Evaluating the Optimality of Dynamic Coupling Strategies in Interdependent Network Systems

I-Cheng Lin
Electrical & Computer Engineering
Carnegie Mellon University
Pittsburgh, PA
ichengl@andrew.cmu.edu

Osman Yağan
Electrical & Computer Engineering
Carnegie Mellon University
Pittsburgh, PA
oyagan@andrew.cmu.edu

Carlee Joe-Wong
Electrical & Computer Engineering
Carnegie Mellon University
Pittsburgh, PA
cjoewong@andrew.cmu.edu

***Abstract*—Cascading failures are a common phenomenon in complex networked systems, where failures at only a few nodes may trigger a process of sequential failures. We investigate the robustness against cascading failures in systems carrying flows or loads that contain multiple interdependent networks, e.g., power grid, transportation system, etc. In these systems, the *coupling* coefficients between the networks, which determine how the flow from failed components gets redistributed across the networks, is a key factor affecting the robustness against cascading failures. Prior work has introduced the step-wise optimization (SWO) strategy that dynamically adjusts the coupling coefficients during the course of the cascading failures in an effort to preserve the network size. SWO has been shown to have good performance against cascading failures on synthetic data. In this paper, we show the optimality of the SWO strategy under certain conditions on the flow and capacity distributions of the nodes. We also show, via simulations, that the SWO strategy performs well under various real-world network topologies as well.**

***Index Terms*—cascading failure, interdependent networks**

I. INTRODUCTION

Modern systems such as the communication networks connecting the Internet of Things (IoT), power grids, and urban transportation networks form *interdependent* large-scale network systems. Recently, there has been significant research interest in the analysis and optimization of robustness of such systems, with a major focus on the phenomenon called *cascading failures*. This phenomenon represents cases where failures initiated in a small part of the network trigger a process of sequential failures that may eventually cause a disastrous impact on the whole system. Many real-world systems, including the power grid [1], railway networks [2], communication networks, IoT [3], etc., can be subject to cascading failures, e.g., see [4], [5] for additional examples.

Much of the existing literature on cascading failures focuses on the robustness of a single network [4], [6], [7]. However, cascading failures are particularly likely to take place in systems consisting of *multiple*, interdependent networks. For example, urban transportation systems may include road networks, bus systems, a subway systems, other railway systems, and bike-sharing systems. These networks are interdependent in the sense that a failure in one of them will likely lead to an increased load (e.g., by passengers using alternatives modes of transportation) in others. The robustness of such interdependent networks was studied in [8]–[14], where it was shown that interdependent networks can be more vulnerable to cascading failures than isolated networks [15].

In this paper, we focus on networks that carry a load (or, a flow) and apply a flow/load redistribution model to investigate the robustness of the interdependent networks. The coupling coefficients that define the portion of flow/load to be transferred between networks are a crucial factor determining the robustness of such systems [16]. We are interested in cases where these coefficients can be adjusted *dynamically* in response to network failures. For example, passengers can be given time-varying incentives to take one mode of transportation instead of another after a road or subway link fails, as studied in [17]. Intuitively, dynamic coupling policies are expected to outperform static ones (where the coupling across networks remains the same throughout the cascading failures). However, it is difficult to analyze the optimality of dynamic coupling policies, as we need to track the system's dynamics and not just its asymptotic outcome, as done in prior work [18]. This analysis is particularly difficult in real networks, where we cannot necessarily characterize the system in terms of the mean statistics over several nodes: real-world transportation networks, for instance, have relatively few nodes, which may invalidate the common use of the law of large numbers to track system dynamics.

In this work, we consider the step-wise optimization (SWO) strategy introduced in [18], which minimizes the total extra load (from the failed nodes) that needs to be redistributed in the subsequent step. Our previous work [18] creates a framework to dynamically adjust the coupling coefficients and characterize the resulting final system size. In this work, we go a step further to show the effectiveness of our proposed strategy with the following contributions:

- We show that the SWO strategy maximizes the final surviving number of nodes when the distribution of *free-space* (i.e., capacity minus initial load) in each node follows a Uniform distribution (see Section IV).
- We show that the SWO strategy achieves comparable performance to that of the optimal fixed-coupling coefficients (FCC) strategy (see Section V) in extensive simulations.
- We use real-world transportation networks from several cities around the world to show the effectiveness of the SWO strategy in real networked systems (see Section V).

II. Network Model

Consider the set of networks $\mathcal{N} = \{A, B\}$. There are three characteristics for each node i, which are its capacity C_i (indicating the maximum load the node can handle), load L_i, and free space $S_i = C_i - L_i$.

The initial load and free space of the nodes in both networks are assumed to be drawn from given (possibly different) probability distributions. We assume that both networks can be subject to an *initial* failure that leads to the removal of a certain fraction of their nodes, chosen uniformly at random. We list the initial parameters defining the system in Table I.

List of initial parameters	
Notation	Definition
p_A, p_B	Initial portion of nodes failed in networks A and B. (Equivalent to the probability that each node fails.)
N_A, N_B	Initial number of nodes of network A and B.
L_A, L_B	Initial load of a node in network A and B.
S_A, S_B	Free space of a node in network A and B, each following certain probability distribution.

TABLE I: Initial Parameters of the System

When a node fails, it is assumed that the load it carries will be redistributed to other nodes. For example, suppose a subway station shuts down due to construction. In that case, the load carried by that node must then be redistributed to other adjacent nodes, becoming part of their load, e.g., passengers must instead walk to nearby stations, which must now serve them. This redistribution might lead to further node failures by virtue of their increased load exceeding their capacity, i.e., their free-space falling below zero. For example, a station may become so crowded that passengers cannot board trains there.

In an interdependent network system, the extra load, i.e., the load carried by those nodes that recently failed, can be redistributed in the same network or in other networks, which captures the phenomenon of the load being shed to the coupled networks. We formalize this idea by considering a sequence of discrete time-steps $t = 1, 2, \ldots$. We define the coupling coefficients at time t as $1 - \alpha(t)$ and $1 - \beta(t)$, which respectively define the portion of load that is redistributed from network A (resp. network B) to network B (resp. network A).

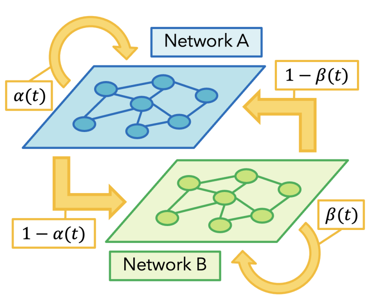


Fig. 1: Illustration of the two networks system.

A. Cascading Failure Process

In this work, we are interested in the phenomenon "cascading failure," which is the sequential failure triggered by the initial failure of a small portion of the network. We track the node failures in each time step $t = 1, 2, \ldots$. A node will fail if its total load after receiving the extra load exceeds the node's capacity. That is, for the i-th node in network A to fail at time step t, the total load it carried before receiving the extra load $(L_{A_i}(t-1))$ should be less than its capacity C_{A_i}, ensuring it does not fail before time step t. Further, $L_{A_i}(t) > C_{A_i}$ since node i's total load exceeds its capacity after receiving the extra load. If the node fails at time t, its load will be redistributed to other surviving nodes at the next time step $t+1$. This process continues until no more nodes fail, either because there is no extra load to be redistributed or all nodes have already failed.

The relation between the initial attack size and the final system size (i.e., number of surviving nodes) is a monotonic decreasing function. When the initial attack size reaches a certain threshold, prior work [4], [16] has shown that the final size of the system transitions between two phases: one where some nodes survive and one where none survive. We define the critical attack size as our metric for the network robustness, which specifies the minimal initial attack size (the expected initial attack portion) for which all nodes eventually fail.

B. Time-Dependent Parameters and Load Update Rules

With the initial system parameters given in Table I, we now define various time-dependent system parameters that specify the system's status. These parameters are listed in Table II for network A. The same notation is used for network B by changing the subscript from A to B.

We follow the widely used [6], [19], [20] fiber-bundle model, which has been used to investigate the breakdown of a broad class of systems [4], [16], [21], and we assume both networks are fully connected. Our prior work [18] shows that this model generalizes to networks that are not fully-connected. Under these conditions, the load to be redistributed in network A at time t will consist of the internally redistributed load, and the load from network B. $\Delta L_A(t)$, the extra load that a single node received in network A, will be this sum of this load divided by the number of surviving nodes. That is:

$$\Delta L_A(t) = \frac{F_{At}\alpha(t) + F_{Bt}(1-\beta(t))}{N_{At}} \tag{1}$$

Thus, for a single node j in network A, the load update is:

$$L_{A,j}(t) = L_{A,j}(t-1) + \frac{F_{At}\alpha(t) + F_{Bt}(1-\beta(t))}{N_{At}} \tag{2}$$

Similar equations can be derived for the extra load experienced by nodes in network B.

III. Mathematical Background

Under our assumptions of fully-connected networks and an equal redistribution model with a large number of nodes, we can use mean-field analysis to specify the system status by tracking the surviving portion and average extra load of nodes in each network. We give a brief overview of mathematical analysis step by step in this section, following our prior work [18]. In the next section, we use this analysis to analyze the optimality of the SWO strategy.

List of parameters of the system at time t for network A	
Notation	Definition
t	Current time step. Initially $t = 0$, after each iteration, the time-step is incremented by one.
$\alpha(t), \beta(t)$	The in-net ratio at time t for networks A and B respectively.
N_{At}	Number of surviving nodes in network A.
f_{At}	Fraction of failed nodes up to time t in network A.
F_{At}	Total extra load from network A.
Q_{At}	Cumulative average extra load distributed at a single node in network A up to time t.
ΔQ_{At}	Average extra load distributed at a single node in network A at time t.

TABLE II: System Parameters at time t

A. Initial Conditions

Following the notation in Table I, Definition 3.1 summarizes the initial condition of the network after the initial attack.

Definition 3.1 (Initial Conditions): Consider the system after the initial attack. For $\mathcal{X} \in \{A, B\}$, the status of the system, including the failing portion $f_{\mathcal{X}0}$, number of surviving nodes $N_{\mathcal{X}0}$, total extra load $F_{\mathcal{X}0}$ and average extra load being redistributed per node $Q_{\mathcal{X}0}$ can be written as:

$$\begin{cases} f_{\mathcal{X}0} &= p_{\mathcal{X}} \\ N_{\mathcal{X}0} &= (1 - f_{\mathcal{X}0})N_{\mathcal{X}} \\ F_{\mathcal{X}0} &= N_{\mathcal{X}} \cdot f_{\mathcal{X}0} \cdot \mathbb{E}[L_{\mathcal{X}}] \\ Q_{A0} &= \Delta Q_{A0} = \frac{\alpha(0) \cdot F_{A0} + (1-\beta(0) \cdot F_{B0})}{(1-f_{A0}) \cdot N_A} \\ Q_{B0} &= \Delta Q_{B0} = \frac{\beta(0) \cdot F_{B0} + (1-\alpha(0) \cdot F_{A0})}{(1-f_{B0}) \cdot N_B} \end{cases} \tag{3}$$

$\alpha(0)$ and $\beta(0)$ denote the internal load redistribution portion for network A and network B respectively at time 0. Detailed derivations can be found in our prior work [18].

B. Recursive Equations

To trigger the cascading failure process, at least one of the two networks should have their average extra load exceed the minimum of the nodes' free spaces, denoted as S_{A0} and S_{B0}. In other words, the necessary and sufficient condition for the cascading failure process to start is that at least one of the conditions $Q_{A0} \geq S_{A0}$ or $Q_{B0} \geq S_{B0}$ is satisfied.

From time step $t \geq 1$, we have the following Theorem 3.1.

Theorem 3.1 (Recursive Equations for the System): For $\mathcal{X} \in \{A, B\}$, time step $t \geq 1$, the status of the system can be written as the following recursive equations.

$$\begin{cases} f_{\mathcal{X}t} &= 1 - (1 - p_{\mathcal{X}}) \cdot P[S_{\mathcal{X}} \geq Q_{\mathcal{X}(t-1)}] \\ N_{\mathcal{X}t} &= (1 - f_{\mathcal{X}t})N_{\mathcal{X}} \\ F_{\mathcal{X}t} &= N_{\mathcal{X}} \cdot (f_{\mathcal{X}t} - f_{\mathcal{X}(t-1)}) \cdot \mathbb{E}[L_{\mathcal{X}} + Q_{\mathcal{X}(t-1)}] \\ &= N_{\mathcal{X}} \cdot (1 - p_{\mathcal{X}}) \cdot P[Q_{\mathcal{X}(t-2)} < S_{\mathcal{X}} \leq Q_{\mathcal{X}(t-1)}] \cdot \mathbb{E}[L_{\mathcal{X}} + Q_{\mathcal{X}(t-1)}] \\ \Delta Q_{At} &= \frac{\alpha(t) \cdot F_{At} + (1-\beta(t) \cdot F_{Bt})}{(1-f_{At}) \cdot N_A} \\ \Delta Q_{Bt} &= \frac{\beta(t) \cdot F_{Bt} + (1-\alpha(t) \cdot F_{At})}{(1-f_{Bt}) \cdot N_B} \\ Q_{\mathcal{X}t} &= Q_{\mathcal{X}(t-1)} + \Delta Q_{\mathcal{X}t} \end{cases} \tag{4}$$

The final state of the system after the cascading failure process is when the system size (i.e., number of surviving nodes) does not change. More formally:

$$N_{At} - N_{A(t-1)} = 0, \quad N_{Bt} - N_{B(t-1)} = 0 \tag{5}$$

C. Step-wise Optimization & the Objective Function

When we attempt to stop a cascade of failures, we wish to maximize the number of surviving nodes (i.e., minimize the number of failed nodes). We first note that at any given time, the total free space and the number of surviving nodes have already been determined by prior failures. It has been shown in [18] that step-wise optimization strategy minimizing the total extra load to be redistributed at the subsequent time step is an effective algorithm to prevent cascading failures. The optimization problem can be written as Definition 3.2:

Definition 3.2: (*The Step-wise Optimization (SWO) Problem*): In each step of the cascading failure process, greedily minimize the expected extra load in the next time step. The optimization problem at each time step can be formulated as:

$$\min_{\alpha,\beta} \left(\sum_{\mathcal{X}=A,B} \mathbb{E}\left[L_{\mathcal{X}} + Q_{\mathcal{X}(t+1)}\right] \Delta N_{\mathcal{X}(t+1)} \right) \tag{6}$$

subject to:

$$0 \leq \alpha(t) \leq 1, \qquad 0 \leq \beta(t) \leq 1 \tag{7}$$

Here the expectation of the load on a single node is:

$$\mathbb{E}\left[L_{\mathcal{X}} + Q_{\mathcal{X}(t+1)}\right] = \mathbb{E}\left[L_{\mathcal{X}} + Q_{\mathcal{X}t} + \Delta Q_{\mathcal{X}(t+1)}\right] \tag{8}$$

Based on different choices of the free space distribution, the optimization problem can be either convex or not [18]. When the free space distribution is Uniform, the SWO problem in Eqs. (6-7) is shown in [18] to be convex.

IV. Optimality of the Step-Wise Optimization (SWO) Strategy

In this section, we will show how to specify the extra load in the subsequent time steps and show the optimality of SWO in a special uniform distribution case.

A. Optimality and the Expression of Subsequent Extra Load

At time step t_0, we define the expected loads and free space distributions of Network A and Network B as $\mathbb{E}[L_{0A}], \mathbb{E}[L_{0B}], f_A, f_B$ respectively.

We define L_{tA}^G, L_{tB}^G as the extra load being generated at time step t for networks A and B respectively; the total load being generated at time step t is $L_t^G = L_{tA}^G + L_{tB}^G$. We then define L_{tA}^R, L_{tB}^R as the extra load being redistributed at time step t for networks A and B respectively; the total load being redistributed at time step t is $L_t^R = L_{tA}^R + L_{tB}^R$. Since the extra load generated at the current time step will be redistributed at the next time step, we have the relation that $L_{t+1}^R = L_t^G$.

At an arbitrary time step t_0, for simplicity, we define the current time step to be 1 and the previous time step to be 0. For the corresponding $L_0^G = L_1^R$, setting the load to be redistributed in network A as L_{1A}^R, we have L_1^G:

$$L_1^G = L_{1A}^G + L_{1B}^G \tag{9}$$

where:

$$\begin{cases} L_{1A}^G = \mathbb{E}\left[S_A | L_A \le S_A < L_A + \frac{L_{1A}^R}{N_{1A}}\right] \cdot N_{A0}\left(F_A\left(L_A + \frac{L_{1A}^R}{N_{1A}}\right) - F_A(L_A)\right) \\ L_{1B}^G = \mathbb{E}\left[S_B | L_B \le S_B < L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}}\right] \cdot N_{B0}\left(F_B\left(L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}}\right) - F_B(L_B)\right) \end{cases}$$

(N_{1A}, N_{1B}) represent the current numbers of nodes in Network A and Network B respectively.

Following this, we denote the fraction of total extra load to be redistributed in the next time step $t = 2$ in Network A as L_{2A}^R. The total extra load being generated in $t = 2$ can be represented as:

$$L_2^G = L_{2A}^G + L_{2B}^G \tag{10}$$

where:

$$\begin{cases} L_{2A}^G = \mathbb{E}\left[S_A | L_A + \frac{L_{1A}^R}{N_{1A}} \le S_A < L_A + \frac{L_{1A}^R}{N_{1A}} + \frac{L_{2A}^R}{N_{2A}}\right] \\ \qquad \cdot N_{0A}\left(F_A\left(L_A + \frac{L_{1A}^R}{N_{1A}} + \frac{L_{2A}^R}{N_{2A}}\right) - F_A\left(L_A + \frac{L_{1A}^R}{N_{1A}}\right)\right) \\ L_{2B}^G = \mathbb{E}\left[S_B | L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}} \le S_B < L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}} + \frac{L_1 - L_{2A}^R}{N_{2B}}\right] \\ \qquad \cdot N_{0B}\left(F_B\left(L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}} + \frac{L_1 - L_{2A}^R}{N_{2B}}\right) - F_B\left(L_B + \frac{L_0^G - L_{1A}^R}{N_{1B}}\right)\right) \end{cases} \tag{11}$$

Again, (N_{2A}, N_{2B}) represent the current number of nodes in Network A and Network B respectively.

To prove the optimality of SWO, we need to show that $L_1^{G*} + L_2^{G*} = L^*$, where:

$$\begin{cases} L_1^{G*} = min_{L_{1A}^R} L_1^G \\ L_2^{G*} = min_{L_{2A}^R} L_2^G \\ L^* = min_{L_{1A}^R, L_{2A}^R}(L_1^G + L_2^G) \end{cases} \tag{12}$$

In other words, for any two consecutive time steps, the possible minimum total extra load is equal to the sum of myopically minimizing the extra load in each time step. This result will imply that for any two consecutive time steps, we cannot do better than myopically minimizing the extra load being redistributed in each time step. By induction, starting from the last two time steps, we can then guarantee that no other strategy can have less total extra load than step-wise minimization of the extra load.

B. Optimality of SWO under Uniform Distribution

We suppose that the free space of both networks follows uniform distributions with the same parameters, so that:

$$\begin{cases} S_A \sim U(S_0, S_0 + d) \\ S_B \sim U(S_0, S_0 + d) \end{cases} \tag{13}$$

The objective function of the SWO problem in Definition 3.2 can then be rewritten, and [18] derives closed-form solutions for the resulting optimal coupling coefficients.

Under uniform free space distributions, the SWO strategy can reach the global optimum, i.e., it minimizes the accumulated total extra load (and thus maximizes robustness to cascading failures). In this section, we are going to prove this statement.

Theorem 4.1: (*SWO reaches the global optimum under Uniformly distributed free space*) The solution that myopically minimizes the total extra load in each time step is equivalent to the globally optimal policy that minimizes the accumulated total extra load across all time steps.

Proof sketch: Following Section IV-A, we compare the sum of the total extra load by SWO for two consecutive time steps ($t = 1$ and $t = 2$) and the minimal total extra load that can be reached for these two consecutive time steps.

We define the remaining width of the uniform distribution of networks A and B as D_A, D_B, respectively, where $D_A, D_B \le d$ in the original uniform distribution in Equation (13). These widths change through out the cascading failure process and are determined by the average extra load a node receives in each round. Since we assume equal redistribution, when the cumulative average extra load is less than S_0 in Equation (13), the width remains unchanged, that is, $D_A = D_B = d$. Once the cumulative average extra load is more than S_0, the width starts to decrease and nodes fail. The remaining number of nodes is proportional to the current width. The distribution of free space remains uniform since the extra load is equally redistributed to all nodes.

Since by adjusting coupling coefficients α and β, we can either send all the loads to network A ($\alpha = 1, \beta = 0$) or send all the loads to network B ($\alpha = 0, \beta = 1$), any load redistribution in between these two cases can be achieved (i.e., partial redistribution to A and partial redistribution to B.). It is therefore sufficient to optimize L_{1A}^G and L_{1B}^G instead of optimizing α, β.

Following Equation (9), we have:

$$\begin{aligned} L_1^G &= L_{1A}^G + L_{1B}^G \\ &= \frac{L_0^G \mathbb{E}[L_{0A}]}{D_A} + L_{1B}^R\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_A}\right) \end{aligned} \tag{14}$$

The target variable L_{1B}^R means the load to be redistributed in Network B. The load to be redistributed in Network A is simply $L_{1B}^R = L_0^G - L_{1B}^R$, which is determined by L_{1B}^R. Since the total extra load generated L_1^G is a simple linear equation of the target variable L_{1B}^R, the optimal point of the coupling coefficients are located at the boundary, i.e., either $L_{1B}^R = 0$ or $L_{1B}^R = L_0$, which means either $\alpha(1) = 1, \beta(1) = 0$ or $\alpha(1) = 0, \beta(1) = 1$. Without loss of generality, we assume that $\frac{\mathbb{E}[L_{0B}]}{D_B} \ge \frac{\mathbb{E}[L_{0A}]}{D_A}$, where considering SWO as in Definition 3.2, the L_{1B}^R that minimizes L_1^G will be 0. Considering SWO in the following time step $t = 2$, we separate our analysis into two cases.

Case 1 ($L_{2B}^R = L_1^G$ minimizes L_2^G):

L_2^G, the total extra load being generated in $t = 2$ is then:

$$L_2^G = \frac{L_1^G \mathbb{E}[L_{1B}]}{D_B} \tag{15}$$

The total extra load when using SWO to myopically minimize the extra load in each time step can be written as:

$$\begin{aligned} L_{(SWO)}^G &= L_{1(SWO)}^G + L_{2(SWO)}^G \\ &= \frac{L_0^G \mathbb{E}[L_{0A}]}{D_A} + \frac{L_1^G \mathbb{E}[L_{1B}]}{D_B} \end{aligned} \tag{16}$$

Suppose that at $t = 1$ we take arbitrary $L^R_{1B(arb)}$ such that $0 < L^R_{1B(arb)} \leq L^G_0$. Then the extra load of the first step $L^G_{1(arb)}$ will be:

$$L^G_{1(arb)} = \frac{L^G_0 \mathbb{E}[L_{0A}]}{D_A} + L^R_{1B}\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_A}\right) > \frac{L^G_0 \mathbb{E}[L_{0A}]}{D_A} = L^G_{1(SWO)} \tag{17}$$

When taking $L^R_{1B(arb)}$, the optimal $L^R_{2B(arb)}$ that minimizes the load generated in $t = 2$ will be:

$$L^R_{2B(arb)} = L^G_{1(arb)} \tag{18}$$

Then the total extra load of $t = 2$ will be:

$$L^G_{2(arb)} = \frac{L^G_{1(arb)} \mathbb{E}[L_{1B(arb)}]}{D_{1B(arb)}} > L^G_{2(SWO)} \tag{19}$$

Thus in this case, altering the choice of the first step by choosing $0 < L^R_{1B(arb)} \leq L^G_0$ will cause the sum of the total extra load to exceed the one from SWO. In other words, no other method can do better than SWO in this case.

Case 2 ($L^R_{2B} = 0$ minimizes L^G_2):

In this case, all the extra load is still distributed to network A. Then we have the following relation:

$$\frac{\mathbb{E}[L_{0A}]N_A + L^G_0}{D_A N_A - L^G_0} < \frac{\mathbb{E}[L_{0B}]}{D_B}, \tag{20}$$

which can be written as:

$$D_A N_A \mathbb{E}[L_{0B}] > D_B N_A \mathbb{E}[L_{0A}] + L^G_0 D_B + L^G_0 \mathbb{E}[L_{0B}] \tag{21}$$

Considering the sum of the total extra load of the two steps by SWO, we have:

$$L^G_{(SWO)} = L^G_{1(SWO)} \frac{D_A + \mathbb{E}[L_{0A}]}{D_A - \frac{L^G_0}{N_A}} \tag{22}$$

Consider another sum of the total extra load of the two steps if we set an arbitrary value $0 < L^R_{1B(arb)} \leq L^G_0$, we have:

$$L^G_{(arb)} = \left(L^G_{1(arb)} + L^R_{1B(arb)}\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_B}\right)\right) \cdot \left(\frac{D_A + \mathbb{E}[L_{0A}]}{D_A - \frac{L^G_0}{N_A}} + \frac{L^R_{1B(arb)}}{N_A}\right) \tag{23}$$

The difference between the total extra loads of these two different strategies will be:

$$\begin{aligned} & L^G_{(SWO)} - L^G_{(arb)} \\ &= L^G_{1(SWO)} \frac{D_A + \mathbb{E}[L_{0A}]}{D_A - \frac{L^G_0}{N_A}} \\ &\quad - \left(L^G_{1(SWO)} + L^R_{1B(arb)}\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_B}\right)\right) \cdot \left(\frac{D_A + \mathbb{E}[L_{0A}]}{D_A - \frac{L^G_0}{N_A}} + \frac{L^R_{1B(arb)}}{N_A}\right) \\ &= \frac{L^R_{1B(arb)}(D_A + \mathbb{E}[L_{0A}])}{\left(D_A - \frac{L^G_0}{N_A}\right)\left(D_A - \frac{L^G_0}{N_A} + \frac{L^R_{1B(arb)}}{N_A}\right)} \left(\frac{L^G_{1(SWO)}}{N_A} - \left(D_A - \frac{L^G_0}{N_A}\right)\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_A}\right)\right) \end{aligned} \tag{24}$$

We only need to verify the sign of the last term which is:

$$\left(\frac{L^G_{1(SWO)}}{N_A} - \left(D_A - \frac{L^G_0}{N_A}\right)\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_A}\right)\right) \tag{25}$$

The above equation can be written as:

$$\left(\frac{L^G_0 \mathbb{E}[L_{0A}]}{D_A N_A} - \left(D_A - \frac{L^G_0}{N_A}\right)\left(\frac{\mathbb{E}[L_{0B}]}{D_B} - \frac{\mathbb{E}[L_{0A}]}{D_A}\right)\right) \tag{26}$$

This can be further simplified as:

$$\frac{L^G_0 \mathbb{E}[L_{0B}]}{D_A N_A D_B} + \frac{\mathbb{E}[L_{0A}]D_A}{D_A} - \frac{D_A \mathbb{E}[L_{0B}]}{D_B} \tag{27}$$

Multiplying both sides by $D_A N_A D_B$, we have:

$$\mathbb{E}[L_{0A}]N_A D_B - \mathbb{E}[L_{0B}]D_A N_A + L^G_0 \mathbb{E}[L_{0B}] \tag{28}$$

Substituting Equation (20) into the above equation we have $-L^G_0 D_B < 0$. Thus for this case the total extra load of the two steps with arbitrary $0 < L^R_{1B(arb)} \leq L^G_0$ will yield more total extra load compared to that of using the SWO strategy.

Combining the above two cases, as stated in Section IV-A, we can conclude that the SWO strategy under uniform free space distribution achieves the minimum extra load.

V. Numerical Results

To evaluate the SWO strategy, we first show its results under fully-connected networks, verifying that SWO can outperform other strategies under uniform free space distributions. Then we show that in different real-world transportation systems, the SWO strategy is still an effective algorithm to prevent cascading failures.

A. Simulation under Fully-Connected Networks

In this subsection, all experiments are done with 1 million nodes in both networks, and each data point is an average over 100 experiments. The surviving curve indicates the relationship between the initial attack size and fraction of nodes that ultimately survive. The transition from a system with some surviving nodes to 0 indicates the critical attack size. All initial attacks occurred in network A.

First, we compare SWO to the optimal Fixed Coupling Coefficients (FCC) strategy and the Size-Based Dynamics (SBD) strategy. SBD always redistributes the same amount of load to all nodes in the system. The specific settings are:

- **Uniform Distribution (identical):** $S_A, S_B \sim U(20, 180)$, $\mathbb{E}[L_A] = \mathbb{E}[L_B] = 75$.
- **Uniform Distribution (imbalanced number of nodes):** $S_A, S_B \sim U(20, 180)$, $\mathbb{E}[L_A] = \mathbb{E}[L_B] = 75$, Network B has half the number of nodes as Network A.
- **Uniform Distribution (non-identical):** $S_A \sim U(20, 180)$, $S_B \sim U(40, 280)$, $\mathbb{E}[L_A] = \mathbb{E}[L_B] = 75$.

The experiment results are shown in Table III. We can see that the SWO strategy has the best robustness compared to all other baseline strategies. To be specific, under the identical and imbalanced numbers of nodes settings, all three strategies can reach the same performance. However, under the non-identical setting, the SBD strategy performs worse than FCC or SWO. The FCC strategy can also reach the optimal performance under all these settings. However, searching for the optimal coupling coefficients of the FCC strategy requires brute force search, which scales poorly. Finding the optimal coefficients to

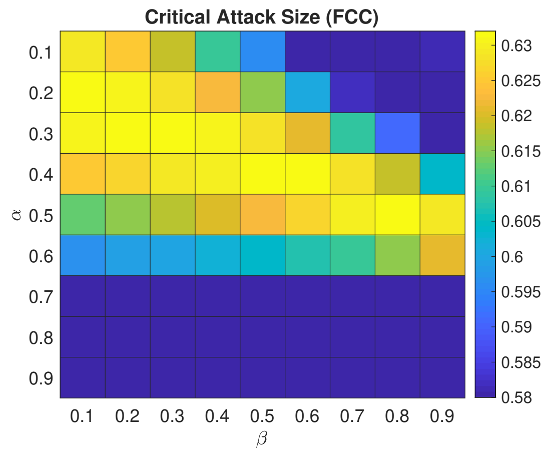


Fig. 2: Heatmap for the critical attack size of different FCC settings. The color threshold is set to be 0.58, if the critical attack size is less than 0.58 it will shows blue in the block.The optimal critical attack size is 0.632.

a precision of degree d requires d^2 more time in a two-network system. Once the number of interdependent networks grows, the number of coupling coefficients will grow in exponential order, as well as the running time. Figure 2 shows a heatmap of the robustness under different FCC coefficients.

Strategies	Identical	Imbalanced Number of Nodes	Non-Identical
SWO	0.524	0.431	0.631
SBD	0.524	0.431	0.607
Best FCC	0.524	0.431	0.631

TABLE III: Critical Attack Size under Different Settings

We proved the optimality of SWO under a uniform free space distribution in Section III by showing that optimizing the load for 2 steps is equal to the myopically optimizing the load 1 step at a time. Here we extend this result to Exponential free space distribution by running a simulation comparing 2-step optimization (2-step SWO) and 1 step at a time. In the symmetric settings, where both networks have the same number of nodes, average load, and free space distribution, both SWO and 2-step SWO take $\alpha = \beta = 0.5$. (i.e. when initial extra load $L_0 = 100000$, number of nodes $N_A = N_B = 10000, S_A, S_B \sim$ Exponential($\frac{1}{120}$), $\mathbb{E}[L_A] = \mathbb{E}[L_B] = 20$, both solutions yield $\alpha(1) = \beta(1) = \alpha(2) = \beta(2) = 0.5$, and the total extra load is $L_1 + L_2 = 18350.6 + 3880.6 = 22231.2$). In other non-symmetric settings, both SWO and 2-step SWO yield coupling coefficients of 0 and 1 (i.e., when initial extra load $L_0 = 100000$, number of nodes $N_A = 20000, N_B = 10000, S_A, S_B \sim$ Exponential($\frac{1}{120}$), $\mathbb{E}[L_A] = 10, \mathbb{E}[L_B] = 20$, both solution yields $\alpha(1) = \alpha(2) = 1, \beta(1) = \beta(2) = 0$, and the total extra load is $L_1 + L_2 = 10188 + 1293.2 = 11481.2$). We conjecture that as with uniformly distributed free space, the SWO strategy will also minimize the accumulated extra load in a system with Exponentially distributed free space. Generalizing this result to other distributions, however, may be difficult as some distributions do not have a direct relation of remaining free space and the number of surviving nodes. In such cases, minimizing the accumulated extra load maximizes the total remaining free space of the system, which can be viewed as an alternate measure of robustness as it measures the ability of the system to accommodate future failures.

B. Simulation of Real World Networks

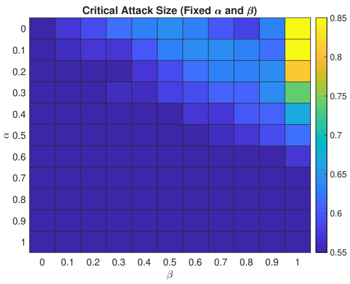
We use the subway network topologies of some of the largest cities in the world [22] to drive our simulation. First, we take the Tokyo subway as an example. Under the FCC strategy, the heatmap of the critical attack size of different coupling coefficient settings is shown in Figure 3a. We can see that the highest critical attack size is 0.85. The critical attack size of our SWO strategy is 0.75 instead.

These results imply that though our SWO has better critical attack size than most FCC settings, it performs worse than the optimal FCC ones. However, since load is redistributed according to the (not fully connected) network topology, the load is not being equally redistributed to all nodes in the network. Instead, some nodes fail more easily than others (e.g., when a node with only one neighbor fails, its entire load will be redistributed to its neighbor, which is then also likely to fail.). Keeping extra load in one network might save the other one and could guarantee that some part of the system survives, but most of the system may still fail. We plot the survival curve of SWO and the best FCC strategies, as shown in Figure 3b. We can see that before the whole system fails, our SWO strategy allows more nodes to survive than the FCC strategy. The SWO strategy exchanges the extra load between 2 networks at each time step and makes sure not to sacrifice any part of the system. From this viewpoint, our SWO strategy balances the flow in two networks and has a higher surviving portion compared to the FCC strategy before the system fails.

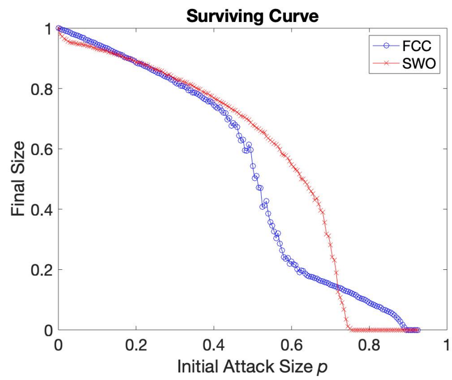
Last in Figure 3c, we compare the results of the subway systems in different cities. The SWO strategy works for all of them and yields similar results for all cities. This may be because subway systems, no matter where they located, tend to follow the same design concepts. Tokyo has a slightly higher robustness compared to other cities, which may be due to the Tokyo subway comprising only the core, denser part of the overall east Japan railway system.

VI. Conclusion

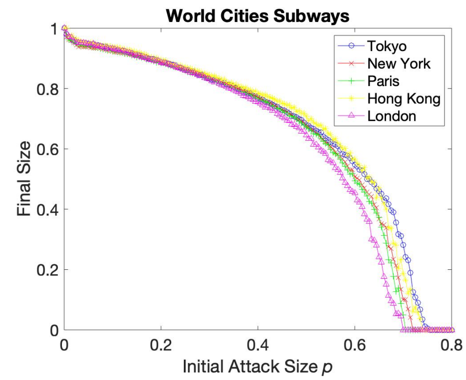
In this work, we studied the robustness of interdependent networks under different coupling strategies based on a flow-redistribution model. Based on the previous work [18], we prove the optimality of the SWO strategy to dynamically adjust the coupling coefficients between networks according to the current system state. Furthermore, we verify that the SWO strategy can provide good performance against cascading failures in topologies based on real world urban transportation systems. Evidence in Section V shows that networks with an Exponential free space distribution may have the same SWO and 2-step SWO solutions. We therefore conjecture that the



(a) Heatmap for the critical attack size of different FCC. The color threshold is set to be 0.55. If the critical attack size is less than 0.55, it will shows blue in the block. The optimal critical attack size is 0.85.



(b) Surviving curves SWO and the FCC that yields the highest critical attack size.



(c) Comparison of different cities. All curves have similar trend & critical attack size.

Fig. 3: Results for real-world network topologies.

SWO strategy may be optimal when the free space distributions have monotonically decreasing probability density functions (PDFs), since the trend of the PDF (decreasing or increasing) does not change when we change our reference point. A direct future work will be proving this conjecture. Moreover, as seen in [1], [3], cascading failures in power grids and IoT systems or other networked systems can be modeled as kind of load carrying systems. Applying and adapting our dynamic coupling strategies to these different real-world networked systems can be another useful future direction.

Acknowledgement

This work was supported in part by the National University Transportation Center for Improving the Mobility of People and Goods (Mobility21), the Scott Institute for Energy Innovation, and the Air Force Office of Scientific Research (AFOSR) through Grant #FA9550-22-1-0233. A portion of this research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-20-2-0175. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

[1] D. Liu, X. Zhang, and C. K. Tse, "A tutorial on modeling and analysis of cascading failure in future power grids," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 1, pp. 49–55, 2021.

[2] Y. Wang and C. Tian, "Measure vulnerability of metro network under cascading failure," *IEEE Access*, vol. 9, pp. 683–692, 2021.

[3] L. Xing, "Cascading failures in internet of things: Review and perspectives on reliability and resilience," *IEEE Internet of Things Journal*, vol. 8, no. 1, pp. 44–64, 2021.

[4] Y. Zhang and O. Yağan, "Optimizing the robustness of electrical power systems against cascading failures," *Scientific reports*, vol. 6, no. 1, pp. 1–15, 2016.

[5] G. Sun, C.-C. Chen, and S. Bin, "Study of cascading failure in multisubnet composite complex networks," *Symmetry*, vol. 13, no. 3, p. 523, 2021.

[6] O. Yağan, "Robustness of power systems under a democratic-fiber-bundle-like model," *Physical Review E*, vol. 91, no. 6, p. 062811, 2015.

[7] A. Scala, P. G. D. S. Lucentini, G. Caldarelli, and G. D'Agostino, "Cascades in interdependent flow networks," *Physica D: Nonlinear Phenomena*, vol. 323, pp. 35–39, 2016.

[8] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature physics*, vol. 8, no. 1, pp. 40–48, 2012.

[9] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 048701, 2010.

[10] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.

[11] D. Duan, C. Lv, S. Si, Z. Wang, D. Li, J. Gao, S. Havlin, H. E. Stanley, and S. Boccaletti, "Universal behavior of cascading failures in interdependent networks," *Proceedings of the National Academy of Sciences*, vol. 116, no. 45, pp. 22452–22457, 2019.

[12] 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.

[13] M. Di Muro, S. Buldyrev, H. Stanley, and L. Braunstein, "Cascading failures in interdependent networks with finite functional components," *Physical Review E*, vol. 94, no. 4, p. 042304, 2016.

[14] Y. Zhang and O. Yağan, "Robustness of interdependent cyber-physical systems against cascading failures," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 711–726, 2019.

[15] A. Vespignani, "The fragility of interdependency," *Nature*, vol. 464, no. 7291, pp. 984–985, 2010.

[16] Y. Zhang, A. Arenas, and O. Yağan, "Cascading failures in interdependent systems under a flow redistribution model," *Physical Review E*, vol. 97, no. 2, p. 022307, 2018.

[17] I.-C. Lin, C. Joe-Wong, and O. Yagan, "Dynamic coupling strategy for interdependent network systems against cascading failures," 2022. https://arxiv.org/abs/2203.01295.

[18] I.-C. Lin, O. Yağan, and C. Joe-Wong, "Dynamic coupling strategy for interdependent network systems against cascading failures," *IEEE Transactions on Network Science and Engineering*, pp. 1–18, 2023.

[19] J. V. Andersen, D. Sornette, and K.-t. Leung, "Tricritical behavior in rupture induced by disorder," *Physical Review Letters*, vol. 78, no. 11, p. 2140, 1997.

[20] S. Pahwa, C. Scoglio, and A. Scala, "Abruptness of cascade failures in power grids," *Scientific reports*, vol. 4, no. 1, pp. 1–9, 2014.

[21] Y. Moreno, A. Correig, J. Gómez, and A. Pacheco, "A model for complex aftershock sequences," *Journal of Geophysical Research: Solid Earth*, vol. 106, no. B4, pp. 6609–6619, 2001.

[22] C. Roth, S. M. Kang, M. Batty, and M. Barthelemy, "A long-time limit for world subway networks," *Journal of The Royal Society Interface*, vol. 9, no. 75, pp. 2540–2550, 2012.