Reliability Engineering and System Safety 174 (2018) 41–52

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Measuring the vulnerability of community structure in complex networks

Daijun Wei[a,b,1], Xiaoge Zhang[b,1], Sankaran Mahadevan[b,*]

[a] School of Science, Hubei University for Nationalities, Enshi 445000, China

[b] Department of Civil and Environmental Engineering, School of Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Keywords:
Complex networks
Community
Vulnerability
Connectivity

ABSTRACT

This paper develops a quantitative method to measure the vulnerability of community structure with emphasis on both internal and external connectivity characteristics of the community. In particular, the number of links between communities and the strength of links connecting two communities are considered as external factors, while the connection density, the degree of gateway nodes, as well as the strength of links within each community are treated as internal factors. A non-linear weighted function is used to combine the internal factors with external factors. Then the developed method is used to illustrate the vulnerability analysis of community structure of a power transmission grid, a karate club network, and an air transportation network. The results reveal that the proposed measure is effective in differentiating the vulnerability level of community structure in a variety of networks.

1. Introduction

Community structure is a widely acknowledged characteristic in a large body of research on complex networks, as highlighted in recent studies [1–3]. Typically, a community is a subnetwork consisting of a group of nodes with a higher connection density of edges within the group than between the groups [2]. In general, there are two fundamental issues related to the community study in complex networks. The first one is how to discover community structure in networks, and the other one is what properties these communities have in common. Numerous approaches have been developed to detect the community structure in a variety of networks. For example, since it is common for two communities to share some nodes in complex networks, Orman et al. developed a novel method to detect overlapping communities [4]. Rocco et al. analyzed the effects of multi-state links on community detection [5]. To account for the dynamic behavior of each node in evolving networks, Orman et al. developed an innovative community detection method, in which the evolution of topology, nodal attributes, and community structure over time were considered [6]. In [7], they also characterized the role of each node by studying the evolution of its neighborhood based on the assumption that the neighborhood changes reveal the importance of the node in the entire network. The aforementioned approaches have been studied for several real-world networks, including scientific collaboration networks, and a network of Jazz listeners extracted from LastFM [6,7]. Meanwhile, several algorithms have been developed for discovering community structure in networks [8,9]. For example, a hierarchical agglomerative algorithm (HAA) for detecting community structure in unweighted network was proposed by Newman et al. [2,10] and the algorithm was further extended for weighted networks [11,12], which will be introduced in Sections 2.3 and 2.4, respectively.

With respect to the second issue, the vulnerability evaluation on community structure in complex networks has received increasing attention. For example, Zio et al. identified three elements related to the vulnerability of infrastructure systems, namely the degree of loss caused by a hazard, the degree of exposure to the hazards, and the degree of ability to recover to a stable state, and they demonstrated the analysis procedures for Critical Infrastructures (CIs) [13]. Holme et al. defined vulnerability as the reduction of network functionality due to selected removal of certain vertices or links, and investigated which attack strategy (based on measures such as degrees of nodes and betweenness centrality) is most effective [14]. Torres-Vera et al. measured the vulnerability of a pipeline system as the amount of the damage suffered by a structure due to a seismic activity [15].

Recently, a qualitative metric to measure the vulnerability of community structure was developed in Ref. [16], in which the vulnerability is quantified in terms of the connectivity between any two communities. The vulnerability between communities x and y, denoted as v_{xy}, is a metric to provide a measure of the degree to which two communities are disconnected. This is a qualitative metric; however, it offers some insights regarding the relative strength of each community. The developed community structure vulnerability measure has been illus-

* Corresponding author.
E-mail address: sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

[1] D. Wei and X. Zhang contributed equally to this work.

https://doi.org/10.1016/j.ress.2018.02.001
Received 10 May 2017; Received in revised form 29 January 2018; Accepted 1 February 2018
0951-8320/

trated in several networks, such as telephone network, power network, and terrorist network [16]. Considering that vulnerability denotes the ability of an individual or group to withstand and cope with the impact of natural or man-made hazards [17], connectivity is certainly an important consideration in measuring a community's vulnerability. Therefore, the proposed method in this paper also focuses on connectivity. A community that has more connections to other communities is more capable of coping with the impact of hazardous events, due to the opportunity to draw assistance and resources from neighboring communities. Similarly, within a community, the more the number of connections, the stronger is its ability to cope with the hazard, due to the increased ability to get in contact with other community members for receiving support. One study worthy of mention is Carrington et al., which considered the inter and intra links among communities to measure group centrality scores [18]; they stated that the number of connections alone cannot fully characterize the vulnerability of community structure. The quality or strength of the connections, as well as the structure of the community, are also important considerations. Actually, the vulnerability metric developed in Ref. [16] is a special case of the group centrality score developed in Ref. [18].

To illustrate the above points, consider a ground invasion scenario where the objective of the invading force is to defeat a community. The community is less vulnerable to the attack if it is well connected and able to draw support from other neighboring communities. The more the number of connections, the less vulnerable the community – this is the implication of the metric in Ref. [16]. Therefore, the attacking force might try to isolate the targeted community from its neighbors by severing or blocking its transportation and communication links to the neighbors. However, different links might have different strengths and thus might require different amounts of resources in order to be destroyed or blocked. Thus, in addition to the number of connections, the strength of the connections is also a significant consideration in affecting a community's vulnerability.

When a community is being attacked, its structure and topology will also determine its ability to withstand the attack. If the community has higher connection density and connection strengths, when a specific zone is attacked, it can receive more help from other zones, which in turn enhances its ability to withstand the attack. Another factor is the connectivity of the gateway nodes (i.e., nodes that connect the community to other communities) to other nodes within the community. Such nodes play an important role in distributing food, water, fuel, energy, and information received from other communities. The adequate operation of resource distribution within the community is vital in strengthening its ability to cope with the attack.

Considering the above factors, the number of external connections to neighboring communities alone cannot differentiate the vulnerability of different communities. For example, the telephone network in Belgium is divided into 7 communities in Ref. [16]. Five of the seven communities are different from each other in topology and structure, but all of them have the same number of connections to their neighboring communities. Based on the number of external connections alone, all five communities would be judged to have the same level of vulnerability. Such a measure has a low resolution in differentiating community vulnerability, which limits its usefulness in practical applications.

The goal of this paper is to develop a generalized measure that has the ability to distinguish the degree of vulnerability among separate communities. The proposed measure fuses the information regarding connectivity characteristics of the community, both internal and external. Our major idea is described as follows: vulnerability of community structure is not only related to the outer connectivity of a community, but also its inner structure. Two communities may have different vulnerability even if they have the same number of edges connecting with other communities. Besides the number of external connections, the degree of the gateway nodes, the connection density within the community, the strengths of edges connecting with other communities, and the strengths of edges within the community are also important considerations when measuring the vulnerability of community structure. Thus, in this paper, we consider 5 factors - 2 external factors (number of links between communities, and strengths of the links connecting two communities), and 3 internal factors (connection density, degrees of gateway nodes, and strengths of the links within each community). All five factors are combined into one composite qualitative metric in this paper. Five parameters are used to indicate the weights related to the factors we considered. As a result, the qualitative vulnerability measure in Ref. [16] becomes a special case of the vulnerability measure proposed in this paper. The proposed method is illustrated in three networks, namely a power transmission grid [19], a karate club network [20], and an air transport network [21]. The results show the effectiveness of the proposed method in differentiating the vulnerability values of separate individual communities.

The rest of the paper is organized as follows: the basic concepts are introduced in Section 2. In Section 3, we propose a generalized metric to measure the vulnerability of community structure. In Section 4, we illustrate the proposed vulnerability metric with three different networks. In Section 5, we provide concluding remarks.

2. Preliminaries

In this section, basic concepts such as the degree of a node, connection density, and strength of edges in complex networks, are introduced. In addition, HAA [2,10], a classical method for detecting community structure in unweighted networks and its extension to detect community in weighted networks, are described.

2.1. Degree and connection density in unweighted networks

Consider a complex network $G(E, N)$, where $E = (1, 2, \cdots, m)$ is the set of edges and $N = (1, 2, \cdots, n)$ is the set of nodes, where m and n are the numbers of edges and nodes in network G, respectively.

Definition 2.1. The degree of a node i in an unweighted network, denoted as k_i, is defined as [22]:

$$k_i = \sum_{j=1}^{N} x_{ij} \tag{1}$$

where x_{ij} represents the connection between node i and node j. $x_{ij} = 1$ if node i is connected to node j, and $x_{ij} = 0$ otherwise.

It can be observed that the degree of a node is the number of links connecting itself with other nodes. The average degree of all the nodes is $\overline{d} = \frac{1}{n}\sum_{i=1}^{n} k_i$, where n is the number of nodes in network G.

The network density describes the fraction of the potential connections in a network that are actual connections [23]. By contrast, a potential connection is a connection that could potentially exist between two nodes (regardless of whether or not it) [24].

Definition 2.2. The network density, denoted as ρ, is defined as,

$$\rho = \frac{m}{C_n^2} \tag{2}$$

where m and n are the number of edges and nodes in the network, respectively. C_n^2 is the number of links if any two nodes are connected in the network.

From this definition, we have $0 \le \rho \le 1$. $\rho = 0$ means there is no connection in the network, while $\rho = 1$ reveals that the network is fully connected. Network density is a significant indicator on the connectivity of a network.

2.2. Strengths of edges in weighted networks

Generally, a weighted network can be modeled as: $G(E, N, W)$, where $W \in R^+$ is the weight of an edge. If $W \equiv 1$ for all the links, then the weighted network degenerates to an unweighted network. As mentioned