Physica A 453 (2016) 290–297

Contents lists available at ScienceDirect

Physica A

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

Using mapping entropy to identify node centrality in complex networks

Tingyuan Nie [a,*], Zheng Guo [a], Kun Zhao [a], Zhe-Ming Lu [b]

[a] Institute of Communication and Electronic Engineering, Qingdao University of Technology, No. 11, Fushun Road, Qingdao, 266033, Shandong, China

[b] School of Aeronautics and Astronautics, Zhejiang University, No. 38, Zheda Road Hangzhou, 310027, Zhejiang, China

HIGHLIGHTS

- Mapping entropy is proposed based on the knowledge of a node and its neighbors.
- Mapping entropy centrality is more efficient than the traditional centralities.
- Mapping entropy centrality identifies the node importance well in complex network.
- Dynamic attack using mapping entropy centrality is more efficient than static attack.

ARTICLE INFO

Article history:
Received 12 October 2015
Received in revised form 1 February 2016
Available online 11 February 2016

Keywords:
Mapping entropy
Centrality
Complex network
Invulnerability

ABSTRACT

The problem of finding the best strategy to attack a network or immunize a population with a minimal number of nodes has attracted much current research interest. The assessment of node importance has been a fundamental issue in the research of complex networks. In this paper, we propose a new concept called mapping entropy (ME) to identify the importance of a node in the complex network. The concept is established according to the local information which considers the correlation among all neighbors of a node. We evaluate the efficiency of the centrality by static and dynamic attacks on standard network models and real-world networks. The simulation result shows that the new centrality is more efficient than traditional attack strategies, whether it is static or dynamic.

1. Introduction

The study of complex systems is an active area of scientific research inspired largely by the empirical study of real-world networks. The expanding research fields include mathematics, physics, biology, telecommunications, computer science, sociology, epidemiology, and others [1]. In a complex network, the problem of finding the best strategy to attack a network or immunize a population with a minimal number of nodes has attracted a lot of current research interest. There are some important nodes in the network which play a key role. These have influential effects on network dynamic processes, such as network synchronization [2–4], disease propagation [5,6], traffic navigation [7], and cascading failures [8–10]. Hence identifying the importance of a node becomes an efficient way to understand the relationship between the structure and the functionality of a network. Based on this research, the reliability, the robustness, and the performance of a network can be improved dramatically.

* Corresponding author.
E-mail address: tynie@qtech.edu.cn (T. Nie).

http://dx.doi.org/10.1016/j.physa.2016.02.009

In the literature, many researchers have put forward various criteria to identify node importance. The most popular algorithms in data mining are PageRank [11] and Hypertext Induced Topic Search (HITS) [12], which determine the importance of a website mainly according to the number of visits to a node in the diffusion process.

At the same time, many centrality measures are proposed to evaluate node importance in complex networks. The proposed centrality measures consist of degree centrality (DC) [13], betweenness centrality (BC) [14], closeness centrality (CC) [15], eigenvector centrality (EC) [16], second order centrality [17], Shannon-Parry measure [18], and so on. Generally, centrality measures characterize different aspects of a specific problem, so they reflect node importance in different directions. A centrality is optimal for one application, yet is often sub-optimal for a different application.

In information theory, entropy appears as a basic concept. It is well known that Shannon entropy and Von Neumann entropy are related to the information present in classical and quantum systems. In complex network research, a number of different entropy measures have been introduced [19–24]. Traditionally, entropy is used to analyze the statistical behavior or the structural features of a given real network.

Early research by Callaway et al. and Cohen et al. found that the scale-free network is resilient to random attack but sensitive to intentional attack [25,26]. Gallos et al. proved that an intentional attack, even with little knowledge on high degree nodes, can reduce the threshold drastically compared with the random case. This implies that an appropriate strategy may lead to efficient attack effects [27]. Recently, Chen et al. presented a graph-partitioning immunization strategy to substantially improve the efficiency of intentional attack [28]. Kitsak et al. proposed an attack strategy via identifying the most efficient spreaders in a network, and the simulation result showed that this is positive and more efficient [29].

In this paper, we propose a novel definition called mapping entropy (ME) to evaluate the importance of a node in the network. The value of ME reflects the correlation between a node and its neighbor nodes. We evaluate the efficiency of the centrality by static attacks and dynamic attacks on three standard network models and three real-world networks. The simulation result shows that ME is more efficient than traditional attack strategies on most of networks, and it benefits from identifying the node importance in the network.

2. Mapping entropy

Given a network $G(N, L)$ with N nodes and L links, the information entropy of the network is defined as follows:

$$E_i = -\sum_i I_i \log I_i = -\sum_i DC_i \log DC_i \tag{1}$$

where I_i is the importance of a node and is usually replaced by degree centrality (DC). A node and its neighbors construct a sub-network. Here we define the local entropy (LE) of the sub-network originated from node v_i, as shown in formula (2).

$$LE_i = -\sum_{j=1}^{M} DC_j \log DC_j \tag{2}$$

where DC_j is the degree centrality of node v_j, which belongs to the neighbor set M of node v_i.

Consider the mapping relation between a node and its neighbors, we define the mapping entropy (ME) by interleaving the degree centralities of node v_i and v_j.

$$ME_i = -DC_i \sum_{j=1}^{M} \log DC_j \tag{3}$$

where DC_i is the degree centrality of node v_i, and DC_j is the degree centrality of one of its neighbor nodes. We think the definition considers both the degree of the node and the degrees of its neighbors. This may be useful to identify the importance of the node.

3. The networks

We evaluate the efficiency of the proposed centrality on both real-world networks and network models. For this, we use three network models. The ER random network and the WS small-world network are exponential networks, in which each vertex has approximately the same degree. Another network is the power-law BA scale-free network, whose functionality is often determined by a relatively small number of highly connected vertices. The first real-world power grid network is an undirected network representing the topology of the Western States Power Grid of the United States [30]. The second is a social network which reflects the co-authorship of research papers [31]. The third is the Protein network, which represents protein-protein interaction [31]. The used scale-free networks display different levels of assortative mixing: the Coauthor network behaves assortatively, the Protein network behaves disassortatively, and the BA network behaves neutrally. The average degrees of the networks are calculated as: (a) the ER random network: 5.96; (b) the BA scale-free network: 5.65; (c) the WS small-world network: 6; (d) the power-grid network: 2.67; (e) the Coauthor network: 3.45; and (f) the Protein network: 5.83. The rewiring probability for the WS small-world network is set to 0.01. We show the degree distribution of the networks in Fig. 1.