Optimal Allocation of Interconnecting Links in Cyber-Physical Systems: Interdependence, Cascading Failures and Robustness

Osman Yağan, *Student Member*, *IEEE*, Dajun Qian, *Student Member*, *IEEE*, Junshan Zhang, *Fellow*, *IEEE*, and Douglas Cochran, *Senior Member*, *IEEE*

Abstract—We consider a cyber-physical system consisting of two interacting networks, i.e., a cyber network overlaying a physical network. It is envisioned that these systems are more vulnerable to attacks since node failures in one network may result in (due to the interdependence) failures in the other network, causing a cascade of failures that would potentially lead to the collapse of the entire infrastructure. The robustness of interdependent systems against this sort of catastrophic failure hinges heavily on the allocation of the (interconnecting) links that connect nodes in one network to nodes in the other network. In this paper, we characterize the *optimum* inter-link allocation strategy against random attacks in the case where the topology of each individual network is unknown. In particular, we analyze the "regular" allocation strategy that allots exactly the same number of bidirectional internetwork links to all nodes in the system. We show, both analytically and experimentally, that this strategy yields better performance (from a network resilience perspective) compared to all possible strategies, including strategies using random allocation, unidirectional interlinks, etc.

Index Terms—Interdependent networks, cascading failures, robustness, resource allocation, random graph theory.

1 Introduction

Today's worldwide network infrastructure consists a web of interacting cyber networks (e.g., the Internet) and physical systems (e.g., the power grid). There is a consensus that integrated cyber-physical systems will emerge as the underpinning technology for major industries in the 21st century [9]. The smart grid is one archetypal example of such systems where the power grid network and the communication network for its operational control are coupled together and depend on each other; i.e., they are *interdependent*. While interdependency allows building systems that are larger, smarter, and more complex, it has been observed [17] that interdependent systems tend to be more fragile against failures, natural hazards and attacks. For example, in the event of an attack to an interdependent system, the failures in one of the networks can cause failures of the dependent nodes in the other network and vice versa. This process may continue in a recursive manner and hence lead to a cascade of failures causing a catastrophic impact on the overall cyber-physical system. In fact, the cascading effect of even a partial Internet blackout could disrupt major national infrastructure networks involving Internet services, power grids, and financial markets [3]. Real-world examples include the 2003 blackout in the northeastern United States and southeastern Canada [17] and the electrical blackout that affected much of Italy on 28 September 2003 [3].

1.1 Background and Related Work

Despite recent studies of cascading failures in complex networks, the dynamics of such failures and the impact across multiple networks are not well understood. There is thus a need to develop a new network science for modeling and quantifying cascading failures, and to develop network management algorithms that improve network robustness and ensure overall network reliability against cascading failures. Most existing studies on failures in complex networks consider only the single network case. A notable exception is the very recent work of Buldyrev et al. [3] in which a "one-to-one correspondence" model for studying the ramifications of interdependence between two networks is set forth. This model considers two networks of the same size, say networks A and B, where each node in network A depends on one and only one node in network B and vice versa. In other words, each node in network A has one bidirectional *interedge* connecting it to a *unique* node in network B. Furthermore, it is assumed that a node in either network can function *only if* it has support from the other network; i.e., it is connected (via an interedge) to at least one functioning node from the other network.

The robustness of the one-to-one correspondence model was studied in [3] using a similar approach to that of the works considering single networks [5], [7]. Specifically, it is assumed that a random attack is launched upon network A, causing the failure of a fraction $1 - p$ of the nodes; this was

- *O. Yağan is with the Department of Electrical and Computer Engineering and CyLab, Carnegie Mellon University, 4742 Centre Avenue, Apt. 705, Pittsburgh, PA 15213. E-mail: oyagan@andrew.cmu.edu.*
- *D. Qian, J. Zhang, and D. Cochran are with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706. E-mail: {dqian, junshan.zhang, cochran}@asu.edu.*

Manuscript received 14 Sept. 2011; revised 14 Dec. 2011; accepted 30 Jan. 2012; published online 3 Feb. 2012.
Recommended for acceptance by S. Papavassiliou, N. Kato, Y. Liu, and C.-Z. Xu.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDSSI-2011-09-0632.
Digital Object Identifier no. 10.1109/TPDS.2012.62.

modeled by a random removal of a fraction $1-p$ of the nodes from network A. Due to the interdependency, these initial failures lead to node failures from network B, which in turn may cause further failures from network A thereby triggering an avalanche of cascading failures. To evaluate the robustness of the model, the size of the functioning parts of both networks are computed at each stage of the cascading failure until a *steady state* is reached; i.e., until the cascade of failure ends. One of the important findings of [3] was to show the existence of a critical threshold on p, denoted by p_c, above which a considerable fraction of nodes in both networks remain functional at the steady state; on the other hand, if $p < p_c$, both networks go into a complete fragmentation and the entire system collapses. Also, it is observed in [3] that interdependent network systems have a much larger p_c compared to that of the individual constituent networks; this is compatible with the observation that interdependent networks are more vulnerable to failures and attacks.

The original work of Buldyrev et al. [3] has received much attention and spurred the study of interdependent networks in many different directions; e.g., see [4], [6], [10], [14], [15], [16]. One major vein of work, including [4], [14], [16], aims to extend the findings of [3] to more realistic scenarios than the one-to-one correspondence model. More specifically, in [4] the authors consider a one-to-one correspondence model with the difference that mutually dependent nodes are now assumed to have the same number of neighbors in their own networks; i.e., their intradegrees are assumed to be the same. In [14], the authors consider the case where only a fraction of the nodes in network A depends on the nodes in network B, and vice versa. In other words, some nodes in one network are assumed to be *autonomous*, meaning that they do not depend on nodes of the other network to function properly. Nevertheless, in [14] it was still assumed that a node can have *at most* one supporting node from the other network. More recently, Shao et al. [16] pointed out the fact that, in a realistic scenario, a node in network A may depend on more than one node in network B, and vice versa. In this case, a node will function as long as at least one of its supporting nodes is still functional. To address this case, Shao et al. [16] proposed a model where the interedges are unidirectional and each node supports (and is supported by) a *random* number of nodes from the other network. In a different line of work, Schneider et al. [15] adopted a design point of view and explored ways to improve the robustness of the one-to-one correspondence model by letting some nodes be autonomous. More precisely, they assume that the topologies of networks A and B are known and propose a method, based on degree and centrality, for choosing the autonomous nodes properly in order to maximize the system robustness.

1.2 Summary of Main Results

In this study, we stand in the intersection of the two aforementioned lines of work. First, we consider a model where interedges are allocated *regularly* in the sense that all nodes have exactly the *same* number of *bidirectional* interedges, assuming that no topological information is available. This ensures a uniform support-dependency relationship where each node supports (and is supported by) the same number of nodes from the other network. We analyze this new model in terms of its robustness against random attacks via characterizing the steady state size of the functioning parts of each network as well as the critical fraction p_c. In this regard, our work generalizes the studies on the one-to-one correspondence model and the model studied by Shao et al. [16]. From a design perspective, we show analytically that the proposed method of *regular* interedge allocation improves the robustness of the system over the random allocation strategy studied in [16]. Indeed, for a given *expected* value of interdegree (the number of nodes it supports *plus* the number of nodes it depends upon) per node, we show that 1) it is better (in terms of robustness) to use bidirectional interlinks than unidirectional links, and 2) it is better (in terms of robustness) to deterministically allot each node exactly the same number of bidirectional interedges rather than allotting each node a random number of interedges.

These results imply that if the topologies of networks A and B are unknown, then the optimum interlink allocation strategy is to allot exactly the same number of bidirectional interedges to all nodes. Even if the statistical information regarding the networks is available, e.g., say it is known that network A is an Erdős-Rényi (ER) [2] network and network B is a scale-free network [1], regular interedge allocation is still the best strategy in the absence of the detailed topological information; e.g., in the case where it is not possible to estimate the nodes that are likely to be more important in preserving the connectivity of the networks, say nodes with high betweenness [8]. Intuitively, this makes sense because without knowing which nodes play a key role in preserving the connectivity of the networks, it is best to treat all nodes "identically" and give them equal priority in interedge allocation.

The theoretical results in this paper are also supported by extensive computer simulations. Numerical results are given for the case where both networks are Erdős-Rényi and the optimality of the regular allocation strategy is verified. To get a more concrete sense, assume that A and B are ER networks with N nodes and average degree 4. When interedges are allocated regularly so that each node has exactly two bidirectional interedges, the critical threshold p_c is equal to 0.43. However, for the same networks A and B, if the number of interedges follows a Poisson distribution with mean 2, the critical p_c turns out to be equal to 0.82. This is a significant difference in terms of robustness, since in the former case the system is resilient to the random failure of up to 57 percent of the nodes while in the latter case, the system is resilient to the random failure of up to only 18 percent of the nodes.

To the best of our knowledge, this paper is the first work that characterizes the robustness of interdependent networks under regular allocation of bidirectional interedges. Also, it is the first work that determines analytically and experimentally the optimum interedge allocation strategy in the absence of topological information. We believe that our findings along this line shed light on the design of interdependent systems.

1.3 Structure of the Paper

The paper is organized as follows. In Section 2, we introduce the system model and present an overview of

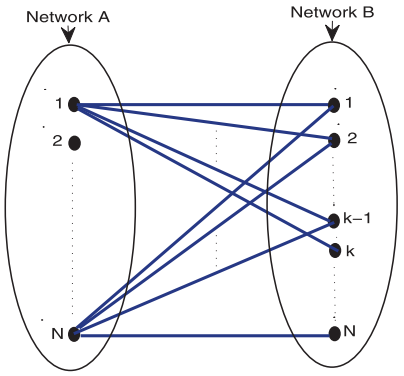


Fig. 1. A sketch of the proposed system model, namely the *regular* allocation strategy of bidirectional interedges: each node in A is connected to exactly k nodes in B, and vice versa.

cascading failures. The behavior of this model under random attacks is analyzed in Section 3, where the functional size of the networks is characterized. Section 4 is devoted to proving the optimality of regular interlink allocation, while in Section 5 we give numerical examples and simulation results. Possible future research is explained and the paper is concluded in Section 6.

2 System Model

We consider a cyber-physical system consisting of two interacting networks, say networks A and B. For simplicity, both networks are assumed to have N nodes and the vertex sets in their respective graphical representations are denoted by $\{v_1, \ldots, v_N\}$ and $\{v'_1, \ldots, v'_N\}$. We refer to the edges connecting nodes within the same network as *intraedges* and those connecting nodes from two different networks as *interedges*. Simply put, we assume that a node can function *only if* it is connected (via an interedge) to at least one functioning node in the other network [3], and we will elaborate further on this. Clearly, the interdependency between two networks is intimately related to the interedges connecting them. In this study, interedges are assumed to be bidirectional so that it is convenient to use an $N \times N$ interdependency matrix $\mathbf{C}$ to represent the bidirectional interedges between networks A and B. Specifically, for each $n, m = 1, \ldots, N$, let

$$(\mathbf{C})_{nm} = \begin{cases} 1 & \text{if } v_n \text{ and } v'_m \text{ depend on each other} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

We also assume that interedges are allocated *regularly* so that each node has exactly k interedges, where k is an integer satisfying $k \leq N$. Without loss of generality, this strategy can be implemented in the following manner: for each $n = 1, 2, \ldots, N$, let the interdependency matrix be given by

$$(\mathbf{C})_{nm} = \begin{cases} 1 & \text{if } m = n, n \oplus 1, \ldots, n \oplus (k-1) \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where we define

$$n \oplus l = \begin{cases} n + \ell & \text{if } n + l \leq N \\ n + \ell - N & \text{if } n + l > N, \end{cases}$$

for each $\ell = 0, 1, \ldots, N-1$; see also Fig. 1.

We are interested in evaluating the network robustness in the case of random node failures (or equivalently random attacks). Specifically, in the dynamics of cascading failures, we assume that a node is *functioning* at Stage i if the following conditions are satisfied [3], [16]: 1) the node has at least one interedge with a node that was functioning at Stage $i-1$; 2) the node belongs to the giant (i.e., the largest) component of the subnetwork formed by the nodes (of its own network) that satisfy condition 1. For both networks, a giant component consisting of functioning nodes will be referred to as a *functioning giant component*.

We assume that the cascade of failures is triggered by the failure of a fraction $1-p$ of the nodes in network A. We further assume that these $(1-p)N$ nodes are chosen (say by the attacker) uniformly at random among all nodes in network A. By the definitions given above, it can be seen that after the initial attack, only nodes in the functioning giant component of A can operate properly. As a result of that, in the next stage, some of the nodes in network B may end up losing all of their interconnections and turn dysfunctional. In that case, the nodes that can function properly in network B will only be those in the functioning giant component of B. But, this fragmentation of network B may now trigger further failures in network A due to nodes that lose all their B-connections. Continuing in this manner, the cascade of failures propagates alternately between A and B, eventually (i.e., in steady state) leading to either: 1) *residual functioning giant components in both networks*, or 2) *complete failure of the entire system*. For an illustrative example, see Fig. 2 where a cascading failure is demonstrated for a pair of interdependent networks with $N = 6$ nodes, $k = 2$, and $p = 2/3$.

3 Analysis of Cascading Failures under Regular Allocation of Interedges

In this section, we analyze the dynamics of cascading failures in two interacting networks. A principal objective of this study is to quantify the effectiveness of the regular allocation strategy for network robustness, by means of: 1) characterizing the size of the remaining giant components in networks A and B after the cascade has reached a steady state, and 2) finding the corresponding critical threshold p_c. To these ends, we will use the technique of generating functions [12], [13] to analyze the sizes of functioning giant components in the two networks at each stage. For convenience, the notation used in the calculations is summarized in Table 1.

3.1 Stage 1: Random Failure of Nodes in Network A

Following the failures of a fraction $1-p$ of randomly selected nodes in network A, the remaining network $\bar{A}_1$ has size pN; since we eventually let N grow large, pN can be approximated as an integer. As in [3], [12], [13], [16], we use the technique of generating functions to quantify the fraction of the functioning giant component $A_1 \subset \bar{A}_1$. Specifically, let the function $P_A(p)$ determine the fraction of the giant component in a random subgraph that occupies a fraction p of the nodes in network A (the exact calculation of $P_A(p)$ will be elaborated later). It follows that the functioning giant component has size

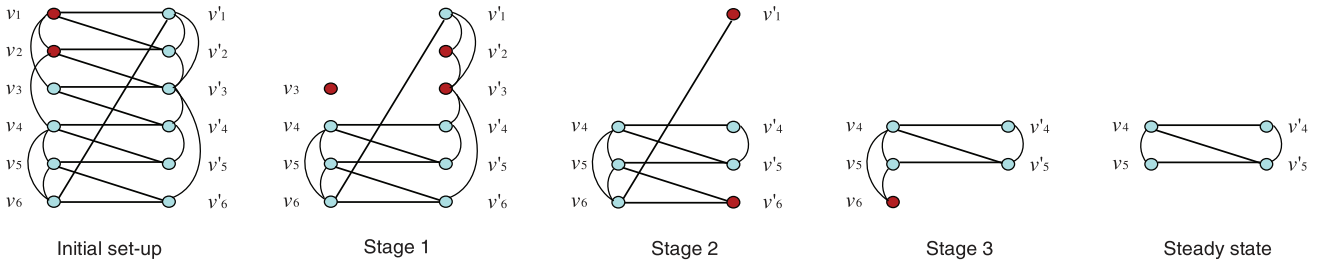


Fig. 2. An illustration of cascading failures in two interdependent networks. Network A with nodes $\{v_1, v_2, \ldots, v_6\}$ and network B with nodes $\{v'_1, v'_2, \ldots, v'_6\}$ are interdependent with each node having exactly two bidirectional interedges. Initially, a random attack causes the failure of nodes v_1 and v_2. In stage 1, v_1 and v_2 are removed from the system along with all the links (inter and intra) that are incident upon them. As a result, node v_3 becomes disconnected from the functioning giant component of network A, and thus fails. These failures then cause the nodes v'_2 and v'_3 to fail as they lose all their supports; i.e., all the interedges that are incident upon them are removed. In stage 2, we see the effect of removing v'_2 and v'_3 from network B: nodes v'_1 and v'_6 fail as they become disconnected from the functioning giant component. The failure of nodes v'_1 and v'_6 then leads to the failure of node v_6 in stage 3, since v_6 was being supported solely by v'_1 and v'_6. By removal of the node v_6, the failures stop and the system reaches steady state.

$$|A_1| = pP_A(p)N := p_{A1}N. \tag{3}$$

As shall become apparent soon, at the end of each stage it is necessary to determine not only the size of the functioning giant component, but also the specific interedge distribution over the functioning nodes; i.e., the numbers of functioning nodes having particular numbers of interedges. *Indeed, this is what makes the analysis of the regular allocation model more complicated than the models considered in [3], [14], [16].* Here, at the end of Stage 1, each node in A_1 still has k interedges from network B since network B has not changed yet.

3.2 Stage 2: Impact of Random Node Failures in Network A on Network B

As the functioning part of network A fragments from A to A_1 (in Stage 1), some of the interedges that were supporting B-nodes would be removed. Observe that the probability of removal can be approximated by $1 - |A_1|/|A| = 1 - p_{A1}$ for each interedge. With this perspective, a B-node loses $k - j$ of its interedges with probability $\binom{k}{j} p_{A1}^j (1 - p_{A1})^{k-j}$. Moreover, it stops functioning with probability $(1 - p_{A1})^k$ due to losing all k of its interedges. As a result, with $\bar{B}_2$ denoting the set of nodes in B that retain at least one interedge, we have

$$|\bar{B}_2| = (1 - (1 - p_{A1})^k)N = p'_{B2}N, \tag{4}$$

where $p'_{B2} = 1 - (1 - p_{A1})^k$. Also, the distribution of interedges over the nodes in $\bar{B}_2$ is given by

$$|\bar{B}_2|_j = \binom{k}{j} p_{A1}^j (1 - p_{A1})^{k-j} N, \quad j = 1, 2 \ldots, k, \tag{5}$$

with $|\bar{B}_2|_j$ denoting the number of nodes in $\bar{B}_2$ that have j interedges.

As in Stage 1, the size of the functioning giant component $B_2 \subset \bar{B}_2$ can be predicted by

TABLE 1
Key Notation in the Analysis of Cascading Failures

A_i, B_i	the functioning giant components in A and B at stage i
p_{Ai}, p_{Bi}	the fractions corresponding to functioning giant components at stage i, $\lvert A_i\rvert = p_{Ai}N$, $\lvert B_i\rvert = p_{Bi}N$
$\bar{A}_i$, $\bar{B}_i$	the remaining nodes in A and B retaining at least one inter-edge at stage i.

$$|B_2| = p'_{B2}P_B(p'_{B2})N = p_{B2}N, \tag{6}$$

where $P_B(\cdot)$ is defined analogously to the definition of $P_A(\cdot)$ given in Section 3.1. Obviously, each node in $\bar{B}_2$ can survive as a functioning node in B_2 with probability $P_B(p'_{B2})$. Thus, for each $j = 1, 2, \ldots, k$, the number of nodes in B_2 that have j interedges is given (in view of (5)) by

$$|B_2|_j = P_B(p'_{B2})\binom{k}{j} p_{A1}^j (1 - p_{A1})^{k-j} N. \tag{7}$$

3.3 Stage 3: Further A-Nodes Failures Due to B-Node Failures

Due to the fragmentation of the functional part of network B from $\bar{B}_2$ to B_2 (not B to B_2), some of the nodes in A_1 may now lose all their interedges and stop functioning. To compute the probability of this event, first observe that each interedge from $\bar{B}_2$ to A_1 will be removed with an approximate probability of $1 - |B_2|/|\bar{B}_2| = 1 - P_B(p'_{B2})$. Hence, the probability that a node in A_1 will lose all of its interedges is given by $(1 - P_B(p'_{B2}))^k$. It also follows that the size of the network $\bar{A}_3 \subset A_1$ comprised of the nodes that did not lose all their interconnections is given via

$$|\bar{A}_3| = p_{A1}(1 - (1 - P_B(p'_{B2}))^k)N. \tag{8}$$

In other words, in passing from A_1 to $\bar{A}_3$, a fraction $1 - |\bar{A}_3|/|A_1| = (1 - P_B(p'_{B2}))^k$ of the nodes have failed. As previously, the next step is to compute the size of the functioning giant component $A_3 \subset \bar{A}_3$. However, this a challenging task as noted in [3]. Instead, we view the joint effect of the node failures in Stages 1 and 3 as equivalent (in terms of the size of the resulting functional giant component; i.e., $|A_3|$) to the effect of an initial random attack that targets an appropriate fraction (to be determined later) of the nodes. Intuitively, the node failures in A_1 at Stage 3 (i.e., the removal of a fraction $(1 - P_B(p'_{B2}))^k$ of nodes from A_1) have the same effect as taking out the same portion from $\bar{A}_1$ [3]. In other words, it is equivalent to the removal of a fraction $p(1 - P_B(p'_{B2}))^k$ of the nodes from A. Recalling also that a fraction $1 - p$ of the nodes in network A failed as a result of the initial attack at Stage 1, we find that the fragmentation of A to $\bar{A}_3$ can as well be modeled (with respect to the size of A_3) by an initial attack targeting a fraction

$$1-p+p(1-P_B(p'_{B2}))^k = 1-p(1-(1-P_B(p'_{B2}))^k)$$

of the nodes. It is now a standard step to conclude that, with $p'_{A3} = p(1-(1-P_B(p'_{B2}))^k)$, the size of the functioning giant component A_3 is given by

$$|A_3| = p'_{A3}P_A(p'_{A3})N = p_{A3}N. \tag{9}$$

3.4 Stage 4: Further Fragmentation of Network B

Due to the network fragmentation from $\bar{A}_3$ to A_3 in Stage 3, each interedge supporting a B_2-node will be disconnected with probability that equals the proportion nodes in $\bar{A}_3$ that did not survive to A_3; i.e., $1-|A_3|/|\bar{A}_3| = 1-P_A(p'_{A3})/P_A(p)$ by (8) and (9). Consequently, a node in B_2 with j interedges will stop functioning with probability $(1-P_A(p'_{A3})/P_A(p))^j$. Recalling also the interedge distribution (7), the fraction L of node failures in B_2 is given by

$$\begin{aligned}
L &= \frac{1}{N}\sum_{j=1}^{k}|B_2|_j\left(1-\frac{P_A(p'_{A3})}{P_A(p)}\right)^j \\
&= P_B(p'_{B2})\sum_{j=1}^{k}\binom{k}{j}p_{A1}^j(1-p_{A1})^{k-j}\left(1-\frac{P_A(p'_{A3})}{P_A(p)}\right)^j \\
&= P_B(p'_{B2})\left(\left(1-p_{A1}\frac{P_A(p'_{A3})}{P_A(p)}\right)^k-(1-p_{A1})^k\right) \\
&= P_B(p'_{B2})((1-pP_A(p'_{A3}))^k-(1-p_{A1})^k).
\end{aligned}$$

Since $|\bar{B}_4| = |B_2| - LN$, it follows that:

$$|\bar{B}_4| = P_B(p'_{B2})(1-(1-pP_A(p'_{A3}))^k)N. \tag{10}$$

In order to compute the size of the functioning giant component $B_4 \subset \bar{B}_4$, we proceed as in Stage 3. Specifically, we view the joint effect of node removals in Stages 2 and 4 as equivalent to that of an initial random attack which targets an appropriate fraction of the nodes. To determine this fraction, first observe that the failures in Stage 3 have triggered further node failures in B_2 resulting a fraction

$$1-|\bar{B}_4|/|B_2| = 1-(1-(1-pP_A(p'_{A3}))^k)/p'_{B2} \tag{11}$$

of the nodes' failure. Next, note that the effect of these failures on $|B_4|$ is equivalent to that of taking out the same fraction of nodes from $\bar{B}_2$ [3]. Moreover, it has the same effect as taking out a fraction $p'_{B2}\{1-(1-(1-pP_A(p'_{A3}))^k)/p'_{B2}\}$ of the nodes in B. Now, recalling that a fraction $1-p'_{B2}$ of nodes in B have failed in Stage 2, we conclude that the joint effect of cascading failures in Stages 2 and 4 (on $|B_4|$) is identical to that of an initial random attack which targets a fraction

$$\begin{aligned}
&1-p'_{B2}+p'_{B2}\left(1-\frac{1-(1-pP_A(p'_{A3}))^k}{p'_{B2}}\right) \\
&\quad = (1-pP_A(p'_{A3}))^k
\end{aligned}$$

of nodes. As previously, with $p'_{B4} = 1-(1-pP_A(p'_{A3}))^k$ we conclude that the size of the functioning giant component B_4 is given by $|B_4| = p'_{B4}P_B(p'_{B4})N = p_{B4}N$.

3.5 Cascading Dynamics of Node Failures

As mentioned earlier, the main goal of this section is to characterize the size of the functional giant components in steady state. Indeed, along the lines outlined above, one can obtain the sizes of all functioning giant components $A_1 \supset A_3 \supset \cdots \supset A_{2m+1}$ and $B_2 \supset B_4 \supset \cdots \supset B_{2m}$ for any integer m. However, it is easy to observe the pattern in the expressions obtained so far and conclude that with $p'_{A1} = p$ the size of all giant components are given by the recursive relations

$$\begin{aligned}
p_{Ai} &= p'_{Ai}P_A(p'_{Ai}), \\
p'_{Ai} &= p(1-(1-P_B(p'_{Bi-1}))^k), \quad i = 3,5,7\ldots
\end{aligned} \tag{12}$$

and

$$\begin{aligned}
p_{Bi} &= p'_{Bi}P_B(p'_{Bi}), \\
p'_{Bi} &= 1-(1-pP_A(p'_{Ai-1}))^k, \quad i = 2,4,6,\ldots.
\end{aligned} \tag{13}$$

This recursive process stops at an "equilibrium point" where we have $p'_{B2m-2} = p'_{B2m}$ and $p'_{A2m-1} = p'_{A2m+1}$ so that neither network A nor network B fragments further. Setting $x = p'_{A2m+1}$ and $y = p'_{B2m}$, this yields the transcendental equations

$$x = p(1-(1-P_B(y))^k) \quad y = 1-(1-pP_A(x))^k. \tag{14}$$

The analysis carried out up to this point is valid for *all* networks, irrespective of their intrastructures. In principle, for specific intrastructures of networks A and B (which determine the functions P_A and P_B, respectively), the system (14) of equations can be solved for given p and k. The steady-state fractions of nodes in the giant components can then be computed by using the relations $\lim_{i\to\infty}p_{Ai} := P_{A_\infty} = xP_A(x)$ and $\lim_{i\to\infty}p_{Bi} := P_{B_\infty} = yP_B(y)$. Indeed, in Section 5, we consider a special case where both networks A and B are Erdős-Rényi graphs [2] and give solutions of the system (14) for several values of p and k.

4 Optimality of Regular Allocation Strategy

In this section, we show analytically that the regular allocation strategy always yields stronger robustness than other strategies and thus it is optimal in the absence of intratopology information. In the following, we refer to the system that uses the regular allocation strategy as System 1. Specifically, we consider two networks A and B where each node is uniformly supported by k bidirectional interedges. For convenience, we denote the fractions in the recursive relations (12)-(13) as $p'_{Ai}(p;k)$ and $p'_{Bi}(p;k)$, where $1-p$ is the initial fraction of failed nodes in network A. Also, we let $P_{A^1_\infty}(p;k)$ and $P_{B^1_\infty}(p;k)$ be the steady-state fractions of functional giant components of the two networks, respectively. Finally, we use $p_{c_1}(k)$ to denote the critical threshold associated with System 1.

In what follows, we first investigate the dynamics of cascading failures in the auxiliary System 2, where bidirectional interedges are distributed *randomly* among nodes. The analysis is carried out under a *generic* interdegree distribution so that all possible (bidirectional) interlink allocation strategies are covered. By making use of the convexity property and Jensen's inequality, we show that for a fixed mean interdegree, System 2 achieves the highest robustness against random attacks when its interdegree distribution degenerates; i.e., when all nodes have exactly the same number of interedges so that System 2 is equivalent to System 1. Therefore, we conclude that regular allocation

yields the strongest robustness among all possible (bidirectional) interlink allocation strategies. Next, we show that systems with bidirectional interedges can better combat the cascading failures compared to the systems with unidirectional interedges [16]. Together, these results prove the optimality of the interlink allocation strategy in System 1; i.e., regular allocation of bidirectional interedges.

4.1 Analysis of Random Allocation Strategy

We now introduce the auxiliary System 2. Consider two arbitrary networks A and B, each with N nodes, and a discrete probability distribution $F : \mathbb{N} \to [0, 1]$, such that

$$F(j) = \alpha_j, \quad j = 0, 1, \ldots, \tag{15}$$

with $\sum_{j=0}^{\infty} \alpha_j = 1$.

To allocate the interdependency links, we first partition each network randomly into subgraphs with sizes $\alpha_0 N, \alpha_1 N, \alpha_2 N, \ldots$ [1] By doing so, we can obtain subgraphs $\{S_{A_{\alpha_0}}, S_{A_{\alpha_1}}, S_{A_{\alpha_2}}, \ldots\}$ and $\{S_{B_{\alpha_0}}, S_{B_{\alpha_1}}, S_{B_{\alpha_2}}, \ldots\}$, such that

$$|S_{A_{\alpha_j}}| = |S_{B_{\alpha_j}}| = \alpha_j N, \quad j = 0, 1, \ldots.$$

Then, for each $j = 0, 1, \ldots$, assume that each node in the subgraphs $S_{A_{\alpha_j}}$ and $S_{B_{\alpha_j}}$ is assigned j bidirectional interedges. This ensures that the interdegree of each node is a *random* variable drawn from the distribution F; i.e., an arbitrary node will have j interedges with probability α_j, for each $j = 0, 1, \ldots$ It is worth noting that the interdegrees of the nodes are *not* mutually independent since the total number of interedges is fixed at $E = \sum \alpha_j j N$ for both networks.

We have a few more words on the possible implementation of the above random allocation strategy. Observe that each bidirectional edge can be treated equivalently as two unidirectional edges. In this way, there are a total of $2E$ unidirectional interedges in the system, where E edges are going outward from network A and the other E edges are going outward from network B. We randomly match each unidirectional edge going outward from A to a unique edge going outward from B and combine them into a single bidirectional edge. To this end, let the edges going outward from A and B be separately labeled as $\boldsymbol{e} = \{e_1, \ldots, e_E\}$ and $\boldsymbol{e}' = \{e'_1, \ldots, e'_E\}$, respectively. Next, use the Knuth shuffle algorithm [11] to obtain random permutations $\bar{\boldsymbol{e}} = \{\bar{e}_1, \ldots, \bar{e}_E\}$ and $\bar{\boldsymbol{e}}' = \{\bar{e}'_1, \ldots, \bar{e}'_E\}$ of the vectors $\boldsymbol{e}$ and $\boldsymbol{e}'$, respectively. Finally, for each $i = 1, \ldots, E$, match the unidirectional interedges $\bar{e}_j$ and $\bar{e'}_j$ to obtain E bidirectional interedges.

We now analyze the dynamics of cascading failures in System 2 using an iterative approach similar to that in Section 3. For brevity, we skip most of the details and give only an outline of the arguments that lead to the sizes of functional giant components. The main difference from the analysis of Section 3 is that the fractions of nodes in A and B retaining at least one interedge, i.e., the fractions $\bar{A}_i$ and $\bar{B}_i$, need to be calculated differently from (8) and (10) due to the random interdegree of each node.

Owing to the fragmentation from $\bar{B}_{i-1}$ to B_{i-1}, each interedge supporting A could be disconnected with probability $1 - |B_{i-1}|/|\bar{B}_{i-1}|$, triggering further failures in network A at step i. With this insight, the aggregate effect of the failures in B up to stage i can be treated equivalently (with respect to the size of A_i) as removing each interedge supporting A with probability $1 - u_i$. According to Section 3, u_i can be derived as follows:

$$u_i = \prod_{\ell=1}^{(i-1)/2} \frac{|B_{2\ell}|}{|\bar{B}_{2\ell}|} = P_B(p'_{Bi-1}) \quad i = 3, 5, 7 \ldots \tag{16}$$

Similarly, the aggregate effect of node failures in A before step i can be viewed as equivalent to removing each interedge supporting B with probability $1 - v_i$ (with respect to the size of B_i) such that

$$v_i = \frac{|A_1|}{|A|} \prod_{\ell=1}^{i/2-1} \frac{|A_{2\ell+1}|}{|\bar{A}_{2\ell+1}|} = pP_A(p'_{Ai-1}) \quad i = 2, 4, 6 \ldots. \tag{17}$$

In System 2, each node is supported by j interedges with probability α_j. In view of this, at step i, a node in network A would retain at least one interedge with probability $1 - \sum_{j=0}^{\infty} \alpha_j (1 - u_i)^j$. Recalling also that a fraction $1 - p$ of the nodes had already failed before the onset of the cascading failure, the equivalent remaining fraction of network A at stage i is given by

$$\begin{aligned} p'_{Ai} &= p\left(1 - \sum_{j=0}^{\infty} \alpha_j (1 - u_i)^j\right) \\ &= p\left(1 - \sum_{j=0}^{\infty} \alpha_j (1 - P_B(p'_{Bi-1}))^j\right). \end{aligned}$$

Similarly, the equivalent remaining fraction of network B turns out to be

$$\begin{aligned} p'_{Bi} &= 1 - \sum_{j=0}^{\infty} \alpha_j (1 - v_i)^j \\ &= 1 - \sum_{j=0}^{\infty} \alpha_j (1 - pP_A(p'_{Ai-1}))^j. \end{aligned}$$

Hence, the fractional sizes of the giant components at each stage are given (with $p'_{A1} = p$) by

$$\begin{aligned} p_{Ai} &= p'_{Ai} P_A(p'_{Ai}), \\ p'_{Ai} &= p\left(1 - \sum_{j=0}^{\infty} \alpha_j (1 - P_B(p'_{Bi-1}))^j\right), \end{aligned} \tag{18}$$

for $i = 3, 5, 7 \ldots$, and by

$$\begin{aligned} p_{Bi} &= p'_{Bi} P_B(p'_{Bi}), \\ p'_{Bi} &= 1 - \sum_{j=0}^{\infty} \alpha_j (1 - pP_A(p'_{Ai-1}))^j, \end{aligned} \tag{19}$$

for $i = 2, 4, 6, \ldots$ We next show that System 1 is always more robust than System 2 against random attacks by comparing the recursive relations (12)-(13) and (18)-(19).

4.2 Regular Allocation versus Random Allocation

We now compare Systems 1 and 2 in terms of their robustness against random attacks. For convenience, we use a vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots)$ to characterize the interdegree distribution F, where $F(j) = \alpha_j$. Next, we denote the fractions in the recursive relations (18)-(19) as $p_{Ai}(p; \boldsymbol{\alpha})$, $p'_{Ai}(p; \boldsymbol{\alpha})$ and

1. For N large enough, each of these subgraph sizes can be well approximated by an integer.

$p_{Bi}(p;\boldsymbol{\alpha}), p'_{Bi}(p;\boldsymbol{\alpha})$. Also, we let $P_{A^2_\infty}(p;\boldsymbol{\alpha})$ and $P_{B^2_\infty}(p;\boldsymbol{\alpha})$ be the respective steady-state fractions of the functional giant components in the two networks where $1-p$ is the fraction of initially failed nodes in network A. In other words, we set $\lim_{i\to\infty} p_{Ai}(p;\boldsymbol{\alpha}) := P_{A^2_\infty}(p;\boldsymbol{\alpha})$ and $\lim_{i\to\infty} p_{Bi}(p;\boldsymbol{\alpha}) := P_{B^2_\infty}(p;\boldsymbol{\alpha})$. Finally, we denote the critical threshold associated with System 2 by $p_{c_2}(\boldsymbol{\alpha})$.

Assume that network A (respectively, network B) of Systems 1 and 2 have the same size N and the same intradegree distribution such that the functions P_A (respectively, P_B) are identical for both systems. The next result shows that if the two systems are "matched" through their mean interdegrees, i.e., if $k=\sum_{j=0}^{\infty}\alpha_j j$, System 1 always yields stronger robustness than System 2 against random node failures.

Theorem 4.1. *Under the condition*

$$k=\sum_{j=0}^{\infty}\alpha_j j, \tag{20}$$

we have

$$\begin{aligned} P_{A^1_\infty}(p;k) &\geq P_{A^2_\infty}(p;\boldsymbol{\alpha}),\\ P_{B^1_\infty}(p;k) &\geq P_{B^2_\infty}(p;\boldsymbol{\alpha}); \end{aligned} \tag{21}$$

and furthermore

$$p_{c_1}(k)\leq p_{c_2}(\boldsymbol{\alpha}). \tag{22}$$

Proof. Since P_A and P_B are monotonically increasing functions [13], a sufficient condition ensuring (21) will hold is

$$\begin{aligned} p'_{Ai}(p;k) &\geq p'_{Ai}(p;\boldsymbol{\alpha}), \quad i=3,5,7\ldots,\\ p'_{Bi}(p;k) &\geq p'_{Bi}(p;\boldsymbol{\alpha}), \quad i=2,4,6\ldots, \end{aligned} \tag{23}$$

where $p'_{Ai}(p;k)$, $p'_{Bi}(p;k)$, and $p'_{Ai}(p;\boldsymbol{\alpha})$, $p'_{Bi}(p;\boldsymbol{\alpha})$ denote the fractions in the recursive relations (12)-(13), and (18)-(19), respectively. We establish (23) by induction. First observe that $p'_{A1}(p;k)=p'_{A1}(p;\boldsymbol{\alpha})=p$ and the inequality (23) is satisfied for $i=1$. In view of (12)-(13) and (18)-(19), condition (23) for $i=2$ will be satisfied if

$$(1-pP_A(p'_{A1}(p;k)))^k \leq \sum_{j=0}^{\infty}\alpha_j(1-pP_A(p'_{A1}(p;\boldsymbol{\alpha})))^j,$$

or equivalently

$$(1-pP_A(p))^k \leq \sum_{j=0}^{\infty}\alpha_j(1-pP_A(p))^j. \tag{24}$$

Under (20), the convexity of $(1-pP_A(p))^x$ implies (24) by Jensen's inequality. Hence, we get that $p'_{B2}(p;k)\geq p'_{B2}(p;\boldsymbol{\alpha})$ and the base step is completed.

Suppose that the condition (23) is satisfied for each $i=1,2,\ldots,2m-1,2m$. We need to show that (23) holds also for $i=2m+1$ and $i=2m+2$. For $i=2m+1$, the first inequality will be satisfied if it holds that

$$(1-P_B(p'_{B2m}(p;k)))^k \leq \sum_{j=0}^{\infty}\alpha_j(1-P_B(p'_{B2m}(p;\boldsymbol{\alpha})))^j.$$

By the induction hypothesis, we have $P_B(p'_{B2m}(p;k))\geq P_B(p'_{B2m}(p;\boldsymbol{\alpha}))$ since $p'_{B2m}(p;k)\geq p'_{B2m}(p;\boldsymbol{\alpha})$. As a result, the above inequality is satisfied if

$$(1-u)^k \leq \sum_{j=0}^{\infty}\alpha_j(1-u)^j \tag{25}$$

with $u=P_B(p'_{B2m}(p;\boldsymbol{\alpha}))$. As before, under (20), (25) is ensured by the convexity of $(1-u)^x$ in view of Jensen's inequality. The condition $p'_{A2m+1}(p;k)\geq p'_{A2m+1}(p;\boldsymbol{\alpha})$ is now established.

Now let $i=2m+2$. The desired condition $p'_{B2m+2}(p;k)\geq p'_{B2m+2}(p;\boldsymbol{\alpha})$ will be established if

$$\begin{aligned} &(1-pP_A(p'_{A2m+1}(p;k)))^k\\ &\quad\leq \sum_{j=0}^{\infty}\alpha_j(1-pP_A(p'_{A2m+1}(p;\boldsymbol{\alpha})))^j, \end{aligned}$$

or equivalently

$$(1-v)^k \leq \sum_{j=0}^{\infty}\alpha_j(1-v)^j, \tag{26}$$

where we set $v=pP_A(p'_{A2m+1}(p;\boldsymbol{\alpha}))$. The last step follows from the previously obtained fact that $p'_{A2m+1}(p;k)\geq p'_{A2m+1}(p;\boldsymbol{\alpha})$. Once more, (26) follows by the convexity of $(1-v)^x$ and Jensen's inequality. This establishes the induction step and the desired conclusion (21) is obtained.

We next prove the inequality $p_{c_1}(k)\leq p_{c_2}(\boldsymbol{\alpha})$ by way of contradiction. Assume toward a contradiction that $p_{c_2}(\boldsymbol{\alpha})<p_{c_1}(k)$ and fix p such that $p_{c_2}(\boldsymbol{\alpha})<p<p_{c_1}(k)$. Then, let a fraction $1-p$ of the nodes randomly fail in network A of both systems. Since p is less than p_{c_1}, the node failures will eventually lead to complete fragmentation of the two networks in System 1; i.e., we get $P_{A^1_\infty}(p;k)=P_{B^1_\infty}(p;k)=0$. On the other hand, the fact that p is larger than the critical threshold p_{c_2} ensures $P_{A^2_\infty}(p;\boldsymbol{\alpha})>0$ and $P_{B^2_\infty}(p;\boldsymbol{\alpha})>0$ by definition. This clearly contradicts (21) and therefore it is always the case that $p_{c_1}(k)\leq p_{c_2}(\boldsymbol{\alpha})$ under (20). □

We have now established that the regular allocation of bidirectional interedges always yields stronger robustness than any possible random allocation strategy that uses bidirectional links. In the following section, we show that using bidirectional interedges leads to a smaller critical threshold and better robustness than using unidirectional interedges.

4.3 Bidirectional Interedges versus Unidirectional Interedges

We now compare the robustness of System 2 with that of the model considered in [16], hereafter referred to as System 3. As mentioned earlier, the model considered in [16] is based on the random allocation of *unidirectional* interedges and can be described as follows. As with System 2, consider two arbitrary networks A and B, each with N nodes, and a discrete probability distribution $F:\mathbb{N}\to[0,1]$ such that (15) holds. Assume that each node is associated with a random number of supporting nodes from the other network, and

that this random number is distributed according to F. In other words, for each $j = 0, 1, \ldots$, a node has j *inward* interedges with probability α_j. The supporting node for each of these inward edges is selected randomly among all nodes of the other network ensuring that the number of *outward* interedges follows a *binomial* distribution for all nodes.

System 3 was studied in [16] using similar methods to those of Sections 3 and 4.1. This time, after an initial failure of a fraction $1 - p$ of the nodes in network A, the recursive relations for the fractions of giant components at each stage turns [16] out to be (with $p'_{A1} = p$)

$$\begin{aligned} p_{Ai} &= p'_{Ai} P_A(p'_{Ai}), \\ p'_{Ai} &= p\left(1 - \sum_{j=0}^{\infty} \alpha_j (1 - p'_{Bi-1} P_B(p'_{Bi-1}))^j\right), \end{aligned} \tag{27}$$

for $i = 3, 5, 7 \ldots$, and

$$\begin{aligned} p_{Bi} &= p'_{Bi} P_B(p'_{Bi}), \\ p'_{Bi} &= 1 - \sum_{j=0}^{\infty} \alpha_j (1 - p'_{Ai-1} P_A(p'_{Ai-1}))^j, \end{aligned} \tag{28}$$

for $i = 2, 4, 6, \ldots$

Next, we compare Systems 2 and 3 using the recursive relations (18)-(19) and (27)-(28). In doing so, we use the same notation to define the fractions in the recursive relations (18)-(19) as used in Section 4.2, while the fractions in (27)-(28) will be denoted by $p^3_{Ai}(p;\boldsymbol{\alpha}), p'^3_{Ai}(p;\boldsymbol{\alpha})$ and $p^3_{Bi}(p;\boldsymbol{\alpha}), p'^3_{Bi}(p;\boldsymbol{\alpha})$. We let $P_{A^3_\infty}(p;\boldsymbol{\alpha})$ and $P_{B^3_\infty}(p;\boldsymbol{\alpha})$ be the steady-state fractions of functional giant components in System 3 if a fraction $1 - p$ of the nodes initially fail in network A. In other words, we set $\lim_{i\to\infty} p^3_{Ai}(p;\boldsymbol{\alpha}) := P_{A^3_\infty}(p;\boldsymbol{\alpha})$ and $\lim_{i\to\infty} p^3_{Bi}(p;\boldsymbol{\alpha}) := P_{B^3_\infty}(p;\boldsymbol{\alpha})$. Finally, we denote by $p_{c_3}(\boldsymbol{\alpha})$ the critical threshold for System 3.

The next result shows that System 2 is always more robust than System 3 against random node failures.

Theorem 4.2. *We have that*

$$\begin{aligned} P_{A^2_\infty}(p;\boldsymbol{\alpha}) &\geq P_{A^3_\infty}(p;\boldsymbol{\alpha}), \\ P_{B^2_\infty}(p;\boldsymbol{\alpha}) &\geq P_{B^3_\infty}(p;\boldsymbol{\alpha}), \end{aligned} \tag{29}$$

and furthermore,

$$p_{c_2}(\boldsymbol{\alpha}) \leq p_{c_3}(\boldsymbol{\alpha}). \tag{30}$$

Proof. Since $P_A(x)$ and $P_B(x)$ are monotonically increasing [13], a sufficient condition ensuring (29) is given by

$$\begin{aligned} p'_{Ai}(p;\boldsymbol{\alpha}) &\geq p'^3_{Ai}(p;\boldsymbol{\alpha}), \quad i = 1, 3, 5 \ldots, \\ p'_{Bi}(p;\boldsymbol{\alpha}) &\geq p'^3_{Bi}(p;\boldsymbol{\alpha}), \quad i = 2, 4, 6 \ldots. \end{aligned} \tag{31}$$

We establish (31) by induction. First, observe that for $i = 1$, $p'_{A1}(p;\boldsymbol{\alpha}) = p'^3_{A1}(p;\boldsymbol{\alpha}) = p$ and condition (31) is satisfied. Next, for $i = 2$, we see from (19) and (28) that the inequality

$$p'_{B2}(p;\boldsymbol{\alpha}) \geq p'^3_{B2}(p;\boldsymbol{\alpha})$$

will hold if

$$\begin{aligned} &\sum_{j=0}^{\infty} \alpha_j (1 - p P_A(p'_{A1}(p;\boldsymbol{\alpha})))^j \\ &\quad \leq \sum_{j=0}^{\infty} \alpha_j \big(1 - p'^3_{A1}(p;\boldsymbol{\alpha}) P_A\big(p'^3_{A1}(p;\boldsymbol{\alpha})\big)\big)^j. \end{aligned} \tag{32}$$

Since $p'_{A1}(p;\boldsymbol{\alpha}) = p'^3_{A1}(p;\boldsymbol{\alpha}) = p$, it is immediate that (32) is satisfied with equality and this completes the base step of the induction.

Suppose now that condition (31) is satisfied for all $i = 1, 2, \ldots, 2m-1, 2m$. We will establish (31) for $i = 2m+1$ and $i = 2m+2$ as well. Comparing (18) and (27), it is easy to check that for $i = 2m+1$, (31) will hold if

$$\begin{aligned} &\sum_{j=0}^{\infty} \alpha_j (1 - P_B(p'_{B2m}(p;\boldsymbol{\alpha})))^j \\ &\quad \leq \sum_{j=0}^{\infty} \alpha_j \big(1 - p'^3_{B2m}(p;\boldsymbol{\alpha}) P_B\big(p'^3_{B2m}(p;\boldsymbol{\alpha})\big)\big)^j. \end{aligned} \tag{33}$$

By the induction hypothesis, (31) holds for $i = 2m$ so that $P_B(p'^3_{B2m}(p;\boldsymbol{\alpha})) \leq P_B(p'_{B2m}(p;\boldsymbol{\alpha}))$. It is now immediate that (33) holds, since we always have $p'^3_{B2m}(p;\boldsymbol{\alpha}) \leq 1$. This establishes (31) for $i = 2m+1$; i.e., that

$$p'^3_{A2m+1}(p;\boldsymbol{\alpha}) \leq p'_{A2m+1}(p;\boldsymbol{\alpha}). \tag{34}$$

For $i = 2m+2$, we see from (19) and (28) that condition (31) will be satisfied if

$$\begin{aligned} &\sum_{j=0}^{\infty} \alpha_j (1 - p P_A(p'_{A2m+1}(p;\boldsymbol{\alpha})))^j \\ &\quad \leq \sum_{j=0}^{\infty} \alpha_j \big(1 - p'^3_{A2m+1}(p;\boldsymbol{\alpha}) P_A\big(p'^3_{A2m+1}(p;\boldsymbol{\alpha})\big)\big)^j. \end{aligned} \tag{35}$$

In view of (34) and the fact that $p'^3_{A2m+1}(p;\boldsymbol{\alpha}) \leq p$, we immediately obtain (35) and the induction step is now completed. This establishes condition (31) for all $i = 1, 2, \ldots$, and we get (29).

The fact that (29) implies (30) can be shown by contradiction, as in the proof of Theorem 4.1. □

Summarizing, it can be seen from Theorem 4.2 that using bidirectional interedges (System 2) always yields stronger system robustness compared to using unidirectional interedges (System 3). This being valid under an arbitrary distribution $\boldsymbol{\alpha}$ of interedges, we conclude that regular allocation of bidirectional interedges leads to the strongest robustness (among all possible strategies) against random attacks as we recall Theorem 4.1.

5 Numerical Results: the Erdős-Rényi Networks Case

To get a more concrete sense of the above analysis results, we next look at some special cases of network models. In particular, we assume both networks are Erdős-Rényi networks [2], with mean intradegrees a and b, respectively. For this case, the functions $P_A(x)$ and $P_B(y)$ that determine the size of the giant components can be obtained [13] from

$$P_A(x) = 1 - f_A \quad \text{and} \quad P_B(y) = 1 - f_B, \tag{36}$$

where f_A and f_B are the unique solutions of

$$f_A = \exp\{ax(f_A - 1)\} \quad \text{and} \quad f_B = \exp\{by(f_B - 1)\}. \tag{37}$$

In what follows, we derive numerical results for the steady-state giant component sizes as well as critical p_c values. Specifically, we first study System 1 by exploiting the recursive relations (12)-(13) using (36) and (37). Similarly, we derive numerical results for System 2 by using the recursive relations (18)-(19). For both cases, we use extensive simulations to verify the validity of the results obtained theoretically.

5.1 Numerical Results for System 1

Reporting (36) into (14), we get

$$x = p\big(1 - f_B^k\big) \qquad y = 1 - (1 - p(1 - f_A))^k. \tag{38}$$

It follows that the giant component fractions at steady state are given by

$$\begin{aligned} P_{A_\infty} &= p\big(1 - f_B^k\big)(1 - f_A), \\ P_{B_\infty} &= (1 - (1 - p(1 - f_A))^k)(1 - f_B). \end{aligned} \tag{39}$$

Next, substituting (38) into (37) we obtain

$$\begin{aligned} f_A &= \exp\{ap\big(1 - f_B^k\big)(f_A - 1)\}, \\ f_B &= \exp\{b(1 - (1 - p(1 - f_A))^k)(f_B - 1)\}. \end{aligned} \tag{40}$$

We note that the system of equations (40) always has a trivial solution $f_A = f_B = 1$, in which case the functional giant component has zero fraction for both networks. More interesting cases arise for large values of p when there exist nontrivial solutions to (40). In particular, we focus on determining the critical threshold p_c; i.e., the *minimum* p that yields a nontrivial solution of the system. Exploring this further, we see by elementary algebra that (40) is equivalent to

$$\begin{aligned} f_B &= \sqrt[k]{1 - \frac{\log f_A}{(f_A - 1)ap}} \quad \text{if } 0 \le f_A < 1; \ \forall f_B \text{ if } f_A = 1 \\ f_A &= 1 - \frac{1 - \sqrt[k]{1 - \frac{\log f_B}{(f_B - 1)b}}}{p} \quad \text{if } 0 \le f_B < 1; \ \forall f_A \text{ if } f_B = 1. \end{aligned} \tag{41}$$

In general, it may be difficult to derive an explicit expression for p_c. Instead, we can solve (41) graphically for a given set of parameters a, b, k, p and infer the critical threshold p_c using numerical methods. For instance, Fig. 3 shows the possible solutions of the system for several different p values when $a = b = 3$ and $k = 2$. In Figs. 3a, 3b, and 3c, we have $p < p_c$ and there is only the trivial solution $f_A = f_B = 1$ so that both networks go into a complete fragmentation at steady state. In Fig. 3d, we have $p = p_c$ and there exists one nontrivial solution, since the two curves intersect tangentially at one point. In Figs. 3e and 3f, we have $p > p_c$ and there exist two nontrivial intersection points corresponding to two sets of giant component sizes. In these cases, the solution corresponding to the cascading failures should be the point that yields the larger giant component size. In other words, the solution corresponds to the intersection point that is closer to the starting point of the iterative process (see (39)).

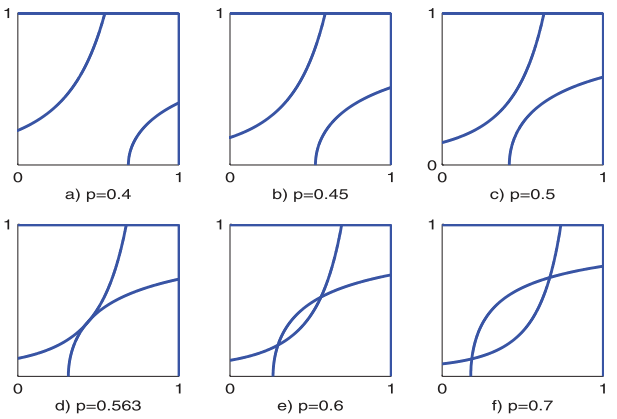


Fig. 3. Possible solutions of the system (41) are depicted for several different p values when $a = b = 3$ and $k = 2$. In all figures, the x-axis represents f_A while the y-axis represents f_B. The critical p_c corresponds to the case where there is only one nontrivial solution to the system, i.e., the case when the two curves are tangential to each other.

In the manner outlined above, we can find the critical threshold p_c for any fixed values of the parameters a, b, and k. As illustrated in Fig. 3, we can further add the tangential condition

$$\frac{df_A}{df_B} \times \frac{df_B}{df_A} = 1 \tag{42}$$

to (41) since the critical p_c value corresponds to the tangent point of the two curves given by (41). Thus, the critical values f_{A_c}, f_{B_c}, and p_c can be computed (numerically) for any given set of parameters through the following system of equations:

$$f_B = \sqrt[k]{1 - \frac{\log f_A}{(f_A - 1)ap}} \quad \text{if } 0 \le f_A < 1; \tag{43}$$

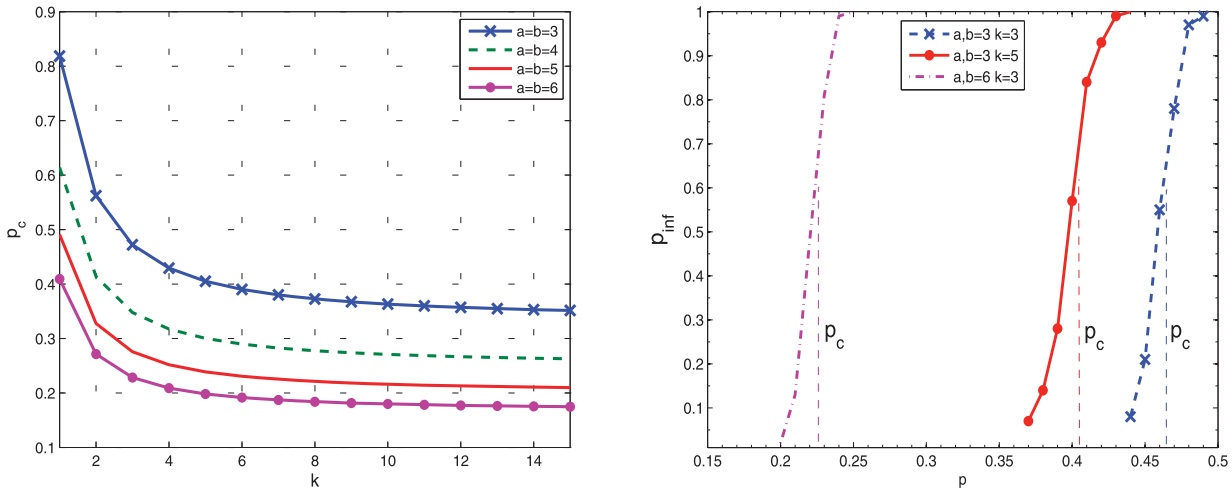
$$f_A = 1 - \frac{1 - \sqrt[k]{1 - \frac{\log f_B}{(f_B - 1)b}}}{p} \quad \text{if } 0 \le f_B < 1; \tag{44}$$

$$\left.\frac{df_A}{df_B}\right|_{(44)} \times \left.\frac{df_B}{df_A}\right|_{(43)} = 1. \tag{45}$$

The analysis results are now corroborated by simulations. In Fig. 4a, we show the variation of p_c with respect to k for different values of $a = b$, where the critical p_c values are obtained by solving the system (45) graphically. To verify these findings, we pick a few sets of values a, b, and k from the curves in Fig. 4a and run simulations with $N = 5{,}000$ nodes to estimate the probability p_{inf} of the existence of a functional giant component in steady state. As expected [3], in all curves we see a sharp increase in p_{inf} as p approaches a critical threshold p_c. It is clear that the estimated p_c values from the sharp transitions in Fig. 4b are in good agreement with the analysis results given in Fig. 4a.

5.2 Numerical Results for System 2

As in System 1, the recursive process (18)-(19) of System 2 stops at an "equilibrium point" where we have $p'_{B2m-2} = p'_{B2m} = x$ and $p'_{A2m-1} = p'_{A2m+1} = y$. This yields the transcendental equations



(a) (b)

Fig. 4. (a) The critical p_c value versus k for the regular allocation strategy (System 1). The plots are obtained by solving the system (45) graphically for various a, b values. It can be seen that as k increases the robustness of the system increases and the critical fraction p_c approaches that of a single network; i.e., $\frac{1}{a}$ [2]. (b) Experimental results for the regular allocation strategy (System 1) with $N = 5{,}000$ nodes. A fraction $1 - p$ of the nodes are randomly removed (from network A) and the corresponding empirical probability p_{inf} for the existence of a functional giant component at steady state is plotted. As expected, in all cases there is a sharp increase when p approaches a critical threshold p_c; for $(a = b = 3, k = 3)$, $(a = b = 3, k = 5)$, and $(a = b = 6, k = 3)$, the critical p_c values are roughly equal to 0.47, 0.41, and 0.23, respectively. Clearly, these p_c values are in close agreement with the corresponding ones of Fig. 4a which are obtained analytically.

$$\begin{aligned} x &= p\left(1 - \sum_{j=0}^{\infty} \alpha_j (1 - P_B(y))^j\right), \\ y &= 1 - \sum_{j=0}^{\infty} \alpha_j (1 - pP_A(x))^j. \end{aligned} \tag{46}$$

The steady-state fraction of nodes in the giant components can be computed by using the relations $\lim_{i\to\infty} p_{Ai} := P_{A_\infty} = xP_A(x)$ and $\lim_{i\to\infty} p_{Bi} := P_{B_\infty} = yP_B(y)$.

In particular, we assume that the interdegree distribution F at each node is a Poisson distribution with mean k, and hence

$$\alpha_j = e^{-k}\frac{k^j}{j!}, \quad j = 0, 1, 2, \ldots, \infty. \tag{47}$$

Substituting (36) and (47) into (46), we get

$$x = p\left(1 - \sum_{j=0}^{\infty} \frac{k^j}{j!} e^{-k} {f_B}^j\right) = p(1 - e^{-k(1-f_B)}), \tag{48}$$

and

$$y = 1 - \sum_{j=0}^{\infty} \frac{k^j}{j!} e^{-k}(1 - p(1 - f_A))^j = 1 - e^{-kp(1-f_A)}. \tag{49}$$

Next, putting (48) and (49) into (37), we find

$$\begin{aligned} f_A &= 1 + \frac{1}{pk}\ln\left(1 + \frac{\ln f_B}{b(1 - f_B)}\right), \quad \text{if } 0 \le f_B < 1; \\ f_B &= 1 + \frac{1}{k}\ln\left(1 + \frac{\ln f_A}{ap(1 - f_A)}\right), \quad \text{if } 0 \le f_A < 1; \\ &\forall f_A \text{ if } f_B = 1; \quad \forall f_B \text{ if } f_A = 1. \end{aligned} \tag{50}$$

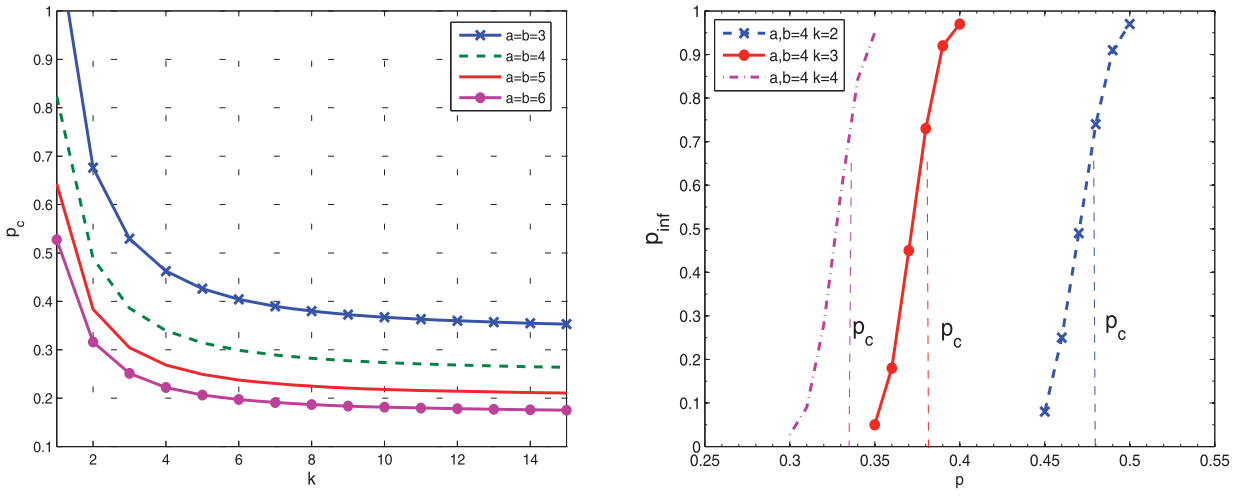
As in the case for System 1, the critical threshold p_c for System 2 corresponds to the tangential point of the curves given by (50), and can be obtained by solving (50) graphically.

We now check the validity of these analytical results via simulations. In Fig. 5a, we show the variation of analytically obtained p_c values with respect to average interdegree k for different values of $a = b$. To verify these results, we pick a few sets of values a, b, and k from the curves in Fig. 5a and run simulations with $N = 5{,}000$ nodes to estimate the probability p_{inf} of the existence of a functional giant component in steady state. As expected [3], in all curves we see a sharp increase in p_{inf} as p approaches a critical threshold p_c. It is also clear from Fig. 5b that, for all parameter sets, such sharp transition occurs when p is close to the corresponding p_c value given in Fig. 5a.

5.3 A Comparison of System Robustness

In Sections 4.2 and 4.3, we have analytically proved that the regular allocation of bidirectional interedges leads to the strongest robustness against random attacks. To get a more concrete sense, we now numerically compare the system robustness of these strategies in terms of their critical thresholds p_c. Specifically, we consider coupled Erdős-Rényi networks with mean intradegrees a and b. For the sake of fair comparison, we assume that the mean interdegree is set to k for all systems; in both Systems 2 and 3, the interdegree distribution F at each node is assumed to be Poisson. The critical threshold value p_c corresponding to all three strategies are compared under a variety of conditions. For Systems 1 and 2, we use the numerical results derived in Sections 5.1 and 5.2, respectively, while for System 3 we use the numerical results provided in [16].

First, we compare System 1 with System 3 to see the difference between the proposed regular interedge allocation strategy and the strategy in [16]. Fig. 6a depicts p_c as a function of mean interdegree k for various values of $a = b$,



(a) (b)

Fig. 5. a) The critical p_c value versus k for the random allocation strategy (System 2). The plots are obtained by solving the system (50) graphically for various a, b values. It is seen that the critical p_c can be larger than one in some cases (e.g., for $a = b = 3$ and $k = 1$) meaning that the system collapses already without any node being attacked. This is because, due to the random allocation of interedges, a nonnegligible fraction of the nodes receive no interedges and become automatically nonfunctional even if they are not attacked. b) Experimental results for System 2 with $N = 5{,}000$ nodes. A fraction $1 - p$ of the nodes are randomly removed (from network A) and the corresponding empirical probability p_{inf} for the existence of a functional giant component at the steady state is plotted. As expected, in all cases there is a sharp increase when p approaches to a critical threshold p_c; for $(a = b = 4, k = 2)$, $(a = b = 4, k = 3)$ and $(a = b = k = 4)$, the critical p_c values are roughly equal to 0.480, 0.380, and 0.335, respectively. Clearly, these p_c values are in close agreement with the corresponding ones of Fig. 4a which are obtained analytically.

while Fig. 6b depicts the variation of p_c with respect to $a = b$ for different k values. In all cases, it is seen that regular allocation of bidirectional interedges yields a much smaller p_c (and thus, a more robust system) than random allocation of unidirectional interedges. For instance, for $a = b = k = 4$, System 3 [16, Figure 2] gives $p_c = 0.43$, whereas, as seen via Fig. 6a, System 1 yields a critical threshold at 0.317. This is a significant difference since it means that System 3 can have a functioning giant component despite a random failure of at most 57 percent of the nodes, whereas System 1, which uses the regular interedge allocation scheme proposed in this paper, is resistant to a random failure of up to 68 percent of the nodes. Indeed, in some cases, our strategy can outperform that in [16] even with half the (mean) interdegree per node. For instance, when $a = b = 4$, our strategy yields $p_c = 0.414$ with only $k = 2$ as compared to $p_c = 0.43$ of the System 3 with $k = 4$.

We also compare System 1 with System 2 in order to see the improvement in allocating bidirectional edges regularly rather than randomly. Fig. 6c depicts p_c as a function of mean interdegree k for various values of $a = b$. It is seen that, in all cases, System 1 yields a lower p_c (and thus a more robust system) than System 2. For example, when $a = b = 3$ and $k = 2$, we get $p_c = 0.56$ for System 1, while for System 2, we find that $p_c = 0.68$. The difference is significant in that it corresponds to a resiliency against a random failure of up to 44 percent of the nodes in System 1 as compared to 32 percent in System 2.

Finally, in order to better illustrate the optimality of System 1 in terms of system robustness, we depict in Fig. 6d the variation of p_c with respect to $a = b$ for different values of k in all three systems. It is clear that the proposed regular allocation strategy in System 1 always yields the lowest p_c and thus provides the best resiliency against random attacks. We also see that System 2 always outperforms System 3, showing the superiority (in terms of robustness) of using bidirectional interedges rather than unidirectional edges.

We believe that the drastic improvement in robustness against random attacks seen in System 1 has its roots as follows. First, in the absence of intratopology information, it is difficult to tell which nodes play more important roles in preserving the connectivity of the networks. Thus, in order to combat *random* attacks, it is reasonable to treat all nodes equally and give them equal priority in interedge allocation. Second, in Systems 2 and 3, there may exist a nonnegligible fraction of nodes with no interedge support from the other network. Those nodes are automatically nonfunctional even if they are not attacked. But, the regular allocation scheme promises a guaranteed level of support, in terms of interedges, for all nodes in both networks. Finally, using bidirectional interedges ensures that the amount of support provided is equal to the amount of support being received for each node. Thus, the use of bidirectional interedges increases the *regularity* of the support-dependency relationship relative to unidirectional interedges, and this may help improve the system robustness.

6 Conclusion and Future Work

We study the robustness of a cyber-physical system in which a cyber network overlays a physical network. To improve network robustness against random node failures, we develop and study a regular allocation strategy that allots a fixed number of internetwork edges to each node. Our findings reveal that the proposed regular allocation strategy yields the optimal robustness among all strategies when no information regarding the intratopologies of the

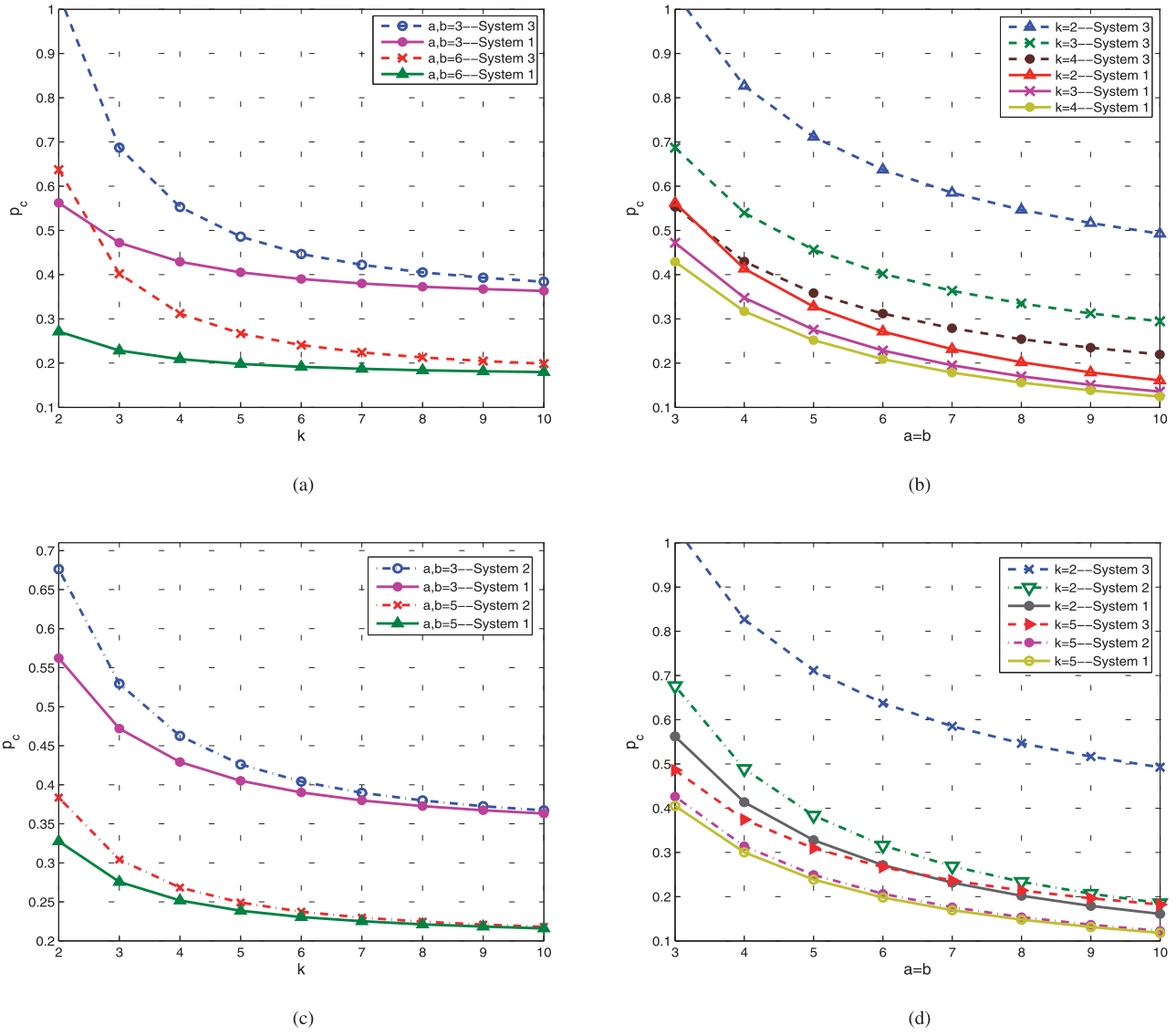


Fig. 6. A comparison of Systems 1, 2, and 3 in terms of their critical p_c values when expected interdegree of any node is set to k. For Systems 2 and 3, the distribution of the number of interedges is assumed to be Poisson. In all figures, dashed lines correspond to System 3, dash-dot lines represent System 2, and solid lines stand for System 1. (a) p_c versus k is depicted for different values of $a = b$ in Systems 1 and 3. (b) p_c versus $a = b$ is depicted for various k values in Systems 1 and 3. In all cases, we see that the regular allocation of bidirectional interedges yields a smaller p_c than the Poisson distribution of unidirectional interedges with the same mean value k. (c) p_c versus k is depicted for different values of $a = b$ in Systems 1 and 2. It is clear that System 1 yields a lower p_c (and thus a higher robustness) than System 2 in all cases. (d) p_c versus $a = b$ is depicted for various k values in Systems 1, 2, and 3. In all cases, System 1 yields the lowest p_c (i.e., highest robustness), while System 3 has the highest p_c (i.e., lowest robustness) and System 2 stands in between.

individual networks is available. For future work, we conjecture that in the presence of such information, the topology of the networks can be exploited to further improve the robustness of cyber-physical systems against cascading failures.

It is also of interest to study models that are more realistic than the existing ones. For instance, in a realistic setting, one can expect to see a certain correlation between the interedges and the intraedges of a system owing to the geographical locations of the nodes. Also, some of the nodes may be *autonomous*, meaning that they do not depend on nodes of the other network to function properly; in that case, one can expect the regular allocation strategy to still be the optimum if the nodes that are autonomous are not known. Clearly, there are still many open questions centered around network interdependence in cyber-physical systems. We are currently investigating related issues along this avenue.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their careful reading of the original manuscript; their comments helped improve the final version of this paper. They also thank Prof. Armand Makowski for his insightful comments. Part of this material was presented in [18]. This research was supported in part by the US National Science Foundation (NSF) grants No. CNS-0905603, CNS-0917087, and the DTRA grant HDTRA1-09-1-0032.

REFERENCES

[1] A.L. Barabási and L. Albert, "Emergence of Scaling in Random Networks," *Science,* vol. 286, pp. 509-512, 1999.

[2] B. Bollobás, *Random Graphs,* Cambridge Studies in Advanced Math. Cambridge Univ. Press, 2001.

[3] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, and S. Havlin, "Catastrophic Cascade of Failures in Interdependent Networks," *Nature,* vol. 464, pp. 1025-1028, 2010.

[4] S.V. Buldyrev, N.W. Shere, and G.A. Cwilich, "Interdependent Networks with Identical Degrees of Mutually Dependent Nodes," *Physical Rev. E,* vol. 83, p. 016112, 2011.

[5] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, "Network Robustness and Fragility: Percolation on Random Graphs," *Physical Rev. Letters,* vol. 85, no. 25, pp. 5468-5471, 2000.

[6] W. Cho, K.I. Goh, and I.M. Kim, "Correlated Couplings and Robustness of Coupled Networks," arXiv:1010.4971v1[physics. data-an], 2010.

[7] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Physical Rev. Letters,* vol. 85, no. 21, pp. 4626-4628, 2000.

[8] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function.* Cambridge Univ. Press, 2010.

[9] CPS Steering Group, "Cyber-Physical Systems Executive Summary," http://varma.ece.cmu.edu/summit/CPS-Executive-Summary.pdf, 2008.

[10] X. Huang, J. Gao, S.V. Buldyrev, S. Havlin, and H.E. Stanley, "Robustness of Interdependent Networks under Targeted Attack," *Physical Rev. E,* vol. 83, p. 065101, 2011.

[11] D.E. Knuth, *The Art of Computer Programming,* vol. 2. Addison-Wesley, 1981.

[12] M.E.J. Newman, "Spread of Epidemic Disease on Networks," *Physical Rev. E,* vol. 66, no. 1, p. 16128, 2002.

[13] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, "Random Graphs with Arbitrary Degree Distributions and Their Applications," *Physical Rev. E,* vol. 64, no. 2, p. 26118, 2001.

[14] 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 Rev. Letters,* vol. 105, p. 048701, 2010.

[15] C.M. Schneider, N.A.M. Araujo, S. Havlin, and H.J. Herrmann, "Towards Designing Robust Coupled Networks," arXiv:1106.3234v1 [cond-mat.stat-mech], 2011.

[16] J. Shao, S.V. Buldyrev, S. Havlin, and H.E. Stanley, "Cascade of Failures in Coupled Network Systems with Multiple Support-Dependent Relations," *Physical Rev. E,* vol. 83, p. 036116, 2011.

[17] A. Vespignani, "Complex Networks: The Fragility of Interdependency," *Nature,* vol. 464, pp. 984-985, Apr. 2010.

[18] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "On Allocating Interconnecting Links against Cascading Failures in Cyber-Physical Networks," *Proc. Third Int'l Workshop Network Science for Comm. Networks (NetSciCom '11),* Apr. 2011.

Osman Yağan (S'07) received the BS degree in electrical and electronics engineering from the Middle East Technical University, Ankara (Turkey) in 2007, and the PhD degree in electrical and computer engineering from the University of Maryland, College Park, MD, in 2011. He was a visiting postdoctoral Scholar at Arizona State University during Fall 2011. Since December 2011, he has been a postdoctoral research fellow in the Cyber Security Laboratory (CyLab) at the Carnegie Mellon University. His research interests include wireless network security, dynamical processes in complex networks, percolation theory, random graphs and their applications. He is student member of the IEEE.

Dajun Qian received the BS and MS degrees in electrical engineering from Southeast University, Nanjing, China, in 2006 and 2008, respectively. He is currently working toward the PhD degree from the Department of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ. His research interests include wireless communications, social networks and cyber-physical systems. He is a student member of the IEEE.

Junshan Zhang (S'98-M'00-F'12) received the PhD degree from the School of ECE at Purdue University in 2000. He joined the EE Department at Arizona State University in August 2000, where he has been professor since 2010. His research interests include communications networks, cyber-physical systems with applications to smart grid, stochastic modeling and analysis, and wireless communications. His current research focuses on fundamental problems in information networks and network science, including network optimization/control, smart grid, cognitive radio, and network information theory. He is the recipient of the ONR Young Investigator Award in 2005 and the US National Science Foundation (NSF) CAREER award in 2003. He received the Outstanding Research Award from the IEEE Phoenix Section in 2003. He served as TPC cochair for WICON 2008 and IPCCC '06, TPC vice chair for ICCCN '06, and the general chair for IEEE Communication Theory Workshop 2007. He was an associate editor for *IEEE Transactions on Wireless Communications*. He is currently an editor for the *Computer Network journal* and *IEEE Wireless Communication Magazine*. He coauthored a paper that won IEEE ICC 2008 best paper award, and one of his papers was selected as the INFOCOM 2009 Best Paper Award Runner up. He is TPC cochair for INFOCOM 2012. He is a fellow of the IEEE.

Douglas Cochran (S'86-A'90-M'92-SM'96) received the MS and PhD degrees in applied mathematics from Harvard University, Cambridge, MA, and degrees in mathematics from the Massachusetts Institute of Technology, Cambridge, and the University of California, San Diego. Since 1989, he has been on the faculty of the School of Electrical, Computer and Energy Engineering, Arizona State University (ASU), Tempe, and is also affiliated with the School of Mathematical and Statistical Sciences. Between 2005 and 2008, he served as assistant dean for research in the Ira A. Fulton School of Engineering at ASU. Between 2000 and 2005, he was a program manager for mathematics at the US Defense Advanced Research Projects Agency (DARPA) and he held a similar position in the US Air Force Office of Scientific Research between 2008 and 2010. Prior to joining the ASU faculty, he was senior scientist at BBN Systems and Technologies, Inc., during which time he served as resident scientist at the DARPA Acoustic Research Center and the Naval Ocean Systems Center. He has been a visiting scientist at the Australian Defence Science and Technology Organisation and served as a consultant to several technology companies. He was general cochair of the 1999 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-99) and cochair of the 1997 US-Australia Workshop on Defense Signal Processing. He has also served as associate editor for book series and journals, including the *IEEE Transactions on Signal Processing*. His research is in applied harmonic analysis and statistical signal processing. He is senior member of the IEEE.