



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 xxx

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.

© 2015 Elsevier B.V. All rights reserved.

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.

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]:

$$a_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}, \quad i = 1, \dots, m; j = 1, \dots, n. \quad (1)$$

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

$$b_{ij} = w_j \times a_{ij}, \quad i = 1 \dots, m; j = 1 \dots, n \quad (2)$$

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

3. Proposed method

The standard method TOPSIS has been applied to identify influential nodes in complex networks by Du et al. [37] in 2014. It synthesizes variety of measures and can evaluate the influential node well. But in the original TOPSIS, for simplicity, it just set the weights to the attributes are equal to $\frac{1}{n}$ (n is the total number of the attributes). We think it is not reasonable, because different attributes are from different centrality measures. Each centrality measure has more or less limitations. According to the accuracy of each centrality measure for node importance evaluation, the weights should not be the same. Thus, how to give the weight to these attributes is an important issue to us. In order to overcome the deficiency of the original TOPSIS, we propose a new method to identify influential nodes in complex networks, namely weighted technique for order performance by similarity to ideal solution (weighted TOPSIS). In our method, we present a new algorithm to calculate the weight of each attribute. Moreover, this is a dynamically weighted algorithm, which means with the changes in the original metrics, the corresponding weights will change. Because the epidemic spreading process is different in different networks, one centrality measure should have different performance in different system. Since the weights should be different from one system to another system.

In order to achieve the dynamic weighting, we use the *Susceptible–Infected–Recovered* (SIR) model [34]. SIR model is widely used in the verification of node influence [6,2,39]. In SIR model, each node has three discrete states: (i) Susceptible $S(t)$ represents the number of individuals susceptible to (not yet infected) the disease; (ii) Infected $I(t)$ represents the number of individuals which have been infected and are able to spread the disease to susceptible individuals; (iii) Recovered $R(t)$ denotes the individuals who have been recovered from the disease and will not be infected. In each step, only one node is select to be infected and then let it spread the disease or information to random susceptible individuals with probability α . And then, the infected nodes recovered from the disease with probability β (in this paper, we set $\beta = 1$). The sum of infected and recovered nodes at time t , denoted by $F(t)$, can be considered as an indicator to evaluate the influence of the initially infected node at time t . Obviously, $F(t)$ increases with the increase of t , and it will remain stable at last. Hence, for different initially infected nodes, when reaching the steady state, higher $F(t)$ represents a larger influence. Thus we match the attributes to $F(t)$, the higher the matching degree, the bigger the weight. In this paper, to obtain the $F(t)$ value, we set the spreading probability $\alpha = 0.3$. Meanwhile, in order to ensure the state when we obtain the $F(t)$ value is the stable state in each network, we set $t = 100$ and $F(t)$ is obtained by over 100 independent runs.

In recent years, various centrality measures have been proposed to rank the spreading ability of the nodes, such as degree centrality, betweenness centrality, closeness centrality, semi-local centrality, leader rank and so on. In this paper, we choose best known centralities as the multi-attribute in our method, including degree, betweenness and closeness centralities. Degree centrality is the earliest and simplest metric. But it has failed to consider the global structure of the network. Closeness centrality was defined as the inverse sum of the shortest distance to all other nodes. But it also have some limitations. Betweenness centrality considers both the global structure of the networks and the disconnected components, but it ignores the nodes' local feature. Thus, we take the three centralities as the multi-attribute in our methods.

3.1. The algorithm of the proposed method

Fig. 1 shows the flow chart of our method. The detailed description of the proposed method follows:

Step 1, consider a matrix $D = (x_{mn})$, the first $n - 1$ columns represent the value of different centrality measures. The last column represents the results of SIR model ($F(t)$):

$$D(x_{mn}) = \begin{pmatrix} DC_1 & BC_1 & \cdots & F_1(t) \\ DC_2 & BC_2 & \cdots & F_2(t) \\ \vdots & \vdots & \cdots & \vdots \\ DC_n & BC_n & \cdots & F_n(t) \end{pmatrix}_{m \times n} \quad (12)$$

Step 2, construct a submatrix $M = (y_{mk})$ of $D = (x_{mn})$ ($k = n - 1$). The submatrix M is consisted of the first $n - 1$ columns in matrix D :

$$M(y_{mk}) = \begin{pmatrix} DC_1 & BC_1 & \cdots & CC_1 \\ DC_2 & BC_2 & \cdots & CC_2 \\ \vdots & \vdots & \cdots & \vdots \\ DC_m & BC_m & \cdots & CC_m \end{pmatrix}_{m \times k} \quad (13)$$

Step 3, normalize the matrix $D = (x_{mn})$:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n. \quad (14)$$

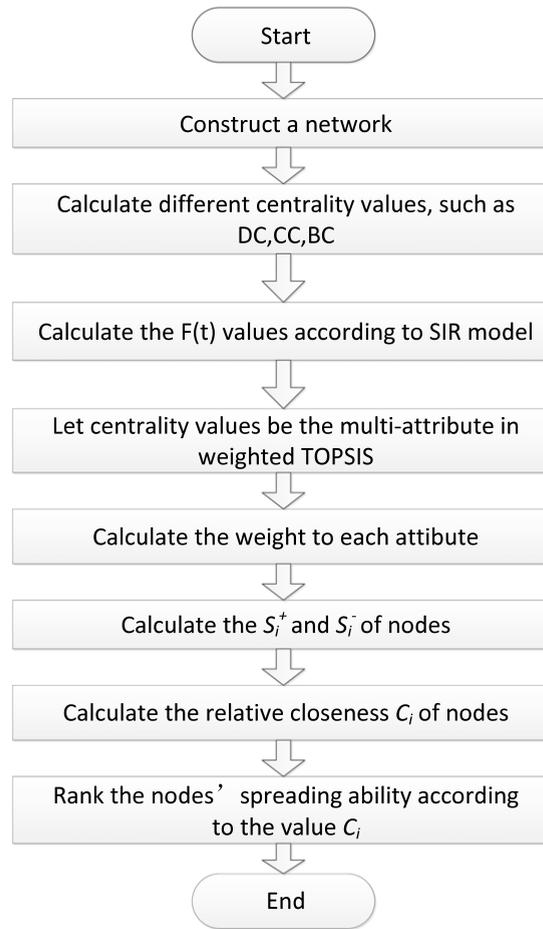


Fig. 1. The flow chart of the proposed method.

Step 4, match the attribute to $F(t)$ as follows:

$$v_{ij} = \frac{1}{|r_{ij} - r_{in}|}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (15)$$

where r_{in} represents $F(t)$.

Step 5, calculate the weight of each attribute:

$$e_j = \sum_{i=1}^m v_{ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (16)$$

$$w_j = \frac{e_j}{\sum_{j=1}^n e_j}, \quad j = 1, 2, \dots, n. \quad (17)$$

Step 6, normalize the matrix $M = (y_{mk})$ [36]:

$$h_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^m y_{ij}^2}}, \quad i = 1, \dots, m; j = 1, \dots, k. \quad (18)$$

Step 7, multiply the columns of the normalized matrix by the associated weights to obtain the weighted decision matrix $B = (b_{mk})$ [36]:

$$b_{ij} = w_j \times h_{ij}, \quad i = 1 \dots, m; j = 1 \dots, k \quad (19)$$

where w_j is the weight for j criterion.

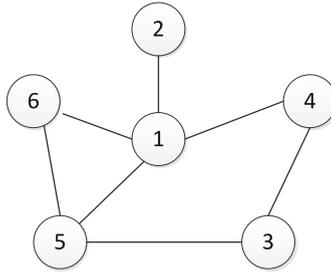


Fig. 2. An example network with 6 nodes and 6 edges.

Table 1
Degree centrality (DC), betweenness centrality (BC), closeness centrality (CC) and $F(t)$ values of each node. $F(t)$ is obtained by over 100 independent runs.

Node	DC	CC	BC	$F(t)$
1	4	0.1667	0.55	1.5
2	1	0.1	0	1.1
3	2	0.1111	0.05	1.4
4	2	0.125	0.1	1.1
5	3	0.1429	0.2	1.3
6	2	0.125	0	1.1

Step 8, confirm the positive ideal solution and the negative ideal solution. The positive ideal solution is denoted by A^+ , and the negative ideal solution is denoted by A^- . They are defined as follows [36]:

$$A^+ = \{b_1^+, b_2^+, \dots, b_k^+\} = \{(\max_i b_{ij} | j \in K_b) (\min_i b_{ij} | j \in K_c)\} \tag{20}$$

$$A^- = \{b_1^-, b_2^-, \dots, b_k^-\} = \{(\min_i b_{ij} | j \in K_b) (\max_i b_{ij} | j \in K_c)\} \tag{21}$$

where K_b is the set of benefit criteria and K_c is the set of cost criteria.

Step 9, obtain the separation measures of the existing alternatives from the positive ideal and the negative ideal solutions. The separation measures based on Euclidean distance, S_i^+ and S_i^- , are respectively derived from the follows [36]:

$$S_i^+ = \sqrt{\sum_{j=1}^n (b_j^+ - b_{ij})^2}, \quad i = 1, \dots, m; j = 1, \dots, k \tag{22}$$

$$S_i^- = \sqrt{\sum_{j=1}^n (b_j^- - b_{ij})^2}, \quad i = 1, \dots, m; j = 1, \dots, k. \tag{23}$$

Step 10, calculate the relative closeness to the ideal solution [36]:

$$C_i = \frac{S_i^-}{S_i^- + S_i^+}, \quad i = 1, \dots, m. \tag{24}$$

Higher C_i represents a larger influence. So the spreading ability of the node depends on the value C_i .

3.2. Example explanation

In order to get a better understanding of our method, here we give a simple example to explain how the weighted TOPSIS works. We take Fig. 2 as an example.

Step 1, calculate DC, CC and BC values of the network. Meanwhile, using SIR model to obtain the $F(t)$ values. These values are shown in Table 1. Then, consider the matrix $D = (x_{mn})$:

$$D = \begin{pmatrix} 4 & 0.1667 & 0.55 & 1.5 \\ 1 & 0.1000 & 0.00 & 1.1 \\ 2 & 0.1111 & 0.05 & 1.4 \\ 2 & 0.1250 & 0.10 & 1.1 \\ 3 & 0.1429 & 0.2 & 1.3 \\ 2 & 0.1250 & 0.00 & 1.1 \end{pmatrix}.$$

Step 2, construct a submatrix $M = (y_{mk})$ of $D = (x_{mn})$ ($k = n - 1$):

$$M = \begin{pmatrix} 4 & 0.1667 & 0.55 \\ 1 & 0.1000 & 0.00 \\ 2 & 0.1111 & 0.05 \\ 2 & 0.1250 & 0.10 \\ 3 & 0.1429 & 0.2 \\ 2 & 0.1250 & 0.00 \end{pmatrix}.$$

Step 3, normalize the matrix $D = (x_{mn})$ by Eq. (14). For example:

$$r_{11} = \frac{x_{11}}{\sum_{i=1}^6 x_{i1}} = \frac{4}{4 + 1 + 2 + 2 + 3 + 2} = 0.2875.$$

Then, we will obtain the normalized matrix $R = (r_{mn})$:

$$R = \begin{pmatrix} 0.2857 & 0.2163 & 0.6111 & 0.2000 \\ 0.0714 & 0.1298 & 0 & 0.1467 \\ 0.1429 & 0.1442 & 0.0556 & 0.1867 \\ 0.1429 & 0.1622 & 0.1111 & 0.1467 \\ 0.2143 & 0.1854 & 0.2222 & 0.1733 \\ 0.1429 & 0.1622 & 0 & 0.1467 \end{pmatrix}.$$

Step 4, match the attribute to $F(t)$ by Eq. (15). For instance:

$$v_{11} = \frac{1}{|r_{11} - r_{14}|} = \frac{1}{|0.2857 - 0.2|} = 11.6667.$$

Step 5, calculate the weight of each attribute by Eqs. (16) and (17). For example:

$$e_1 = \sum_{i=1}^6 v_{i1} = 597.2025$$

$$w_1 = \frac{e_1}{\sum_{j=1}^3 e_j} = 0.5824.$$

Then, we can get the weights.

$$W = (0.5824 \quad 0.0705 \quad 0.3471).$$

Step 6, normalize the matrix $M = (y_{mk})$ by Eq. (18). For instance:

$$h_{11} = \frac{y_{11}}{\sqrt{\sum_{i=1}^6 y_{i1}^2}} = \frac{1}{\sqrt{4^2 + 1^2 + 2^2 + 2^2 + 3^2 + 2^2}} = 0.6489.$$

Step 7, multiply the columns of the normalized matrix by the associated weights to obtain the weighted decision matrix $B = (b_{mk})$ by Eq. (19). For instance:

$$b_{11} = w_1 \times h_{11} = 0.5824 \times 0.6489 = 0.3779.$$

Through this step, the weighted decision matrix $B = (b_{mk})$ is obtained:

$$B = \begin{pmatrix} 0.3779 & 0.0368 & 0.3204 \\ 0.0945 & 0.0221 & 0 \\ 0.1890 & 0.0245 & 0.0291 \\ 0.1890 & 0.0276 & 0.0583 \\ 0.2834 & 0.0316 & 0.1165 \\ 0.1890 & 0.0276 & 0 \end{pmatrix}.$$

Step 8, according to Eqs. (20) and (21), we can get the positive ideal solution and the negative ideal solution:

$$A^+ = \{0.3779 \quad 0.0368 \quad 0.3204\}$$

$$A^- = \{0.0945 \quad 0.0221 \quad 0\}.$$

Table 2
The ranking list according to the value C_i .

Node	C_i	Rank
1	1	1
2	0	6
3	0.2216	4
4	0.2558	3
5	0.4971	2
6	0.2028	5

Table 3
The basic topological features of the four real networks. n and m are the total numbers of nodes and links, respectively. $\langle k \rangle$ and k_{\max} denote the average and the maximum degree. C is the clustering coefficient [40].

Network	n	m	$\langle k \rangle$	k_{\max}	C
Yeast	2 375	11 693	9.85	118	0.1529
Blog	3 982	6 803	3.42	189	0.1409
Router	5 022	12 516	2.4	106	0.2409
PGP	10 680	24 316	4.55	205	0.2659

Step 9, obtain the separation measures of the existing alternatives from the positive ideal and the negative ideal solutions by Eqs. (22) and (23). For instance:

$$S_1^+ = \sqrt{(0.3779 - 0.3779)^2 + (0.0368 - 0.0368)^2 + (0.3204 - 0.3204)^2} = 0$$

$$S_1^- = \sqrt{(0.0945 - 0.379)^2 + (0.0221 - 0.0368)^2 + (0 - 0.3204)^2} = 0.4280.$$

Step 10, by Eq. (24), the relative closeness to the ideal solution can be obtain. For example:

$$C_1 = \frac{S_1^-}{S_1^- + S_1^+} = \frac{0.4280}{0.4280 + 0} = 1.$$

Finally, we can get ranking list according to the value C_i . The result is shown in Table 2. By using the proposed method, we can see that Node 1 is the most important node in this network. Obviously, our proposed method can rank the nodes spreading ability correctly.

4. Experimental analysis

4.1. Data

In order to verify whether the proposed method (weighted TOPSIS) is better than the original method (TOPSIS), we use four real networks to evaluate the performance. (i) Yeast: this is a network of yeast and protein interactions. Each node represents a protein and they are connected by the side if there is interaction between proteins. The data can be downloaded from <http://snap.stanford.edu/data/>; (ii) Blog: the communication relationships between owners of Blog on the MSN (Windows Live) Spaces website [41]. The data can be downloaded from <http://www.cs.bris.ac.uk/steve/peacockpaper/>; (iii) Router: this is a routing network. It consists of 5022 nodes and 12516 edges [42]; (iv) PGP: this is a encrypted communication network. In order to protect privacy between peers, Pretty-Good- Privacy algorithms have been developed. It can be downloaded from <http://snap.stanford.edu/data/>. The basic topological properties of these four real networks are shown in Table 3.

4.2. Effectiveness

Tables 4 and 5 shows the top-10 lists generated by weighted TOPSIS and TOPSIS respectively. In principle, if the method is effective, its ranking list should be as consistent as possible with the ranking list which is generated by the real spreading process. To evaluate the performance, we use SIR model which is described in Section 3 to simulate the real spreading process. Firstly, we investigate the correlations between node's influence measured by the two methods and $F(t)$ via Kendall's tau coefficient τ [43]. Kendall's tau coefficient τ is widely used for correlation analysis [6,44,45]. The Kendall's tau coefficient considers a set of joint observations from two random variables X and Y . Any pair of observation (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree, that is, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. Meanwhile, they are said to be discordant if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ and $y_i = y_j$, the pair is neither concordant nor discordant. The Kendall's tau coefficient, τ , is defined as [46,47]:

$$\tau = \frac{n_c - n_d}{0.5n(n - 1)} \quad (25)$$

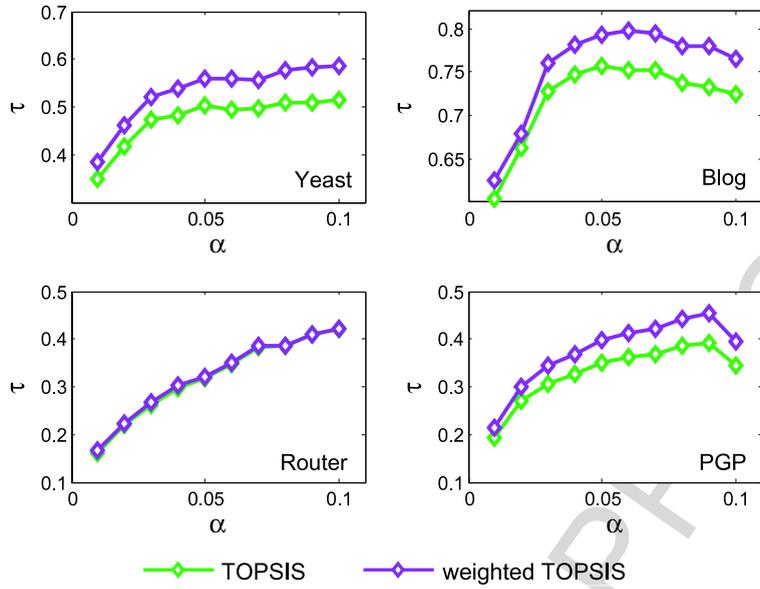


Fig. 3. The Kendall's tau τ obtained by comparing the ranking list generated by the two methods and the ranking list generated by SIR model on four real networks. The results are obtained by over 100 independent runs where the spreading probability α is ranging from 0.01 to 0.1.

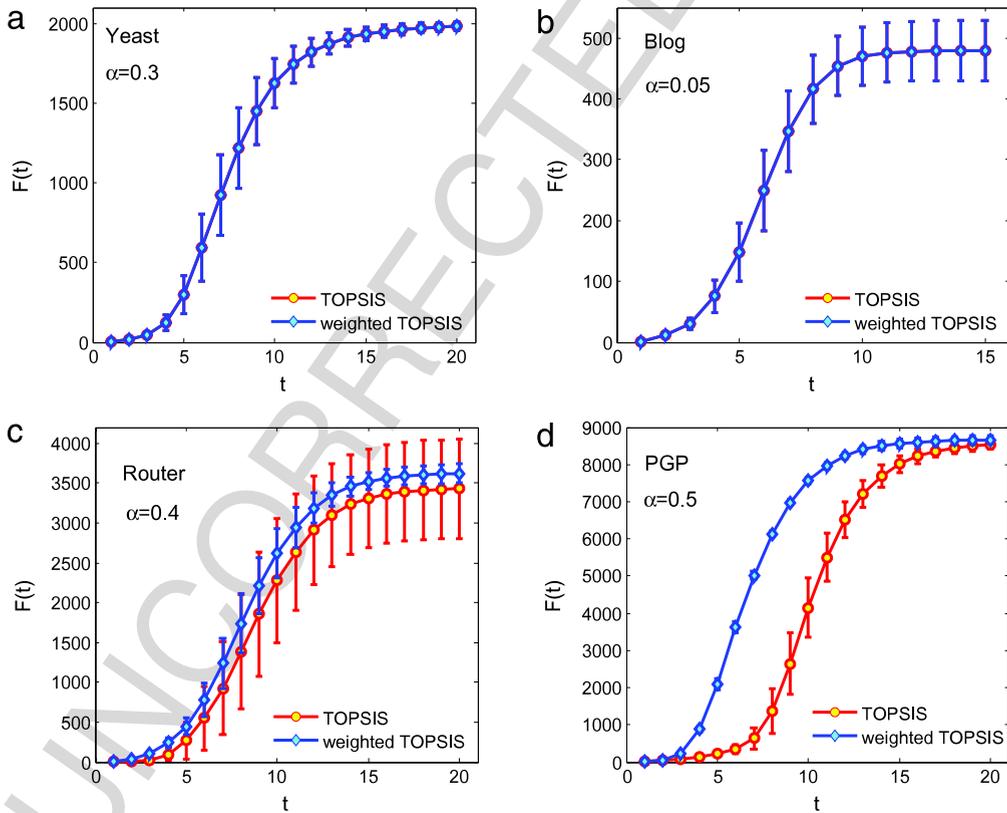


Fig. 4. The vertical axis represents the cumulative number of infected nodes. In each simulation, we only compare the nodes that either appear in the top-10 list the proposed method or the original TOPSIS measures. Results are obtained by averaging over 100 implementations.

1 outperforms the original TIPSIS on strongly positive correlation with real spreading process in Yeast, Blog and PGP. The
 2 Kendall's tau value τ is as high as 0.8 when $\alpha = 0.06$ in PGP. In Router, our method and the TIPSIS almost have the same
 3 performance. But our method is slightly better than TIPSIS before $\alpha = 0.05$.

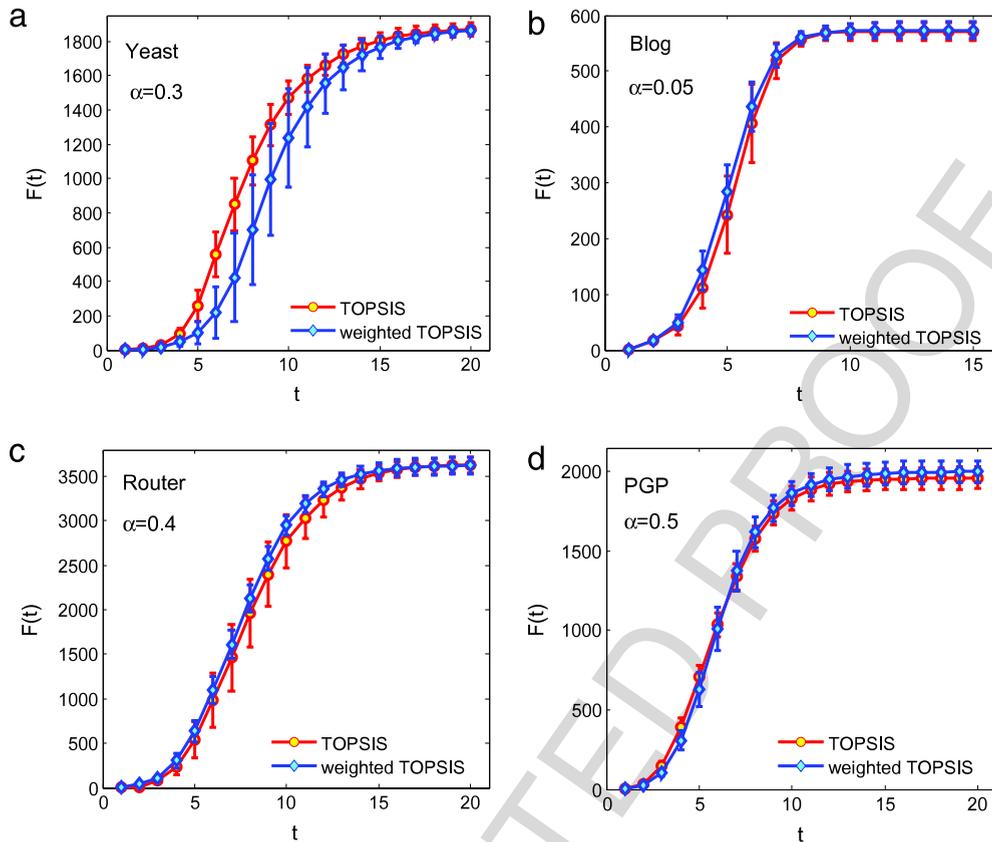


Fig. 5. The vertical axis represents the cumulative number of infected nodes. In each simulation, we compare these nodes with different ranking orders that both appear in the same top-10 lists. Results are obtained by averaging over 100 implementations.

Since the correlations between $F(t)$ and the two methods can only display whether the methods can identify the influential nodes correctly, it cannot evaluate the real spreading ability of the nodes. In order to make sure whether the proposed method can identify the influential nodes more accurately, we compare the influence of the nodes that either appear in the top-10 list by proposed method or the original TOPSIS. The two methods can be well distinguished since we do not consider the effects of common nodes in their ranking lists. The lists of top-10 influential nodes on four networks are given in Tables 4 and 5. In Yeast, comparing with the proposed method and TOPSIS, they have the all same nodes in the top-10 lists. The situation is same in Blog. In Router, the number of the same nodes in the top-10 lists between the proposed method and TOPSIS are eight. In PGP, the proposed method and TOPSIS only have one different node. The simulation results on the cumulative number of infected nodes, $F(t)$, are shown in Figs. 4 and 5.

In Yeast and Blog, our method and TOPSIS have same performance since they have all same top-10 nodes. The proposed method performs much better than TOPSIS in Router and PGP. In order to make the performance of our method more persuasive, we compare these nodes with different ranking orders that both appear in the same top-10 lists by using SIR model. In Yeast, the top-6 nodes have the same ranking orders: Node 321 > Node 375 > Node 174 > Node 2084 > Node 433 > Node 288, where “>” represents “better than” or “more influential than”. However, Node 1194 and Node 120 have different ranking orders. In the proposed method, Node 120 > Node 1194, but the situation is opposite in TOPSIS. Thus, in order to clarify the performance more persuasively, we compare the influence of Node 120 and Node 1194 as shown in Fig. 5(a). Although our method does not perform as good as Topsisid before $t = 17$, after $t = 17$, they reach the same steady state. In Blog, by using our method, Node 12 > Node 304, but the situation is also opposite in TOPSIS. The simulation result between the two nodes is shown in Fig. 5(b). It shows that Node 12 can spread the information (or disease) faster than Node 304, which means the proposed method outperforms TOPSIS. In Router, by using our method, Node 3670 > Node 1480, but in Topsisid, Node 3670 < Node 1480. So, we compare the influence of the two nodes. Fig. 5(c) shows the simulation result between Node 1480 and Node 3670 in Router. Our method still outperforms the original TOPSIS. The comparing between Node 6656 and Node 6556 in PGP, is shown in Fig. 5(d). we can see that our method also performs slightly better than TOPSIS.

Figs. 4 and 5 show that our method outperforms the original TOPSIS when α is equal to 0.3 in Yeast, 0.05 in Blog, 0.4 in Router and 0.5 in PGP. In order to make the results more convincing, we compare the simulation results when the spreading probability α is ranging from 0.1 to 0.5. In each simulation, we only compare the influence of the nodes that either appear in the top-10 list by proposed method or the original TOPSIS. In Yeast and Blog, the two methods have the same top-10 lists,

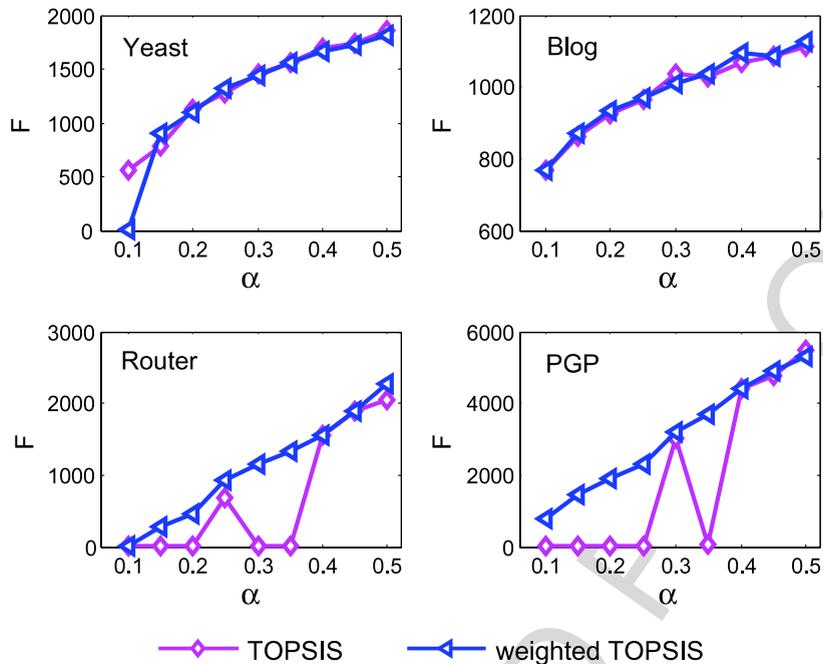


Fig. 6. F is the total number of infected nodes when the spreading process reaches the steady state in SIR model. It is obtain by over 100 independent runs where the spreading probability α is ranging from 0.1 to 0.5.

as same as the previous experiment, we compare these nodes with different ranking orders that both appear in the same top-10 lists. In SIR model, when reaching the steady state, we set $F(t) = F$. Hence higher F represents a larger influence. Simulation results on four networks are shown in Fig. 6. In Yeast, TOPSIS performs better than our method before $\alpha = 0.15$, but after $\alpha = 0.15$, our method almost have the same performance comparing with TOPSIS. In Blog, our method performs slightly better than TOPSIS. In Router and PGP, we can clearly see that our method performs much better than TOPSIS.

As discussed above, the proposed method performs better than the original TOPSIS. It can rank the spreading ability of nodes in complex networks more accurately.

5. Conclusion

In this paper, we proposed a new method to rank the node's spreading ability based on TOPSIS. We consider several different centrality measures as the multi-attribute to the network and give the corresponding weights to each attribute according to the matching degree with $F(t)$. To evaluate the performance, we used the SIR model to estimate the spreading ability of the top-ranked nodes by our method and the original TOPSIS. The experimental results on four real networks indicate that our method can identify influential nodes well. Compared with the original TOPSIS, the results show that our method has better performance.

Acknowledgments

The authors greatly appreciate the reviews' suggestions and the editor's encouragement. The work is partially supported by National High Technology Research and Development Program of China (863 Program) (Grant No. 2012AA041101), National Natural Science Foundation of China (Grant Nos. 61174022, 71271061, 61573290), China State Key Laboratory of Virtual Reality Technology and Systems, the General Research Program of Natural Science of Sichuan Provincial Department of Education (Grant No. 14ZB0322), the Self-financing Program of State Ethnic Affairs Commission of China (Grant No. 14SCZ014).

References

- [1] Q. Jin, L. Wang, C.-Y. Xia, Z. Wang, Spontaneous symmetry breaking in interdependent networked game, *Sci. Rep.* 4 (2014).
- [2] S. Gao, J. Ma, Z. Chen, G. Wang, C. Xing, Ranking the spreading ability of nodes in complex networks based on local structure, *Physica A* 403 (2014) 130–147.
- [3] E. Bozzo, M. Franceschet, Resistance distance, closeness, and betweenness, *Social Networks* 35 (2013) 460–469.
- [4] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, H.A. Makse, Identification of influential spreaders in complex networks, *Nat. Phys.* 6 (2010) 888–893.
- [5] W. Wang, M. Tang, H. Yang, Y. Do, Y.-C. Lai, G. Lee, Asymmetrically interacting spreading dynamics on complex layered networks, *Sci. Rep.* 4 (2014).

- [6] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, T. Zhou, Identifying influential nodes in complex networks, *Physica A* 391 (2012) 1777–1787.
- [7] W.-J. Yuan, J.-F. Zhou, Q. Li, D.-B. Chen, Z. Wang, Spontaneous scale-free structure in adaptive networks with synchronously dynamical linking, *Phys. Rev. E* 88 (2013) 022818.
- [8] F. Bauer, J.T. Lizier, Identifying influential spreaders and efficiently estimating infection numbers in epidemic models: A walk counting approach, *Europhys. Lett.* 99 (2012) 68007.
- [9] J. Liu, Y. Li, Z. Ruan, G. Fu, X. Chen, R. Sadiq, Y. Deng, A new method to construct co-author networks, *Physica A* 419 (2015) 29–39.
- [10] D.-B. Chen, H. Gao, L. Lü, T. Zhou, Identifying influential nodes in large-scale directed networks: the role of clustering, *PLoS One* 8 (2013) e77455.
- [11] X. Zhang, J. Zhu, Q. Wang, H. Zhao, Identifying influential nodes in complex networks with community structure, *Knowl.-Based Syst.* 42 (2013) 74–84.
- [12] L.C. Freeman, Centrality in social networks conceptual clarification, *Soc. Networks* 1 (1979) 215–239.
- [13] G. Sabidussi, The centrality index of a graph, *Psychometrika* 31 (1966) 581–603.
- [14] L. Lü, Y.-C. Zhang, C.H. Yeung, T. Zhou, Leaders in social networks, the delicious case, *PLoS One* 6 (2011) e21202.
- [15] B. Hou, Y. Yao, D. Liao, Identifying all-around nodes for spreading dynamics in complex networks, *Physica A* 391 (2012) 4012–4017.
- [16] J. Ugander, L. Backstrom, C. Marlow, J. Kleinberg, Structural diversity in social contagion, *Proc. Natl. Acad. Sci.* (2012) 201116502.
- [17] F. Cuzzolin, A geometric approach to the theory of evidence, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 38 (2008) 522–534.
- [18] L.V. Utkin, A new ranking procedure by incomplete pairwise comparisons using preference subsets, *Intell. Data Anal.* 13 (2009) 229–241.
- [19] V.-N. Huynh, T.T. Nguyen, C.A. Le, Adaptively entropy-based weighting classifiers in combination using dempster–shafer theory for word sense disambiguation, *Comput. Speech Lang.* 24 (2010) 461–473.
- [20] W. Jiang, Y. Yang, Y. Luo, X. Qin, Determining basic probability assignment based on the improved similarity measures of generalized fuzzy numbers, *Int. J. Comput. Commun. Control* 10 (2015) 333–347.
- [21] Y. Deng, Generalized evidence theory, *Appl. Intell.* 43 (2015) 530–543.
- [22] C. Fu, K.S. Chin, Robust evidential reasoning approach with unknown attribute weights, *Knowl.-Based Syst.* 59 (2014) 9–20.
- [23] X. Su, S. Mahadevan, P. Xu, Y. Deng, Dependence assessment in human reliability analysis using evidence theory and AHP, *Risk Anal.* 35 (2015) 1296–1316.
- [24] X. Deng, Y. Hu, Y. Deng, S. Mahadevan, Supplier selection using AHP methodology extended by D numbers, *Expert Syst. Appl.* 41 (2014) 156–167.
- [25] D. Wei, X. Deng, X. Zhang, Identifying influential nodes in weighted networks based on evidence theory, *Physica A* 392 (2013) 2564–2575.
- [26] Y. Deng, W. Jiang, R. Sadiq, Modeling contaminant intrusion in water distribution networks: A new similarity-based dst method, *Expert Syst. Appl.* 38 (2011) 571–578.
- [27] Y. Deng, R. Sadiq, W. Jiang, S. Tesfamariam, Risk analysis in a linguistic environment: a fuzzy evidential reasoning-based approach, *Expert Syst. Appl.* 38 (2011) 15438–15446.
- [28] Y. Liu, A method for 2-tuple linguistic dynamic multiple attribute decision making with entropy weight, *J. Intell. Fuzzy Syst.: Appl. Eng. Technol.* 27 (2014) 1803–1810.
- [29] W. Jiang, Y. Luo, X. Qin, J. Zhan, An improved method to rank generalized fuzzy numbers with different left heights and right heights, *J. Intell. Fuzzy Syst.* 28 (2015) 2343–2355.
- [30] C. Kahraman, S.C. Onar, B. Oztaysi, Fuzzy multicriteria decision-making: A literature review, *Int. J. Comput. Intell. Syst.* 8 (2015) 637–666.
- [31] Y. Deng, A threat assessment model under uncertain environment, *Math. Probl. Eng.* 2015 (2015) 878024. <http://dx.doi.org/10.1155/2015/878024>.
- [32] S.-B. Tsai, M.-F. Chien, Y. Xue, L. Li, X. Jiang, Q. Chen, J. Zhou, L. Wang, Using the fuzzy dematel to determine environmental performance: A case of printed circuit board industry in taiwan, *PLoS One* 10 (2015) e0129153.
- [33] K. Zhou, A. Martin, Q. Pan, Z.ga Liu, Median evidential c-means algorithm and its application to community detection, *Knowl.-Based Syst.* 74 (2015) 69–88.
- [34] R.S. Koff, Infectious diseases of humans: Dynamics and control. by r.m. anderson and r.m., *Hepatology* 15 (1992) 169. 757 pp. oxford: Oxford university press, 1991. 95.00.
- [35] Y. Deng, Y. Liu, D. Zhou, An improved genetic algorithm with initial population strategy for symmetric TSP, *Math. Probl. Eng.* 2015 (2015) 212794.
- [36] K.P. Yoon, C.L. Hwang, Multiple Attribute Decision Making: An Introduction, Vol. 104, Sage Publications, 1995.
- [37] Y. Du, C. Gao, Y. Hu, S. Mahadevan, Y. Deng, A new method of identifying influential nodes in complex networks based on topsis, *Physica A* 399 (2014) 57–69.
- [38] T. Opsahl, F. Agneessens, J. Skvoretz, Node centrality in weighted networks: Generalizing degree and shortest paths, *Social Networks* 32 (2010) 245–251.
- [39] J.-G. Liu, Z.-M. Ren, Q. Guo, Ranking the spreading influence in complex networks, *Physica A* 392 (2013) 4154–4159.
- [40] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* 393 (1998) 440–442.
- [41] A. Chin, M. Chignell, A social hypertext model for finding community in blogs, in: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, ACM, pp. 11–22.
- [42] R. Mahajan, N.T. Spring, D. Wetherall, Measuring ISP topologies with rocketfuel, in: Proc ACM SIGCOMM, Vol. 32, 2002, pp. 133–145.
- [43] M.G. Kendall, The treatment of ties in ranking problems, *Biometrika* 33 (1945).
- [44] K. Jin, W. Liang, C. Yan, J. Wang, L. Xia, H. Yong, Characterizing dynamic functional connectivity in the resting brain using variable parameter regression and kalman filtering approaches, *Neuroimage* 56 (2011) 1222–1234.
- [45] R. Criado, E. García, F. Pedroche, M. Romance, A new method for comparing rankings through complex networks: Model and analysis of competitiveness of major European soccer leagues, *Chaos* 23 (2013) 043114.
- [46] W. Hoeffding, A non-parametric test of independence, *Ann. Math. Statist.* 19 (1994) 546–557.
- [47] M. Hauskrecht, R. Pelikan, M. Valko, J. Lyons-Weiler, Feature selection and dimensionality reduction in genomics and proteomics, in: Fundamentals of Data Mining in Genomics and Proteomics, 2006, pp. 149–172.