served load, while without intervention, the entire system would fail. In the same scenario, our backward recovery approach *Max-R-Backward* performs full recovery, serving on average 20% more accumulative load than our baseline algorithm *Max-R-Greedy*.

II. Related Works

The existing works on cascading failures in interdependent networks, of interest to our approach, can be broadly classified into two categories: 1) those, that study the interaction through percolation theory and stochastic analysis [10–13], 2) studies that try to identify the most vulnerable nodes [14, 15], and aim at finding the root cause of failures and identify performance degradation [16–18].

The approaches of the first category rely on prior knowledge of the probabilistic model of failure propagation, which is hard to obtain. In addition, real systems usually have a deterministic failure propagation. For example, if a power line fails, a certain number of communication routers will certainly stop working.

Concerning the works of the second category, we underline that finding the root cause of the propagating failure is key to the design of failure-resilient systems, but does not provide a mitigation solution.

Another line of research addresses the problem of cascading failure in the power grid and studies the peculiarity of the propagation across power lines. Cascading failures in power grids can be due to a permanent failure, e.g. a tree falls on a transmission line etc., or to a temporary failure, e.g. a temporary short circuit in a transmission line. When a short circuit happens in a transmission line, a protective overcurrent relay sends a "trip" signal to the breakers and the breakers set open. Then the relay attempts to re-close the breaker a few times. In case of a permanent failure, auto-reclosing fails and the breaker stays open circuit. After a line fails in the system, the power re-distributes according to Kirchhoff's and Ohm's laws. This can cause the overload of other lines, trigger new failures, spreading over the entire network.

Unlike the approach proposed in [19], that re-distributes the power flow evenly over all transmission lines, we use the DC power flow model [20, 21]. Notice that this model is widely used in studies of cascading failures in power grids and is acknowledged to be a good approximation of the AC power flow model.

The operation and reliability of today's power grid is highly dependent on the operation of the communication network that provides the necessary information needed by the Supervisory Control and Data Acquisition\Energy Management System (SCADA\EMS) and more recently, the Wide-Area Monitoring, Protection, and Control (WAMPAC) system to respond to emergency situations. The required data is measured and gathered at the substations from the Intelligent Electronic Devices (IEDs), fault recorders, breaker status monitors, and Phasor Measurement Units (PMUs) [22, 23]. While substation automation is increasing the intelligence and autonomy of local protection units, it is facilitating the trend of centralized wide-area protection popularly called the Remedial Action Scheme (RAS) or the Special Protection Scheme (SPS). In the latter, data is communicated to a centralized location, which can initiate corrective actions such as generation re-dispatch or load shedding to remote locations. A nice description can be found in [24] and [25] and references therein. The security of such centralized control systems is itself an important challenge on the reliability of the power grids (e.g. a compromised RAS can send an anomalous load shedding signal to disrupt the power grid) [26–29].

Concerning the problem of restoring network functionalities, previous work only considers one homogeneous network [5–9]. Our recovery approach takes inspiration from the works in [6, 8, 9] and aims at restoring the functionality of multiple interdependent networks in a progressive manner, depending on resource and incremental knowledge availability. Our recovery approach represents a step ahead with respect to these previous contributions in that it deals with heterogeneous interdependent networks, and includes a run-time feedback control of the flow during the recovery process. To the best of our knowledge, the problem of jointly mitigating and recovering from cascading failures, during the transient regime of the propagation process, was never studied extensively before as we do in the present work. We address such a problem with particular focus on failures in the power grid and the interdependent communication network.

III. Network Model and Background

We model the interdependency between the power grid and the communication system for which some failures are detected while the propagation is still in progress. We propose a mitigation strategy to minimize further cascades and a recovery plan to entirely restore the power grid functionality, while maximizing the accumulative amount of delivered power during multiple stages of progressive recovery. Our approach can be extended to the use of other measures related to the operation of the grid such as the total number of working power lines, etc. Table I shows the notation used in this paper.

A. Notation

The power and communication networks are modeled as undirected graphs $\mathcal{G}_p = (V_p, E_p)$ and $\mathcal{G}_c = (V_c, E_c)$, respec-

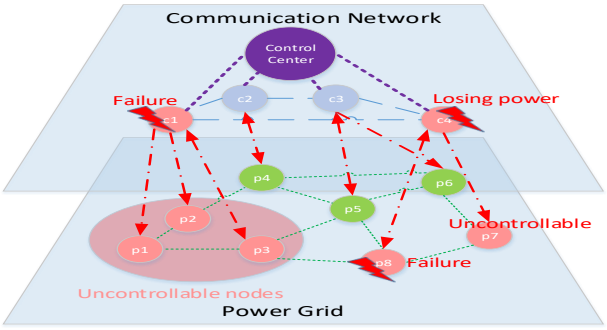


Fig. 1: Interdependency model between a power grid and a communication network.

TABLE I: Summary of notations.

Notation	Explanation
$\mathcal{G}_p = (V_p, E_p)$	undirected graph modeling the power grid. V_p is the set of nodes and E_p is the set of links
$\mathcal{G}_c = (V_c, E_c)$	undirected graph modeling the communication network. V_c is the set of nodes and E_c is the set of links
$G \subset V_p$	set of generator nodes G_i where power is generated
$L \subset V_p$	set of load nodes L_i where the power is consumed
$J \subset V_p$	set of junction nodes J_i where the power just flows
$E_p^{B,t} \subseteq E_p$	set of broken (red) edges
$E_p^{U,t} \subseteq E_p$	set of unknown status (grey) edges
$E_p^{W,t} \subseteq E_p$	set of working (green) edges
F_{ij}^t	power flow in line (ij) at time t
θ_i^t	voltage angle of node i at time t
x_{ij}	series reactance of line (ij)
P_i^t	power generated/consumed at node i at time t
Y^t	nodal admittance matrix at time t
w_{G_i}	cost of increasing the power in generator G_i
w_{L_i}	cost of shedding the power of load L_i
$P_{G_i}^{max}$	maximum power that can be generated in G_i
$P_{L_j}^{demand}$	demand load at L_j
F_{ij}^{max}	maximum capacity of the line (ij)
E_k^R	set of restored edges at stage k
$\delta_{(ij),k}$	decision to repair $(ij) \in E_p^{B,t}$ at the kth stage
r_{ij}	resources needed for repairing (ij)
R_k	available resource at stage k of the recovery
α_i	a constant factor showing the reduction of the new power distribution in node i, where $0 \leq \alpha_i \leq 1$.

tively. Each node i in the power grid is monitored by several sensors deployed nearby. The monitoring data is then sent to the node of the communication network which hosts the control functionalities related to node i of the power grid. In addition, control commands are sent to the dependent communication node for generator re-dispatch or load shedding. We acknowledge that several other emergency actions could be taken (e.g. system separation, dynamic braking, fast valving, etc.), which are not considered here.

The set of nodes V_p of the power grid is composed by three disjoint sets of generators G where the power is generated, loads L where the power is consumed, and junctions J where power flows by, with $V_p = G \cup J \cup L$.

As node failures are less likely [30], we hereby assume that initial failures only occur in power lines (E_p). A node in the power grid is considered failed if it is not able to deliver the required power to the loads. Further, we consider the interdependency between the power grid and the communication network such that (i) failures in the communication network lead to lack of information and controllability of power grid in the control center, and (ii) failures in the power grid can cause further failures in the communication system due to lack of power. The edges in the power grid graph $\mathcal{G}_p$ can be in three different states: 1) the set $E_p^{B,t} \subseteq E_p$ is the set of **broken** edges (hereby denoted as red edges) at time t [1]; 2) the set $E_p^{U,t} \subseteq E_p$ is the set of edges with **unknown** status (denoted as grey edges) at time t; 3) the set $E_p^{W,t} \subseteq E_p$ is the set of **working** edges (denoted as green edges) at time t.

B. Interdependency Model

To clarify the interdependency model between the communication network and the power grid, consider the example shown in Figure 1. The figure shows the interdependency model between a communication network with 4 nodes $\{c_1, .., c_4\}$, and a power grid with 8 nodes $\{p_1, ..., p_8\}$. The red arrows show the interdependency between the two networks. For example, c_1 controls three power nodes $\{p_1, p_2, p_3\}$ and gets power from p_3. Now consider a failure in one of the communication nodes c_1. In this case three power grid nodes $\{p_1, p_2, p_3\}$ become uncontrollable as the controller cannot send the power adjustment control commands to them. Next, consider a failure in a node in the power grid p_8. In this case, the communication node c_4 that gets power from p_8 loses power and consequently the dependent power grid node p_7 becomes uncontrollable. In this work, we consider three types of interdependency models between the communication network and the power grid: the one-way interdependency model (Fig. 2a); the location-based interdependency model (Fig. 2b); the random interdependency model (Fig. 2c).

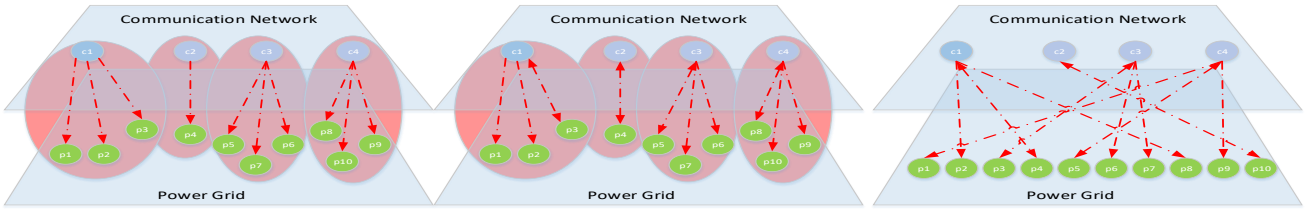
In the one-way interdependency model we assume the power lines are monitored and controlled by the closest communication node, while if a grid node fails, the communication nodes get backup power from an external source (e.g., battery). This is the case for many telecommunication deployments with battery backup. In the one-way interdependency model, failures in the power grid will not cascade in the communication network. However, the disruption in the communication network is observed as lack of knowledge and uncontrollability in the power grid.

As in the one-way model, in the location-based interdependency model, each power line is monitored and controlled by the closest communication node. Nevertheless, in the location-based model we assume that each node in the communication network gets power from the closest node in the power grid.

Both in the one-way and location-based interdependency models, the communication network and power grid are divided into dependent regions and the failure in one region does not cascade or affect other nodes in a different region.

In the random interdependency model, we assume each power line is monitored and controlled by a random communication node and each communication node gets power

[1]Notice that in order to be able to assess an edge failure, the edge must be connected to a working communication node in $\mathcal{G}_c$. The working communication node provides the failure status of the edge to the central controller and can send power adjustment commands to the connected loads or generators.



(a) One-way (b) Location-based (c) Random

Fig. 2: Three interdependency model between a power grid and a communication network: a) One-way, 2) Location-based and 3) Random.

from a random substation in the power grid. In the random interdependency model, the failure in one node may cascade and spread over the entire network.

Notice that most prior works on the interdependency between the power grid and the communication network use a random interdependency model (see examples in [11, 31–33]) and focus on the ping-pong failures from the power grid to the communication network and vice versa.

While a random interdependency model is easy to analyze and simulate on synthetic graphs, it fails to capture most real network topologies. Nevertheless, we include this model in our analysis for the sake of completeness in providing comparisons with previous work and because we consider this as a stress-test due to its potentiality to produce larger cascades.

By comparing the experiments conducted on the random interdependency model with the ones using location-based and one-way interdependency models, in Section VI-A we will show that the extent of the cascading failure phenomenon is mostly related to the lack of controllability consequent to failures in the communication network.

C. DC Power Flow Model

We model the cascading failure in a power system using a DC load flow model [20]. Let F_{ij}^t be the power flow in line (ij) at time t, x_{ij}^t be the series reactance of line (ij) and θ_i^t and θ_j^t be the voltage angles of node i and j, at time t[2]. The DC power flow model provides a linear relationship between the active power flowing through the lines and the power generated/consumed in the nodes, which can be formulated as follows.

$$F_{ij}^t = \frac{\theta_i^t - \theta_j^t}{x_{ij}^t}\ , \tag{1}$$

The power flow of node i can be found by summing up the power flows of all its adjacent power lines:

$$P_i^t = \sum_j F_{ij}^t \tag{2}$$

We can re-write the power flow model as a linear system of equations as follows:

$$P^t = Y^t \theta^t, \tag{3}$$

[2]Notice that the reactance x_{ij}^t of line (i,j) varies with t as a consequence of failures or recovery events.

where Y^t is nodal admittance matrix at time t, $y_{ij}^t = -\frac{1}{x_{ij}^t}$ for $i \neq j$, and $y_{ii}^t = \sum_k \frac{1}{x_{ik}^t}$ [34]. Once events like over-current are detected in a transmission line (ij), a protective relay trips a circuit breaker ($x_{ij} = \infty$) and the power is redistributed according to the DC model. In particular, if the current flow exceeds the maximum threshold on another line $(i'j')$, in a cascading manner, the transmission line $(i'j')$ may also trip.

Notice that in order to determine the power flowing through each line after one or more failures, we need to solve the system of equations 1, which requires the solution of Equation 3 to obtain the values of the vector θ^t, for each connected component of the power grid graph.

Remark 1. [35] The system of equations (3) has a feasible solution for each connected component of the power graph.

Discussion on remark 1. Let us consider a unique connected component. The nodal admittance matrix Y of a connected graph with n nodes has always rank $(n-1)$ because one can construct a graphic matroid where the nodal admittance matrix is a weighted incident matrix. It is known that the rank of a weighted incident matrix is equal to the rank of any basis (tree) in the graph, which is $(n-1)$ [35, 36]. To make this equation solvable, one of the equations is removed and the corresponding node can be chosen as a reference node. Without loss of generality, the removal of the first equation implies $\theta_0 = 0$ and the other $n-1$ values of the vector θ can be calculated inverting the system of equations 3 reduced after the removal of the first equation, according to the technique in [37, 38]. The reduced admittance matrix has full rank and thus invertible. If instead the power grid graph has c connected components due to the disruption of several lines, then the admittance matrix will have rank $n-c$ and each component must be addressed by means of the same technique. Let Y'^t be the admittance matrix of a connected component, and θ'^t its phasor vector, let also P'^t be its power vector, at time t. Then the power flow system of equations of the considered connected component is $P'^t = Y'^t \theta'^t$, which can be solved independently of the other connected components, in the way described for the case of a unique connected component, with the removal of one equation and the introduction of a reference phasor vector, as described in [37, 38]. Therefore, the DC power flow model for each connected component of the graph has a feasible solution. □

IV. Cascade Mitigation and Network Recovery

In this section, we address the problem of cascading failure mitigation and related recovery process in an interdependent network formed by a power grid and a communication network. Figure 3 illustrates our two-phase approach. As described in the figure, whenever a new failure event shows up, a preliminary monitoring activity is performed to localize the failure sites. After the failure assessment it follows a first phase in which further cascades are mitigated or prevented by means of a combination of load shedding and adjustment of the generated power. Once the cascade is stopped, a progressive recovery activity follows, using either a greedy or backward approach. Recovery is performed in multiple stages according to resource availability. After the system is recovered, the monitoring activity restarts, until new failures occur.

1) Cascade mitigation: Once we detect an outage of a transmission line, we readjust power and load according to the optimization problem described in the following.

For clarity of presentation, we hereby assume that the power grid is formed by a unique connected component. Notice that in the presence of multiple connected component, all the following techniques are still valid, when applied to each connected component, independently. The Minimum Cost Flow Assignment (*Min-CFA*) optimization problem minimizes the total cost of reducing the load or generator's power. Let w_{G_i} be the weighted cost of reducing the power in generator G_i and w_{L_i} be the weighted cost of decreasing the power of load L_i. The *Min-CFA* problem to prevent the cascading failures can be formulated as follows:

$$\text{minimize} \quad \sum_{G_i \in G, L_j \in L} w_{G_i}(P^0_{G_i} - P^t_{G_i}) - w_{L_j}(P^t_{L_j} - P^0_{L_j})$$

$$\text{subject to} \quad 0 \leqslant P^t_{G_i} \leqslant P^0_{G_i}, \ \ \forall G_i \in G \tag{4a}$$

$$0 \leqslant P^t_{L_j} \leqslant P^0_{L_j}, \ \ \forall L_j \in L \tag{4b}$$

$$-F^{max}_{ij} \leqslant F^t_{ij} \leqslant F^{max}_{ij}, \ \ \forall (ij) \in E^t_p \tag{4c}$$

$$\sum_{G_i \in G} P^t_{G_i} + \sum_{L_j \in L} P^t_{L_j} = 0 \tag{4d}$$

$$P^t_{G_i} = \sum_{j:(G_i,j) \in E^t_P} F^t_{ij}, \ \ \forall G_i \in G \tag{4e}$$

$$-P^t_{L_i} = \sum_{j:(L_i,j) \in E^t_P} F^t_{ij}, \ \ \forall L_i \in L \tag{4f}$$

$$P^t = Y^t \theta^t \tag{4g}$$

$$F^t_{ij} = \frac{(\theta^t_i - \theta^t_j)}{x_{ij}}, \ \ \forall (ij) \in E^t_p \tag{4h}$$

The decision variables in *Min-CFA* are the continuous values of power in the generators ($P^t_{G_i}$) and loads ($P^t_{L_j}$).

Constraint 4a indicates that the power generated at each generator at time t cannot exceed the initial power observed at time 0. In case of full knowledge of the location of failures, we could have a more relaxed constraint to increase the power at some of the generators without violating a maximum threshold. However, under uncertain failure we reduce our solution space to decrease the possibility of consequent cascades due to imperfect knowledge. Constraint 4b shows that the reduced load cannot exceed the demand. Constraint 4c shows that the power flowing through each line cannot exceed the maximum capacity of the line (*thermal constraint*). Constraint 4d is the power conservation condition, i.e. the total power generated in the generators should be equal to the total power consumed in the loads. Constraints 4e and 4f show that the total power generated/consumed at each node should be equal to the total power flowing through its edges. Constraints 4g and 4h reflect the DC power flow model, to be solved according to Remark 1 [38].

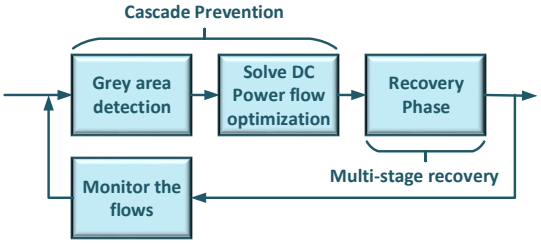


Fig. 3: Recovery Process: 1) Cascade mitigation phase, and 2) Recovery phase.

We next provide the sufficient condition for each connected component of the power grid, to ensure that the solution to our power re-distribution model under uncertain knowledge of the failure does not increase the power flowing through any line in the system, making it exceed the thermal threshold given by constraint 4c. Suppose that in the new power assignment for each node i, P_i is reduced by a factor α_i, i.e. $P^t_i = \alpha_i P^0_i$, for $i = 0, \ldots, n-1$, where $0 \leq \alpha_i \leq 1$. Without loss of generality, let the reference voltage angle be at node 0 ($\theta_0 = 0$). Also, using the DC power flow model, we can calculate the voltage angles at each node as follows:

$$\theta_i = \sum_{m=1}^{n-1} k_{im} P_m, \qquad \text{with} \qquad k_{im} \geq 0, \tag{5}$$

where n is the number of nodes in the power grid, and k_{im} reflects impedance values, that are always non negative [37, 38].

Let $\alpha^- \triangleq min\{\alpha_i : i = 1, ..., n-1\}$ and $\alpha^+ \triangleq max\{\alpha_i : i = 1, ..., n-1\}$. A sufficient condition for not having cascade propagation is the following.

Theorem 1. *Sufficient condition for no additional cascading failures in Min-CFA problem is as follows:*

$$\frac{1-\alpha^-}{1-\alpha^+} \leq \frac{\sum_{m:G_m \in G}(k_{im} - k_{jm})P_m}{\sum_{m:L_m \in L}(k_{im} - k_{jm})(-P_m)} \quad \forall (ij) \in E^t_p \tag{6}$$

Proof of Theorem 1. Without loss of generality, we assume $F_{ij} \geq 0$, otherwise we can simply interchange i and j index. In order to check whether the new power flow $F_{ij,new}$ provided by our control approach is not larger than the previous flow F_{ij}, we need to check if $F_{ij} \geq F_{ij,new}$, which is equivalent to

check if $(\theta_i - \theta_j) \geq (\theta_{i,new} - \theta_{j,new})$. Note that the new and previous voltage angles at each node is computed as follows:

$$\theta_i = \sum_{m=1}^{n-1} k_{im} P_m, \text{ where } k_{im} \geq 0$$

$$\theta_{i,new} = \sum_{m=1}^{n-1} k_{im}\alpha_m P_m, \text{ where } 0 \leq \alpha_m \leq 1, \quad k_{im} \geq 0$$

Therefore, we have to verify the following inequality:

$$(\sum_{m=1}^{n-1} k_{im} P_m - \sum_{m=1}^{n-1} k_{jm} P_m) \geq (\sum_{m=1}^{n-1} k_{im}\alpha_m P_m - \sum_{m=1}^{n-1} k_{jm}\alpha_m P_m)$$

To check if the flow in F_{ij} does not increase after our power flow adjustment, we need to verify the following inequality, for each edge $(i,j) \in E_p^t$:

$$\sum_{m:G_m \in G} (k_{im} - k_{jm}) P_m (1 - \alpha_m) \geq \sum_{m:L_m \in L} (k_{im} - k_{jm}) P_m (\alpha_m - 1) \tag{7}$$

Now we show that if condition 6 is satisfied for every flow, we can make sure that the new power flow at each line is no more than the power flow without any adjustment of generator or load's power. We have:

$$\sum_{m:G_m \in G} (k_{im} - k_{jm}) P_m (1 - \alpha_m) \geq$$
$$\sum_{m:G_m \in G} (k_{im} - k_{jm}) P_m (1 - \alpha^+) =$$
$$(1 - \alpha^+) \sum_{m:G_m \in G} (k_{im} - k_{jm}) P_m \geq$$
$$(1 - \alpha^-) \sum_{m:L_m \in L} (k_{im} - k_{jm})(-P_m) =$$
$$\sum_{m:L_m \in L} (k_{im} - k_{jm})(-P_m)(1 - \alpha^-) \geq$$
$$\sum_{m:L_m \in L} (k_{im} - k_{jm})(-P_m)(1 - \alpha_m)$$

Therefore, if equation 6 holds, the inequality 7 is also verified, which implies that the new power set at each line does not exceed the previous value. □

2) Recovery Phase: In this paragraph we address the problem of scheduling recovery interventions in order to maximize the total accumulative flow absorbed by the loads during K stages of recovery. The number of stages can be set according to the assumption that at least one edge can be repaired at each stage. Therefore K can be set equal to the number of broken edges. We hereafter refer to the multiple stages of progressive recovery shortly with the word *stages*.

Let $\delta_{(ij),k}$ be a binary variable representing the decision to repair edge (i,j) at time $k = 1, \ldots, K$. Namely, if edge (i,j) is being repaired at time k, $\delta_{(ij),k} = 1$ and $\delta_{(ij),k'} = 0$, $\forall k' \neq k$. Similarly, if an edge (i,j) had never failed, we set $\delta_{(ij),0} = 1$ to keep into account that it must not be scheduled for repair. For shortness of notation we also define the decision matrix Δ_k, whose ij-th element corresponds to the decision $\delta_{(ij),k}$.

The recovery of a broken line (ij) requires r_{ij} recovery resources. At each recovery stage, R_k resources are available for recovery interventions. We assume to have a budget rollover, so that resources that have not been consumed until the end of stage $k-1$ are available at the k-th recovery stage, and summed up to the R_k newly available.

Notice that in our model, the power grid graph at each stage k includes all the edges repaired according to the repair schedule performed until time k. Hence, we denote the power absorbed by load L_j at time k by $P_{L_j}^k(\Delta_1, \ldots, \Delta_k)$, calculated by means of the iterative solution of problem *Min-CFA* at stage k.

The maximum recovery (*Max-R*) optimization problem aims at maximizing the accumulative delivered power over K recovery stages. The *Max-R* recovery problem is formulated as follows:

$$\text{maximize} \quad \sum_{k=1}^{K} \sum_{L_j \in L} P_{L_j}^k(\Delta_1, \ldots, \Delta_k)$$

$$\text{s.t.} \sum_{m=1}^{k} \sum_{(ij) \in E_p} \delta_{(ij),m} \cdot r_{ij} \leq \sum_{m=1}^{k} R_m, \ k = 1, ..., K \tag{8a}$$

$$\sum_{m=0}^{k} \delta_{(ij),m} \leq 1, \ \forall (ij) \in E_p, \ k - 1, \ldots, K \tag{8b}$$

$$\delta_{(ij),k} \in \{0,1\}, \ \forall (ij) \in E_p, \ k = 1, ..., K \tag{8c}$$

where $\delta_{(ij),k}$ is the decision variable to repair $(ij) \in E_p^B$ at the k-th stage of the algorithm. Constraint 8a indicates that at stage k of the recovery schedule, R_k new recovery resources become available and resources rollover from the previous stages so that if some available resources are not used until the k-th stage of the recovery, they are still available in the following stages. Constraint 8b shows that each broken line can be repaired only in one stage of the recovery schedule.

Notice that *Max-R* is a combinatorial and nonlinear optimization problem. In fact, the objective function is the accumulative power flow measured at the loads in the K steps of execution of the algorithm. We underline that the total delivered power appearing in the objective function depends on the recovery decisions adopted at each stage of the recovery schedule. With $P_{L_j}^k(\Delta_1, \ldots, \Delta_k)$ we denote the power absorbed by load L_j when the recovery decisions given by the decision matrix Δ_m are made according to the schedule up to step $m = 1, \ldots, k$. Such an optimization problem is combinatorial and non-linear, in that it requires the solution of the *Min-CFA* optimization problem to find $\sum P_{L_j}^k(\Delta_1, \ldots, \Delta_k)$, since the set of working lines at stage k changes based on the current and previous decisions of the recovery schedule Δ_m, for $m = 1, \ldots, k$.

Note that in the recovery phase, we remove the generator's power reduction constraint and the generator and load's power increases gradually until all demand loads are all satisfied.

Theorem 2. *The problem of Max-R is NP-Hard.*

Proof of Theorem 2. We prove the NP-hardness of the *Max-R* problem showing that it generalizes the Knapsack problem. We recall that the Knapsack problem considers a set of items I, each item $i \in I$ has a size S_i and a value $V_i > 0$. The problem is to find a subset $I' \subseteq I$ such that $S(I') \leq S$ and $V(I')$ is *maximized*, where $S(I') = \sum_{i \in I'} S_i$ and $V(I') = \sum_{i \in I'} V_i$.

In the following we show how we can build, in polynomial time, an instance of a single stage ($K = 1$) of *Max-R* problem whose solution corresponds to the solution of the generic formulation of the Knapsack problem given above.

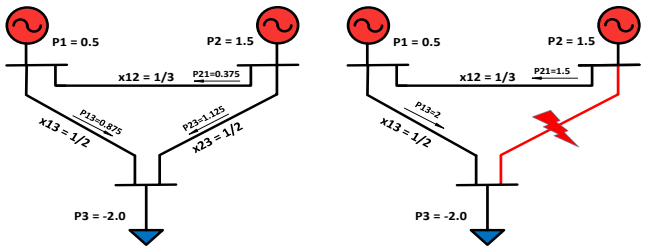
Since we consider a single stage of the *Max-R* problem, we assume R resources are available to repair all disrupted lines $(ij) \in E_p^{B,t}$. We also assume that we have complete information about the disrupted lines. Let us consider a set of generators I, each generator corresponding to an element $i \in I$ of the Knapsack problem, producing a flow equivalent to the value V_i of the element. Each generator $i \in I$ is connected to a unique common load L with a broken line, whose repair cost is equivalent to the size S_i of the corresponding Knapsack element. We also assume that the load L has a demand of at least the summation of all flows ($\sum_{i \in I} V_i$). We set the recovery budget of *Max-R* equal to S, the size of the Knapsack. This instance of *Max-R* can be defined in polynomial time starting from any instance of Knapsack. Solving this instance of *Max-R*, corresponds to finding a list of links to be recovered with cost limited by S, such that the flow reaching the common load L is maximized, which is equivalent to selecting the Knapsack subset $I' \subseteq I$ with maximum value, and bounded size S, which completes the proof that any instance of the Knapsack problem can be polynomially reduced to the solution of an instance of *Max-R*, which implies the NP-hardness of *Max-R*, showing that *Max-R* is at least as difficult as the Knapsack problem. □

There exists no polynomial-time solution for general instances of knapsack problem. Therefore, each stage of Max-R is NP-hard. We also note that the maximum recovery problem is a combinatorial optimization and the total flow that each line can add to the final solution of the problem is unknown in advance and depends on the recovery schedule of other lines in the previous recovery stages. The marginal flow that each line can add to the current solution of the problem can be found by solving the *Min-CFA* problem introduced in section IV-1 which itself is a linear programming optimization.

As *Max-R* is NP-hard, and due to the broadness of the feasible region of the problem, which includes all possible permutations of repair interventions, in the following Section V we propose polynomial heuristic approaches to stop failure propagation and recover the network under uncertain failure localization.

V. Heuristics for Cascade Mitigation and Recovery under Uncertain Knowledge

In this section, we first apply a linear algebraic approach to increase our incomplete knowledge of the phasor voltages as much as possible, and then we describe the consistent failure set approach to detect the status of grey lines. The grey lines, as described in Section III are the set of edges whose working status is unknown to the controller. Then, we describe two heuristic algorithms to solve *Max-R*. Inspired by the approach proposed in [9] that finds a progressive recovery schedule in a data communication network, we first propose a greedy approach with polynomial time complexity and then propose a polynomial time backward approach that solves a single stage of the problem and finds the solution for all stages in reverse recovery scheduling order.



(a) All lines working. (b) Failure in line (23).

Fig. 4: An example of a 3-bus network where active power and reactances are in pu.

A. Identifiability of voltage phasors

When the network is divided into a known and unknown part we can re-write the DC power flow equations as follows:

$$\begin{pmatrix} Y_{known} \\ Y_{unknown} \end{pmatrix} \times \begin{pmatrix} \theta_{known} \\ \theta_{unknown} \end{pmatrix} = \begin{pmatrix} P_{known} \\ P_{unknown} \end{pmatrix} \quad (9)$$

Therefore, some of the unknown voltage phasors can be found by solving the following linear set of equations:

$$Y_{known} \times \begin{pmatrix} \theta_{known} \\ \theta_{unknown} \end{pmatrix} = P_{known} \quad (10)$$

Let $Null(Y_{known})$ denote the null space of Y_{known}, i.e., for any vector $v \in Null(Y_{known})$, we have $Y_{known} \cdot v = 0$ [18].

Theorem 3. *[39] For an arbitrary matrix Y_{known}, let $Null(Y_{known})$ represent the null space of Y_{known}. Voltage phasor θ_i is identifiable if and only if $\forall v \in Null(Y_{known})$ we have $v_i = 0$.*

Therefore, in order to find a set of identifiable voltage phasors, we can first compute the null space of Y_{known} and find all indexes with zero values in the null space. The value of the identifiable voltage phasors can be found by solving the linear system of equations (10).

B. Finding a Consistent Failure Set (CFS)

In this section, we first consider an illustrative example showing the impact of incomplete knowledge on the extent of the failure propagation and uncontrollability. We then propose a Consistent Failure Set (*CFS*) algorithm to cope with lack of knowledge.

Example. *Consider the network given in Figure 4, using the DC power flow model to calculate the power flows in the lines, where the reference angle is $\theta_1 = 0$, we have:*

$$\begin{pmatrix} \theta_2^0 \\ \theta_3^0 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1.5 \\ -2 \end{pmatrix} = \begin{pmatrix} 0.125 \\ -0.4375 \end{pmatrix} \quad (11)$$

The power flow through each line is then computed as follows:

$$F_{12}^0 = \frac{\theta_{12}^0}{x_{12}} = 3 \times (0 - 0.125) = -0.3750, \quad (12)$$

$$F_{13}^0 = \frac{\theta_{13}^0}{x_{13}} = 2 \times (0 - (-0.4375)) = 0.875, \quad (13)$$

$$F_{23}^0 = \frac{\theta_{23}^0}{x_{23}} = 2 \times (0.125 - (-0.4375)) = 1.125. \quad (14)$$

If the power line 23 *gets disrupted as in Figure 4b, the power redistributes according to DC power flow model, where* $F_{21}^1 = 1.5$ *and* $F_{13}^1 = 2$*. Suppose that the maximum power that each line can tolerate is* $F_{ij}^{max} = 1.3$*. Therefore, after the first line gets disrupted, the whole system collapses and the demand load cannot be satisfied. However, if we know the exact location of the failure, the RAS/SPS may reduce the generator's power to satisfy a degraded quality of service. One trivial solution of Min-CFA to this problem is to reduce the second generator's power to* $P_2^1 = 0.8$ *and reduce the load to* $P_3^1 = -1.3$ *without violating the maximum power on each line. However, under the uncertainty of the exact location of the failure, the controller fails to make appropriate decisions and the whole network collapses.*

In order to have a correct damage assessment, and solve the uncertainty in the grey area we propose the following *CFS* algorithm. We assume that there is no local load shedding. Hence, the powers at the generators and loads are only controlled through the central controller unit, so the power generated/consumed in each generator/load or junction is known.

To explain our algorithm, we first consider the nodes that have the smallest number of grey links.

Lemma 1. *In the power grid graph* G_p*, if there exists a node* $v \in V_p$ *which has only one grey incident link* $(v, w) \in E_p^t$*, the exact status of the grey link* (v, w) *is identifiable.*

Proof of Lemma 1. The exact status of a single grey link incident to a node v can be determined using the power flow Equation 2. In fact, under the assumption of only centralized load shedding, the power generated/consumed at node v is known from the previous stage. Hence, the power flowing through the grey line can be found by solving the power flow Equation 2 where the only unknown variable is F_{vw}^t. □

Lemma 2. *If the grey area does not contain a cycle and there exists at least one edge in the power grid graph* G_p *whose status is known, the exact status of all grey edges can be found in* $O(|E_p^{U,t}|)$*.*

Proof of Lemma 2. If the grey area does not contain any cycles, it forms a tree. Hence there exists at least one node that has only one grey incident link whose statuse can be identified according to lemma 1. This procedure can be repeated to find the status of all grey links in $O(|E_p^{U,t}|)$. □

Therefore, we can find the state of the network for all grey links, which are not inside a cycle. In the presence of a grey cycle, the *CFS* algorithm breaks the cycle by selecting one arbitrary link within the cycle, and considering the two potential statuses, broken or working. *CFS* then finds one or multiple consistent failure sets, namely sets of status assignments to each grey link of the cycle, which are consistent with the available observations.

If the consistent failure set algorithm finds multiple solutions, *CFS* provides local inspection to determine the actual failure set.

The *CFS* algorithm is described in Algorithm 1. In more details, for the case study of a graph G_p which has one or multiple cycles in its grey area, *CFS* starts by considering a case in which there are no grey link cycles (**line 2**), finding the status of grey links by considering them iteratively, starting from the nodes with only one incident grey link, according to the function `CFS-Cycle-Free`, described in Algorithm 2.

If all nodes have at least two grey links in the graph, i.e. there exists a cycle in the grey area (**line 4**), *CFS* picks an arbitrary edge within a cycle (**line 5**).

The algorithm tries to solve the unknown status of the grey edges by assuming one edge at each cycle to be working or not working according to a decision tree, henceforth generating 2^C possible link status combinations. *CFS* then uses Algorithm 2 to determine which combination is consistent with the observation (**line 6**) and provides the corresponding status of the remaining links of the cycles.

In cases where there exists multiple consistent failure sets, *CFS* requires a local inspection of a link which appears with a different status in any two solutions to determine which solution is consistent with the result of the local inspection.

Observation 1. *Assuming the grey area becomes a tree by removing* C *edges, CFS algorithm runs in* $O(2^C|E_p^{U,t}|)$*.*

Proof of Observation 1. We can find out the status of all grey links which are not within a cycle in $O(|E_p^{U,t}|)$ according to Lemma 2.

If there are cycles of grey links, we break the cycles by selecting C links. By making an assumption on the status of each of these C links *CFS* builds a decision tree, where at each node it assumes whether each link is either working or broken. Such a decision tree will have 2^C leaves corresponding to different failure sets whose consistency is validated separately by means of Algorithm 1, which will also provide the state of the remaining grey links in the broken cycles, in the case of consistency.

Note that the consistent failure set is not unique and sometimes we may end up having multiple failure sets, which are all consistent with our partial knowledge. □

Example. *Figure 5 shows an example of a network with 6 grey links and shows different steps of the CFS algorithm. In the first step, we evaluate the status of all grey links, which are the unique grey incident links of a node and therefore are not part of a cycle.*

In the second step, we evaluate the status of grey links forming a cycle. For link (23) *we build a decision tree to remove the existing cycle. The decision tree rooted in the link* (23) *will have two branches, corresponding to the two potential different state of link* (23)*, working or not working. Assuming edge* $(23) \in E_p^{B,t}$ *was broken, we do not find a consistent feasible solution, which lead us to evaluate the other branch,*

Algorithm 1: Consistent Failure Set (*CFS*) algorithm.

Data: A set of grey lines $(ij) \in E_p^{U,t}$ whose failure status is unknown, the graph of the network $G_p = (V_p^t, E_p^t)$, the power generated at each generator $P_{G_i}\ \forall G_i \in V_p^t$, the power consumed at each load $P_{L_i}\ \forall L_i \in V_p^t$

Result: The status of edges in the grey area $(ij) \in E_p^{U,t}$, which can be failed or working.

$C =$ `Number of edges in` $E_p^{U,t}$ `that need to be removed to make the grey area cycle-free`
if $C = 0$ **then**
  run CFS-Cycle-Free($E_p^{U,t}, G_p, P_{G_i}, P_{L_i}$).
else if $C > 0$ **then**
  pick an edge at each cycle to generate a cycle-free grey area
  for all 2^C combination of the chosen edges at each cycle, run CFS-Cycle-Free($E_p^{U,t}, G_p, P_{G_i}, P_{L_i}$) to find a consistent failure set
return $E_p^{B,t}, E_p^{W,t}$

assuming that (23) *is working, namely* $(23) \in E_p^{W,t}$*. The last graph shows a consistent failure set of broken and working edges. In this example, we derived only one consistent failure set. If we had multiple consistent failure sets, we would have performed a local inspection of the edges whose failure status is different in the possible consistent solutions, and picked the solution consistent with the local inspection.*

Algorithm 2: CFS-Cycle-Free

Function *CFS-Cycle-Free* ($E_p^{U,t}, G_p, P_{G_i}, P_{L_i}$)
  $greys = argmin|(n_i j) \in E_p^{U,t}|$;
  while $greys = 1$ **do**
    Select a node $i \in V_p^t$ with one grey neighbor $greys = argmin|(ij) \in E_p^{U,t}|$;
    detect whether $(ij) \in E_p^{U,t}$ is working or not using equation 2.;
    if *there exists no solution from equation 2* **then**
      return `INCONSISTENT`;
      break ;
    if $(ij) \in E_p^{U,t}$ *is working* **then**
      $E_p^{W,t} = E_p^{W,t} \cup (ij)$ and $E_p^{U,t} = E_p^{U,t} \setminus (ij)$;
    else
      $E_p^{B,t} = E_p^{B,t} \cup (ij)$ and $E_p^{U,t} = E_p^{U,t} \setminus (ij)$;
  return `CONSISTENT`, $E_p^{B,t}, E_p^{W,t}$;

Remark 2 (Discussion on the number of grey edges.)**.** Table II shows the average number of grey edges which are part of a cycle in the graph of the Italian power grid network [3, 34, 40] when the size of the disrupted communication network (GARR) increases from 10% to 90% for 100 different random selection of disrupted communication nodes. Assuming all possible failures within a cycle are consistent with known information, we only need a maximum of 10% local inspection of the grey edges.

C. Max-R-Greedy

In this paragraph we introduce the baseline heuristic *Max-R-Greedy* to solve the *Max-R* problem. We recall that the objective function of the problem *Max-R* is the total accumulative flow in K stages of recovery. *Max-R-Greedy* greedily selects the links to repair based on the marginal value of their contribution to the accumulative flow.

TABLE II: Average number of local inspections needed as the size of the grey area increases in the Italian power grid network.

Percentage of disrupted monitors	Average number of grey edges in the Italian power grid	Average # of grey edges within a cycle
10	25.25	3.74
20	62.49	7.15
30	92.06	10.04
40	124.16	13.35
50	157.6	16.84
60	193.29	21.52
70	227.59	26.95
80	265.49	32.71
90	303.5	40.06

Algorithm 3: *Max-R-Greedy* recovery algorithm.

Data: A set of failed lines $(ij) \in E_p^{B,t}$, A set of demand loads $L_j \in V_p$ and generators $G_i \in V_p$, limit on the tolerable power of each transmission line F_{ij}^{max}, the nodal admittance matrix B, the required resources to repair each line r_{ij}

Result: The recovery schedule of the failed transmission lines $\delta_{(ij),k}$

1: $R = 0$
2: **for** $k \in \{1, ..., K\}$ **do**
3: $R = R + R_k$
4: **while** $\exists (ij) \in E_p^{B,t}$ that $r_{ij} \leqslant R$ **do**
5: Select an un-repaired line $(ij)^* = argmax_{ij} \frac{F_{(ij)}}{r_{(ij)}}$
6: $\delta_{(ij)^*,k} = 1$
7: $R = R - r_{(ij)^*}$
8: **return** $\delta_{(ij)^*,k}$

At each stage k, *Max-R-Greedy* repairs the transmission lines that add the maximum to the total delivered power over the required resource, i.e. it repairs the line arg $\max_{(ij)\in E_p^{B,k}}(F_{ij}/r_{ij})$ among the broken ones, until the total available resources for stage k are used. Algorithm 3 shows different stages of the *Max-R-Greedy* algorithm.

More in details, the algorithm works iterating repair interventions through stages. At each stage k, it first updates the current value of the rollover budget (**line 3**). Then, until the available budget allows, it selects a new broken link $(i, j)^*$ for recovery, based on the marginal contribution to the accumulative flow, with respect to the repair cost (**line 5**). Finally it schedules $(i, j)^*$ for recovery at stage k (**line 6**) and updates the available recovery resources accordingly (**line 7**).

D. Max-R-Backward

As an alternative heuristic to compute a more accurate solution of the *Max-R* problem, we use *Max-R-Backward*. The algorithm finds the recovery schedule in K stages as a list of K sets $\mathcal{S}_k,\ k = 1, \ldots, K$, of links to be repaired up to stage k. The algorithms works backward from stage K to stage 1, by considering a decreasing budget at each stage. It starts by solving a version of the problem which assumes $R = R_1 + ... + R_K$ resources are available (**line 4**). The solution of this execution of the algorithm produces a list of edges to be recovered from the start until stage K, named $\mathcal{S}_K$. Initially, $\mathcal{S}_K$ is the entire set of broken links E_p^B. In order to calculate the recovery schedule for each stage, the algorithm proceeds by selecting from $\mathcal{S}_K$ all the lines that exceed the budget available for the first $K - 1$ stages, and which contribute the minimum marginal value of flow over cost (**line 7**) . The selected lines

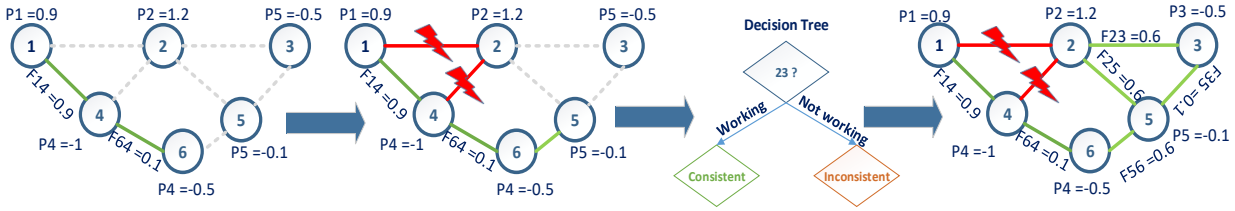


Fig. 5: An example of a 6-bus network with 6 grey edges and different steps of *CFS* algorithm.

Algorithm 4: *Max-R-Backward* recovery algorithm.

Data: A set of failed lines $(ij) \in E_p^{B,t}$, A set of demand loads $L_j \in V_p$ and generators $G_i \in V_p$, limit on the tolerable power of each transmission line F_{ij}^{max}, the nodal admittance matrix Y^t, the required resources to repair each line r_{ij}

Result: The recovery schedule of the failed transmission lines $\delta_{(ij),k}$

1: solve DC power flow model to find F_{ij}, assuming all lines are working
2: $\mathcal{S}_{K+1} = E_p^B$
3: **for** $k = K$ downto $k = 1$ **do**
4: $\quad R = \sum_{m=1}^{k} R_m$
5: $\quad \mathcal{S}_k = \mathcal{S}_{k+1}$
6: $\quad$ **while** $\sum_{(ij)\in\mathcal{S}_k} r_{ij} > R$ **do**
7: $\quad\quad$ Select a line with minimum flow per cost $(ij)^* = argmin_{ij} \frac{F_{(ij)}}{r_{(ij)}}$
8: $\quad\quad \delta_{(ij)^*,k+1} = 1$
9: $\quad\quad \mathcal{S}_k = \mathcal{S}_k \setminus (ij)^*$
10: $\quad$ solve DC power flow model to find F_{ij}, assuming $(ij) \in \mathcal{S}_k$ are working.
11: **return** $\delta_{(ij)^*,k}$

will be scheduled for recovery at stage K (**line 8**), and will be excluded from the recovery schedule of any previous stage (**line 9**). Before the end of each stage, the algorithm must solve the DC power flow problem taking account of the scheduled repairs, to update the values of the flows F_{ij}, for all the links of E_p (**line 10**).

This procedure repeats until the repair schedule of all stages is found.

VI. Evaluation

In this section, we perform an experimental evaluation of our algorithms in a real network setting. We consider the Italian power grid network, called HVIET, shown in Figure 6a consisting of 310 nodes, 113 generators and 97 demand loads. The network has 361 power lines. For the communication network we use the GARR network, shown in Figure 6b consisting of 39 nodes and 50 edges [34, 40]. Transmission lines in the power grid are monitored by several sensors deployed nearby. Based on the interdependency model, the aggregated data are then sent to the closest/random communication node and to the control center. In addition, the control commands for power adjustment are sent to the closest/random communication node. Therefore, each node in the communication network might monitor and control several lines and nodes in the power grid. In our evaluation, we assume the Italian power grid network to be purely inductive (lossless) with zero reactive injection, so that the DC power flow is actually accurate. We also note that, since the DC power flow model is an approximation to the AC power flow, applying our model to a real coupled system, can result in a lower performance. We implement our cascade prevention and recovery algorithms in Python and used the Gurobi optimization toolkit [41], on a 120-core, 2.5 GHz, 4TB RAM cluster.

In our experiments we vary the interdependency model and we randomize the results running 10 different trials with randomized selection of failed transmission lines.

Summary of observations. The key observations are as follows. First, we observed that the random interdependency model has a larger impact on the number of disrupted/uncontrollable nodes in the two networks compared to a one-way and a location-based interdependency model. This is due to the fact that in one-way and location-based interdependency model, the cascade is limited to the geographical interdependent regions in the two network while in a random interdependency model the cascade spreads from any part of the network to the other parts.

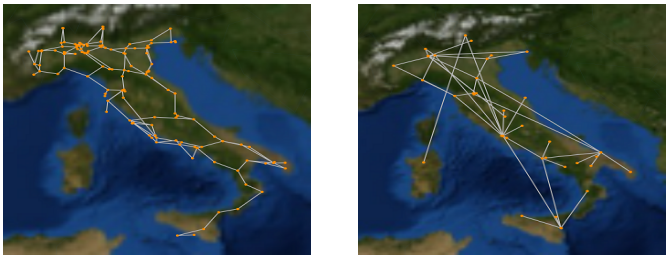
Second, we observed that our cascade prevention approach could still provide service, though potentially degraded, differently from the case without countermeasures.

Finally, we observed that the backward recovery approach performs better than the greedy approach. This is due to the fact that the greedy approach does not consider the correlation between different steps of the recovery approach and makes a repair decision at each stage independently.

A. Impact of Interdependency

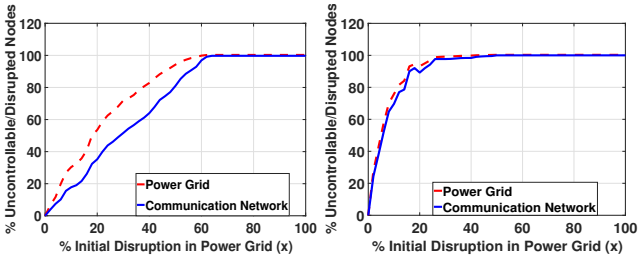
We first evaluate the impact of different interdependency models on the extent of failure propagation within the power grid and the communication network.

We assume the initial failures are in the power grid. In this experiment we gradually increase the percentage x of randomly failed lines in the power grid and we observe how the number of uncontrollable power grid nodes and disrupted communication node varies accordingly. In the first scenario, we connect the HVIET and GARR networks following the location-based interdependency model described in Section III-B. Figure 7a shows the percentage of uncontrollable nodes in the power grid and disrupted nodes in the communication network which lose power due to the failures in the power grid, when the propagation stops naturally. We observe that in a location-based interdependency model the failure does not spread much and the percentage of failed power grid nodes and communication nodes is linear with respect to the initial failures in the power grid.



(a) HVIET. (b) GARR.

Fig. 6: a) The Italian high-voltage (380 kV) transmission grid (HVIET), and b) its communication network (GARR).



(a) Location-based. (b) Random.

Fig. 7: a) Location-based and b) Random interdependency model in the Italian power grid (HVIET) and its communication network (GARR).

In the second scenario, we consider the random interdependency model. Figure 7b shows the simulation results for this scenario. As shown, compared to location-based interdependency model, in a random interdependency model, the failures spread more in both the power grid and the communication network. This is due to the fact that in a random interdependency model, the failures are not limited to geographical interdependent regions and can spread from any part of the network to the other. Hence, we consider this model as a stress-test for our algorithms.

B. Impact of incomplete knowledge

In this section, we investigate the impact of incomplete knowledge of the exact location of failures. Initially, x% of the communication nodes get disrupted. We consider the one-way interdependency model where the consequences of failures in the communication network are lack of information and loss of controllability in the power grid. Figure 8 shows the simulation results of this experiments. When 100% of the communication network and 20% of the power grid is disrupted, the total delivered power can drop by 10.48 power units (pu). Assuming the maximum unitary profit of 26.6 €$/MW$ according to [42], the total profit loss, due to uncertainty of failure location can be as high as 209076 € $= 10.48pu \times 750MW/pu \times 26.6$ €$/MW$ which could be avoided using a detection algorithm and a cascade prevention approach.

C. Preventing the cascade (Min-CFA)

In this paragraph we evaluate the performance of our cascade prevention approach, namely the *Min-CFA* algorithm, with the case in which no cascade countermeasure is available.

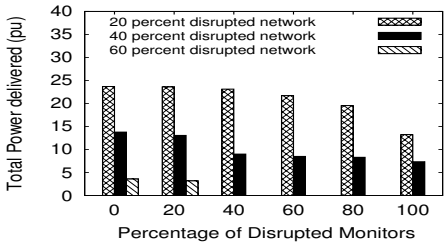


Fig. 8: Total delivered power (pu) versus the percentage of failures on the monitoring nodes in the Italian power grid network.

We consider the one-way interdependency model described in Section III-B. We assume the communication network gets power from an external source in case of a failure in the power grid. The disruption in the communication network is observed as lack of knowledge and uncontrollability in the power grid.

The performance metric considered in this experiment is the total delivered demand power.

Similar to [20], we assume all loads have the same priority and give a high penalty for not being able to satisfy the demand. We assume the weighted cost of decreasing power of load L_j is 100, i.e. $w_{L_j} = 100, \quad \forall L_j \in L$, while the normalized weighted cost of generators is 1, i.e. $w_{G_i} = 1, \quad \forall G_i \in G$.

In the first set of simulations we set the disruption percentage of the power grid to $x = 60\%$ and run *Min-CFA* to find the optimal flow assignment. Figure 9 shows the total delivered power during different time steps of the algorithms with *Min-CFA* cascade prevention and without it. As shown in the figure, *Min-CFA* can save 54% of the total power that would be delivered if the power grid were not disrupted. On the other hand, if we do not run a cascade prevention algorithm, the failed transmission lines lead to more lines failing and this process continues until the whole system fails.

In the next set of simulations, we use a continuous cascade prevention, meaning that the decision variable, $P^t_{L_j}$ in Equation 4 can decrease continuously. Then, we consider a discrete cascade prevention scenario, where the decision variable, $P^t_{L_j}$ in Equation 4 can either be equal to each load's demand power which should be satisfied or set to zero (i.e., loads are turned off); and finally, we consider a scenario, where there is no monitoring technique to reschedule the power flow or avoid the cascade and the failed transmission lines can trigger multiple cascade.

In this experiment we gradually increase the percentage x of randomly failed lines in the power grid and observe the total amount of load served. Figure 10 shows the simulation results for the three cases. As shown, the continuous cascade prevention approach saves more power compared to the discrete power optimization and to the case without cascade prevention.

Notice that in absence of cascade prevention measures, an initial failure that disrupts more than 60% of the power grid lines is sufficient to make a black out of the entire system, due to a full propagation of the failure.

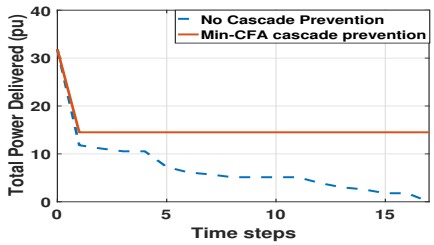


Fig. 9: Total delivered power (pu) during time when we use *Min-CFA* cascade prevention algorithm and without any cascade prevention in the Italian power grid network.

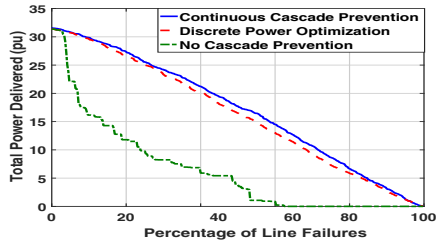


Fig. 10: Total delivered power (pu) versus the percentage of line failures in the Italian power grid.

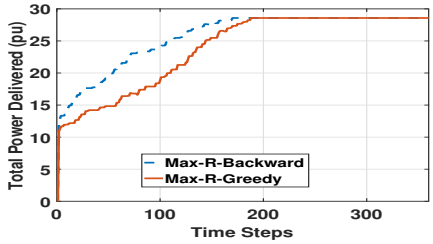


Fig. 11: Total delivered power (pu) flow over time for *Max-R-Backward* and *Max-R-Greedy* in the Italian power grid.

D. Recovery phase (Max-R)

In the next set of experiments, we compare the recovery performance of the proposed heuristics (*Max-R-Greedy* and *Max-R-Backward*). Figure 11 shows the total delivered power flow over different stages of progressive recovery intervention, when using the two algorithms. As shown, the greedy approach does not consider the correlation between different steps of the recovery approach and tries to maximize the added flow at each iteration step. On the other hand, the backward algorithm solves the problem using all repair resources in the beginning and removes the repair edges with less profit ($F_{ij}/r_{(ij)}$) from the schedule of previous stage until all repair schedules are determined. Therefore, *Max-R-Backward* performs better than the *Max-R-Greedy* approach with larger total area behind the curve in Figure 11. We next increase the number of resources at each stage and study the normalized accumulative delivered power for the two recovery approaches. Table III shows the results of this scenario.

TABLE III: Normalized accumulative delivered power for *Max-R-Backward* and *Max-R-Greedy* approaches.

Recovery Resources	*Max-R-Backward*	*Max-R-Greedy*
1	0.8062	0.6791
2	0.9012	0.8377
6	0.9645	0.9435

VII. Conclusion and Future Directions

This paper studies the problem of mitigating propagating failures and performing progressive recovery interventions to restore the functionality of an interdependent power grid and communication network, under incomplete localization of failures. We formulate an optimization problem to stop the cascading failures and, due to high complexity of the recovery problem, we propose two heuristic approaches (i) a baseline greedy and (ii) a backward heuristic, to restore the power grid functionality. By means of extensive simulations, we show that since the backward algorithm takes account of the combined impact of repaired component, it significantly outperforms the baseline recovery algorithm in terms of accumulative delivered power. Our detection mechanism and recovery approach with incomplete information opens up the avenues of new research on improving power grid reliability and resiliency under incomplete or noisy information. Future research directions include: (1) more accurate power grid representation (e.g. AC power flow model), (2) more accurate communication network models for SCADA (e.g. including power-line carrier) and WAMPAC, and (3) mitigating the disruptions to electric power grids caused by malicious attacks.

References

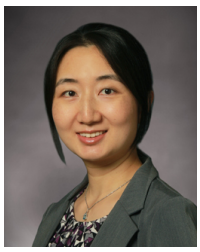
[1] D. Z. Tootaghaj, N. Bartolini, H. Khamfroush, and T. La Porta, "Controlling cascading failures in interdependent networks under incomplete knowledge," in *IEEE SRDS*, 2017.
[2] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature physics*, 2012.
[3] M. Parandehgheibi, E. Modiano, and D. Hay, "Mitigating cascading failures in interdependent power grids and communication networks," in *IEEE SmartGridComm*, 2014.
[4] D. Bienstock, *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*. SIAM, 2015.
[5] N. Bartolini, S. Ciavarella, T. F. La Porta, and S. Silvestri, "Network recovery after massive failures," in *IEEE/IFIP DSN*, 2016.
[6] K. Al Sabeh, M. Tornatore, and F. Dikbiyik, "Progressive network recovery in optical core networks," in *IEEE RNDM*, 2015.
[7] D. Z. Tootaghaj, H. Khamfroush, N. Bartolini, S. Ciavarella, S. Hayes, and T. La Porta, "Network recovery from massive failures under uncertain knowledge of damages," in *IFIP Networking*, 2017.
[8] J. Wang, C. Qiao, and H. Yu, "On progressive network recovery after a major disruption," in *IEEE INFOCOM*, 2011.
[9] S. Ciavarella, N. Bartolini, H. Khamfroush, and T. La Porta, "Progressive damage assessment and network recovery after massive failures," in *IEEE INFOCOM*, 2017.
[10] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, 2010.
[11] H. Khamfroush, N. Bartolini, T. La Porta, A. Swami, and J. Dillman, "On propagation of phenomena in interdependent networks," *IEEE Transactions on network science and engineering*, 2016.
[12] L. M. Shekhtman, M. M. Danziger, and S. Havlin, "Recent advances on failure and recovery in networks of networks," *Chaos, Solitons & Fractals*, 2016.
[13] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 048701, 2010.
[14] M. Parandehgheibi and E. Modiano, "Robustness of interdependent networks: The case of communication networks and the power grid," in *IEEE GLOBECOM*, 2013.

[15] A. Sen, A. Mazumder, J. Banerjee, A. Das, and R. Compton, "Identification of k most vulnerable nodes in multi-layered network using a new model of interdependency," in *IEEE INFOCOM WKSHPS)*, 2014.
[16] A. Das, J. Banerjee, and A. Sen, "Root cause analysis of failures in interdependent power-communication networks," in *IEEE MILCOM*, 2014.
[17] A. Das, C. Zhou, J. Banerjee, A. Sen, and L. Greenwald, "On the smallest pseudo target set identification problem for targeted attack on interdependent power-communication networks," in *IEEE MILCOM*, 2015.
[18] D. Z. Tootaghaj, T. He, and T. La Porta, "Parsimonious tomography: Optimizing cost-identifiability trade-off for probing-based network monitoring," *ACM SIGMETRICS Performance Evaluation Review*, 2018.
[19] E. Chatziafratis, Y. Zhang, and O. Yagan, "On the robustness of power systems: optimal load-capacity distributions and hardness of attacking," in *ITA*, 2016.
[20] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos: An interdisciplinary journal of nonlinear science*, 2002.
[21] M. J. Eppstein and P. D. Hines, "A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Transactions on Power Systems*, 2012.
[22] N. Kayastha, D. Niyato, E. Hossain, and Z. Han, "Smart grid sensor data collection, communication, and networking: a tutorial," *Wireless communications and mobile computing*, 2014.
[23] D. Pendarakis, N. Shrivastava, Z. Liu, and R. Ambrosio, "Information aggregation and optimized actuation in sensor networks: enabling smart electrical grids," in *IEEE INFOCOM*, 2007.
[24] J. Bertsch, C. Carnal, D. Karlson, J. McDaniel, and K. Vu, "Wide-area protection and power system utilization," *Proceedings of the IEEE*, 2005.
[25] M. Begovic, D. Novosel, D. Karlsson, C. Henville, and G. Michel, "Wide-area protection and emergency control," *Proceedings of the IEEE*, 2005.
[26] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber–physical system security for the electric power grid," *Proceedings of the IEEE*, 2012.
[27] S. Soltan, M. Yannakakis, and G. Zussman, "Joint cyber and physical attacks on power grids: Graph theoretical approaches for information recovery," in *ACM SIGMETRICS*, 2015.
[28] S. Soltan, D. Mazauric, and G. Zussman, "Analysis of failures in power grids," *IEEE Transactions on Control of Network Systems*, 2017.
[29] S. Soltan and G. Zussman, "Power grid state estimation after a cyber-physical attack under the ac power flow model," *Proc. IEEE PES-GM*, vol. 17, 2017.
[30] D. Bienstock, *Electrical Transmission System Cascades and Vulnerability: An Operations Research Viewpoint*. SIAM, 2016, vol. 22.
[31] J. Shao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Cascade of failures in coupled network systems with multiple support-dependence relations," *Physical Review E*, 2011.
[32] Z. Huang, C. Wang, S. Ruj, M. Stojmenovic, and A. Nayak, "Modeling cascading failures in smart power grid using interdependent complex networks and percolation theory," in *IEEE ICIEA*, 2013.
[33] X. Huang, J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of interdependent networks under targeted attack," *Physical Review E*, 2011.
[34] V. Rosato, L. Issacharoff, G. Gianese, and S. Bologna, "Influence of the topology on the power flux of the italian high-voltage electrical network," *arXiv preprint arXiv:0909.1664*, 2009.
[35] A. M. Kettner and M. Paolone, "On the properties of the power systems nodal admittance matrix," *arXiv preprint arXiv:1702.07235*, 2017.
[36] J. G. Oxley, *Matroid theory*. Oxford University Press, USA, 2006, vol. 3.
[37] G. Andersson, "Modelling and analysis of electric power systems," *EEH-Power Systems Laboratory, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland*, 2004.
[38] N. Biggs, *Algebraic graph theory*. Cambridge university press, 1993.
[39] Y. Chen, D. Bindel, H. Song, and R. H. Katz, "An algebraic approach to practical and scalable overlay network monitoring," in *ACM SIGCOMM Computer Communication Review*. ACM, 2004.
[40] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, and R. Setola, "Modelling interdependent infrastructures using interacting dynamical models," *International Journal of Critical Infrastructures*, 2008.
[41] "Gurobi," `http://www.gurobi.com/`, accessed in 2017.
[42] M. Dicorato, A. Minoia, R. Sbrizzai, and M. Trovato, "A simulation tool for studying the day-ahead energy market: the case of italy," in *Power Engineering Society Winter Meeting*. IEEE, 2002.

Diman Zad Tootaghaj is a Ph.D. student in the department of computer science and engineering at the Pennsylvania State University. She received B.S. and M.S. degrees in Electrical Engineering from Sharif University of Technology, Iran in 2008 and 2011 and a M.S. degree in Computer Science and Engineering from the Pennsylvania State University in 2015. Her current research interests include computer network, recovery approaches, distributed systems, and stochastic analysis.

Novella Bartolini received the graduate degree with honors and the PhD degree in computer engineering from the University of Rome, Italy, in 1997 and 2001, respectively. She is now associate professor with Sapienza University of Rome and visiting professor with Penn State University, since 2014. She was program chair and program committee member of several international conferences. She has served on the editorial board of Elsevier the Computer Networks and the ACM Wireless Networks.

Hana Khamfroush is an assistant professor in the department of Computer Science, University of Kentucky. Before joining University of Kentucky, she was a research associate in the Department of Computer Science and Engineering, Penn State University. She received her PhD degree with honor in telecommunication engineering from the University of Porto, Portugal. She was named a rising star in EECS by MIT and CMU in 2015 and 2016. She was selected as one of the 200 young researchers for the Heidelberg Laureate Forum, Germany.



Ting He received the B.S. degree in computer science from Peking University, China, in 2003, and the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2007. From 2007 to 2016, she was a Research Staff Member with the Network Analytics Research Group, IBM T. J. Watson Research Center, Yorktown Heights, NY, USA. She is currently an Associate Professor with the School of Electrical Engineering and Computer Science, The Pennsylvania State University, University Park, PA, USA.

Nilanjan Ray Chaudhuri received the Ph.D. degree in power systems from Imperial College London, London, U.K., in 2011. He is a Member of the IEEE Power and Energy Society. He holds seven US patents and three European Patents with numerous pending applications. He is the lead author of the book Multi-Terminal Direct Current Grids: Modeling, Analysis, and Control (Wiley/IEEE Press, 2014), and an Associate Editor of the IEEE TRANSACTIONS ON POWER DELIVERY. He received the National Science Foundation CAREER Award in 2016.

Thomas La Porta is the Wlliam E Leonhard chair professor in the Department of Computer Science and Engineering, Penn State University. He is the director of the Institute for Networking and Security Research. Prior to joining Penn State, he was with Bell Laboratories, where he was director of the Mobile Networking Research Department. He received the Bell Labs Distinguished Technical Staff Award, and an Eta Kappa Nu Outstanding Young Electrical Engineer Award in 1996. He also won Thomas Alva Edison Patent Awards in 2005 and 2009. He was the founding editor-in-chief of the IEEE Transactions on Mobile Computing, and served as editor-in-chief of the IEEE Personal Communications Magazine. He is a fellow of the IEEE and the Bell Labs.