Physica A 419 (2015) 622–629

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Vulnerability of complex networks under path-based attacks

Cun-Lai Pu [*], Wei Cui

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

HIGHLIGHTS

- We study the longest-path attacks on complex networks.
- We propose two approximating algorithms for finding the approximately longest path.
- Homogeneous networks are fragile to longest-path attacks.

ARTICLE INFO

Article history:
Received 17 July 2014
Received in revised form 2 September 2014
Available online 14 October 2014

Keywords:
Vulnerability
Path
Complex networks

ABSTRACT

We investigate vulnerability of complex networks including model networks and real-world networks subject to path-based attacks. Specifically, we remove approximately the longest simple path from a network iteratively until there are no paths left in the network. We propose two algorithms, the random augmenting approach (RPA) and the Hamilton-path based approach (HPA), for finding the approximately longest simple path in a network. Results demonstrate that steps of longest-path attacks increase with network density linearly for random networks, while exponentially increasing for scale-free networks. The more homogeneous the degree distribution is, the more fragile the network, which is different from the previous results of node or edge attacks. HPA is generally more efficient than RPA in the longest-path attacks of complex networks. These findings further help us understand the vulnerability of complex systems, better protect complex systems, and design more tolerant complex systems.

1. Introduction

Vulnerability is one of the fundamental properties in many natural and man-made complex systems [1–4]. For instance, genetic defects induce cell lesions [5], router failures interrupt the Internet [6], incidents or terrorist attacks cause collapse of public transport systems [7], and malfunction of a power station results in large-scale blackouts [8]. The vulnerability of complex systems is determined by the underlying networks in which nodes represent individuals in complex systems and edges represent the interactions of individuals. In past decades, many breakthroughs have been made on understanding individual components' disturbance's effects on the overall function of complex networks. Albert et al. [9] found that the Internet and WWW (World Wide Web) are very robust against random failures, but quite fragile in intended attacks. This robust yet fragile property was confirmed in much larger maps of the WWW as well as many other scale-free networks [1], and percolation processes on model networks were introduced to explain this property [10,11]. Since then many researchers studied the robustness of model networks and real-world networks subject to node attacks and edge attacks [1,2], in which nodes or edges are removed in order of degree [12,13] or other centrality measures such as betweenness [14,15],

[*] Corresponding author.
E-mail address: pucunlai@njust.edu.cn (C.-L. Pu).

http://dx.doi.org/10.1016/j.physa.2014.10.038

eigenvector [16], PageRank [17], *etc.* A few localized failures or attacks may cause cascading failures and lead to breakdown of the whole system which has been observed in Internet [18,19], power grids [20–25], financial systems [26], *etc.* Recently, researchers studied robustness of temporal or time-varying networks by extending the measures of centrality and robustness with temporal properties [27,28]. Furthermore, current attention focuses on the percolation processes on multiplex or interdependent networks which better model catastrophic events in power grids, transport systems and many other interdependent systems [8,29,30].

However, attacks are not merely limited to node or edge attacks that have been widely studied in the literature. For instance, hurricanes always have a large-scale effect on the public transport networks, terrorists usually prefer much larger scale attacks, and drugs always affect many targets. Therefore, we need to understand the effect of attacks on larger parts of complex networks other than nodes and edges, like path attacks. A simple path composed of nodes and edges is a common subpart of a network. In this context, "path" always means "simple path" which indicates that a node appears at most once in a path. This restriction is consistent with the attacks problem. In a network, a straightforward measure of paths is path length. Therefore, a natural question is how removals of the longest paths affect the function of a network. Finding the longest paths in general graphs is a well-known NP-hard problem in the literature [31]. Many approximation algorithms are proposed to approximate the longest paths including color-coding method [32], divide and conquer approach [33], algebraic approach [34], *etc.* In this paper, we propose two algorithms, RPA and HPA, to find the approximately longest paths during the attack processes. The efficiency of the two algorithms in the attacks is determined by the decay rates of the largest components in networks. The robustness of model networks and real-world networks to longest-path attacks are reflected by the steps of the iterative attacks.

2. Model of longest-path attacks

In a large network, there is usually a huge number of paths which is much larger than the number of nodes and edges. In some sense, paths can be thought of as the combinations of nodes and edges. When attacking a network, we prefer to remove the critical paths so that the removal of these paths significantly degenerate the function of the network. There are plenty of centrality measures in literature [1], but most of them are for nodes and edges, few are connected with paths. Path length, a natural measure for a path, is defined as the number of edges in a path. In our path-attack process, each time we remove the longest path from the network based on specified algorithms. Note that when attacking a path, we just remove all its edges from the network, but keep all its nodes in the network. The removal of longest path continues until there are no paths in the networks. In fact there are no edges in the final network, since edges are paths of length 1, and they would be removed from the network if there were any left.

3. Measures of longest-path attacks

The size of the largest component in a network reflects the communication capability of a network. Therefore we compute the size of largest component S during each step of the longest-path attacks, which indicates the efficiency of an approximating algorithm. The larger the decay rate of the largest component, the more efficient the approximating algorithm. Also, we count the total step of the longest-path attacks T, which reflects the tolerance of a network subject to longest-path attacks, and more steps means a more robust network against longest-path attacks.

4. Algorithm for approximating the longest path

In each iteration of the attacks process, we need to find the longest path in the corresponding network, which is a famous NP-hard problem. Here we present two algorithms, random path augmenting approach (RPA) and Hamilton-path based augmenting approach (HPA), to find the approximate longest path in a network.

4.1. Random path augmenting approach

RPA is a very simple algorithm for constructing a path in a network. First, RPA randomly selects a node from the network as the root node. Second, RPA chooses a neighboring node of the root node as the second node of the path, and then chooses a neighboring node of the second node as the third node of the path, and so on. The path augmenting process continues until it cannot find a new appropriate node. The main procedure of RPA is as follows:

(1) Maintain a path P which is a node sequence (initially null), and node set V_0 (initially null).
(2) Randomly select a node r as the root node.
(3) Append r to the end of P, and add r to V_0.
(4) Find a neighboring node of r named q, such that $q \notin V_0$.
(5) IF q exists, take q as r, return to (3). Otherwise, the algorithm stops.

4.2. Hamilton-path based augmenting approach

HPA employs the idea of approximating the Hamilton path in a network [35]. Initially, HPA randomly selects a node as the root node, and augments a path as long as possible in both directions of the root node. In each iteration, HPA first generates a