Characterizing the Robustness of Complex Networks

Ali Sydney, Caterina Scoglio, Mina Youssef, and Phillip Schumm

Electrical and Computer Engineering Department
Kansas State University
Manhattan, KS USA
Email:{asydney,caterina,mkamel,pbschumm}@ksu.edu

Abstract

With increasingly ambitious initiatives such as GENI and FIND that seek to design future internets, it becomes imperative to define the characteristics of robust topologies, and build future networks optimized for robustness. This paper investigates the characteristics of network topologies that maintain a high level of throughput in spite of multiple attacks. To this end, we select network topologies belonging to the main network models and some real world networks. We consider three types of attacks: removal of random nodes, high degree nodes, and high betweenness nodes. We use elasticity as our robustness measure and, through our analysis, illustrate that different topologies can have different degrees of robustness. In particular, elasticity can fall as low as 0.8% of the upper bound based on the attack employed. This result substantiates the need for optimized network topology design. Furthermore, we implement a tradeoff function that combines elasticity under the three attack strategies and considers the cost of the network. Our extensive simulations show that, for a given network density, regular and semi-regular topologies can have higher degrees of robustness than heterogeneous topologies, and that link redundancy is a sufficient but not necessary condition for robustness.

Key words: Complex Networks, Robustness, Optimization, Attack, Tradeoff, Topology, Heterogeneity, Characteristic Path Length

1. Introduction

Why study future network topologies? For one, we have experienced several moderate sized failures and thus, large failures are inevitable. In particular, the 2006 earthquake in Taiwan disrupted undersea fiber optic communication lines and as a result, banks from South Korea to Australia suffered massive interruptions [1]. Though this represents a direct network failure, failures can also occur indirectly. For example Code Red, a computer virus that incapacitated numerous networks, resulted in a global loss of 2 billion US dollars [2]. Furthermore, in 2004, Sassar virus disruptions accounted for the halt on maritime operations in the UK, the halt on railway operations in Australia, and interruptions in hospital facilities in Hong Kong [3]. The US General Accounting Office estimated 250,000 annual attacks on Department of Defense networks [4]. Objectives range from theft to immobilization of entire networks. Another riveting example stems from a series of cascading failures in 2003 that resulted in a blackout in the Northeastern states [5]. A similar phenomenon occurred the very same year in Italy, and left 56 million residents without power for 9 hours [6]. Our daily routines would cease to exist should network topologies disintegrate. Thus, as failures and attacks increase, it is imperative to design future topologies robust against unforeseen catastrophes for future network initiatives.

Amongst other definitions, a network can be robust if disconnecting components is difficult. However, we define robustness as the ability of a network to maintain its total throughput under node and link removal. The former definition is based on topological characteristics, while the latter also considers flows within the network such as IP packets.

Approaches for determining the robustness of graphs have evolved from simple graph theoretic concepts that highlight the connectivity of a graph [7] to more recent concepts that consider the spectrum of a graph [8]. However, these measures are unable to capture our definition of robustness. For this reason, we use elasticity as our measure of robustness; it meets the functional requirements of capturing throughput under node and link removal.

The importance of this paper stems from our objective to extract the characteristics of robust networks. With these results, we seek to produce future robust network topologies. Thus, to realize our first goal, we 1) use the metric elasticity as a measure of robustness of a network, 2) establish the upper bound for elasticity, 3) assess elasticity for diverse network models, 4) present correlations between elasticity and selected network

metrics, 5) develop and implement a function that considers the tradeoff between elasticity and network cost, and 6) extract characteristics of networks that make them robust.

The rest of this paper is structured as follows. Section 2 reviews measures of robustness based on the structure and behavior of the network. Section 3 presents the network models from which networks will be selected to assess their elasticity. In Section 4, we review elasticity, our robustness measure, and provide analytical and numerical approaches to obtain the upper bound. In Section 5, we assess the elasticity of each network, implement a tradeoff function that considers elasticity under the three removal strategies and discuss the characteristics which make a network robust. Finally, we discuss the benefits and shortcomings of elasticity and highlight our future initiatives to characterize the robustness of complex networks in Section 6.

2. Background and Related Work

The classical approach for determining robustness of networks entails the use of basic concepts from graph theory. For instance, the connectivity of a graph is an important, and probably the earliest, measure of robustness of a network [7]. Node (link) connectivity, defined as the size of the smallest node (link) cut, determines in a certain sense the robustness of a graph to the deletion of nodes (links). However, the node or link connectivity only partly reflects the ability of graphs to retain certain degrees of connectedness after deletion. Other improved measures were introduced and studied, including super connectivity [9], conditional connectivity [10], restricted connectivity [11], fault diameter [12], toughness [13], scattering number [14], tenacity [15], expansion parameter [16], and isoperimetric number [17]. In contrast to node (link) connectivity, these new measures consider both the cost to damage a network and how badly the network is damaged.

Subsequent measures consider the size of the largest connected component as nodes are attacked [18]. Furthermore, percolation models were used to assess the damage incurred by random graphs [19]. From spectral analysis, experimentalists consider the second smallest Laplacian eigenvalue as a measure of how difficult it is to break the network into components [8].

The measures reviewed thus far consider the network structure to assess robustness. However, more recent efforts have incorporated the behavior of the network [20,21]. More precisely, the authors maximized flows in the network while imposing constraints on routers and links.

Other metrics in networking literature include the average node degree [22], betweenness [23], heterogeneity [24], and characteristic path length [25]. In this paper, our results show significant corelations between elasticity and some of these metrics which will be used to characterize the robustness of networks.

3. Network Models

This section reviews the six models from which 18 topologies were selected. They include networks from random models, Watts-Strogatz models, preferential attachment models, near-regular models, trade-off and optimization models, and real-world models. For each topology, some of the more common properties are shown in Table 1.

Table 1
Network characteristics where ASP is the average shortest path and Het is heterogeneity

Networks	♯ Nodes	♯ Links	Density	Diameter	ASP	Het
Gi-dense	1000	4505	0.00902	7	3.391	0.331
MySpace	955	10976	0.02409	4	2.013	2.027
Watts-Strogatz 1	1000	3000	0.00601	7	4.14	0.301
PA 2	1000	2964	0.00593	6	3.534	1.109
Gi-sparse	1000	2009	0.00402	12	5.154	0.491
PA 1	1000	1981	0.00397	8	4.177	1.185
Watts-Strogatz 2	1000	2000	0.004	9	5.294	0.37
YouTube	1089	1576	0.00266	12	5.096	1.319
Flickr	967	1515	0.0032	12	4.624	1.394
meshcore	1000	1275	0.00255	3	2.911	3.796
near-regular 2	992	3781	0.00769	31	14.706	0.133
HOT 2	1000	1049	0.0021	12	7.144	1.892
ringcore	1000	1000	0.002	14	8.196	3.122
HOT 1	939	988	0.00224	10	6.812	2.032
PA-sparse	1000	1049	0.0021	14	5.793	1.892
Abilene	886	896	0.00229	10	6.95	2.09
near-regular 1	992	1921	0.00391	61	21	0.089

3.1. *Random models*

A random graph is obtained by random addition of links between n vertices. Two notable properties are 1) the average node degree determines the connectivity of the graph and 2) the node degree can be approximated using a Poisson distribution. Erdos-Renyi's (ER) stochastic model is one of the most studied of these models. In the construction of an ER graph $G(N, E)$, E edges are connected at random to N nodes [19]. However, this paper considers the Gilbert (Gi) model $G(N, p)$, a modified version of the ER model where edges are connected to vertices with a probability of p. For the Gi-dense and Gi-sparse networks used in this paper, $p = 0.0091$ and 0.004094 respectively [26]. Figure 1 shows the Gi-sparse network.

3.2. *Watts-Strogatz Models*

The Watts-Strogatz model is constructed by interpolating between a regular ring lattice and a random network [19]. Each node is connected to its k nearest neighbors and random

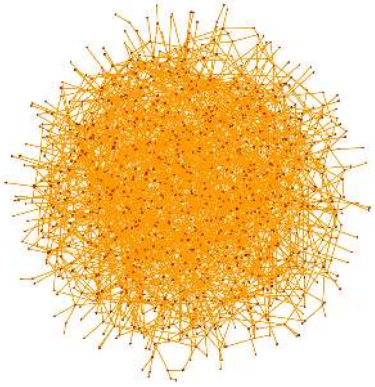

Fig. 1. The Gi-sparse network with size $N = 1000$ and average degree $\bar{k} = 4.018$

rewiring occurs with a probability of p. For intermediate values of p, Watts-Strogatz models produces a Small-world network which captures the high clustering properties of regular graphs and the small characteristic path length of random graph models. For the Watts-Strogatz (W-S) 1 and 2 networks used, the rewiring probability was 0.3 and 0.5 [26]. Figure 2 shows the W-S 1 network.

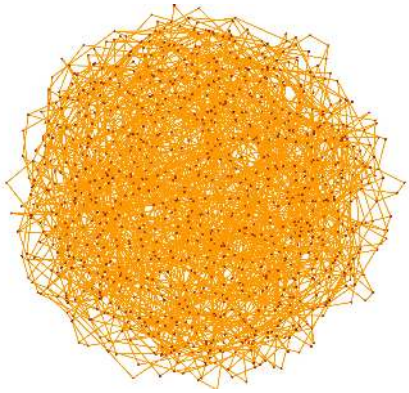

Fig. 2. The W-S 1 network with $N = 1000$ and $\bar{k} = 4$

3.3. *Preferential Attachment Models*

From their origin, preferential attachment (PA) models have been considered vulnerable to targeted attacks while robust to random failures and have a heavy tail distribution [27]. This model constitutes popular nodes called "hubs" that have a large number of neighbors compared to other nodes with few neighbors. At each time step, nodes with a higher degree have a higher probability of attracting new nodes than nodes with a lower degree. For this work, the PA 1, PA 2, and PA-sparse networks were constructed using the Barabasi-Albert Scale-free model [26,20]. Figure 3 shows the PA-sparse network.

3.4. *Near-Regular Models*

The near-regular (n-r) networks are best visualized in a planar, grid-like fashion. The n-r 1 network is composed of a 31 by 32 grid where node i is connected to node j if j is a distance

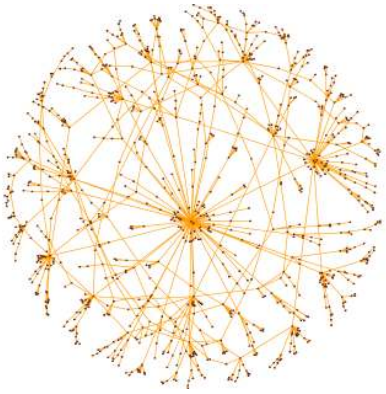

Fig. 3. The PA-sparse network with $N = 1000$ and $\bar{k} = 2.098$

$d = 1$ unit: 1 unit is the regular distance among nodes in the grid. The structure of n-r 2 is similar to that of the regular. However, in addition to $d = 1$ unit, all nodes within a distance of $d = \sqrt{2}units$ are connected. Figure 4 shows the n-r 1 network.

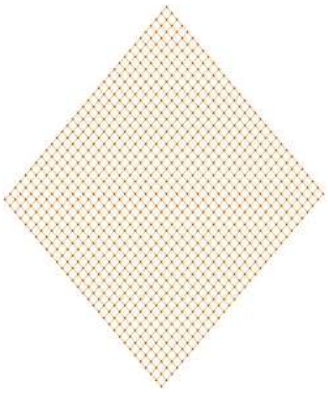

Fig. 4. The n-r 1 network with $N = 992$ and $\bar{k} = 3.87$

3.5. *Trade-off and Optimization Models*

The authors of [28] introduce networks with bimodal degree distributions optimized to minimize the impact of random attacks. The meshcore and ringcore topologies shown in Figures 5 and 6 represent this model. The Heuristically Optimized Trade-off (HOT) network presents a simple model for Internet growth [29,20]. The HOT 1 and 2 networks represent this model. Figure 7 shows the HOT 2 network.

3.6. *Real-World Models*

Online social networking connects individuals with common interests. This paper features the MySpace, YouTube, and Flickr networks. These networks were obtained via snowball sampling and have been rescaled [30]. The Abilene network in Figure 8 was built using the Abilene core while customers and peer networks were each replaced with a gateway router [20].

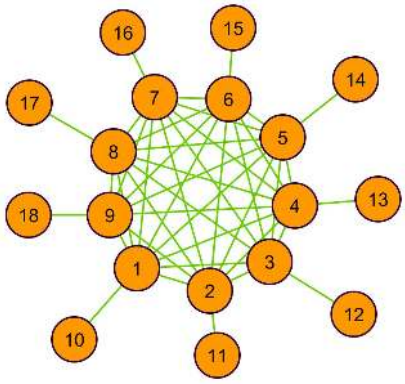

Fig. 5. The meshcore network with $N = 1000$ and $\bar{k} = 2.55$

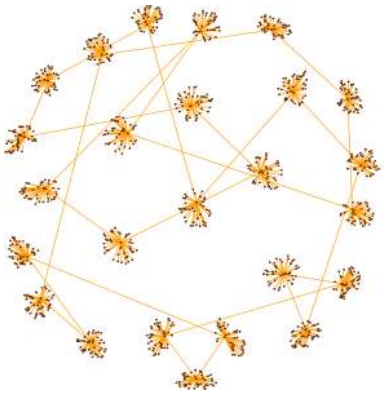

Fig. 6. The ringcore network with $N = 1000$ and $\bar{k} = 2$

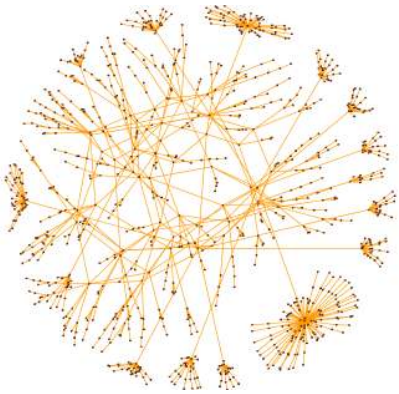

Fig. 7. The HOT 2 network with $N = 1000$ and $\bar{k} = 2.098$

4. Robustness Metric

The study of robustness is fundamental to numerous network research problems using approaches that amplify internal behaviors of a network. To this end, we use elasticity as our measure of robustness, obtain its upper bound and finally, select the most feasible routing algorithm for elasticity.

4.1. *Elasticity*

For a network G, having no loops or parallel links, elasticity $E(G)$ is a measure of the overall robustness. As shown in Figure 9, elasticity is the area under the curve of throughput versus the

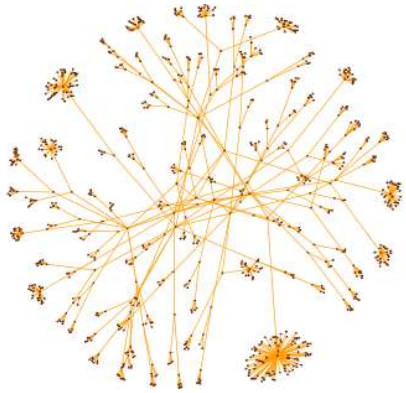

Fig. 8. The Abilene network with $N = 1000$ and $\bar{k} = 2.022$

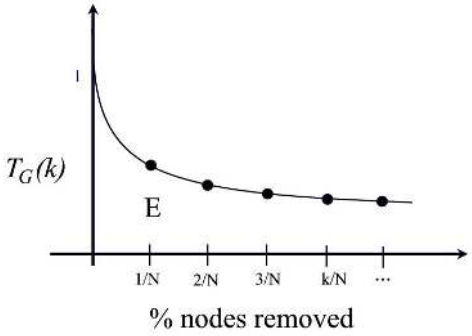

Fig. 9. The evaluation of elasticity

percentage of nodes removed. The throughput is normalized to compare networks of different magnitudes and at each iteration, it is recalculated at the removal of each node. Initially, $T_G(0) = 1$ which accounts for the normalized throughput. This value decreases as $\frac{k}{N}\%$ of nodes are removed and therefore, elasticity (E) provides a measure of robustness at any point of node removal.

Therefore, when ζ nodes have been removed, elasticity can be computed as

$$E\left(\frac{\zeta}{N}\right) = \frac{1}{2N}\sum_{k=0}^{\zeta}\left(T_G\left(\frac{k}{N}\right) + T_G\left(\frac{k+1}{N}\right)\right) \quad (1)$$

where $T_G(\frac{k}{N})$ is the throughput at each interval when k nodes are removed. N is the total number of nodes in the network and $0 \leq (\ zeta, k) \leq N$. At each iteration, the throughput is computed as

$$T_G(t) = \frac{\max_\rho \sum_{i,j} X_{i,j}(t)}{\alpha} \quad s.t. \quad LX \leq B(t) \quad (2)$$

where $t = \frac{k}{N}$ and ρ is a constant used to vary the proportion of flows in network. α is the unnormalized initial throughput and $X_{i,j}(t)$ is the traffic flow between source node i and destination node j. L is the routing matrix, X is a vector of all $X_{i,j}(t)$ flows, and $B(t)$ is a vector of all link bandwidth capacities.

4.2. *Upper bound for Elasticity*

4.2.1. *Analytical results*

In this section, we consider the mesh network as the topology which provides the highest elasticity under all attack strategies for any given network. We assume homogeneous flows where each flow has a value of 1. Additionally, each link has a capacity of 1 and $X_{ij}(t)$ can be 1 or 0 depending on whether or not a flow exists between nodes i and j. With these assumptions, we proceed to determine the upper bound for elasticity.

Theorem. Given a mesh network with N nodes, and assuming homogeneous flows and link capacities of 1, then $\lim_{N\to\infty} E(N) = \frac{1}{3}$.

Proof. Elasticity can be formulated using both discrete and continuous approaches. At each iteration when a node is removed, the throughput is given by

$$T_G(t) = \frac{(N-k)(N-k-1)}{N(N-1)} \tag{3}$$

where $t = \frac{k}{N}$.

Discrete Elasticity (trapezoidal integration). For a given network of size N, Equation 4 computes elasticity when ζ nodes have been attacked.

$$E(\zeta) = \frac{1}{N}\left(\frac{1}{2} + \sum_{k=1}^{\zeta-1}\beta + \delta\right) \tag{4}$$

where $\beta = \frac{(N-k)(N-k-1)}{N(N-1)}$, $\delta = \frac{(N-\zeta)(N-\zeta-1)}{2N(N-1)}$, and $\zeta \leq N-1$. Equation 5 computes the total elasticity for a network with N nodes when all N nodes are progressively removed.

$$E(N) = \frac{1}{N}\left(\frac{1}{2} + \sum_{k=1}^{\zeta-1}\frac{(N-k)(N-k-1)}{N(N-1)}\right) \tag{5}$$

Continuous Elasticity Equation 6 gives the formulation of elasticity for the continuous case. Similar to the discrete case, Equation 7 computes elasticity for a given mesh network with size of N where ζ nodes have been removed and Equation 8 computes the total elasticity for a mesh network with N nodes. As the size of the network grows, Equation 9 then provides the upper bound on elasticity when all N nodes are removed.

$$E(t) = \int_0^t T_G(\tau)\,d\tau, \quad 0 \leq t \leq 1 \tag{6}$$

$$E(\zeta) = \frac{N(N-1)\zeta + \frac{1}{2}(1-2\zeta)\zeta^2 + \frac{1}{3}\zeta^3}{N^2(N-1)} \tag{7}$$

$$E(N) = \frac{1}{3} - \frac{1}{6N} - \frac{1}{6N^2} \tag{8}$$

Therefore,

$$\lim_{N\to\infty} E(N) = \frac{1}{3} \tag{9}$$

Q.E.D.

4.2.2. *Numerical Results*

Figure 10 compares the convergence rate of the discrete and continuous cases when ζ nodes have been attacked from a network where $N = 20$. As depicted, both approaches converge at the onset of node removal.

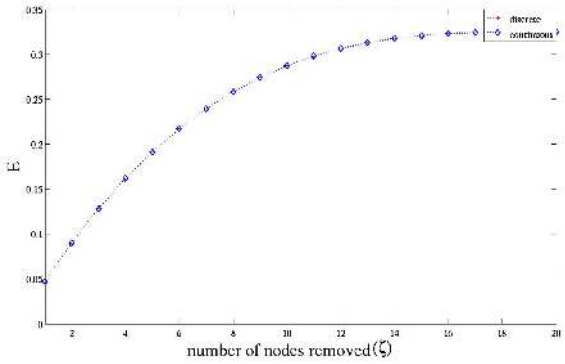

Fig. 10. Comparison of the convergence rates of elasticity, from Equations 4 and 7, where ζ nodes have been attacked.

Figure 11 compares the convergence rate of elasticity for the discrete case in Equation 5 to the continuous case in Equation 8. As shown, both cases converge for a network with 10 nodes.

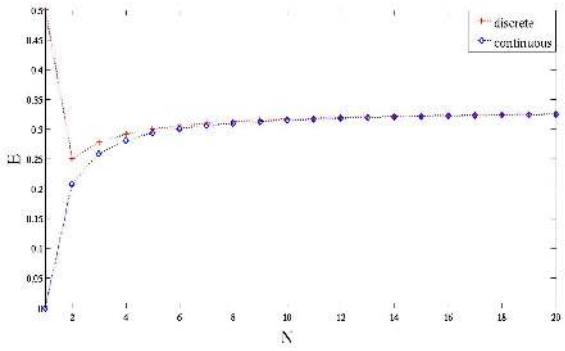

Fig. 11. Comparison of the convergence rates of elasticity, from Equations 5 and 8, for a network of size N

These convergence rates are significant because they necessitate few iterations. More importantly, the discrete approaches can be abandoned for the continuous approaches to simplify calculations without compromising accuracy.

4.3. *Routing Algorithm for Elasticity*

Elasticity depends on the routing algorithm selected. For this reason, three routing approaches are explored: 1) Optimization (heterogeneous traffic matrix), 2) Dijkstra's Algorithm (heterogeneous traffic matrix), and 3) Dijkstra's Algorithm (homogeneous traffic matrix). All approaches assumed homogeneous link capacities of 1.

4.3.1. *Optimization (Heterogeneous traffic matrix)*

The Objective Function (Function 10) of the optimization problem maximizes the individual flow between any pair of

nodes. Equations 11-14 are the main constraints to the optimization problem. Equation 11 ensures that each node sends δ unit of traffic to every node, while Equation 12 represents the balance of the incoming and outgoing traffic demands through any node in the network. Inequality 13 represents the capacity constraint on each link, and Equation 14 computes the utilization of each link.

$$Maximize\ \delta \tag{10}$$

Subject to

$$\sum_{j \in N} flow_{s,j,s} = \delta(N-1)\ \forall s \tag{11}$$

$$\sum_{i \in N} (flow_{i,j,s} - flow_{j,i,s}) = \delta\ \forall s, j, j \neq s \tag{12}$$

$$\sum_{s \in N} flow_{i,j,s} \leq capacity_{i,j}\ \forall i, j, i \neq j \tag{13}$$

$$utilization_{i,j} = \sum_{s \in N} flow_{i,j,s}\ \forall i, j \tag{14}$$

Algorithm 1 provides elasticity using the optimization approach discussed previously.

Algorithm 1 Optimization

while $Connected := True$ **do**
  $capacity_{i,j} := 1$
  $demand_{i,j} := 0$
  while $\sum_{i,j} capacity_{i,j} \neq 0$ **do**
    Solve the optimization problem
    Update the demand between nodes that are connected with non-zero capacity links
    $demand_{i,j} := demand_{i,j} + \delta$
    $capacity_{i,j} := capacity_{i,j} - utilization_{i,j}$
  end while
  Remove one node (or a group of nodes)
end while

4.3.2. *Dijkstra's algorithm (heterogeneous traffic matrix)*

The second approach realizes Dijkstra's algorithm. As shown in Algorithm 2, flows traverse the shortest path from source to destination. This algorithm has a running time $O(n^2)$. However, when the heterogeneous traffic matrix is considered, the running time increases to $O(n^3)$.

4.3.3. *Dijkstra's algorithm (homogeneous traffic matrix)*

This approach also revolves around Algorithm 2 and likewise, has a running time $O(n^2)$. However, a homogeneous traffic matrix was implemented. Given these three models, Subsection 4.4 evaluates each and selects the most feasible.

4.4. *Evaluation of Routing Models*

Figure 12 shows the three networks for which elasticity was computed: Net 1, Net 2, and Net 3. For these three networks,

Algorithm 2 Dijkstra's algorithm

begin
$S := 0; \bar{S} := N$
$d(i) := \infty$ for each node $i \in N$
$d(s) := 0$ and pred(s) $:= 0$
while $|S| < n$ **do**
begin
let $i \in \bar{S}$ be a node for which $d(i) = \min\{d(j) : j \in \bar{s}\}$
$S := S \cup \{i\}$;
$\bar{S} := \bar{S} - \{i\}$;
for each $(i, j) \in A(i)$ **do**
  if $d(j) > d(i) + c_{ij}$ **then** $d(j) := d(j) + c_{ij}$ and pred(j) $:= i$
end for
end

we compare the results of elasticity provided by each routing algorithm targeting first, nodes with the highest degree and second, nodes with highest betweenness.

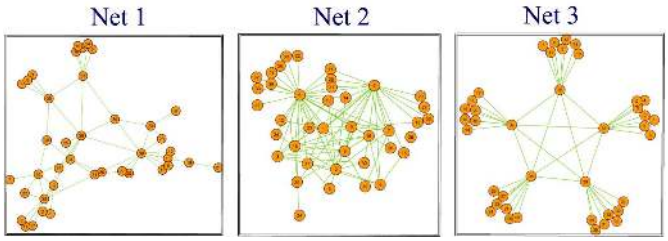

Fig. 12. Three networks for which elasticity was evaluated

Figure 13 shows the throughput degradation as nodes with highest degree are attacked in Net 1. As depicted, the optimization approach produces the highest elasticity, followed by Dijkstra's heterogeneous approach and finally, Dijkstra's homogeneous approach. This trend was observed for each network under both attack strategies. However, under certain circumstances where the network has low connectivity, the elasticity results were identical for both Dijkstra's "heterogeneous" and optimization approach.

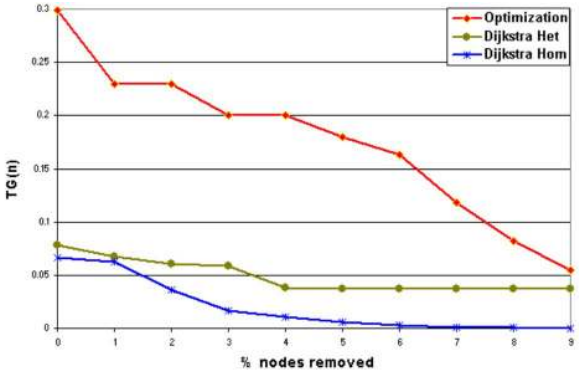

Fig. 13. Throughput degradation as nodes with highest degree are attacked for Net 1. "Het" represents a heterogeneous traffic matrix and "Hom" represents homogeneous traffic matrix.

For each of the three routing approaches, each network was given a rank of 1, 2 or 3, based on its value for elasticity: 1 as

the highest and 3 as the lowest. Table 2 displays the rankings for each network under highest node degree attack. As shown, elasticity was highest for Net 1, followed by Net 2, and finally, Net 3, for each approach. Though the values were different for the highest betweenness attack strategy (not shown), the rankings were similar to that of Table 2.

Table 2
Elasticity comparison for all networks under **highest node degree** attack

Algorithm	Rank 1	Rank 2	Rank 3
1	Net 1	Net 2	Net 3
2	Net 1	Net 2	Net 3
3	Net 1	Net 2	Net 3

Furthermore, we observed that the criteria for node addition to the shortest path could potentially affect the results of elasticity. More specifically, in Algorithm 2, nodes are added to the shortest path if the following optimality condition is satisfied:

$$d(j) > d(i) + c_{ij} \tag{15}$$

where $d(j)$ is the distance label at node j and c_{ij} is the cost of moving from node i to j.

However, if there are several nodes j, such that each node satisfies this condition, the next node added to the shortest path is selected sequentially. To investigate the impact of this constraint on elasticity, we modify Algorithm 2 to relax the sequential constraint by randomly selecting the next node j that will be added to the shortest path. Algorithm 3 reflects these changes.

Algorithm 3 Dijkstra's "Modified" algorithm

begin
$S := 0; \bar{S} := N; X = 0$
$d(i) := \infty$ for each node $i \in N$
$d(s) := 0$ and pred(s) $:= 0$
while $|S| < n$ **do**
begin
let $i \in \bar{S}$ be a node for which $d(i) = \min\{d(j) : j \in \bar{s}\}$
$S := S \cup \{i\}$;
$\bar{S} := \bar{S} - \{i\}$;
for each $(i, j) \in A(i)$ **do**
$X_i = j, \forall j \in N$ which satisfy the optimality condition
$j_{selected} = rand(X_i)$
then $d(j) := d(j) + c_{ij}$ and pred(j) $:= i$
end for
end

We conducted 100 sample runs and averaged elasticity for each network under highest degree and highest betweenness attacks. Our results show a negligible difference between the elasticity results for Algorithm 2 and 3. Hence, the rankings shown in Table 2 remain the same.

From the three algorithms, we select Dijkstra's algorithm, using a homogeneous traffic matrix, as the most feasible because it produces qualitatively comparable results to the other two algorithms and has the least costly running time: $O(n^2)$.

5. Experimental Results

In this Section, we evaluate elasticity for a set of selected topologies. First, we compute the elasticity of all networks under each attack strategy and second, we implement a tradeoff function that combines the elasticity obtained for each attack strategy and penalizes networks for having excess links.

5.1. *Elasticity of Networks Under Three Attack Strategies*

In the subsequent sections, Elasticity R, Elasticity D, and Elasticity B refer to elasticity under the following three attack strategies:

(i) removal of random nodes (Elasticity R)
(ii) removal of highest degree nodes (Elasticity D)
(iii) removal of highest betweenness nodes (Elasticity B)

Table 3 ranks all networks in descending order of magnitude based on the number of links and the scores for elasticity under the three strategies. As shown, the mesh network is the most robust under all strategies. This is expected, as it sets the upper bound on elasticity. Under random attacks, the elasticity for the Gi-dense and MySpace networks are in proximity to that of the mesh network. As cost is a critical factor in network design, it is financially sensible to implement the latter two topologies rather than the mesh because Table 3 shows that the MySpace and Gi-dense networks can provide about 94% of the elasticity that the mesh provides while only using about 1% of the links.

Table 3
Networks sorted in descending order for number of links, Elasticity R (Elas. R), Elasticity D (Elas. D), and Elasticity B (Elas. B)

Nets.	**links**	**Nets.**	**Elas. R**	**Nets.**	**Elas. D**	**Nets.**	**Elas. B**
mesh	499500	mesh	0.3333	mesh	0.3333	mesh	0.3333
MySpace	10976	MySpace	0.3119	n-r 2	0.2426	Gi-dense	0.2390
Gi-dense	4505	Gi-dense	0.3111	Gi-dense	0.2082	W-S 2	0.1770
n-r 2	3781	PA 2	0.2743	MySpace	0.1721	MySpace	0.1719
W-S 2	3000	W-S 2	0.2703	W-S 2	0.1640	W-S 1	0.1260
PA 2	2964	PA 1	0.2677	n-r 1	0.1342	Gi-sparse	0.1010
Gi-sparse	2009	Gi-sparse	0.2520	W-S 1	0.1170	PA 2	0.0719
W-S 1	2000	W-S 1	0.2490	Gi-sparse	0.1143	PA 1	0.0558
PA 1	1981	n-r 2	0.2316	PA 2	0.0644	YouTube	0.0332
n-r 1	1921	Flickr	0.2211	PA 1	0.0535	Flickr	0.0315
YouTube	1576	YouTube	0.2132	YouTube	0.0371	n-r 2	0.0246
Flickr	1515	meshcore	0.1997	Flickr	0.0285	n-r 1	0.0178
meshcore	1275	HOT 2	0.1623	HOT 1	0.0129	meshcore	0.0083
HOT 2	1049	PA-sparse	0.1537	HOT 2	0.0095	HOT 1	0.0059
PA-sparse	1049	HOT 1	0.1405	Abilene	0.0093	HOT 2	0.0048
ringcore	1000	ringcore	0.1290	meshcore	0.0083	PA-sparse	0.0039
HOT 1	988	Abilene	0.1280	PA-sparse	0.0045	Abilene	0.0031
Abilene	896	n-r 1	0.1016	ringcore	0.0040	ringcore	0.0026

The subsequent Subsections show correlations for elasticity under the specified attack strategy.

5.1.1. *Correlation Between Elasticity and Number of Links*

From Table 3 it is notable that for all removal strategies the MySpace, Gi, PA, and Watts-Strogatz networks all vie for the highest elasticity. This phenomenon can be explained by considering the large number of links of these networks. Figures 14, 15, and 16 confirm this propensity and depict elasticity under random, targeted, and highest betweenness attacks respectively. In these figures, each network is assigned a symbol representative of two classes of networks: 1) The heterogeneous class with graphic or unshaded symbols represents networks with a power-law distribution, and 2) the semi-regular class, further broken down into deterministic and random networks, are the blocked, shaded symbols and is indicative of networks with a Poisson degree distribution. Furthermore, each symbol within a class can be one of two sizes: The large symbols correspond to the networks shown in "Full caps" in the legend and the small symbols correspond to networks in "Lower caps." These Figures show that the tendency for elasticity to increase as the number of links increase is not always the case. Thus, a large number of links is not a necessary condition even if it is a sufficient condition for high elasticity.

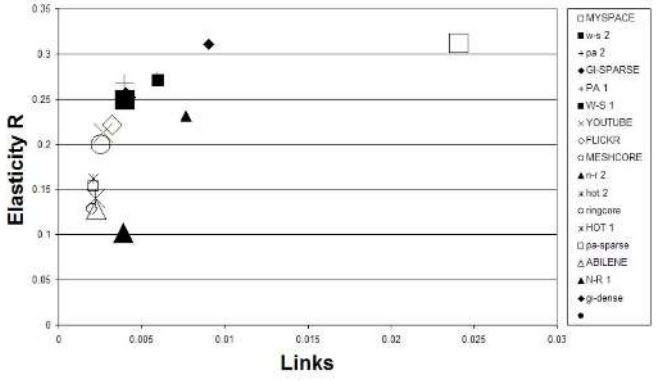

Fig. 14. Elasticity R vs number of links for each network in Table 3

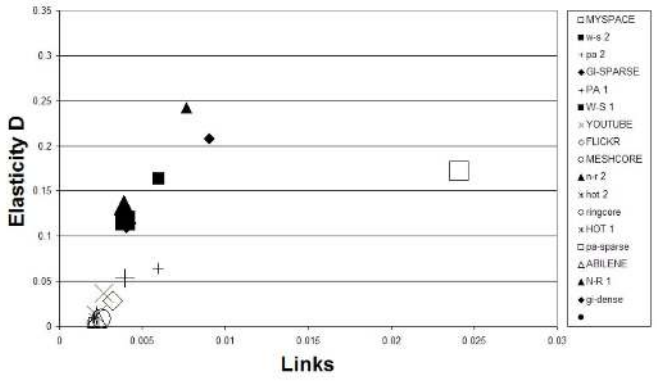

Fig. 15. Elasticity D vs number of links for each network in Table 3

Table 3 shows that under random attack, elasticity can be as low as 30.5% of the upper bound. This sharply declines to

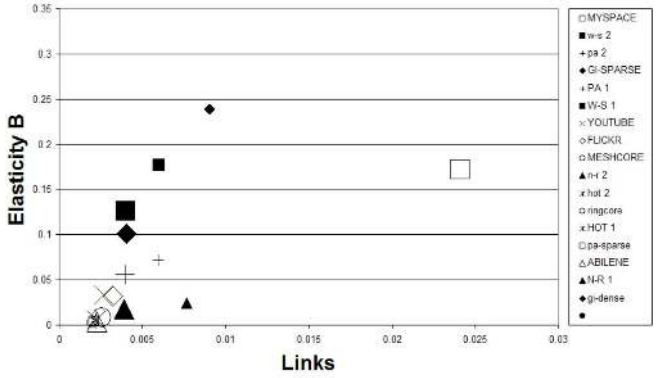

Fig. 16. Elasticity B vs number of links for each network in Table 3

1.2% for highest degree attacks and 0.8% for highest betweenness attacks. For this reason, the design of a robust topology is of utmost importance to obtain high elasticity. For example, the HOT 1 and PA-sparse networks have the same number of links, the same number of nodes, and almost identical degree distributions. However, their response to attacks differ [20]. Under random attacks, the PA-sparse provides 9.76% more elasticity than the HOT 1 network. In the PA-sparse network, low degree nodes outnumber high degree nodes (hubs) and hence, the probability of attacking hubs is lower than that of attacking other nodes. This is also the case for the HOT 1 network. However, the ratio of low degree nodes to hubs is higher in the PA-sparse network than in the HOT 1 network. As a result, Elasticity R for the PA-sparse network is higher than the HOT 1 network. For highest degree attack the elasticity provided by both networks decreases. However, the HOT 1 topology provides about three times Elasticity D as the PA-sparse network. The PA-sparse network is more susceptible to this attack because the hubs in this network facilitate interconnection and are vital to the elasticity of the network. However, the hubs in the HOT 1 network are located on the periphery and are less critical to interconnections [18].

For highest betweenness attack, the elasticity of both networks decreases even more. It is notable that from highest degree to highest betweenness attack, the elasticity provided by HOT 1 exhibits a 54.3% decrease whereas that provided by PA 1 exhibits a much smaller decrease of 13.3%. This can be interpreted from Figures 17 and 18 that show the betweenness distribution for the PA-sparse and HOT 1 networks. For the PA-sparse, nodes with the highest degrees have the highest betweenness. Thus, damage incurred under highest betweenness attacks is almost similar to that under highest degree attack. However, for the HOT 1 network there is a large decrease in elasticity from highest degree attack to highest betweenness attacks because nodes with highest betweenness tend to have lower degrees and facilitate interconnection within the network. Thus, attacks on these nodes are more detrimental than high degree attacks.

5.1.2. *Correlation Between Elasticity and Heterogeneity*

Figures 19, 20, and 21 illustrate the effect of heterogeneity on the elasticity of a network. The interpretation of these Figures is

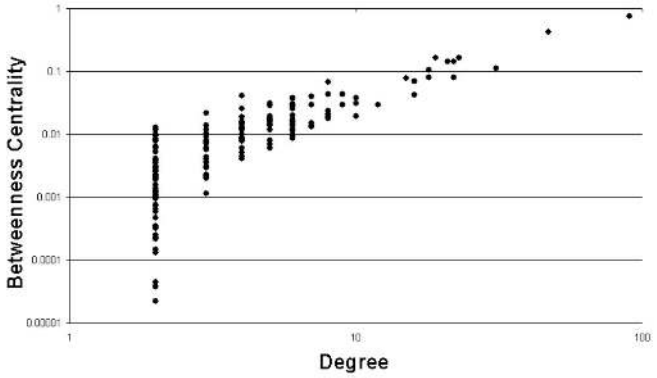

Fig. 17. Betweenness distribution for the PA-sparse network

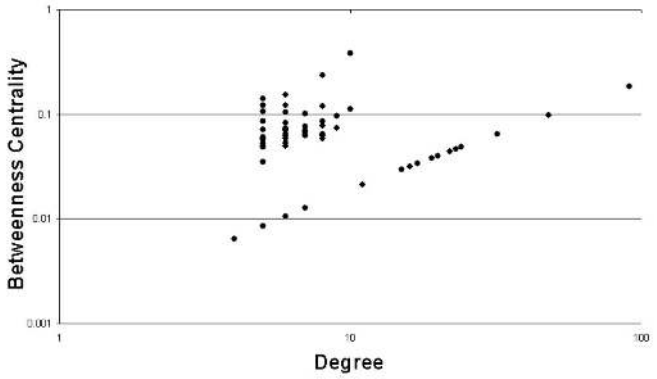

Fig. 18. Betweenness distribution for the HOT 1 network

that homogeneous networks have a proclivity for higher levels of elasticity. These include the variations of Watts-Strogatz's small world models, the random models, and the near-regular models. The implications of these results are far reaching where network structure is concerned. For example, the W-S 2 network is a representative of the random, semi-regular class of topologies where the majority of nodes tend to have a degree close to the average degree. Therefore, the damage incurred under highest node degree and highest betweenness attacks is comparable. For example, from Table 3 W-S 2 has elasticity scores of 0.164 and 0.177 for highest degree and highest betweenness attacks.

A representative of the deterministic, semi-regular networks, n-r 1 maintains its elasticity under random and high degree attacks. This result can be understood by the almost constant node degree. Thus, random attacks in addition to highest degree attacks, result in similar throughput degradation. However, as nodes are removed under highest betweenness attacks, core nodes appear and are destroyed. For n-r 1, elasticity decreases considerably from highest degree attacks to highest betweenness attacks by 35%. Thus, although these topologies are sufficiently costly, in addition to the fact that they may fail to capture the properties of some real world networks, their topological structures offer remarkable resilience to attacks.

Figures 22 and 23 compare the degree distribution for W-S 2, a representative of the semi-regular class of networks, and Abilene, a representative of the heterogeneous class of networks. As discussed previously, the almost constant degree for

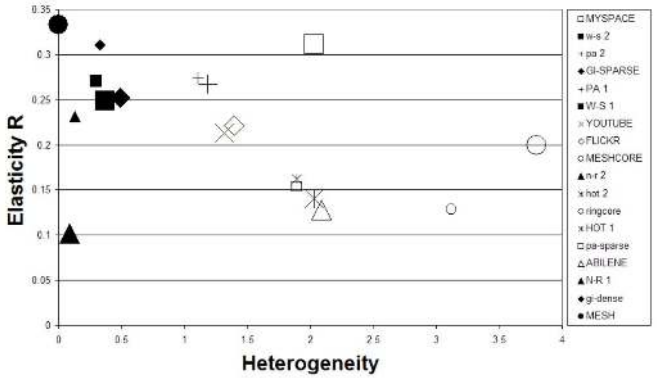

Fig. 19. Elasticity R vs heterogeneity for each network in Table 3

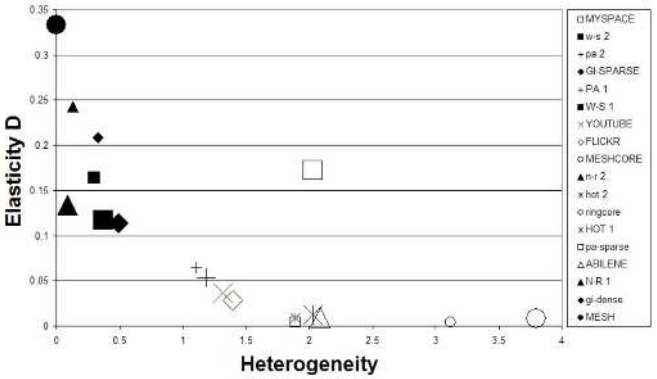

Fig. 20. Elasticity D vs heterogeneity for each network in Table 3

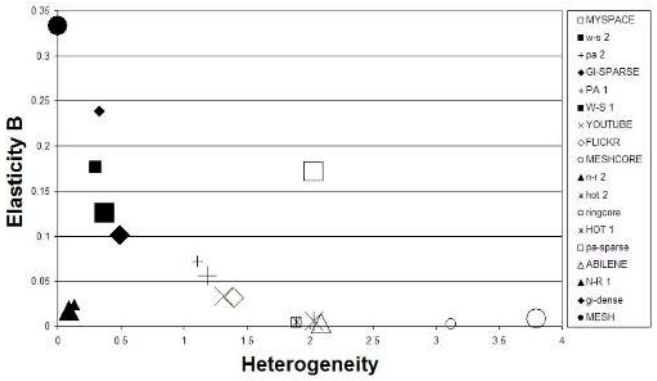

Fig. 21. Elasticity B vs heterogeneity for each network in Table 3

the semi-regular class results in high elasticity scores. However, heterogeneous networks span a wide range of degrees and behave differently under attacks. More precisely, based on the "type" of heterogeneous network under investigation, the impact of highest degree attacks can vary. On the one hand, networks like Abilene avoid cataclysmic damage under high degree attack because the hubs are located on the periphery of the network and thus, highest degree attack has minimal effect on the overall operation of this network. However, heterogeneous networks like PA-sparse are severely damaged because the hubs are critical and hold the network together. Thus, homogeneity has far reaching implications in the robustness of networks.

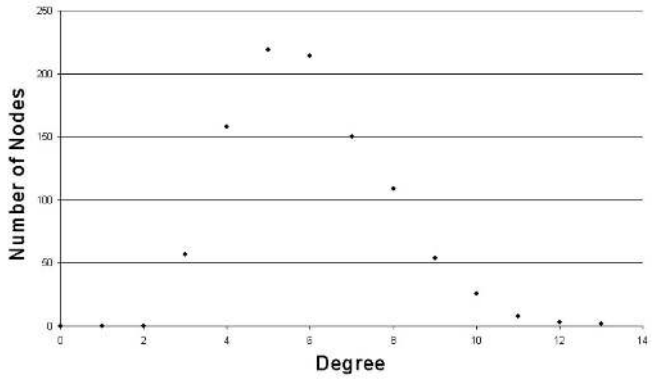

Fig. 22. Node degree distribution of Watts-Strogatz 2 network

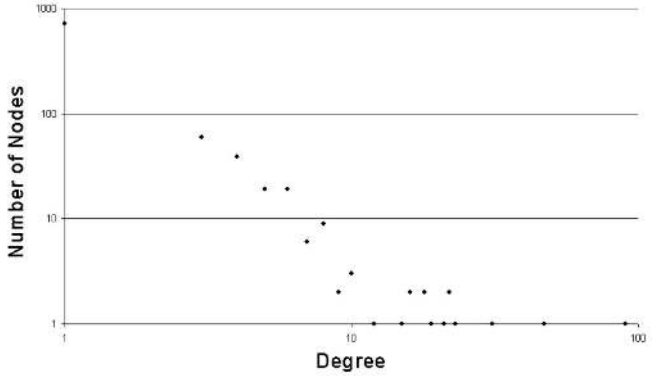

Fig. 23. Node degree distribution of Abilene network

5.1.3. *Correlation Between Elasticity and Characteristic Path Length*

The characteristic path length tells the expected distance, in number of hops, from a given source node s to a destination node t. Figures 24, 25, and 26 show that the characteristic path length tends to be negatively correlated with elasticity. This is not a necessary condition as these Figures provide instances where a network with high characteristic path length can have a higher elasticity than a network with a smaller characteristic path length. However, if the number of nodes in a given network is kept constant as the number of links increase, path diversity will eventually increase. As a result, network congestion decreases which ultimately increases elasticity.

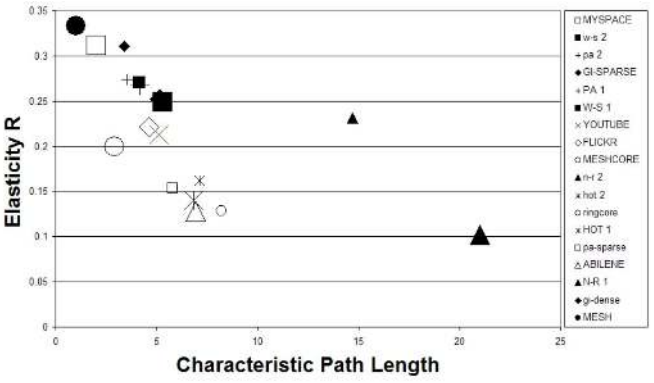

Fig. 24. Elasticity R vs characteristic path length for each network in Table 3

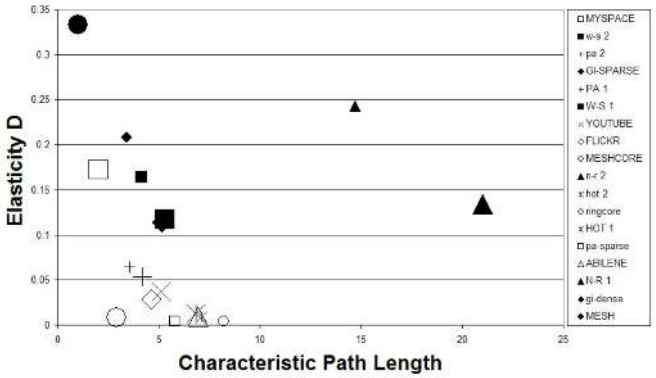

Fig. 25. Elasticity D vs characteristic path length for each network in Table 3

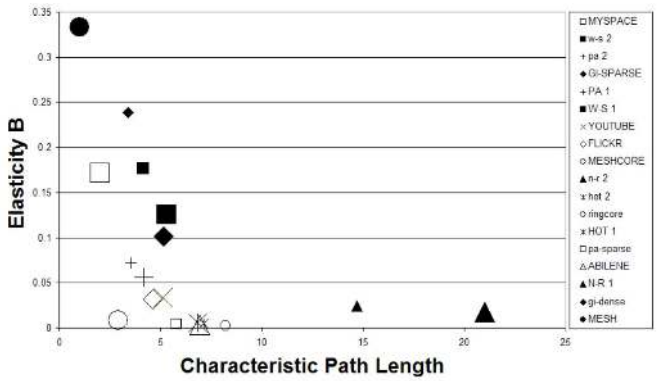

Fig. 26. Elasticity B vs characteristic path length for each network in Table 3

5.2. *Elasticity of Networks with Tradeoff Function Applied*

To compensate for the tradeoff between elasticity and number of links, we introduce a tradeoff function $Re\,(G)$ that provides robustness with respect to elasticity. For a given network G, our robustness measure can be computed as

$$Re\,(G) = \alpha A + \beta B + \delta C - \gamma density' \tag{16}$$

where A, B, C, represent Elasticity R, Elasticity D, and Elasticity B. $0 \leq (\alpha, \beta, \delta, \gamma) \leq 1$, $0 \leq density' \leq 1$ and $density' = 1 - e^{-\frac{1}{2}\frac{(M-(N-1))}{N}}$. The $\frac{1}{2}$ factor determines the rate at which $density'$ changes. α, β,δ, and γ are tolerance parameters and as such, represent the tolerance of a network towards random, targeted, and highest betweenness attacks. M is the total number of links and $M-(N-1)$ represents the number of excess links in a network: these are links which exceed the threshold necessary to obtain 1 connected component with N nodes.

This function facilitates independence for constructing networks based on a projected need. Thus, a network engineer who envisions persistent, random attacks would consider a high value of α. Similarly, β or δ would dominate where targeted attacks or highest betweenness attacks respectively are predominant. Moreover, γ could be varied based on financial constraints.

Table 4 depicts the rankings of each topology with their respective number of links and Re scores. For each network,

Table 4
Ranking of networks after implementing the cost function Re

Networks	Number of links	Re
HOT 2	1049	0.1519
PA-sparse	1049	0.1374
ringcore	1000	0.1351
Abilene	896	0.1342
HOT 1	988	0.1330
W-S 1	2000	0.0982
meshcore	1275	0.0874
YouTube	1576	0.0828
Gi-sparse	2009	0.0708
Flickr	1515	0.0340
PA 1	1981	-0.0110
W-S 2	3000	-0.0210
Gi-dense	4505	-0.0684
near-regular 1	1921	-0.1206
PA 2	2964	-0.2150
near-regular 2	3781	-0.2561
MySpace	10976	-0.3388

we obtained Re for tolerance values of $\alpha = \beta = \gamma = \delta = 1$. The common tolerance values facilitate an unbiased analysis of robustness by providing equal likelihood of occurrence to each attack strategy. In addition, these rankings represent the case where networks are completely penalized for having excess links and as a result, the structure of the network plays a more significant role to determine the robustness of the network.

From this analysis, HOT 2 was the highest ranked network. This network has only 50 excess links and thus, it virtually avoids the penalty for the existence of excess links. Furthermore, though it exhibits power-law properties, the hubs are located on the periphery of the network and hence, HOT 2 has an admirable structure against targeted attacks but becomes vulnerable under highest betweenness attacks. However, considering the values for the tolerance parameters and number of links discussed previously, the HOT 2 network is the most suitable.

5.3. *Tradeoff Between Characteristic Path Length and Heterogeneity*

The ideal network to provide high elasticity tends to exhibit a low score for heterogeneity and a short characteristic path length. In all networks, the mesh has the shortest characteristic path and the lowest score for heterogeneity and hence, it features the highest elasticity. However, this high elasticity comes at a very high cost which network designers are unwilling to consider. For this reason, it is imperative to consider a tradeoff between a short characteristic path length and a low score for heterogeneity. Figure 27 shows a plot of heterogeneity against characteristic path length. The colorbar (to the left) provides the third dimension to this plot of elasticity. This plot can be interpreted as a decrease in the characteristic path length such that the network becomes more homogeneous increases elasticity.

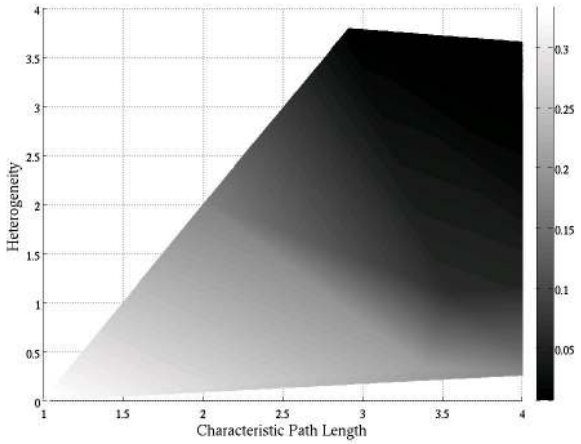

Fig. 27. Elasticity increases as characteristic path length and network heterogeneity decrease

6. Conclusions and Outlook

This paper endeavors to extract the characteristics of robust complex networks. As our measure for robustness, we used elasticity, which measures the ability of a network to maintain its total throughput under increasing removal of nodes with respective links, and theoretically derived its upper bound of $\frac{1}{3}$. We then illustrated its utility on 18 networks from six different network models under random, highest degree, and highest betweenness attacks and then implemented a tradeoff function which computes robustness with respect to elasticity.

Elasticity is defined and computed under simple assumptions. As an example, it is dependent on the routing algorithm used, which can perhaps alter current network rankings. However, elasticity provides benefits which are far-reaching. More precisely, it identifies key characteristics of robust complex networks: A short characteristic path length, low heterogeneity, and strategically located links to facilitate a "homogeneous" core such that if hubs should be added, they should be placed on the periphery of the network to provide added resilience against targeted attacks.

For our future work, we intend to incorporate expander graphs in our evaluation and formulate a working definition of the core and periphery to include details about the size and characteristics. Armed with this knowledge, we seek to combine particular graphs to determine the essential components to increase elasticity. Finally, we will develop heuristics to build graphs such that elasticity is maximized.

7. Acknowlegements

This research was supported by National Science Foundation (NSF) under award number 0841112. We would like to thank Dr. Robert Kooij for contributing to the success of this work.

References

[1] ITPRO. [Online]: http://www.itpro.co.uk/100966/taiwan-quake-exposes-weakness-of-undersea-data-lines, 2006.

[2] CNN. [Online]: http://archives.cnn.com/2001/TECH/2001.

[3] BBC. [Online]: http://news.bbc.co.uk/2/hi/technology/3682537.stm, 2004.

[4] Computer Security. [Online]: http://ftp.fas.org/irp/gao/aimd-96-108.htm, 1996.

[5] US Canada Power System Outage Task Force. August 14th blackout: Causes and recommendations, (2003).

[6] Swiss Federal Office of Energy. Report on the Blackout in Italy on September 28 (2003).

[7] H. Frank and I. Frisch, "Analysis and design of survivable networks," in *IEEE Transactions on Communications Technology COM-18*, p. 567, 1970.

[8] A. Jamakovic and S. Uhlig, "Influence of the network structure on robustness," in *Networks, 2007. ICON 2007. 15th IEEE International Conference*, pp. 278–283, 2007.

[9] D. Bauer, F. Boesch, C. Suffel, and R. Tindell, "The theory and application of graphs," pp. 89–98, 1981.

[10] F. Harary, "On conditional edge-connectivity of graphs," in *Networks*, vol. 13, p. 346, 1981.

[11] A. H. Esfahanian and S. Hakimi, "On computing a conditional edge-connectivity of a graph," in *Journal of Information processing Letters*, vol. 27, p. 195, 1988.

[12] M. Krishnamoorth and B. Krishnamirthy, "Fault diameter of interconnection networks," in *Computers and Mathematics with Applications*, vol. 13, p. 577, 1987.

[13] V. Chvatal, "Tough graphs and hamiltonian circuits," in *Discrete Math*, vol. 5, p. 215, 1973.

[14] H. Jung, "On a class of posets and the corresponding comparability graphs," in *Journal of Combinatorial Theory B*, vol. 24, p. 125, 1978.

[15] M. Cozzen, D. Moazzami, and S. Stueckle, "Seventh international conference on the theory and applications of graphs," p. 11111122, 1995.

[16] M. Alon, "Eigenvalues and expanders," in *Combinatorica*, vol. 6, p. 83, 1986.

[17] B. Mohar, "Isoperimetric numbers of graphs," in *Journal of Combinatorial Theory Series B*, vol. 47, p. 274, 1989.

[18] R. Albert and H. J. A.-L. Barabasi, "Error and attack tolerance of complex networks," in *Nature*, vol. 406, pp. 378–382, 2000.

[19] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Process on Complex Networks*. Cambridge University Press, 2008.

[20] D. Alderson, L. Li, and C. D. W. Willinger, "Understanding internet topology: Principles, models, and validation," in *IEEE/ACM TRANSACTIONS ON NETWORKING*, vol. 13, pp. 1205–1218, 2005.

[21] A. Sydney, C. Scoglio, P. Schumm, and R. Kooij, "Elasticity: Topological characterization of robustness in complex networks," in *IEEE/ACM Bionetics*, 2008.

[22] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Hauffaker, X. Dimitropoulos, K. Claffy, and A. Vahdat, "Lessons from three views of the internet topology," tech. rep., CAIDA, 2005.

[23] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat, "Systematic topology analysis and generation using degree correlations," in *ACM/SIGCOMM*, vol. 47, p. 274, 2006.

[24] J. Dong and S. Horvath, "Understanding network concepts in modules," in *BMC Systems Biology*, vol. 1, 2007.

[25] D. L. Alderson, "Catching the network science bug: Insight and opportunity for the or researcher," in *INFORMS*, vol. 56, p. 10471065, 2008.

[26] NWBTEAM. Network Workbench Tool, Indiana University, Northeastern University, and University of Michigan. http://nwb.slis.indiana.edu, 2008.

[27] A.-L. Barabsi and E. Bonabeau, "Scale-free networks," in *Scientific American*, vol. 288, pp. 60–69, 2003.

[28] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley, "Optimization of network robustness to waves of targeted and random attacks," in *PHYSICAL REVIEW E*, vol. 71, 2005.

[29] A. Fabrikand, E. Koutsoupias, and C. Papadimitriou, "Heuristically optimized trade-offs: A new paradigm for power laws in the internet," in *ICALP*, vol. 2380, pp. 110–122, 2002.

[30] P. Mahadevan, C. Hubble, D. Krioukov, B. Huffaker, and A. Vahdat, "Orbis: Rescaling degree correlations to generate annotated interet topologies," in *ACM/SIGCOMM*, 2007.