# Recovery of Infrastructure Networks via Importance-based Multicentric Percolation Processes

Bowen Fu
*Graduate Student, Dept. of Civil and Environmental Engineering, Rice University, Houston, USA*

Leonardo Dueñas-Osorio
*Associate Professor, Dept. of Civil and Environmental Engineering, Rice University, Houston, USA*

ABSTRACT: Recovery processes across infrastructure systems after disasters are critical to improve their resilience, yet poorly understood. The common assumption of prioritizing the size of the reconnected network as the goal for recovery in many algorithms today is impractical, given that satisfaction of demands is more important for the functional recovery of infrastructure systems such as power grids. Mixed-integer programming formulations that guarantee optimality under practical resource and time constraints continue advancing, but become computationally intractable even for systems with only hundreds of elements. Algorithms approximating the optimal solution with lower computational cost are in need, including competitive percolation or surrogate models. We propose a method based on statistical mechanics that exhibits phase transitions, as when restoring networked systems. Our importance-based multicentric percolation recovery strategy for spatially distributed engineered networks, approximates optimal restoration solutions with a substantially lower computational cost. Small tree-like clusters form first in the network, which then interconnect into bigger components gradually mirroring optimal restoration and aligning with field practices. A key observation is that the formation of large connected components is suppressed during the recovery process, which enables balancing computational efficiency and accuracy. The proposed strategy is very close to optimization-based methods and methods based on competitive percolation, particularly when load is homogeneous and the fraction of generators is small; illustrative examples showcase the adequate trade-off between computation cost and accuracy relative to competing alternatives.

Catastrophic disasters, such as earthquakes and hurricanes, put modern society at tremendous risk of losing resources and lives if recovery is not expedient. Timely recovery from catastrophic damage is thus crucial to communities, including their lifelines such as transportation, water, power, and communications, as well as their overall well-being. Particular to engineered systems, lifelines usually connect or affect each other, shaping them into truly interdependent networks.

The functions of many networked systems rely on connectedness or the existence of a giant connected component (GCC), which is a subset of the entire network that spans its global scale. Examples of this object exist across transportation systems (Li et al., 2015) and communication networks (Glauche et al., 2003), among others. Hence, many proposed topology-based methods aim at recovering systems by firstly forming a GCC (Hu et al., 2016; Shang, 2016). However, for some other systems, such as water distribution networks and power grids, the principal goal is not to form a GCC, but to satisfy demands with supplies in the absence of a dominant component, except at local scales (Rudnick et al., 2011). In such systems, customers can be served if they are connected to some local sources, such as the 'islanding' technique in power grids (Dueñas-Osorio and Kwasinski, 2012; Panteli et al., 2016) and sectorization in water distribution networks (Di Nardo et al., 2013).

In commodity transport systems, commodity deficit usually costs the most. Optimization-based recovery strategies, such as the time-dependent interdependent network design problem (td-INDP) (González et al., 2016a,b) guarantee the optimality of the recovery process, but fail on systems of practical size, due to their expensive computational complexity (at least NP-complete). A recent decentralized recovery method based on competitive percolation (Smith et al., 2018) approximates the optimal solutions well, with a much cheaper computational cost. It selectively restores components, based on their contributions to reducing commodity deficit. Results from leading methods, such as INDP and competitive percolation, indicate that when demand satisfaction becomes the primary concern of the recovery process, the formation of a GCC is suppressed until demands are mostly satisfied locally. This is consistent with field constraints, as both, awareness of the overall situation, and distant transportation of resources, are usually not available at the early stages of system restoration.

To recognize limited global information and difficulty in resource mobilization, we propose an importance-based multicentric (local) percolation recovery (IMPR) process to mimic the restoration of systems as seen in the field. This strategy approximates the solutions given by optimization-based methods, based on the principles of 'demand satisfaction first' and 'suppression of the GCC', which are common in both optimal recovery solutions and practical recovery experience. In the proposed IMPR process, we first partition the original network into several clusters with single supplier. Next, the importance of the links are evaluated by efficient load-weighted betweenness. Then, clusters are restored locally in a multicentric local percolation way according to the importance of links, before trees in a graph forest are connected into a spanning tree of the original network.

The rest of the paper is organized as follows: Section 1 describes the proposed (IMPR) process. Section 2 provides results of recovery experiments on actual power transmission grids in the United States. Section 3 concludes the paper and provides ideas for future research.

## 1. IMPORTANCE-BASED MULTICENTRIC LOCAL PERCOLATION RECOVERY PROCESS

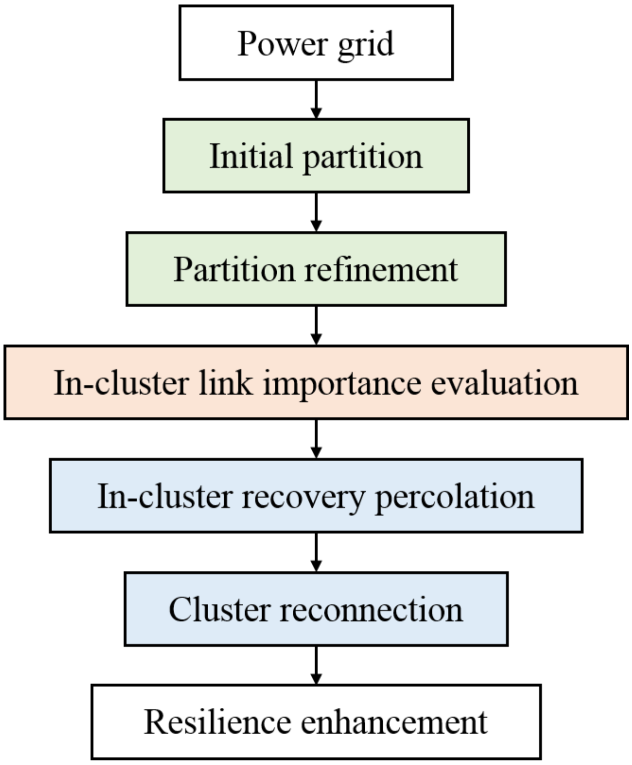


*Figure 1: Flowchart of proposed IMPR strategy.The green, yellow and blue boxes represent network partition, link importance evaluation and percolation recovery, respectively.*

The proposed IMPR strategy consists of three stages: network partition, link importance evaluation and percolation recovery (see Figure 1). We perform an example simulation on the topology of the power transmission grid in Amarillo, TX (47 nodes and 62 links as in Figure 2). The four large nodes in blue are generators assigned manually. Generators are assumed to have the same capacity, and the demands of consumers vary uniformly in [8,12], as a generic example. Performance comparison under different levels of load variations are shown later. This unitless choice reflects the fact that systems are sensitive to the *relative* demand variation.

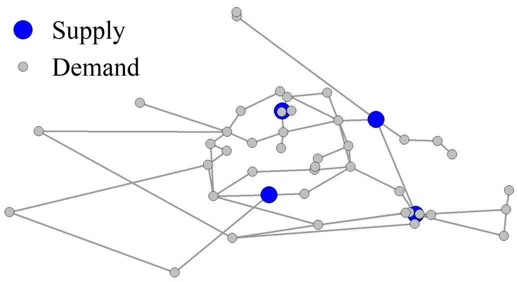


*Figure 2: Power transmission network topology of Amarillo, TX.*

### *1.1. Network partition*

#### *1.1.1. Tails contraction*

We contract the tails in the network before partitioning it. Tails are sequences of degree-two nodes with a degree-one extremity and a degree greater than two for the other. Figure 3 shows an example of tail contraction. The tail in the green box on the left is contracted into a single meganode E' on the right. The demand of meganode E' equals to the sum of demands of node E, F, G. Contracting the tails facilitates the network partition. Note that when a tail contains suppliers, only the outer part of the tail is contracted to the outermost suppliers.

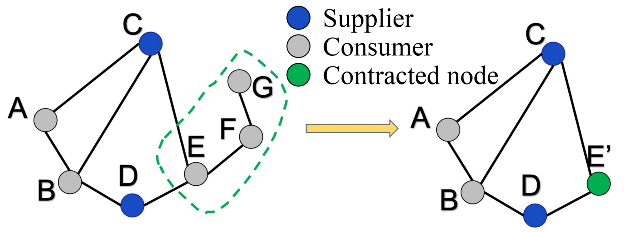


*Figure 3: Contraction of tails. The tail consist of nodes E, F and G, contracted into a single meganode E'.*

#### *1.1.2. Cluster growth*

The cluster growth is the first step to assign demand nodes to suppliers. It is completed in the following steps:

1. Find the cluster with the smallest total demand in the cluster list as $C_m$. Stop when the cluster list is empty.
2. Find all the neighbor demand nodes of cluster $C_m$, and remove the ones with 'too big' demand. The demand node is 'too big' if the total demand of cluster $C_m$ after the demand node is assigned to the cluster, is greater than the capacity of the supplier in the cluster.
3. If there is no eligible neighbor demand node, remove the cluster from the cluster list and go back to 1. Or find the demand node with the largest demand $L_m$.
4. If $C_m$ is the closest cluster(s) to $L_m$, assign $L_m$ to $C_m$, update the total demand of $C_m$, and go back to 1. If not, remove $L_m$ temporarily and find the next largest demand node.
5. If $C_m$ is the closest cluster to none of the demand nodes in step 4, find the demand node having farthest distance to its closest cluster, assign it to $C_m$, update the total demand of $C_m$, and go back to 1.

During the above cluster growth process, both of the connectivity of clusters and the total demand limit are ensured (so as to not exceed the capacity of the supplier).

#### *1.1.3. Remaining node assignment*

After the cluster growth, demand nodes may still be unassigned. We iteratively assign the demand nodes with the largest demand to the nearest suppliers among all the clusters its neighbors come from. This is to promote that large demand nodes are assigned to the cluster closest to them at least in the initial partition stage. The larger the demand of a node is, the faster we want to connect it to a cluster, to reduce commodity deficit as much as possible. So it will be better if large demand nodes are close to suppliers.

After we assign remaining nodes, each demand node is linked to a connected cluster, where there is only one supplier. However, the total demand of clusters might be significantly changed in this step, so we refine the initial partition in the next stage by a heuristic node exchanging algorithm, to make the supply and demand balance properly in every cluster.

### *1.2. Partition refinement*

After the initial partition of the network, the distribution of total demand of clusters can be uneven. Some clusters may have disproportionately high total demands, yet others may have insufficient demand. Refinement of the initial partition is hence necessary to make the distribution of total demands of clusters more uniform.

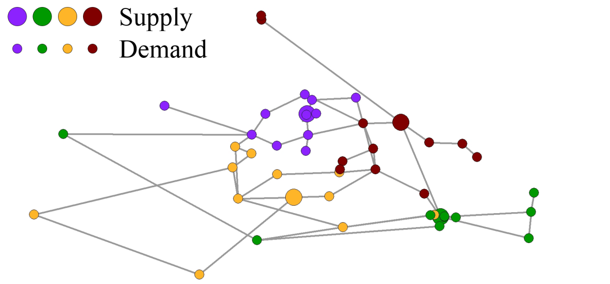


*Figure 4: Clusters after network partition (Section 1.2)*

The partition refinement is done on the cluster graph, where each cluster in the original network is represented with a node. A directed link exists from cluster A to cluster B, when some demand nodes in cluster A is connected to any node in cluster B. Thus, a directed link means possible demand node exchange between clusters. The following steps are based on both, the cluster graph and the original network:

1. For each pair of clusters in the cluster graph, find the shortest path from the cluster with largest total demand to the closer with smallest total demand.
2. Transfer a demand node between clusters along the shortest path. At each transfer, the demand node with smallest demand is used. Record the network partition with the smallest average total demand.
3. Partition the network with the recorded network partition (smallest average total demand across clusters). If the recorded network partition is not superior, stop the algorithm. Otherwise, go back to 1.

After the partition refinement, the clusters are adjusted to have similar total demands, and each has an almost-balanced supply and demand.

An example partition is shown in Figure 4. Note that there is only one supplier in each cluster. The distribution of total demand is shown in Figure 5.

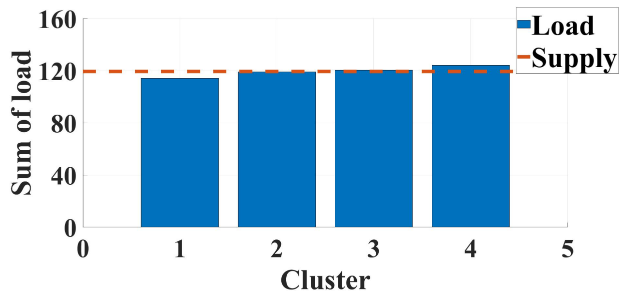


*Figure 5: Distribution of total demand across clusters.*

### *1.3. Link importance evaluation*

This step ranks the importance of links inside each cluster according to a demand-weighted betweenness, so as to prioritize for recovery in the next stage. Importance is calculated as follows:

1. Within each cluster, find all the shortest paths from the (only) supplier to all its demand nodes.
2. Multiply all shortest paths with their demand, which is the amount of 'traffic or flow packets' that all the links on the shortest path will gain, due to this demand node.
3. Sum up all the traffic that each link obtained from each demand node, as the importance of the link.

To maximally reduce the total demand deficit cost, we need to recover the consumers with large demands as early as possible, which is consistent with practice (Liu et al., 2016). Then links on the path from the supplier to large demand consumers should have relatively high recovery priority. However, connecting distant consumers with large demands to the suppliers requires the links on the path from the supplier to the consumers to be recovered first. The demand-weighted betweenness happens to meet both of the criteria, because it considers not only the distance of demand nodes from the supplier, but also how much demand will pass through each individual link. The higher the traffic on a link has, the more important the link is, and the higher priority it gets for the next stage. This mechanism aims to minimize the demand deficit cost by satisfying large demands early enough.

### 1.4. Multicentric percolation recovery

In this stage, all the clusters are recovered internally by forming a spanning tree in each cluster. At each step, find all in-cluster links connecting a failed node and the restored part within all clusters. Then, we randomly recover one of the in-cluster links, based on the probability distribution proportional to the importance of those links. Links with higher importance will be recovered earlier in a probabilistic perspective.

The multicentric percolation recovery mimics the early restoration in reality, when the network is mostly recovered locally to satisfy demands of consumers. Given that the importance of links are evaluated by demand-weighted betweenness, the links with higher demand 'traffic' will be restored earlier in probability, thus reducing total commodity deficit as much as possible. In this step, a spanning tree is recovered in each cluster, and all the trees forms a forest in the whole network.

### 1.5. Clusters re-connection

After all the local clusters are recovered internally , imbalance still exists since each cluster has unequal demand and supply. This imbalance can only be reduced by reconnecting clusters together to form a spanning component in the entire network. Links between clusters are selected to be recovered consecutively. To minimize demand deficit cost, at each step, only the between-cluster link that reduces the most overall imbalance is recovered.In this way, we link the sub-trees in the forest into a minimum spanning tree at the network level, achieve the balance of supply and demand over the whole network. The total number of links restored from the very beginning till then equals to $n-1$, where $n$ is the number of nodes.

## 2. EXPERIMENTS AND RESULTS

The performance of the proposed IMPR strategy is compared against competitive percolation recovery (CP), *iterative*-INDP (*i*-INDP), and random recovery (RR).

For the competitive percolation recovery, 20% of links in the network are selected as candidates, among which the one that yields maximum demand deficit reduction is recovered (Smith et al., 2018).

The *i*-INDP method is an iterative variation of *td*-INDP (González et al., 2016a). Instead of seeking a global optimal solution over the whole recovery process as in *td*-INDP, *i*-INDP iteratively finds the local optimal solution within $t$ time-steps to approximate to *td*-INDP. The solution by *i*-INDP approaches the optimal solution, and at the same time significantly reduces the computational cost. In this simulation, $t = 1$ is used, and the solution is equivalent to competitive percolation when at each step, the selection is performed among all of the links.

For the random recovery, one link is selected randomly from all the failed links to be recovered.

For the random recovery, one link is selected randomly from all the failed links to be recovered.

We perform different recovery process on the topology of the power transmission grid in Amarillo, TX in the United States (see Figure 2). We assume that the network suffers a catastrophic contingency, after which all the links in the network are damaged (the most challenging case for an optimizer). The comparison of component excess supply, total commodity deficit and size of GCC by different recovery strategies are shown in Figure 6, Figure 7 and Figure 8. Figure 6 shows the component excess supply during different recovery process. The excess supply of most components is reduced at the early stage when the proposed IMPR recovery is used, which is similar to *i*-INDP and competitive percolation. However, in the proposed IMPR strategy, there is hardly any component with big excess supply after the early stages, which is not the case in *i*-INDP and competitive percolation. This shows that *i*-INDP and competitive percolation optimize the recovery process based on total excess surplus without much attention to local dynamics. But the proposed IMPR also takes care of local situations. Figure 7 shows the total commodity deficit during recovery processes. The proposed IMPR strategy is hardly different from the competitive percolation and *i*-INDP. To compare the recovery performance, we define the cost ratio cr be-

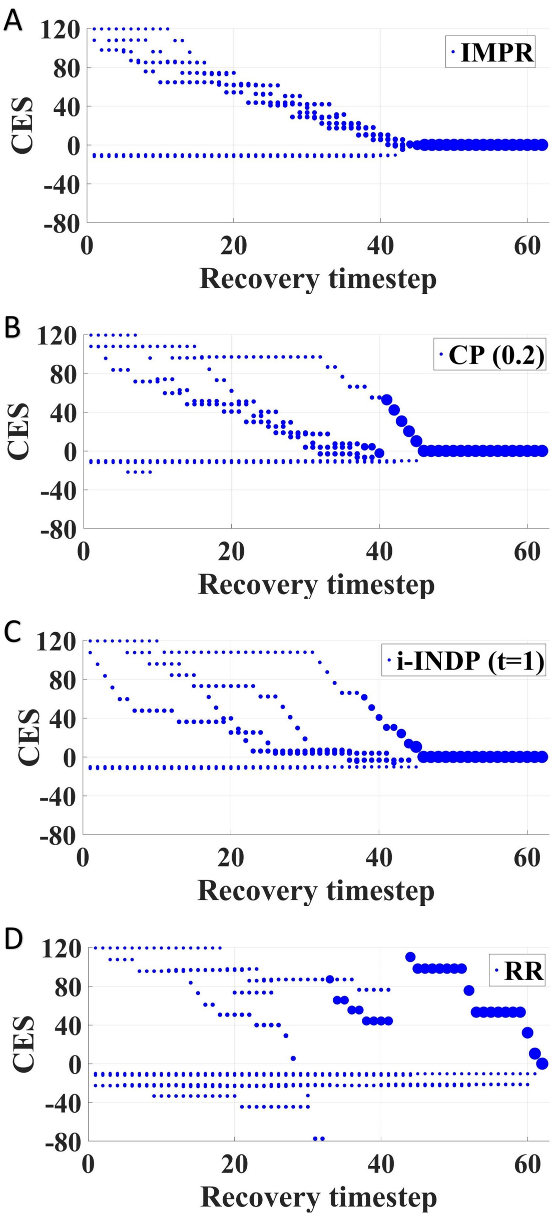


*Figure 6: Component excess supply (CES): (A) Proposed IMPR strategy. (B) Competitive percolation recovery (CP), where 20% of the total number of links are selected as candidates per step. (C) i-INDP with window length $t = 1$. (D) Random recovery (RR).*

tween two recovery processes as below:

$$cr = \frac{\int ts_I}{\int ts_{II}} \tag{1}$$

where $ts_I$ and $ts_{II}$ are the total commodity deficit function over time. The cost ratio between two recovery processes show how much more cost method $I$ induced, compared to method $II$. The

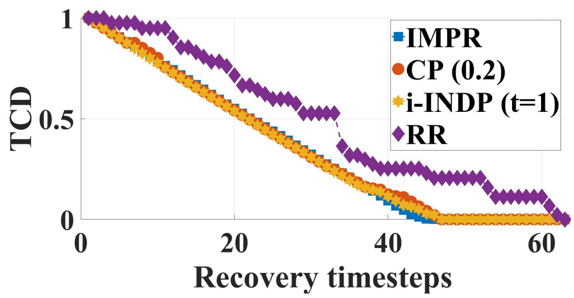


*Figure 7: Total commodity deficit (TCD) during recovery.*

closer $cr$ is to 1, the more similar the performance of the two recovery processes is. When $cr$ is smaller than 1, method $I$ is better than method $II$, and vice versa.

The cost ratio between IMPR and CP for this realization is 0.9925, slightly smaller than 1, indicating that the proposed IMPR strategy outperforms competitive percolation in this realization. The cost ratio between IMPR and $i$-INDP ($t = 1$) is 1.0098, meaning that the cost of IMPR is slightly higher (by 0.98%) than the time-constrained optimal cost in this realization. As is also seen in Figure 7, IMPR is significantly better than random recovery, with a $cr$ of 0.6974.

Recovery of the function relies on the recovery of topology. We find that the formation of the GCC is delayed in leading algorithms such as such as INDP and competitive percolation, corresponding to field practices such as islanding. We investigate the formation of GCC in Figure 8, showing the relationship between topology and function. The formation of the GCC is suppressed until the last 3 stages, later than both of competitive percolation and $i$-INDP, demonstrating the similarity of IMPR to the local recovery practice in practice (Panteli et al., 2016; Rudnick et al., 2011).

To demonstrate the performance of the IMPR more generally, we perform sensitivity analysis on a larger network, the power transmission grid of Columbia, TN, with 163 nodes and 240 links.

Recall the generator fraction is the ratio between generators and all the nodes. Generators are selected at random with the same capacity. The total supply of generators balance the total demand of consumers in the network.

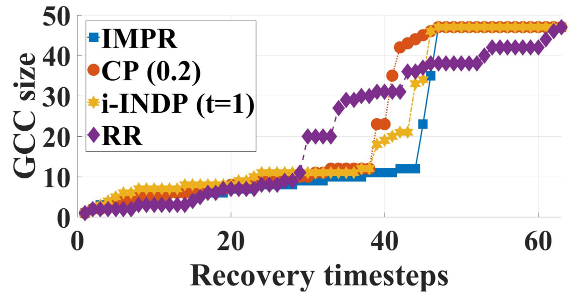


*Figure 8: GCC size during the recovery processes.*

The demands of consumers are uniformly distributed in $[(1-\delta)\cdot L_b, (1+\delta)\cdot L_b]$, where $\delta$ is the demand variation and $L_b$ is the average demand. For each pair of parameters, the results are averaged over 100 realizations (Figure 9).

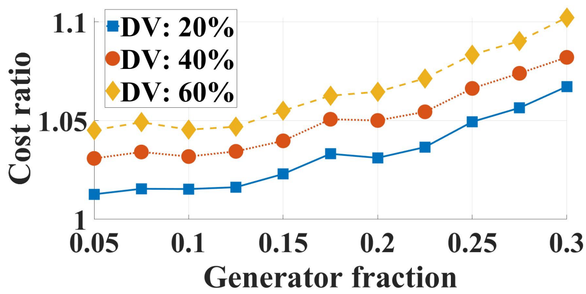


*Figure 9: Cost ratio between IMPR and i-INDP $(t = 1)$.*

Figure 9 shows the cost ratio between IMPR and $i$-INDP $(t = 1)$. The cost ratio is very close to 1, meaning the proposed IMPR is almost as efficient as $i$-INDP, the step-wise optimal solution, itself a good approximator to the global optimal solution. Moreover, the IMPR strategy requires much less computation. IMPR resembles field practices in that it suppresses the formation of the GCC the most (prioritizing local stability). In particular, the performance of IMPR approaches local optimal solutions when both generator fractions and demand variations are small, around 0.1. This value is close to realistic power transmission grids, such as 0.112 for the eastern interconnect (EI) and 0.183 for the western one (WECC) in North America (Birchfield et al., 2017).

Overall, the proposed IMPR strategy shows almost as good efficiency as the $i$-INDP $(t = 1)$ benchmark. According to additional sensitivity analyses, the IMPR strategy is particularly suitable to solve problems with small demand fraction and small demand variation. We suggest that it is because small demand fractions and demand variations create regular topological properties for appropriate network partitions. However, we will need to quantify (and proof) in future work how close the performance of the proposed strategy can be to the optimal solution, especially with respect to networks with specific topological and graph theoretic characteristics, such as infrastructure systems (Karger, 2001; Rosenthal, 1977).

## 3. CONCLUSIONS

In this paper, we propose an importance-based multicentric (local) percolation recovery (IMPR) strategy for the recovery of engineered systems from catastrophic hazards. This method mimics the recovery processes observed for engineered networks in the field, which is consistent with local-level recovery priority to satisfy customer demands quickly. Then, it connects the local clusters into a spanning component as time progresses. We show and compare the restoration evolution patterns of component excess supply, the total commodity deficit and the GCC size with $i$-INDP, competitive percolation and random recovery. We demonstrate that the proposed IMPR strategy approximates the step-wise mixed-integer optimal solution well, with much lower computational cost and a more realistic recovery logic relative to other approximation methods. Furthermore, through sensitivity analyses on generator fraction and demand variation, we find that IMPR is ever closer to optimal recovery as supplier fractions and demand variation get smaller. These parameters tend to be small across realistic engineered systems. Hence, carefully selecting the locations for suppliers and consumers, and assigning proper capacity to suppliers, one can promote small demand variation and supplier fractions to naturally enhance systems' resilience.

Small supplier fractions and demand variations render topological regularities, which are exploited in the partition of the networks. They also make islanding feasible to prevent cascading failures, and facilitate local recovery. Interestingly, the best generator fraction we found is consistent with balanced sizes of clusters, which brings the power of the partition and importance evaluation into full play. The proposed IMPR method is meaningful in supporting decision making for the recovery of infrastructure systems after catastrophic natural or deliveate hazards. Due to its realistic restoration patterns, it also provides ideas for decentralized restoration strategies to expedite performance recovery. Finally, network partition is critical in solving significant problems in many systems, such as transportation systems, water distribution systems, and supply chain systems. It also makes network science techniques applied to engineering desirable for their good compromise between computational cost and accuracy.

## 4. ACKNOWLEDGEMENT

The authors gratefully acknowledge the support by the U.S. Department of Defense (Grant W911NF-13-1-0340) and the U.S. National Science Foundation (Grants CMMI-1436845 and CMMI-1541033).

## 5. REFERENCES

Birchfield, A. B., Schweitzer, E., Athari, M. H., Xu, T., Overbye, T. J., Scaglione, A., and Wang, Z. (2017). "A Metric-Based Validation Process to Assess the Realism of Synthetic Power Grids." *Energies*, 10(8), 1233.

Di Nardo, A., Di Natale, M., Santonastaso, G. F., Tzatchkov, V. G., and Alcocer-Yamanaka, V. H. (2013). "Water network sectorization based on graph theory and energy performance indices." *Journal of Water Resources Planning and Management*, 140(5), 620–629.

Dueñas-Osorio, L. and Kwasinski, A. (2012). "Quantification of lifeline system interdependencies after the 27 February 2010 Mw 8.8 offshore Maule, Chile, earthquake." *Earthquake Spectra*, 28(S1), S581–S603.

Glauche, I., Krause, W., Sollacher, R., and Greiner, M. (2003). "Continuum percolation of wireless ad hoc communication networks." *Physica A: Statistical Mechanics and its Applications*, 325(3-4), 577–600.

González, A. D., Dueñas-Osorio, L., Medaglia, A. L., and Sánchez-Silva, M. (2016a). "The time-dependent interdependent network design problem (td-INDP) and the evaluation of multi-system recovery strategies in polynomial time." *The 6th Asian-Pacific Symposium on Structural Reliability and Its Applications, Shanghai, China*, 544–550.

González, A. D., Dueñas-Osorio, L., Sánchez-Silva, M., and Medaglia, A. L. (2016b). "The interdependent network design problem for optimal infrastructure system restoration." *Computer[U+2010]Aided Civil and Infrastructure Engineering*, 31(5), 334–350.

Hu, F., Yeung, C. H., Yang, S., Wang, W., and Zeng, A. (2016). "Recovery of infrastructure networks after localised attacks." *Scientific reports*, 6, 24522.

Karger, D. R. (2001). "A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem." *SIAM review*, 43(3), 499–522.

Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H. E., and Havlin, S. (2015). "Percolation transition in dynamical traffic network with evolving critical bottlenecks.." *Proceedings of the National Academy of Sciences of the United States of America*, 112(3), 669–72.

Liu, Y., Fan, R., and Vladimir, T. (2016). "Power system restoration: a literature review from 2006 to 2016." *Journal of Modern Power Systems and Clean Energy*, 4(3), 332–341.

Panteli, M., Trakas, D. N., Mancarella, P., and Hatziargyriou, N. D. (2016). "Boosting the power grid resilience to extreme weather events using defensive islanding." *IEEE Transactions on Smart Grid*, 7(6), 2913–2922.

Rosenthal, A. (1977). "Computing the reliability of complex networks." *SIAM Journal on Applied Mathematics*, 32(2), 384–393.

Rudnick, H., Mocarquer, S., Andrade, E., Vuchetich, E., and Miquel, P. (2011). "Disaster management." *IEEE Power and Energy Magazine*, 9(2), 37–45.

Shang, Y. (2016). "Localized recovery of complex networks against failure." *Scientific reports*, 6, 30521.

Smith, A. M., Pósfai, M., Rohden, M., González, A. D., Dueñas-Osorio, L., and D'Souza, R. M. (2018). "Competitive percolation processes for recovery in engineered networks.