Network Design Methods for Mitigation of Intentional Attacks in Scale-Free Networks

Walid K. Ghamry
Department of Information Systems Engineering
National Research Center
walidghamry@gmail.com

Khaled M. F. Elsayed
Department of Electronics and Communications Engineering
Cairo University
khaled@ieee.org

Abstract: Network robustness and network reliability are important issues in the design of Internet Service Provider’s topologies. In this paper, we examine the structural characteristics of network topologies that affect robustness and reliability. We examine the interplay between the structural characteristics of network topologies and the resource capacity over-provisioning strategies when the network breakdowns subject to practical constraints (router technology) and economic considerations (link costs). We study the robustness of the Internet connectivity under node intentional harmful attack using two attack strategies: static degree-based and static load-based. We find that the robustness of network topologies is affected by the variation of their structural characteristics. In our proposed approach, we show that highly-heterogeneous topologies have less robustness compared with lightly-heterogeneous topologies. The observations from the robustness study provide us useful insights for proposing multiple efficient preventive resource capacity over-provisioning strategies for mitigation of intentional attacks. The proposed strategies utilize the structural properties by calculating the excess traffic in case of single global cascading failure for each node and measure its influence on the other nodes as well as locally. The results show that our proposed strategies can significantly enhance the robustness and increase the resilience of network topology. Due to our proposed approach results, we also show that highly-heterogeneous topologies have high resilience compared with lightly-heterogeneous topologies. By using real data from the Sprint network at the router level, we provide further empirical evidence in support of the proposed approach.

Keywords: resource over-provisioning, network robustness, network resilience, mitigation of attacks, scale-free networks.

1 Introduction

Recently, there has been increased interest in studying large-scale real-world systems including the Internet, World-Wide Web (WWW), protein-protein reactions, and social networks. When formulating these systems into network models, it is found that they share some common features. Such observations have spawned a new research area called complex networks [1]. The large-scale topological structure of the Internet can be made on routers in the router-level graph [2] or entire subnetworks (Autonomous Systems) in the AS-level graph [3]. The most important complex network model is the scale-free network [1] in which the nodal-degree distribution is described as

$$P(k) \propto k^{-g}, \qquad (1)$$

where $P(k)$ denotes the fraction of nodes with degree k, g is the exponent (the slope of the line that represents the power law scaling relationship between the node degree k and its node rank).

The main focus of research on information transportation in large scale networks is the efficient flow of information. The efficient performance for communication systems is affected by the system reliability (the ability of the system to maintain end-to-end paths in the presence of node or link losses and to return to its normal state after a disturbance as quick as possible). It is desirable to maximize the network robustness (*the ability to withstand strains and to maintain its functionality in some way*) and resilience (*the ability to return to its normal state after a disturbance*) while minimizing the loss of information.

The most important results for scale-free networks are that they are highly robust against random failures and they are fragile under intentional attacks that bring down network nodes in a decreasing order of their nodal degrees [4, 5]. In this paper, we propose various resource capacity over-provisioning methods to increase robustness and resilience of the network under attack. Also, we attempt to answer the questions: how does the network reliability depend on the network topology and how does it vary with the interplay between the resources capacity over-provisioning strategies and the network structure, and is there any relation between the efficacy of the proposed resource over-provisioning strategies and the network structure?

This paper is organized as follows: In section 2 we present previous work related to our study. Section 3 provides an overview of network graph theory, presents the network topology models used in our study, and lastly describes the network measurement metrics for our work. In section 4 we present the first part of our study focusing on attack strategies and network robustness. We propose our resource over-provisioning capacity methods for network robustness and reliable communication in section 5. Section 6 presents a performance evaluation study of the proposed methods. Section 7 concludes our study.

2 Related Work

The first numerical studies on the robustness of real networks are those reported in [4]. Albert et al. have studied how the properties of the Internet and of a sample of the World Wide Web change when a fraction of nodes are removed. The network performance is measured by the size of the largest connected components. They found that the Internet and the WWW break into small fragments for a threshold fraction of node removal under attack (i.e. targeted node removal related to its degree). Under failures, giant component persists for high rates of random node removal for both the Internet and the World Wide Web.

Crucitti et al. [6] have used the concept of network efficiency at both global and local scales to study the effects of errors and attacks on Barabasi-Albert (BA) structured scale-free networks. Global and local efficiency are measured by the average inverse geodesic length. The numerical results indicate that both the global and the local efficiency of scale-free networks are unaffected by the removal of some (up to 2%) of the nodes chosen at random (under failure), but rapidly decrease when the nodes removed are those with the higher connectivity (under attack).

Magoni [7] propose an in-depth study of the Internet topology robustness to attacks at the network layer. Several attacking techniques (static and dynamic) related to node importance (degree and the size of the largest connected components) are presented. Also, a study of their effects on the connectivity of the Internet is presented. New metrics are introduced related to the size of the largest connected components (classifying them into clusters related to their size) which evaluate the amount of fragmentation in the network. He shows that although the removal of a small fraction of nodes (less than 10%) can damage the Internet connectivity, such a node removal attack would still require a large amount of work to be carried out in terms of detailing interactions between the intradomain and interdomain levels of the Internet.

Doyle et al. [8] argue that the high-degree nodes on the router level of the Internet are mostly edge nodes with a large number of low-capacity connections. Thus, loss of edge routers disconnects only low-bandwidth users and has no other effect on the overall connectivity and robustness of the network. The authors did not study the loss of high load routers which are distributed through the gateway and core levels whose loss result in more harmful effect on the overall connectivity.

Holme et al. [9, 10] have proposed a model for breakdowns triggered by changing nodes [9] or edges/links[1] [10] load in an evolving network (growing with time and not fixed). The work in [9] used the BA model with the preferential attachment and investigated the development of networks where the nodes and/or edges might break down due to overload. The load was defined in terms of the betweenness centrality. The mechanism is that the nodes with the highest loads are also the nodes with the highest degrees, and accordingly removing such nodes will thus decrease the average degree and increase the betweenness on other nodes maximally. They found that, in large time equilibrium, the network consists of many isolated chain-like clusters and for the network to be connected the capacity of the nodes to relay connections have to increase with the size of the network. The same conclusion was found by [10] in case of edge removal.

Holme et al. [11] extend the study of [9, 10] to attacks on both nodes and edges using two different importance measures. The first is the node degree while the second is the betweenness centrality. The edges were then removed in decreasing degree order, while the number of nodes remained constant. Since only one edge is removed at every time step, edge attacks should not be as harmful as node attacks. Holme et al. [11] consider both cases when the distributions are recalculated after every removed node or edge (dynamic model), and the case when the information is limited to the initial distributions (static model). The latter may sometimes be the more realistic scenario, as information on the structure of the network may be hard to obtain and the measures, the betweenness in particular, may be expensive to calculate.

[1] We use edge or link interchangeably to refer to the physical connection between nodes in the network.

There are preventive strategies which enhance network robustness before attack happens. Motter et al. [12] introduce a simple model for cascades of overload failures. The model shows that even a small fraction of highly loaded nodes can trigger global cascades in networks with heterogeneous distribution of loads. It is based on the assumption that each node is characterized by a finite capacity, defined as the maximum load that the node can handle. The capacity is assumed to be proportional to the initial load of the node with the tolerance parameter of the network representing the ability of nodes to handle the increased load thereby resisting perturbations. This is classified as an over-provisioning method.

Xiao et al. [13, 14] show that incomplete global information has different impacts on the intentional attack in different circumstances, while local information-based attacks can be actually highly efficient. They study two different effects: (i) the accurate intentional attack is practically impossible; and (ii) most attacks propagate locally. The proposed method includes hiding or partially hiding hub nodes. They show that their insights would be helpful to develop efficient network protection schemes.

Beygelzimer et al. [15] present empirical results that show how robustness, as measured either by the size of the largest connected components or by the average path length is affected by several different strategies that alter the network by rewiring a fraction of the edges or by adding new edges. They enhance the network robustness by link insertion between low-degree nodes without changing its nodal-degree distributions. Xiao et al. [16] consider two different rewiring methods that change the assortative mixing of the network. They study the effects of connections patterns on the robustness of scale-free networks and show that the scale-free networks have different levels of robustness.

There are reactive strategies which take immediate actions during or after the attack to mitigate the damage: Xiao et al. [17] propose a simple local repairing strategy by restoring some of the links that have been cut when a network node is crashed by rewiring each of them to another node. Their results show that their simple strategy can enhance network robustness. Another work similar to this work was introduced by Wang et al. [18].

In this paper, it is the first time to study the effects of various resource capacity over-provisioning methods (preventive strategies) on network topologies having the same node degree sequence, subject to practical considerations (router technology) and economic considerations (link costs). We use two different static attack techniques: the static high degree (SHD) and the static high load (SHL) attacks.

We classify the nodes of the network topology models subject to their locations in the network topology: hosts, access routers, gateway routers, and backbone core routers. We propose resource capacity over-provisioning methods that utilize the network structural properties by calculating the increase of traffic in case of single global/local failures for each node. The proposed strategies can be applied to any combination of network topology and associated routing methods.

3 Background and Foundations

3.1 Network Graph Theory and Network Topology Models

In this section we provide some necessary definitions related to graphs and network topology models used subsequently in the paper.

3.1.1 Network Graphs

A graph is a mathematical model to represent networks that have a certain structure (or topology) and can have additional quantitative information. The structure might be directed or undirected. Quantitative information about types, weights or other attributes for nodes and edges might exist.

An undirected graph G_U is defined by a pair of sets $G_U = (V , E)$, where V is a non-empty countable set of elements, called nodes or vertices and E is a set of unordered pairs of different nodes, called edges or links. A directed graph D, or digraph, is defined by a non-empty countable set of nodes V and a set of ordered pairs of different nodes E_D that are called directed edges.

Many real networks display a large variability in the intensity (e.g. link capacity in a communication network) values of edges. Therefore, it is desirable to go beyond the mere topological representation and construct a weighted graph where each edge is associated with a weight representing the intensity or capacity of the edge/link.

3.1.2 Network Types

Networks types can be divided into three classes: homogeneous, heterogeneous, and sparse. Homogenous networks mean that all nodes have a similar number of links. Homogeneity in the topology structure means that almost all nodes are topologically equivalent, like in regular lattices or in random graphs. In these latter ones, the degree distribution is binomial or Poisson in the limit of large graph size (peaked around the average value).

Heterogeneous networks are characterized by having a highly inhomogeneous degree distribution which results in the simultaneous presence of a few nodes (the hubs) linked to many other nodes, and a large number of poorly connected elements. A lightly heterogeneous network has a large number of high-degree nodes connected to a significant fraction of all nodes in the network. A highly heterogeneous network has a small number of high-degree nodes connected to a significant fraction of all nodes in the network. The degree distribution follows a power-law.

A sparse network has an average degree that is much smaller than the size of the network, that is, $k << m$ where k is the average node degree and m is the total number of nodes.

3.1.3 Network Topology Models

We rely on measurements on data of five networks constructed explicitly to have the same node degree sequence. Figure 1 depicts these five networks and its constituents are explained as follows:

(a) The power-law type degree sequence of all five networks.

(b) A graph constructed from the Preferential Attachment (BA) model [19].

(c) A construction based on the General Random Graph (GRG [20]) or Power-Law Random Graph (PLRG [21]) method by using the degree sequence of the BA network as the expected node degree to generate a random graph using the GRG method.

(d) Heuristically Optimal Topology (HOT [22]): constructed by using a heuristic, nonrandom, degree-preserving rewiring of the links and routers in the BA graph to produce a network having a mesh-like core with hierarchical aggregation from the edge to the core.

(e) Sub-optimal Topology [22]: heuristically designed network that has been intentionally constructed to have poor throughput and purpose for comparison.

(f) Abilene-Inspired Topology [22]: Inspired by the publicly available actual data of the Abilene network [23].

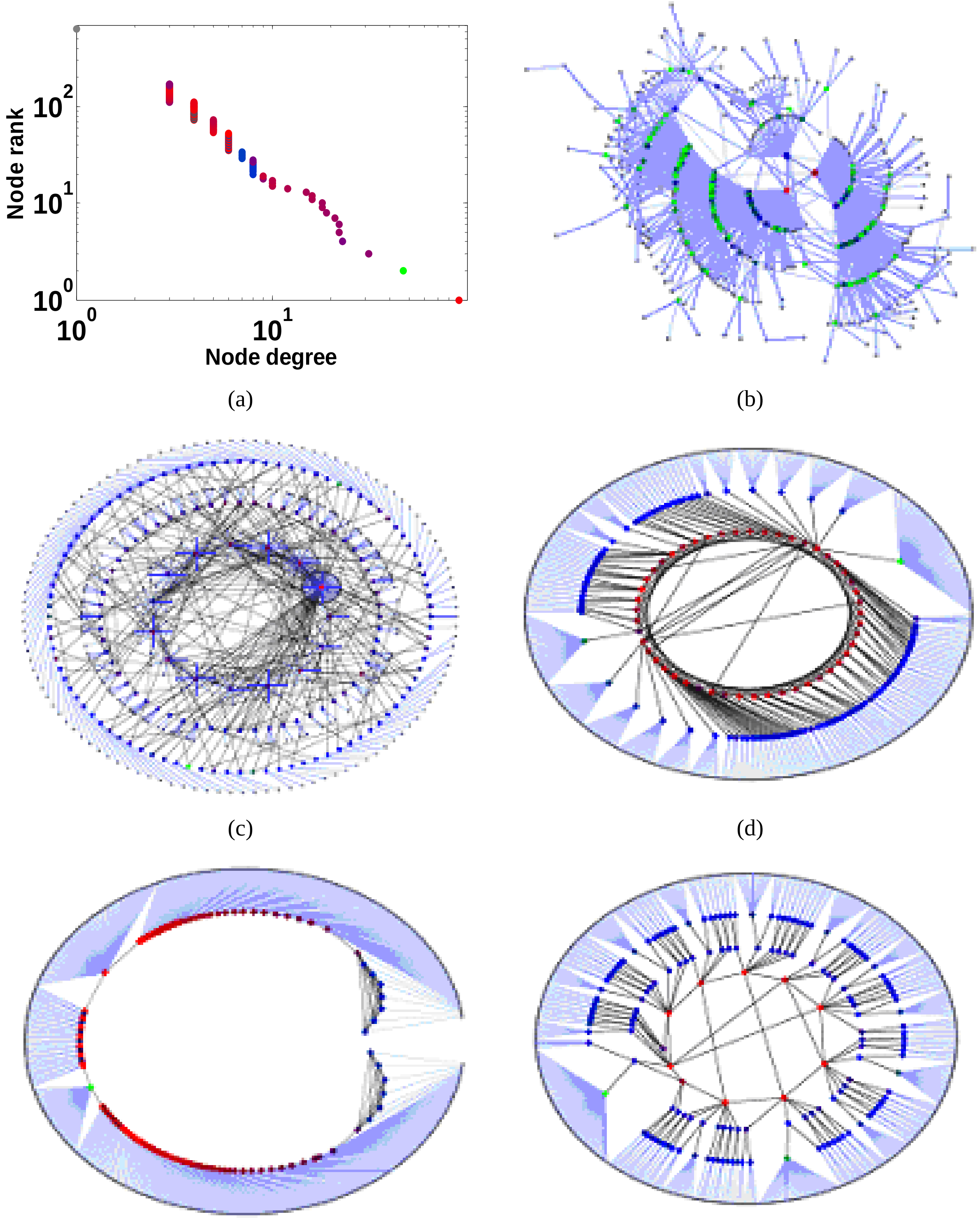
These five topologies are available from [24] and constructed by Li et al. [22]. Our empirical evidence in support of the proposed approach for our work relies on a network topology at the router level for real data of the Sprint network [25]. The Sprint network is composed of 17 PoPs scattered throughout the US. The Sprint network with 466 nodes and 1283 edges is shown in Figure 2.

3.2 Network Measurement Metrics

Based on the technology used in the cross-connection fabric of the router itself, a router has a maximum number of packets that can be processed in any unit of time. This constrains the number of link connections (i.e. node degree) and connection speeds (i.e. bandwidth) at each model type. This results in what is called the “efficient frontier” of possible bandwidth-degree combinations available for each router as shown in Figure 3.

The Cisco 12416 GSR at the core has 16 line card slots and its maximum throughput (denoted as BW) is 320 Gbps (full duplex). When the number of ports is less than 16, the throughput of each port is limited by the maximum speed of supported line cards (10 GE or OC-192) and the router’s maximum throughput increases with the number of ports. When the number of ports is greater than 16 the maximum router throughput decreases.

At the edge, the Cisco 7500 GSR has 11 line card slots and its maximum throughput is 44 Gbps (full duplex). When the number of ports is less than 11, the throughput of each port is limited by the maximum speed of supported line cards (OC-24 (1.25 Gbps)) and the router’s maximum throughput increases with the number of ports. When the number of ports is greater than 11, the maximum router throughput decreases as the total number of ports increases up to a maximum of 27 possible direct connections. The Cisco 7600 GSR series aggregation router with 8 slots: (8 x 8) port OC12 channelized into OC3 (155 Mbps) can support up to a maximum of 256 possible connections with maximum throughput 40 Gbps [27]. Hence, when the degree of access routers at the edge network exceeds 27 we specify this router as Cisco 7600 GSR series with each connected link having OC3 link speed.



(a)

(b)

(c)

(d)

(e)

(f)

Figure 1: Five networks having the same node degree distribution. (a) Degree distribution (degree versus rank on log-log scale) (b) BA (c) GRG (d) HOT (e) SUB ; (f) Abilene real data network.

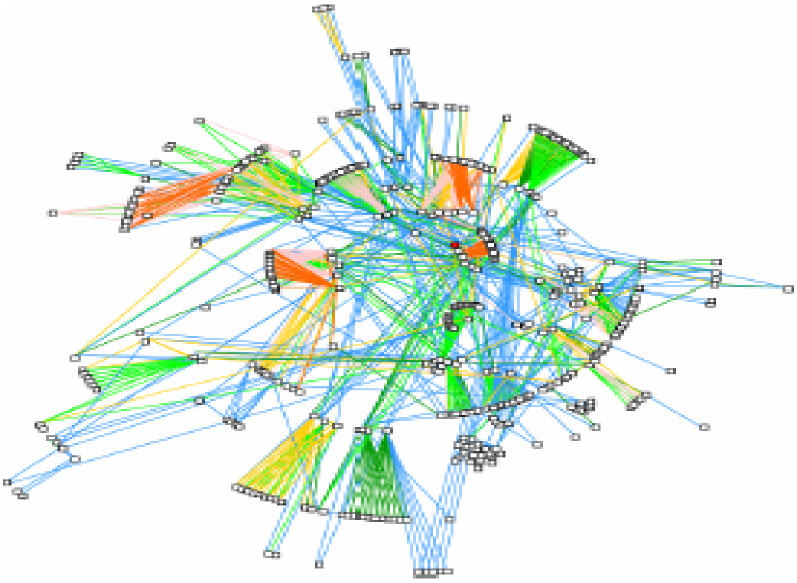


Figure 2: Sprint network using Pajek program [26]

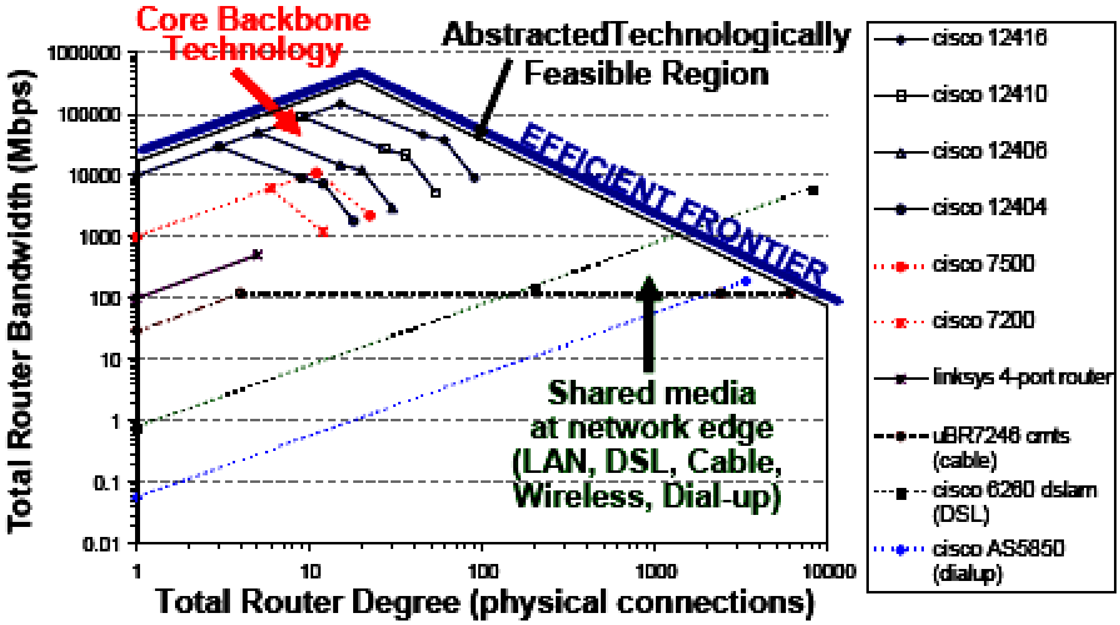


Figure 3: Aggregate picture of router technology constraints [22]

In our experiments we allocate the capacities of routers based on the technology constraints imposed by the Cisco 12416 GSR for all non edge routers and by the Cisco 7500 GSR and Cisco 7600 GSR series aggregation router at the network edge (access routers) [22, 27]. Due to a lack of publicly available information on traffic demand for each point, we assume the bandwidth demand at a router is proportional to the aggregated demand of any end hosts connected to it. That is, starting at the network edge and based on the gravity model [28] we consider the demand for traffic by an access router to be the aggregate connectivity bandwidth of its end hosts (nodes with degree one; a single computer connected directly to a router or an IP end-point address in a subnet belonging to a LAN and connected directly to a router). Then, to determine the flow of traffic across the network core, we consider flows on all source-destination pairs of access routers, such that the amount of flow $X_{i\,j}$ between source *i* and destination *j* is proportional to the product of the traffic demand X_i and X_j at end points *i* and *j*.

$$X_{i\,j} = r\ X_i X_j \tag{2}$$

$$\text{Subject to } RX \le B\,, \tag{3}$$

where r is some global constant and is otherwise uncorrelated from all other flows, R is the routing matrix (defined such that R_{kl} = {0;1} depending on whether or not flow *l* passes through router *k*) and X is a vector obtained by stacking all the flows $X_{i\,j}$ by equation (2), indexed to match the routing matrix R. We use shortest path (SP) routing to get the routing matrix, and define *B* as the vector consisting of all router bandwidths according to the degree bandwidth constraint (Figure 3).

3.2.1 Excess Traffic

The excess traffic is measured as the total amount of traffic exceeding the configured (not maximum) bandwidth constraint for the most loaded routers under breakdown. The amount of excess traffic is considered as the network tolerance (the ability of nodes to handle increased load traffic) against intentional attacks. A low value of excess traffic is a good indication of network robustness and resilience against failures. Thus, for each network, we target the worst-case (worst-case means highest degree or highest-load) node that has not yet been deleted based on the attack strategy. After each router loss, we compute the amount of traffic that exceeds the configured bandwidth for routers which can still be served by the remaining network, possibly after some re-routing through unharmed parts of the network using shortest path (SP).

3.2.2 Largest Connected Component (LCC)

The LCC is the maximum number of nodes that can connected in a given graph. This is a very interesting metric because it gives an upper bound on the number of nodes that can communicate. The damage caused by the intentional attack is quantified in terms of the relative size of the largest connected component

$$RLCC = n'/n, \tag{4}$$

where n and n' are the numbers of nodes in the largest component before and after the attack respectively.

3.2.3 Clustering Coefficient

A common feature of many complex networks is clustering. Node clustering is a measure of how well the neighbors of a given node are locally interconnected (how close a node’s neighbors are to forming a clique). Node clustering coefficient is defined as the ratio of existing edges E_i between its k_i neighbors to the maximum possible number of such edge connections $k_i(k_i-1)/2$. The node clustering coefficient is defined as

$$cc_i = \frac{2E_i}{k_i(k_i-1)} \tag{5}$$

The clustering coefficient of the whole network $\langle cc \rangle$ is the average of cc_i over all the nodes in the network

$$\langle cc \rangle = \frac{\sum_{i=1}^{m} c_i}{m}, \qquad (6)$$

where m is the total number of nodes. The clustering coefficient is a practical metric that measures the local robustness in the graph: the higher the local clustering of a node, the more connected are its neighbors, thus increasing the path diversity locally around the node.

4 Attack Strategies and Network Robustness

The attack strategies considered in this work are based on node intentional harmful attack. The intentional strategies are classified depending on whether they are based on degree or load (router utilization) and whether the order of node attack is determined statically or dynamically. The static high degree-based (SHD) attack is the method in which the target nodes are selected one by one in descending order of their initial high degrees. The static high load-based (SHL) attack is the method in which the target nodes are selected one by one in descending order of their initial loads. We use both strategies in our paper.

The metrics measured under components loss (nodes/links) is clearly not to maximize the throughput or router utilization, but to provide good level of reliability and robustness. The robustness can be defined as the ability of the network to maintain its functionality when nodes or links are damaged in some way.

4.1 Effects of Attack Strategies on Network Robustness

We use C++ and MATLAB to simulate the schemes and obtain the results. The average clustering coefficient $\langle cc \rangle$ for the five topologies is shown in Table 1. The clustering coefficients for HOT, Abilene, and BA have the smallest values and therefore the removal of any constant fraction of nodes disconnects most node pairs. The clustering coefficients for GRG and SUB are larger than the other topologies, and thus the removal of nodes disconnects less node pairs.

Table 1: Average clustering coefficient for all topologies

BA	Abilene	HOT	GRG	SUB
0.0006	0.0076	0.0012	0.07	0.22

Figure 4 and Figure 5 show the excess traffic for all topologies under SHD and SHL intentional attack respectively. The excess traffic decreases with the increase of the removed nodes fraction due to the disconnection of a large number of node pairs and the network breaks down into isolated clusters. The excess traffic under SHD attack is larger than under SHL attack.

The reason is that all deleted nodes under SHL attack are located internally at the router level, while for SHD attack some topologies (Abilene and HOT) have all higher degree nodes located at the network edge.

Therefore, most node pairs are disconnected under SHL attack compared to SHD attack and for the Abilene topology the destruction is more harmful as it has few nodes at the core structure.

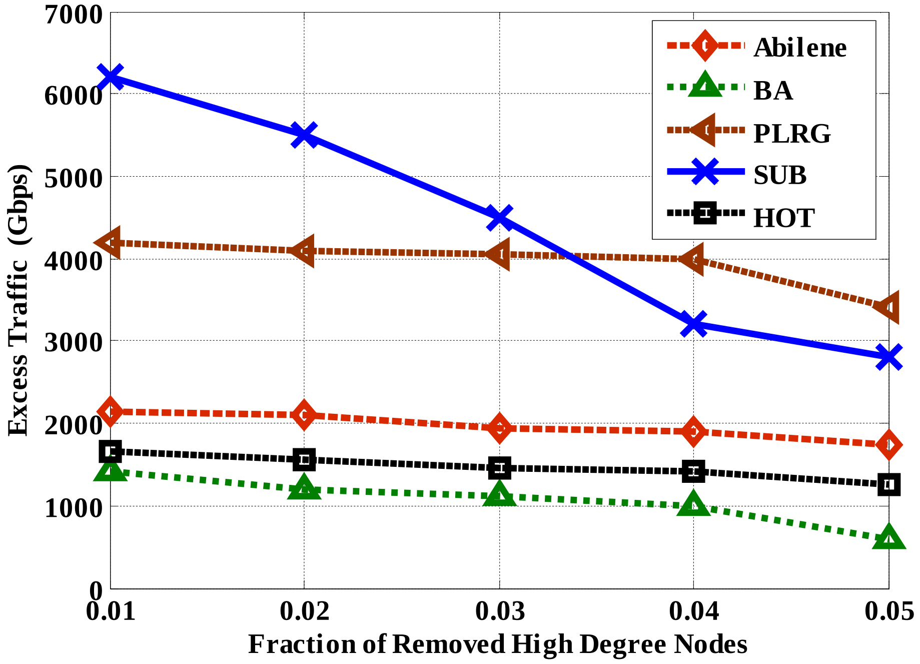


Figure 4: Excess traffic for all topologies under SHD attack

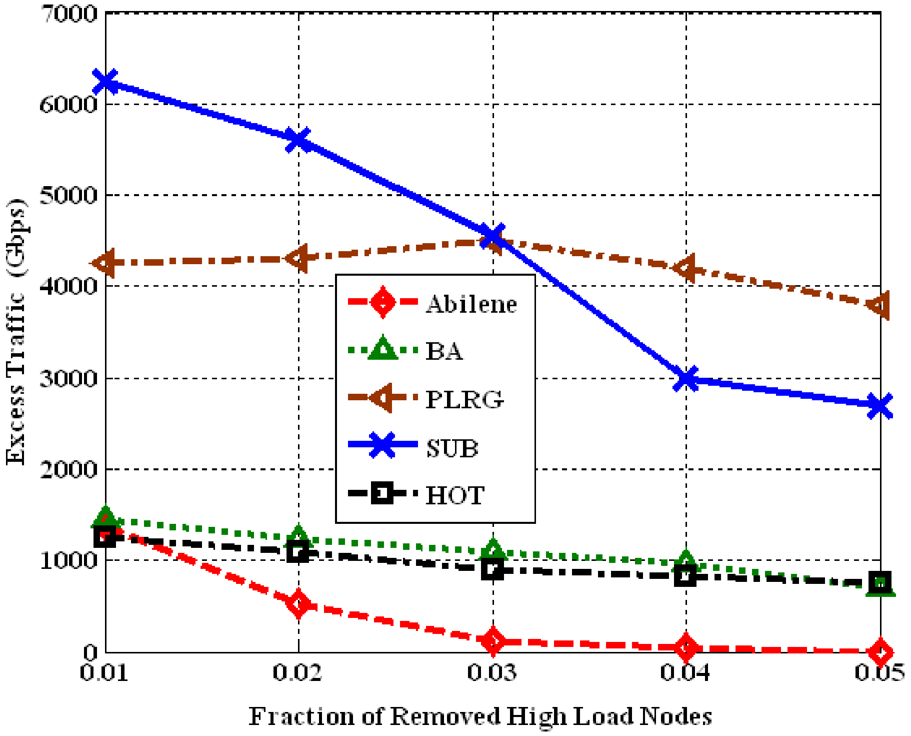


Figure 5: Excess traffic for all topologies under SHL attack

The excess traffic for PLRG and SUB has the largest values compared to BA, Abilene and HOT since the removal of nodes disconnects less node pairs as these topologies have the largest average clustering coefficient. Also PLRG has a highly heterogeneous degree distribution and accordingly the removal of its hub nodes results in large amount of excess traffic for the remaining nodes. The networks structures for Abilene and HOT are lightly heterogeneous represented by a mesh-like core supported by a hierarchical tree-like structure at the edges. Therefore, the removal of any constant fraction of nodes internally (not from the network edge) disconnects most nodes and results in small amount of excess traffic for the remaining nodes. Thus, the order of load-tolerant network topologies under SHD is BA > HOT > Abilene > PLRG > SUB (where the ">" indicates better than). While for SHL attack, the order is HOT > BA > PLRG > SUB > Abilene. As observed above, the excess traffic introduced by the SHD attack for HOT and Abilene is larger than SHL, however, for all the other topologies, the SHL attack results in larger excess traffic compared with SHD.

Figure 6 and Figure 7 show the relative LCC for all topologies under SHD and SHL intentional attack respectively. The LCC decreases with the increase of the removed nodes fraction. The Abilene and HOT topologies have the largest LCC compared to the other topologies under SHD attack. The reason is that all deleted nodes for these topologies are located at the network edge and do not affect the majority of the node members of the LCC which are located internally at the router level (gateway and core levels).

The case is different under SHL attack as all deleted nodes are internally at the gateway and core levels. The LCC for Abilene topology which has a few nodes at the core declines more rapidly with the increase of the removed nodes. Therefore, LCC under SHD attack is larger than under SHL attack. The order of load-tolerant network topologies under SHD attack is Abilene and HOT > PLRG > BA > SUB. While for SHL attack the order is PLRG > HOT > SUB > BA > Abilene.

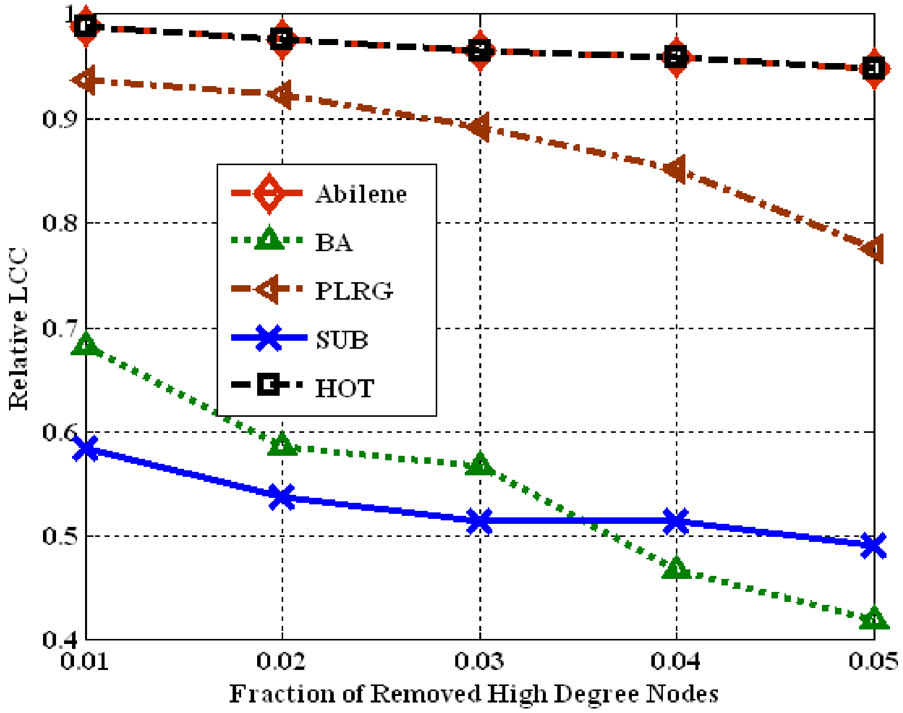


Figure 6: Relative LCC for all topologies under SHD attack

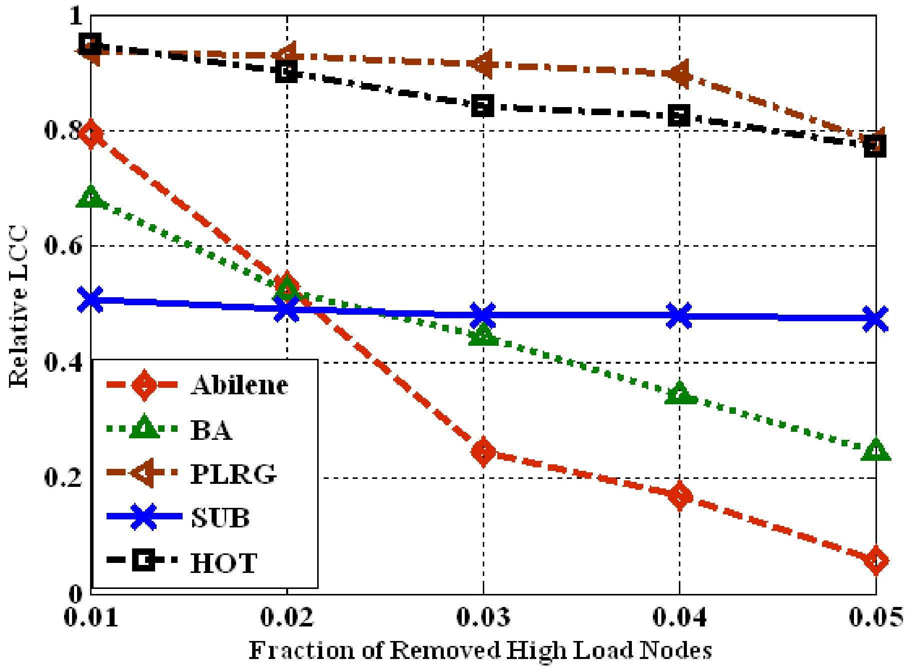


Figure 7: Relative LCC for all topologies under SHL attack

5 Resource Over-provisioning for High Robustness and Reliable Communication for Mitigation of Intentional Attacks

An important issue in the design of ISP topologies is survivability as represented by the reliability in the presence of equipment failure or intentional attacks. Network survivability is quantified in terms of the ability of the network to maintain end-to-end paths in the presence of node or link losses. For real ISPs, the objective is clearly not to maximize throughput but to provide good level of reliability by considering network resource over-provisioning. The network resource over-provisioning is a preventive method which enhances network robustness before an attack is started.

In this section, we introduce three preventive resource over-provisioning capacity strategies that utilize the structural properties of the network topology to improve robustness and increase its resilience (the ability to return to a normal state after a disturbance). Also, we compare the results obtained from these proposed strategies with the results obtained from the original network resource over-provisioning model introduced by Motter et al. [12]. Our proposed strategies can be applied to any combination of network topology and associated routing methods. Furthermore, our approach addresses the practical considerations (router technology) and economic indicators (link costs).

5.1 The Motter Resource Over-provisioning Model

Motter et al. [12] have introduced a simple network resource over-provisioning model to deal with cascades of overload failures. The model assumes that each node is characterized by a finite capacity, defined as the maximum load that the node can handle. The capacity is assumed to be proportional to the initial load of the node with tolerance parameter that represents the ability of nodes to handle increased load thereby resisting perturbations. The model is defined for our work as

$$C_i = d\,B_i\ ,\ \ i = 1, 2, \ldots N, \tag{7}$$

where C_i is the resource over-provisioning capacity associated with each router i which cannot exceed its maximum throughput $B_i^{(\max)}$ as outlined in section 3, $d \geq 1$ is the over-provisioning ratio (tolerance parameter), B_i is the configured bandwidth for router i constrained to its original feasible region, and N is the number of routers. The value for d depends on the impact of the excess traffic obtained when the network breaks down under failure. If the impact is extremely high, d must be high enough to absorb the effect of the redistribution of flows and attempting to keep the network operation unaffected by the failure.

5.2 The Proposed Resource Over-provisioning Strategies

5.2.1 Adjacent Resource Over-provisioning Method Based on Local Information

Our first proposed method (P1) assumes that most attacks in communication networks are propagated from one node to its adjacent node(s). This will cause the network to redistribute the load of the crashed node over its neighbors. Hence, we propose network resource over-provisioning method as follows:

$$C_i = d \max_j B_j \ \ \forall\, i = 1, 2, \ldots N,\ j \in NGH(i) \tag{8}$$

where *NGH(i)* is the set of neighbors for router i, C_i is the resource capacity over-provisioning associated to each router i which cannot exceed its maximum throughput $B_j^{(\max)}$ as outlined in section 3, $d \geq 1$ is the over-provisioning ratio (tolerance parameter), B_j is the configured bandwidth for all adjacent routers to router i constrained to their original feasible region, and N is the number of routers. Thus, the capacity of a router i is proportional to the maximum capacity of the neighbor that has the largest influence under network breakdown.

5.2.2 Multiplicative Compensation Over-provisioning

Our second proposed method (P2) utilizes the structural properties of the topology by computing the excess traffic in case of single global cascade failure for each router and its influence on all other routers of the network. The observations about the excess traffic for the topologies under intentional attacks in section 4 provide us useful insights for selecting the important routers that should be removed from the network and monitoring the progression of overloading of other routers according to the influence of removing these important routers.

Thus, we remove 30 % of the routers (a fraction of removed routers enough to crash the network) at the router level in their node index. In the HOT and Abilene topologies, the routers are selected on the degree, whereas for the BA, GRG, and SUB topologies, routers are selected based on load. We choose this value for the following : (1) from practical experience [7, 8, 13, 15, 16], about 15-20% of node removal is enough to fragment the network into isolated clustered components. We choose a larger value to confirm that the attacks would breakdown the whole network and (2) We observe that about 20 % to 25% for nodes have high degree or high load and their removal result in high excess traffic for the remaining nodes, while the other nodes have lower degree or lower load and their removal result in low excess traffic effect. Hence, we decided to choose 30 % of the routers to be able to remove a large portion of those critical nodes.

Our approach should be able to deal with the more harmful attacks and accordingly our approach will do better against less harmful attacks. Thus, we remove router i from the network and calculate the traffic loads for all the remaining routers. We denote by $L_i(k)$ the traffic load of router k after router i removal and $\Delta L_i(k)$ for the load difference (excess traffic) before and after router i removal.

$$\Delta L_i(k) = L_i(k) - B_k \ , \qquad (9)$$

where B_k is the initial configured bandwidth for router k constrained to its original feasible region. We restore the removed router and repeat this procedure for each router of the 30 % removed routers, i.e. we calculate $\Delta L_i(k)$ for every router k with router i removal. Then, we construct a matrix $\Delta L = (L_{ij})$ as

$$\Delta L = \begin{bmatrix} \Delta L_1(1) & \ldots & \Delta L_1(j) & \ldots \\ \vdots & \ddots & \vdots & \\ \Delta L_i(1) & \ldots & \Delta L_i(j) & \\ \vdots & & & \ddots \end{bmatrix}, \qquad (10)$$

The matrix is $Z \times N$, where $Z = \mathrm{int}(0.3 * N)$ and N is the number of routers. The element L_{ij} represents how much the removal of router i influences router j . We arrange every column in descending order of their traffic load difference and get the largest load difference for each router.

We assume that the amount of excess traffic caused by the failure of M routers is equal to M times the maximum excess traffic of single cascade global failure. The second proposed method network resource capacity over-provisioning for each router is then defined as

$$C_i = M \Delta L(i) + B_i \ , \ i = 1, 2, \ldots N , \qquad (11)$$

where C_i is the resource over-provisioning capacity associated to each router i which cannot exceed its maximum throughput $B_i^{(\max)}$ as outlined in section 3, M denotes the number of failures to be assumed to occur, $\Delta L(i)$ is the maximum load difference for router i , and B_i is the initial configured bandwidth for router i constrained to its original feasible region.

5.2.3 Successive Compensation Over-provisioning

Our third proposed method (P3) utilizes the structural properties of the topology by computing the excess traffic in case of single global cascade failure for each router and its influence on the other routers as described above in section 5.2.2. For this method, we set the amount of excess traffic caused by the failure of MS routers to be the same as the summation of the largest MS cases of a single cascade global failure. Thus, MS here indicates the number of strongest different successive single cascade global failure (i.e. if

$MS = 3$, we get the first 3 largest traffic load difference for each router (for example $\Delta L_1(1)$, $\Delta L_2(1)$, and $\Delta L_3(1)$ for router 1) from equations (10) and (9)). Hence, the network resource capacity over-provisioning for each router is defined as

$$C_i = \sum_{j=1}^{MS} \Delta L_j(i) + B_i^{(0)}\ ,\ i = 1, 2, \mathbf{K}\, N, \qquad (12)$$

6 Performance Evaluation

We evaluate the performance of our proposed strategies in sections 6.1 and 6.2. We define two new network measurement metrics to compare the performance of the Motter method and our proposed strategies:

Over-provisioning cost: We define the over-provisioning cost as the sum of the configured bandwidths for all routers in the network (i.e. total configured capacity of the network). It measures the amount of total configured router capacities required when intentional attacks occur as a cost for over–provisioning strategy against attacks.

Relative excess traffic cost: The relative excess traffic cost is defined as the ratio between the excess traffic times the over-provisioning cost for a given over-provisioning strategy and the same quantities for a reference over-provisioning strategy. We introduce the formal description of our relative excess traffic cost function as:

$$RETC = \frac{(Excess\ traffic * Over\ \ provisioning\ \cos t)_{for\ strategy\ x}}{(Excess\ traffic * Over\ \ provisioning\ \cos t)_{for\ reference\ \ strategy}} \qquad (13)$$

The *RETC* measures the relative cost between the amount of additional router capacity required to decrease the excess traffic when a certain over-provisioning strategy is used compared to the initial amount of additional configured router capacity required to decrease the excess traffic for the reference strategy at each node removal fraction.

6.1 Network Robustness and Resilience under Resource Provisioning Strategies

Figure 8 and Figure 9 show the excess traffic for Abilene topology with the Motter over-provisioning method with d =1, 2, and 3 denoted by o(1), o(2), and o(3) respectively, P1 with d = 2 and 3 denoted by P1(2) and P1(3) respectively, P2 with M = 2 and 3 denoted by P2(2) and P2(3), and P3 with MS = 2 and 3 denoted by P3(2) and P3(3). The values for d = 2, 3, M = 2, 3 and MS = 2, 3 are examined to give indication about the comparison between our proposed approaches and the Motter model in case the impact of excess traffic is extremely high.

The excess traffic decreases by a small amount using the Motter resource over-provisioning method and also the resilience increases slightly when using proposal 1 (p1(2) and p1(3)). The resilience increases significantly as the excess traffic decreases by a large amount when using proposal 2 and proposal 3. P2 and P3 seem to have the same effect, but P2 is better than P3. The same behavior is achieved by using our resource over-provisioning methods with the HOT, PLRG and BA topologies.

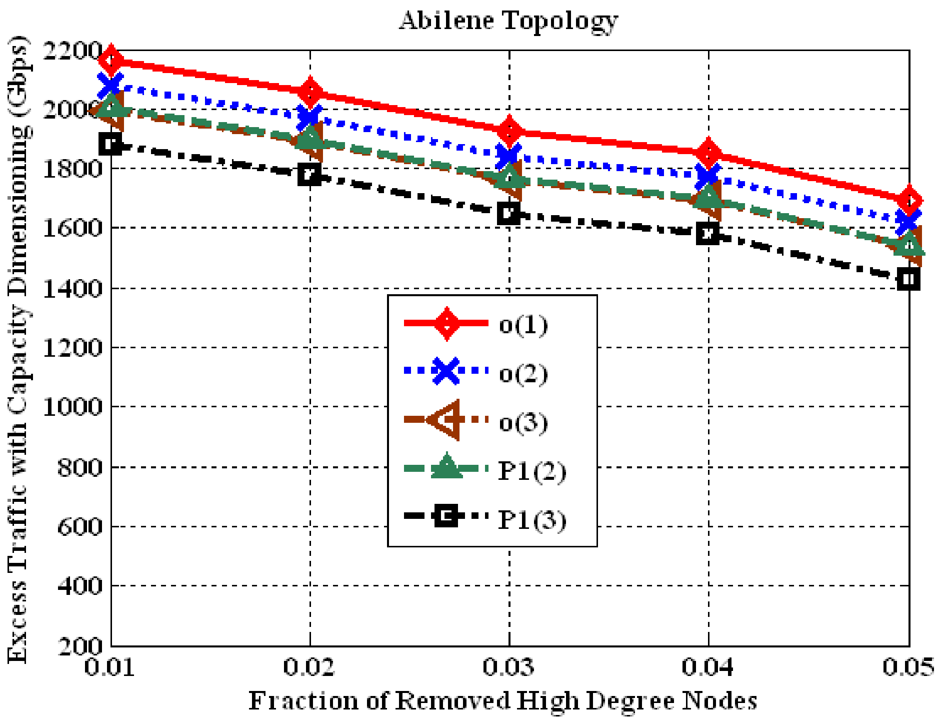


Figure 8: Excess traffic for Abilene under SHD with capacity over-provisioning using P1

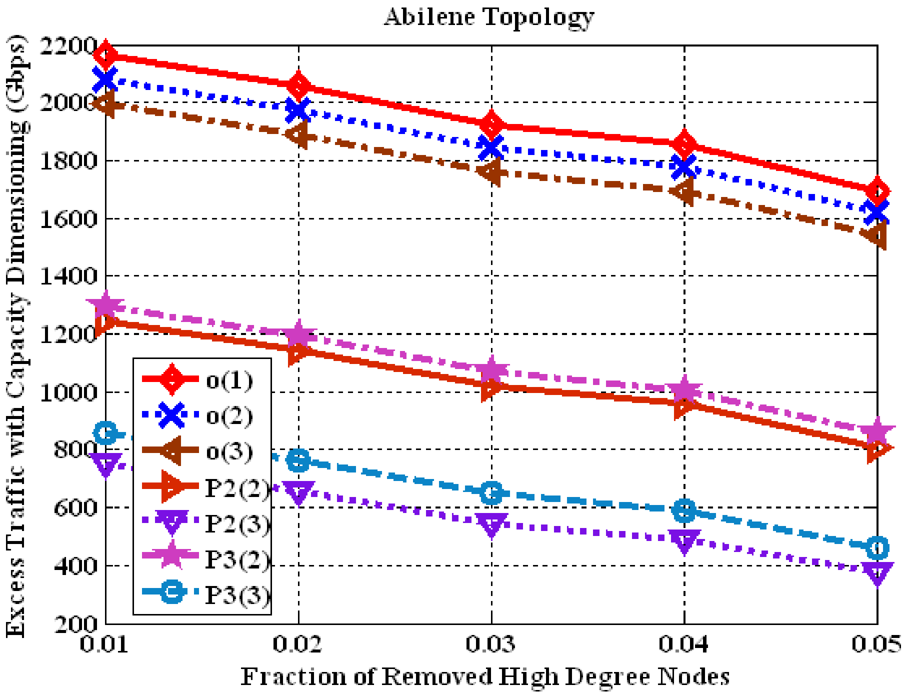


Figure 9: Excess traffic for Abilene under SHD with capacity over-provisioning using P2 and P3

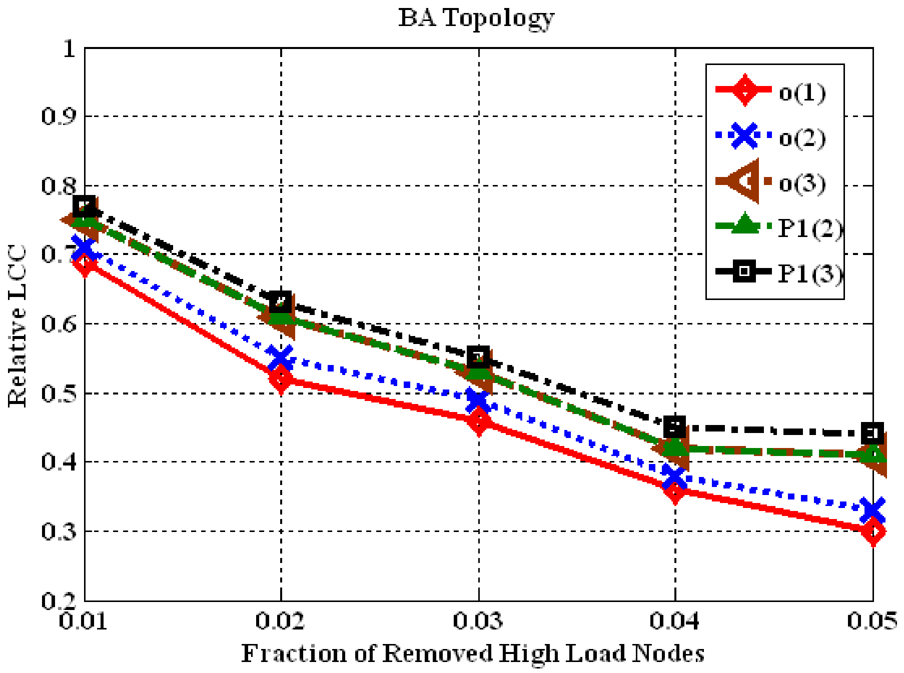


Figure 10: Relative LCC for BA under SHL attack with capacity over-provisioning using P1

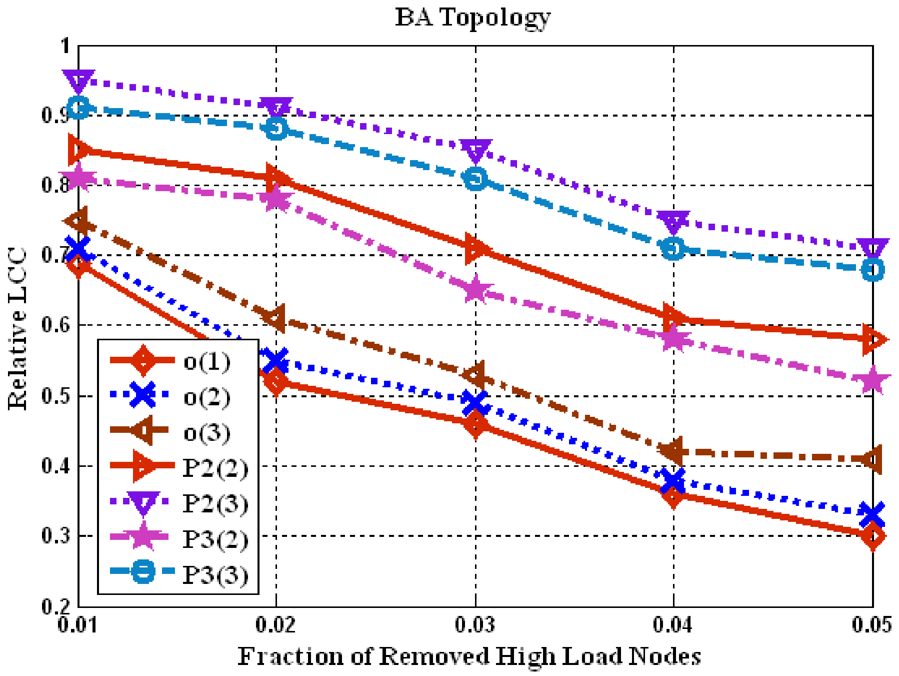


Figure 11: Relative LCC for BA under SHL attack with capacity over-provisioning using P2 and P3

The decrease of the excess traffic with the increase of the removed nodes fraction should not be interpreted as an enhancement of network functionality, but as an indication that the disruptions of the isolated network are so severe. Therefore, disconnection of large node pairs occurs and the network becomes fragmented into clusters which results in the decrease of the excess traffic. The network LCC increases according to the enhancement of the resilience. Figure 10 shows that the relative LCC for BA increases by a small amount using the Motter resource over-provisioning method and proposal 1. Figure 11 shows that the relative LCC increases significantly when using proposal 2 and proposal 3 and the network exhibits better survivable characteristics than the Motter method.

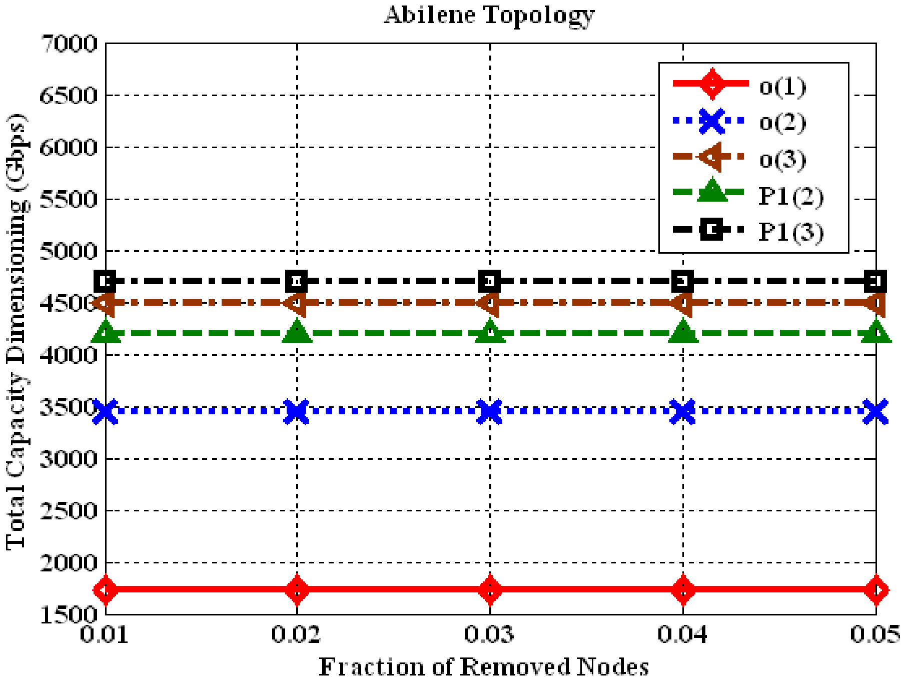


Figure 12: Over-provisioning costs for Abilene real network using P1

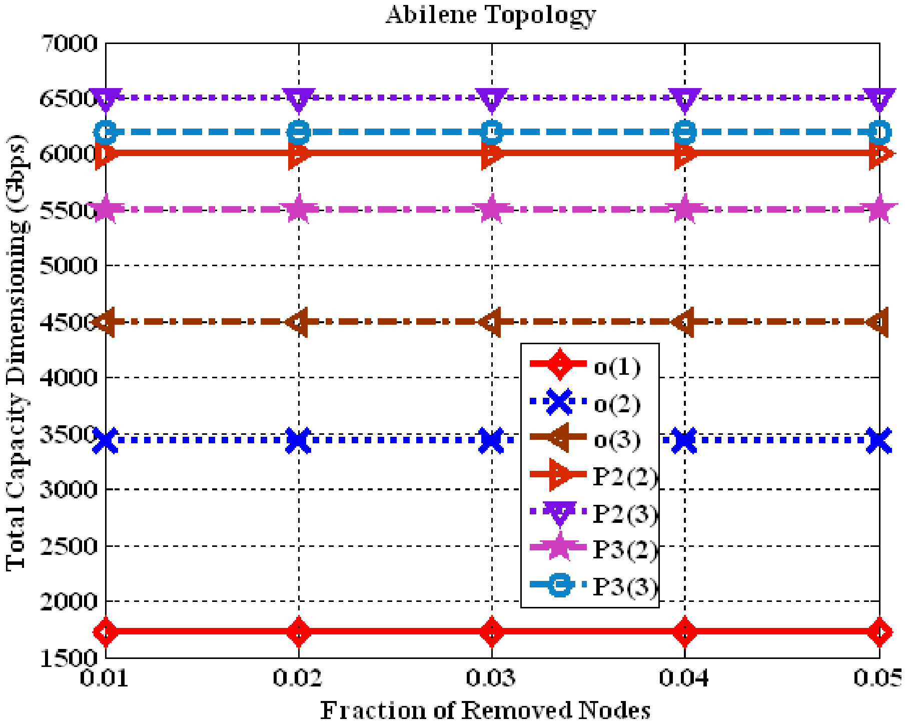


Figure 13: Over-provisioning costs for Abilene real network using P2 and P3

We consider the over-provisioning costs and RETC for Abilene real data network. Figure 12 and Figure 13 show the over-provisioning costs for the Motter method, P1, P2, and P3. The figures show that the smaller over provisioning ratio (tolerance parameter) makes the over-provisioning cost in the network smaller. The figures show a wide difference in over-provisioning costs between (the Motter method, P1) and (P2, P3). The figures show that P2(3) has the largest over-provisioning cost compared to the other different strategies. The figures also show that the amounts of allocated bandwidths are almost the same in o(3) and P1(3), P3(3) and P2(2), and different between o(2) and P1(2).

Although the over-provisioning cost is almost the same for o(3) and P1(3), the amount of excess traffic of P1(3) is smaller because P1(3) allocates more capacity for the most overloaded nodes whereas o(3) increases capacity for all nodes identically (triple the original capacity).

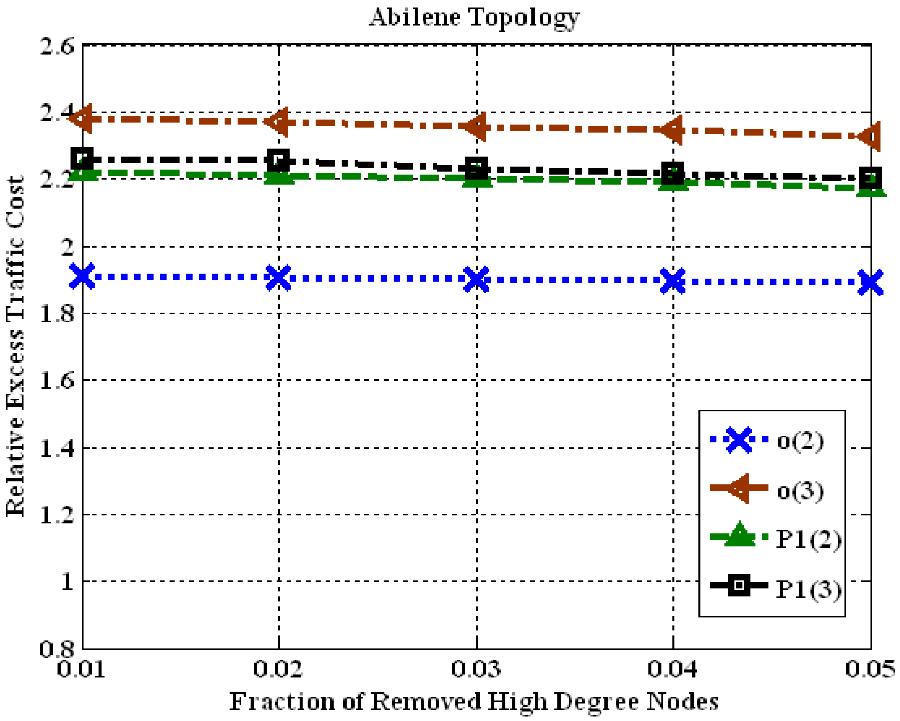


Figure 14: *RETC* for Abilene real network using P1

Figure 14 and Figure 15 show the relative excess traffic costs for the Motter method (o(2), o(3)), (P1, P2, P3) relative to the initial excess traffic costs for over-provisioning with o(1). The figures show that the Motter method and P1 have larger relative cost values compared to P2, P3 and o(3) has the largest relative cost value. The Figures show that the relative cost value decreases for all strategies simultaneously with the decrease in excess traffic at each node removal fraction.

The figures also show that P2(3) and P3(3) have the smallest relative cost values compared to the other strategies and o(2) strategy has almost the same relative cost for P2(2) and P3(2). Although the relative cost is almost the same for o(2) and P2(2) and P3(2), the amount of excess traffic of P2(2) and P3(2) is smaller because P2(2) and P3(2) allocates more capacity for the most overloaded nodes whereas o(2) increases capacity for all nodes identically (double the original capacity).

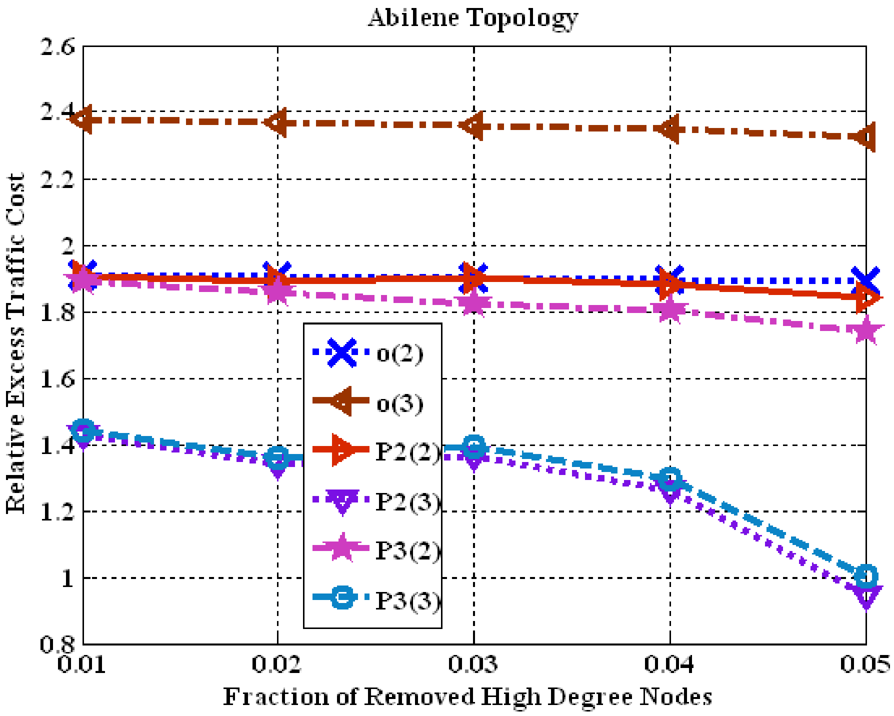


Figure 15: *RETC* for Abilene real network using P2 and P3

6.2 Empirical Evidence using Sprint Network for Network Robustness and Resilience under Resource Provisioning Strategies

We use the data from Sprint network at the router level as further evidence that our work can be applied to any network topology structure. Figure 16 shows the excess traffic for Sprint under SHD and SHL attacks. The amount of excess traffic for Sprint is large since the Sprint topology has a highly heterogeneous degree distribution and a large average clustering coefficient (0.65).

Accordingly, the Sprint network has a tendency to concentrate traffic on certain nodes since it has redundant links and any removal of its nodes results in large amount of excess traffic for the remaining nodes.

Figure 17 shows the relative LCC for Sprint under SHD and SHL attacks. The LCC decreases with the increase of the removed nodes fraction and LCC under SHL attack is larger than under SHD attack. The reason is that the majority of deleted nodes under SHL attack are not members of the LCC. Therefore, survivability of the network under the SHL attack is better than under the SHD attack.

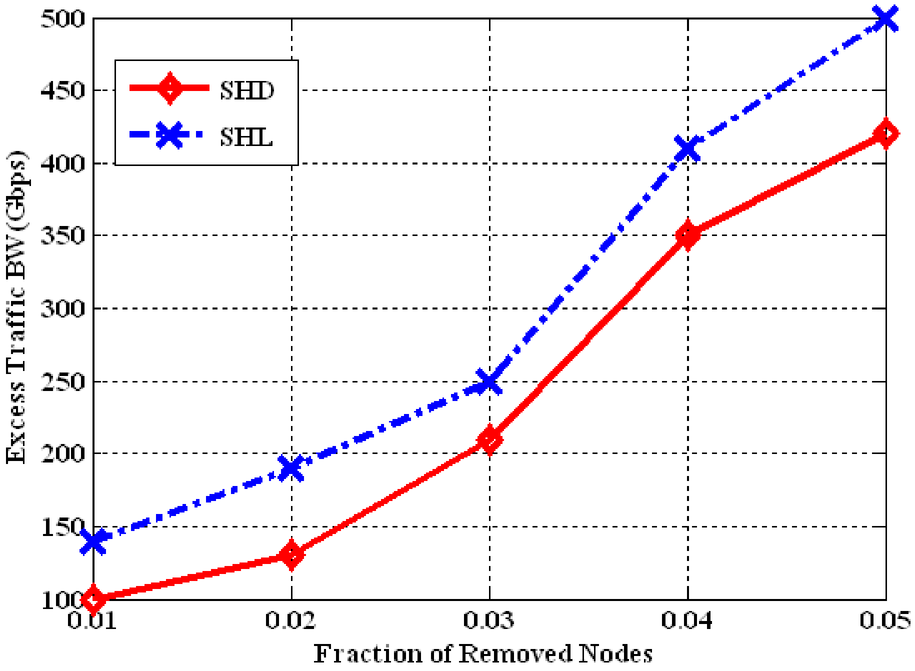


Figure 16: Excess traffic for Sprint under attacks

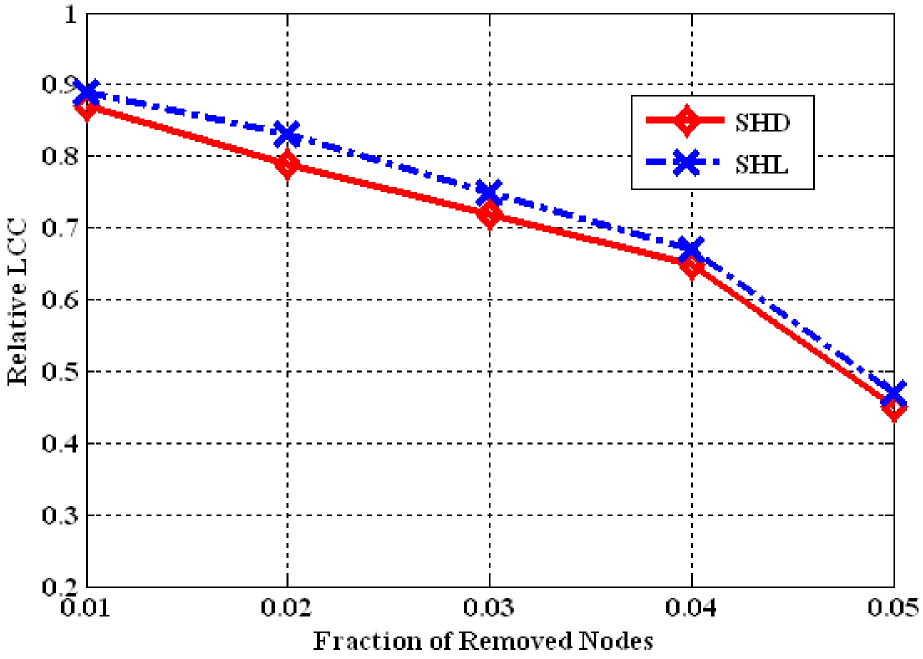


Figure 17: Relative LCC for Sprint under attacks

Figure 18 and Figure 19 show the excess traffic with resource over-provisioning strategies under SHL attack. The excess traffic decreases by a small amount using the Motter resource over-provisioning method and P1 The resilience increases significantly when using P2 and P3 and the network becomes highly-tolerant and can return to its normal functions better than the Motter method. In addition, P2 and P3 seem to have the same effect, but P2 is better than P3.

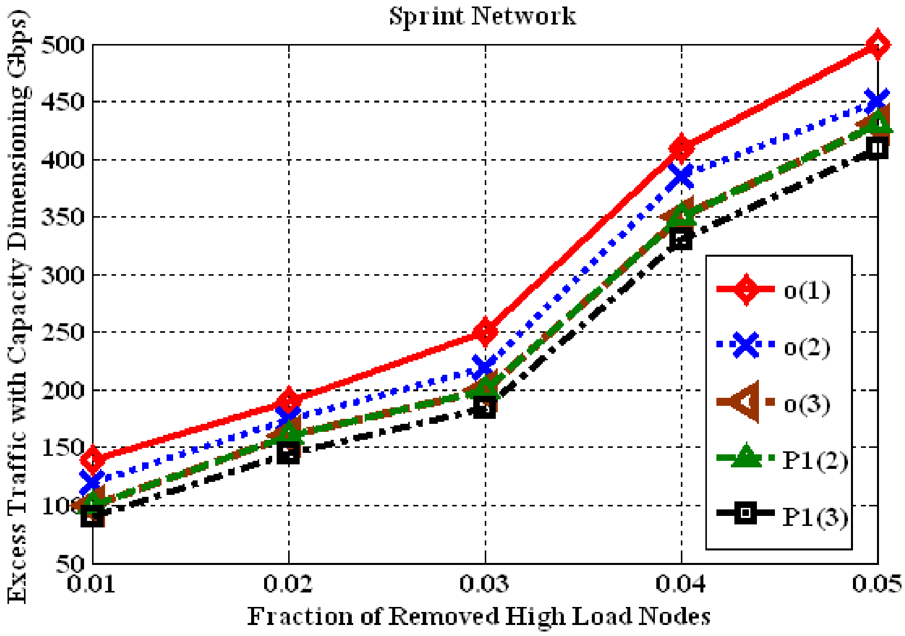


Figure 18: Excess traffic for Sprint under SHL with capacity over-provisioning using P1

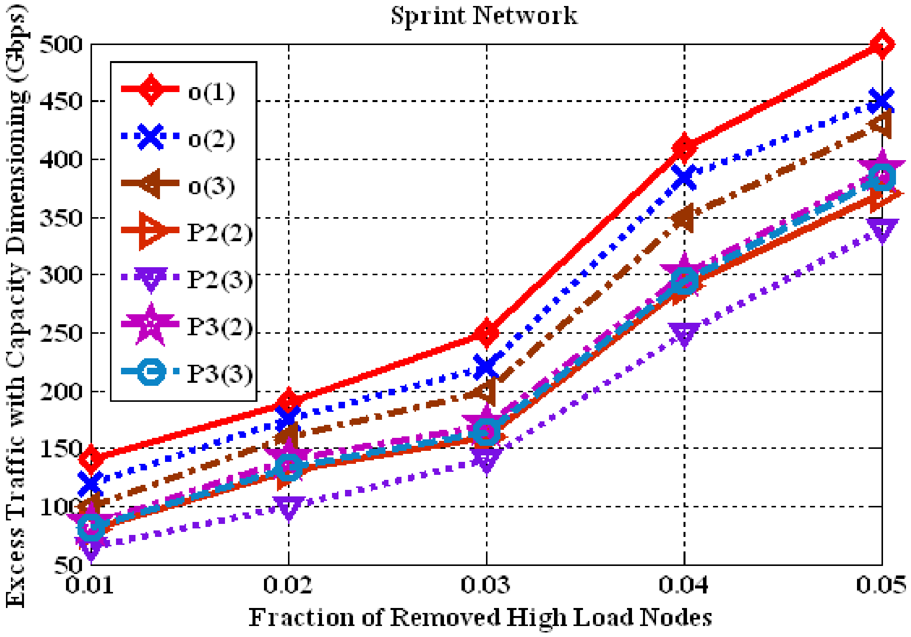


Figure 19: Excess traffic for Sprint under SHL with capacity over-provisioning using P2 and P3

The same behavior is achieved by using our resource over-provisioning methods to increase the survivability of the Sprint network. Figure 20 and Figure 21 show that the relative LCC for Sprint under SHD attack increases by a small amount using the Motter resource over-provisioning method and proposal 1. The relative LCC increases significantly when using P2 and P3 and the network becomes more efficient compared with the Motter method.

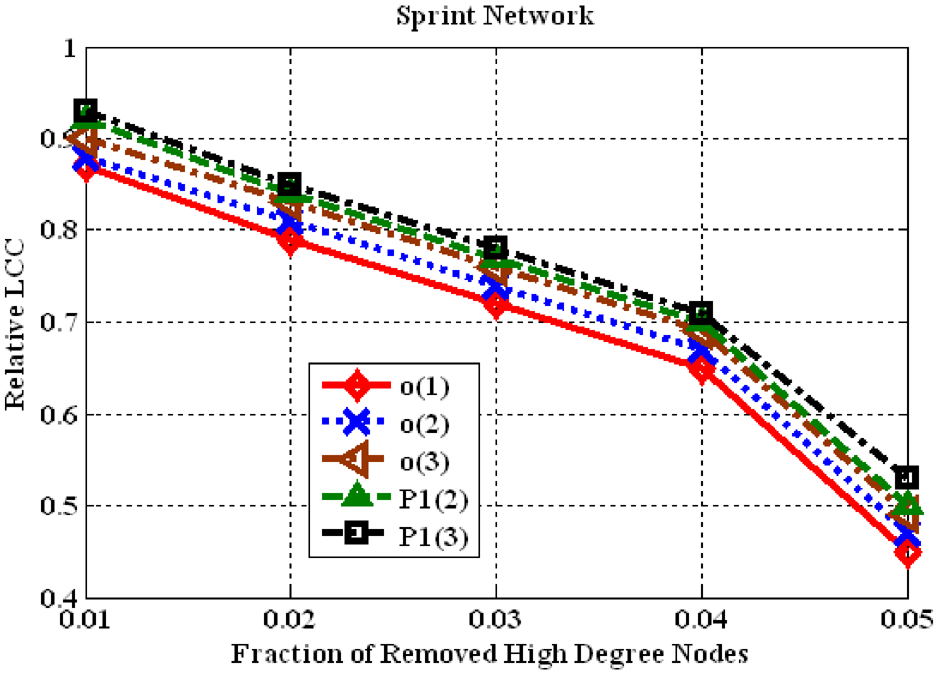


Figure 20: Relative LCC for Sprint under SHD with capacity over-provisioning using P1

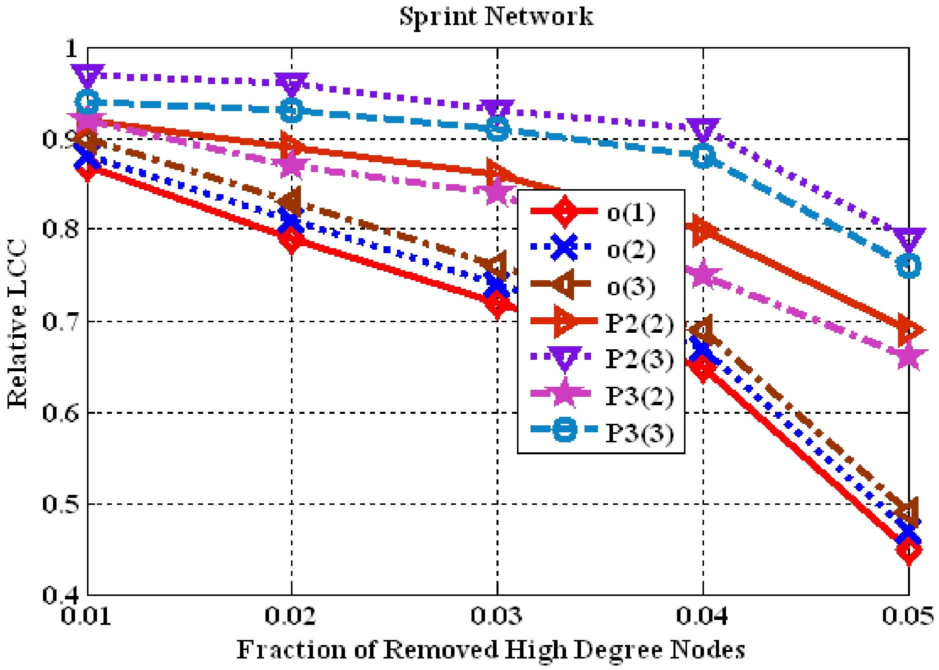


Figure 21: Relative LCC for Sprint under SHD with capacity over-provisioning using P2 and P3

Figure 22 and Figure 23 show the relative excess traffic costs for the Motter method (o(2), o(3)), (P1, P2, P3) relative to the initial excess traffic costs for over-provisioning with o(1). The figures show that the Motter method and P1 have larger relative cost values compared to (P2, P3) and o(3) has the largest relative cost value.

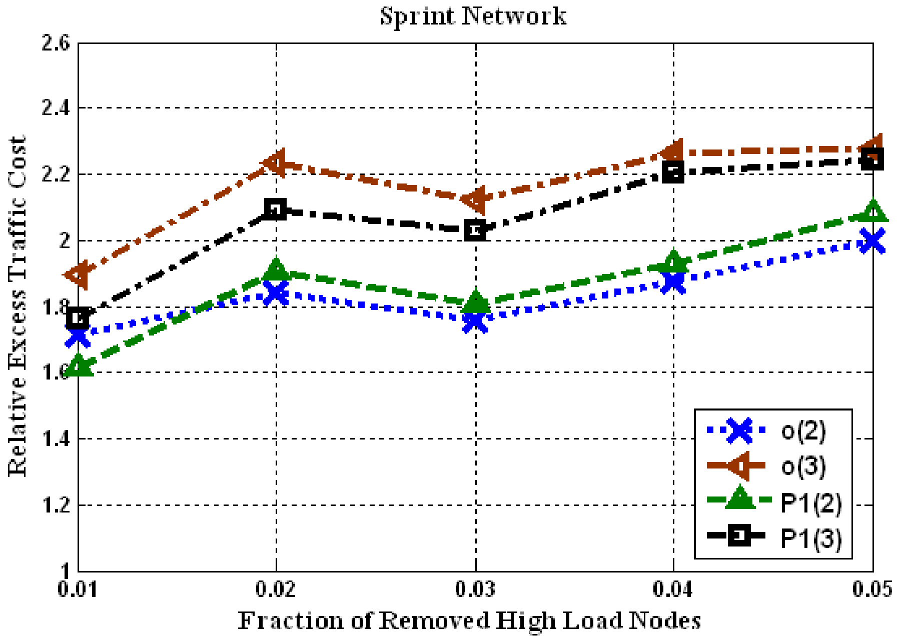


Figure 22: *RETC* for Sprint network using P1

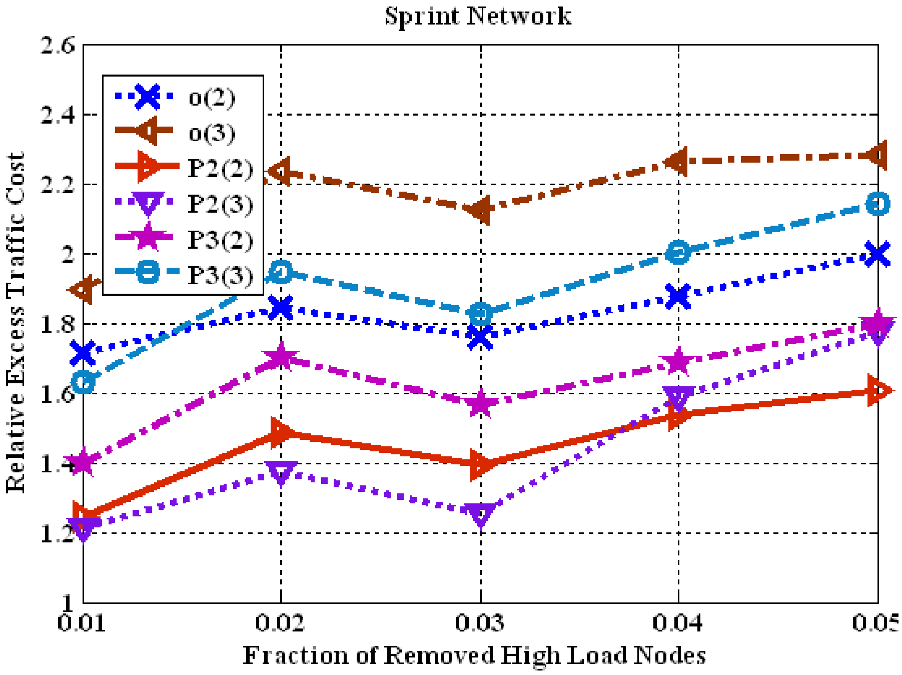


Figure 23: *RETC* for Sprint network using P2 and P3

The figures show that the relative cost value increases for all strategies simultaneously with the increase in each node removal fraction except when node removal percentage is equal to 0.03. The reason is that there is a little amount of excess traffic at the 0.03 node removal compared to the excess traffic at 0.02 and it is small compared to the excess traffic at its following node removal fraction at 0.04.

The figures also show that P2(2) and P2(3) have the smallest relative cost values compared to the other strategies. Although the relative cost is almost the same for o(2) and P1(2) and P3(3), the amount of excess traffic of P3(3) is smaller because P3(3) allocates more capacity for the most overloaded nodes whereas o(2) and P1(2) increase capacity for all nodes identically (double the original capacity).

The results show that there is a correlation between the network topology structure and the amount of resource capacity over-provisioning needed to absorb the large amount of excess traffic that results in network attacks (SHD or SHL). The amount of excess traffic due to attacks in highly heterogeneous networks is large and accordingly the amount of resource capacity over-provisioning needed to absorb the large amount of excess traffic is high. The opposite occurs in lightly heterogeneous networks where the amount of excess traffic is small.

From the above, we can conclude that highly heterogeneous networks exhibit more resilience and can return to normal functions faster than lightly heterogeneous networks. In addition, P2 and P3 seem to have the same effect, but P2 is better than P3 and both of them are better than the Motter method and also P2 seem to have the best effect on highly heterogeneous network topologies compared to P3 and the Motter method.

7 Conclusions

We show that the physical connectivity is very important for router-level related issues such as throughput, reliability, and robustness to router loss. The effect of the same resource over-provisioning mechanism for reliability varies for different topologies depending on the underlying network structure. We also show that rerouting in case of component failure requires additional resource capacity on the network to absorb the excess traffic and a more detailed description of the underlying physical structure; including, link bandwidths and router capacities are required for assessing network vulnerability to intentional physical attacks. We proposed a resource capacity over-provisioning methods that utilize the network structural properties by calculating the increase of traffic in case of single global/local failures for each node.

Our proposed strategies can be applied to any combination of network topology and associated routing methods. Evaluation results show that our proposed resource capacity over-provisioning strategies enhance the robustness and increase the resilience compared to the Motter resource over-provisioning model. Also, it is found that there is a strong correlation between the amount of resource capacity over-provisioning needed to absorb the blockade of excess traffic which results in network failure (SHD and SHL attacks) and the network structure topology.

We also showed that the highly-heterogeneous networks become more resilient and can return to their normal operation faster than the lightly-heterogeneous networks. The effect of resource capacity over-provisioning strategies under attack also depends on the degree distribution of the network structure for the remaining nodes after deletion.

References

[1] S. Bornholdt and H. G. Schuster (Eds), “Handbook of Graphs and Networks: from the Genome to the Internet,” Wiley-VCH, 2003.

[2] R. Govindan and H. Tangmunarunkit, “Heuristics for Internet Map Discovery,” in Proc. IEEE INFOCOM, Tel Aviv., Israel, 2000, pp. 1371–1380.

[3] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, “Toward Capturing Representative AS-Level Internet Topologies,” in Proc. ACM SIGMETRICS, Marina Del Rey, CA, Jun. 2002, pp. 280–281.

[4] R. Albert, H. Jeong, and A.-L. Barabasi, “Error and Attack Tolerance of Complex Networks,” Nature Vol. 406, pp. 378-382, July 2000.

[5] R. Cohen, K. Erez, D. Ben-Avraham, and S. Halvin, “Breakdown of the Internet under Intentional Attack,” Phy. Review Letter. Vol. 86, No. 16, pp. 3682-3685, Apr. 2001.

[6] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, “Error and Attack Tolerance of Complex Networks,“ Physica A320 (2003) 622.

[7] D. Magoni, “Tearing Down the Internet,” IEEE Journal on Selected Areas in Communications, Vol. 21, No. 6, pp. 949-960, Aug. 2003.

[8] J. C. Doyle, D. L. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka, and W. Willinger, “The “Robust yet Fragile,” Nature of the Internet”, Proceedings of the National Academy of Sciences of the United States of America, Vol. 102, No. 41, pp. 14497-14502, Oct. 2005.

[9] P. Holme and B. J. Kim, “Vertex Overload Breakdown in Evolving Networks,” Physical Review Letters, E 65 (026139), Vol. 65, June 2002.

[10] P. Holme, “Edge Overload Breakdown in Evolving Networks,” Physical Review Letters, E 65 (066109), Vol. 65, Sept. 2002.

[11] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, “Attack Vulnerability of Complex Networks,” Physical Review E65, 056109(2002).

[12] A. E. Motter and Y. Lai, “Cascade-Based Attacks on Complex Networks,” Phys. Rev. E66 (2002), 065102(R).

[13] S. Xiao and G. Xiao. “On Intentional Attacks and Protections in Complex Communication Networks,” in Proc. IEEE Globecom’06, Nov.-Dec. 2006.

[14] S. Xiao, G. Xiao, and T. H. Cheng, “Tolerance of Intentional Attacks in Complex Communication Networks,” IEEE Communications Magazine, Vol. 46, No. 1, pp. 146-152., Jan. 2008.

[15] A. Beygelzimer, G. M. Grinstein, R. Linsker, and I. Rish, “Improving Network Robustness by Edge Modification,” Physica A, Vol. 375, No. 3-4, pp. 593-612, Apr. 2005.

[16] S. Xiao, G. Xiao, and T. H. Cheng, “Robustness of Complex Communication Networks under Rewiring Operations,” in Proc. IEEE ICCS’06, Oct-Nov. 2006.

[17] S. Xiao and G. Xiao, “On Local Link Repairing in Complex Communication Networks under Intentional Attack,” Proc. ICICS, Dec. 2007.

[18] Y. Wang, G. Xiao, T. H. Cheng, S. Xiao, and X. Fu, “Robustness of Complex Communication Networks under Link Attacks,” in Proc. ICAIT’08, July 2008.

[19] R. Albert and A. L. Barabási, “Statistical Mechanics of Complex Networks,” Rev. Modern Phys., Vol. 74, no. 1, pp. 47–97, Jan. 2002.

[20] F. Chung and L. Lu, “The average distance in a random graph with given expected degrees,” Internet Math., Vol. 1, pp. 91–113, 2003.

[21] W. Aiello, F. Chung, and L. Lu, “A Random Graph Model for Massive Graphs,” Proc. STOC 2000.

[22] L. Li, D. Alderson, W. Willinger, and J. Doyle, “A First-Principles Approach to Understanding the Internet’s Router-Level Topology,” ACM SIGCOMM Computer Communication Review, Vol. 34, No. 4, pp. 3–14, Oct. 2004.

[23] Abilene Network, “Detailed Information about the Objectives, Organization, and Development of the Abilene Network,” [Online]. Available: http://www.Internet2.edu/abilene, 2004

[24] L. Li, D. Alderson, W. Willinger, and J. Doyle, “Topology Information for Five Networks,” [Online]. Available: http://hot.caltech.edu/topology/toynets.html, 2004.

[25] Rocketfuel maps and data, “An ISP Topology Mapping Engine,” [Online]. Available: http://www.cs.washington.edu/research/networking/rocketfuel, 2003.

[26] Networks Software and Data, “Program for Large Network Analysis,” [Online]. Available: http://vlado.fmf.uni-lj.si/pub/networks/pajek, 2009.

[27] D. Alderson, L. Li, W. Willinger, and J. Doyle, “Understanding Internet Topology: Principles, Models, and Validation,” IEEE/ACM Transactions on Networking, Vol. 13, No. 6, pp. 1205-1218, DEC. 2005.

[28] Y. Zhang, M. Roughan, C. Lund, and D. Donoho, “An Information-Theoretic Approach to Traffic Matrix Estimation,” Proc. ACM SIGCOMM, Comput. Commun. Rev., Vol. 33, pp. 301–312, 2003.