An Experimental Evaluation of Robustness of Networks

Hyoungshick Kim and Ross Anderson

***Abstract*—Models of conflict in networks provide insights into applications ranging from epidemiology to guerilla warfare. Barabási, Albert, and Jeong modeled selective attacks on networks in which an attacker targets high-order nodes to destroy connectivity; Nagaraja and Anderson extended this to iterated attacks where the attacker and defender take turns removing and rebuilding nodes and edges according to given strategies. We extend the iterative model by introducing the cost required to perform network operations. This gives a much finer granularity than previous models, whether we are interested in network resilience against random failures or intentional attacks. We empirically study how to design more effective attacks and/or defenses through intensive simulation on several well-known network topologies, including the three real-world networks. In particular, an effective defense against many attacks is to add new links connecting low-centrality nodes to maintain the overall balance of network centrality.**

***Index Terms*—Iterative attacks and defenses, network resilience, network robustness.**

I. Introduction

MANY IMPORTANT phenomena depend on networks, from social interactions between people to explicit networks such as the Internet and supply chains. Recent advances in the theory of networks have provided us with mathematical and computational tools to understand them better [1], [2]. Often the topology of a network has distinctive features, such as vertex order distribution, clustering and characteristic path lengths, which can be explained in terms of its evolution and which in turn explain some aspects of its behavior. For example, networks that grow by preferential attachment may acquire a power-law distribution of vertex order that in turn makes them robust against random node failures—yet this distribution also makes them vulnerable to attacks targeted on high-degree nodes. Insights like this can inform activities from epidemiology to policing. Doctors may first vaccinate those individuals who are likely to come into contact with most others, while police forces tackle criminal gangs by placing the most highly connected criminals under arrest or surveillance. They also apply to technological networks such as the Internet, the electrical power grid, and transportation networks; these are also robust to random failures but vulnerable to targeted attacks [3].

Network failure models are not limited to the one-shot case. When an attack occurs, a defender tries to minimize the damage by deploying new resources, while the attacker may then follow through by causing further damage. In other words, we need to consider dynamic interaction between an attacker and a defender over multiple periods. Nagaraja and Anderson [4] developed a framework to explore iterated attack and defense operations: an attacker removes k_a nodes from a network at each attack round, and a defender adds k_d nodes to the network at each defense round. The attacker's aim is to decrease network connectivity or efficiency, which can be measured as the size of the largest connected component or the average shortest path length in the network, while the defender's aim is the opposite [4]. This models how a network will likely evolve under continuous attack, based on evolutionary game theory [5]. It enables us to investigate what sort of attack and defense strategies might prevail in counterrevolutionary warfare: a conflict in which peacekeepers identify and arrest rebel ring leaders, while the rebels constantly recruit and reorganize themselves. However, Nagaraja and Anderson's [4] work has two limitations.

1) It does not model the costs of creating new nodes and edges realistically, as the defender is allowed to create a fixed number of new nodes at each round plus an arbitrary number of edges. In practice the cost of establishing edges is not zero, so it would be preferable to enable the defender to allocate his budget to nodes and edges with some fixed marginal cost of substitution.
2) Their experimental results were limited to a single scenario, namely a Barabási–Albert scale-free network consisting of 400 nodes with most of the experiments involving $k_a = k_d = 10$ and a few dozen rounds.

We extend their work into a more generalized model; while Nagaraja and Anderson simply assumed a newly recruited node could form the right number of new connections to pursue any given defense strategy, we vary the budgets to limit newly added nodes and their connections. We note that to make a network highly robust against node removal attacks, a simple strategy is just to increase its network density. In this paper, we seek to answer a simple question: "how much does a connection cost?" We want a quantitative understanding of the correlation between network density and resilience to

Manuscript received September 26, 2011; revised March 20, 2012; accepted July 26, 2012. Date of publication January 17, 2013; date of current version April 17, 2013. The work of H. Kim was funded by Northrop Grumman Corporation, Falls Church, VA.

H. Kim is with the University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: hyoung@ece.ubc.ca).

R. Anderson is with the Computer Laboratory, University of Cambridge, Cambridge, CB2 1TN, U.K. (e-mail: ross.anderson@cl.cam.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSYST.2012.2221851

random failures or attacks as well as insight into the evolution of networks by iterative growth and shrinkage processes.

We therefore empirically analyze the effects of attack/defense strategies with more realistic budgets on well-known network topologies: two Erdős-Rényi random graphs, two Barabási–Albert scale free networks [6], two Chord networks [7], two Hypergrid graphs [8], a Transit-Stub graph [9], a Watts-Strogatz small world network [10], a Content-Addressable network [11], a PRU network [12], and the three real-world networks [13]–[15].

Our experimental results show that the strategy of connecting low centrality nodes produces the best overall performance for maintaining network connectivity for a given budget. Also, even simple defense strategies (e.g., adding nodes and their connections randomly) can be effective enough to fight against sophisticated attacks (e.g., removing nodes with high centrality in high priority) if we can increase the number of connections per node past a certain threshold.

II. Related Work

Albert *et al.* [3] showed that attacks targeted on high-degree nodes in scale-free networks are much more effective than random attacks; the size of the largest connected component is rapidly reduced. This is because scale-free networks get much of their connectivity from few nodes of high degree. It is hard to remove enough of those hub nodes in a random attack, but if they are targeted deliberately, then connectivity decreases dramatically. Holme *et al*. [16] got similar experimental results by doing such attacks on edges, and also suggested using centrality as an alternative to degree for targeting. Zhao *et al*. [17] studied the circumstances under which a scale-free network suffers cascading breakdown caused by the successive failures of hub nodes.

Nagaraja and Anderson [4] extended this by introducing a framework from evolutionary game theory to explore the effectiveness of iterated attack and defense operations. They showed that a defender can make a network resilient to attacks by replacing highly connected nodes with cliques—small groups of vertices that are fully connected to each other and which share the outgoing edges that previously went to a single highly connected node. This strategy, however, requires the modification of the existing connections in a network and may have high implementation costs if adding an extra edge is expensive. Clique-based defense strategies are likely to be nontrivial in some environments such as wired networks.

Recently, Domingo-Ferrer has been extending Nagaraja and Anderson's model to weighted and directed networks [18]. He also found that the costs of attacks/defenses were not clearly defined in [4] and discussed the economic aspects of the attack and defense strategies by considering the cost of node destruction/replenishment. However, it is hard to evaluate the usefulness of the estimated cost functions; as already noted, the cost of rewiring edges is usually not zero. We therefore set out to refine the iterated attack/defense model to take account of the cost of edges added or changed as well as the number of nodes added during the defense phase. We also extended the modeling to a much larger range of graph topologies.

III. Model

Our framework can formally be represented as a game on a graph G by iterating attack and defense operations for a certain number of rounds. Here, an attacker's objective is to maximize disruption to the network while a defender tries to minimize it.

Each round consists of an attack phase followed by a defense phase. In an attack phase, an attacker picks the existing k_a nodes from the graph G according to her attack strategy and then removes the selected nodes and their associated edges. In a defense phase, a defender creates k_d new nodes and then adds them by sequentially connecting a new node v with m edges to m different nodes already present in G according to his strategy. Unlike Nagaraja and Anderson's model, we do not allow the defender to rewire the existing edges in the graph G—this may be expensive compared to establishing new edges in some real environments such as wired networks since an edge rewiring operation can actually be treated as a combination of destroying existing edges and establishing new edges. We assume that the defender has no knowledge of which nodes and their connections are disappeared (otherwise the best strategy may be to restore the last status of the network when $k_a = k_d$).

In order to measure the effectiveness of attacks and defenses, we use the size of the largest connected component after a certain number of rounds as the metric, as other authors in this field have done.

A. Attack Strategies

An attack strategy is a strategy (an algorithm) to select k_a nodes to be removed from a graph $G = (V, E)$ in an attack phase. We here assume k_a is a constant. We consider the following three strategies.

1) Random removal: Pick a node randomly from G and remove it and its associated edges. Repeat this process k_a times.
   a) This strategy is very simple and efficient: An attacker does not need any knowledge of the network topology. The total running time is $O(d(G))$ if we ignore the cost of random selection where $d(G)$ is the average node degree in the graph G.
2) High-degree removal: Pick the highest degree node from G and remove it and its associated edges. Repeat this process k_a times.
   a) This strategy requires global knowledge of the node degree. The total running time is $O(|V| \log |V|)$ since the nodes are sorted in decreasing order with respect to their degree.
3) High-centrality removal: Pick the highest-betweenness centrality node from G and remove it and its associated edges. Repeat this process k_a times. Here, we only consider betweenness centrality since this is known to be more related to network connectivity than closeness or eigenvector centrality. Betweenness centrality $b(u)$ is calculated for a node u as the proportion of shortest paths between all node pairs in the network that pass

through u

$$b(u) = \frac{1}{(|V|-1)\cdot(|V|-2)} \sum_{s\neq u, t\neq u \in V} \frac{\sigma_{s,t}(u)}{\sigma_{s,t}} \quad (1)$$

where $\sigma_{s,t}$ is the total number of shortest paths from source node s to destination node t, and $\sigma_{s,t}(u)$ is the number of shortest paths from source node s to destination node t which actually pass through node u.

a) This strategy requires knowledge of the network topology and has total running time $O(|V|\cdot|E| + |V|\log|V|)$ (the nodes are sorted in decreasing order with respect to their betweenness centrality which can be computed in $O(|V|\cdot|E|)$ time [19].)

We use $\mathbf{A}^{random}$, $\mathbf{A}^{degree}$, and $\mathbf{A}^{central}$ to denote random removal, high-degree removal, and high-centrality removal attack strategies, respectively. We note that all of these attack strategies remove the same number of k_a nodes at every round if the remaining nodes of the graph G is greater than or equal to k_a.

B. Defense Strategies

A defense strategy is a strategy (algorithm) to connect k_d newly recruited nodes to the existing nodes in a graph $G = (V, E)$ in a defense phase. We here assume k_d is a constant. A new node v is connected with m new edges to m different nodes. We consider the following three strategies.

1) Random replenishment: Create a new node and add it to G such that the node is connected with m new edges to m randomly selected different nodes. Repeat this process k_d times.
   a) This strategy requires no knowledge of the network topology. The total running time is $O(m)$ if we ignore the cost of random selection.
2) Preferential replenishment: Create a new node and add it to G such that the node is connected with m new edges to m different nodes with probability proportional to their degree (i.e., the node is connected to an existing node u with the probability $p(u) = d(u)/\sum_{v\in V} d(v)$ where $d(u)$ is the node u's degree in G). Repeat this process k_d times.
   a) This strategy requires knowledge of the node degree. The total running time is $O(|V|)$ if we ignore the cost of random selection.
3) Balanced replenishment: Create a new node and add it to G such that the node is connected with m new edges to m different nodes with probability inversely proportional to their betweenness centrality (i.e., the node is connected to an existing node u with the probability $p(u) = (b(u)+\epsilon)^{-1}/\sum_{v\in V}(b(v)+\epsilon)^{-1}$ where $b(u)$ is the node u's betweenness centrality in G and ϵ is a very small constant to prevent division by zero). Repeat this process k_d times.
   a) This strategy requires knowledge of the network topology. The total running time is $O(|V|\cdot|E|)$ if we ignore the cost of random selection.

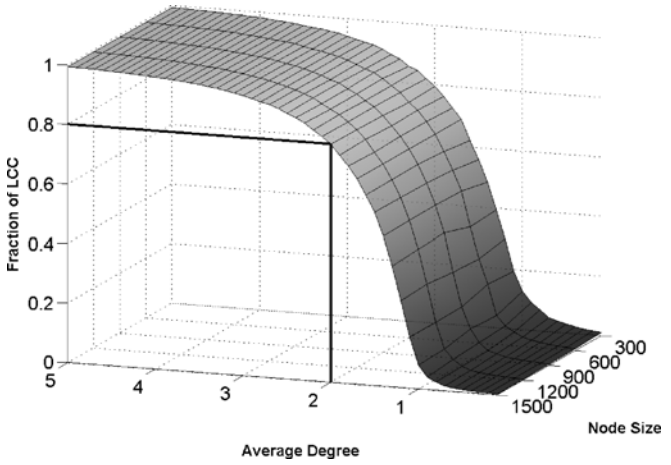

Fig. 1. Average size of the largest connected components over Erdős-Rényi random graphs. With each pair of node size (from 300 to 1500) and average degree (from 0 to 5), we generate 100 random graphs and compute the average size of the largest connected components over these graphs. All decay rates of the largest connected component in random graphs increase dramatically at $d(G) = 2$.

We use $\mathbf{D}^{random}$, $\mathbf{D}^{prefer}$, and $\mathbf{D}^{balance}$ to denote random replenishment, preferential replenishment, and balanced replenishment defense strategies, respectively. We note that all of these defense strategies add the same numbers of k_d nodes and $m\cdot k_d$ edges at every round.

IV. The Necessary Network Density For Robust Networks

A fundamental question is whether we can make a network resilient against node failures or attacks by just increasing the number of edges. If a graph has too few edges, it is necessarily disconnected. As a network becomes better connected so its robustness will in general increase. And when providing network robustness we want to know the optimal edge budgets for newly recruited nodes.

To get an insight into this problem, we generated Erdős-Rényi random graphs by varying the parameters (from 300 to 1500 for the number of nodes n and from 0 to 5 for the average node degree $d(G)$) and found the largest connected component in each graph. We repeated this 100 times for each tuple $(n, d(G))$ and computed the average size of the largest connected components over the sample. Fig. 1 demonstrates the fraction of nodes remaining in the largest connected component with $d(G)$.

The decay rates of the largest connected components in all Erdős–Rényi random graphs show almost the same pattern regardless of n. The curve has a gentle slope until $d(G) = 2$ then plunges toward 0 when $d(G) < 2$. As a selective attack is at least as effective as random edge removal, we can always expect a significant number of nodes to be disconnected from the network if $d(G) < 2$. And when $d(G) = 2$, the largest connected component in G has a tendency to form a tree-like graph structure which can be easily decomposable or be already reduced to a small component (see Fig. 2).

In Section V we will explore the relationship between the size of the largest connected component and the average

Fig. 2. Example graphs with $n = 21$ and $d(G) = 2$. The largest connected component in graph G has (a) a tree-like graph structure or is already reduced to (b) a small component when $d(G) = 2$.

degree through intensive simulation results on various network topologies.

V. Simulation Results

Because of highly nonlinear characteristics of network structures, it is very difficult to establish mathematical models with a closed form solution. So, we use the simulation model as an alternative to the theoretical model. We experimentally tested the attack and defense strategies discussed in Sections III-A and III-B on three real-world and 12 synthetic networks for evaluating the performance of the strategies as follows.

1) Random graphs ($G_R^{0.005}, G_R^{0.01}$—we denote as G_R^x the random graph with the linking probability x): Random graphs are fundamental and useful for modeling problems in many applications.
2) Barabási–Albert scale free networks [6] (G_{BA}^2, G_{BA}^4—we denote as G_{BA}^x the Barabási–Albert network where each new node is connected to x existing nodes): Scale-free networks are abundant in nature and society, describing such diverse systems as the world wide web, the web of human sexual contacts, or the chemical network of a cell. Albert *et al.* [3] showed that scale-free networks are resistant to random failures but vulnerable to targeted attacks since a few hubs dominate their topology.
3) Chord networks [7] (G_C^2, G_C^4—we denote as G_C^x the Chord network where x is the minimum degree of each node): Chord network is a typical structured peer-to-peer overlay network. Chord network is simple and useful to build a fault-tolerant and decentralized peer-to-peer structure.
4) Hypergrid graphs [8] (G_H^4, G_H^8—we denote as G_H^x the hypergrid network where x is the maximum degree of each node): Hypergrid graph is built for peer-to-peer systems by enforcing low graph diameter and fixed node degree.
5) Transit–Stub graph [9] (G_{TS}): The Transit–Stub model is a hierarchical graph generation model that produces graphs having a structure similar to the Internet.
6) Watts–Strogatz small world network [10] ($G_{WS}^{4,0.1}$—we denote as $G_{WS}^{x,y}$ the Watts-Strogatz small world network with the initial x neighbours and the linking probability y): The Watts–Strogatz model is a random graph generation model that produces graphs with small-world properties, including short average path lengths and high clustering.

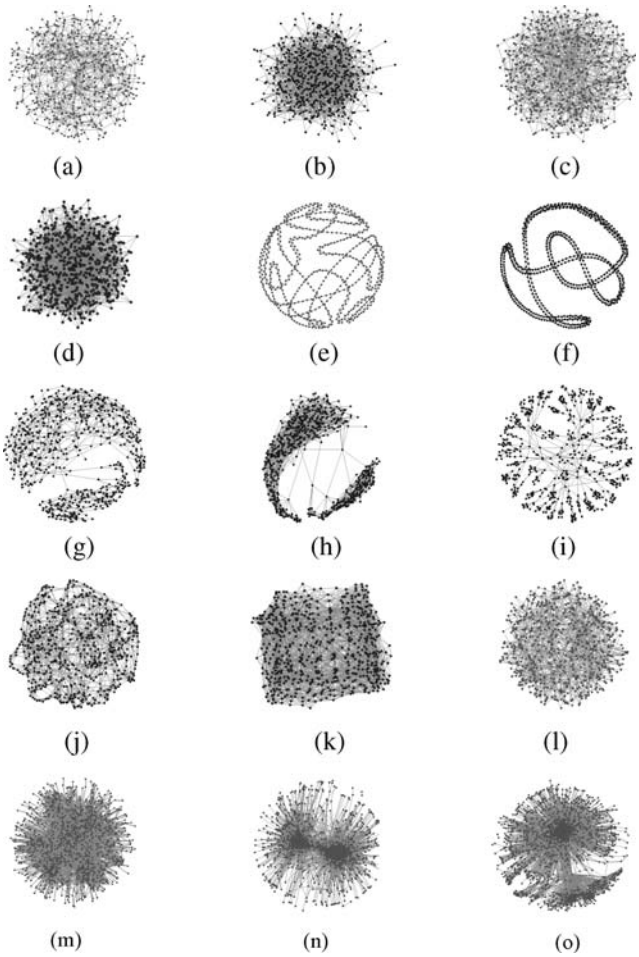

Fig. 3. Networks used in the experiments. (a) $G_R^{0.005}$. (b) $G_R^{0.01}$. (c) G_{BA}^2. (d) G_{BA}^4. (e) G_C^2. (f) G_C^4. (g) G_H^4. (h) G_H^8. (i) G_{TS}. (j) $G_{WS}^{4,0.1}$. (k) G_{CA}. (l) $G_{PRU}^{50,12,2}$. (m) G_{MAIL}. (n) G_{BLOG}. (o) G_{AIR}.

7) Content-addressable network [11] (G_{CA}): Content-addressable network is designed for a distributed and scalable peer-to-peer systems.
8) PRU network [12] ($G_{PRU}^{50,12,2}$—we denote as $G_{PRU}^{x,y,z}$ the PRU network with the initial c nodes in cache, the minimum degree y of nodes and the maximum degree z of nodes): PRU networks are suggested by Pandurangan et al. to produce graphs having a structure similar to the unstructured P2P networks.
9) E-mail network [13] (G_{MAIL}): The e-mail network is obtained by collecting mutual email communication interactions through email logs from a company.
10) Blog network [14] (G_{BLOG}): The blog network is obtained by analyzing the network structure of political blogs published around of the time the 2004 U.S. presidential election.
11) Airport network [15] (G_{AIR}): The airport network is obtained by analyzing routes between all the United States airports in 2010.

The network topologies are shown in Fig. 3.

We summarize the properties of the networks used in the experiments in Table I. Given a graph G, let $s(G)$ and $d(G)$

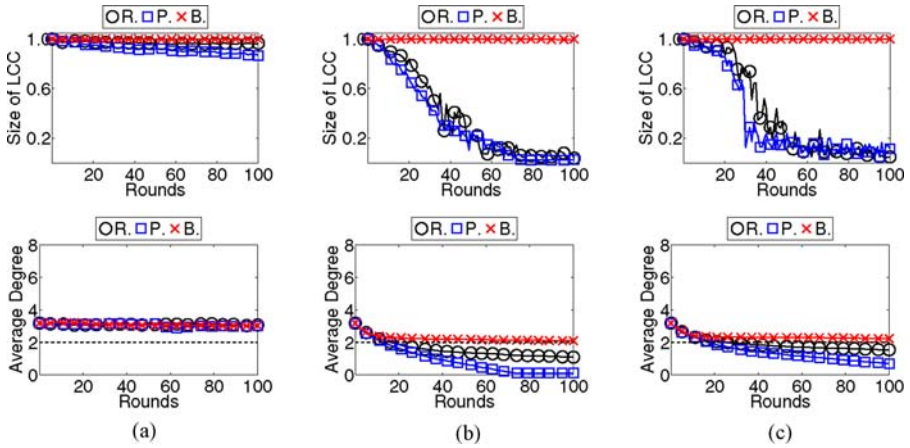

Fig. 4. Changes in the size of the largest connected component and the average degree in $G_R^{0.005}$ over rounds. The first row graphs show the changes in the size of the largest connected component and the second row graphs show the changes in the average degree. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

TABLE I
PROPERTIES OF THE NETWORKS

	$\|V\|$	$\|E\|$	Diameter	Density	$d(G)$	$s(G)$
$G_R^{0.005}$	584	929	13	0.005	3.182	5.786
$G_R^{0.01}$	599	1778	7	0.010	5.937	3.796
G_{BA}^{2}	600	1182	7	0.007	3.940	3.986
G_{BA}^{4}	600	2345	5	0.013	7.817	3.033
G_C^2	600	1200	150	0.007	4.000	75.376
G_C^4	600	2400	39	0.013	8.000	19.720
G_H^4	600	1099	11	0.006	3.663	7.318
G_H^8	600	2126	7	0.012	7.087	4.768
G_{TS}	600	1228	14	0.007	4.093	7.060
$G_{WS}^{4,0.1}$	600	2400	8	0.013	8.000	4.683
G_{CA}	600	5401	9	0.030	18.003	4.283
$G_{PRU}^{50,12,2}$	600	1245	7	0.007	4.143	4.075
G_{MAIL}	1133	10 903	8	0.009	9.622	3.606
G_{BLOGS}	1224	19 025	–	0.022	27.312	2.738
G_{AIR}	1574	28 236	–	0.014	21.874	3.115

TABLE II
SUMMARY OF PARAMETERS IN SIMULATION

Parameters	Description
k_a	The number of the removed nodes in an attack phase
k_d	The number of the added nodes in a defense phase
w	The edge construction weight
$d(G)$	The average degree in a graph G
m	The number of the added edges per node in a defense phase: $\texttt{Round}(w \cdot d(G))$

When $k_a = k_d$ and $w = 1.0$, an attacker's damage ability is approximately equal to a defender's repair ability in terms of the number of connections.

be the average shortest path length among all pairs of vertices and the average degree, respectively. Network diameter is the maximum distance between nodes in the network [20]. Network density is a normalized version of the average number of neighbors, which indicates the overall level of interaction between all nodes in a network [21].

For our simulations, starting with an original graph G in Fig. 3, at each round: 1) we remove k_a nodes and their adjacent edges following the attack strategy and then 2) we add k_d new nodes such that each of the added nodes is connected with m new edges to m existing nodes. We set m to be the nearest integer rounded from $w \cdot d(G)$ where w is the edge construction weight. These parameters are summarized in Table II. The aim of the experiments is to evaluate feasibility and usefulness of each strategy and to find the optimal parameter values (e.g., w) of each strategy at the same time.

With fixed k_a, k_d, and m, we can observe how the size of the largest connected component and the average degree in a graph evolve. For example, with $k_a = k_d = 10$ and $w = 1.0$, Fig. 4 shows how these values in $G_R^{0.005}$ are changed under iterated attack and defense operations. The size of the largest connected component in each round is normalized by dividing by the size of the largest connected component in the original graph. From this figure, we can see that $\mathbf{D}^{balance}$ only performed well against $\mathbf{A}^{degree}$ or $\mathbf{A}^{central}$: the size of the largest connected component in the graph remained unchanged during the 100 rounds against these attacks while $\mathbf{D}^{random}$ and $\mathbf{D}^{prefer}$ are not effective—within 40 rounds the size of the largest connected component has fallen by a half. Interestingly, there exists a relationship between the size of the largest connected component and the average degree as we discussed in Section IV. The size of the largest connected component in $G_R^{0.005}$ started to drop dramatically when the average degree of the network falls below 2.

In this paper, our research interest is finding the best attack and defense strategies with varying k_a, k_d, and w rather than observing how network connectivity evolves in a particular network with fixed k_a, k_d, and w. Even if an attacker finds the ideal attack strategy for a given budget k_a, a defender may block this attack even with a naive defense strategy if she can increase k_d and/or w without limit. In this context, an attacker's goal should be interpreted to find an attack strategy maximizing the defense costs (k_d and/or w) while the defender wishes to find a defense strategy minimizing them.

With the edge construction weight w ranging from 0.5 to 1.5, we first discuss the effects of w for each of the fifteen graphs in Fig. 3. To demonstrate this we fix $k_a = k_d = 10$

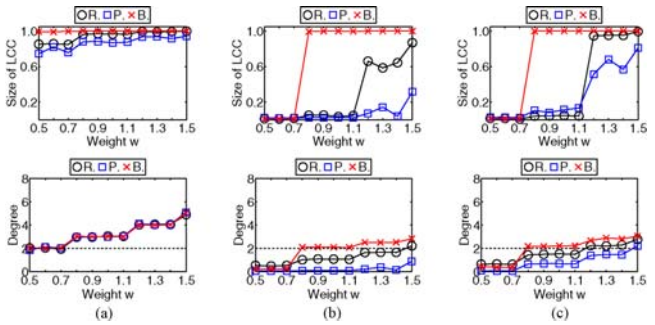

Fig. 5. Random graph 1 with w: changes in the size of the largest connected component in $G_R^{0.005}$ with w. In this figure, and in Figs. 6–16, the first row graphs show the changes in the size of the largest connected component and the second row graphs shows the changes in the average degree. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

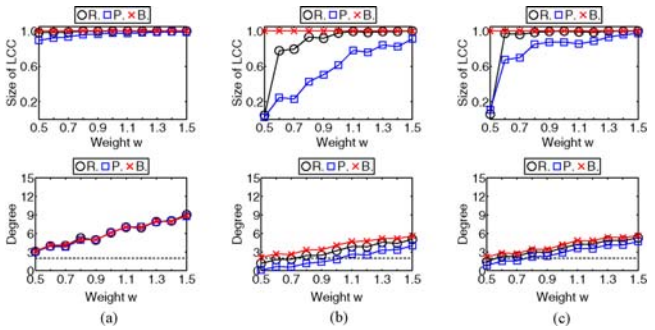

Fig. 6. Random graph 2 with w: changes in the size of the largest connected component in $G_R^{0.01}$ with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

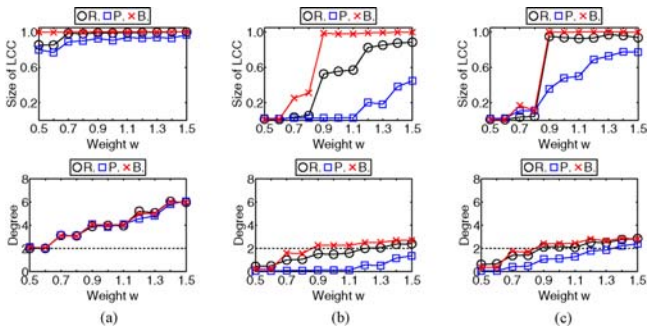

Fig. 7. Scale-free graph 1 with w: changes in the size of the largest connected component in G_{BA}^2 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

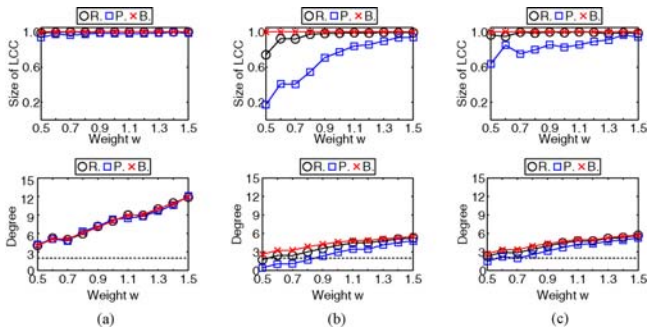

Fig. 8. Scale free graph 2 with w: changes in the size of the largest connected component in G_{BA}^4 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

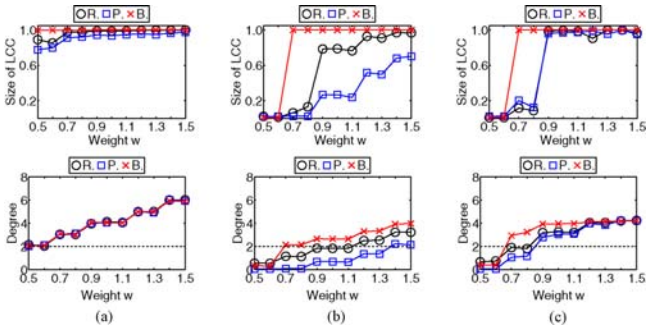

Fig. 9. Chord graph 1 with w: changes in the size of the largest connected component in G_C^2 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

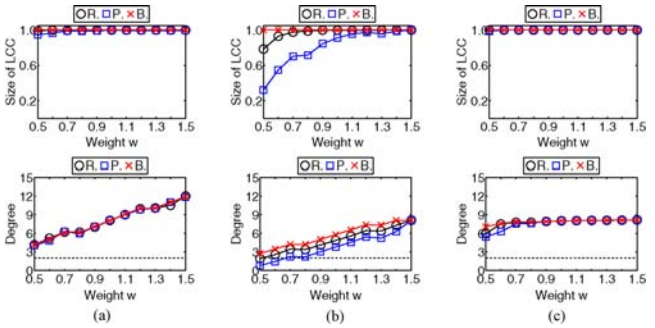

Fig. 10. Chord graph 2 with w: changes in the size of the largest connected component in G_C^4 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

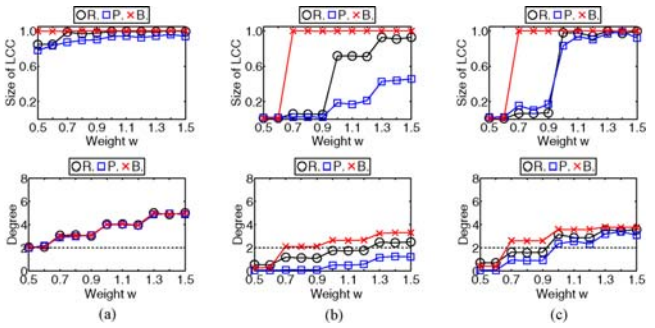

Fig. 11. Hypergraph 1 with w: changes in the size of the largest connected component in G_H^4 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

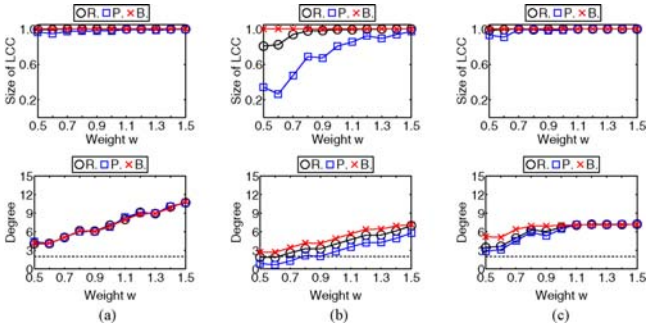

Fig. 12. Hypergraph 2 with w: changes in the size of the largest connected component in G_H^8 with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

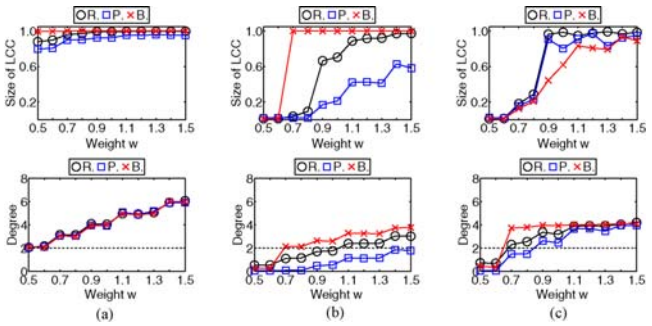

Fig. 13. Transit-stub graph with w: changes in the size of the largest connected component in G_{TS} with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

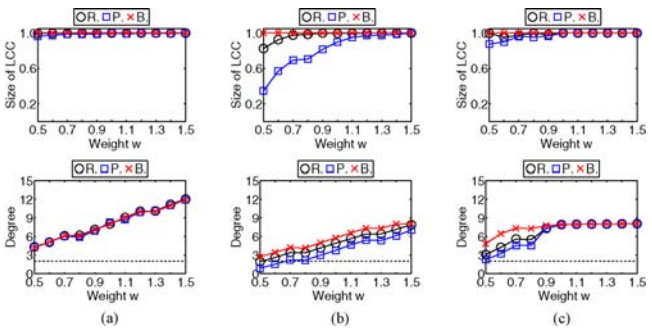

Fig. 14. Small-world graph with w: changes in the size of the largest connected component in $G_{WS}^{4,0.1}$ with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

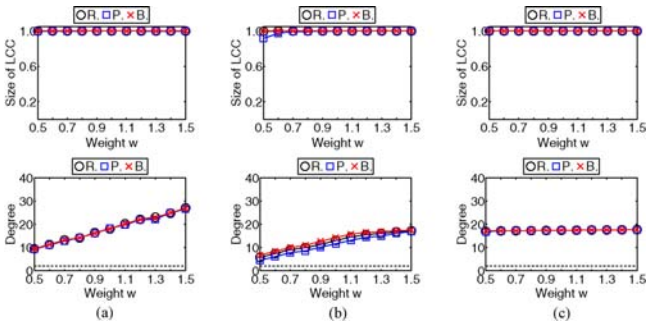

Fig. 15. CAN with w: changes in the size of the largest connected component in G_{CA} with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

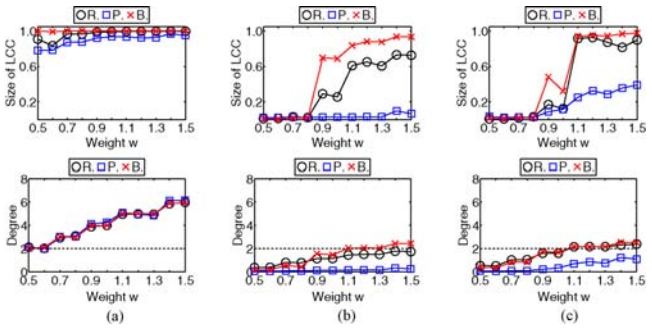

Fig. 16. PRU graph with w: changes in the size of the largest connected component in $G_{PRU}^{50,12,2}$ with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

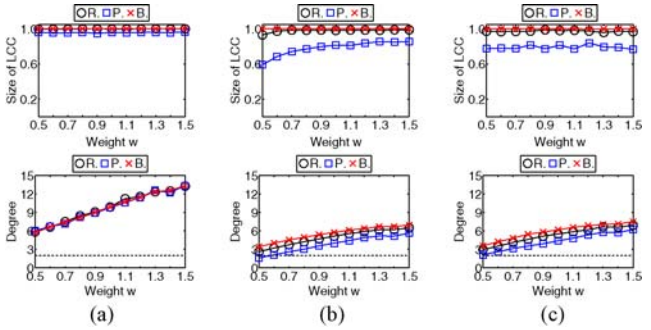

Fig. 17. Email network with w: changes in the size of the largest connected component in G_{MAIL} with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

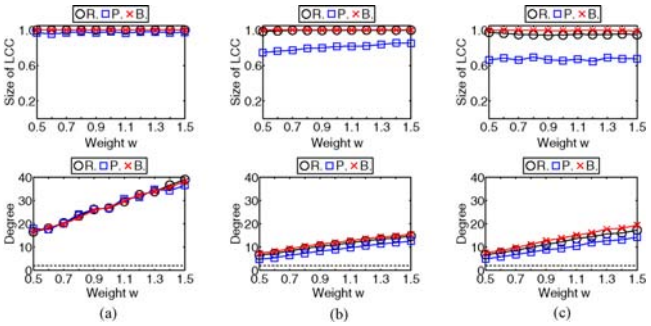

Fig. 18. Blog network with w: changes in the size of the largest connected component in G_{BLOG} with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

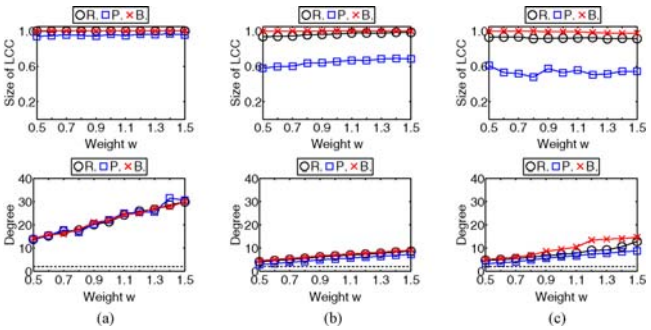

Fig. 19. Airport network with w: changes in the size of the largest connected component in G_{AIR} with w. (a) $\mathbf{A}^{random}$. (b) $\mathbf{A}^{degree}$. (c) $\mathbf{A}^{central}$.

and analyse the size of the largest connected component and the average degree in each graph after the 100th round. The experimental results for each network are shown, respectively, in Figs. 5–19. The results for best attack and defense strategies with w are summarized in Table III.

From these figures, we can see $\mathbf{D}^{balance}$ performed well except for the case against $\mathbf{A}^{central}$ in G_{TS} [see Fig. 13(c)]. When a defender uses $\mathbf{D}^{balance}$ even with a small $w \leq 1.0$, most nodes in all networks except for G_{TS} and $G_{PRU}^{50,12,2}$ remain connected to each other. However, $\mathbf{D}^{random}$ and $\mathbf{D}^{prefer}$ are not sufficiently effective against $\mathbf{A}^{degree}$ or $\mathbf{A}^{central}$ in many network topologies. For example, when $w = 1.0$, the graph $G_R^{0.005}$ is totally disconnected if $\mathbf{D}^{random}$ or $\mathbf{D}^{prefer}$ is used against either $\mathbf{A}^{degree}$ or $\mathbf{A}^{central}$ [see Fig. 5(b) and (c)]. In fact, even when $w = 1.5$, there are not enough edges to defend against $\mathbf{A}^{degree}$ [see Fig. 5(b)]. In particular, $\mathbf{D}^{prefer}$ performed badly

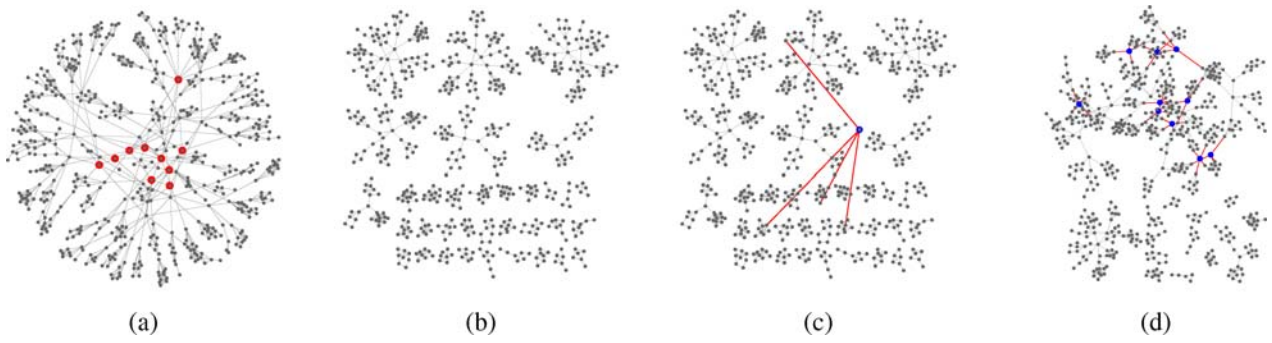

Fig. 20. Performing $\mathbf{A}^{central}$ and $\mathbf{D}^{balance}$ on G_{TS} in the first round in the transit-stub graph. (a) Ten nodes (large circle) are selected to be removed for $\mathbf{A}^{central}$. (b) Graph is totally disconnected after the first attack. (c) New node and its connections are added by $\mathbf{D}^{balance}$. This node connects different clusters. (d) Ten nodes and their connections are newly created after the first defense. However, many nodes are still disconnected from the largest connected component (the above subnetwork) in the graph.

against $\mathbf{A}^{degree}$. Since most real networks exhibit preferential connectivity [6], a real-world network may be very vulnerable to high-degree node attacks even if the network has grown continuously up with new nodes over time. Our results on real networks supported this conjecture [see Figs. 17(b), 18(b), and 19(b)].

On the other hand, $\mathbf{A}^{random}$ is not effective from the attacker's point of view; the size of the largest connected component remained unchanged and the average degree is still greater than 2 after 100 rounds if any defense strategy is used with $w = 1.0$. This is natural enough; the damage done by random failures is not essentially greater than the level of repair by random replenishment. So network connectivity will be maintained well on most popular network topologies under random node failure or removal if the same number of new nodes can be continuously recruited.

To maintain network connectivity, a possible approach is to increase network density. In our experiments, the half of network topologies with a high network density ≥ 0.009 ($G_R^{0.01}$, G_{BA}^4, G_C^4, G_H^8, $G_{WS}^{4,0.1}$, G_{CA}, G_{MAIL}, G_{BLOG}, and G_{AIR}—see Table I) is resilient against any attack strategies when $\mathbf{D}^{balance}$ is used with $w = 0.5$ only (see Figs. 6, 8, 10, 12, 14 and 15). In fact, $\mathbf{D}^{random}$ is also adequate in these networks except G_{BLOG} and G_{AIR} if a defender can increase w to 1.0. For the two real networks, G_{BLOG} and G_{AIR}, $\mathbf{D}^{random}$ is not perfect against $\mathbf{A}^{central}$ even with a large $w = 1.5$.

Interestingly, $\mathbf{D}^{balance}$ is worse than the other defense strategies against $\mathbf{A}^{central}$ in G_{TS} [see Fig. 13(c)]. This is because a small number of nodes connect different clusters in G_{TS}. Since these nodes have high betweenness centrality, transit-stub graphs are inherently very vulnerable to $\mathbf{A}^{central}$. Unfortunately, $\mathbf{D}^{balance}$ cannot change this weakness of transit-stub graphs since nodes newly recruited in the defense phase generally play a role as new gateway nodes by connecting separated clusters. This trend can be observed in Fig. 20.

Another interesting observation is the relationship between the size of the largest connected component and the average degree in a graph. As we discussed in Section IV, in all experiments, we can see that the size of the largest connected component is not maintained well when the average degree of the network falls below 2, regardless of defense strategy, but the opposite is not true [see the counter example in Fig. 13(c)].

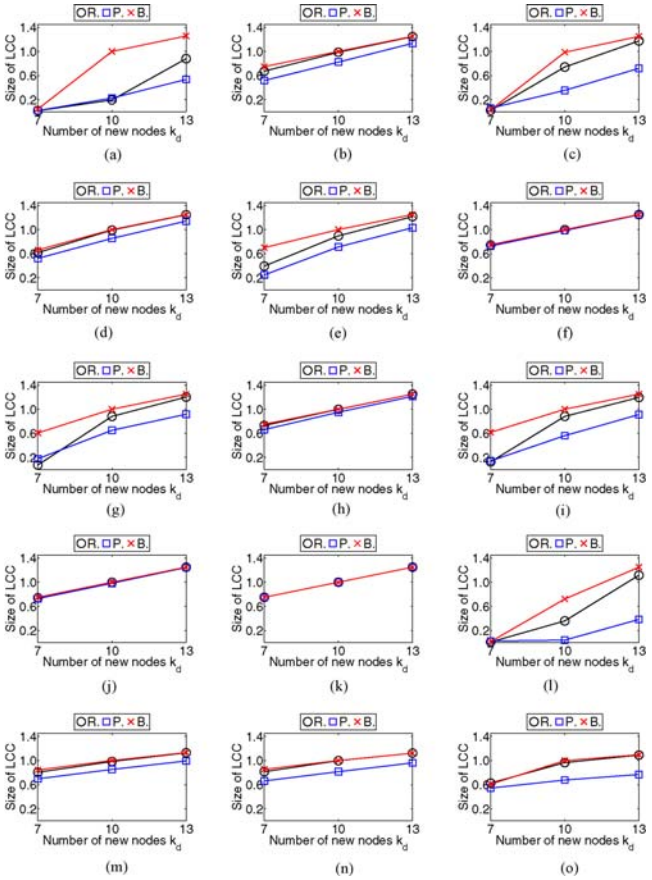

Fig. 21. Size of the largest connected component with k_d against high-degree node attacks. (a) $G_R^{0.005}$. (b) $G_R^{0.01}$. (c) G_{BA}^2. (d) G_{BA}^4. (e) G_C^2. (f) G_C^4. (g) G_H^4. (h) G_H^8. (i) G_{TS}. (j) $G_{WS}^{4,0.1}$. (k) G_{CA}. (l) $G_{PRU}^{50,12,2}$. (m) G_{MAIL}. (n) G_{BLOG}. (o) G_{AIR}.

Finally, we discuss how the performance of attack and defense strategies may change when $k_a \neq k_d$. As k_d increases, network connectivity between nodes will increase over rounds. Figs. 21 (against high-degree node attacks) and 22 (against betweenness centrality attacks) show the effects of varying k_d from 7 to 13 with $k_a = 10$ and $w = 1.0$. To demonstrate this we plot the size of the largest connected component in a graph at the 50th round.

TABLE III
BEST ATTACK AND DEFENSE STRATEGIES FOR NETWORKS

	Best attack	Best defense	w	Damage
$G_R^{0.005}$	$\mathbf{A}^{degree}$, $\mathbf{A}^{central}$	$\mathbf{D}^{balance}$	0.8	None
$G_R^{0.01}$	–	$\mathbf{D}^{balance}$	0.5	None
G_{BA}^2	$\mathbf{A}^{degree}$, $\mathbf{A}^{central}$	$\mathbf{D}^{balance}$	0.9	None
G_{BA}^4	–	$\mathbf{D}^{balance}$	0.5	None
G_C^2	$\mathbf{A}^{degree}$, $\mathbf{A}^{central}$	$\mathbf{D}^{balance}$	0.7	None
G_C^4	–	$\mathbf{D}^{balance}$	0.5	None
G_H^4	$\mathbf{A}^{degree}$, $\mathbf{A}^{central}$	$\mathbf{D}^{balance}$	0.7	None
G_H^8	–	$\mathbf{D}^{balance}$	0.5	None
G_{TS}	$\mathbf{A}^{central}$	$\mathbf{D}^{random}$	0.9	Small
$G_{WS}^{4,0.1}$	–	$\mathbf{D}^{balance}$	0.5	None
G_{CA}	–	$\mathbf{D}^{random}$, $\mathbf{D}^{balance}$	0.5	None
$G_{PRU}^{50,12,2}$	$\mathbf{A}^{degree}$	$\mathbf{D}^{balance}$	1.5	Small
G_{MAIL}	–	$\mathbf{D}^{balance}$	0.5	None
G_{BLOG}	–	$\mathbf{D}^{balance}$	0.5	None
G_{AIR}	–	$\mathbf{D}^{balance}$	0.5	None

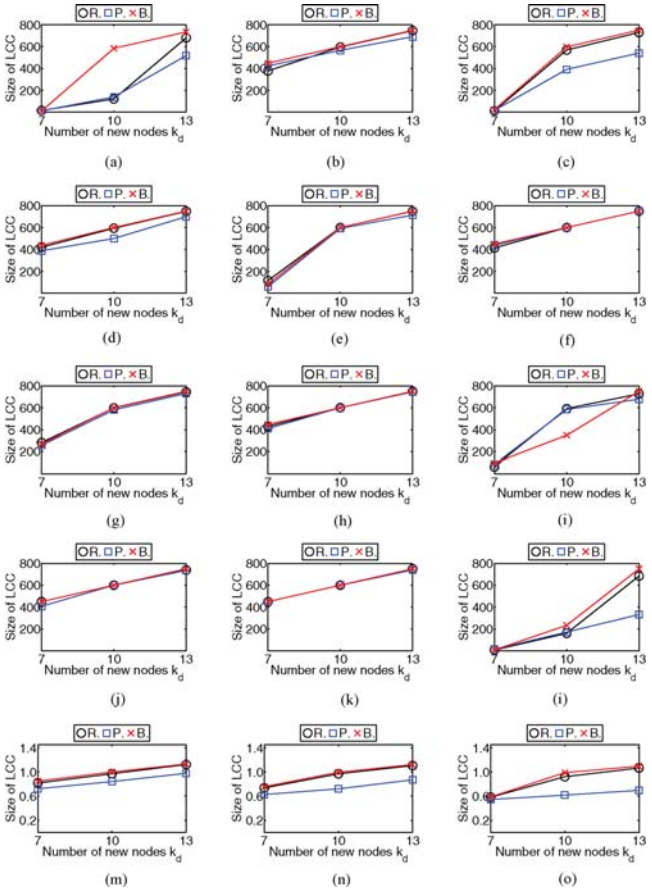

Fig. 22. Size of the largest connected component with k_d against betweenness centrality attacks. (a) $G_R^{0.005}$. (b) $G_R^{0.01}$. (c) G_{BA}^2. (d) G_{BA}^4. (e) G_C^2. (f) G_C^4. (g) G_H^4. (h) G_H^8. (i) G_{TS}. (j) $G_{WS}^{4,0.1}$. (k) G_{CA}. (l) $G_{PRU}^{50,12,2}$. (m) G_{MAIL}. (n) G_{BLOG}. (o) G_{AIR}.

As k_d increases, so does the size of the largest connected component. This is natural enough, and is particularly clear in low density networks ($G_R^{0.005}$, G_{BA}^2, G_C^2, G_H^4, G_{TS}, and $G_{PRU}^{50,12,2}$).

The performance of $\mathbf{D}^{balance}$ is still better than those of the other two defense strategies and is highly scalable in terms of k_d: the gap between them is clearly shown in low density networks when $k_d > k_a$ and in some networks ($G_R^{0.01}$, G_{BA}^4, G_C^2, G_H^4, and G_{TS}) when $k_d < k_a$. Interestingly, $\mathbf{D}^{balance}$ produced the best result in G_{TS} against $\mathbf{A}^{central}$ when $k_d > k_a$; it is different from the case when $k_d = k_a$.

VI. CONCLUSION

Barabási, Albert, and Jeong showed that, while small-world networks were resilient against random node failure, they were very vulnerable to targeted attacks. Nagaraja and Anderson extended this single-shot analysis to the dynamic case so that attack and defense strategies could be measured against each other. We extended their work to a wide range of network topologies, including some real-world networks, and to account for the costs of replacing edges as well as nodes. In summary, we have the following.

1) The best defense strategy in general is balanced replenishment, $\mathbf{D}^{balance}$. For high density networks with network density ≥ 0.009, it is enough to set $w = 0.5$.
2) The best attack strategy in general targets vertex order or betweenness centrality, that is $\mathbf{A}^{degree}$ or $\mathbf{A}^{central}$, in the sense that it maximizes the cost of defense. However, when the network has a hierarchical tree-like structure, it will often be better to use $\mathbf{A}^{central}$.
3) It is necessary but not sufficient for the defender to maintain the average node degree ≥ 2 to maintain connectivity.
4) A real-world network may be very vulnerable to $\mathbf{A}^{degree}$ or $\mathbf{A}^{central}$ even if the network has grown continuously up with new nodes and connections over time.

In future work, we plan to develop better models of the adversary. We may consider not only an adversary with global knowledge of network topology but also a weaker adversary with limited information (e.g., a local police force). For example, we expect that $\mathbf{D}^{random}$ is secure against any adversary with no knowledge of the network topology at all; what strategies suffice against an attacker whose knowledge is is local? Also, while adding more edges to a network may be a viable strategy for a disease pathogen, it may not help an insurgent group as a better-connected network may be more vulnerable to insider threats.

As an extension to this paper, we plan to consider a theoretical study to formally generalize and verify our results. We will also employ more advanced centrality metrics such as bridging centrality [22] to improve the performance of strategies.

ACKNOWLEDGMENT

The contents of this paper do not necessarily express the views of Northrop Grumman Corporation, Falls Church, VA. The authors also acknowledge support from the FKPPL International Associated Laboratory, Korea, for grid computing resources.

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, nos. 4–5, pp. 175–308, Feb. 2006.

[2] M. E. J. Newman, A. L. Barabási, and D. J. Watts, Eds., *The Structure and Dynamics of Networks*. Princeton, NJ: Princeton Univ. Press, 2010.
[3] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.
[4] S. Nagaraja and R. Anderson, "The topology of covert conflict," in *Proc. 5th WEIS)*, Jul 2006.
[5] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.
[6] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
[7] I. Stoica, R. Morris, D. Liben-Nowell, D. R. Karger, M. F. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup protocol for internet applications," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 17–32, Feb. 2003.
[8] F. Saffre and R. Ghanea-Hercock, "Beyond anarchy: Self-organized topology for peer to peer networks," *Complexity*, vol. 9, no. 2, pp. 49–53, 2003.
[9] K. Calvert, M. Doar, and E. Zegura, "Modeling internet topology," *IEEE Commun. Mag.*, vol. 35, no. 6, pp. 160–163, Jun. 1997.
[10] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton Studies in Complexity), illustrated edition. Princeton, NJ: Princeton Univ. Press, Nov. 2003.
[11] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Schenker, "A scalable content-addressable network," in *Proc. Conf. Appl. Technol. Arch. Protocols Comput. Commun.*, 2001, pp. 161–172.
[12] G. Pandurangan, P. Raghavan, and E. Upfal, "Building low-diameter peer-to-peer networks," *IEEE J. Select. Areas Commun.*, vol. 21, no. 6, pp. 995–1002, Aug. 2003.
[13] R. Guimerà, L. Danon, D. A. Guilera, F. Giralt, and A. Arenas, "Self-similar community structure in a network of human interactions," *Phys. Rev. E*, vol. 68, no. 6, p. 065103, Dec. 2003.
[14] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 U.S. election: Divided they blog," in *Proc. 3rd Int. Workshop Link Discovery*, 2005, pp. 36–43.
[15] T. Opsahl. (2011). *Why Anchorage Is Not (That) Important: Binary Ties and Sample Selection* [Online]. Available: http://toreopsahl.com/2011/08/12/why-anchorage-is-not-that-important-binary-ties-and-sample-selection
[16] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, "Attack vulnerability of complex networks," *Phys. Rev. E*, vol. 65, no. 5, p. 056109, May 2002.
[17] L. Zhao, K. Park, and Y.-C. Lai, "Attack vulnerability of scale-free networks due to cascading breakdown," *Phys. Rev. E*, vol. 70, no. 3, p. 035101, Sep. 2004.
[18] J. Domingo-Ferrer and Úrsula González-Nicolás, "Decapitation of networks with and without weights and direction: The economics of iterated attack and defense," *Comput. Netw.*, vol. 55, no. 1, pp. 119–130, 2011.
[19] U. Brandes, "A faster algorithm for betweenness centrality," *J. Math. Sociol.*, vol. 25, no. 2, pp. 163–177, 2001.
[20] H. Per and H. Frank, "Eccentricity and centrality in networks," *Social Netw.*, vol. 17, no. 1, pp. 57–63, 1995.
[21] J. Dong and S. Horvath, "Understanding network concepts in modules," *BMC Syst. Biol.*, vol. 1, no. 1, p. 24, 2007.
[22] W. Hwang, T. Kim, M. Ramanathan, and A. Zhang, "Bridging centrality: Graph mining from element level to group level," in *Proc. 14th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, 2008, pp. 336–344.

Hyoungshick Kim received the B.S. degree from the Department of Information Engineering, Sungkyunkwan University, Seoul, Korea, the M.S. degree from the Department of Computer Science, KAIST, Daejeon, Korea, and the Ph.D. degree from the Computer Laboratory, University of Cambridge, Cambridge, U.K., in 1999, 2001, and 2012, respectively.

He is currently a Post-Doctoral Fellow with the University of British Columbia, Vancouver, BC, Canada. He was with Samsung Electronics as a Senior Engineer from May 2004 to September 2008. He has also served as a member of DLNA and Coral standardization for DRM interoperability. His current research interests include social computing and usable security.

Ross Anderson is currently a Professor of security engineering with the Computer Laboratory, University of Cambridge, Cambridge, U.K. He is one of the founders of a vigorously growing new academic discipline, the economics of information security. He has also been a seminal contributor to peer-to-peer systems, hardware tamper resistance, emission security, copyright marking, and the robustness of application programming interfaces. He authored the standard textbook, *Security Engineering—A Guide to Building Dependable Distributed Systems*.

Dr. Anderson is a fellow of the Royal Society, the Royal Academy of Engineering, the IET, the Institute of Physics, and the IMA.