Toward Quantifying Vulnerabilities in Critical Infrastructure Systems

Pravallika Devineni*, Bill Kay*, Hao Lu*, Anika Tabassum†, Supriya Chintavali*, and Sangkeun (Matt) Lee*
Oak Ridge National Laboratory* Virginia Tech†
Email: {devinenip, kaybw, luh1, chintavalis, lees4}@ornl.gov, anikat1@vt.edu

***Abstract*—Modern society is increasingly dependent on the stability of a complex system of interdependent infrastructure sectors. Vulnerability in critical infrastructures (CIs) is defined as a measure of system susceptibility to threat scenarios. Quantifying vulnerability in CIs has not been adequately addressed in the literature. This paper presents ongoing research on how the authors model CIs as network-based models and propose a set of metrics to quantify vulnerability in CI systems. The size and complexity of the CIs make this a challenging task. These metrics could be used for planning and efficient decision-making during extreme events.**

***Index Terms*—critical infrastructure, risk analysis, interdependencies, metrics**

I. Introduction

Critical infrastructures (CIs) are the systems, assets, and services upon which the economy and society generally depend to function [1]. Examples of CIs include telecommunications, electric power systems, natural gas and oil, banking and finance, transportation, water supply systems, government, and emergency services [2]. CIs have complex interdependencies in which a threat can have cascading effects across other CIs and could seriously harm society as a whole [3]. For example, Hurricane Sandy in 2012 caused enormous power utility damage, which led to the non-operation of major fuel pipelines, telecommunication infrastructure, and water and sewage facilities. Then, the transportation system collapsed due to electric and fuel outages, and the health care centers shut down because none of the infrastructure systems functioned anymore, paralyzing New York [4]. Clearly, society cannot function if large portions of the CI are disrupted or destroyed.

Given the potential consequences, vulnerability analysis is an essential tool for subject matter experts to quantify and identify subsystems that are critical (i.e., important and vulnerable). Quantifying infrastructure criticality informs researchers in how best to allocate resources to protect against accidents or malicious threats, resulting in subsystems that are resilient to cascading failures in an all-hazards environment.

CI systems can be viewed as a network in which each component or asset is a node, and there is an edge $a \rightarrow b$ if b is dependent on a. Hence, a failure in node a will likely cause a failure in node b. Failure propagation through this network model is precisely the cascading effect through CI that was described. Labels can also be assigned to the nodes according to their role in the network, and criticality can be quantified within certain types of subsystem (e.g., which power plants are the most vulnerable).

This paper's goal is to model CI systems as network-based models and identify metrics to quantify vulnerabilities effectively. Traditional models have proposed theoretical frameworks or dealt with identifying vulnerabilities in one or two CIs at a time without fully considering the system dynamics. Very few efforts focused on numerically quantifying vulnerable components across many CIs. For example, several researchers defined centrality indices based on betweenness to analyze power system vulnerabilities [5]. Network robustness studies have suggested that networks become less resilient to attacks specifically focusing on high-degree nodes [6] and have validated this conclusion on the available US power grid data [7], [8]. These studies highlight the impact of highly connected nodes on the robustness of a network and their role in cascading failures in interdependent systems [9].

The contributions of this paper are as follows. First, the relevant terminology commonly used in the CI literature is clearly defined. Second, multiple CIs are modeled as network-based models by using a publicly available dataset, and a non-exhausting set of potential research questions is listed. Lastly, metrics are presented to quantify vulnerabilities, and the importance of these vulnerabilities in these CI networks is discussed.

II. Background

This section reviews the existing studies and categorizes them into two parts. Then, the terminology pertaining to CI is presented.

A. Related Work

Infrastructure Vulnerability Analysis. This is a well-researched topic that is mostly understood in the context of risk analysis [10] and natural hazard assessment [11]. Vulnerability indices measure the negative consequences of extreme shock events (e.g., floods, storms, earthquakes) [11], [12]. The vulnerability assessment frameworks represent CIs as complex

socio-technological systems that rely on other CIs. These interdependencies trigger failure propagation mechanisms across systems, thereby amplifying disruption impacts [2]. Modeling approaches for infrastructure vulnerability analysis include empirical studies, Leontief input-output-based models, network models, and agent-based models [13]. Network models represent CIs as nodes and edges, using their structure to assess vulnerabilities and understand their impact on cascading failures [14], [15].

Vulnerability Assessment Metrics. Because it is computationally expensive to evaluate the negative consequences for the exhaustive list of failures, several vulnerability studies use smaller strategic sets of failure scenarios [16], [17]. Several approaches consider the topological and flow-based characteristics of the network to identify critical nodes and links, such as the degree centrality score to estimate the connectedness of a node. Centrality measures have mainly been used in network-based approaches for vulnerability analysis in power networks [5]. These networks are then tested for random failures of key critical components with high-degree distributions [18]. To deal with the high computational complexity of characterizing large graphs, researchers proposed the use of network topology local information to identify node and link importance [19]. In this work, the authors' approach is to identify metrics for network-based infrastructure models to assess and quantify infrastructure vulnerabilities should they be exploited.

B. Terminology

CI literature provides many definitions for its terminology, so the following terms are defined in the context of this work. The authors' goal is that these definitions will facilitate discussions across sectors and academic disciplines.

- *Infrastructure*: The Oxford dictionary defines *infrastructure* as "the basic physical and organizational structures and facilities (e.g., buildings, roads, power supplies) needed for the operation of a society or enterprise" [20].
- *Asset/component*: An infrastructure asset sustains key services, such as water, transport, electricity, and waste disposal. Some examples of assets include natural gas compressor stations, electric transmissions lines, railroads, and wastewater treatment plants. This paper uses the terms *asset* and *component* interchangeably.
- *Critical Infrastructure (CI)*: *CIs* are infrastructures that are so vital that their incapacitation or destruction would have a debilitating impact on defense or economic security [21]. Prolonged disruptions in CIs could cause significant military and economic dislocation. The Presidential Policy Directive 21 [22] identified 16 CI sectors that are considered vital for the functioning of society.
- *Dependencies and interdependencies:* CI systems are dependent and interdependent in multiple ways. *Dependency* refers to a unidirectional relationship, and *interdependency* indicates a bidirectional relationship between assets [2].
  In a dependency, the operations of Asset A affect the operations of Asset B. For example, a water treatment plant depends on communications services that support the supervisory control and data acquisition (SCADA) systems required for the control of plant operations. In an interdependency, the operations of Asset A affect the operations of Asset B, and the operations of Asset B then affect the operations of Asset A. For example, the water treatment plant requires communications for its SCADA system, and in turn it provides water that is used by the communications system to cool its equipment.
  - *Order of dependencies*: Dependencies between infrastructures can be of first order, second order, or of multi-order. First-order dependencies are direct and often easy to spot. For example, Asset B depends on Asset A. However, multi-order dependencies can be indirect and are not often obvious. For example, if Asset C is dependent on Asset B and Asset B is dependent on Asset A, then a second-order dependency exists between Assets A and C. Order of dependency is often used to assess the effect of disruption to consequent infrastructures.
  - *Degree of dependency*: Dependencies can be classified as tight or loose, depending on the relative degree of coupling. *Tight coupling* refers to agents or infrastructures that are highly dependent on one another. *Loose coupling* implies that the infrastructures are only weakly correlated to or independent of the state of the other. Additionally, interdependencies between infrastructures can be symmetric or asymmetric. For example, Asset A can be highly dependent on Asset B, whereas Asset B has weak or no dependence on Asset A.
  - *Types of dependencies*: One frequently used taxonomy by Rinaldi et al. [2] proposed four interedependency categories: physical, cyber, geographic, and logical. However, due to the complexity of identifying interdependencies, the authors of this present document chose to use interdependencies differentiated into two types [23]: geographic (i.e., two or more systems that are co-located in a physical space) and functional (i.e., physical, cyber, and logical).
- *Threat/hazard*: A *hazard* is a natural or artificial source of

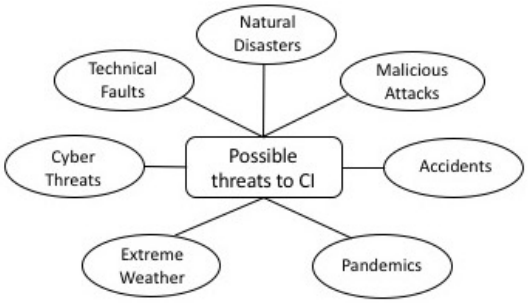

Figure 1. Threats to CIs.

Figure 2. Elements of risk in CIs.

harm or difficulty. For example, an improperly maintained chemical storage tank presents a potential hazard. A *threat* is a natural or artificial occurrence, individual, entity, or action that could harm life, information, operations, the environment, and/or property [24]. Threat is generally estimated as the likelihood that a hazard will manifest. A threat is directed at an entity, asset, network, or geographic area, but a hazard is not targeted. Figure 1 presents the different kinds of threats to CIs.

- *Vulnerability*: A *vulnerability* is a weakness of an asset or control that can be exploited by one or more threats. More simply, it is the susceptibility of the infrastructure to threat scenarios.
- *Risk*: A *risk* is the chance of something going wrong as a result of a hazard or a threat that has an impact on operations [24]. It is traditionally defined as a function of three elements: the threats to which an asset is susceptible, the vulnerabilities of the asset to the threat, and the consequences potentially generated by the degradation of the asset, which is shown in Figure 2. For example, organizations analyze intelligence reports, vulnerability assessments, and consequence models to calculate the risk of an attack.
- *Failure*: Given the complex interdependencies between CI components, threats to CI systems lead to disruptions or failures in different subsystems [25]. CI failures are categorized as common-cause (i.e., two or more infrastructures disrupted at the same time), cascading (i.e., disruption in one infrastructure affects one or more components in another infrastructure), and escalating (i.e., exacerbated disruptions from one infrastructure to another with increasing severity and recovery time). The terms *failure* and *disruption* are used interchangeably in this paper.
- *Resilience*: *Resilience* is the ability to prepare for and adapt to changing conditions [24]. This means being able to withstand and recover rapidly from disruptions, attacks, or incidents. Resilience can be factored into vulnerability and consequence estimates when measuring risk.

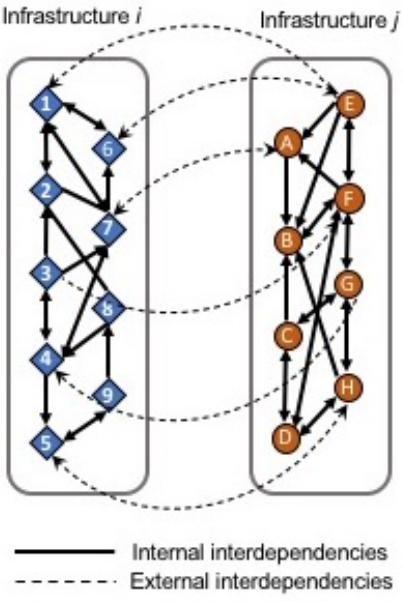

Figure 3. CI as a graph-based network.

III. Modeling Critical Infrastructure Systems

This section introduces the basic concepts of complex networks, how multiple CIs are modeled as network-based models, and presents the potential research questions that might help identify vulnerabilities in CI systems.

A. DHS HIFLD Dataset

To conduct research on interconnected CI networks, access to real-world data was needed to validate the authors' assumptions about CI network structure and function. For this purpose, the authors chose the US Department of Homeland Security (DHS) Homeland Infrastructure Foundation-Level Data (HIFLD) dataset [26], which provides open, public domain geospatial data on CIs in a variety of formats. According to their webpage [26], the intent of the DHS HIFLD is "to support research and development efforts focused on community preparedness, resiliency, research, and more." The HIFLD Open Data Catalog has over 500 layers pertaining to 16 CI sectors.

B. CI as a Network

A CI system can be represented by using an interconnected network in which nodes represent components and edges mimic the physical and relational connections among them. An infrastructure network has topological and functional aspects. For example, an electrical supply network can be expressed as a topology of electric substations and transmission lines as well as annotations (capacities and gradients), and the functional aspect can be represented as network flows.

1) CI Graph: An infrastructure network, I, is a set of nodes related to each other by a common function. The network can be connected or disjoint. It can be directional, bidirectional,

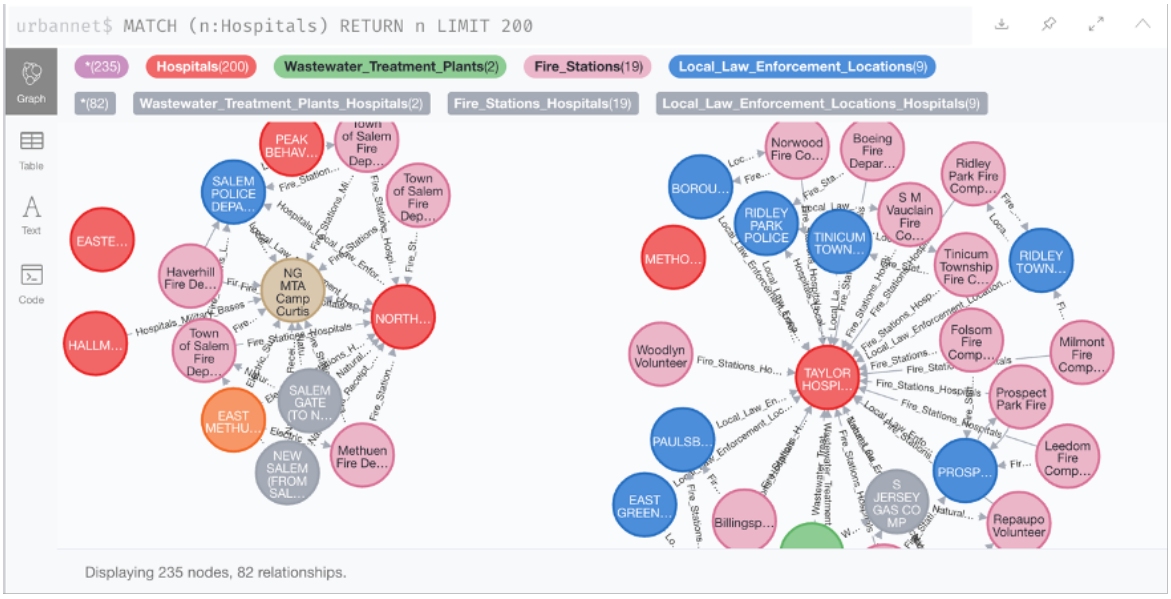

Figure 4. Graphical representation of HIFLD layers in the Neo4j graph database.

or have elements of both. Internal dependencies within the infrastructure I are represented by edges (a, b) with $a, b \in I$.

Given that I_i and I_j are infrastructure networks with $i \neq j, a \in I_i$ and $b \in I_j$, an *interdependency* is defined as a relationship between infrastructures and is represented as the edge (a, b), which implies that node b is dependent upon node a. That is, if a fails, then b is likely to fail. Depending on the nature or type of the relationship, this relationship can be reflexive in that $(a, b) \rightarrow (b, a)$.

Figure 3 demonstrates two independent infrastructures, i and j, as well as their internal and external interdependencies.

2) HIFLD Neo4j Network: The authors downloaded the HIFLD data layers [26] as shapefiles, which are a geospatial vector data format. Common geospatial component shapes include points (e.g., power plants and hospitals), multiline strings (e.g., road transportation and electric transmission lines), and polygons (e.g., airports and military bases). The authors used their previously developed software modules to construct a CI network [15] and imported this network into the Neo4j database.

Neo4j is an open-source graph database that is highly scalable and schema-free. It models the data in the form of a graph in which nodes depict the graph entities and edges depict the association of these nodes. Neo4j conveniently provides a declarative query language called Cypher and the Cypher output can be represented as a graph visualization. The current Neo4j database comprises 33 HIFLD layers and 104 interdependency links between these layers. Some layers include electric substations, microwave links, ethanol plants, and hospitals. The database has 1,941,484 nodes and 1,083,180 edges, and the nodes represent infrastructure assets, such hospitals, wastewater treatment plants, and transmission, as depicted in Figure 4. The current Neo4j instance is housed on Oak Ridge National Laboratory's (ORNL's) Compute and Data Environment for Science (CADES) cloud environment and uses about 2.4 GB of storage space.

C. Potential Research Questions

The objective of an attacker is to cause maximum damage with minimum attack cost. Attacker profiling, taking into account the topological structures of networks, is an effective technique for identifying and managing CI vulnerabilities. In this context, nodes in a CI graph can be categorized as: (1) vulnerable nodes, which are nodes that are likely to fail if any other node in the graph fails; (2) important nodes, which are nodes whose failure causes many other important nodes to be likely to fail; and (3) critical nodes, which are nodes that are vulnerable and important. A similar categorization could be applied to edges in the network. It is important to determine whether a node is vulnerable, important, and/or critical and then quantify it.

The authors identified a non-exhaustive list of potential research questions that must be addressed to identify vulnerable components in CI. The authors expect to use these questions to quantify vulnerabilities in the HIFLD infrastructure network.

1) How can interdependencies between CI components be modeled?
2) How can vulnerable/important/critical nodes be quantified?
3) How can vulnerable/important/critical paths be quantified?
4) How can vulnerable/important/critical geographical regions be identified?
5) How can nodes that expose the system to maximum failure be identified?
6) What is the minimum number of functioning nodes required to maintain network efficiency above a certain

threshold? How can efficiency be quantified in this context?

7) How can the network be made more resilient to threats? Which nodes should be chosen to introduce mitigation measures in case of disruptions?
8) Which metrics capture the impact of a disruption when one node or a large portion of the network is disrupted?

IV. Performance Metrics

CI vulnerability closely aligns with network weakness and consequences of failure. Regardless of the methodology used, it is important to develop metrics that yield reproducible results while reducing subjectivity and ambiguity. The authors identified two kinds of metrics: (1) network topology metrics that identify vulnerable nodes and edges taking into account the network structure and (2) network performance metrics that quantify the performance of a given network when some of its components are disrupted.

A. Topological Metrics

Network topology refers to how various nodes in a network are physically or logically arranged in relation to each other. Topology metrics include network properties, such as degree, distance, centrality, clustering, and robustness. Centrality measurements are used to estimate the relative "importance" or role of a node in a graph. Intuitively, nodes with high centrality scores are expected to highly affect the overall risk [27]. For example, when simulating how removing a node can trigger cascading failures in a CIS, centrality measures help assess whether removing a node with high significance causes the same amount of damage as removing a random node.

- *Degree centrality*: The *node degree* describes the number of neighbors that a node has. Intuitively, it is the estimate of a node's influence in a graph from the size of its immediate environment [28]. A directed graph has two variants of degree centrality: in-degree and out-degree. For a node u, the degree centrality is defined as:

$$\begin{aligned} C_d(u) &= d_u^{in}, \\ C_d(u) &= d_u^{out}, \\ C_d(u) &= d_u^{in} + d_u^{out}, \end{aligned} \tag{1}$$

  where d_u^{in} and d_u^{out} are the count of incoming and outgoing edges to and from node u, respectively, and the overall degree centrality of node u is the sum of in-degree and out-degree. The running time required to compute C_d for N nodes is $O(N^2)$. A node with a high in-degree corresponds to an asset that relies on many other assets. For example, a fire station's functionality might depend on many utilities, such as power, water, and the nearest hospital. Hence, a fire station is *vulnerable* because the failure of any of these utilities can cause operational difficulty to the fire station. This corresponds to the fire station having many incoming edges in the infrastructure network, and thus the in-degree is a measure of a node's vulnerability

  A node with a high out-degree corresponds to an asset that supports many other assets. For example, an electric substation supplies to utilities that cannot function without power, such as hospitals, schools, and police stations. Hence, a power plant is *important* because its failure can cause operational difficulty to any of its dependent utilities. This corresponds to a power plant having many outgoing edges in the infrastructure network, making out-degree a measure of a node's importance.

  Given these observations, a node with high in-degree and high out-degree will have high degree centrality and is thus both vulnerable and important (i.e., critical).
- *Closeness centrality*: Closeness centrality captures the idea of communication speed between nodes. It is based on geodesic distances and quantifies the average farness of a node to all other nodes in a 2D region. The node that is closest to all others receives the highest score. It is defined as:

$$C_c(u) = \frac{1}{\sum\limits_{v \in V(G)} \delta(u, v)}, \tag{2}$$

  where $\delta(u, v)$ is the length of the shortest path between the nodes u and v. The running time required to compute C_c for all N nodes using the Floyd algorithm is $O(N^3)$ [29].

  A node with a high closeness centrality corresponds to an asset whose failure does not need to cascade far to reach many other assets. In some sense, closeness centrality generalizes out-degree. A node with high out-degree will frequently have high closeness centrality. For example, a power plant could be connected to a single substation, which in turn distributes electricity to many assets. One of these assets could be a hospital, which services several other assets in the network. Although the power plant has out-degree 1, its failure cascades every asset in 1, 2, or 3 hops (i.e., they are all close).
- *Betweenness centrality*: Betweenness centrality is based on the idea that a node is central if it lies between many other nodes in the sense that is traversed by many of the shortest paths connecting pairs of nodes. It is defined as:

$$\begin{aligned} C_b(u) &= \sum_{u \neq i \neq j \in V} \delta_{ij}(u), \\ \text{s.t. } \delta_{ij}(u) &= \frac{\sigma_{ij}(u)}{\sigma_{ij}}, \end{aligned} \tag{3}$$

  where $\sigma_{ij}(u)$ is the number of geodesic distances from i to j in which node u is present, and σ_{ij} is the number of geodesic distances from i to j. The running time required to compute C_b for all nodes using the Floyd algorithm is $O(N^3)$ [30].

  A node with a high betweenness centrality corresponds to an asset that lies on many shortest paths between other pairs of resources. For example, a fire station depends on assets that supply water and power, and in turn it serves assets such as schools, hospitals, and power plants. One of

the shortest paths between the two sets of assets is likely through the fire station, and a disruption to an asset that requires services from the fire station is likely to cascade to the remainder of the network. Hence, a node with a high betweenness centrality is more likely to be critical.

- *Pagerank score*: The *Pagerank score* is a measure of node importance that results from a random walk on the network. It assigns probability distributions to each node, denoting the importance of the node by measuring the probability of being at that node during the random walk. Pagerank centrality is based on link analysis and is defined as:

$$C_p(u) = \frac{1-d}{N} + d \sum_{v \in C_{id}(v)} \frac{C_p(v)}{C_{od}(v)}, \tag{4}$$

where u is the node under consideration; $C_{id}(v)$ and $C_{od}(v)$ are the set of nodes inbound and outbound to v, respectively; N is the total number of nodes; and d is the random jumping probability.

- *Network Robustness Index (NRI)*: NRI assesses link criticality by combining topological and functional metrics and has been used in transportation networks [31]. The NRI of link a measures its importance according to the total delay experienced in the network if link a is disrupted. Comparing NRI values across links in the network helps quantify the worst consequences experienced in a connected network. NRI is defined as:

$$NRI(a) = \sum_{i \neq a} t_i x_i - c, \tag{5}$$

where x_i is the flow of the network on link i without link a, t_i is travel time, and c is the system-wide travel time cost without removing link a.

Pagerank and NRI metrics represent the global properties of the network. In an electric grid network, a power plant node with a high Pagerank represents an important node and helps evaluate and identify critical power grid nodes. In a transportation network, an NRI metric for links helps identify critical links based on network topology, helping capacity expansions [31].

Much criticality can reside on few nodes, and the above metrics help locate a group of important nodes and links that are difficult to identify otherwise. For example, a node with low degree and high betweenness indicates that the node's few ties are crucial for network flow. A node with low closeness and high degree indicates how the node is embedded in a community that is far from the rest of the network, whereas a node with high degree and low betweenness indicates that the node's connections are redundant.

B. Network Performance Metrics

The method commonly used in vulnerability analysis is to take a variety of infrastructure networks and subject them to random and/or targeted attacks. In a random attack, random nodes are removed until the graph is no longer connected, and in a targeted attack, most connected nodes are removed first. The following metrics could be used to compare the response of CI networks with attacks and failures and are usually measured before and after the occurrence of an attack.

- *Maximum connectivity coefficient*: The maximum connectivity coefficient reflects the size of the maximum connected component of the network, which is attacked by removing node. Mathematically, it reads:

$$G = R/N,$$

where R is the size of the giant component after node removal, and N is the number of nodes in the network. Clearly, the smaller the G, the better the attack strategy.

- *Efficiency of the network*: This is used to measure how efficiently resources are exchanged in a network [32]. The efficiency η of the network is defined as:

$$\eta = \frac{1}{N(N-1)} \sum_{i,j \in G, i \neq j} \epsilon_{ij}, \tag{6}$$

where N is the total number of nodes in graph G, and ϵ_{ij} is the shortest path length between between nodes i and j.

- *Average clustering coefficient*: This is the average of the local clustering coefficient of all nodes in the network [33]. It is used to measure the average degree to which all nodes are clustered together, which quantify the density of the network composition. It is formulated as:

$$\eta = \frac{1}{N} \sum_{i=1}^{N} \frac{\textit{number of triangles connected to node } i}{\textit{number of triplets centered on node } i}. \tag{7}$$

- *Change in network performance*: Taking into account the individual failed nodes and combinations of failed nodes, the resultant percentage (e_{loss}) change in network performance is evaluated according to:

$$e_{loss} = \Big(1 - \frac{\eta_{final}}{\eta_{initial}} \times 100\%\Big), \tag{8}$$

where $\eta_{initial}$ is the initial network performance, η_{final} is the final performance after node failures, and $\eta_{initial}$ and η_{final} are computed based on the respective network performance metric. e_{loss} will be positive, indicating the criticality of the failed node.

V. Conclusions

This paper discusses the development of techniques to identify vulnerabilities in a CI network in which the failure of some part of the infrastructure would have negative consequences on access to specific components and overall system performance. The relevant terminology is clearly defined to facilitate cross-sector discussions. Multiple infrastructures were modeled as network-based models, and the authors theorize that the most critical components in the infrastructure graphs tend to involve nodes with high centrality measures. However, there are multiple topology and network performance metrics that can be applied to the CI network, and these contribute to the overall vulnerability analysis in various degrees.

Future work will focus on developing methods to answer the set of research questions identified in this paper while taking the context of the CI system into account. The future direction of this work will be to demonstrate the use of topological quantity used to rank the importance of nodes and identify parts of the network that are very vulnerable to failures.

ACKNOWLEDGMENT

This material is based upon work supported by DOE's Office of Cybersecurity, Energy, Security, and Emergency Response. This research used resources of CADES at ORNL, which is supported by the US Department of Energy's (DOE's) Office of Science under contract no. DE-AC05-00OR22725.

REFERENCES

[1] B. Obama, "Executive order – improving critical infrastructure cybersecurity," 2013.

[2] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *IEEE control systems magazine*, vol. 21, no. 6, pp. 11–25, 2001.

[3] F. Petit, D. Verner, D. Brannegan, W. Buehring, D. Dickinson, K. Guziel, R. Haffenden, J. Phillips, and J. Peerenboom, "Analysis of critical infrastructure dependencies and interdependencies," Argonne National Lab.(ANL), Argonne, IL (United States), Tech. Rep., 2015.

[4] M. Haraguchi and S. Kim, "Critical infrastructure interdependence in new york city during hurricane sandy," *International Journal of Disaster Resilience in the Built Environment*, 2016.

[5] G. Oliva, A. E. Amideo, S. Starita, R. Setola, and M. P. Scaparra, "Aggregating centrality rankings: A novel approach to detect critical infrastructure vulnerabilities," in *International Conference on Critical Information Infrastructures Security*. Springer, 2019, pp. 57–68.

[6] D. S. Callaway, M. E. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Physical review letters*, vol. 85, no. 25, p. 5468, 2000.

[7] Z. Wang, A. Scaglione, and R. J. Thomas, "The node degree distribution in power grid and its topology robustness under random and selective node removals," in *2010 IEEE International Conference on Communications Workshops*. IEEE, 2010, pp. 1–5.

[8] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[9] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[10] T. Aven, "A unified framework for risk and vulnerability analysis covering both safety and security," *Reliability engineering & System safety*, vol. 92, no. 6, pp. 745–754, 2007.

[11] W. N. Adger, "Vulnerability," *Global environmental change*, vol. 16, no. 3, pp. 268–281, 2006.

[12] S. F. Balica, N. G. Wright, and F. Van der Meulen, "A flood vulnerability index for coastal cities and its use in assessing climate change impacts," *Natural hazards*, vol. 64, no. 1, pp. 73–105, 2012.

[13] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability engineering & System safety*, vol. 121, pp. 43–60, 2014.

[14] S. Puuska, K. Kansanen, L. Rummukainen, and J. Vankka, "Modelling and real-time analysis of critical infrastructure using discrete event systems on graphs," in *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, 2015, pp. 1–5.

[15] S. Lee, L. Chen, S. Duan, S. Chinthavali, M. Shankar, and B. A. Prakash, "Urban-net: A network-based infrastructure monitoring and analysis system for emergency management and public safety," in *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, 2016, pp. 2600–2609.

[16] S. M. Wilkinson, S. Dunn, and S. Ma, "The vulnerability of the european air traffic network to spatial hazards," *Natural hazards*, vol. 60, no. 3, pp. 1027–1036, 2012.

[17] R. Pant, J. W. Hall, and S. P. Blainey, "Vulnerability assessment framework for interdependent critical infrastructures: case-study for great britain's rail network," *European Journal of Transport and Infrastructure Research*, vol. 16, no. 1, 2016.

[18] T. Verma, W. Ellens, and R. E. Kooij, "Context-independent centrality measures underestimate the vulnerability of power grids," *International Journal of Critical Infrastructures 7*, vol. 11, no. 1, pp. 62–81, 2015.

[19] H. Yang and S. An, "Critical nodes identification in complex networks," *Symmetry*, vol. 12, no. 1, p. 123, 2020.

[20] "infrastructure, n." [Online]. Available: https://www.lexico.com/en/definition/infrastructure

[21] R. Marsh, "Critical foundations: protecting america's infrastructure," 1997.

[22] B. Obama, "Presidential policy directive 21: Critical infrastructure security and resilience," 2013.

[23] P. Pederson, D. Dudenhoeffer, S. Hartley, and M. Permann, "Critical infrastructure interdependency modeling: a survey of us and international research," *Idaho National Laboratory*, vol. 25, p. 27, 2006.

[24] R. S. Committee, "threat."

[25] P. Kotzanikolaou, M. Theoharidou, and D. Gritzalis, "Cascading effects of common-cause failures in critical infrastructures," in *International Conference on Critical Infrastructure Protection*. Springer, 2013, pp. 171–182.

[26] D. of Homeland Security, "Dhs homeland infrastructure foundation-level data (hifld)," https://hifld-geoplatform.opendata.arcgis.com/, online; accessed 17 September 2020.

[27] G. Stergiopoulos, M. Theocharidou, P. Kotzanikolaou, and D. Gritzalis, "Using centrality measures in dependency risk graphs for efficient risk mitigation," in *International Conference on Critical Infrastructure Protection*. Springer, 2015, pp. 299–314.

[28] R. Zafarani, M. A. Abbasi, and H. Liu, *Social media mining: an introduction*. Cambridge University Press, 2014.

[29] R. W. Floyd, "Algorithm 97: shortest path," *Communications of the ACM*, vol. 5, no. 6, p. 345, 1962.

[30] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social networks*, vol. 1, no. 3, pp. 215–239, 1978.

[31] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo, "Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks," *Journal of Transport Geography*, vol. 14, no. 3, pp. 215–227, 2006.

[32] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Physical review letters*, vol. 87, no. 19, p. 198701, 2001.

[33] A. Kemper, *Valuation of network effects in software markets: A complex networks approach*. Springer Science & Business Media, 2009.