Enhancing the Robustness and Efficiency of Scale-free Network with Limited Link Addition

Li Li[1,2], Qing-Shan Jia[3] , Xiaohong Guan[1,3] and Hengtao Wang[3]
[1] SKLMS Lab and MOE KLINNS Lab, Xi'an Jiaotong University
Xi'an, 710049, China
[2] College of Computer Science, Shaanxi Normal University
Xi'an, 710062, China
[3] CFINS, Department of Automation, TNLIST, Tsinghua University
Beijing, 100084, China
[e-mail: lili@snnu.edu.cn, jiaqs@tsinghua.edu.cn]
*Corresponding author: Qing-Shan Jia

Received January 6, 2012; revised March 6, 2012; accepted April 25, 2012; published May 25, 2012

Abstract

The robustness of a network is usually measured by error tolerance and attack vulnerability. Significant research effort has been devoted to determining the network design with optimal robustness. However, little attention has been paid to the problem of how to improve the robustness of existing networks. In this paper, we investigate how to optimize attack tolerance and communication efficiency of an existing network under the limited link addition. A survival fitness metric is defined to measure both the attack tolerance and the communication efficiency of the network. We show that network topology reconfiguration optimization with limited link addition (NTRLA) problem is NP-hard. Two approximate solution methods are developed. First, we present a degree-fitness parameter to guide degree-based link addition method. Second, a preferential configuration node-protecting cycle (PCNC) method is developed to do trade-off between network robustness and efficiency. The performance of PCNC method is demonstrated by numerical experiments.

Keywords: Network topology reconfiguration, limited link addition, attack tolerance, communication efficiency

This work is supported in part by the National Natural Science Foundation under Grants (Nos.60921003, 60704008, and 61174072), 863 High Tech Development Plan (Nos.2007AA01Z475 and 2007AA01Z480), the Fundamental Research Funds for the Central Universities (No.GK201102008), the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20070003110), Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation, and 111 International Collaboration Program (No. B06002) of China.
The authors would like to thank Dr. Yong He and Dr. Jianghai Li, Mr. Chaobo Yan, and Mrs. Xiaoyan Xu for their helpful discussions.

http://dx.doi.org/10.3837/tiis.2012.05.005

1. Introduction

Recent studies have shown that scale-free network (SFN) has provided a good model for many real-world networked systems, such as the Internet, airline route, transportation networks, electrical power grids, and other cyber-physical systems [1][2][3][4]. As important infrastructures in modern society, these systems are expected to be sufficiently robust against unpredictable breakdown of nodes and links in order to function constantly [3]. The network topology of SFN is known to be effective in terms of both the average path length and the robustness against random failures of nodes. However, if the hub nodes are intentionally attacked, SFN is found vulnerable [1][5][6]. It is thus important to reconfigure an existing network to improve both attack tolerance and communication efficiency while maintaining the networked system performance [4].

Network functionality is often measured by the connectivity and the average path length. On the one hand, the connectivity is a fundamental requirement for communication network topologies. The communication efficiency of a network, on the other hand, depends of the average path length (indicating communication delay) between nodes. When a network undergoes reconfiguration, either because of link additions or node/link failures, the changes will affect the performance of the network. Understanding the effects of such changes is a vital task of network topology reconfigurations.

Several topology reconfiguration methods [7][8][9][10][11][12][1314] have been proposed to strengthen the network reliability and robustness under various types of breakdown and attack. In most of the cases, these methods did not consider the resources constraints [7][8][9][10][11]. Some methods improved the network connectivity but did not consider the communication efficiency of the network [8][9][10]. Overall, existing methods have limited abilities to strike a balance between network robustness and communication efficiency in a dynamic network environment. It is in general still a challenging task to reconfigure the topology of a networked system in order to improve the robustness and the efficiency with limited link additions.

In this paper, we consider this important problem, which is called network topology reconfiguration optimization with limited link addition (NTRLA). Different from the Maximum Diameter Edge Addition (MDEA) problem [14], the NTRLA problem focuses on improving both network robustness and communication efficiency. We make the following major contributions. First, we show that the NTRLA is NP-hard. Second, two approximate solution methods are developed. One method is a new degree-based link addition method using a degree-fitness parameter. The other method is a preferential configuring node-protecting cycle (PCNC) method. Simulation results show that PCNC method outperforms degree-based link addition methods both in improving the attack tolerance and in improving the communication efficiency of the network for small and medium scale networks.

The rest of the paper is organized as follows. In section 2, we review the related works. In section 3, the problem is mathematically formulated and the combined survival fitness metric is discussed. In section 4, the NTRLA problem is shown to be NP-hard, and the node-protecting cycle method is analyzed theoretically. In section 5, we introduce the degree-based link addition methods and the PCNC method. In section 6, the performance of the PCNC method is demonstrated by experimental results. We briefly conclude in section 7.

2. Related Work

An important property of networked systems is their robustness against removal of network nodes, through either random node failure or targeted attack. Albert et al. [1] studied how the properties of the Internet and the properties of a sample of the World Wide Web change when a fraction of nodes are removed. They pointed out that the topological weaknesses of the current communication networks are rooted in their heterogeneous connectivity distribution which seriously reduces their survivability under attack.

Several groups of authors, e.g., Shargel et al. [15], Paul et al. [16], Valente et al. [4] have considered the problem of designing networks with optimal robustness or good tradeoff between random failures and targeted attacks. However, the cost of design from scratch is very high in the real world, and a real network is the result of many different processes which may have little effects on robustness against removal of network nodes. Though the topologies of large real-world networks can be changed to improve the performance, the modification of the network should be as small as possible due to economic and various other concerns. For example, it is almost impossible to abandon the existing Internet to build a new one. But it is possible to add a small number of links to achieve higher robustness and efficiency. A series of methods with link addition have been developed [7][8][9][10][11][12][13][14]. For example, Beygel et al. [7] discussed several degree-based topology modification methods, including random addition, random rewiring, preferential addition, and preferential rewiring. They compared the effectiveness of various modification methods under two measures: the fraction of nodes remaining in the largest connected component, and the average inverse shortest path length. They found that random addition is better than random rewiring, and preferential addition performs the best.

Chi [8] proposed a repair method for complex networks under attacks. She studied the stability and correlation properties of Erdos-Renyi (ER) random graphs, Watts-Strogatz (WS) small-world networks, and Barabasi-Albert (BA) scale-free networks under the repair method. Zhao et al. [9] proposed a new parameter α to guide the process of enhancing the robustness and the experiments showed that the effect of enhancement is better when $\alpha < 0$. Their experiment results showed that the strategy of establishing new links between nodes with the lowest degree can greatly enforce the attack survivability of the network without reducing the error tolerance. Sato et al. [10] investigated whether the current Internet is optimized in both aspects of communication efficiency and attack tolerance. They found that the current topology is not appropriate, and that a more suitable topology can be achieved by reducing the value of scaling exponent γ. They have proposed four methods for re-organizing a network topology. These methods focused on improving network robustness but did not consider the communication efficiency of the network.

Sekiyama et al. [11] presented a dynamic reconfiguration process of the network topology, which was an extension from the conventional preferential linking model. Their local evaluation indicator and control parameters were introduced to regulate a balance between efficiency and robustness. Their results suggested that a well-balanced network topology can be created via reconfiguration according to various intentional attack patterns. However, the aforementioned works [7][8][9][10][11] did not consider the constraint of limited link resources. The reconfiguration of a practical network usually adds limited new links due to economic concerns.

Wang and Piet [12] investigated how to optimize a network for a given dynamic process via minor topological modifications. Two link addition methods were proposed. They compared two methods with random link addition in three classes of networks: the ER random

graph, the BA model, and the *k*-ary tree. However, the random link addition method usually was poor performance and may not be a good benchmark.

In the wireless network topology configuration, Ranjitkar and Ko [17] investigated how to construct a robust and efficient topology. They designed an algorithm to construct degree constrained topology which can reduce processing complexity and maintaining the network connectivity.

Schoone et al. [14] proved that the Maximum Diameter Edge Addition (MDEA) problem is NP-complete. In this paper, leveraging the results of [14], we explore that the network topology reconfiguration with limited link addition (NTRLA) problem under the survival fitness metric which taking both the network robustness and the diameter into consideration.

When robustness and efficiency requirements are both important, the node protecting cycle (np-cycle) structure is simple and effective. Compared with node-encircling p-cycle structure that is applied in mesh and lattice networks [18][19], the np-cycle structure can be configured for any network to increase the robustness and the efficiency of the network. We will develop our solution method for the NTRLA problem based on the optimal configuration of np-cycles.

3. Optimal Network Topology Reconfiguration with Link Addition

To obtain an appropriate network topology, it is necessary to take into account both network robustness and communication efficiency. Therefore, we introduce the survival fitness metric to measure both the robustness and the efficiency of a network under various types of attacks. Before proposing the evaluation metric and the formal model of NTRLA problem, we first privede a list of notations for the convenience of the readers.

3.1 Notations

We first summarize the notations which will be used throughout this paper in **Table 1**.

Table 1. Notations

G	A graph with n nodes
V	The set of all nodes in a network, e.g., node $i, j, k \in V$
E	The set of links in a netwok, e.g., link $e \in E$
E'	The set of links to be added in a network, e.g., link $\ell \in E'$
η^R	The robustness metric of a graph
η^E	The efficiency metric of a graph
α	A given environment parameter , $0 \le \alpha \le 1$
Γ	The survival fitness metric of a graph, i.e., $\Gamma = \alpha \cdot \eta^R + (1-\alpha) \cdot \eta^E$
x_k	A node-protecting cycle (np-cycle) for protecting node k
$CO(x_k)$	The order of the np-cycle x_k , i.e., the number of neighboring nodes of k on x_k
$NCE(x_k)$	The normalized efficiency of the np-cycle x_k
F_i	The degree-fitness of node i
$Ctr(\ell)$	The link contribution metric of the adding link ℓ

3.2 Evaluation Metric for Network Topology

The survival fitness metric is based on the robustness metric and the efficiency metric. Let us introduce some terms in graph theory first [20]. The node degree or degree of a node is the number of links connected to the node. The path length or node-to-node distance of a graph is the minimal number of steps it takes to go from one node to the other. A graph is said to be connected if there is a path beteen any two nodes. We assume that all nodes and links are of equal importance.

Robustness Metric

Before introducing the robustness metric, we define the following notations [21][22][23[24][25]. Reachability shows whether there exists a path between two nodes in the graph. If there is a path between node i and j, reachability r_{ij} equals 1. Otherwise r_{ij} equals 0. Let G_k be the graph after removing node k from G. The robustness with respect to the removal of node k is defined as

$$\eta_k^R = \frac{1}{(n-1)(n-2)} \sum_{j \neq i \in G_k} r_{ij} \ , \tag{1}$$

which is the reachability of G_k divided by the maximal possible reachability of G. Let p_k denote the removal probability of node k. If $p_k = 1/n$, $\forall k \in V$, this means random failure, namely, all nodes fail with equal probability [21]. For targeted attack, the nodes with high degree would fail with higher probability [22]. Then the robustness η^R of a graph G can be defined as

$$\eta^R = \sum_{k \in G} p_k \cdot \eta_k^R \ . \tag{2}$$

By the definition, we have $0 \leq \eta^R \leq 1$. Robustness shows the ratio of the number of available nodes after a single node failure to the number of all nodes in a network. If we can communicate between any pair of nodes in a network, robustness η^R becomes 1. Note that a larger value of η^R means that the network is more robust under node failures. For more complicated analysis it is possible to initiate multiple nodes failures. However, exhaustive examination of set failures increases computational complexity of implementation.

Efficiency Metric

We follow the definition of efficiency in [23] in this paper. Let d_{ij} and $\delta_{ij} = 1/d_{ij}$ denote the shortest path length and the efficiency between nodes i and j, respectively. The efficiency η^E of a graph G is defined as

$$\eta^E = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \delta_{ij} \ . \tag{3}$$

In Eq. (3), the value of η^E decreases as d_{ij} increases. Note that if there does not exist any path between nodes i and j, we have $d_{ij} = +\infty$ and $\delta_{ij} = 0$. And η^E is still well defined.

Survival Fitness Metric

The survival fitness Γ is defined as a weighted sum of η^R and η^E, i.e.,

$$\Gamma = \alpha \cdot \eta^R + (1-\alpha) \cdot \eta^E \ , \tag{4}$$

where α is a constant, $0 \le \alpha \le 1$. The parameter α models the environmental pressure on the network. When α equals 1, the survivability of the network depends entirely on its robustness with no regard for efficiency. When α equals 0, the survivability is determined entirely by its efficiency with no regard for robustness. For the other values of α, the network should be both robust and efficient with the specified weights.

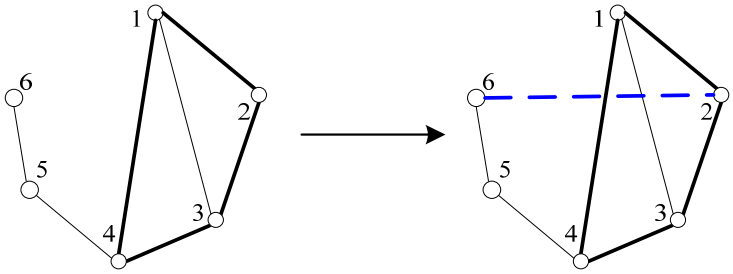
We focus on the survival of a network against random failures and targeted attacks. Depending on the functional goal of a network and its survival environment, we discuss "good case" and "bad case" survival fitness of a network. By "good case" survival we mean that the network reconfigures available resources to maximize its survival fitenss for random failure scenario. That is for 'good case' survival fitness, the calculation of Γ is discussed for $\alpha \le 0.5$. In general, compared with random failures, targeted attacks will produce a great harm to network survivability. Only the increase in path length can guarantee at least basic network communication, while network partitions mean that the communication between nodes is completely interrupted. So, for "bad case" survival fitness against targeted attacks, the calculation of Γ is discussed for $\alpha > 0.5$.

3.3 Model Outline

Let $G=(V,E)$ be a graph, where V denotes the set of nodes and E denotes the set of undirected links. Let $n=|V|$ and $m=|E|$. Define the survival fitness of G as $\Gamma(G) = \alpha \cdot \eta^R + (1-\alpha) \cdot \eta^E$. Then the network topology reconfiguration optimization problem with limited link additions (NTRLA) can be formulated as

$$\begin{aligned} &\mathrm{Max}_{E'} \ \Gamma(G'(V, E \cup E')) \\ &\text{s.t.} \ \ |E'| \le q \end{aligned}$$

where E' is the set of links to add, and q is a given positive integer. For example in **Fig. 1**, we have $V = \{1,2,3,4,5,6\}$ and $E = \{(1,2),(1,3),(1,4),(2,3),(3,4),(4,5),(5,6)\}$. The removal of node 4 or 5 will disconnect the initial graph (**Fig. 1(a)**). Whereas, when adding a link between nodes 2 and 6, the removal of any single node in **Fig. 1(b)** does not affect the connectivity of the graph. Such removal of any single node also has less effect on the shortest paths in **Fig. 1(b)**. The highest degree in **Fig. 1(b)** is not increased by the link addition, and therefore more nodes and links can survive under targeted attacks on the highest degree nodes. The NTRLA problem with $q=1$ looks for a link adding which can maximize the survival fitness of the resulting graph. The investigation on adding one link to improve the survival fitness will also provide insights on how to dynamically add a set of limited links to increase the survival fitness at the most.



(a) An initial graph (b) An illustration by adding a proper link

Fig. 1. An example of network topology reconfiguration with one link addition

3.4 Np-cycle Structure

The star structure is efficient (the average shortest path length is small) but fragile w.r.t. the removal of the central node. On the contrary, the circle structure is robust w.r.t. the removal of any single node but inefficient (the average shortest path length is large). As will be shown in the following, the node-protecting cycle (np-cycle) offers a desirable combination of the circle and the star, and can be used to improve network robustness and the efficiency when the removal of the protected node.

Denote B_k as the set of neighbors of node k : $B_k = \{u \mid (u,k) \in E\}$. Following the notations in [24], we make the following definitions.

Definition 1 Node-protecting Cycle (np-cycle): A cycle x_k is said to be an np-cycle for protecting node k ($k \in V$) if and only if: (1) it contains node k ; and (2) for any pair of its neighboring nodes u and v ($u, v \in B_k$), there exists at least one path between node u and v that does not pass node k .

Definition 2 np-cycle order (*CO*): The order of an np-cycle is defined as $CO(x_k) = \sum_{u \in B_k} \varepsilon_{x_k}^u$, where $\varepsilon_{x_k}^u$ is an indicator representing whether or not node $u \in B_k$ belongs to x_k . If u belongs to x_k , $\varepsilon_{x_k}^u = 1$; otherwise, $\varepsilon_{x_k}^u = 0$.

By definition, if node u occurs more than once on x_k , it is counted only once when calculating $CO(x_k)$. $CO(x_k)$ indicates the number of neighboring nodes of k on x_k . It reflects the protection ability of x_k for node k .

Definition 3 np-cycle efficiency (*CE*): Let L_{x_k} denote the number of links on the np-cycle x_k , then the efficiency of x_k is defined as $CE(x_k) = \sum_{u \in B_k} \varepsilon_{x_k}^u / L_{x_k}$.

Definition 4 Eligible np-cycle: If the order of x_k is $|B_k|$, i.e., $CO(x_k) = |B_k|$, then the np-cycle x_k for node k is called eligible np-cycle.

Definition 5 Perfect np-cycle: If the np-cycle x_k for node k is an eligible np-cycle, and it further satisfies $CE(x_k) = |B_k| / (|B_k| + 1)$, then the x_k for node k is called a perfect np-cycle.

We define the normalized efficiency as $NCE(x_k) = \left(\sum_{u \in B_k} \varepsilon_{x_k}^u / L_{x_k} \right) \Big/ \left(|B_k| / (|B_k| + 1) \right)$.

For example in **Fig. 1(a)**, node 5 has no np-cycle, node 4 has no eligible np-cycle, and node 3 is fully protected by the perfect np-cycle 3-2-1-4 with $CO = 3$ and $NCE = 1$. So, the removal of node 3 has no effect on network connectivity and shortest paths length, while the removal of node 5 or node 4 would change network connectivity and the shortest paths length in **Fig. 1(a)**.

4. Theoretical Bases of Topology Reconfiguration Methods

In this section, we show several desirable properties of the network topology derived by limited link addition. We also show that the network topology reconfiguration optimization with limited link addition (NTRLA) problem is NP-hard. Therefore, two approximate solution methods are discussed in the next section.

4.1 Theoretical Analysis of Optimal Topology Reconfiguration

The primary goal of optimal topology reconfiguration is to make a network more robust under

"bad case" environment. So the topology reconfiguration should firstly solve the connectivity problem of a network.

For $\kappa \geq 2$, we say that a graph G is κ-connected if either G is a complete graph $K_{\kappa+1}$ or else it has at least κ+2 nodes and no set of κ-1 nodes separates it. Clearly, a graph is 2-connected iff it is connected, has at least 3 nodes, and not be separated by any one node. We have the following three important characteristics for *k*-connected graph [20]:

1. A graph is κ-connected if and only if it has at least two nodes and any two nodes can be connected by κ independent paths.
2. A graph is κ-connected if and only if any κ nodes are on a circle.
3. The least number of links of a κ-connected graph with n nodes is $\lceil \kappa n/2 \rceil$, where $\lceil \cdot \rceil$ is a ceiling function.

According to the characteristics, we have:

Theorem 1. For an arbitrary graph G, if every node has at least an eligible np-cycle, then G is 2-connected.

Proof. Because every node has at least an eligible np-cycle, by definition we know that all the nodes are on a circle. That means, any two nodes are on a circle in G. According to above the second characteristic for κ-connected graph, G is 2-connected. □

An important characteristic of SFN is the heterogeneity of the degree distribution, which makes SFN tolerant to random failures but extremely vulnerable to malicious attacks [1]. So we try to reduce the heterogeneity of the degree distribution of a network by applying topology reconfiguration methods with limited link addition.

Theorem 2. For any network topology, if the topology reconfiguration with the minimum number of link addition can reduce the heterogeneity of the degree distribution of the network at the most, then the topology reconfiguration is optimal for improving the connectivity of the network.

Proof. For a κ-connected graph, we have at least two nodes and any two nodes can be connected by κ independent paths. We know the minimal number of links of the *k*-connected graph is $\lceil \kappa n/2 \rceil$. Therefore, we have the degree of each node must be κ (n or κ is even) or at most one $\kappa+1$ (n and κ are both odd). From the relationship of the number of links and node degree in a given graph, we obtain that if node degree can be the minimum, then the number of links can be the minimum. This complete the proof. □

Theorem 2 tells us the degree distribution has a significant effect on network robustness. This is in agreement with the results of Albert et al. [1]. Thus, a basic design idea of topology reconfiguration algorithm based on the optimal configuration of np-cycles is to reduce the heterogeneity of the degree distribution in a reconfigured network.

As describe above, we introduce the survival fitness metric Γ which can be used to evaluate the effectiveness of topology reconfiguration methods with limited link addition. To provide analysis of the improvement on the robustness and the efficiency in the preferential configuring node-protecting cycle (PCNC) method (where the PCNC method is described in details in section 5.3), the following theorem provides the formal proof of Γ's qualities regarding the optimal configuration of np-cycles.

Theorem 3. $\Gamma(G)$ monotonically increases when any np-cycle order increases.

Proof. Let $k \in V$ be some node which has an np-cycle x_{k_i} and its order $CO(x_{k_i}) = r$ contains r neighbors. Thus, there is some node $b \in B_k$ for which $d_G(k,b) = 1$ (i.e., the shortest path length between node k and b in G is 1) and some node $u \in V$ for which $d_G(u,b) = d_G(u,k) + 1$ (specifically, this is true for some node $u \in B_k$), and no other path connects node u to b. Since

node b cannot be reached by any other nodes in V, it follows $d_{G_k}(v,b)=\infty$, $\forall v \in V-\{k\}$.

Let $\ell \notin E$ be a link that is added which extends the np-cycle x_{k_i} to form a new np-cycle $x_{k_{i+1}}$ containing node b. Thus, node k has an np-cycle $x_{k_{i+1}}$ in $G'=(V,E\cup\{\ell\})$ and the np-cycle $x_{k_{i+1}}$ order is $r+1$ (i.e., $CO(x_{k_i})=r+1$). To prove the above it is enough to show that $\Gamma(G)\le\Gamma(G')$.

Denote $d_{G'_k}(u,b)$ as the shortest path length between nodes u and b which bypasses node k in G'. As there exists a path on the np-cycle $x_{k_{i+1}}$ which bypasses node k between nodes u and b, we have $d_{G'_k}(u,b)\le d_{G_k}(u,b)=\infty$, $\forall u \in V-\{k\}$. It follows that $\Gamma(G)\le\Gamma(G')$. □

4.2 Complexity Analysis

A version of network topology reconfiguration optimization with limited link addition (NTRLA) problem can be stated following the NTRLA decision problem (NTRLA_D). Given a graph $G(V,E)$ with the available link resources set E' ($|E'|\le q$, $E'\cap E=\varnothing$), can we obtain a new graph $G'(V,E\cup E')$ such that $\Gamma(G')\ge\Gamma_0$?

We will show that NTRLA_D problem is NP-complete, and then show that the NTRLA problem is NP-hard.

Theorem 4. The NTRLA_D problem is NP-complete.

Proof. Given any instance graph G^* from G with limited link additions to the NTRLA_D problem, we need verify the calculation of $\Gamma(G^*)$ in polynomial time. According to the definition of survival fitness metric Γ, an $O(n^3\cdot m)$ test algorithm of computing the metric is proposed, which can confirm the above conjecture. So the NTRLA_D problem is in NP. That the NTRLA_D problem is NP-complete follows with the Restriction method, using the following Lemma 1 which considers the special case of finding an optimal configuration in a given connected graph. □

Define the special NTRLA_D problem (s_NTRLA_D) as: given a connected graph G with the available link resources set E' ($|E'|=q$), let the survival fitness $\Gamma=\eta^E$ (i.e., $\alpha=0$), can we obtain a new graph $G'(V,E\cup E')$ s.t. $\Gamma(G')\ge\Gamma_0$?

We use the MDEA problem [14] in our proof. The MDEA problem can be formulated: given a connected graph G and $k,D\in\mathrm{N}$, can we add k links to G to obtain a new graph G' s.t. G' has a diamter smaller than or equal to D?

Lemma 1. The s_NTRLA_D problem is NP-complete.

Proof. For any instance of MDEA problem, we suppose that a graph $G'(V,E\cup E')$ from $G(V,E)$ by adding the link set E' is obtained, then G' has the diameter D' and $D'\le D$. For $D'\le D$, there is $1\le d_{G'}(u,v)\le D$, $u\ne v\in V$. We have $\frac{1}{D}\le\frac{1}{d_{G'}(u,v)}\le 1$. So, $\Gamma(G')\ge\frac{1}{D}$. Therefore, the s_NTRLA_D problem can be regarded as an extension of the MDEA problem and it is NP-complete. □

We show that the network topology reconfiguration with limited link addition (NTRLA) problem is NP-hard. Therefore, two approximate link addition methods are developed in the following section.

5. Topology Reconfiguration Methods with Link Addition

Existing degree-based link addition methods [7][13] are limited to strike a balance between the robustness and the efficiency, especially considering the cost of link addtion resources (see simulation results in section 6). Hence, we propose two methods: 1. adding a link between the node with the highest degree-fitness and the node with the lowest degree; 2. adding a link by applying the preferential configuration node-protecting cycle (PCNC).

5.1 Degree-Based Link Addition Methods

We review four existing degree-based link addition methods [7][9] and [11][12][13]. All these methods require only local information, i.e., node degree.

Lowest Degree Preference Addition (LDP) [7][9]

The LDP method adds a link by connecting two unconnected nodes that have the lowest degrees in the network. From the analysis in Zhao et al. [9], we know that the LDP method is a good way to enhance the robustness because it can drastically reduce the heterogeneity of degree distribution when adding links.

Random and Lowest Degree Preference Addition (RLP) [12]

The PLP method adds a link between the node with the lowest degree and a random other one. The idea is to improve the potential communication bottleneck. We know that there are a lot of nodes with low degree in the scale-free network, but a few nodes may have very high degrees. It is intuitive that connecting high-degree nodes are beneficial to the rapid improvement in communication efficiency compared with low-degree nodes.

Random and Highest Degree Preference Addition (RHP) [13]

The RHP method adds a link between the node with the highest degree and a random other node. It is possible that, after several link additions, the highest degree node is connected with all the other network nodes. Hence, no link can be added any more. To prevent this situation a modification of the method is done, if the maximum degree node is already fully connected with the other nodes, links will be added one by one between the second highest degree node and a random other node [13].

Degree Probability Preference Addition (DPP) [11]

The DPP method adds a link based on the preferential attachment which is the basic element of the conventional preferential linking model. The link addition selects a node with the probability proportional to the degree of the node. This means that centralized nodes are more likely to be selected than the other nodes. Let k_i be the degree of node i. Then the preferential connection probability of the link from node i to j is

$$P_i = \frac{(k_i)^\tau}{\sum_j (k_j)^\tau}, \tag{5}$$

where τ indicates the preference. When $\tau = 1$ Eq. (5) reduces to the conventional preferential linking. In our simulation experiment $\tau = 2$.

5.2 Degree-fitness Based Link Addition Method

As afore mentioned, link addition methods that prefer high-degree nodes improve network

efficiency but have low robustness under "bad case". On the contrary, link addition methods that prefer low-degree nodes improve robustness but do not improve the communication efficiency effectively. In order to make a trade off, we present a degree-fitness parameter to guide the degree-based link additions. Define the degree-fitness of node i as

$$F_i = \frac{\sum_{j \in B_i} k_j}{|\langle k \rangle - k_i| + \delta}, \tag{6}$$

Where B_i is the set of direct neighbors of node i, k_i is the degree of node i, and $\langle k \rangle$ is the mean value of node degree in the network, δ is a given arbitrarily small positive value. From the definition of degree-fitness of a node, we can conclude that the node with large degree-fitness is the direct neighbor of a high-degree node. Whereas high-degree nodes are easy targets under targeted attacks, their low-degree neighbors will not be easily removed. The node with large degree-fitness is comparatively more important because it has high transmission efficiency.

As an example, according to Eq. (6), F of nodes in **Fig. 2** are shown in **Table 2** ($\delta = 0.0001$). To demonstrate the characteristics of F, degrees of these nodes are also given. For the robustness, we see that node 12 with a larger degree-fitness is more important than the node with the same degree (such as node 9) and a higher degree (such as node 10).

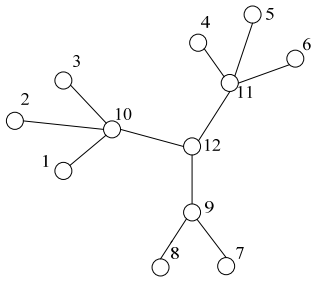


Fig. 2. An example of degree-fitness of nodes

According to the introduction of the degree-fitness parameter, we develop a degree-fitness based preference addition (DFP) method. The DFP method adds a link between the node with the highest degree-fitness and a node with the lowest degree in the network. The idea is to connect the nodes with high communication efficiency with low-degree nodes in order to improve both the robustness and the efficiency of the network.

Table 2. Comparison of degree-fitness and degree of nodes

Node No.	Degree-fitness (F) of Nodes	Degree of Nodes
1,2,3	4.800	1
4,5,6	4.800	1
7,8	3.600	1
9	4.285	3
10,11	2.769	4
12	9.428	3

5.3 Preferential Configuration Np-Cycle Link Addition Method

We develop a preferential configuration node-protecting cycle (PCNC) method to optimize network topology reconfiguration with limited link addition. We first discuss the link contribution metric used in the PCNC. Then we describe PCNC method for sequential link additions in details.

5.3.1 Link contribution metric

Different link addition process will produce different results in improving network robustness and the efficiency. We introduce a metric called the link contribution which measures the contribution of a link in improvement the protection ability of np-cycles. In the following, we formally define the link contribution metric in the PCNC.

Given a graph $G=(V,E)$, the E' is the set of links to add. Let v_k denote the importance of node $k \in V$. Centrality measures are generally used to characterize the improtance of the node, such as degree centrality, closeness centrality, betweenness centrality, etc [23].

Define the link contribution metric of a link ℓ as $Ctr(\ell)$, which indicates the contribution of link ℓ to improve the protection ability of np-cycles that contain link ℓ.

$$Ctr(\ell)=\sum_{k\in V} v_k \cdot W(\pi_k^{\ell}), \tag{7}$$

where π_k^{ℓ} is a set of np-cycles for protecting node k, and each candidate np-cycle $\pi_{k_i}^{\ell} \in \pi_k^{\ell}$ contains link ℓ. $W(\pi_k^{\ell})$ is the total weight of np-cycles that can be used to protect node k and contain link ℓ, i.e.,

$$W(\pi_k^{\ell})=\sum_{i\in\{1,2,\ldots,|\pi_k^{\ell}|\}}\sum_{e\in\pi_{k_i}^{\ell}} w_e \ , \tag{8}$$

where w_e denotes the weight of a link $e\in(E\cup E')$, which can be used to characterize the improtance of the link. $|\pi_k^{\ell}|$ is the number of np-cycles in π_k^{ℓ}, and $\sum_{e\in\pi_{k_i}^{\ell}} w_e$ is the total weight of links that are contained in the np-cycle $\pi_{k_i}^{\ell} \in \pi_k^{\ell}$.

From above, we see that a higher link contribution metric means a higher utilization rate of a link. This motivates the PCNC method that adds links according to Eq. (7).

5.3.2 Description of PCNC

The idea of PCNC method is to add links to construct np-cycles so that more nodes are protected. We model the NTRLA problem basing link contribution of links that are added as follows:

$$\text{Max} \sum_{\ell\in E'} \frac{Ctr(\ell)}{c(\ell)} \tag{9}$$

$$\text{s.t.} \ \sum_{\ell\in E'} c(\ell) \le C_0 \ , \tag{10}$$

where $c(\ell)$ is the cost to add link ℓ, $Ctr(\ell)/c(\ell)$ is the contribution efficiency of link ℓ for configuring np-cycles, and C_0 is the given link resource constraint. The objective function Eq. (9) maximizes the total contribution efficiency of added links, and constraint Eq. (10) ensures that the total cost is within the given resource constraint.

We propose a heuristic algorithm based on the idea of PCNC method. In the algorithm, the process of link additions has three purposes. The first is to enhance the robustness against targeted attacks. The second is to improve the efficiency of the network. The third reducing the reconfigured cost is also an important factor in the process of link addition.

Before running the algorithm, we need to do some preprocessing of the set of target nodes (protected nodes), and remove the nodes with less than 2 degree (the node with only one degree obviously can not be covered by any np-cycle) from the set of target nodes. The main steps of the algorithm are described below:

Step 1: First add the set of links $E^{'}$ to generate an extended topology structure $G^{'}\left(V, E \cup E^{'}\right)$ (maybe obtain a complete graph). Then predefine the weight w_e for every link $e \in \left(E \cup E^{'}\right)$ in $G^{'}$, which is as the following equation,

$$w_e = \begin{cases} 0 & e \in E \\ weight(e) & e \in E^{'} \end{cases}.$$

Here, $weight(e)$ represents the weight of link e, which is assumed to be the inverse of the average degree of two end-nodes of link e in this paper. In addition, we define that $w_e\,(e \in E)$ is zero to make links of G with priority being selected to configure np-cycles, which can reduce the extra reconfigurable resource consumption.

Step 2: Search the set of np-cycle π_k^{ℓ} which can be used to protect node k and contains link $\ell \in E^{'}$. First we get a target node and its adjacency nodes, and search the minimal weight path, the sub-minimal weight path,…, between any two nodes of adjacency nodes of the target node. Then we connect target nodes with paths between their adjacency nodes to constitute np-cycles.

Step 3: Computer $Ctr(\ell)$ of every link $\ell \in E^{'}$. Because the np-cycle for every target node obtained by step 2 is not only one, we rank links in $E^{'}$ according to their link contribution metric.

Step 4: Select and configure preferential links with high $Ctr(\ell)$ into the network to improve network robustness and efficiency.

The flowchart of the above algorithm description is in **Fig. 3**.

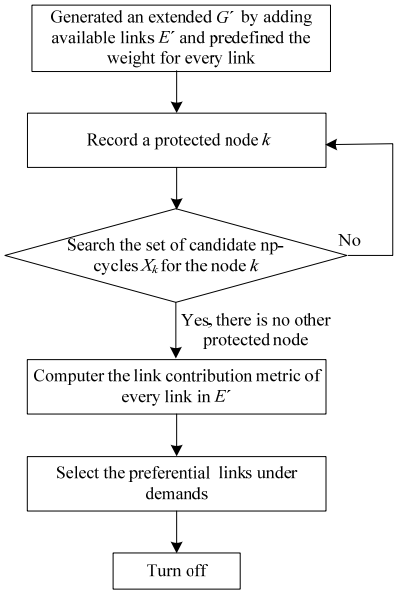


Fig. 3. The flowchart of PCNC algorithm

The number of possibilities of adding a link to a graph with n nodes and m links is $\binom{n}{2}-m$. For large realistic (hence sparse) networks, it is impracticable to compare all these possibilities and to find the optimal one when add a set of links. The NTRLA problem is NP-hard, and we present a heuristic algorithm to solve this problem approximately. Compared with searching the complete solution space of link additions which is exponential time complexity, the heuristic algorithm search local np-cycles for target nodes, of which the time complexity is $O(n^3)$.

6. Simulation Results

In this section, we compare the six link addition methods numerically via simulation experiments. The number constraint of link addtion is same for different link addition method in the simulation. For a given network of n nodes and m links, the constraint of adding links is a few percent of $\binom{n}{2}-m$. Small-scale network topologies other than large-scale network topologies are used to evaluate different link addition methods due to two main factors listed as follows. (1) For large-scale networks, it is impracticable to compare all possibilities of adding links and find the optimal the addition. We also showed that the network topology reconfiguration optimization problem with limited link additions is NP-hard; (2) The larger the network, the longer the set of np-cycles will take to be computed, and the lower algorithm performance will take to be. So, the PCNC method is suitable for small and medium scale networks. Two small-scale networks are considered, namely a router network and a BA network.

6.1 Router Network Topology

We adopt the network manipulator (NEM), which generates the topology most similar to the realistic Internet shown in [26]. A route network topology is generated using the NEM as shown in **Fig. 4**. There are 36 nodes and 37 links. Suppose that we can add at most $\lceil 2\% \times \binom{36}{2} - 37 \rceil = 12$ links. It noted that every curve in this subsection represents the averages over 100 replications.

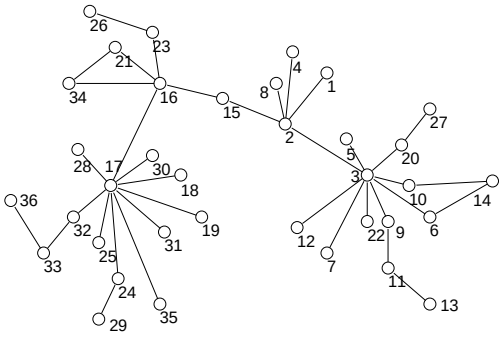


Fig. 4. A router network topology

We apply the six methods in section 5 to add links and show their effectiveness under the survival fitness metric for $\alpha = 0, 0.2, 0.5, 0.8, 1$ as showed in **Fig. 5**. We make the following remarks. Remark 1. RHP method and DPP method are superior to PCNC method, and the PCNC is superior to RLP, DFP, LDP methods when $\alpha = 0$ in **Fig. 5(a)**. Remark 2. LDP

method and RLP method are superior to RHP method and DPP method, the DFP has advantage over LDP, RLP, RHP and DPP methods, and the PCNC is better than these degree-based methods when $\alpha = 1$ in **Fig. 5(e)**. Remark 3. When α increases, RHP method and DPP method have poor survival fitness compared with other methods; The survival fitness of RLP method and LDP method can be improved gradually; The DFP can be enhanced significantly compared with other degree-based methods, and the PCNC has a advantage compared with these degree-based methods under the increasingly bad environment.

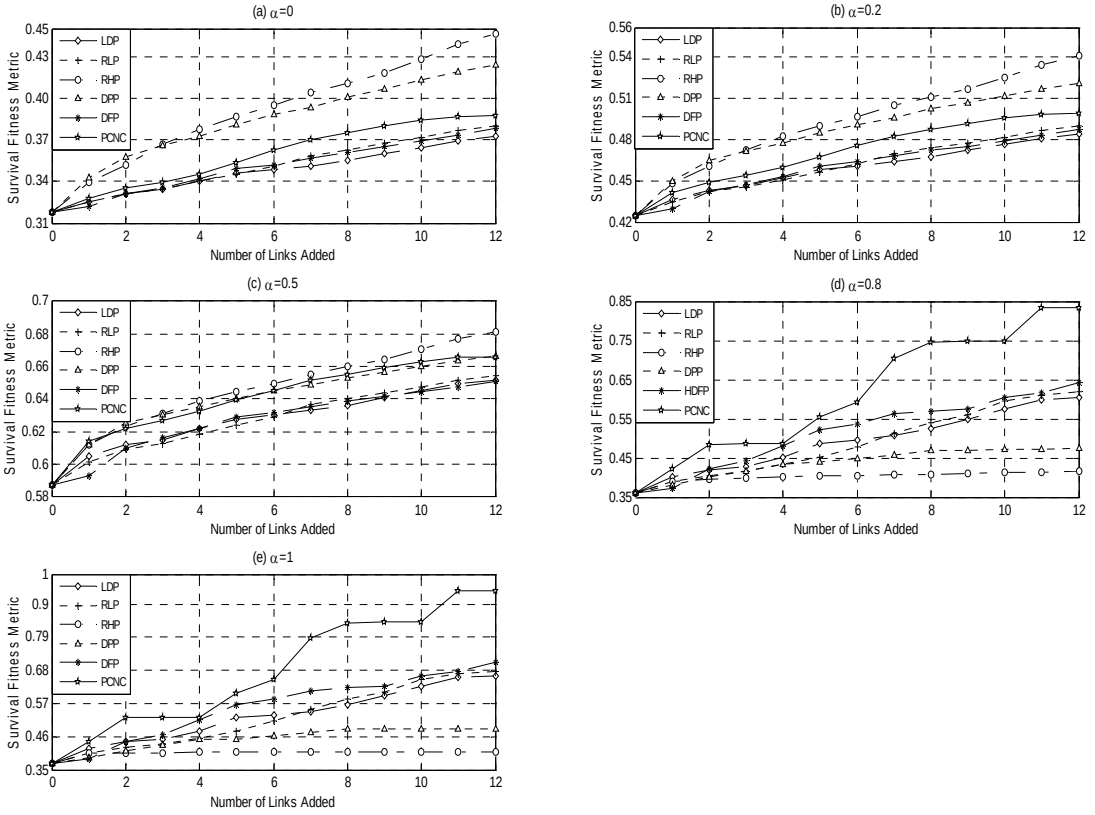


Fig. 5. Variation of the survival fitness by applying different link addition methods

In **Fig. 6** we compare the variation of the survival fitness of different link addition methods based on different environment parameter α .We make the following remarks. Remark 1. When $\alpha \le 0.5$, the upstrend of the survivial fitness is slow for these methods. Remark 2. When $\alpha > 0.5$, the upstrend of the survival fitness become apparent for LDP, RLP, DFP and PCNC methods. Especially for PCNC method, there exists increase advantage for "bad case" survival fitness ($\alpha > 0.5$), while DPP method has a very slow growth and RHP method is almost no increase in **Fig. 6(f)**.

We use the Standard Deviation of Node-degree (S) [27] to measure the heterogeneity of the degree distribution of the resulting networks by applying different link addition method. **Fig. 7** shows the variation of S metric of these methods with the increase of the number of links. We make the following remark. The S metric of RHP method and DPP method increase, and other four methods, LDP, RLP, DFP and PCNC decrease as the number of links increases. It appears an interesting fact that the downtrend of the S of the PCNC is near to the LDP. This observation is consistent with the principal idea of PCNC method which

preferentially selects candidate links between low-degree nodes.

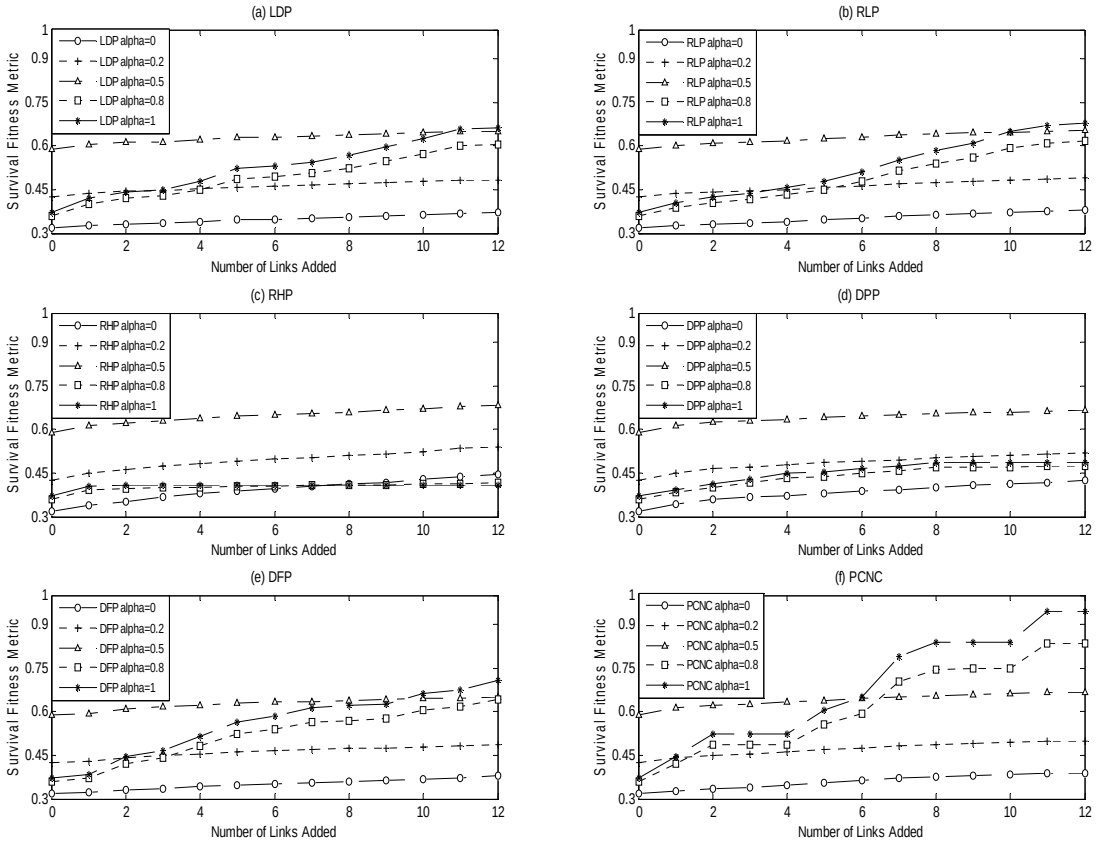


Fig. 6. Variation of the survival fitness based on different environment parameter α

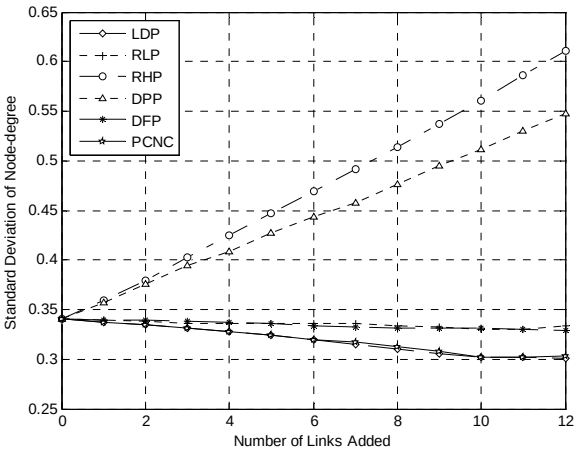


Fig. 7. Variation of Standard Deviation of Node-degree by applying different link addition methods

6.2 BA Model Network Topology

We use BA model to generate a network which starts with m_0 nodes. The preferential linking

is constructed in two stages to get a sparse network topology. The first is to add each node with m_0 links, and the second is to add each node with $m_0 - 1$ links. The preferential linking will select a node by the probability which is proportional to the degree of the node.

We perform the same experiment in a BA network with $n = 80$ and $m_0 = 2$. For the BA network with 80 nodes and 90 links, the constraint of adding links is a few percent of $\binom{80}{2} - 90$. Suppose that we can add at most $\lceil 1\% \times \binom{80}{2} - 90 \rceil = 31$ links. Every curve in this subsection represents the averages over 50 replications.

Fig. 8 shows the variation of the survival fitness metric for $\alpha = 0, 0.2, 0.5, 0.8, 1$ by applying different link addition methods. When $\alpha = 0, 0.2, 0.5$, the remarks from the BA network are consistent with the remarks from the router network in **Fig. 5**. When $\alpha = 0.8, 1$, PCNC method has little superiority with the same number of links in **Fig. 8(e)**, while it has obvious superiority in the router network. For DFP method, we have the same remark that the DFP has certain superiority in degree-based link addition methods under the increasingly bad environment.

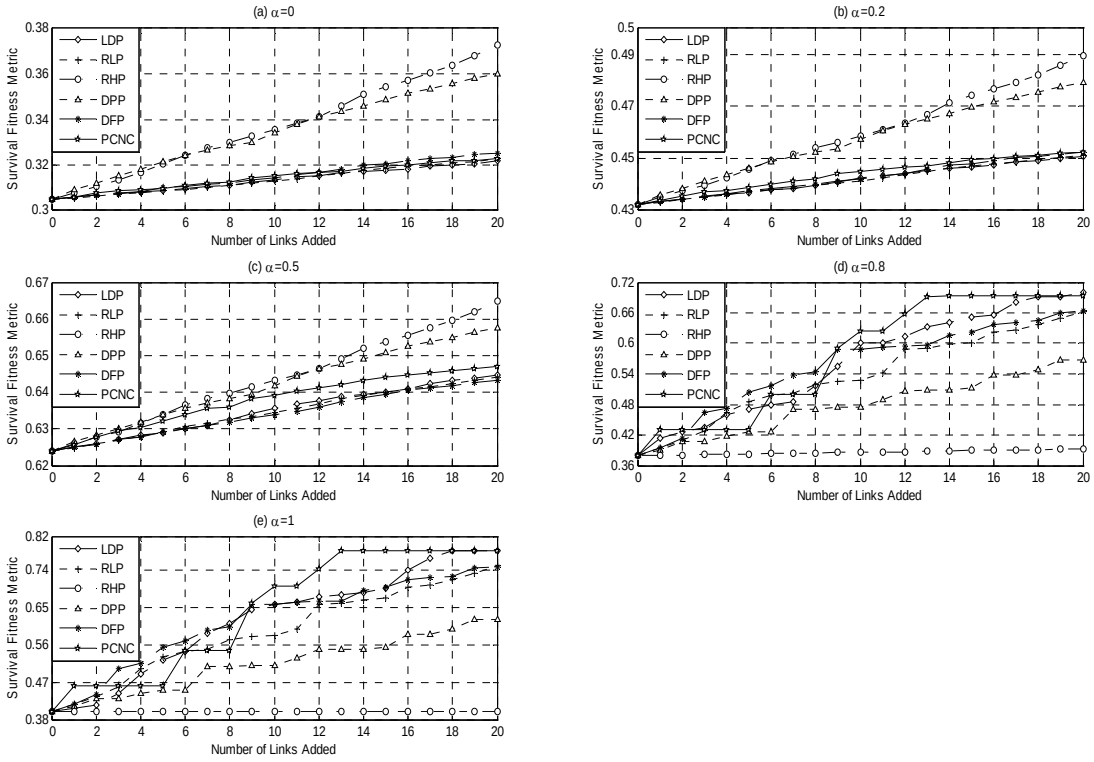


Fig. 8. Variation of the survival fitness by applying different link addition methods

We show the variation of the survival fitness of the six methods based on different environment parameter α in **Fig. 9**. When α increases from 0 to 1, the remarks from the BA network are consistent with the remarks from the router network in **Fig. 6**.

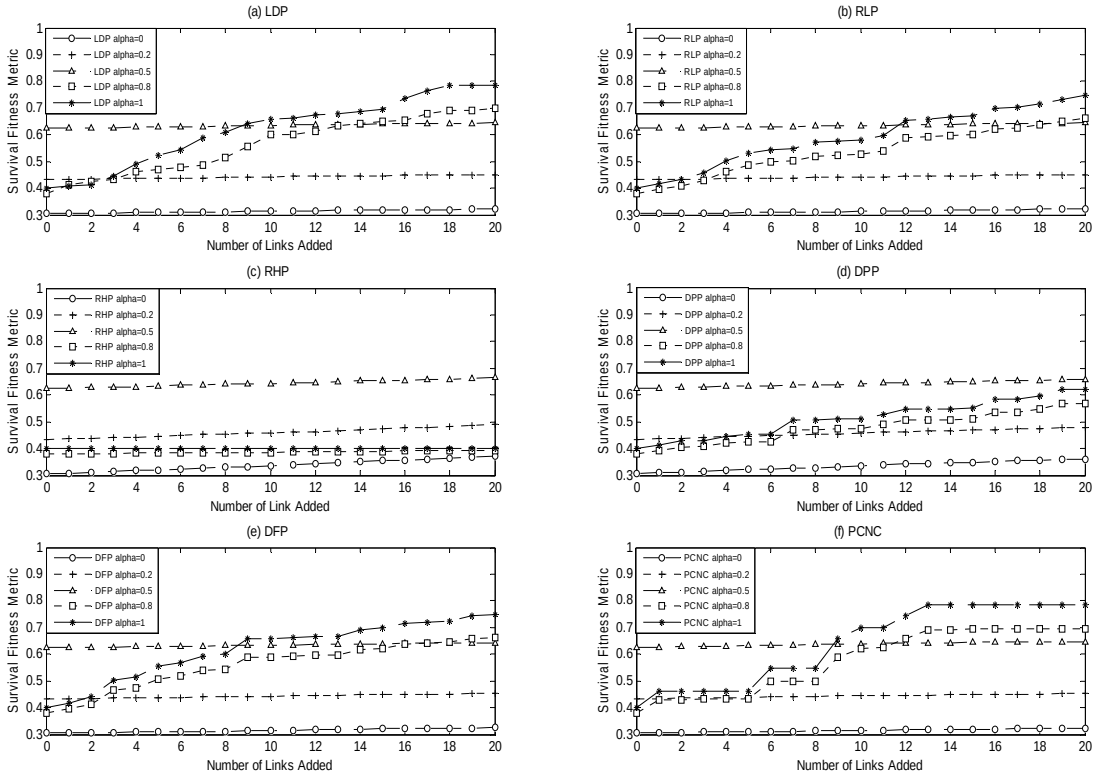


Fig. 9. Variation of the survival fitness based on different environment parameter α

Fig. 10 shows the variation of S metric of different link addition methods with the increase of the number of links. We obtain the same remark from the BA network compared with the remark from **Fig. 7**. We see that the downtrend of the S by applying PCNC method is near to LDP method for the given BA model network topology. We can conclude that PCNC method has the same superiority with LDP method in reducing the heterogeneity of the degree distribution, while RHP method and DPP method will make the network more vulnerable under targeted attacks.

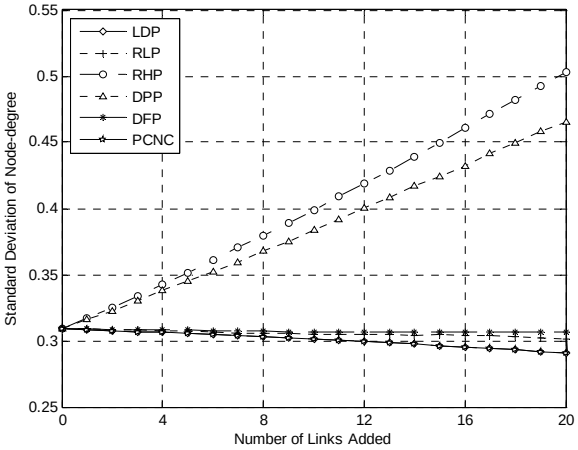


Fig. 10. Variation of Standard Deviation of Node-degree by applying different link addition methods

6.3 Discussion

From the above simulation results, we see that preferential adding links between high-degree nodes (as RHP method and DPP method) are beneficial to improve the efficiency. However, such link addition increase the dependence of a network on its hubs, making it more fragile under targeted attacks. In constant, preferential adding links between low-degree nodes (as LDP method and RLP method) are beneficial to enhance robustness under attacks. The heterogeneity of the degree distribution can be reduced by adding links between low-degree nodes in **Fig. 7** and **10**, which do enhance network robustness [1].

We see that DFP method is better than LDP, RLP, RHP and DPP methods with respect to the survival fitness under the increasingly bad environment in **Fig. 5** and **8**. We can conclude that high-degree nodes are easily removed and their neighbors with high degree-fitness are surviving against targeted attacks. This demonstrates that the robustness and the efficiency can be improved by applying DFP method with limited link addition under targeted attacks.

From **Fig. 5** and **8**, we see that PCNC method has significant advantages for regulating the robustness and the efficiency when compared with degree-based link addition methods. We observe that the PCNC can effectively reduce the heterogeneity of the degree distribution in **Fig. 7** and **10**. In addition, the PCNC improves the efficiency with limited link addition in **Fig. 5(a)** and **8(a)**. In summary, PCNC method can efficiently enhance the attack tolerance and the communication efficiency of scale-free networks at the cost of least link addition resources.

7. Conclusion

In this paper, we investigated how to optimize an existing network in both aspects of attack tolerance and communication efficiency with limited link addition. For this purpose, we defined the survival fitness metric to quantitatively characterize the capability of a network both the attack tolerance and the communication efficiency. We showed that the network topology reconfiguration optimization with limited link addition (NTRLA) problem is NP-hard. A degree-fitness based preference addition (DFP) method and a preferential configuration node-protecting cycle (PCNC) method are then developed. Simulation results show that PCNC method outperforms degree-based link addition methods, LDP, RLP, RHP, DPP and DFP, in terms of improvement on the robustness and the efficiency with limited link addition. It is also shown that DFP method is better than LDP, RLP, RHP and DPP methods under targeted attacks, and PCNC method performs the best in both aspects of attack tolerance and communication efficiency at the cost of least link addition resources.

References

[1] R. Albert, H. Jeong and A.L. Barabasi, “Error and attack tolerance of complex networks, ” *Nature*, vol.406, pp.378-382, Jul.2000. Article (CrossRef Link)

[2] S.H. Strogtz, “Exploring complex networks,” *Nature*, vol.410, pp.268–276, 2001. Article (CrossRef Link)

[3] US Dept of Homeland Security United States, “National emergency communications plan,” United States, August, 2008. Article (CrossRef Link)

[4] A.X.C.N. Valente, A. Sarkar and H.A. Stone, “Two-peak and three-peak optimal complex networks,” *Phys. Rev. Lett.*, vol.92, no.2, 2004. Article (CrossRef Link)

[5] J. Carlson and J. Doyle, “Highly optimized tolerance: Robustness and power laws in complex systems,” *Phys. Rev. Lett.*, vol.84, no.11, pp.2529-2532, 2000. Article (CrossRef Link)

[6] B. Bollobas and O. Riordan, "Robustness and vulnerability of scale-free random graphs," *Internet Math.*, vol.1, no.1, pp.1-35, 2003. Article (CrossRef Link)
[7] A. Beygelzimer, G. Grinstein, R. Linsker and I. Rish, "Improving network robustness by edge modification," *Physica A*, vol.357, pp.593-612, 2005. Article (CrossRef Link)
[8] L.P. Chi, "On the repair strategy and correlation properties of complex networks under attacks," *Ph.D. Dissertation*, Central China Normal University, 2006. Article (CrossRef Link)
[9] J.C. Zhao and K. Xu, "Enhancing the robustness of scale-free networks," in *Journal of Physics A: Mathematical and Theoretical*, vol.42 , no.19, 2009. Article (CrossRef Link)
[10] Y. Sato, S. Ata and I. Oka, "A strategic approach for re-organizing the Internet topology by applying social behavior dynamics," in *Journal of Network and Systems Management*, vol.17, pp.208-229, 2009. Article (CrossRef Link)
[11] K. Sekiyama and H. Araki, "Network topology reconfiguration against targeted and random attack," in *International Workshop on Self-Organizing System*, pp.119-130, 2007. Article (CrossRef Link)
[12] H.J. Wang and P.V. Mieghem, "Algebraic connectivity optimization via link addition," in *Proc. of International Conference on Bio-Inspired Models of Network, Information, and Computing Systems*, Nov.2008. Article (CrossRef Link)
[13] M.C.S. Martinez, "Robustness optimization via link additions," *Ph.D. Dissertation* , Jul.2009. Article (CrossRef Link)
[14] A.A. Schoone, H.L. Bodlaeder and J.V. Leeuwen, "Diameter increase caused by edge deletion," in *Journal of Graph Theory*, vol.11, no.3, pp.409-427, 1987. Article (CrossRef Link)
[15] B. Shargel, H. Sayama, I. Epstein and Y. Bar-Yam, "Optimization of robustness and connectivity in complex networks," *Phys. Rev. Lett.*, vol.90, no.6, Feb.2003. Article (CrossRef Link)
[16] G. Paul, T. Tanizawa, S. Havlin and H. Stanley, "Optimization of robustness of complex networks," *Eur. Phys. J. B*, vol.38, pp.187–191, 2004. Article (CrossRef Link)
[17] A. Ranjitkar and Y.B. Ko, "A Distributed Web-Topology for the Wireless Mesh Network with Directional Antennas," *KSII Transactions on Internet Information Systems*, vol.5, no.1, Jan.2011. Article (CrossRef Link)
[18] W.D. Grover and D. Stamatelakis, "Cycle-Oriented distributed preconfiguration: ring-like speed with mesh-like capacity for self-planning network restoration," in *Proc. of IEEE International Conference on Communications*, pp.537-543, Jun.1998. Article (CrossRef Link)
[19] T.F. Zhao, "Research on pre-configuration cycle algorithms in survivable mesh networks," *Ph.D. Dissertation*, University of Electronic Science and Technology of China, 2007. Article (CrossRef Link)
[20] B. Bollobas, Modern Graph Theory, first ed., Beijing World Publishing Corporation, Beijing, 2003. Article (CrossRef Link)
[21] Y. Singer, "Dynamic Measure of Network Robustness," in *2006 IEEE 24th Convention of Electrical and Electronics Engineers*, 2006. Article (CrossRef Link)
[22] S.H. Kim and B-h. Roh, "Fast detection of distributed global scale network attack symptoms and patterns in high-speed backbone networks," *KSII Transactions on Internet Information Systems*, vol.2, no.3, Jun.2008. Article (CrossRef Link)
[23] V. Latora and M. Marchiofi, "A measure of centrality based on network efficiency," *New Journal of Physics*, Jun.2007. Article (CrossRef Link)
[24] L. Li, Q-S. Jia, H.T. Wang, R.X. Yuan and X.H Guan, "A Near Optimal Solution for Network Topology Reconfiguration with Limited Link Resources," in *World Conference on Intelligent Control and Automation*`, pp.489-494, Jun.2011. Article (CrossRef Link)
[25] V. Venkatasubramanian, S. Katare, P.R. Patkar and F.P Mu, "Spontaneous emergence of complex optimal networks through evolutionary adaptation," *Computers and Chemical Engineering*, pp.1-21, 2004. Article (CrossRef Link)
[26] D. Magoni, "Nem: a software for network topology analysis and modeling," in *10th International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, USA, pp.364-371, Oct.2002. Article (CrossRef Link)

[27] T. Nishikawa, A.E. Motter, Y.C. Lai and F.C. Hoppensteadt, "Heterogeneity in oscillator network: are samaller worlds easier to synchronize?," *Phys. Rev. Lett.*, vol.91, no.1, Jul.2003. Article (CrossRef Link)

Li Li received the B.S. and M.S. degrees in computer science from Shanxi Normal University, Xi'an, China, in 2002, 2005 respectively. She is currently Ph.D. candidate with the Systems Engineering Institute and SKLMS Laboratory of Xi'an Jiaotong University. Her research interests include network security, reconfigurable network and survivable network.

Qing-Shan Jia received the B.E. degree in automation in July 2002 and the Ph.D. degree in control science and engineering in July 2006, both from Tsinghua University, Beijing, China. He is currently an Associate Professor at the Center for Intelligent and Networked Systems (CFINS), Department of Automation, Tsinghua University, Beijing, China. He was a Visiting Scholar at Harvard University in 2006, and a Visiting Scholar at the Hong Kong University of Science and Technology in 2010. His research interests include theories and applications of discrete event dynamic systems (DEDS's) and simulation-based performance evaluation and optimization of complex systems.

Xiaohong Guan received the B.S. and M.S. degrees in automatic control from Tsinghua University, Beijing, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs, in 1993. From 1993 to 1995, he was a Consulting Engineer at PG&E. From 1985 to 1988, he was with the Systems Engineering Institute, Xi'an Jiaotong University, Xi'an, China. From January 1999 to February 2000, he was with the Division of Engineering and Applied Science, Harvard University, Cambridge, MA. His research interests include allocation and scheduling of complex networked resources, network security, and sensor networks.

Hengtao Wang received the B.S. degree in Automation from Xi'an Jiaotong, Xi'an, China, in 2007. He is currently a Ph.D. candidate with the Center for Intelligent and Networked Systems of Tsinghua University. His research interests include information fusion, reconfigurable network and survivable network.