Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

LEVERAGING NETWORK INTERDEPENDENCIES TO OVERCOME INACCESSIBLE CIVIL INFRASTRUCTURE DATA

Brigham Moore
David Jacques
Steven Schuldt

Dept. of Systems Engineering and Management
Air Force Institute of Technology
2950 Hobson Way
WPAFB, OH 45433, USA

ABSTRACT

Data-driven decision making and expansion of smart city infrastructure require massive amounts of data that might not be available. The lack of infrastructure data can make it challenging to recover interdependent infrastructure systems following a disruption. Interdependent infrastructure systems are often modeled as networks with an interdependency parameter. Researchers can partially overcome gaps in data associated with the individual networks by modifying interdependency parameters to include interdependency type and coupling strategy information. Overcoming missing telecommunications data is illustrated using a combined network design and scheduling problem with a modified interdependency parameter. The modified parameter allows analysis without a full dataset and removes the necessity of adding constraints and variables to handle complex infrastructure relationships. The difference in system operability results from partial and full datasets is less than or equal to 2.6%. This modeling method provides an interim solution to full data acquisition and may be suitable in other applications.

1 INTRODUCTION

Changes in infrastructure management and protection are evident in current trends with Industry 4.0, Smart Cities, and City of the Future initiatives (American Society of Civil Engineers 2019; Rutgers and Sniderman 2018). Collectively these changes require massive quantities of data which can be hard to acquire or access. While certain strategies have been employed to overcome data access challenges, it remains a significant problem.

The work presented herein addresses current work that proves useful when encountering access issues for modeling and simulation of interdependent infrastructure systems. It explores two of the six dimensions of interdependent infrastructure systems called interdependency type and coupling (Rinaldi et al. 2001). This paper then leverages these concepts of interdependency type and coupling as a way to overcome incomplete data. This is accomplished by modifying a commonly used interdependency parameter to incorporate these two elements, which allows for complex interdependencies to be created based on the available infrastructure data. This work's applicability is shown by comparing the results of recovery following a disaster for a network with all required infrastructure data and a network with a significant portion of the data missing.

2 LITERATURE REVIEW

This section details relevant literature for two topics important to this research. The first topic concerns current methods to overcome incomplete data for infrastructure modeling and simulation (M&S). The second topic is handling operational interdependencies by the use of an interdependency parameter in network-based mathematical programs. This work combines these two topics by showing how interdependency parameters and coupling strategies can help overcome partial or incomplete infrastructure data.

2.1 Overcoming Inaccessible Infrastructure Data for M&S

Issues with access to infrastructure data typically stem from one of three reasons: the data is sensitive, proprietary, or lacks sufficient quality (Ouyang 2014). Sensitive infrastructure information is the type of information that could cause security concerns for a community if mishandled or inappropriately used. Geospatial coordinates of water storage access points are an example of this. Proprietary infrastructure information is the type of information that allows a private company providing an infrastructure service (e.g., drinking water, electricity, etc.) some business advantage for sole ownership and control of the information. Proprietary information is also not specifically mandated for public disclosure. Data quality concerns may stem from sparse or randomly collected data, lack of standardization in data collection, and subjective data. This last issue is deeply concerning, seeing how emerging technology uses data to inform so many decisions, and data quality is not always readily apparent.

The Cybersecurity and Infrastructure Security Agency (CISA) of the United States Government has taken steps towards securing a data repository, and they have incentivized critical infrastructure information (CII) sharing through the Protected CII (PCII) Program (CISA 2005). The PCII was initiated in 2002, with the passing of the CII Act, and updated in 2006 when additional regulations were added to ensure proper handling and use of CII (CISA). However, industries and communities are still reluctant to exchange data and/or relinquish proprietary data (Peretti 2014).

These governmental efforts are commendable; however, the data is also not widely available for use or research. Therefore, researchers and practitioners in the area of infrastructure M&S have come up with different ways to overcome the access to data issues. Ouyang (2014), in a review article on M&S for critical interdependent infrastructure, identified three workarounds: 1) empirical data harvesting from historical events, 2) random or characteristic-specific generated networks, and 3) representative data that seeks to take real systems and remove sensitive or proprietary information. While none of these are ideal, they have made substantial research and improvements possible. The present work uses the third option by using a representative dataset.

2.2 Operational Interdependency Parameter in Network-Based Programming

Rinaldi et al. (2001) identified six dimensions of infrastructure interdependent relationships; however, only two are critical for the present work. These two are interdependency type and coupling. In their study, they provided a useful classification of the types of interdependencies that affect network operations. These are physical (i.e., dependency based on the flow of materials), cyber (i.e., dependency based on the flow of information), geospatial (i.e., dependency based on proximity), and logical (i.e., any other dependency). These authors also described the coupling as being either tight or loose and either linear or complex. Tight coupling suggests a strict interdependency between systems (e.g., an electrically driven water pump). A loose coupling suggests there is an effect of one system on another, but it may not be directly felt (e.g., mining operation disruption may slow road repair and maintenance, but not immediately due to the buffer of raw material). Linear relationships behave proportionally, while complex relationships are not proportionally related or the proportions change over time.

González et al. (2016) introduced an idea of how to view these two dimensions in their presentation of the interdependent network design problem (INDP). While the authors only presented one coupling strategy, they described four variations that can cover most situations. These four strategies can be described as one-to-one, one-to-any, one-to-all, and one-to-many couplings. The authors suggested that multiple interdependency types and coupling strategies could be implemented if necessary; however, the method for employing multiple types and coupling strategies was to make independent sets of constraints with new variables and new interdependency parameters related to different types and coupling strategies. The four different coupling strategies will be discussed in more depth in the following section.

Other authors modeling interdependent infrastructure recovery have also used an interdependency parameter to describe whether infrastructure systems are interdependent. Lee et al. (2007) used a series of connector parameters which allowed them to establish node-to-node and node-to-arc relationships, both types being a one-to-one style of coupling. This formulation was a build-as-you-go type of formulation depending on what relationships were needed, and it also used special sets extensively. Cavdaroglu et al. (2013) used a binary variable equal to 1 if the slack of unmet demand at the parent node was zero, allowing the child node to be operable. This parameter did not include various interdependency types or couplings, thereby reflecting only a one-to-one relationship. Sharkey et al. (2015) used a binary variable similar to Cavdaroglu et al., except that their binary variable was arc-to-arc instead of node-to-node and didn't require all demand to be met, but rather a sufficient amount of demand. This, in essence, allowed for some degradation of service before the interdependency rendered the child node inoperable. This also represents a one-to-one relationship. In contrast to these methods, Almoghathawi et al. (2019) and Karakoc et al. (2019) used an operability variable instead of an interdependency parameter to relate physically interdependent infrastructure systems. These examples also represent a one-to-one and node-to-node relationship.

There is currently no model that employs both interdependency types and coupling strategies as an inherent part of the interdependency parameters or constraints. This paper proposes a way to implement such an integration. This modified interdependency parameter is then used to show how it can help overcome situations with partial infrastructure data.

3 METHODOLOGY

This section pulls together the formalization of coupling strategies and integrates those strategies in a combined network design and scheduling problem. First, the general notation used in the mixed-integer program (MIP) is given. Second, coupling strategies are explained in detail, given a mathematical expression, and provided with anecdotal context. Third, the MIP integrates the coupling strategies and interdependency types into the formulation to addresses the combined network design and scheduling problem.

3.1 General Notation for MIP

The combined network design and scheduling problem is based on a graph, $\mathcal{G}(\mathcal{N},\mathcal{A})$, comprised of nodes and arcs divided into layers indexed by $k \in \mathcal{K}$. Each infrastructure layer has one or more commodities indexed by $l \in \mathcal{L}^k$. The network is assumed to be damaged, which means that subsets of nodes and arcs within each layer have become inoperable. These nodes and arcs must be repaired by assigning work crews and repairing the nodes and arcs. Table 1 summarizes the relevant notation for the MIP. An additional parameter and set dealing with the integration of the coupling strategies and the interdependency types are detailed in the following subsection.

Table 1: General notation for the MIP comprising sets, variables, costs, and other parameters for flow and scheduling.

Sets	Description	Variables	Description
$\mathcal{K}$	The set of infrastructure layers.	x_{ijlt}^{k}	The variable of flow of l in arc.
$\mathcal{N}$	The set of nodes, indexed as i.	α_{ijwt}^{k} or α_{iwt}^{k}	Binary variable equal to 1 if work crew w assigned to repair arc or node.
$\mathcal{N}^{k}$	The subset of nodes.*		
$\mathcal{N}'^{k}$	The subset of damaged nodes.*	β_{ijwt}^{k} or β_{iwt}^{k}	Binary variable equal to 1 if arc or node was repaired by work crew w.
$\mathcal{A}$	The set of arcs, indexed (i,j).		
$\mathcal{A}^{k}$	The subset of arcs.*		
$\mathcal{A}'^{k}$	The subset of damaged arcs*.	y_{ijt}^{k} or y_{it}^{k}	The variable between 0 and 1 of operability of node or arc.
$\mathcal{L}^{k}$	The set of commodities.*		
$\mathcal{W}^{k}$	The set of work crews.*	$x_{ilt}^{-,k}$	The variable of unmet demand of l at node.
Ψ	The set of interdependency types, indexed as ψ.	$x_{ilt}^{+,k}$	The variable of surplus of l at node.
Ξ	The set of coupling strategies, indexed as ξ.		
$\mathcal{T}$	The set of T time periods evaluated, indexed as t.	<u>Parameters</u>	
		b_{ilt}^{k}	The amount of supply or demand of l.
<u>Costs</u>			
c_{ijlt}^{k}	The cost of flow of l in arc.	u_{ijt}^{k}	The capacity of arc for all commodities.
a_{wt}^{k}	The cost rate of assigning w.		
q_{ijt}^{k} or q_{it}^{k}	The cost of repairing arc or node.	p_{ij}^{k} or p_{i}^{k}	The processing time for repair of arc or node.
μ_{ijt}^{k} or μ_{it}^{k}	The value (cost equivalent priority) of arc or node.	μ_A or μ_B	Priority weight between 0 and 1 for objectives A and B.

* Superscript k means in infrastructure layer $k \in \mathcal{K}$; subscript t means at time period $t \in \mathcal{T}$. Asterisk is used only for sets but pertains to variables and parameters as well.

3.2 Operational Interdependency Parameter and Coupling Strategies

Let $\gamma_{i\tilde{\imath}\psi\xi}^{k\tilde{k}}$ be a parameter that takes on a value from 0 to 1, describing a parent-child relationship between parent node $i \in \mathcal{N}^{k}$ and child node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ based on some interdependency type $\psi \in \Psi$ and coupling strategy $\xi \in \Xi$. This operational interdependency parameter effectively integrates the elements of previous work, which allows for a node-to-node pairing. This parameter expands upon previous work by adding characterization of interdependency type and coupling. This means that a node can have more than one type of interdependency relationship between node pairs. This also expands the interdependency relationship of a child node to one or more parent nodes.

Before describing the coupling strategies in depth and describing how they affect the interdependency parameter, it is worthwhile to define the sets Ψ and Ξ. The set $\Psi = \{physical,\ cyber,\ logical,\ geospatial\}$, which encompasses the operational interdependency types identified by Rinaldi et al. (2001). The set $\Xi = \{one2one, one2any, one2all, one2many\}$, where each of these relationships is explained below. An additional subset, used as a filtering set, is advantageous in the programming of the MIP. Let $\mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ be a subset of nodes in a given network $k \in \mathcal{K}$, that have an operational interdependent relationship of some type ψ with another node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ in a different network $\tilde{k} \in \mathcal{K}$ based on some coupling ξ, where $\mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}} \subseteq \mathcal{N}$.

The one2one coupling describes when a child node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ can be functional only if a parent node $i \in \mathcal{N}^{k}$ is functional. This effectively means that when $\xi = one2one$, $\mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ is a singleton set for a given interdependency type ψ (Figure 1). The one2any coupling is the case when at least one of any number of nodes $i \in \mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ must be functional for the child node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ to be functional (Figure 2). The one2all coupling is where all nodes $i \in \mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ must be functional for the child node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ to be functional. This means that each one2all parent-child relationship receives an equal portion of the interdependency parameter, where the sum of all parts equals one (Figure 3). Finally, one2many coupling means that a portion (not necessarily equal) of the interdependency parameter is associated with each parent-child relationship, where the sum of all parts equals one (Figure 4). Therefore, let $\omega^{k\tilde{k}}_{i\tilde{\imath}\psi t}$ be the portion of functionality or weight between nodes $i \in \mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ and $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$, where $\sum_{i\in\mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}} \omega^{k\tilde{k}}_{i\tilde{\imath}\psi t} = \left|\mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}\right|$, $\forall \psi \in \Psi, t \in \mathcal{T}$ when $\xi = one2many$.

Each one of these coupling relationships will also depend on the operability or functionality of the parent nodes. This is represented by y^{k}_{it}, which in the present work is allowed to take on a value between 0 and 1. A parent node is inoperable when $y^{k}_{it} = 0$, partially operable when $0 < y^{k}_{it} < 1$, and fully operable when $y^{k}_{it} = 1$. In Figures 1 – 4 below, interdependent relationships are illustrated with either inoperable or fully operable nodes. Partial operability in parent nodes is reflected by partial operability in child nodes.

Table 2 summarizes these relationships and provides the mathematical representation of the interdependency parameter. The MIP is presented following this summary. It is important to note that although parent node(s) may be functional, that does not directly equate to the child node's functionality. The child node must also have its demand met, must not be damaged, or if damaged, must be repaired to be functional; therefore, in the following figures, it is stated that the child node may or may not be functional.

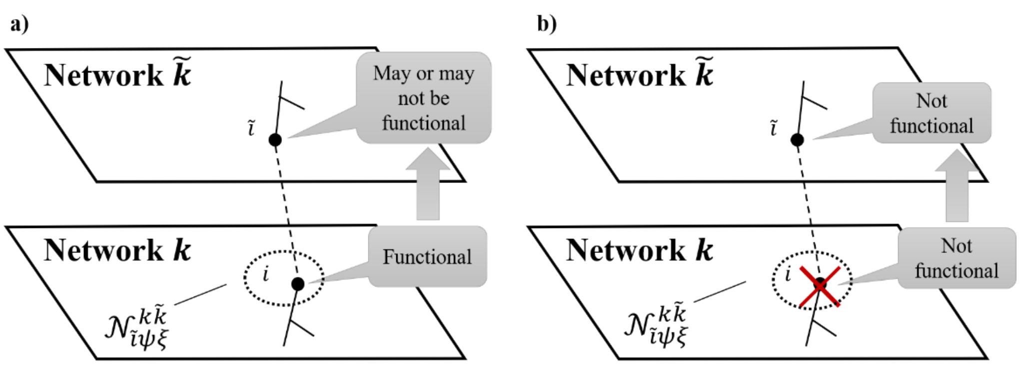

Figure 1: Illustration of one2one coupling between two infrastructures k and $\tilde{k}$; a) when node $i \in \mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ is functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ may be functional depending on other conditions; b) when node $i \in \mathcal{N}^{k\tilde{k}}_{\tilde{\imath}\psi\xi}$ is not functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ is not functional based on the interdependent relationship.

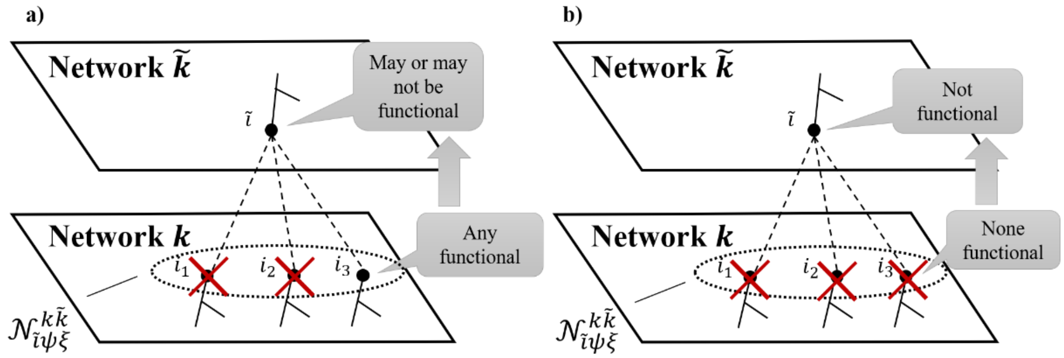

Figure 2: Illustration of one2any coupling between two infrastructures k and $\tilde{k}$; a) when any node(s) $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ may be functional depending on other conditions; b) when all nodes $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are not functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ is not functional based on the interdependent relationship.

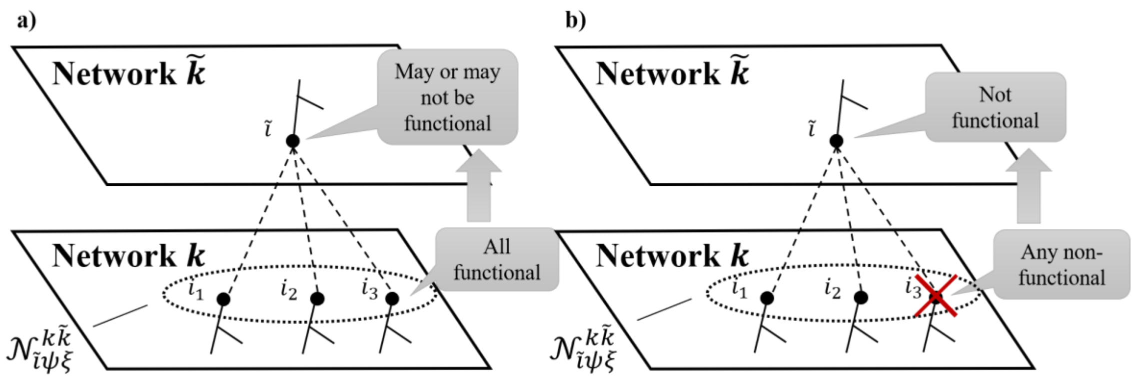

Figure 3: Illustration of one2all coupling between two infrastructures k and $\tilde{k}$; a) when all nodes $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ may be functional depending on other conditions; b) when any node $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ is not functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ is not functional based on the interdependent relationship.

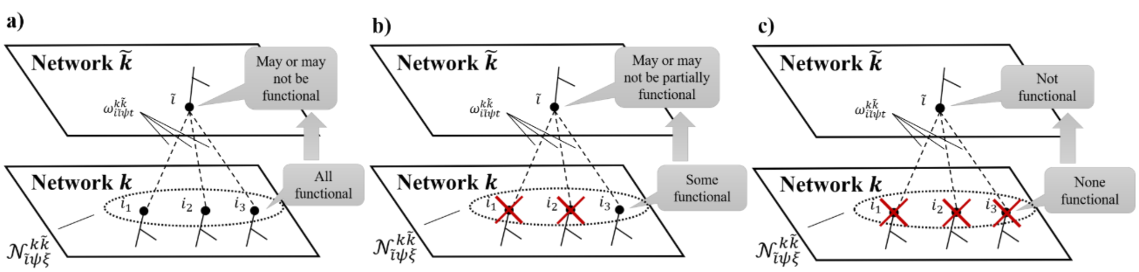

Figure 4: Illustration of one2many coupling between two infrastructures k and $\tilde{k}$; a) when all nodes $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ may be functional depending on other conditions; b) when some nodes $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ may be partially functional based on a weighting factor ($\omega_{i\tilde{\imath}\psi t}^{k\tilde{k}}$) and depending on other conditions; c) when all nodes $i \in \mathcal{N}_{\tilde{\imath}\psi\xi}^{k\tilde{k}}$ are not functional, then node $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ is not functional based on the interdependent relationship.

Of note, strict adherence to the one2all coupling relationship is most effectively achieved with binary restrictions on operability. Another method of modeling is based on the understanding that one2all relationships are multiple one2one relationships and is discussed in greater detail in the results section.

Table 2: Interdependency coupling strategies ξ affects the interdependency parameter $\gamma_{i\tilde{i}\psi\xi t}^{k\tilde{k}}$ by changing the possible values.

Coupling, ξ	Description	$\gamma_{i\tilde{i}\psi\xi t}^{kk}$ *
One2one	$\tilde{i} \in \mathcal{N}^k$ is only functional when a specific singular node $i \in \mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$ .is functional and $\mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$ is a singular set.	1
One2any	$\tilde{i} \in \mathcal{N}^k$ is functional when at least one node of a subset is functional, namely some node $i \in \mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$.	1
One2all	$\tilde{i} \in \mathcal{N}^k$ is functional only if every node from a subset $\mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$ is functional.	$\frac{1}{\lvert\mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}\rvert}$
One2many	$\tilde{i} \in \mathcal{N}^k$ depends partially on the functionality of a subset of nodes $i \in \mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$; each node $i \in \mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}$ provides a fraction of the functionality.	$\frac{\omega_{i\tilde{i}\psi t}^{k\tilde{k}}}{\lvert\mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}\rvert}$

* This holds for all $i \in \mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}, \tilde{i} \in \mathcal{N}^{\tilde{k}}, k, \tilde{k} \in \mathcal{K}, \psi \in \Psi, t \in \mathcal{T}$.

3.3 MIP Formulation

The following presentation describes the multiple objectives used in a weighted objective function followed by the applicable constraints.

$$\mathit{Cost\ Objective}\text{: } A = \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \Big(\sum_{w\in\mathcal{W}^k} \Big[\sum_{(i,j)\in\mathcal{A}'^k} \big(q_{ijt}^k \alpha_{ijwt}^k + a_{wt}^k p_{ij}^k \alpha_{ijwt}^k \big) + \sum_{i\in\mathcal{N}'^k} \big(q_{it}^k \alpha_{iwt}^k + a_{wt}^k p_i^k \alpha_{iwt}^k \big) \Big] + \sum_{l\in\mathcal{L}^k} \sum_{(i,j)\in\mathcal{A}^k} c_{ijlt}^k x_{ijlt}^k \Big) \quad (1)$$

$$\mathit{Operability\ Objective}\text{: } B = \sum_{t\in\mathcal{T}} \sum_{k\in\mathcal{K}} \Big(\sum_{i\in\mathcal{N}^k} \mu_{it}^k y_{it}^k + \sum_{(i,j)\in\mathcal{A}^k} \mu_{ijt}^k y_{ijt}^k \Big) \quad (2)$$

Minimize

$$\mu_A A - \mu_B B. \quad (3)$$

Subject to

$$\sum_{j:(i,j)\in\mathcal{A}^k} x_{ijlt}^k - \sum_{j:(j,i)\in\mathcal{A}^k} x_{jilt}^k = b_{ilt}^k + x_{ilt}^{-,k} - x_{ilt}^{+,k},\ \forall i \in \mathcal{N}^k, l \in \mathcal{L}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (4)$$

$$\sum_{l\in\mathcal{L}^k} x_{ijlt}^k \le u_{ijt}^k y_{it}^k,\ \forall (i,j) \in \mathcal{A}^k, i \in \mathcal{N}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (5)$$

$$\sum_{l\in\mathcal{L}^k} x_{ijlt}^k \le u_{ijt}^k y_{jt}^k,\ \forall (i,j) \in \mathcal{A}^k, j \in \mathcal{N}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (6)$$

$$\sum_{l\in\mathcal{L}^k} x_{ijlt}^k \le u_{ijt}^k y_{ijt}^k,\ \forall (i,j) \in \mathcal{A}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (7)$$

$$y_{it}^k \le \sum_{w\in\mathcal{W}^k} \sum_{\tau=1}^{t} \beta_{iw\tau}^k,\ \forall i \in \mathcal{N}'^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (8)$$

$$y_{ijt}^k \le \sum_{w\in\mathcal{W}^k} \sum_{\tau=1}^{t} \beta_{ijw\tau}^k,\ \forall (i,j) \in \mathcal{A}'^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (9)$$

$$\sum_{t\in\mathcal{T}} \sum_{w\in\mathcal{W}^k} \beta_{iwt}^k \le 1,\ \forall i \in \mathcal{N}'^k, k \in \mathcal{K}. \quad (10)$$

$$\sum_{t\in\mathcal{T}} \sum_{w\in\mathcal{W}^k} \beta_{ijwt}^k \le 1,\ \forall (i,j) \in \mathcal{A}'^k, k \in \mathcal{K}. \quad (11)$$

$$\sum_{t\in\mathcal{T}} \sum_{w\in\mathcal{W}^k} \alpha_{iwt}^k \le 1,\ \forall i \in \mathcal{N}'^k, k \in \mathcal{K}. \quad (12)$$

$$\sum_{t\in\mathcal{T}} \sum_{w\in\mathcal{W}^k} \alpha_{ijwt}^k \le 1,\ \forall (i,j) \in \mathcal{A}'^k, k \in \mathcal{K}. \quad (13)$$

$$\beta_{iwt}^k \le \sum_{\tau=1}^{\min[T,t-p_i^k]} \alpha_{iw\tau}^k,\ \forall i \in \mathcal{N}'^k, w \in \mathcal{W}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (14)$$

$$\beta_{ijwt}^k \le \sum_{\tau=1}^{\min[T,t-p_{ij}^k]} \alpha_{ijw\tau}^k,\ \forall (i,j) \in \mathcal{A}'^k,\ w \in \mathcal{W}^k, k \in \mathcal{K}, t \in \mathcal{T}. \quad (15)$$

$$\sum_{\tau=1}^{\min[T,t+p_i^k-1]}\sum_{i\in\mathcal{N}'^k}\alpha_{iw\tau}^k+\sum_{\tau=1}^{\min[T,t+p_{ij}^k-1]}\sum_{(i,j)\in\mathcal{A}'^k}\alpha_{ijw\tau}^k\leq 1+\sum_{\tau=p_i^k+1}^{t}\sum_{i\in\mathcal{N}'^k}\beta_{iw\tau}^k+\sum_{\tau=p_{ij}^k+1}^{t}\sum_{(i,j)\in\mathcal{A}'^k}\beta_{ijw\tau}^k,\ \forall w\in\mathcal{W}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (16)$$

$$\sum_{i\in\mathcal{N}_{\tilde{i}\psi\xi}^{k\tilde{k}}}\gamma_{i\tilde{i}\psi\xi t}^{k\tilde{k}}y_{it}^k\geq y_{\tilde{i}t}^{\tilde{k}},\ \forall\tilde{i}\in\mathcal{N}^{\tilde{k}},\tilde{k}\in\mathcal{K},\psi\in\Psi,\xi\in\Xi,t\in\mathcal{T}. \quad (17)$$

$$y_{it}^k b_{ilt}^k\geq b_{ilt}^k+x_{ilt}^{-,k},\ \forall i\in\mathcal{N}_D^k,l\in\mathcal{L}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (18)$$

$$0\leq x_{ijlt}^k\leq u_{ijt}^k,\ \forall(i,j)\in\mathcal{A}^k,l\in\mathcal{L}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (19)$$

$$x_{ilt}^{-,k}\geq 0,\ \ \forall i\in\mathcal{N}^k,l\in\mathcal{L}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (20)$$

$$0\leq y_{it}^k\leq 1,\ \forall i\in\mathcal{N}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (21)$$

$$0\leq y_{ijt}^k\leq 1,\ \forall(i,j)\in\mathcal{A}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (22)$$

$$\alpha_{iwt}^k\in\{0,1\},\ \forall i\in\mathcal{N}'^k,w\in\mathcal{W}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (23)$$

$$\alpha_{ijwt}^k\in\{0,1\},\ \forall(i,j)\in\mathcal{A}'^k,w\in\mathcal{W}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (24)$$

$$\beta_{iwt}^k\in\{0,1\},\ \forall i\in\mathcal{N}'^k,w\in\mathcal{W}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (25)$$

$$\beta_{ijwt}^k\in\{0,1\},\ \forall(i,j)\in\mathcal{A}'^k,w\in\mathcal{W}^k,k\in\mathcal{K},t\in\mathcal{T}. \quad (26)$$

Equation (1) includes repair and assignment costs for damaged arcs and nodes, followed by the flow costs. Equation (2) represents a weighted operability, which is set as a competing objective in (3). Basic flow balance is shown in (4). Multicommodity flow is capacitated and flow is restricted in three different ways based on operable start-nodes, end-nodes, and arcs in (5-7), respectively. A repaired asset can become operable, as shown in (8-9). Assets can only be repaired once and assigned to one crew, as shown in constraints (10-11) and (12-13), respectively. Damaged assets are only repaired after they have been assigned and sufficient processing time has occurred, as shown in (14-15). Constraint (16) shows work crews may only be assigned to one repair task at a time.

Constraint (17) represents the operational interdependency constraint, which uses the interdependency parameter to determine child node operability. Constraint (18) suggests that a node is proportionally operable to the met amount of demand. The constraints (19-26) represent the side constraints based on variable definitions.

4 RESULTS AND DISCUSSION

This section describes the infrastructure network, the missing telecommunications data, and the results when comparing optimization results with full and partial datasets.

4.1 Simulated Military Base

Using the CLARC database as a starting point, the data was reduced to about 10% of the original size while still preserving the diversity of operations and asset types (T. Sharkey et al. 2018). This was done to recreate a representative military base with bi-directional system-to-system interdependencies inherent in the CLARC database. The resultant reduced dataset was then constructed in a multiplex fashion, reflecting nodes into layers where they had a demand, supply, or transshipment function.

An issue with the telecommunication infrastructure data was found due to only 4 of 47 different nodal asset types having any communication demand. For example, facilities such as Fire Stations, Police Stations, Schools, Hospitals, and others had no communication connections (i.e., arcs) and no demand. However, these facilities are essential in recovery operations and are controlled largely by communicating with an Emergency Control Center (Lee et al. 2007). This issue represents partial infrastructure data within a given layer, which was overcome using two different methods.

4.2 Overcoming Telecommunications Data Gap with Operational Interdependencies

The first method to overcome the partial telecommunication data represents working with data owners and receiving the necessary data. This was accomplished by creating a geospatial context for the reduced dataset and physically drawing each connection to create a full representation of the complete infrastructure systems. This became the full dataset. The second method used the partial data provided and created various interdependency relationships to influence operability in lieu of acquiring additional infrastructure data. This became the partial dataset with additional interdependencies. The cost to produce such interdependencies is the time to communicate with stakeholders on the actual or perceived connection and dependency to establish the appropriate coupling relationship. The number of additional interdependencies needed will be dependent on the amount of infrastructure data missing.

An example of overcoming missing infrastructure data by using an interdependency is a Fire Station that requires communication to receive 911 emergency calls. If this service is not available, then the emergency responders will not respond because they are unaware of the call. Thus, the operability of one of two telecommunication nodes (part of the partial telecommunications data) would allow the Fire Station to remain as a supply node for the fire and emergency service commodity. However, if both telecommunication nodes were inoperable, then the Fire Station would also be inoperable since this represents no ability to send and receive 911 emergency calls. While actual systems have additional backups, this is used for illustration purposes and as a proof of concept.

This example of the Fire Station depending on the telecommunication network is an example of a one2any coupling based on a cyber (i.e., data and information flow) type interdependency. This process was applied to every node that should have a communications demand within a full dataset. The result was three variations of the network: 1) dataset with full telecommunication data, 2) dataset with partial telecommunication data and additional interdependency relationships, and 3) dataset with partial telecommunication data without additional interdependency relationships. The third set serves as a basis to judge the addition of interdependent relationships to overcome infrastructure data gaps.

4.3 Comparison of Optimization Results

Comparing the full dataset and the partial dataset with additional interdependencies shows the use of interdependencies as a viable option for overcoming partial data. The time horizon for this comparison is 12 8-hour time periods. While not the primary focus of this research, the model was programmed in GAMS v31.1.1 and used CPLEX 12.10. All tests were conducted on a desktop computer with an Intel Xeon CPU E5-1620 operating at 3.60 GHz with 16 GB of RAM. The average computational time for the tests with partial data and additional interdependencies averaged at less than 8 mins, while the tests with the full dataset averaged at 18 mins.

Table 3 summarizes the number of nodes, arcs, and interdependent relationships between the two different simulations. The full dataset represents 227 more nodes and arcs than the partial dataset, whereas the partial dataset with additional interdependencies represents 102 more interdependency relationships than the full dataset. The same damage was simulated in both simulations, even though additional arcs or nodes that were not in the partial dataset could have been damaged in the full dataset.

Table 3: The full dataset represents more nodes and arcs, while the partial dataset represents more interdependency relationships.

Feature	Full Dataset	Partial Dataset
Nodes	507	432
Arcs	886	734
Interdependencies	123	225

The two different datasets were evaluated over varying objective function weights, establishing Pareto optimal values or a Pareto front. Due to the disparity in the number of assets, the overall operability objective value for the full dataset was 1.25 times higher than that of the partial dataset across the Pareto fronts. There was one anomaly when cost was weighted the most and operability the least (i.e., $\mu_A = 0.9, \mu_B = 0.1$), which resulted in the operability objective function value being 1.57 times greater than the partial dataset. After acknowledging the slight difference in the magnitude of the operability objective function values, the overall trends were identical.

In the case with balanced objective functions (i.e., $\mu_A = \mu_B = 0.5$), the full dataset showed an increase in operability from 65.8%, representing immediate operability following the disruption, to 68.1% within the first four time periods. Then the model showed a significant jump in operability at time period 5 to 89.4%, where it remained for the time periods being evaluated. This signifies that the bulk of the optimal recovery trying to balance operability and cost was achieved by time period 5, or 40 hours following the disruption, based on 8-hr time periods. The partial dataset showed similar trends, with slight deviation in the percent operable. The partial dataset showed an increase in operability of 65.6% to 68.2% in the first four time periods and an increase to 91.8% at time period 5 and beyond. The partial dataset deviation from the full dataset in the first four time periods ranged from -0.2% to +0.3%. With the jump in operability at time period 5 the percent deviation also increased to +2.6% from time period 5 on. Partial data without the additional interdependencies underestimated the recovery from as great as -6.0% to as little as -3.8%, never achieving as accurate results as the partial dataset with additional interdependencies. Figure 5 illustrates how the partial dataset with an increased number of interdependency relationships closely approximates the operability of the system during recovery. The final operability percentage in these scenarios ranged from 86.1% to 91.8% and didn't progress to 100% operability due to the presence of redundant flow pathways and the desire to balance cost and operability. Additionally, nodes and arcs that have extremely low value, denoted by μ_{it}^k or μ_{ijt}^k, and high costs repair costs, denoted by q_{it}^k or q_{ijt}^k, tend to be excluded from optimal results. This can be beneficial to emergency repair crews to ensure emphasis on the critical aspects of the system, prior to addressing non-critical components.

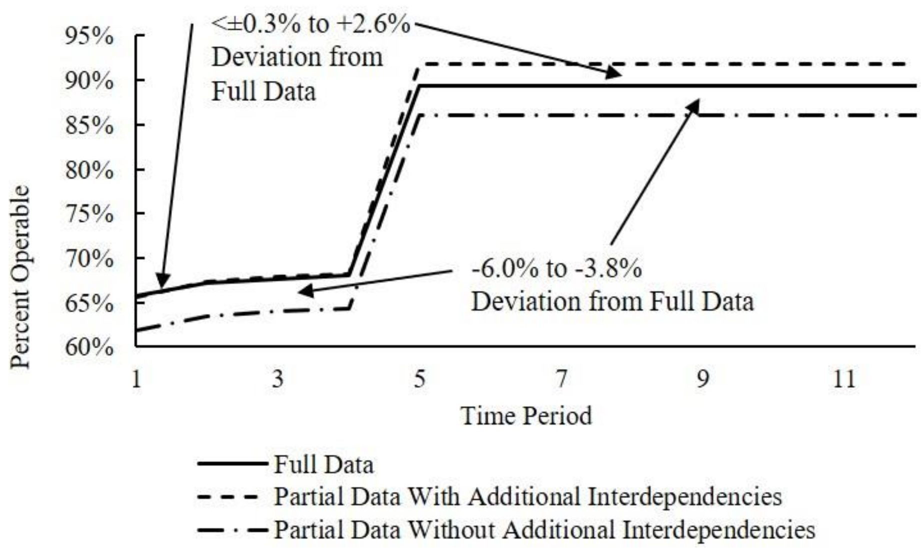

Figure 5: Partial data simulation with additional interdependencies more closely approximated a full dataset than the partial data without additional interdependencies.

The partial dataset employed only one2one and one2any coupling strategies since this most accurately reflected the same relationships that existed in the full dataset. The partial dataset scenario was also run by modifying the MIP to restrict the operability variables, y_{it}^k and y_{ijt}^k, to binary values with no significant changes to the operability objective value, being within 3% at the greatest point of deviation. In fact, the strict adherence to the one2one, one2any, and one2all coupling strategies may be best seen when operability is modeled as binary variables. If operability is modeled as binary variables, the same formulation as presented above holds for all coupling strategies except one2many, which inherently is incompatible with binary operability variables.

In contrast, the inclusion of all the coupling strategies with a non-binary operability variable, as in the current work, also becomes problematic when desiring strict adherence to all the coupling strategies. The use of non-binary operability variables means that child node partial operability is possible based on parent node partial operability. Effectively this creates an upper bound on child node operability based on full or partial parent node operability and the associated coupling strategy. A child node in one2one relationships has an upper bound based on the parent node operability. A child node in one2any relationships may be fully operable so long as one node is fully operable or the sum of all parent nodes' partial operability amount to one or more. A child node in one2all relationships has an upper bound of some fraction of parent node partial operability. A child node in one2many relationships has an upper bound of some partial operability based on the sum of partial operability of the parent nodes.

A comparative example between binary and non-binary operability variables for one2all relationships illustrates the difference. A one2all coupling between three parent nodes and one child node results in an inoperable child node if any one of the three parent nodes is inoperable when operability is binary. In the case of non-binary operability, the node may experience operability up to 2/3 operability based on one node being inoperable and the other two being fully operable. To achieve strict adherence to the one2all coupling strategy with non-binary variables, a modification is made to constraints (17) by removing the summation over the set $\mathcal{N}_{i\tilde{\imath}\psi\xi}^{k\tilde{k}}$. This can be accomplished by employing conditional constraint generation when programming the MIP.

Despite the need to slightly adjust the MIP presentation to accommodate one2all relationships, the use of non-binary operability variables adds a significant level of reality to the simulation. In very few instances will the termination of telecommunication services result in complete inoperability. Therefore, partial operability is a closer approximation to reality. This also allows the use of a pseudo node which can establish a baseline operability level regardless of the loss of service. For example, if an industry is still 80% operable with the loss of internet and telephone services. A one2many relationship can exist between $\tilde{\imath} \in \mathcal{N}^{\tilde{k}}$ and any number of nodes $i \in \mathcal{N}^{k}$, with 80% of the weight times the cardinality of the set $\mathcal{N}_{i\tilde{\imath}\psi\xi}^{k\tilde{k}}$ for some interdependency type $\psi \in \Psi$ residing in the relationship with pseudo node $i^{*} \in \mathcal{N}^{k}$.

An additional scenario was built based on the partial dataset, which included a partial operability baseline of 80% despite lack of telecommunication services except for the emergency responders, which rely on telecommunications to send and receive 911 emergency calls. This scenario resulted in a near-perfect match because only three facilities in the power infrastructure system met the conditions to have 80% operability versus being reduced to zero. A different damage scenario could highlight this better, but consistency for comparison was chosen over introducing a different damage scenario.

This shows the ability to incorporate all the various coupling strategies and leverage the one2many relationship to help model complex relationships that result in some impact to operability but do not render a node inoperable. This effectively assigns a lower bound to operability based on interdependencies.

During the construction of these datasets, it was assumed and then shown in analysis that this model's applicability only worked if the actual known telecommunication nodes were damaged or inoperable. Suppose the service disruption was from a telecommunication node in the partial and full datasets downstream to the point of interest, thereby only belonging to the full dataset. In that case, this method could not show similar disruption as can be seen in the full dataset. This lack of granularity points to the limitations of using interdependencies in lieu of a full dataset.

5 CONCLUSION

This paper detailed issues concerning access to data and then highlighted how interdependencies could be leveraged to overcome partial infrastructure data. This was shown in using a representative full and partial dataset for a military base-sized system of networks. The results showed comparable operability projections between the two methods. Additionally, some flexibility was gained to model complex interactions by using more robust interdependencies. The modification to commonly used interdependency parameters integrated multiple interdependency types and coupling strategies, which had not been done as an inherent part of a model before this work. Some limitations exist in not capturing the same granularity of knowledge on damaged assets that can be gleaned from full datasets.

This research was completed as part of doctoral research by the primary author. The views expressed in this study are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

REFERENCES

Almoghathawi, Y., K. Barker, and L. A. Albert. 2019. "Resilience-Driven Restoration Model for Interdependent Infrastructure Networks." *Reliability Engineering and System Safety* 185:12–23.

American Society of Civil Engineers. 2019. *Future World Vision: Infrastructure Reimagined*. American Society of Civil Engineers.

Cavdaroglu, B., E. Hammel, J. E. Mitchell, T. C. Sharkey, and W. A. Wallace. 2013. "Integrating Restoration and Scheduling Decisions for Disrupted Interdependent Infrastructure Systems." *Annals of Operations Research* 203(1):279–94.

Cybersecurity and Infrastructure Security Agency (CISA). n.d. Protected Critical Infrastructure Information (PCII) Program. https://www.cisa.gov/pcii-program, accessed 26th March 2021.

Cybersecurity and Infrastructure Security Agency (CISA). 2005. "Protected Critical Infrastructure Information (PCII) Program Fact Sheet." CISA Fact Sheet.

González, A. D., L. Dueñas-Osorio, M. Sánchez-Silva, and A. L. Medaglia. 2016. "The Interdependent Network Design Problem for Optimal Infrastructure System Restoration." *Computer-Aided Civil and Infrastructure Engineering* 31(5):334–50.

Karakoc, D. B., Y. Almoghathawi, K. Barker, A. D. González, and S. Mohebbi. 2019. "Community Resilience-Driven Restoration Model for Interdependent Infrastructure Networks." *International Journal of Disaster Risk Reduction* 38:101228.

Lee, E. E., J. E. Mitchell, and W. A. Wallace. 2007. "Restoration of Services in Interdependent Infrastructure Systems: A Network Flows Approach." *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 37(6):1303–17.

Ouyang, M. 2014. "Review on Modeling and Simulation of Interdependent Critical Infrastructure Systems." *Reliability Engineering and System Safety* 121:43–60.

Peretti, K. 2014. "Cyber Threat Intelligence: To Share or Not to Share?" Bureau of National Affairs.

Rinaldi, S. M., J. P. Peerenboom, and T. K. Kelly. 2001. "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies." *IEEE Control Systems Magazine* 21(6):11–25.

Rutgers, V., and B. Sniderman. 2018. "Around the Physical-Digital-Physical Loop: A Look at Current Industry 4.0 Capabilities." In *Deloitte Insights*, edited by A. Rao, B. Hurley, R. Bhat, and A. Khan, 1-5. Deloitte Development LLC.

Sharkey, T. C., B. Cavdaroglu, H. Nguyen, J. Holman, J. E. Mitchell, and W. A. Wallace. 2015. "Interdependent Network Restoration: On the Value of Information-Sharing." *European Journal of Operational Research* 244(1):309–21.

Sharkey, T., N. Ni, R. Little, R. Loggins, W. Wallace, and S. Nurre. 2018. "CLARC: An Artificial Community for Modeling the Effects of Extreme Hazard Events on Interdependent Civil Infrastructure Systems." Design Safe Data Depot. https://doi.org/10.17603/DS2FX2D, accessed 18th February 2021.

AUTHOR BIOGRAPHIES

BRIGHAM MOORE is a Ph.D. candidate in the Department of Systems Engineering and Management at the Air Force Institute of Technology. His research focuses on interdependent infrastructure recovery following disruptive events. His email address is brigham.moore@afit.edu.

DAVID JACQUES is a Professor of Systems Engineering and Management at the Air Force Institute of Technology. His research primarily focuses on concept definition and system analysis, optimal system design, cooperative behavior, and control of semi-autonomous flight vehicles. His email address is david.jacques@afit.edu. His website is https://www.afit.edu/BIOS/bio.cfm?facID=94.

STEVEN SCHULDT is an Assistant Professor of Engineering Management at the Air Force Institute of Technology. His research interests include construction management, multi-objective optimization, infrastructure resilience, and sustainability. His email address is steven.schuldt.1@us.af.mil, and his website is http://scholar.google.com/citations?user=9KimdnMAAAAJ&hl=en.