Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

A modified weighted TOPSIS to identify influential nodes in complex networks

Q1 Jiantao Hu [a,1], Yuxian Du [a,1], Hongming Mo [b], Daijun Wei [c], Yong Deng [a,d,*]

[a] *School of Computer and Information Science, Southwest University, Chongqing 400715, China*
[b] *Department of the Tibetan Language, Sichuan University of Nationalities, Kangding Sichuan 626001, China*
[c] *School of Science, Hubei University for Nationalities, Enshi 445000, China*
[d] *School of Engineering, Vanderbilt University, TN 37235, USA*

HIGHLIGHTS

- A weighted TOPSIS method for ranking node's spreading ability is proposed.
- To improve the original TOPSIS, a dynamically weighted algorithm is proposed.
- Experimental results indicate that our method outperforms the classical method.

ARTICLE INFO

Article history:
Received 27 March 2015
Received in revised form 29 June 2015
Available online xxxx

Keywords:
Complex networks
Influential nodes
Weighted TOPSIS
Centrality measures
SIR model

ABSTRACT

Identifying influential nodes in complex networks is still an open issue. Although various centrality measures have been proposed to address this problem, such as degree, betweenness, and closeness centralities, they all have some limitations. Recently, technique for order performance by similarity to ideal solution (TOPSIS), as a tradeoff between the existing metrics, has been proposed to rank nodes effectively and efficiently. It regards the centrality measures as the multi-attribute of the complex network and connects the multi-attribute to synthesize the evaluation of node importance of each node. However, each attribute plays an equally important part in this method, which is not reasonable. In this paper, we improve the method to ranking the node's spreading ability. A new method, named as weighted technique for order performance by similarity to ideal solution (weighted TOPSIS) is proposed. In our method, we not only consider different centrality measures as the multi-attribute to the network, but also propose a new algorithm to calculate the weight of each attribute. To evaluate the performance of our method, we use the *Susceptible–Infected–Recovered* (SIR) model to do the simulation on four real networks. The experiments on four real networks show that the proposed method can rank the spreading ability of nodes more accurately than the original method.

1. Introduction

How does the diversity of diseases such as AIDS, Ebola and H1N1 spread among human and animals? What is the cause of the traffic jam in the city? Why the impact of united states financial crisis could send shockwaves around the world?

* 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, prof.deng@hotmail.com (Y. Deng).
1 These authors contributed equally.

http://dx.doi.org/10.1016/j.physa.2015.09.028
0378-4371/

Although these problems seem to be different, each problem has a common factor, which is the complex network. Based on the researches of complex networks, we can get some quantitative descriptions and solutions to these problems. The study of influential nodes can help us to get a better understanding of the characteristics of complex networks. Thus, the problem of identifying influential spreaders in complex networks has gained increasing attention in recent years [1–6].

Over the recent years, various centrality indices have been proposed to address this issue [7–11]. Freeman [12] proposed the degree centrality measure to identify influential nodes in complex networks, which used the number of links to the focal node. It is a straightforward and efficient measure but it has failed to consider the global structure of the network. In order to compensate for this deficiency, closeness centrality [13] was proposed. It was defined as the inverse sum of the shortest distance to all other nodes. Kitsak et al. [4] proposed a coarse-grained method by using k-core decomposition, but this method does not work in some cases. For example, in a tree, all nodes are in 1-core and thus expected to have the same influence according to their method [10]. Lü et al. [14] proposed Leader Rank to identify leaders in social networks. It performs well in directed networks, but it cannot be applied in undirected networks. In a word, the development of a reasonable ranking method to identify influential nodes in complex networks is still an open issue [15,16,14]. Ranking influential nodes can be seen as a multi-attribute decision making (MADM) problem to some degree. As a result, to comprehensively combine different attribute data of each node is paid great attention. Due to the efficient modeling and fusion of uncertain information [17–24], evidence theory is widely used in identify influential nodes in complex networks [25]. Recently, with its ability to handle linguistic variables [26–32], fuzzy set theory also is applied to this field [33]. In short, using MADM technology to rank node influence is a promising way.

In this paper, on the basis of technique for order performance by similarity to ideal solution (TOPSIS), a new method is proposed to identify the influential node in complex networks, namely, weighted technique for order performance by similarity to ideal solution (weighted TOPSIS). In our method, we take into account different centrality measures as the multi-attribute to the network. Each centrality measure has more or less limitations, and the existing researches in this field all focused on only one centrality indices. So, we believe that this fusion method will lead to a better results. Meanwhile we propose a new algorithm to calculate the weight of each attribute in this fusion method. Because different attributes are obtained by difference centrality measures, their effects are different in the network absolutely. The proposed method partially overcomes the deficiency of the original method (TOPSIS) which only used a same weight for each attribute. To evaluate the performance of our method, *Susceptible–Infected–Recovered* (SIR) model [34] is used to examine the spreading influence of the nodes ranked by our method (weighted TOPSIS) and the original method (TOPSIS). The experimental results on four real networks show that our method can well identify the influential nodes. Furthermore, comparing our method (weighted TOPSIS) with the original method (TOPSIS), the experimental results indicate that our method performs better than the original TOPSIS.

The rest of the paper is organized as follows. In Section 2, we briefly review the basic theory of TOPSIS and the definitions of centrality measures. In Section 3, we introduce our proposed method with a simple example in detail. The effectiveness of our method and the original method are discussed in Section 4. Finally, we draw our conclusion in Section 5.

2. The basic theory of TOPSIS and centrality measures for node influence

2.1. The basic theory of TOPSIS

Optimization exists every where in our real life [35]. Among so many optimization methods, Technique for order performance by similarity (TOPSIS) was proposed by Hwang et al. [36], which is a simple but effective ranking method. There are two important ideal solution types in this method. One is the positive ideal solution, which minimizes the cost criteria and maximizes the benefit criteria simultaneously. The other is the negative ideal solution. On the contrary, the negative ideal solution maximizes the cost criteria and minimizes the benefit criteria simultaneously. The standard TOPSIS aims to select alternatives who have both the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution. In our method, we apply Euclidean distance to calculate the separation measures in TOPSIS applications. In mathematics, the Euclidean distance or Euclidean metric is the ordinary distance between two points that one would measure with a ruler, and is given by the Pythagorean formula. It is a very simple and effective measure for distance and it has widely used in many fields.

Du et al. [37] used degree centrality, closeness centrality and betweenness centrality as the multi-attribute in TOPSIS to generate the ranking lists to evaluate the node's spreading ability in complex networks. This is the first time for TOPSIS to be applied to identify influential nodes in complex networks. The followings is a detailed elaboration on TOPSIS algorithms.

Firstly, considering a decision matrix $D = (x_{mn})$ and normalizing the matrix $D = (x_{mn})$ [36]:

$$a_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i}^{m} {x^2}_{ij}}}, \quad i = 1, \ldots, m;\ j = 1, \ldots, n. \tag{1}$$

Secondly, multiply the columns of the normalized matrix by the associated weights to obtain a new decision matrix $B = (b_{mn})$ [36]:

$$b_{ij} = w_j \times a_{ij}, \quad i = 1 \ldots,\ m; j = 1 \ldots, n \tag{2}$$

where w_j is equal to $\frac{1}{n}$. The weight for j criterion is the same.

Please cite this article in press as: J. Hu, et al., A modified weighted TOPSIS to identify influential nodes in complex networks, Physica A (2015), http://dx.doi.org/10.1016/j.physa.2015.09.028