# Revisiting Botnet Models and Their Implications for Takedown Strategies

Ting-Fang Yen[1] and Michael K. Reiter[2]

[1] RSA Laboratories, Cambridge, MA
tingfang.yen@rsa.com
[2] University of North Carolina, Chapel Hill, NC
reiter@cs.unc.edu

**Abstract.** Several works have utilized network models to study peer-to-peer botnets, particularly in evaluating the effectiveness of strategies aimed at taking down a botnet. We observe that previous works fail to consider an important structural characteristic of networks — assortativity. This property quantifies the tendency for "similar" nodes to connect to each other, where the notion of "similarity" is examined in terms of node degree. Empirical measurements on networks simulated according to the Waledac botnet protocol, and on network traces of bots from a honeynet running in the wild, suggest that real-world botnets can be significantly assortative, even more so than social networks. By adjusting the level of assortativity in simulated networks, we show that high assortativity allows networks to be more resilient to takedown strategies than predicted by previous works, and can allow a network to "heal" itself effectively after a fraction of its nodes are removed. We also identify alternative takedown strategies that are more effective, and more difficult for the network to recover from, than those explored in previous works.

## 1 Introduction

Graph models from network theory have been applied to study properties of real-world networks, including social, biological, and computer networks. Erdös-Rényi random graphs [13] model networks where the edges are created with uniform probability between every pair of nodes. Watts-Strogatz small-world graphs [38] model networks where the diameter of the network is small, i.e., increasing logarithmically with the size of the network. Barabási-Albert scale-free graphs [2] model networks with a few highly connected "hub" nodes and many leaf nodes. These models can be used to analyze the spread of information (or infection) within a network [30,38] and its resilience to node and edge failures [1,9,15], for example.

Recently, several works have also applied graph models from network theory to study peer-to-peer (P2P) botnets [10,11,41,19]. Each node in the network represents an infected host, and edges reflect communications between the hosts. Properties of the graph can quantify the botnet's "usefulness". For instance, the diameter of the network measures the efficiency of bot communications, and

P. Degano and J.D. Guttman (Eds.): POST 2012, LNCS 7215, pp. 249–268, 2012.

the size of the largest connected component is the number of bots that are reachable by the attacker and can carry out her instructions. Assuming that P2P botnets are structured according to known models, these works aim to assess the effectiveness of strategies to take down a botnet, i.e., decreasing the botnet's "usefulness". For example, one strategy that was found to be effective for some network topologies is to target nodes with high degree, i.e., that communicate with many hosts [10,11,41].

We observe that previous works applying graph models to P2P botnets do not consider an important property of networks — assortative mixing [25]. Assortativity refers to the tendency for a node to attach to other "similar" nodes, and is commonly examined in terms of a node's degree, i.e., high-degree nodes are likely to be neighbors of other high-degree nodes. This property is also referred to as *degree correlation*. The existence of this correlation between neighboring nodes has been observed in many real-world networks [29,27,25]. More importantly, it has been found to be a property of *growing* networks [5,18], where the network increases in size as nodes join over time, as is true in a botnet as more hosts become infected.

In this work, we show that assortativity plays an important role in network structure, such that neglecting it can lead to an over-estimation of the effectiveness of botnet takedown strategies. By generating networks with varying levels of degree correlation, we demonstrate that a higher level of assortativity allows the network to be more resilient to certain takedown strategies, including those found to be effective by previous works. Moreover, we note that bots are dynamic entities that can react and adapt to changes in the network, and so the botnet can potentially "heal" itself after a fraction of its nodes are removed. We specifically explore cases where nodes can compensate for lost neighbors by creating edges to other nearby nodes, e.g., that are within $h$ hops. This is similar to the behavior of a P2P bot contacting known hosts on its peer-list, which the bot maintains by constant exchanges with its neighbors [4,6,31,16,36]. Our simulations show that the graph can recover significantly after takedown attempts, even when $h$ is small, and that higher levels of assortativity can allow the network to recover more effectively.

Another contribution of this work is in identifying alternative takedown strategies that are more effective than those explored by previous works. Specifically, we show that targeting nodes with both high degree and low clustering coefficient will decrease the connectivity and communication efficiency of the network significantly, and also makes it considerably more difficult for the network to recover from the takedown attempt. We further examine the effectiveness of applying this strategy "locally" where only a subset of nodes and edges is visible, such as when traffic from only a single subnet can be observed.

The rest of this paper is organized as follows. Section 2 describes related work. Section 3 defines assortativity, studies this value in botnets, and describes our algorithm for generating networks with varying levels of assortativity. The effect of assortativity on network resilience and "healing" ability is investigated in Sections 4 and 5. Discussion and conclusions are presented in Sections 6 and 7.

## 2 Related Work

***Botnet models.*** Several previous works have studied botnets using network models. Cooke et al. [8] described three potential botnet topologies: centralized, P2P, and random, and qualitatively discussed their design complexity, detectability, message latency, and survivability. Other works [19,10] applied theoretical network models to botnets, including Erdös-Rényi random graphs [13], Watts-Strogatz small world graphs [38], and Barabási-Albert scale-free graphs [2]. This allows the effectiveness of takedown strategies to be quantitatively evaluated using graph properties, such as the network diameter, the average shortest distance between pairs of nodes, and the size of the largest connected component. Davis et al. [11] compared Overnet, which is utilized by the Storm botnet [31,16], with random and scale-free networks to justify the choice of structured P2P networks made by bot-masters. They simulated takedown efforts on the networks by removing nodes at random, in descending order of node degree, or in a "tree-like" fashion by identifying nodes reachable from an initial node, and found Overnet to be more resilient than other graph models.

To our knowledge, no previous work on botnet modeling has considered the effect of *degree assortativity* in networks. This property, defined as the correlation coefficient between the degrees of neighboring nodes [25], has been found to be high in many real-world social, biological, and computer networks [26,29]. It has been studied analytically in the statistical physics literature, and found to be an inherent property of *growing* networks where nodes join and edges are created over time [5,18], since older nodes are likely to have higher degree and tend to connect to each other. Studies in the statistical physics domain focus on understanding the underlying interactions between nodes that would result in a network that matches one empirically measured in the real world. By contrast, a network of bots is elusive and difficult to quantify in practice. Making assumptions about the graph structure or node correlation (e.g., that there is none) is thus unfounded.

***Network takedown strategies.*** The resilience of networks to attacks or failures have been explored in the physics branch of complex networks [1,15,9]. A scale-free network, which consists of a few highly-connected "hub" nodes and many "leaf" nodes, has been found to be particularly vulnerable to attacks where high-degree nodes are removed first. A takedown strategy that targets high-degree nodes is also recommended by previous works that studied botnet models [10,11,41], particularly for unstructured P2P networks where there are "super-peers" present.

Other types of takedown efforts on networks have also been explored in the complex networks literature, such as cascaded node removals [37], removing nodes according to their betweenness centrality, or removing edges instead of nodes [15]. These works focus on the resilience of different network topologies, and do not take assortativity into account. Newman et al. [26] studied the prevalence of assortativity in real-world networks. Even though their focus is on measuring and generating assortative networks, they also showed, through

simulation, that higher assortativity allows a network to have a larger connected component after a small fraction of high-degree nodes are removed. However, they did not explore other takedown strategies, the effect on other graph properties, or the network's ability to "heal" itself. In this work, we explicitly study the effect of assortativity on network resilience and the ability of dynamic networks (such as P2P botnets) to recover from takedown attempts.

## 3 Constructing and Measuring Assortative Networks

We first define degree assortativity, following the definition by Newman et al. [25], and perform empirical analyses on the assortativity of real botnets by simulating networks according to the Waledac botnet protocol [6] and examining a portion of the Storm botnet [31,16,36]. We then describe our algorithm for adjusting the level of assortativity in simulated networks, and the metrics we use to quantify the "usefulness" of a network. The metrics are aimed at capturing notions of communication efficiency between nodes and the number of reachable bots, which are likely to be of importance to the bot-master.

### 3.1 Degree Assortativity

Degree assortativity, defined as the correlation coefficient between the degrees of neighboring nodes, measures the tendency for nodes to be connected to others who are "similar" in terms of their degree. For example, this property is especially significant in social networks, where gregarious people are likely to be friends with each other [27,17]. It is also found to be a property of growing networks, where the network size increases as new nodes join and edges are created [5,18], as is true for botnets as vulnerable nodes become infected.

We define assortativity following the definition of Newman et al. [25]. Let the fraction of nodes in a network graph with degree $k$ be denoted $p_k$. If we choose an edge from the graph at random, and follow it to one of its ends, the probability that the node at which we arrive has a degree of $k$ is proportional to $k$. This is because we are more likely to end up at a node with high degree, which has more edges connected to it. To account for the edge from which we arrived, the *remaining degree* of the node is its degree minus one. The probability $q_k$ that we arrive at a node with remaining degree $k$ is then

$$q_k = \frac{(k+1)p_{k+1}}{\sum_{j=0}^{\infty} jp_j} \tag{1}$$

Let $e_{j,k}$ be the probability that a randomly selected edge connects nodes of remaining degree $j$ and $k$, where $\sum_{j,k} e_{j,k} = 1$. The assortativity $\gamma$ of the network, being the correlation coefficient between the degrees of neighboring nodes, is

$$\gamma = \frac{1}{\sigma_q^2} \sum_{j,k} jk(e_{j,k} - q_j q_k) \tag{2}$$

where $\sigma_q^2$ is the variance of the distribution of $q_k$, i.e., $\sigma_q^2 = \sum_k k^2 q_k - [\sum_k k q_k]^2$. A higher value of $\gamma$ indicates that there is higher correlation between the degrees of two neighboring nodes. In a random graph, where every pair of nodes is connected with uniform probability, no correlation exists and $\gamma = 0$.

### 3.2 Degree Assortativity in Botnets

Even though high assortativity is found in many real-world networks, measuring it in practice can be challenging due to difficulties in observing all interactions between nodes in a large network. This is especially true for P2P botnets, since infected hosts cannot always be identified, and obtaining a comprehensive view of those hosts' communications may require multiple administrative entities to share sensitive information. While researchers have studied P2P botnets via infiltration (e.g., [16]), this provides a limited view of only a subset of the botnet.

We expect that real P2P botnets are likely to be assortative. This is not only because assortativity has been found to be a property of growing networks that increase in size over time (e.g., when vulnerable hosts become infected and join the botnet), but also due to the constant peer-list exchanges that occur between neighboring bots, which makes the "edges" in a botnet far from being random.

We perform two experiments to estimate the level of assortativity in P2P botnets. In the first, we simulate networks where nodes create and delete edges according to the algorithm performed by Waledac bots, as described in previous work that reverse-engineered the Waledac bot binary [4,6]. In the second, we examine network traffic from Storm bots in a honeynet running in the wild.

***Waledac botnet simulations.*** Waledac is a P2P botnet that communicates over the HTTP protocol [4,6,35]. Similar to other P2P bots, each bot maintains a fixed-length list of known peers with which it communicates in order to stay connected to the botnet (and hence to the bot-master). A bot periodically exchanges peer-lists with other peers known to it, i.e., by randomly selecting hosts from its peer-list. This allows the bot to learn about other hosts in the botnet and to remove inactive nodes from its peer-list. As documented by Calvet et al. [6], the Waledac binary comes with a hard-coded list of 200 boot-strapping hosts. As the bot learns about other existing peers, its peer-list grows to a maximum of 500 entries, where each entry includes the IP address of the peer, as well as the time at which activities from that peer was last observed. If the number of known peers exceeds 500, the bot only keeps track of 500 most recently active hosts. During each peer-list exchange, each bot extracts 99 entries from its peer-list, appends its own IP address and the current time to this shortened list, and sends it to a host selected at random from its peer-list. In return, the receiving host also responds with a list of hosts extracted from its own peer-list.

We simulated networks where nodes join and depart over time (e.g., due to hosts becoming infected or patched), creating or deleting edges between each other following the Waledac protocol as described above. Assuming a constant rate of nodes joining the network in each round, we drew each node's lifetime from an exponential distribution [28,21], after which the node was removed from

the network. Each simulated network was allowed to evolve this way until the number of online nodes reached 5,000. This number represents a small botnet, and follows the simulation settings in previous work on modeling botnets [10].

From this experiment, we found the assortativity of such networks to be quite high. Over a total of 50 simulation runs, the average assortativity was 0.39 (with a standard deviation of 0.036), which is higher than that of social networks [26]. This suggests that a botnet may be significantly assortative, and highlights the importance of this property in considering botnet models.

***Traffic from Storm bots in a honeynet.*** In addition to our simulations, we also obtained network traffic from a honeynet running in the wild in late 2007 [14]. This dataset consists of a consecutive 24-hour trace from 13 hosts participating in the Storm botnet [31,16,36].

Figure 1 shows the assortativity measured among the 13 Storm bots, where snapshots of their communications were taken on an hourly basis. The "degree" of a bot is represented by 1) the number of distinct source IP addresses from which it receives packets (the in-degree), 2) the number of distinct destination IPs to which it sends packets (the out-degree), or 3) the total number of distinct IPs with which it interacts. Since the rest of the Storm botnet is not directly observable, we calculated the assortativity of the sub-graph that consisted of the 13 Storm bots, i.e., by considering traffic between only the 13 Storm bots. As shown in Figure 1, this value is quite high, ranging from 0.48 to 0.84.

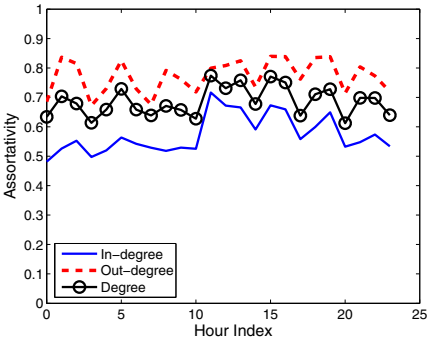



**Fig. 1.** The assortativity for 13 Storm bots in a honeynet running in the wild

That said, we acknowledge that this limited dataset may not be representative of the actual Storm botnet. For example, the high level of assortativity may be due to certain aspects of the honeynet setup; e.g., the observable bots were placed in the same local network and so may have been more likely to communicate with each other. (Such localized measurements may be all that is available in practice to a network administrator who can observe traffic from only a single network. We will discuss the effectiveness of botnet takedown strategies using only local information in Section 6.)

While we recognize the limitations of the above efforts to evaluate assortativity in today's botnets, the results of our analysis in Sections 4 and 5 suggest that a botnet designer would want his botnet to be assortative for added resilience and recoverability, further buttressing our belief that future botnets will leverage this naturally occurring property.

### 3.3 Generating Assortative Networks

To study the effect of assortativity on networks, we need to be able to generate networks with varying levels of assortativity. One method for this is to rewire edges in a given network [40]: At each step, select two edges at random, and shuffle them so that the two nodes with larger remaining degrees are connected, and the two nodes with smaller remaining degrees are connected. Repeating this step will result in the network becoming increasingly assortative. However, rewiring causes the shortest path length between nodes to increase rapidly [40], which may bias the comparison between networks with different levels of assortativity.

We apply another method for constructing assortative networks, similar to Newman et al. [26]. This method takes as input the number of nodes in the network, the desired degree distribution $p_k$, and the edge probabilities $e_{j,k}$. Each node in the network is assigned a degree drawn from $p_k$. The remaining degree distribution $q_k$ can then be calculated from $p_k$, and edges are added by connecting each pair of nodes of remaining degrees $j$ and $k$ with probability $e_{j,k}$.

To control the level of assortativity in the resulting network, we specify $e_{j,k}$ as follows. For a fixed value $j$, assume that $e_{j,k}$ follows a normal distribution centered at $j$, where the standard deviation $\sigma$ is the adjustable knob for tuning the level of assortativity. Figure 2 illustrates $e_{j,k}$ centered at $j$. A smaller $\sigma$ causes the normal distribution to become more peaked, where nodes with remaining degree $j$ have a higher probability of sharing edges with other nodes of remaining degree close to $j$, resulting in a more assortative network.

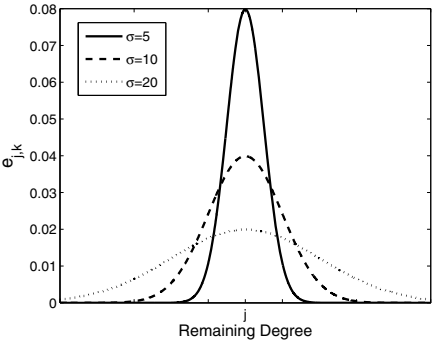



**Fig. 2.** Edge probabilities $e_{j,k}$ as a normal distribution centered at $j$ with different values for the standard deviation $\sigma$

In our simulations, $p_k$ is chosen so that the resulting network is scale-free, specifically, $p_k \sim k^{-3}$. We focus on scale-free networks because it is representative of many real-world networks, including unstructured P2P networks [22]. Empirical analysis by Dagon et al. [10] also suggest that the Nugache P2P botnet [36] has a scale-free structure. We set the number of nodes to 5,000 to represent a small botnet, following the simulation settings in previous work [10]. All of the edges are assumed to be undirected.

### 3.4 Metrics

We utilize the following two graph properties to quantify the "usefulness" of a botnet: 1) the size of the largest connected component, and 2) the inverse geodesic length. These metrics have been used by Dagon et al. [10] to compare the utility of different botnet topologies, and were also used in analyzing the resilience of various networks in the physics literature [15].

The fraction $S$ of nodes in the largest connected component is an upper bound on the number of bots that are directly under the control of the attacker (assuming that she is part of one of the connected components). The more hosts that can carry out the attacker's commands, the larger the scale of the attack that can be launched, e.g., denial-of-service attacks or spamming.

In addition to controlling many infected hosts, another property that is likely to be of importance to the attacker is the efficiency of communication, i.e., how long it takes for messages to be relayed through the botnet. We measure the number of hops between pairs of nodes for this purpose. Specifically, let $N$ be the total number of nodes, $V$ be the set of nodes, $|V| = N$, and $d(u, v)$ be the length of the shortest path between node $u$ and node $v$. The average inverse geodesic length [15] is defined as

$$L^{-1} = \frac{1}{N(N-1)} \sum_{u \in V} \sum_{v \neq u, v \in V} \frac{1}{d(u,v)} \tag{3}$$

Measuring the average inverse geodesic length is particularly useful in cases where the graph may be disconnected, since the distance $d(u, v)$ between two nodes $u$ and $v$ that belong to separate connected components would be infinite (and so its contribution to $L^{-1}$ is zero). The larger $L^{-1}$ is, the shorter the distances between nodes, and hence more efficient their communication. In evaluating the effectiveness of network takedown attempts, we are more interested in measuring the *normalized* average inverse geodesic length, which is defined as

$$\hat{L}^{-1} = \frac{\sum_{u \in V} \sum_{v \neq u, v \in V} \frac{1}{d'(u,v)}}{\sum_{u \in V} \sum_{v \neq u, v \in V} \frac{1}{d(u,v)}} \tag{4}$$

where $d'(u, v)$ is the *modified* length of the shortest path between nodes $u$ and $v$, that is, after takedown efforts or after the network tries to heal itself. Note that both the numerator and denominator in $\hat{L}^{-1}$ are summed over the original set of nodes, $V$. Nodes that are removed have infinite distance to the rest of the network, the inverse of which is zero, and so do not contribute to the sum in Eqn. 4. The value that $\hat{L}^{-1}$ takes ranges from 0 to 1. A smaller value indicates more disruption to network communication and lower communication efficiency.

We measure $\hat{L}^{-1}$ and $S$ of a network before and after takedown to evaluate the effectiveness of the takedown strategy (Section 4), and also measure them after the network attempts to "heal" itself to assess the effectiveness of recovery mechanisms (Section 5).

## 4 Network Resilience

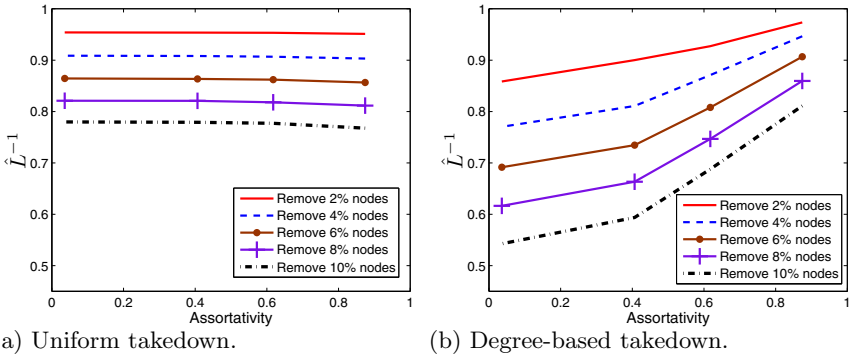
In attempts to take down a P2P botnet, network administrators may wish to prioritize their efforts to focus on the more "important" nodes first, i.e., nodes whose removal will cause the most disruption to botnet operation. Using the two metrics described in Section 3.4, we investigate the effectiveness of botnet takedown strategies, and how they are sensitive to the assortativity of the network.

### 4.1 Uniform and Degree-Based Takedown Strategies

We first focus on strategies explored in previous works that study botnet models [10,11,41,19]:

- **Uniform takedown:** removing nodes from the network by selecting them uniformly at random.
- **Degree-based takedown:** removing nodes from the network in descending order of node degree, that is, targeting high-degree nodes first.

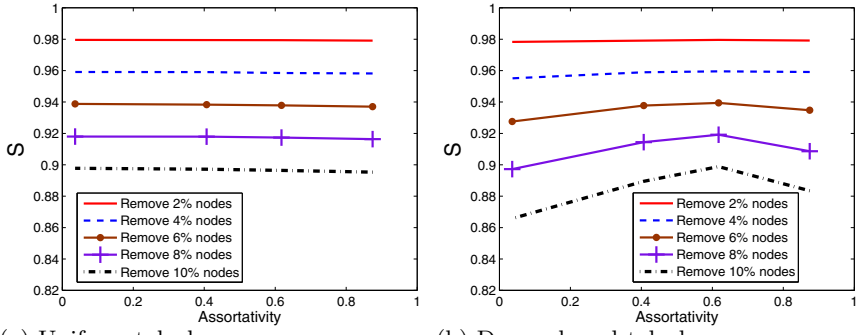
Uniform takedown is similar to the process in which users and network administrators patch infected hosts as they are discovered, without coordinating bot discoveries or patching activities. It has also been used to study random failures in the context of communication networks or biological networks [1]. While most networks are found to be resilient to uniform takedown, many are vulnerable to a degree-based strategy. This targeted takedown strategy is especially effective against scale-free networks, since the few highly-connected "hub" nodes responsible for maintaining the connectivity of the network are removed first, e.g., the "super-peers" that are found in unstructured P2P networks. The degree of a node, interpreted as the number of hosts with which it communicates, has also been used as an indicator of anomalies in network intrusion detection systems (e.g., [23,33,34]). In practice, these takedown strategies do not necessarily require access to the entire network communication graph, but can be applied to takedown efforts within a sub-graph as well, e.g., within a local network. We further discuss implementation challenges in Section 6.




(a) Uniform takedown. (b) Degree-based takedown.

**Fig. 3.** The normalized average inverse geodesic length $\hat{L}^{-1}$ after uniform or degree-based takedown strategies

As described in Section 3.3, we adjust the standard deviation $\sigma$ of the edge probability distribution $e_{j,k}$ to generate networks of varying assortativity. For a scale-free network with 5,000 nodes, we set $\sigma$ to 1, 5, 10, and 15 to obtain networks covering a range of assortativity from 0.04 to 0.87. Figures 3 and 4 show how networks with varying levels of assortativity respond to uniform and




(a) Uniform takedown. (b) Degree-based takedown.

**Fig. 4.** The average fraction $S$ of nodes in the largest connected component after uniform or degree-based takedown strategies

degree-based takedown, when 2%, 4%, 6%, 8%, or 10% of nodes were removed according to each strategy. The numbers are an average of 50 networks generated for each value of $\sigma$. We omit the standard deviations from the plots since they were generally small, that is, within 0.007 for both $\hat{L}^{-1}$ and $S$.

We find the degree-based strategy to be much more effective at taking down a network compared to uniform takedown, in agreement with previous works. However, as shown in Figure 3(b), the effectiveness of the degree-based strategy is highly dependent on the level of assortativity of the network. A lower assortativity, e.g., toward the left of Figure 3(b), results in the network experiencing a larger decrease in $\hat{L}^{-1}$ after takedown attempts. The difference between the decrease in $\hat{L}^{-1}$ for assortative and non-assortative networks grows as more nodes are removed. A similar phenomenon can be observed in Figure 4(b) for the fraction $S$ of nodes in the largest connected component. With the exception of highly assortative networks (e.g., greater than 0.6), the fraction of nodes retained in the largest connected component increases with the level of assortativity. That is, more bots remain reachable to the attacker in moderately assortative networks.

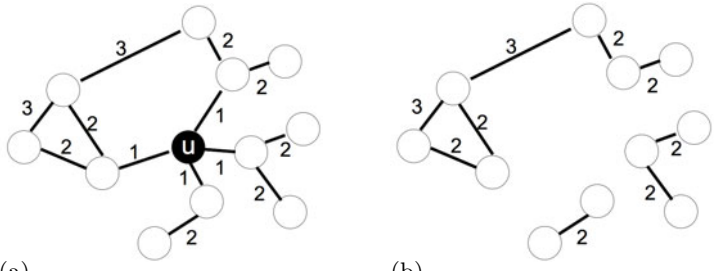
The higher resilience in assortative networks can be attributed to nodes of similar degree "clustering" together. When the high-degree nodes are removed due to the degree-based strategy, only a connected subset of neighboring nodes are lost in effect. Moreover, since high-degree nodes tend to connect to each other, fewer of their edges are attached to nodes of low degree — who would be prone to isolation if their neighbors were removed. However, this also means that there are fewer high-degree nodes that can act as "bridges" between clusters of nodes with varying degrees. As more high-degree nodes are removed, the loss of those "bridging" nodes eventually cancels out other factors contributing to resilience, and the network can disintegrate, as shown on the far right of Figure 4(b). These discrepancies in how networks are affected by the same takedown strategy underline the importance of taking assortativity into account, both in evaluating takedown strategies and in considering botnet network models.

### 4.2 Other Takedown Strategies

While the degree-based strategy is much more effective than the uniform strategy, the former is sensitive to the level of assortativity in the network, as shown in Figures 3 and 4. In the search for a takedown strategy that would be effective even for assortative networks, we explore alternative approaches based on other graph properties, described below.

- **Neighborhood connected components:** We define the local neighborhood of a node $u$ to be those nodes reachable within $h$ hops from it. Figure 5(a) shows an example of the neighborhood of node $u$ within three hops, where the edge labels indicate distances to $u$. If we were to remove $u$ from the network, its local neighborhood would be split into separate "connected components", as shown in Figure 5(b). The number of "connected components" that remains in the neighborhood of a node can be an approximation of its local importance, since communication between components may have to be routed through $u$. Hence, as an alternative takedown strategy, we remove nodes in descending order of the number of connected components in their local neighborhood. A similar metric has also been used to detect hit-list worms [7].
- **Closeness centrality:** Closeness centrality for a node $u$ is defined as the sum of the inverse geodesic distance from $u$ to all other nodes in the network. A larger value indicates that the node is at a more "centered" location, and has more influence over the spread of information within the network. In this strategy, we remove nodes in descending order of their closeness centrality.
- **Clustering coefficient with degree:** The clustering coefficient measures how dense the connections are between the neighbors of a node. For a node $u$, it is defined as the number of edges that exist between $u$'s neighbors, divided by the number of possible edges between $u$'s neighbors. In Figure 6, this value for $u$ is 4/10, while that for all other nodes is 1. A smaller value means that the neighbors of $u$ may be disconnected without $u$. Ignoring nodes with the smallest degrees — in our tests, nodes with degree less than one-fifth of the maximum degree — we remove nodes in increasing order of their clustering coefficient, and among those with the same clustering coefficient, in decreasing order of degree.

Figures 7 and 8 show the normalized average inverse geodesic length $\hat{L}^{-1}$ and the fraction $S$ of nodes in the largest connected component after each of the above takedown strategies, for networks of different levels of assortativity. The results are plotted after removing 2% or 10% of the nodes, and averaged over 50 networks generated for each level of assortativity. The standard deviations are all within 0.02 for both $\hat{L}^{-1}$ and $S$. Compared with the uniform and degree-based strategies discussed earlier, the clustering coefficient strategy is more effective at decreasing the network communication efficiency, as shown in Figure 7, while the connected components strategy seems more effective at lowering the connectivity of the network, as shown in Figure 8. In both of these cases, the alternative takedown strategy out-performs the degree-based strategy that previous works found to be effective [10,11,41].




(a) Before removing node $u$. (b) After removing node $u$.

**Fig. 5.** An example of the connected components within the neighborhood of node $u$. The edge labels indicate number of hops to $u$.

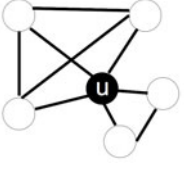
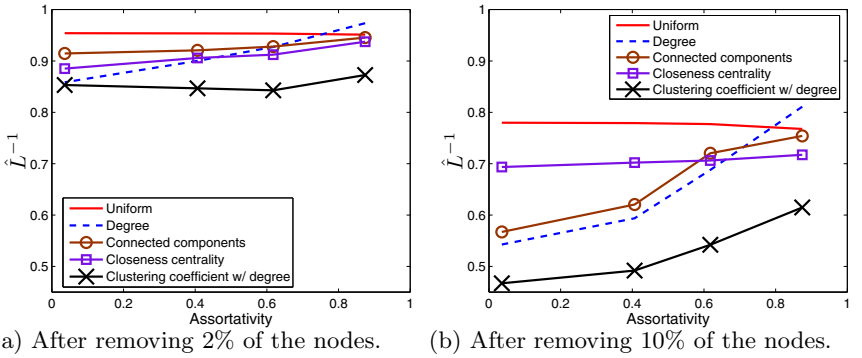


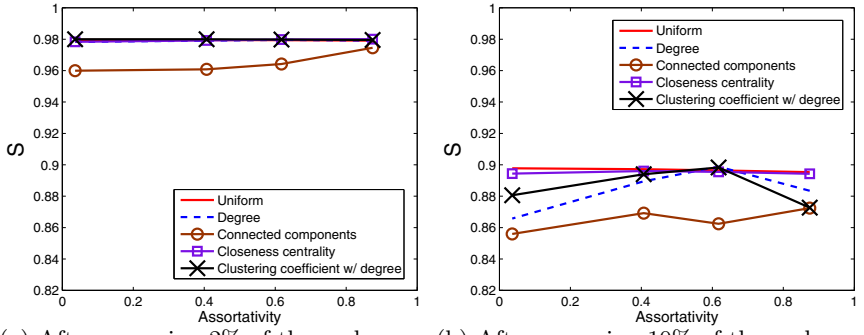

**Fig. 6.** An example of edges between neighbors of node $u$

One of the reasons that the clustering coefficient strategy works well is because nodes that "cluster" together in assortative networks are likely to have higher clustering coefficient as well, since their neighbors also have similar degree. However, while the nodes at the center of a "cluster" may have a clustering coefficient close to 1, this value is likely to be much smaller for those connecting the "cluster" to the rest of the network. For example, all nodes in Figure 6 have a clustering coefficient of 1 except for node $u$, who turns out to be the "bridge" between the two clusters of degree two and three nodes. The removal of nodes with small clustering coefficient in this strategy is hence likely to lower the communication efficiency within the network.




(a) After removing 2% of the nodes. (b) After removing 10% of the nodes.

**Fig. 7.** The normalized average inverse geodesic length $\hat{L}^{-1}$ after removing 2% or 10% of the nodes according to each takedown strategy




(a) After removing 2% of the nodes. (b) After removing 10% of the nodes.

**Fig. 8.** The average fraction $S$ of nodes in the largest connected component after removing 2% or 10% of the nodes according to each takedown strategy

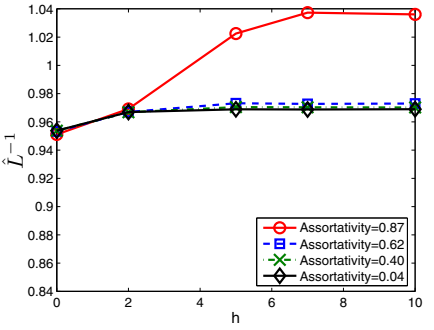
## 5 Network Recovery

The dynamism inherent in P2P networks means that each individual bot is required to adapt to changes in its surroundings, for example, due to newly infected hosts joining the network or current peers going offline, even without takedowns taking place. Such mechanisms would hence also provide opportunities for the network to *recover* itself, i.e., restoring connectivity or reconstructing shortest paths between nodes, in the face of takedown attempts.

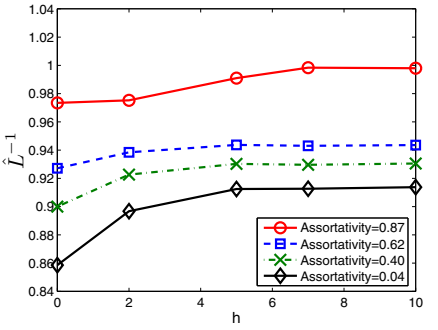
While previous works tend to regard a botnet as a static entity, and evaluate changes to the network immediately after takedown efforts as a measure of their effectiveness, we explicitly consider the ability of dynamic networks to heal themselves. Specifically, we model a recovery process where nodes can "look out" to a distance $h$ and find peers that are within $h$ hops. When a node loses a neighbor, e.g., due to takedown, it compensates for that lost neighbor by creating a new edge to a randomly selected node within distance $h$ from it. This models the edge creation process in a P2P botnet, where nodes discover others that are "close" to it through peer-list exchanges with its neighbors [4,6,31,16,36,35]. The $h$-neighborhood of a node $u$ hence represents hosts on $u$'s peer-list, to which $u$ looks for maintaining connectivity with the rest of the botnet.

### 5.1 Recovering from Uniform and Degree-Based Strategies

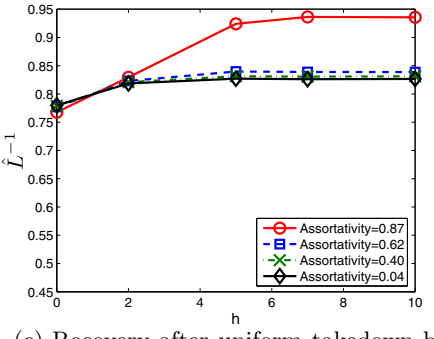
We first consider the ability of botnets to recover after takedown attempts employing the uniform or degree-based strategies described in Section 4.1. We focus on the $\hat{L}^{-1}$ metric, since it better illustrates the difference between networks of varying levels of assortativity. Figure 9 shows the normalized average inverse geodesic distance $\hat{L}^{-1}$ for networks after they attempt to recover from uniform or degree-based takedown strategies, when 2% or 10% of the nodes are removed. The numbers are averaged over 50 runs for each network, where the standard



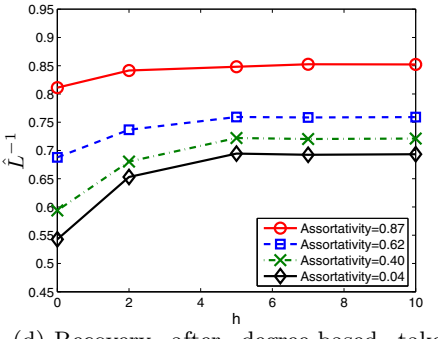

(a) Recovery after uniform takedown by removing 2% nodes.




(b) Recovery by degree-based takedown by removing 2% nodes.




(c) Recovery after uniform takedown by removing 10% nodes.




(d) Recovery after degree-based takedown by removing 10% nodes.

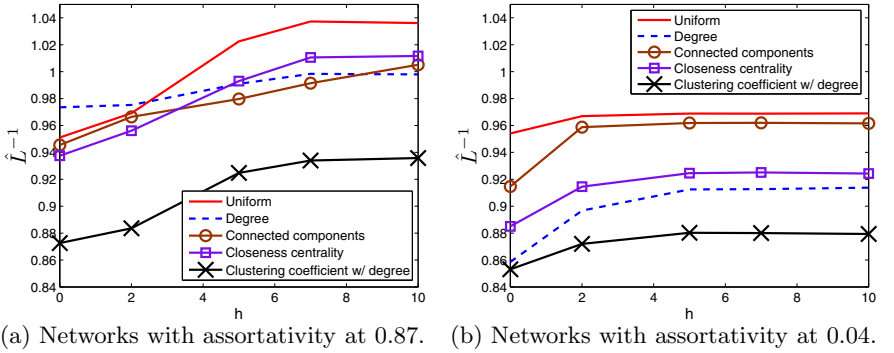
**Fig. 9.** The normalized average inverse geodesic length after recovering from uniform or degree-based takedown, when 2% or 10% of the nodes are removed, for various values of the look-out distance $h$

deviations are all below 0.006. The look-out distance $h$ was set to 2, 5, 7, and 10. As $h$ increases, $\hat{L}^{-1}$ increases as well, even reaching above 1 in Figure 9(a), i.e., the shortest distance between nodes becomes even shorter than before the takedown! However, while the increase in $\hat{L}^{-1}$ for networks with lower assortativity falls flat after a small $h$ (even decreasing slightly, as in Figure 9(d)), the increase for networks with higher assortativity continues.

One reason for the continued recovery benefit enjoyed by assortative networks is high-degree nodes "clustering" together, since nodes tend to connect to others of similar degree. A node that is able to reach a high-degree node upon "looking out" is likely to be able to reach other high-degree nodes as well at a similar distance. This increases the probability that a compensation edge attaches to a high-degree node, hence shortening path lengths within the network and resulting in a higher $\hat{L}^{-1}$. This phenomenon is more pronounced in networks recovering from uniform takedown (see Figures 9(a) and 9(c)), since fewer high-degree nodes remain after the degree-based strategy.

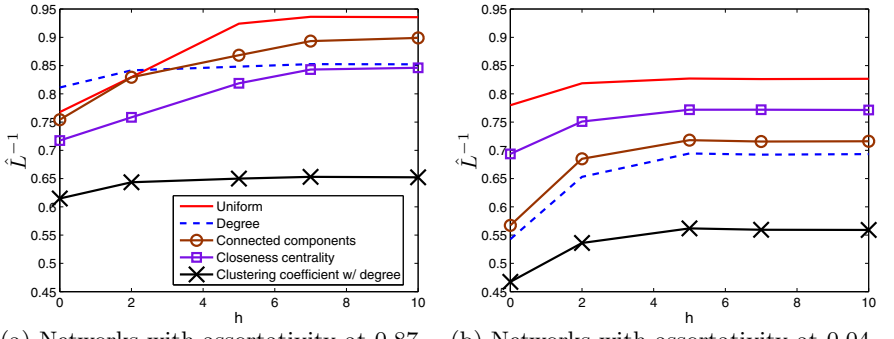
### 5.2 Recovering from Other Takedown Strategies

Figures 10 and 11 show how networks of high and low assortativity recover from those alternative takedown strategies described in Section 4.2, when 2% or 10% of the nodes were removed. The results are an average of 50 networks. The standard deviations are all within 0.009. We observe a trend similar to the recovery from uniform and degree-based strategies, where networks with higher levels of assortativity experience continued recovery benefits with the look-out distance $h$ (Figure 10(a) and 11(a)). Less assortative networks, on the other hand, do not benefit much after a look-out distance of 2 or 3 (Figure 10(b) and 11(b)). Regardless of the takedown strategy, assortative networks have higher communication efficiency after recovery, in terms of $\hat{L}^{-1}$, than less assortative networks.




(a) Networks with assortativity at 0.87. (b) Networks with assortativity at 0.04.

**Fig. 10.** Recovery for networks of high and low assortativity when 2% of the nodes were removed according to each strategy




(a) Networks with assortativity at 0.87. (b) Networks with assortativity at 0.04.

**Fig. 11.** Recovery for networks of high and low assortativity when 10% of the nodes were removed according to each strategy

In addition to being one of the most effective strategies (see Section 4.2), we also find takedown attempts based on clustering coefficient with degree to be the most difficult one for a network to recover from, as shown by low values of $\hat{L}^{-1}$

in Figures 10 and 11. In fact, when 10% of the nodes were removed from the same network, the $\hat{L}^{-1}$ after recovering from the degree-based and the clustering coefficient strategies can differ by 0.2. This shows that the clustering coefficient strategy can be a better alternative to one based solely on degree.

Besides creating compensation edges, a botnet may try to recover from takedowns by re-structuring itself into alternative topologies that are more resilient. Exploring how bots can perform this effectively in practice is part of future work.

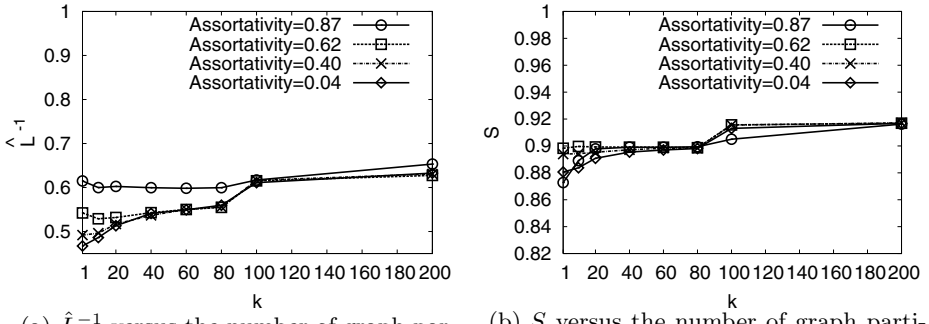
## 6 Discussion

***Applying takedown strategies in practice.*** Perhaps one of the reasons for the widespread study of the degree-based strategy is that it can be applied easily in practice. For example, if the degree of a node is interpreted as the number of hosts with which it communicates in some time interval, then identifying a node's degree can be performed on the basis of flow records (e.g., Cisco Netflows) that are collected from a router (or routers) that its traffic traverses. Notably, a node's degree can be determined solely by observing traffic to and from it, without requiring knowledge about the entity at the other end of the communication.

Other graph properties, however, may not be so straightforward to measure. For instance, takedown strategies based on clustering coefficient or neighborhood connected components depend on observing communications between the neighbors of a node, and may require collaboration between multiple administrative domains. This can be performed using a method similar to that proposed by Xie et al. to trace the origin of worm propagations [39]. Another approach is to examine the peer-lists an infected host receives from its neighbors, assuming that such data can be captured (i.e., it is not sent encrypted, and full packet capture is enabled on the network). If a node $u$ has two neighbors communicating with each other, those nodes should be listed on each other's peer-lists, and so the fact that they communicate with each other can be inferred by identifying overlaps between $u$'s neighbors and peer-lists sent to $u$. Of course, in cases where communications between some neighbors of an infected node are visible neither directly nor by inference, takedown strategies requiring this information can be applied considering only those neighbors for which communications are visible.

To examine the effect of applying takedown strategies locally, we generated networks according to the method described in Section 3.3, and partitioned the network randomly into $k$ equal-sized portions. The clustering coefficient with degree strategy (which we find to be the most effective, see Section 4.2) was then applied separately in each partition, i.e., based on only those edges attached to nodes in each partition. Figure 12 shows the normalized average geodesic length $\hat{L}^{-1}$ and the fraction $S$ of nodes in the largest connected component for varying values of $k$, when 10% of all nodes are removed this way. The numbers were averaged over 50 runs of this experiment. The standard deviations were all within 0.027 for $\hat{L}^{-1}$ and 0.013 for $S$. As shown in the figure, the takedown strategy becomes less effective as the number of network partitions increases, though the difference is small. For example, splitting highly assortative networks

(assortativity at 0.87) into 200 partitions only increases $S$ by 5% compared to the case when the network is not partitioned (i.e., $k$ =1). We hence believe that our suggested takedown strategies can be applied with reasonable effectiveness in practice.




(a) $\hat{L}^{-1}$ versus the number of graph partitions.

(b) $S$ versus the number of graph partitions.

**Fig. 12.** The fraction $S$ of nodes in the largest connected component and the normalized average inverse geodesic length $\hat{L}^{-1}$ after applying the clustering coefficient with degree takedown strategy locally in each of the $k$ network partitions and removing 10% of the nodes

***Modeling networks analytically.*** Rather than assuming a particular network topology, e.g., random, scale-free, or small-world, or a specific level of assortativity, another approach to modeling networks is to specify a set of actions governing the behavior of nodes at each step in time, and analytically determine properties of the resulting network. This type of growing network models have been used extensively in the physics domain of complex networks [3,24,32,18,12]. Given knowledge of individual bot behaviors and how they interact with each other from P2P bot studies [4,6,31,16,36], it seems likely that analytical network models from the physics literature can be adapted to characterize P2P botnets. In fact, a recent work by Li et al. [20] used this approach to derive the degree distribution of a botnet where new nodes joins the network by "copying" the edges of an existing node that it chooses at random.

However, these analytical approaches do make other assumptions about the underlying network that they attempt to model in order to simplify calculations. Specifically, by assuming that both the age of the network $t$ and the network size $N$ is large, $t \to \infty$, $N \gg 1$, all actions experienced by a node are approximated by the *expected* action, e.g., when a node creates one edge at random, the degree of all other nodes increases by $1/N$, where the denominator $N$ is also replaced by the expected value. These assumptions may not be applicable to botnets in practice, since 1) network administrators will be equally, if not more, concerned about infections in the early stages of a botnet when $t$ is small; 2) botnets have been found to consist of a few hundred or thousand nodes only, and are commonly rented out in small numbers, e.g., for sending spam; 3) to

a network administrator managing a local network, $N$ certainly does not grow indefinitely; and 4) approximating aspects of network growth using expected values introduces error that could potentially be magnified by a bot designed counter to assumptions that these approximations imply.

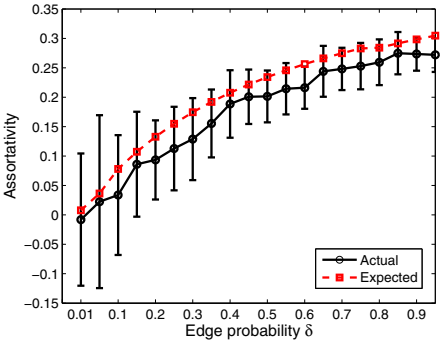



**Fig. 13.** The expected assortativity, shown in the dashed line, versus the actual average value from simulations, with one standard deviation shown with error bars

As a simple demonstration of the separation between analytical models and actual network growth, we examine a derivation by Callaway et al. [5] of the assortativity of a simple network growth model. In each time step, the model assumes that one node joins the network, and with probability $\delta$ an edge forms between two nodes selected at random. Their derivation of the assortativity is based on a rate equation specifying the *expected* increase in the number of edges that connect nodes of remaining degree $j$ and $k$ at each time step, and makes the same assumptions as described above. Figure 13 shows the expected assortativity of the network as approximated by Callaway et al. for various values of $\delta$. The actual average values from simulations are also plotted in the figure, with one standard deviation shown as error bars. To generate these values, we generated 50 networks for each value of $\delta$, and set the number of time steps (i.e., number of nodes) to 1,000. Figure 13 shows that the expected assortativity as predicted by Callaway et al. can differ from the actual average assortativity by an amount that approaches or, in some cases, exceeds one standard deviation. This suggests that the simplifying assumptions typically employed in analytical models may cause nontrivial deviations from practice.

## 7 Conclusion

Peer-to-peer (P2P) botnets, in contrast to their centralized counterparts, do not have a single point-of-failure and are difficult to take down. Identifying and removing those nodes that are "important" to the connectivity or communication efficiency of a botnet is hence critical to disrupting its operation. Toward this goal, several previous works have modeled P2P botnets using theoretical network models [19,10,11]. These works compare the resilience of various network topologies to uniform or degree-based node removals, and quantify the effectiveness of these takedown strategies using graph properties, including the inverse geodesic length or the fraction of nodes in the largest connected component.

We observe that previous works do not consider an important structural property of networks, namely assortativity. Empirical measurements on networks simulated according to the Waledac botnet protocol and on network traffic from a

portion of the Storm botnet suggest that this property can be quite high for botnets in practice. We show that in omitting the presence of assortativity in botnet models, and without considering the effect of dynamic networks actively recovering from node failures, previous works may have over-estimated the effectiveness of recommended takedown strategies. In addition, we identify alternative strategies that are more effective than those in previous works for botnets with high assortativity, and study the application of these strategies in a "local" setting when only a subset of the network is visible.

## References


1. Albert, R., Jeong, H., Barabasi, A.L.: Error and attack tolerance of complex networks. Nature 406 (2000)
2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
3. Barabási, A.L., Albert, R., Jeong, H.: Mean-field theory for scale-free random networks. Physica A 272, 173–187 (1999)
4. Borup, L.: Peer-to-peer botnets: A case study on Waledac. Master's thesis, Technical University of Denmark (2009)
5. Callaway, D.S., Hopcroft, J.E., Kleinberg, J.M., Newman, M.E.J., Strogatz, S.H.: Are randomly grown graphs really random? Phys. Rev. E 64(4), 041902 (2001)
6. Calvet, J., Davis, C.R., Bureau, P.: Malware authors don't learn, and that's good! In: Intl. Conf. Malicious and Unwanted Software (2009)
7. Collins, M.P., Reiter, M.K.: Hit-List Worm Detection and Bot Identification in Large Networks Using Protocol Graphs. In: Kruegel, C., Lippmann, R., Clark, A. (eds.) RAID 2007. LNCS, vol. 4637, pp. 276–295. Springer, Heidelberg (2007)
8. Cooke, E., Jahanian, F., McPherson, D.: The zombie roundup: Understanding, detecting, and disrupting botnets. In: Wksh. Steps to Reducing Unwanted Traffic on the Internet (2005)
9. Crucitti, P., Latora, V., Marchiori, M., Rapisarda, A.: Error and attack tolerance of complex networks. Phys. A 340, 388–394 (2004)
10. Dagon, D., Gu, G., Lee, C.P., Lee, W.: A taxonomy of botnet structures. In: Annual Computer Security Applications Conf. (2007)
11. Davis, C.R., Neville, S., Fernandez, J.M., Robert, J.-M., McHugh, J.: Structured Peer-to-Peer Overlay Networks: Ideal Botnets Command and Control Infrastructures? In: Jajodia, S., Lopez, J. (eds.) ESORICS 2008. LNCS, vol. 5283, pp. 461–480. Springer, Heidelberg (2008)
12. Dorogovtsev, S.N., Mendes, J.F.F.: Scaling properties of scale-free evolving networks: Continuous approach. Phys. Rev. E 63, 056125 (2001)
13. Erdös, P., Rényi, A.: On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5, 17–61 (1960)
14. Gu, G., Perdisci, R., Zhang, J., Lee, W.: BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection. In: USENIX Security Symp. (2008)
15. Holme, P., Kim, B., Yoon, C., Han, S.: Attack vulnerability of complex networks. Phys. Rev. E 65, 056109 (2002)
16. Holz, T., Steiner, M., Dahl, F., Biersack, E., Freiling, F.: Measurements and mitigation of peer-to-peer-based botnets: A case study on Storm worm. In: USENIX Wksh. Large-Scale Exploits and Emergent Threats (2008)

17. Jackson, M.O., Rogers, B.W.: Meeting strangers and friends of friends: How random are social networks? American Economic Review 97(3) (2007)
18. Krapivsky, P., Redner, S.: Organization of growing random networks. Phys. Rev. E 63, 066123 (2001)
19. Li, J., Ehrenkranz, T., Kuenning, G., Reiher, P.: Simulation and analysis on the resiliency and efficiency of malnets. In: Wksh. Principles of Advanced and Distributed Simulation (2005)
20. Li, X., Duan, H., Liu, W., Wu, J.: The growing model of botnets. In: Intl. Conf. Green Circuits and Systems (2010)
21. Liben-Nowell, D., Balakrishnan, H., Karger, D.: Analysis of the evolution of peer-to-peer systems. In: ACM Symp. Principles of Distributed Computing (2002)
22. Matei, R., Iamnitchi, A., Foster, P.: Mapping the Gnutella network. IEEE Internet Computing 6, 50–57 (2002)
23. Mirkovic, J., Prier, G., Reiher, P.: Attacking DDoS at the source. In: IEEE Intl. Conf. Network Protocols (2002)
24. Moore, C., Ghoshal, G., Newman, M.: Exact solutions for models of evolving networks with addition and deletion of nodes. Phys. Rev. E 74, 036121 (2006)
25. Newman, M.: Assortative mixing in networks. Phys. Rev. Lett. 89(20) (2002)
26. Newman, M.: Mixing patterns in networks. Phys. Rev. E 67, 026126 (2003)
27. Newman, M., Park, J.: Why social networks are different from other types of networks. Phys. Rev. E 68, 036122 (2003)
28. Pandurangan, G., Raghavan, P., Upfal, E.: Building low-diameter P2P networks. In: IEEE Symp. Foundations of Computer Science (2001)
29. Pastor-Satorras, R., Vazquez, A., Vespignani, A.: Dynamical and correlation properties of the internet. Phys. Rev. Lett. 87(25) (2001)
30. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. Phys. Rev. Lett. 86(14) (2001)
31. Porras, P., Saidi, H., Yegneswaran, V.: A multi-perspective analysis of the Storm (Peacomm) worm. Tech. rep., Computer Science Laboratory, SRI International (2007)
32. Sarshar, N., Roychowdhury, V.: Scale-free and stable structures in complex ad hoc networks. Physical Review E 69(2), 026101 (2004)
33. Schechter, S.E., Jung, J., Berger, A.W.: Fast Detection of Scanning Worm Infections. In: Jonsson, E., Valdes, A., Almgren, M. (eds.) RAID 2004. LNCS, vol. 3224, pp. 59–81. Springer, Heidelberg (2004)
34. Sekar, V., Xie, Y., Reiter, M.K., Zhang, H.: A multi-resolution approach for worm detection and containment. In: Intl. Conf. Dependable Syst. and Netw. (2006)
35. Sinclair, G., Nunnery, C., Kang, B.B.: The Waledac protocol: The how and why. In: Intl. Conf. Malicious and Unwanted Software (2009)
36. Stover, S., Dittrich, D., Hernandez, J., Dietrich, S.: Analysis of the Storm and Nugache trojans: P2P is here. USENIX; Login 32(6) (2007)
37. Watts, D.J.: A simple model of global cascades on random networks. Natl. Acad. Sci. 99(9) (2002)
38. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature 393 (1998)
39. Xie, Y., Sekar, V., Reiter, M.K., Zhang, H.: Forensic analysis for epidemic attacks in federated networks. In: 14th IEEE Intl. Conf. Network Protocols (2006)
40. Xulvi-Brunet, R., Sokolov, I.: Reshuffling scale-free networks: From random to assortative. Phys. Rev. E 70, 066102 (2004)
41. Yu, J., Li, Z., Hu, J., Liu, F., Zhou, L.: Using simulation to characterize topology of peer to peer botnets. In: Intl. Conf. Computer Modeling and Simulation (2009)