Information Sciences 422 (2018) 290–304

Contents lists available at ScienceDirect

Information Sciences

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

A two-level learning strategy based memetic algorithm for enhancing community robustness of networks

Wenfeng Liu, Maoguo Gong*, Shanfeng Wang, Lijia Ma

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China

ARTICLE INFO

Article history:
Received 30 November 2016
Revised 28 August 2017
Accepted 4 September 2017
Available online 8 September 2017

Keywords:
Complex network
Memetic algorithm
Robustness
Community structure

ABSTRACT

Community structure is a natural and inherent property of complex networks which can reflect their potential functionality. When the robustness of a network is improved, its community structure should be preserved as much as possible. However, most earlier studies only considered enhancing the network robustness and ignored the analysis of the community structure, which may alter the original topological structure and functionality of networks. In this paper, we propose a new memetic algorithm (MA-CR) with a two-level learning strategy to enhance the community robustness of networks, while maintaining the degree distribution and community structure. The proposed MA-CR is a hybrid global-local heuristic search methodology which adopts genetic algorithm as the global search and the proposed two-level learning strategy as the local search. The two-level learning strategy is designed based on the potential characteristics of the node structure and community structure of networks, which aims at mitigating two-level targeted attacks. Experiments on synthetic scale-free networks as well as real-world networks demonstrate the effectiveness and stability of the proposed algorithm as compared with several state-of-the-art algorithms.

1. Introduction

Complex networks are widely adopted as models to explore the properties of complex systems in terms of structure, stability and functionality [2,7,11,19,20,27,35,36]. Among the properties of complex networks, community structure is identified as an intrinsic and important one for reflecting the functionality of complex systems [12]. The community structure represents the functional modules in networks. In the domain of complex networks, communities are described as the subsets of networks. Generally, the connections within communities of a network are denser than those among communities [29,33], and nodes belonging to the same community probably have same or similar properties [39]. For instance, considering a protein-protein interaction network [12], communities are those functional modules of proteins with same or similar functions. In aviation networks, communities correspond to the sets of airlines that have more frequent traffic activities.

Robustness is another important property for reflecting the capability in protecting the security of complex networks. When networks suffer from random failures or malicious attacks, the breakdown of a critical fraction of nodes or edges can lead to the collapse of networks [35]. In recent years, many methods focusing on how to enhance the robustness of networks have been carried out [14,15,35]. Generally, the major methods can be divided into three categories [40]. Initially, the first

* Corresponding author.
E-mail address: gong@ieee.org (M. Gong).

http://dx.doi.org/10.1016/j.ins.2017.09.021

category simply adds edges to existing networks [5,16]. Those methods can enhance network robustness, but adding extra edges to a network may change its degree distribution and community structure. Without adding new edges, the second category makes all nodes have similar degrees by reconnecting edges [17]. This solution may obtain significant enhancement of network robustness, but nearly change all the connections of the network. Moreover, it also influences the degree distribution and the community structure of the network. The third category tends to overcome the obstacles of earlier methods by swapping two randomly chosen edges [24,35], and the network robustness can get significantly enhancement while retaining the degree distribution of the network. Nevertheless, it may still break the community structure of the network. In summary, all these three categories of methods for robustness enhancement fail to preserve the community structure of networks. As we know, altering the community structure of networks may damage its topological structure and functionality, thus we should maintain the original community structure as much as possible.

Traditional measures for the robustness of networks overlook situations in which the networks suffer from a destructive damage without fully collapsing. To solve this problem, Schneider et al. [35] put forward a robustness measure R_n, which considers the size of the largest connected part when nodes are removed gradually, namely $R_n = \frac{1}{N}\sum_{q=1}^{N} s(q)$, where $s(q)$ represents the integrity of nodes in the largest connected component after q nodes are removed. $1/N$ is the normalization factor for comparing the robustness measure of networks with different scales. Generally, a larger R_n indicates a more robust network. After that, Zeng and Liu proposed a new robustness measure, link robustness R_l, to evaluate network robustness under link attacks [41].

Note that the aforementioned measures are both simple and effective, but none of them considers the integrity of community structure. In our previous paper [21], we modeled the malicious attack on networks as a two-level targeted one, with the small-scale node attack and large-scale community attack. Then, we presented a community robustness R_c to measure the community integrity of networks under the two-level targeted attacks. Moreover, in order to preserve the original community structure, we proposed a new constraint that keeps the intracommunity links' amount in every community unchanged for link changes.

Evolutionary algorithms (EAs), due to their inherent global searching abilities, have been found to be very effective for solving many hard optimization problems [13,23,34]. Memetic algorithm (MA) is a very important and popular branch in the domain of EA. The inspiration of MAs derives from the natural systems that introduce the individual learning into the population evolution. The designation of "memetic" derived from Dawkin's concept of a meme, which can perform local refinements [10]. Therefore, MAs efficaciously synthesize the capabilities of the global search and local search, and have been demonstrated to be more efficient than traditional EAs in searching solutions for many optimization problems [8,18,28,30–32,37,43].

In this study, we propose a new memetic algorithm to enhance the community robustness of networks under the two constraints that remain the degree distribution and the intracommunity links' amount of each community unchanged for link changes. The proposed algorithm, termed MA-CR, adopts genetic algorithm (GA) [26] as the global search and the proposed two-level learning strategy as the local search. In standard GAs, the evolution process begins with a population of random individuals. In each iteration, several individuals are firstly selected from the current population according to their fitness values, and then the recombination and mutation operations are performed to construct a new offspring population for the next iteration. The proposed two-level learning strategy is devised based on the potential characteristics of the node structure and community structure of networks, which aims at mitigating the small-scale node attacks and large-scale community attacks, respectively. Experiments on both synthetic scale-free networks and two real-world networks reveal that compared with several state-of-the-art algorithms, the proposed MA-CR can search for much more robust networks under the two constraints and has better stability.

The remainder of the paper is organized as follows. Section 2 introduces the two-level targeted attacks and the community robustness index. Section 3 gives a detailed description of the proposed algorithm MA-CR. Section 4 conducts the experiments on the synthetic scale-free networks with different scales and two diverse real-world networks to demonstrate the effectiveness of MA-CR. Conclusions are summarized in the final section.

2. Community robustness of networks against two-level targeted attacks

Generally, a network can be modeled as a graph $G = (V, E)$, where $V = \{1, 2, \ldots, N\}$ is the set of N nodes, and $E = \{e_{ij} = (i, j) \mid i, j \in V \text{ and } i \neq j\}$ denotes the set of M edges. In this paper, we mainly focus on undirected and unweighted networks. In another perspective, we consider that a network can also be represented as $C = (S, E')$, where $S = \{s_1, s_2, \ldots, s_k\}$ denotes the set of communities of networks. $E' = \{c_{ij} = (s_i, s_j) \mid s_i, s_j \in S \text{ and } s_i \neq s_j\}$ is the set of connections between communities, and the entry of c_{ij} corresponds to the number of intercommunity links between community s_i and s_j.

Before giving a detailed description of our algorithm, we shall summarize our previous work [21] to briefly introduce the conceptions of the two-level targeted attacks and community robustness index.

2.1. Multi-level targeted attacks model

In our previous paper [21], besides the damages from nodes or edges of networks, we considered a possible scenario where attacks could also occur on the communities of networks. Moreover, the damages induced by attacking communities are obviously much greater than those induced by attacking certain nodes or edges. Then, we modeled the malicious attack