Physica A 399 (2014) 57–69

Contents lists available at ScienceDirect

Physica A

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

A new method of identifying influential nodes in complex networks based on TOPSIS

Yuxian Du [a], Cai Gao [a], Yong Hu [b], Sankaran Mahadevan [c], Yong Deng [a,c,*]

[a] School of Computer and Information Science, Southwest University, Chongqing 400715, China
[b] Institute of Business Intelligence and Knowledge Discovery, Guangdong University of Foreign Studies, Guangzhou 510006, China
[c] School of Engineering, Vanderbilt University, TN 37235, USA

HIGHLIGHTS

- A new method of identifying influential nodes in complex networks based on TOPSIS is proposed.
- Different centrality measures are applied to different types of networks.
- It is the first time for TOPSIS to be applied to identify influential nodes in complex networks.
- Experimental results indicate that the proposed method is effective to identify influential nodes.

ARTICLE INFO

Article history:
Received 4 August 2013
Received in revised form 22 October 2013
Available online 2 January 2014

Keywords:
Complex networks
Influential nodes
TOPSIS
MADM
Centrality measure
SI model

ABSTRACT

In complex networks, identifying influential nodes is the very important part of reliability analysis, which has been a key issue in analyzing the structural organization of a network. In this paper, a new evaluation method of node importance in complex networks based on technique for order performance by similarity to ideal solution (TOPSIS) approach is proposed. TOPSIS as a multiple attribute decision making (MADM) technique has been an important branch of decision making since then. In addition, TOPSIS is first applied to identify influential nodes in a complex network in this open issue. In different types of networks in which the information goes by different ways, we consider several different centrality measures as the multi-attribute of complex network in TOPSIS application. TOPSIS is utilized to aggregate the multi-attribute to obtain the evaluation of node importance of each node. It is not limited to only one centrality measure, but considers different centrality measures, because every centrality measure has its own disadvantage and limitation. Then, we use the Susceptible–Infected (SI) model to evaluate the performance. Numerical examples are given to show the efficiency and practicability of the proposed method.

1. Introduction

Complex networks, which have been shown to demonstrate universal features such as small-world [1] and scale-free [2] effects, have been paid much attention in many fields, such as social sciences, computer sciences, biological sciences, and management [3–7]. Node importance is a basic measure in characterizing the structure and dynamics of complex networks [8–13]. Hence, identifying influential nodes has been an open issue and a critical research task in complex networks. Various centrality measures have been proposed over the years to rank the nodes of a graph according to their topological importance [14].

* Corresponding author at: School of Computer and Information Science, Southwest University, Chongqing 400715, China. Tel.: +86 23 68254555; fax: +86 23 68254555.
E-mail addresses: ydeng@swu.edu.cn, professordeng@163.com (Y. Deng).

0378-4371/$ – see front matter
http://dx.doi.org/10.1016/j.physa.2013.12.031

Table 1
The degree, closeness, betweenness value of each node in Kite network.

Degree		Closeness		Betweenness	
Nodes	Value	Nodes	Value	Nodes	Value
Diane	6	Fernando	0.071	Heather	0.622
Fernando	5	Garth	0.071	Fernando	0.370
Garth	5	Diane	0.067	Garth	0.370
Andre	4	Heather	0.067	Ike	0.356
Beverly	4	Andre	0.059	Diane	0.163
Carol	3	Beverly	0.059	Andre	0.037
Ed	3	Carol	0.056	Beverly	0.037
Heather	3	Ed	0.056	Carol	0.000
Ike	2	Ike	0.048	Ed	0.000
Jane	1	Jane	0.034	Jane	0.000

1.1. Previous work

How to identify influential spreaders effectively and efficiently is a big challenge up to now. A large number of centrality indices have been proposed to address this problem [15,16], such as degree centrality, closeness centrality, betweenness centrality [17], and eigenvector centrality [18].

These centrality measures have already been extended to be widely applied in complex networks. Gao et al. [19] proposed a bio-inspired centrality measure model, which combines the Physarum centrality with the K-shell index obtained by K-shell decomposition analysis, to identify influential nodes. Bonacich and Lloyd [20] proposed an extension of the eigenvector centrality measure to deal with the problem when the graphs are asymmetric, the eigenvectors of asymmetric matrices are not orthogonal. Restrepo et al. [21] presented a quantitative, objective characterization of the dynamical importance of network nodes and links in terms of their effect on the largest eigenvalue. Gómez et al. [22] extended the betweenness centrality measure to take into account explicitly several dimensions. A new closeness centrality measure is defined to deal not only with the maximum flow between every ordered pair of nodes, but also with the cost associated with communications. There have been many researches to improve the centrality measures to identify influential nodes in complex networks. However, all these researches focused on only one centrality measure and every centrality measure has its own disadvantage and limitation.

1.2. Problem description

Although these centrality measures have been widely applied in complex networks, there are some limitations and disadvantages: the degree centrality method is very simple such that only the local structure around a node must be known for it to be calculated, since the measure does not take into consideration the global structure of the network. A main limitation of closeness is the lack of applicability to networks with disconnected components: two nodes that belong to different components do not have a finite distance between them. Although the betweenness centrality measure takes the global network structure into consideration and can be applied to networks with disconnected components, it is not without limitations. For instance, a great proportion of nodes in a network generally does not lie on the shortest path between any two other nodes, and therefore receives the same score of 0 [23]. If there are two or more non-identical components the eigenvector of the adjacency matrix will describe only one of the components. In such circumstances the correct solution is to analyze each component separately. That is to say eigenvector centrality cannot be applied to networks that contain multiple components [24].

If only one centrality measure is adopted, then the rankings of identifying influential nodes may be different by using a different centrality measure. For example, Kite network with 10 nodes is designed by Krackhardt [25]. In Kite network as shown in Fig. 1, the node with the largest degree is *Diane*. Although the node *Heather*'s degree is 3, it has the largest betweenness centrality value. *Fernando* and *Garth* have the largest closeness centrality value. In some cases, using different centrality measures may provide different results, even conflicting results (see Table 1).

In this paper, we try to introduce Multiple Attribute Decision Making (MADM) (or called Multi-Criteria Decision-Making, MCDM) problem to explore how to identify influential nodes. MADM methods have been designed to designate a preferred alternative, classify alternatives in a small number of categories, and/or rank alternatives in a subjective preference order [26,27]. It has been extended to many fields [28,29,27,30–32]. Among numerous MADM methods developed to solve real-world decision problems, technique for order preference by similarity to ideal solution (TOPSIS) continues to work satisfactorily across different application areas. Hwang and Yoon [33] originally proposed TOPSIS to help select the best alternative with a finite number of criteria. TOPSIS makes full use of attribute information, provides a cardinal ranking of alternatives, and does not require attribute preferences to be independent [34,35]. As a well-known classical MADM method, TOPSIS has received much interest from researchers and practitioners [36–40].

It is the first time for TOPSIS to be applied to identify influential nodes in complex networks. Firstly, we calculate the value of different centrality measures. Meanwhile, in different networks in which the information goes by different