A Matrix Model to Analyze Cascading Failure in Critical Infrastructures

Assane Gueye[1(✉)], Babacar Mbaye[1], Doudou Fall[2], Alassane Diop[3], and Shigeru Kashihara[2]

[1] University Alioune Diop of Bambey (UADB), Bambey, Senegal
{assane1.gueye, babacar.mbaye}@uadb.edu.sn
[2] Nara Institute of Science and Technology (NAIST), Ikoma, Japan
{doudou-f, shigeru}@is.naist.jp
[3] Université Virtuelle du Senegal (UVS), Dakar, Senegal
alassanediop.dfip@uvs.edu.sn

Abstract. Critical infrastructures are defined as systems and assets, whether physical or virtual, so vital to the nation that their incapacity or destruction would have a debilitating impact on the nation's existence. Although composed of systems that are usually designed/implemented *independently*, critical infrastructures are in reality *interdependent*: hence risks/failures will often *cascade* from one system to another. In this paper, we derive an efficient procedure to fully describe the cascading effects of a node failure in a network of interdependent systems. The procedure is solely based on operations on the adjacency matrix of graph representing the network. We have also shown that the analysis of the cascades can be based on a much smaller matrix that has a DAG structure. This matrix characterization of the cascade and the dimension reduction of the analysis open new opportunities in the study of cascading effects in interdependent networks. Although this paper focuses on the interdependence between the power grid and the communication system, the model presented herein easily generalizes to the interdependence of an arbitrary number of networks.

Keywords: Critical infrastructure · Interdependent networks · Cascading failures · Graph theory · Adjacency matrix

1 Introduction

Critical infrastructures are defined as "systems and assets, whether physical or virtual, so vital to the nation that their incapacity or destruction would have a debilitating impact on security, national economic security, national public health or safety, etc., or any combination thereof [1]". They include (but are not limited to) the communication and information systems, the power grid, the water and gas distribution systems, the transportation systems, the healthcare systems, and the financial systems. These systems can be owned and operated by

J. P. R. Thorn et al. (Eds.): InterSol 2020, LNICST 321, pp. 211–223, 2020.
https://doi.org/10.1007/978-3-030-51051-0_15

both public and private agents. Each system has a certain degree of autonomy, but they are usually all overseen by a common entity (the state: e.g. through laws, regulations, and funding). Although they are mostly in-country, some critical infrastructures might be located outside the country. Finally, each critical infrastructure sector has unique characteristics, operating models, and risk profiles.

More importantly, these systems share common risks and threats and constitute one big system that is globally vulnerable due to the several reasons. One of the most prevalent reasons is the fact that they are interdependent and are often physically and logically interconnected at one or more interconnection points (e.g., where the power grid provides energy to the communication system).

Given the vital role they play in our modern society, critical infrastructures must be secure and able to withstand and rapidly recover from all hazards. Achieving this will require an effective collaboration and coordinated efforts both at the strategic and operation levels, and an efficient information exchange at all levels. These, on the other hands, will need a holistic understanding of the notion of risk at the global level, as risk will often be transferred from one system to the other. The main goal of this paper is to derive models and metrics to quantify the potential impacts as well as cascading effects of an incident or threat scenario in one critical sector on other critical sectors.

The main challenge in the study of interdependent system is to understand how the failure of a particular node (or group of nodes) propagates through the global network. In this paper, we use a graph theory approach to model the interdependence between different networks. More precisely, we use an extended adjacency matrix to capture the connectivity within each network as well as the interdependencies between the networks. The graph we consider is directed, as the interdependencies might not be symmetrical.

The main contribution of this paper is the derivation of an efficient procedure to fully describe the cascading effects of a node failure in the network. The procedure is solely based on operations on the adjacency matrix of the network. We have also shown that the analysis of the cascades in a network can be based on a much smaller graph that has a DAG structure. This matrix characterization of the cascade and the dimension reduction of the analysis open new opportunities in the study on cascading effects in network.

As a use-case, this paper focuses on the interdependence of the power and the communication system. Energy and communications systems are uniquely critical due to the enabling functions they provide across all critical infrastructure sectors. We study the resilience of the combined system against arbitrary node failures.

We would like to stress the fact that the focus on the power grid and communication network is just for illustrative purpose. The model and analysis proposed in the paper easily generalize to the interdependence of an arbitrary number of networks.

The remainder of this paper is organized as follow. In Sect. 2 we present the details of our model as well as the assumptions made for the model. In Sect. 3, we provide the details of procedure to characterize cascading failure using this adjacency matrix. A brief discussion of related work is presented in Sect. 4. This paper ends with a conclusion and perspectives in Sect. 5.

2 Model

In this section, we discuss our assumptions in modeling the power grid, the communication, and the interdependencies between them. These models are simplifications of real-life scenarios. However, we believe that they constitute good first order approximations. More precisely, we assume that:

- The power grid (PG) consists of nodes and power lines, where the nodes are mainly: *Generators* (G) that generate power, *Loads* (L) that consume power and *Buses* (B) that allow the transmission of power through them. Each generator produces a fix quantity of energy and each load demands a fix quantity of energy. A load (L) or bus (B) cannot operate if it does not have a path to at least one generator,
- The power grid is connected to a Communication and Control Network (CCN) that allows monitoring and intelligent control actions to respond to changes in the grid conditions,
- The CCN consists of Control Centers (CC) and routers (R) that are connected by communication links. A router (R) cannot operate if it does not have a path to (at least one) CC,
- Every router (R) receives power from at least one load (L) node via a directed dependency link. When all the load nodes to which a router depend on fail, the router also fails. A router is fully operational if at least one of the load nodes to which it depends is also functional at its full capacity and it has a route to the CC,
- Every bus (B) sends/receives control signals from/to at least one router (R) via a bidirectional dependency link. When all the routers to which a bus depend on fail, the bus also fails. A bus is fully operational if at least one of the dependent routers is operational,
- A load (L) is operational if its entire demand is met by the quantity of energy it receives,
- When a failure occurs in either network, there could be a cascading effect,
- The generators and control centers are autonomous: generators have internal control and control centers have energy backup, so that they will always operate (indeed, if the sources fail, the whole system breaks down: we avoid that trivial case).

Next, we model these systems and their interdependencies using graph theory notions.

2.1 Model of the Power Grid

With the assumptions above, the delivery of energy on the power grid can be modeled by a network flow problem (also called supply-demand model [2]). In a network flow problem, the goal is to carry a fixed amount of goods from a nonempty subset of source nodes G_{en} (the generators) to a nonempty subset L of destination nodes (the loads), using the network links (transmission lines). These links transit through intermediary nodes (the buses). We assume that the subset of sources is disjoint to the subset of destinations (a generator cannot be a load at the same time) and that network links are directed.

With each node in $g \in G_{en}$ (i.e., a generator), we associate a nonnegative number $s(g)$, the "supply" at g; and with each node $l \in L$ (i.e, a load), we associate a nonnegative number $d(l)$, the "demand" at l. At the buses (B), there is no production or consummation. Consequently, the entire amount of goods arriving from the incoming links of a bus are distributed to its outgoing links. In this paper, we assume that there is no capacity constraints: links can carry an arbitrary amount of goods (this assumption shall be relaxed later). Finally, we assume that there is a fixed amount of energy ($\varDelta$) to be moved from the generators to the loads and we consider a balanced network where the supply fully matches the demand (supply = demand = $\varDelta$).

To analyze the model, we define the notion of a *flow* which is essentially a function that assigns a number to each link and satisfies the following properties: (1) the *conservation of flows* (at each node, the difference of incoming and outgoing flows is equal to the locally consumed/produced flow) and (2) the *capacity constraints* (each link carries a quantity of flows that is at most equal to its capacity; note that in this paper capacities of the links are assumed to be infinite), (3) *balanced network*: the total amount of supply is equal to the total amount of demand (in other terms, the total amount of energy produced by the generators is equal to the total amount of energy to be consumed by the loads).

In this paper, we only consider integer flows, where links carry only integer amounts of goods. We also assume that the total amount of goods to be carried from all sources to all destinations ($\varDelta$) is also a positive integer.

Figure 1 (left figure) shows an example of flow model for a simple power grid. The generator produces 3 units of goods and loads L_1 and L_2 have respective demand of 1 and 2 units of goods. The right figure shows an example of flow (i.e., assignment of integer values to links). Links with zero unit of goods are considered inactive and are shown in dotted-line. Notice that, in general, there is a large number of possible flows. We would like to choose one so that the resulting global interdependent system is resilient to failures.

2.2 Model of the Communication Network

In this section, we present the model for the communication network considered in this study. First, notice that in a real-world setting, the communication network is utilized for many different applications. However, in this study we

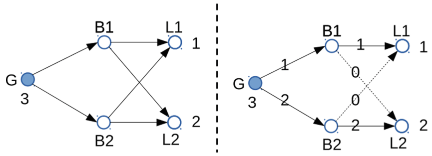

Fig. 1. Example of flow the power grid

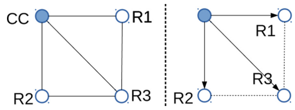

Fig. 2. Example of All-to-One network model

consider that it is exclusively used for the monitoring and control of the power grid. This (rather restrictive assumption) shall be relaxed later.

For the communication network (CCN), we use an All-to-One communication network model [3]. In an All-to-One model, the primary goal of the network (operator) is to enable all routers to communicate with a designated node (the CC for this study). To get all routers connected to the CC, the network (operator) chooses a collection of links that forms a spanning tree. Since the objective with the CCN is to get each router connected to the CC, we can transform the (undirected) spanning tree to a directed one rooted at the CC (with all the links going away from the CC).

Figure 2 shows an example of All-to-One network. The right figure shows an example of rooted spanning. Notice that there is a large number of possible spanning trees. We would like to choose one so that the resulting global interdependent system is resilient to failures.

2.3 Interdependence Model

For their respective operations to be possible, each network needs the other one: the CCN needs energy (for the routers) that is provided by the PG, while the PG needs remote control and monitoring that is enabled by the CCN. Figure 3 shows an example with interdependence links. In real-life situations, these links depends on many factors such as physical co-location and financial options.

2.4 Cascading Failures

For the *global* (interdependent) network to work at its full capacity, the flow chosen for the power grid must be able to carry the total quantity of produced energy and the spanning tree chosen for the communication network must connect all routers to the CC. When a link/node from either network fails, all the nodes in its downstream also fail. This leads to a cascading effect that takes place in the network.

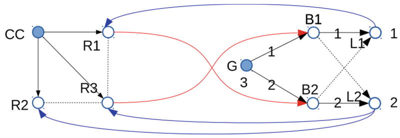

Fig. 3. Example of interdependence

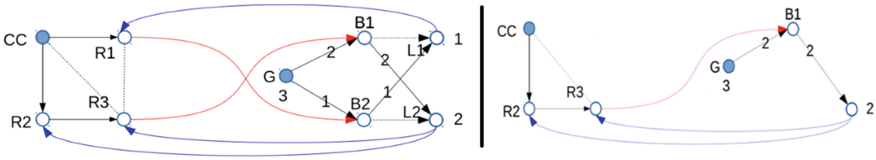

Fig. 4. Example of interdependence with partial cascade

For instance, in the (global) system in Fig. 3, if router R_1 fails, it will cause bus B_2 to lose control and monitoring, hence will not operate normally (i.e., fails). This will then cause load L_2 to fail. But L_2 provides energy to router R_3, which will then fail and cause bus B_1 and load L_1 to fail. At the end of this process, only the generator and the control center (which are considered to be autonomous) are still functioning. This is an example of full cascade in the network.

Figure 4 shows another choice of PG, CCN for the same interdependencies (top figure). In this case, a failure that starts from router R_1 will cascade to only part of the network. The bottom figure shows the residual network after cascade. This network configuration is clearly preferred to the one in Fig. 3.

At the end of the cascade, some part of the flow (f_{lost}) might be lost, and a number of communication nodes (n_{lost}) might become disconnected from the CC. We assume that the cost associated with the failure is a function of these two quantities $H(f_{lost}, n_{lost})$.

Our goal is to derive a procedure that describes the cascading process. Next, we present the mathematical tools needed for such a derivation.

2.5 Adjacency Matrix

Once a (global) network configuration is fixed, we can represent it as a graph whose node set contains the CCN nodes and the nodes from the PG. The links of the graph are composed with the links from those two networks and the interdependence links. The graph is directed, as the interdependencies might not be symmetrical. Equation 1 shows the adjacency matrix of the network configuration shown in Fig. 3. In the next section, we show how to characterize the cascade process by only using operations on this matrix.

$$\mathbf{A} = \left[\begin{array}{cccc|ccccc} 0&1&1&1&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0&0 \end{array}\right] \quad (1)$$

3 Analysis

The main challenge in the study of interdependent systems is to understand how the failure of a particular node (or group of nodes) propagates through the global network. In this section, we show how to fully describe the cascading effects of the failure of a node by only using operations on the adjacency matrix. To our knowledge, this is the first derivation of cascades in interdependent networks that is solely based on operations on the adjacency matrix. The method is also very efficient as it reduces the dimension of the computations.

3.1 Closed Cascade Domains

First, we define the notion of closed cascade domain (CCD) of a network, as *a set of nodes S such that for each pair of nodes $(i, j) \in S$, a failure at i will propagate to j, and vice-versa, a failure at j will propagate to i.*

The following lemma gives a one-to-one correspondence between the CCDs of a network and the strongly connected components (SCCs) of its graph.

Lemma 1. *Every SCC of the graph constitutes a CCD of the network, and vice-versa.*

The proof of the lemma is straightforward. Since a failure at i leads to a failure at j only if there is a path from i to j, SCCs are equivalent to CCDs. In what follows, we will use CCD and SCC interchangeably.

Building upon this correspondence, we now describe the different steps of a cascade process in the network. The general idea is that a failure of some node will eventually cascade to all nodes in the same CCD. It will also then subsequently reach a node belonging to another CCD that is reachable from a node in the first CCD. Once a node in a new CCD is reached, the cascade continues from that node onwards and the same process is repeated.

3.2 Computing the SSC of the Network

The first step is to determine the different CCDs by using the adjacency/configuration matrix A. For that, we use operation in a new *semiring* whose members are the set of matrices with entry values $\{0, 1\}$. We define the operations $\vee$ (the element-wise OR function) and $\wedge$ (the element-wise AND function) (please see [4] for more details). Notice that the adjacency matrix is an element of this semiring. We define a matrix $\mathbf{C}$ (the infinite sum of powers of $\mathbf{A}$) as follow:

$$\mathbf{C} \equiv I \vee \mathbf{A} \vee \mathbf{A}^2 \vee \mathbf{A}^3 + ..., \tag{2}$$

where $\vee$ is the element-wise OR function satisfying $a \vee b = 0$ when both a and b are zero, and $a \vee b > 0$ (which we consider to be 1) if either a and b is nonzero.

The matrices $\mathbf{A}^k$ are the k^{th} powers of the matrix $\mathbf{A}$. Entry (i, j) of the $\mathbf{A}^k$ is positive if there exists a path of length k from i to j (actually, it is the number of paths of length k between i and j). Hence, $\mathbf{C}(i, j) > 0$ if and only if there exists a path from i to j.

It is worth noting that the OR-summation in the definition of $\mathbf{C}$ does not need to go to infinity. However, let us, for the time being, keep the summation infinite for convenience reasons. Later (in Sect. 3.6), we will show an efficient way to compute $\mathbf{C}$.

Now consider the matrix $\mathbf{C} \wedge \mathbf{C}^T$, where $\wedge$ is the element-wise AND function, satisfying $a \wedge b = 0$ when either a or b is zero, and $a \wedge b > 0$ (which we consider to be 1) only when both a and b are nonzero. The following lemma gives a relation between nodes belonging to the same SCC. The lemma is solely based on $\mathbf{C} \wedge \mathbf{C}^T$.

Lemma 2. *Entry (i, j) of $\mathbf{C} \wedge \mathbf{C}^T$ is positive if and only if nodes i and j belongs to the same SCC.*

Corollary 1. *Rows of $\mathbf{C} \wedge \mathbf{C}^T$ corresponding to the nodes belonging to the same SCC are identical.*

The proof of the lemma is quite straightforward. Indeed, if $\mathbf{C} \wedge \mathbf{C}^T(i, j) = \mathbf{C}(i, j) \wedge \mathbf{C}(j, i) > 0$, then each term of the "product" must be positive. This means that there exists a path between i and j and a path between j and i, which implies that i and j belongs to the same SCC. On the other hand, if i and j belongs to the same SCC, both terms of the product are nonzero, and so is the product.

For the corollary, we use the fact that the SCCs are equivalence classes: if i and j belongs to the same SCC and j and k belongs to the same SCC, then i and k belongs to the same SCC. In other terms, if $\mathbf{C}(i, j) \wedge \mathbf{C}(j, i) > 0$ and $\mathbf{C}(j, k) \wedge \mathbf{C}(k, j) > 0$ then $\mathbf{C}(i, k) \wedge \mathbf{C}(k, i) > 0$. By considering all nodes in the same SCC, we can easily use the above lemma to show that their corresponding rows in $\mathbf{C} \wedge \mathbf{C}^T$ have same entry patterns (i.e., are zero in the same entries and nonzero in the same entries).

From the product matrix $\mathbf{C} \wedge \mathbf{C}^T$, we define the matrix $\mathbf{S}$ as its *reduced row echelon form*. Matrix $\mathbf{S}$ is of dimension $K \times N$, where K is the number of SCCs and N is the number of nodes of the network. It can be obtained by performing a series of elementary operations on the rows of $\mathbf{C} \wedge \mathbf{C}^T$. It has a number of nice properties that we list next.

First, remember that each row of the matrix $\mathbf{S}$ corresponds to one SSC of the graph. Let each SCC be indexed by its corresponding row in $\mathbf{S}$ (row 1 corresponds to SCC 1, row 2 to SSC 2, ... and so on). Then, the following lemma gives us a way to label each node with the index of the SSC it belongs to.

Lemma 3 (Labelling nodes with the index of their SCC). *Let $l(i)$ be a function that labels a node with the index of its SCC. Then,*

$$l(i) = [1 : K]\mathbf{S}1_i, \tag{3}$$

where $[1 : K]$ is the row vector $[1, 2, 3, \ldots, K]$ and $1_i = [0 \ldots 1 \ldots 0]^T$ is the column vector with entries equal to zero, except for entry i which is 1.

Corollary 2. *The labels of all nodes are given by the vector*

$$l = [1 : K]\mathbf{SI}, \tag{4}$$

where $\mathbf{I}$ is the identity matrix.

The arguments behind the lemma are as follows: First, notice that each column of $\mathbf{S}$ contains only one nonzero entry (actually a 1). This is the case because each node belongs to one and only one SCC. Also, the product $\mathbf{S}1_i$ is equal to column i of $\mathbf{S}$ (which contains only one nonzero entry at the row corresponding to the SCC containing i). Thus, $l(i)$ is equal to the index of the SCC that contains node i.

3.3 Defining the (Much Smaller) Cascade DAG

Since all the nodes in the same CCD (or SCC) have the same cascading effects, we can group all such nodes in one super-node to form a new residual graph $\tilde{G}$ as follow: contract all edges in the SCC and remove duplicates and multi-edges. The resulting graph $\tilde{G}$ is a directed acyclic graph (DAG). We will call it the cascade DAG. The adjacency matrix of the cascade DAG can be derived as follows.

Lemma 4. *Let the matrix $\mathbf{A}'$ be definded as $\mathbf{A}' = \mathbf{SAS}^T$.*
Then, $\mathbf{A}'(i, j)$ is the number of links (of the original graph) from SCC i to SCC j.

After removing all loops and multi-edges, we obtain the adjacency matrix $\tilde{\mathbf{A}}$ of $\tilde{G}$ as follow:

Lemma 5. $\tilde{\mathbf{A}} = (\mathbf{A}' > 0) - \mathbf{I} = 1_{\mathbf{A}' - diag(\mathbf{A}') > 0}$.

Since the failure of a network node will eventually cause failure of each node in its CCD, we can consider the failure of a super-node in the cascade DAG. Such failure, will propagate to the entire 'descendance' of that super-node (its children, the children of these children, and so on...). Hence, to characterize the effects of a failure of a node, we just need to compute its descendance in the cascade DAG. For that, we define the (cascade) consequence matrix $\mathbf{E}$ as follow:

Definition 1. $\mathbf{E} \equiv \mathbf{I} \vee \tilde{\mathbf{A}} \vee \tilde{\mathbf{A}^2} \vee \tilde{\mathbf{A}^3} + \cdots \vee \tilde{\mathbf{A}^k}$.

As was observed earlier with the matrix $\mathbf{C}$, Sect. 3.6 gives an efficient way to compute the matrix $\mathbf{E}$.

The entry of matrix $\mathbf{E}$ verifies the property: $\mathbf{E}(i,j) > 0$ if super-node j is a descendant of super-node i. In other terms, $\mathbf{E}(i,j) > 0$ if a failure of a node in CCD i will cause failure of all the nodes in CCD j. Hence,

Lemma 6. *Matrix* $\mathbf{E}$ *describes the cascading effect of the failure of each node.*

3.4 Computing Loss Related to Cascade

Now suppose that there is a loss associated with the failure of each node. Let the vector L be the loss vector of the network, such that $L(i)$ is the loss associated with node i. We assume that the loss associated to the failure of a set of nodes S is equal to the sum of the losses of the nodes

$$L(S) = \sum_{i \in S} L(i). \tag{5}$$

Then the losses associated with the failure of the different CCDs are given by

$$L_{DAG} = \mathbf{S}L. \tag{6}$$

In other terms, $L_{DAG}(k)$ is the sum of the losses associated with the failure of all the nodes belonging to the CCD_k (remember that a failure of one node in the CCD will eventually propagate to all the nodes in the CCD).

Now, the losses associated with a cascade that starts from any node on the global network are given by the vector L_{Net}, where the vector L_{Net} is equal to:

Lemma 7 (cascade losses). $L_{net} = \mathbf{E}L_{DAG} = \mathbf{ES}L.$

$L_{net}(k)$ is the total loss incurred after a cascade that starts from any node belonging to the CCD_k. This loss can be written as

$$L_{net}(k) = 1_k L_{net} = 1_k \mathbf{E} L_{DAG} = 1_k \mathbf{ES}L. \tag{7}$$

3.5 Summary of the Cascade Characterization Procedure

Now we summarize the procedure to compute the cascading effects of node failures in the network.

- From the Adjacency matrix $\mathbf{A}$, compute the matrix $\mathbf{C}$ (the next subsection shows an efficient way to compute $\mathbf{C}$)
- Compute the SCC matrix $\mathbf{C} \wedge \mathbf{C}^T$ and compute its reduced row echelon $\mathbf{S}$ to identify the different SCC (hence CCD) of the network
- Label each node of the network with the index of the CCD to which it belongs using: $l(i) = [1:K]S1_i$,
- Compute the adjacency matrix of the cascade DAG

$$\tilde{\mathbf{A}} = (\mathbf{A}' > 0) - \mathbf{I} = 1_{\mathbf{A}' - diag(\mathbf{A}') > 0} \tag{8}$$

where $\mathbf{A}' = \mathbf{SAS}^T$.

– Compute the consequence matrix $\mathbf{E}$ (next subsection shows an efficient way to compute $\mathbf{E}$)
– Compute the cascade loss vector: $L_{net} = \mathbf{E}L_{DAG} = \mathbf{ES}L$
– For (network) node i, compute its associated cascade losses as:

$$L_{net}(i) = 1_{l(i)}L_{net} = 1_{l(i)}\mathbf{E}L_{DAG} = 1_{l(i)}\mathbf{ES}L \quad (9)$$

3.6 Efficient Computation of the Matrix C

Now we provide an efficient method to compute the matrix $\mathbf{C}$. The same method can be used to compute the matrix $\mathbf{E}$.

First, notice that since $\mathbf{C}$ is an infinite sum, directly computing it may be complex. Instead, we define a new matrix $\mathbf{D}$ as follow:

$$\mathbf{D} = \mathbf{I} + (\alpha\mathbf{A}) + (\alpha\mathbf{A})^2 + (\alpha\mathbf{A})^3 + ... \quad (10)$$

We will soon comment on the real number α, but for the time being, just assume that it is a strictly positive number chosen to be sufficiently small.

It is not hard to see that $\mathbf{D}$ has 0 where $\mathbf{C}$ has 0, and $\mathbf{D}$ has nonzero positive entry where $\mathbf{C}$ has 1 (i.e., $\mathbf{C}(i, j) = \mathbf{D}(i, j) > 0$). So, if we can efficiently compute $\mathbf{D}$, we instantly have $\mathbf{C}$. We use the following trick for the computation of $\mathbf{D}$. Let matrix $\mathbf{F}$ be defined as follows:

$$\mathbf{F} = \mathbf{D} - (\alpha\mathbf{AD}) \quad (11)$$

$$= \mathbf{I} + (\alpha\mathbf{A}) + (\alpha\mathbf{A})^2 + (\alpha\mathbf{A})^3 + \ldots \quad (12)$$

$$- (\alpha\mathbf{A}) - (\alpha\mathbf{A})^2 - (\alpha\mathbf{A})^3 - \ldots \quad (13)$$

Hence,

$$\mathbf{F} = \mathbf{D} - (\alpha\mathbf{AD}) = \mathbf{I} \quad (14)$$

By factorizing with $\mathbf{D}$, we get

$$\mathbf{D} = (\mathbf{I} - (\alpha A))^{-1} \quad (15)$$

Now, the matrix $(\mathbf{I} - (\alpha A))^{-1}$ is not always invertible. However, if we choose a value of α sufficiently small, we can always make it invertible. A similar argument can be used for the computation of the consequence matrix $\mathbf{E}$.

In conclusion, the operations on the semiring mentioned earlier were just introduced for convenience reasons. To compute the cascading effects of failure of the nodes, we will only perform classical operations on real-valued matrices.

4 Related Work

Due to the vital role they play in modern societies, there has recently been a surge in the interest to understanding critical infrastructures [1], which are known to be interdependent systems [5].

The authors in [6] present a survey of U.S. and International Research in Critical Infrastructure Interdependency Modeling. The report presents a holistic view that includes political, economic, and social aspects. A more recent survey [7] presents an overview on some modeling approaches and models used to analyze critical infrastructures interdependencies.

Models used to analyze critical infrastructures are largely dominated by *random graphs* [8]. Due to their limitations to capture real-world networks, some authors have turned into simulation models [9]. Hybrid models [5] have also been proposed that combine both random graphs and simulations. The authors in [10] focuses on the interdependence between the power grid and the communication network. The present paper uses some of their assumptions.

Most of these models make assumptions that are often more realistic that the ones we make in this paper. However, because of their advanced models, most of the papers fail to provide a simple way to characterize cascades, and as a consequence, they often end up turning to simulations to analyze cascades. Instead, our paper is a topology-based analysis that considers only the structures of the interconnections of the nodes within and between the different networks. We have purposely taking a step back to look at the system in a more fundamental way and we have derived an efficient characterization of cascades that is solely based on operations on the adjacency matrix. To our knowledge, the is the first result of the kind in the study of cascades in interdependent network.

5 Conclusion

In this paper, we have studied the cascading effects of the failure of a node in an interdependent network. We have presented the derivation of an efficient procedure to fully describe the cascading effects of a node failure in the network. The procedure is solely based on operations on the adjacency matrix of the network. We have also shown that the analysis of the cascades in a network can be based on a much smaller matrix that has a DAG structure.

This matrix characterization of the cascade and the dimension reduction of the analysis open new opportunities in the study of cascading effects in network. In perspective, we plan to leverage our results to propose methods for the design and optimization of resilient interdependent networks.

Acknowledgement. This work was partially accomplished under NIST Cooperative Agreement No.70NANB19H063 with Prometheus Computing, LLC. The authors would like to thank Paul Patrone and Brian Cloteaux (NIST ACM Division) for their useful advice and suggestions.

References

1. US Dept. Homeland Security: Critical infrastructure sectors (2014). https://www.cisa.gov/critical-infrastructure-sectors
2. Smith, D.K., Gueye, A.: Network flows: theory, algorithms, and applications. J. Oper. Res. Soc. **45**, 1340 (1994). https://doi.org/10.1057/jors.1994.208

3. Laszka, A., Gueye, A.: Quantifying All-to-One network topology robustness under budget constraints. ACM SIGMETRICS Perform. Eval. Rev. **41**(4), 8–11 (2014)
4. Kepner, J., Gilbert, J.: Graph Algorithms in the Language of Linear Algebra. Society for Industrial and Applied Mathematics, Philadelphia (2011)
5. Havlin, S.: Catastrophic cascade of failures in interdependent networks. Nature **464**, 1025–1028 (2010)
6. Idaho National Laboratory: Critical infrastructure interdependency modeling: a survey of US and international research (2006). https://www.osti.gov/servlets/purl/911792
7. Banerjee, J., Das, A., Sen, A.: A survey of interdependency models for critical infrastructure networks. In: Examining Robustness and Vulnerability of Networked Systems (2014)
8. D'Souza, R.M., Brummitt, C., Leicht, E.A.: Modeling interdependent networks as random graphs: connectivity and systemic risk. In: D'Agostino, G., Scala, A. (eds.) Networks of Networks: The Last Frontier of Complexity. Understanding Complex Systems, pp. 73–94. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-03518-5_4
9. Rigole, T., Deconinck, G.: A survey on modeling and simulation of interdependent critical infrastructures, January 2006
10. Parandehgheibi, M., Modiano, E.: Robustness of interdependent networks: the case of communication networks and the power grid. In: 2013 IEEE Global Communications Conference (GLOBECOM), pp. 2164–2169 (2013)