Accepted Manuscript

Identifying influential nodes in complex networks based on AHP

Tian Bian, Jiantao Hu, Yong Deng

PII: S0378-4371(17)30219-4
DOI: http://dx.doi.org/10.1016/j.physa.2017.02.085
Reference: PHYSA 18065

To appear in: *Physica A*

Received date : 21 October 2016
Revised date : 8 February 2017

Please cite this article as: T. Bian, et al., Identifying influential nodes in complex networks based on AHP, *Physica A* (2017), http://dx.doi.org/10.1016/j.physa.2017.02.085

Highlights:

- Several different centrality measures are considered as the multi-attribute of complex network in AHP application.
- AHP is used to aggregate the multi-attribute for the evaluation of the influence of each node.
- The global structure and local information in complex networks are both taken into consideration to identify influential nodes.
- The proposed method can rank identify influential nodes accurately.

Identifying influential nodes in complex networks based on AHP

Tian Bian[a], Jiantao Hu[a], Yong Deng[a,*]

[a]*School of Computer and Information Science, Southwest University, Chongqing 400715, China*

Abstract

In the field of complex networks, how to identify influential nodes in the network is still an important research topic. In this paper, a method to identify the influence of the node based on Analytic Hierarchy Process (AHP) is proposed. AHP, as a multiple attribute decision making (MADM) technique has become an important branch of decision making since then. Every centrality measure has its own disadvantages and limitations, thus we consider several different centrality measures as the multi-attribute of complex network in AHP application. AHP is used to aggregate the multi-attribute to obtain the evaluation of the influence of each node. The experiments on four real networks and an informative network show the efficiency and practicability of the proposed method.

Keywords: Complex networks, Influential nodes, AHP, MADM, Centrality measure

*Corresponding author Yong Deng: School of Computer and Information Science, Southwest University, Chongqing 400715, China. Tel. : +86 23 68254555; fax:+86 23 68254555.

Email address: ydeng@swu.edu.cn, prof.deng@hotmail.com (Yong Deng)

1. Introduction

In recent years, complex network theory with advances in the understanding of the highly interconnected nature of various social, biological, and communication systems has gained much attention[1, 2, 3, 4, 5, 6, 7, 8]. Transferring information, trust, ideas, diseases and influences between any two nodes is the key function of complex networks[9, 10, 11]. The information can spread rapidly to a large number of nodes begin with an influential node. Hence, evaluating the influence of the nodes is a significant issue in complex networks [12], such as in the control of the disease and rumor dynamics[13, 14, 15], research on public opinion[16, 17], and creating new marketing tools [18, 19, 20, 21, 22, 23, 24, 25, 26, 27].

Many measurements of node centrality have been used commonly such as Degree centrality (DC)[28], Betweenness centrality (BC)[28, 29, 30], Closeness centrality (CC)[28] and so on. The DC method is very simple but of little relevance, since the measure does not take into consideration the global structure of the network. BC and CC are global metrics which can better identify influential nodes, but they are difficult to apply in large-scale networks due to their computational complexity. Another limitation of CC is the lack of applicability to networks with disconnected components: two nodes that belong to different components but do not have a finite distance between them. Several spectral centrality measures are also available, such as semi-local centrality(SLC)[31], eigenvector centrality (EC)[32], PageRank (PR)[33], and LeaderRank (LR)[34]. In SLC the topo-

logical connections among the neighbors are neglected, only the number of the nearest and the next nearest neighbors of a node is taken into account[35]. EC can not be applied to asymmetric networks in which some positions are unchosen[32]. PR, as well as LR, only has effect in directed networks, it will degenerate to DC in undirected networks[31].

Multiple Attribute Decision Making (MADM) (or called Multi-Criteria Decision-Making, MCDM) methods[36] have been proposed to decide a preferred alternative, classify alternatives in a small number of categories, and prioritize alternatives in a subjective preference order[37, 38, 39, 40]. It has been applied to many fields[41, 42]. Some math tools, such as fuzzy sets [43, 44], evidence theory [45, 46] and D numbers [47, 48], are widely used to address MADM. In this paper, we try to explore how to identify influential nodes by employing MADM methods. Du et al.[49] used D-C, CC, BC as the multi-attribute in TOPSIS to generate the ranking lists to evaluate the node spreading influence. Hu et al.[50] proposed a modified method by considering different weights of these attributes in TOPSIS (W-TOPSIS). Among numerous MADM methods developed to solve real-world decision problems, the Analytic Hierarchy Process (AHP) continues to work satisfactorily across different application areas. Thomas Saaty[51] originally introduced AHP to deal with complex decision making in a systematic and structured way, and may help the decision maker to set priorities and make the best decision with a finite number of criteria. By reducing complex decisions to a series of pairwise comparisons, and then synthesizing the results, the AHP helps to capture both subjec-

tive and objective aspects of a decision. In addition, the AHP incorporates a useful technique for checking the consistency of the decision maker's evaluations, thus reducing the bias in the decision making process[51, 52]. AHP makes full use of attribute information, provides a cardinal ranking of alternatives and it's inability to adequately handle the inherent uncertainty and imprecision in the data is also often criticized[49, 53, 54, 55]. As a well-known classical MADM method, AHP has gained much interest from researchers and practitioners[56, 57].

It is the first time for applying the AHP to identify influential nodes in complex networks. Firstly, we calculate the value of different centrality measures of the network. Notice that in different networks in which the information is transmitted in different ways, the different centrality measures need to be used. Then, AHP is utilized to evaluate the importance of nodes and identify influential nodes by regarding the centrality measures as the multi-attribute of complex networks. To evaluate the algorithmic performance, we use the Susceptible-Infected (SI) model[58] to examine the influence of the nodes ranked by the proposed method.

The paper is organized as follows. Section 2 and 3 begins with a brief overview of centrality measures and an introduction to AHP with a simple example. In Section 4, the method for identifying influential nodes based on AHP is proposed. Then, numerical examples are illustrated to show the efficiency and practicability of the proposed method, the Susceptible-Infected (SI) model and Kendall's tau coefficient is used to evaluate the performance in Section 5. Finally, some conclusions are presented in Sec-

tion 6.

2. Centrality measures

Considering a graph G = (V, E) with n = $|V|$ nodes and m = $|E|$ links. And the node centrality measurement of DC, CC and BC are well defined as follows.

Definition 1. (DC)[49]. The DC of node i, denoted as $CD(i)$, is defined as

$$C_D(i) = \sum_{j}^{N} x_{ij} \tag{1}$$

where i is the focal node, j represents all other nodes, N is the total number of nodes, and x_{ij} represents the connection between node i and node j. The value of x_{ij} is defined as 1 if node i is connected to node j, and 0 otherwise.

Definition 2. (BC)[49]. The BC of node i, denoted as $CB(i)$, is defined as

$$C_B(i) = \sum_{j,k \neq i} \frac{g_{jk}(i)}{g_{jk}} \tag{2}$$

where g_{jk} denotes the number of the shortest paths between node j and k, and $g_{jk}(i)$ means the number of the shortest paths between node j and k that go through node i.

Definition 3. (CC)[49]. The CC of node i, denoted as $CC(i)$, is defined as

$$C_C(i) = \left[\sum_{j}^{N} d_{ij} \right]^{-1} \tag{3}$$

where d_{ij} denotes the distance between node i and node j.

Definition 4. (Semi-local Centrality)[31]. The Semi-local centrality of node i, $C_l(i)$, is defined as

$$Q(v) = \sum_{u \in \Gamma(v)} N(u) \tag{4}$$

$$C_l(i) = \sum_{j \in \Gamma(i)} Q(j) \tag{5}$$

where $\Gamma(v)$ is the set of the nearest neighbors of node v and $N(u)$ is the number of the nearest and the next nearest neighbors of node u.

Definition 5. (EC)[32]. Let $\mathbf{A}$ be an $n \times n$ similarity matrix. The eigenvector centrality x_i of node i is defined as the ith entry in the normalized eigenvector belonging to the largest eigenvalue of $\mathbf{A}$. λ is the largest eigenvalue of $\mathbf{A}$ and n is the number of vertices:

$$Ax = \lambda x, x_i = u \sum_{j=1}^{n} a_{ij} x_j, i = 1, 2, ..., n \tag{6}$$

with proportionality factor $u = \frac{1}{\lambda}$ so that x_i is proportional to the sum of similarity scores of all nodes connected to it.

Definition 6. (PR)[33]. PageRank algorithm[59] is a famous variant of eigenvector centrality and is used to rank websites in Google search engine and other commercial scenarios[60]. Similar to eigenvector centrality, PageRank supposes that the importance of a webpage is determined by both the quantity and the quality of the pages linked to it[61]. Initially, each node gets one unit PR value. Then every node evenly distributes the

PR value to its neighbors along its outgoing links. Mathematically, the PR value of node v_i at t step is

$$PR_i(t) = \sum_{j=1}^{n} a_{ji} \frac{PR_j(t-1)}{k_j^{out}} \tag{7}$$

where n is the total number of nodes in the network, and k_j^{out} is the out-degree of node v_j. The above iteration will stops if the PR values of all nodes reach the steady state.

Definition 7. (LR)[34]. The Leader Rank was proposed by Lu et al.[34] which is a variant of the Page Rank algorithm. In general, they introduced a ground node which bidirectionally connects to all other nodes. Then the random walk process is used to find influential nodes and this process continues until the steady state is attained. The process can be described by the stochastic matrix **P** with a probability $p_{ij} = \frac{a_{ij}}{k_i^{out}}$ that a random walker at node i goes to j the next time, where $a_{ij} = 1$ means node i pointing to node j, and k_i^{out} denotes the out-degree. So the score of node i at time t can be defined as

$$s_i(t+1) = \sum_{j=1}^{N+1} \frac{a_{ji}}{k_j^{out}} s_j(t) \tag{8}$$

The initial scores for all node i is 1 while 0 for the ground node. When the score $s_i(t)$ for all i converges to a unique steady state which can be denoted as $s_i(t_c)$, where t_c is the convergence time. Thus, the final score of a node is defined as

$$S_i = s_i(t_c) + \frac{s_g(t_c)}{N} \tag{9}$$

where $s_g(t_c)$ is the score of the ground node at steady state. Hence, we can rank each node with the S_i value down. (For more detailed information please refer to [34].)

3. The basic theory of AHP

The Analytic Hierarchy Process(AHP), proposed by Saaty[51, 52], are theories of relative measurement of intangible criteria[62]. Saaty proposes the decision makers preferences by formulating the Fundamental Scale(Table 1, from Ref.[52]). Generally, the process to solve an MCDM problem using AHP can be divided into the following subsections:

3.1. Modelling Decision Problem as a Hierarchy

The decision-making problem is established as a hierarchy structure and is recursively decomposed into several levels. The main goal of the decision problem is the top level of the hierarchy structure. The tangible and/or intangible criteria and subcriteria that contribute to the goal are the lower levels.The alternatives to evaluate in terms of the criteria formed the bottom level[63].

3.2. Computing the vector of criteria weights

First, we create a $m \times m$ pairwise comparison matrix $\mathbf{A}$, and m is the number of criteria. In the matrix $\mathbf{A}$, each element a_{ij} represents the importance of the ith criterion relative to the jth criterion. The element a_{ij} satisfy the following properties: (1) $a_{ij}\dot{a}_{ji} = 1$; (2) $a_{ii} = 1$; (3) $a_{ij} > 0$. The numerical

Intensity of importance	Definition
1	Equal importance/preference
2	Weak
3	Moderate importance/preference
4	Moderate plus
5	Strong importance/preference
6	Strong plus
7	Very strong or demonstrated importance/preference
8	Very, very strong
9	Extreme importance/preference

Table 1: Saatys fundamental scale.

rating including 9 scales between two criteria is measured, as shown in Table 1, where it is assumed that the ith criterion is equally or more important than the jth criterion. Finally, the priority weights of criteria $\mathbf{w}$ is built by the equation:

$$Aw = \lambda_{max} w, w = (w_1, w_2, ..., w_n)^T \tag{10}$$

where λ_{max} is the largest eigenvalue of matrix $\mathbf{A}$.

3.3. Computing the matrix of option scores

First, assume a $n \times m$ real matrix $\mathbf{S}$. Each entry s_{ij} of $\mathbf{S}$ represents the score of the ith option with respect to the jth criterion. Then, for each of

the m criteria, a pairwise comparison matrix $\mathbf{B}^{(j)}$ is first built for each of the m criteria, j=1,...,m. The matrix $\mathbf{B}^{(j)}$ is a $n \times n$ real matrix, where n is the number of alternatives. Each element $b_{ih}^{(j)}$ of the matrix $\mathbf{B}^{(j)}$ represents the evaluation of the ith alternative compared to the hth alternative with respect to the jth criterion. The evaluation scales suggested in Table 1 may also be used to in this section.

Second, the same process described for the pairwise comparison matrix $\mathbf{A}$ is applied to each matrix $\mathbf{B}^{(j)}$, then the alternative vectors $s^{(j)}$ (j=1, ..., m) are obtained. The vector $\mathbf{s}^{(j)}$ contains the scores of the evaluated options with respect to the jth criterion. Finally, the score matrix $\mathbf{S}$ is constructed as

$$\mathbf{S} = [\mathbf{s}^{(1)} \cdots \mathbf{s}^{(m)}] \tag{11}$$

3.4. Ranking the options

Finally, the scores $\mathbf{v}$ is obtained by multiplying the weight vector $\mathbf{w}$ and the score matrix $\mathbf{S}$,

$$\boldsymbol{v} = \mathbf{S} \cdot \boldsymbol{w} \tag{12}$$

where the ith element v_i of $\mathbf{v}$ represents the score of the ith alternative calculated by AHP.

4. Proposed method

4.1. Basic method

In different networks information is transmitted in different ways, a great number of centrality measures have been proposed for identifying

influential nodes. However, every measure has its limitations and disadvantages[64]. So the different types of centrality measures need to be taken into account when evaluating a network. Multiple Attribute Decision Making (MADM) problem is a good choice to address this issue. Du et al.[49] has applied TOPSIS to identify influential nodes in complex networks in 2014. Hu et al.[50] proposed a modified method by considering different weights of these attributes in TOPSIS. In this paper, the AHP is introduced to identify the influence nodes in complex network. Degree centrality (DC), closeness centrality (CC) and betweenness centrality (BC) are widely used in identifying the influential nodes in complex network. Thus, in the proposed approach, they are chosen as the criteria for decision making.

The proposed method based on AHP is composed of the following steps(shown in Fig. 4):

Step 1: Calculate the centrality value.

Degree centrality (DC), closeness centrality (CC) and betweenness centrality (BC) are calculated first, which are taken into account as the criteria in AHP application. Assuming the network has m nodes, so we can get a $m \times 3$ matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} DC_1 & BC_1 & CC_1 \\ DC_2 & BC_2 & CC_2 \\ \vdots & \vdots & \vdots \\ DC_m & BC_m & CC_m \end{pmatrix}$$

Step 2: Calculate the weights.

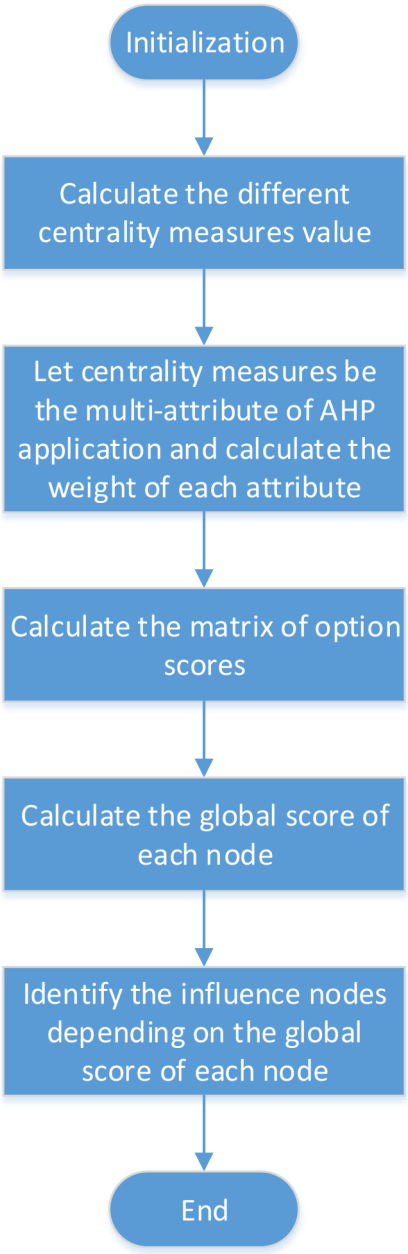

Figure 1: The flowchart of the proposed method.

In order to apply AHP to identify influence nodes of complex networks, we should get the weight of each criterion firstly. However, the method of obtaining the weights in the original AHP is not applicable. Hu's[50] method to calculate the weights of each attribute in w-TOPSIS is effective. In this method the Susceptible-Infected (SI) model[58] is used to calculate the weights. In SI model, $S(t)$ and $I(t)$ are respectively the numbers of susceptible and infected individuals at time t. In each step, only the individuals which have been infected are able to spread the disease to susceptible individuals, and every susceptible individual gets infected with probability λ (for uniformity, we set $\lambda = 0.3$) from the infected neighbor. The total number of infected nodes at time t can be considered as an indicator to evaluate the influence of the initially infected node, namely $F(t)$[31]. Clearly, the number of cumulative infected nodes increases with time t and eventually reaches a steady value. For different nodes initially infected, higher $F(t)$ represents a larger influence. Thus the weight of each attribute can be calculated by matching with $F(t)$. The higher the matching degree, the bigger the weight of the attribute. The process for calculating the weights is illustrated as the following:

Consider a matrix $\mathbf{D}$, the first 3 columns represent the value of different centrality measure. The last column represents the results of SI model($F(t)$)[50]:

$$\mathbf{D} = \begin{pmatrix} DC_1 & BC_1 & CC_1 & F_1(t) \\ DC_2 & BC_2 & CC_2 & F_2(t) \\ \vdots & \vdots & \vdots & \vdots \\ DC_m & BC_m & CC_m & F_m(t) \end{pmatrix} \tag{13}$$

Normalize the matrix[50]:

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}}, i = 1, 2, \ldots, m; j = 1, 2, 3, 4 \tag{14}$$

Match the attribute to $F(t)$ as follows[50]:

$$v_{ij} = \frac{1}{|r_{ij} - r_{i4}|}, i = 1, 2, \ldots, m; j = 1, 2, 3 \tag{15}$$

where r_{i4} represents $F(t)$.

Calculate the value of each criterion e_j by sum the v_{ij} obtained in Eq. (15):

$$e_j = \sum_{i=1}^{m} v_{ij}, i = 1, 2, \ldots, m; j = 1, 2, 3 \tag{16}$$

Finally, the weight of criterion j, w_j, is obtained by normalizing the e_j calculated in Eq. (16).

$$w_j = \frac{e_j}{\sum_{j=1}^{n} e_j}, j = 1, 2, 3 \tag{17}$$

Step 3: Calculate the matrix of option scores

In this step, three pairwise comparison matrix $\mathbf{B}^{(j)}$ is first built for each of the criteria (DC, BC and CC), j=1, 2, 3. Each entry $b_{ih}^{(j)}$ of the matrix $\mathbf{B}^{(j)}$ is calculated by dividing the value of ith node and the value of hth node with

respect to the *j*th criterion. Then the AHP applies to each matrix $\mathbf{B}^{(j)}$ the same two-step procedure described for the pairwise comparison matrix **A** in Section 3.2, thus obtaining the score vectors s(j), j=1, 2, 3. The vector $s^{(j)}$ contains the scores of the nodes with respect to the *j*th criterion. Finally, the score matrix **S** is obtained.

Step 4: Ranking the nodes

By multiplying **S** and **w** obtained above, we can obtain the global score of each node in the network.

4.2. Example explanation

Based on what has been proposed above, a simple example is given to explain how AHP performs in this part. The Kite network is a classical interpersonal relationship network with 10 nodes designed by Krackhardt[65](shown in Fig. 2). The different centrality measure values of each node are shown in Table 2.

Table 2: The different centrality measure values of each node of 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

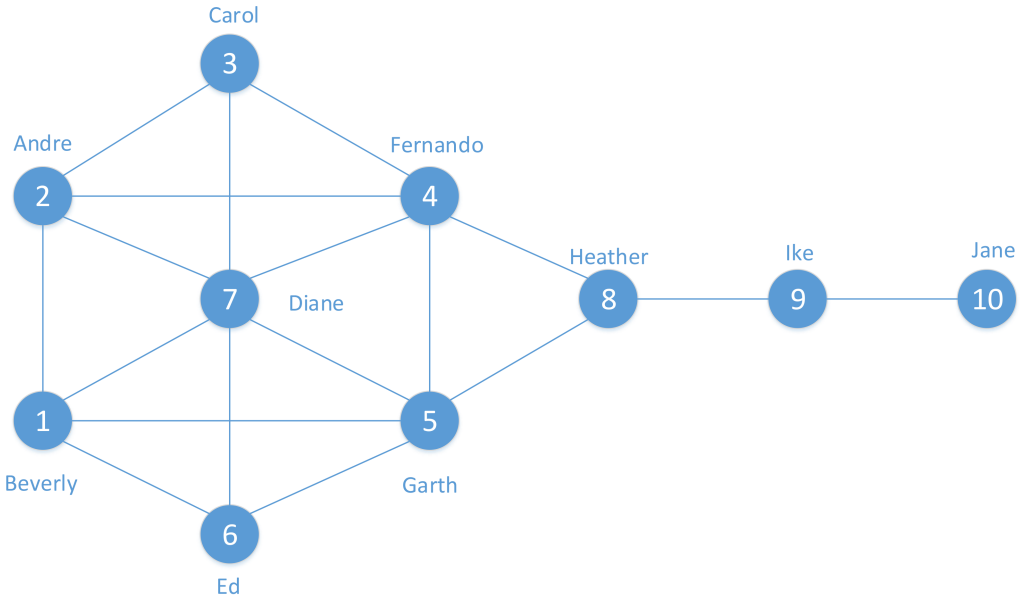

Figure 2: The Kite network.

Firstly, we can construct a matrix **A** from the in Table 2:

$$\mathbf{A} = \begin{pmatrix} & DC & BC & CC \\ Diane & 6 & 0.067 & 0.163 \\ Fernando & 5 & 0.071 & 0.370 \\ Grath & 5 & 0.071 & 0.370 \\ Andre & 4 & 0.059 & 0.037 \\ Beverly & 4 & 0.059 & 0.037 \\ Carol & 3 & 0.056 & 0.000 \\ Ed & 3 & 0.056 & 0.000 \\ Heather & 3 & 0.067 & 0.622 \\ Ike & 2 & 0.048 & 0.356 \\ Jane & 1 & 0.034 & 0.000 \end{pmatrix}$$

Because the Kite network is very small so we let t=6. Then we can get the matrix **D** of which the last column represents the result of $F(t)$.

$$\mathbf{D} = \begin{pmatrix} 6 & 0.067 & 0.163 & 8 \\ 5 & 0.071 & 0.370 & 6 \\ 5 & 0.071 & 0.370 & 6 \\ 4 & 0.059 & 0.037 & 3 \\ 4 & 0.059 & 0.037 & 3 \\ 3 & 0.056 & 0.000 & 1 \\ 3 & 0.056 & 0.000 & 1 \\ 3 & 0.067 & 0.622 & 3 \\ 2 & 0.048 & 0.356 & 2 \\ 1 & 0.034 & 0.000 & 1 \end{pmatrix}$$

Then normalize the matrix **D** by Eq. (14), we can get the matrix **R**:

$$\mathbf{R} = \begin{pmatrix} 0.1667 & 0.1139 & 0.0834 & 0.2353 \\ 0.1389 & 0.1207 & 0.1893 & 0.1765 \\ 0.1389 & 0.1207 & 0.1893 & 0.1765 \\ 0.1111 & 0.1003 & 0.0189 & 0.0882 \\ 0.1111 & 0.1003 & 0.0189 & 0.0882 \\ 0.0833 & 0.0952 & 0 & 0.0294 \\ 0.0833 & 0.0952 & 0 & 0.0294 \\ 0.0833 & 0.1139 & 0.3182 & 0.0882 \\ 0.0556 & 0.0816 & 0.1821 & 0.0588 \\ 0.0278 & 0.0578 & 0 & 0.0294 \end{pmatrix}$$

Next, match the attributes to F(t) by Eq. (15), we can get a new matrix **V**:

$$\mathbf{V} = \begin{pmatrix} 0.1277 & 0.1268 & 0.1263 \\ 0.1706 & 0.1701 & 0.1721 \\ 0.1706 & 0.1701 & 0.1721 \\ 0.3462 & 0.3449 & 0.3354 \\ 0.3462 & 0.3449 & 0.3354 \\ 1.0909 & 1.1053 & 1.0000 \\ 1.0909 & 1.1053 & 1.0000 \\ 0.3429 & 0.3465 & 0.3729 \\ 0.5143 & 0.5213 & 0.5501 \\ 1.0286 & 1.0614 & 1.0000 \end{pmatrix}$$

Calculate the weight of each attribute by Eq. (16) and Eq. (17). We can obtain the weights of all attributes. w=(0.3354 0.3397 0.3249).

Secondly, according to the step 3 described in subsection 4.1, we can get three pairwise comparison matrix $\mathbf{B}^{(j)}$(j=1, 2, 3).

$$\mathbf{B}^{(1)} = \begin{pmatrix} 1 & 1 & 4/3 & 4/5 & 4/5 & 4/3 & 2/3 & 4/3 & 2 & 4 \\ 1 & 1 & 4/3 & 4/5 & 4/5 & 4/3 & 2/3 & 4/3 & 2 & 4 \\ 3/4 & 3/4 & 1 & 3/5 & 3/5 & 1 & 1/2 & 1 & 3/2 & 3 \\ 5/4 & 5/4 & 5/3 & 1 & 1 & 5/3 & 5/6 & 5/3 & 5/2 & 5 \\ 5/4 & 5/4 & 5/3 & 1 & 1 & 5/3 & 5/6 & 5/3 & 5/2 & 5 \\ 3/4 & 3/4 & 1 & 3/5 & 3/5 & 1 & 1/2 & 1 & 3/2 & 3 \\ 3/2 & 3/2 & 2 & 6/5 & 6/5 & 2 & 1 & 2 & 3 & 6 \\ 3/4 & 3/4 & 1 & 3/5 & 3/5 & 1 & 1/2 & 1 & 3/2 & 3 \\ 1/2 & 1/2 & 2/3 & 2/5 & 2/5 & 2/3 & 1/3 & 2/3 & 1 & 2 \\ 1/4 & 1/4 & 1/3 & 1/5 & 1/5 & 1/3 & 1/6 & 1/3 & 1/2 & 1 \end{pmatrix}$$

$$\mathbf{B}^{(2)} = \begin{pmatrix} 1 & 1 & 0 & 5/48 & 5/39 & 0 & 5/6 & 5/78 & 1/9 & 0 \\ 1 & 1 & 0 & 5/48 & 5/39 & 0 & 5/6 & 5/78 & 1/9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 48/5 & 48/5 & 0 & 1 & 16/13 & 0 & 8 & 8/13 & 16/15 & 0 \\ 39/5 & 39/5 & 0 & 13/16 & 1 & 0 & 13/2 & 1/2 & 13/15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6/5 & 6/5 & 0 & 1/8 & 2/13 & 0 & 1 & 1/13 & 2/15 & 0 \\ 78/5 & 78/5 & 0 & 13/8 & 2 & 0 & 13 & 1 & 26/15 & 0 \\ 9 & 9 & 0 & 15/16 & 15/13 & 0 & 15/2 & 15/26 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{B}^{(3)} = \begin{pmatrix} 1 & 1 & 18/17 & 14/17 & 14/17 & 18/17 & 15/17 & 15/17 & 21/17 & 29/17 \\ 1 & 1 & 18/17 & 14/17 & 14/17 & 18/17 & 15/17 & 15/17 & 21/17 & 29/17 \\ 17/18 & 17/18 & 1 & 7/9 & 7/9 & 1 & 5/6 & 5/6 & 7/6 & 29/18 \\ 17/14 & 17/14 & 9/7 & 1 & 1 & 9/7 & 15/14 & 15/14 & 3/2 & 29/14 \\ 17/14 & 17/14 & 9/7 & 1 & 1 & 9/7 & 15/14 & 15/14 & 3/2 & 29/14 \\ 17/18 & 17/18 & 1 & 7/9 & 7/9 & 1 & 5/6 & 5/6 & 7/6 & 29/18 \\ 17/15 & 17/15 & 6/5 & 14/15 & 14/15 & 6/5 & 1 & 1 & 7/5 & 29/15 \\ 17/15 & 17/15 & 6/5 & 14/15 & 14/15 & 6/5 & 1 & 1 & 7/5 & 29/15 \\ 17/21 & 17/21 & 6/7 & 2/3 & 2/3 & 6/7 & 5/7 & 5/7 & 1 & 29/21 \\ 17/29 & 17/29 & 18/29 & 14/29 & 14/29 & 18/29 & 15/29 & 15/29 & 21/29 & 1 \end{pmatrix}$$

to which correspond the score vectors

$\mathbf{s}^{(1)} = [0.1111, 0.1111, 0.0833, 0.1389, 0.1389, 0.0833, 0.1667, 0.0833, 0.0556, 0.0278]^T$,

$\mathbf{s}^{(2)} = [0.0155, 0.0155, 0, 0.1487, 0.1208, 0, 0.0186, 0.2416, 0.1394, 0]^T$,

$\mathbf{s}^{(3)} = [0.1002, 0.1002, 0.0946, 0.1217, 0.1217, 0.0946, 0.1136, 0.1136, 0.0811, 0.0587]^T$.

Hence, the score matrix $\mathbf{S}$ is obtained,

$$\mathbf{S} = [\mathbf{s}^{(1)} \quad \mathbf{s}^{(2)} \quad \mathbf{s}^{(3)}] = \begin{pmatrix} 0.1111 & 0.0155 & 0.1002 \\ 0.1111 & 0.0155 & 0.1002 \\ 0.0833 & 0 & 0.0946 \\ 0.1389 & 0.1487 & 0.1217 \\ 0.1389 & 0.1208 & 0.1217 \\ 0.0833 & 0 & 0.0946 \\ 0.1667 & 0.0186 & 0.1136 \\ 0.0833 & 0.2416 & 0.1136 \\ 0.0556 & 0.1394 & 0.0811 \\ 0.0278 & 0 & 0.0587 \end{pmatrix}$$

Finally, by multiplying $\mathbf{S}$ and $\mathbf{w}$ obtained above, we can obtain the global score vector $\mathbf{v} = [0.0762 \quad 0.0762 \quad 0.0599 \quad 0.1363 \quad 0.1272 \quad 0.0599 \quad 0.1005 \quad 0.1451 \quad 0.0915 \quad 0.0291]^T$.

The ranking results of each node with different centrality measure are shown in Table 3. From the Table 3, we can know that node 8(Heather) is ranked as the most influential node in the proposed method. In the Kite network, node Heather is like an important connection point between the kite and the line, so it's not too much to be ranked first. And node 4(Fernando), node 5(Grath) and node 7(Diane) are like a kite's support, they should be ranked in the following. Fernando and Grath is closer to the connection point, so they are more important. Node 1(Beverly), node 2(Andre), node 3(Carol), node 6(Ed) and node 10(Jane) are on the edge of the kite network so it should be ranked in the back of the node 9(Ike),

Table 3: The ranking results of each node with different centrality measure.

DC	CC	BC	Proposed method
7	4	8	8
4	5	4	4
5	7	9	5
1	8	5	7
2	1	7	9
3	2	1	1
6	3	2	2
8	6	3	3
9	9	6	6
10	10	10	10

which plays the role of the kite line. Therefore the proposed method is more reasonable than other centrality measures.

5. Experimental analysis

5.1. Data

In order to verify the efficiency of the proposed method, we use four real network to evaluate the performance. (i) Email:the network of e-mail interchanges between members of the University Rovirai Virgili(Tarragona)[66]. The data can be downloaded from http://www.cs.bris.ac.uk/steve/peacockpaper; (ii) Groad: It includes the adjacency matrix (value 1 if two locations are directly connected by a highway) and the labels of all 1,168 nodes[67]. (iii) 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/; (iv) USair97 network: It has 322 airports and the air line between two air port can be denoted as a connection between two nodes in the network. The data can be downloaded from http://pajek.imfm.si/doku.php?id=data:pajek:vlado&s[]=air. The basic topological properties of these four networks are shown in Table 4.

5.2. Effectiveness

In this section, DC, CC, BC, EC, SLC, PR and LR are taken to compare with the proposed method. The lists of top-10 influential nodes for four networks are shown in Table 5 and Table 6. Because PR, as well as LR, only

Table 4: The basic topological features of the four real networks. n and m are the total number of nodes and links, respectively. $< k >$ and k_{max} denote the average and the maximum degree. C is the clustering coefficient[68].

Network	n	m	$< k >$	k_{max}	C
Email	1133	5451	9.62	71	0.2202
Groad	1168	1243	2.13	12	0.0006
Yeast	2375	11693	9.85	118	0.1529
USair97	332	2126	12.81	139	0.3964

has effect in directed networks, we don't compare the result of PR and LR with the proposed method in Email, Groad and Yeast networks. In Email, the results between the proposed method and BC or CC have the same nine members in the top-10 lists; the proposed method and SLC have the same eight members in the top-10 lists; the proposed method and DC have the same seven members in the top-10 lists. In Groad, there are the same eight members in the top-10 lists by comparing the proposed method and BC, the proposed method and DC have the same five members in the top-10 lists. In Yeast, in spite of the fact that the results of different measures have many differences, the proposed method and DC, CC, BC still have 2, 3 and 7 same nodes in the top 10 nodes. In USAir97, the number of the same members in the top-10 lists between the proposed method and other centrality measures are 8, 7 8, 5, 9 and 6 respectively in DC, CC, BC, SLC, PR and LR.

The SI model is also used to compare the proposed method with W-

Table 5: The top-10 ranked nodes by degree centrality (DC), closeness centrality (CC), betweenness centrality (BC), eigenvector centrality (EC), semi-local centrality(SLC), W-TOPSIS and the proposed method in Email, Groad and Yeast.

Email							
Rank	DC	CC	BC	EC	SLC	W-TOPSIS	Proposed method
1	105	333	23	105	333	176	105
2	333	23	105	16	23	333	23
3	42	105	333	196	41	23	333
4	23	42	76	204	233	76	42
5	16	41	42	42	52	42	76
6	41	76	578	49	42	41	41
7	196	233	135	56	3	233	135
8	233	52	41	116	135	578	578
9	76	135	52	333	54	135	52
10	21	378	355	3	76	355	233
Groad							
Rank	DC	CC	BC	EC	SLC	W-TOPSIS	Proposed method
1	693	698	219	219	198	693	219
2	403	219	693	217	217	403	693
3	300	450	543	267	219	219	543
4	217	565	758	198	267	758	758
5	373	331	403	207	312	373	403
6	410	763	373	331	331	543	373
7	758	267	698	565	450	698	698
8	207	663	565	236	654	331	565
9	219	729	556	295	236	565	556
10	331	347	763	231	127	300	763
Yeast							
Rank	DC	CC	BC	EC	SLC	W-TOPSIS	Proposed method
1	68	321	321	176	321	321	321
2	176	1302	174	933	1302	375	375
3	933	288	375	192	246	174	174
4	91	246	2084	949	1335	2084	2084
5	192	1335	433	518	1360	433	433
6	206	1360	120	175	410	288	120
7	518	375	1711	924	288	120	1711
8	949	11	1194	177	144	1711	68
9	1101	892	11	112	375	1194	1101
10	112	410	892	186	400	246	288

Table 6: The top-10 ranked nodes by degree centrality (DC), closeness centrality (CC), betweenness centrality (BC), eigenvector centrality (EC), semi-local centrality(SLC), PageRank (PR), LeaderRank (LR), W-TOPSIS and the proposed method in USair97.

USAir97									
Rank	DC	CC	BC	EC	SLC	PR	LR	W-TOPSIS	Proposed method
1	118	118	118	332	313	261	118	118	230
2	261	261	8	14	8	118	248	261	255
3	255	67	261	15	47	152	261	8	201
4	152	255	47	12	67	182	47	47	47
5	182	201	201	10	118	8	201	67	67
6	230	182	67	11	201	201	67	201	152
7	166	47	313	9	248	255	166	182	182
8	67	166	13	18	65	230	255	152	8
9	112	248	182	17	112	67	258	255	261
10	201	122	152	19	144	166	311	230	118

TOPSIS[50]. According to the SI model, the more outstanding infectious ability the node is, the more important it is. The top-10 lists generated by weighted TOPSIS and the proposed method respectively are shown in Table 5 and Table 6. We compare the influence of the nodes that either appear in the top-10 list by the proposed method or W-TOPSIS. The simulations on the cumulative infected nodes for these four networks, are shown in Fig. 3. In Email, it's obvious that proposed method outperforms W-TOPSIS. In Groad, the result for the proposed method and W-TOPSIS is similar, because the lines of the proposed method and W-TOPSIS almost overlap as shown in Fig. 3(b). We also can find that the average number of infected nodes by the proposed method is slightly better than W-TOPSIS at each step. In Yeast, we can also see that the proposed method outper-

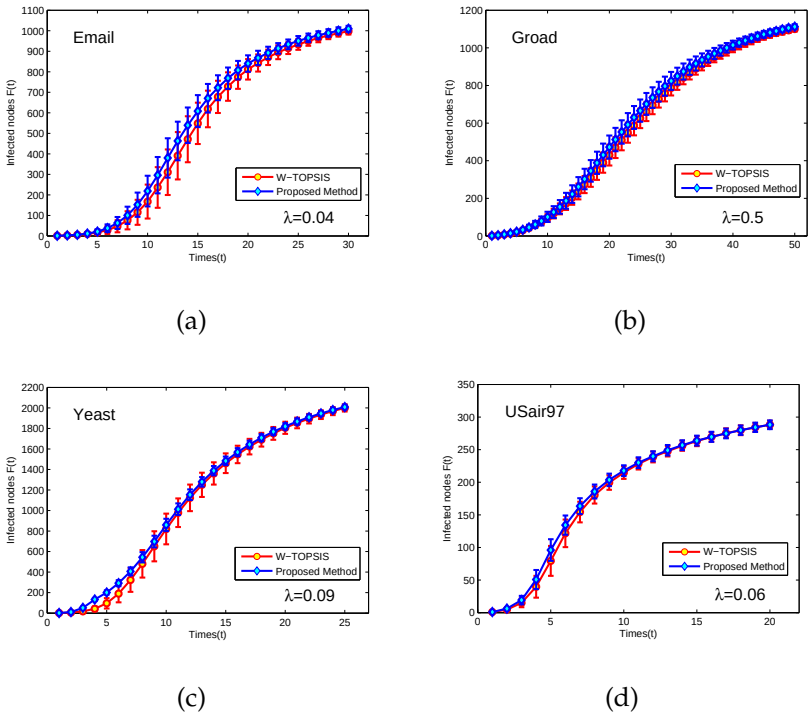

Figure 3: The cumulative number of infected nodes as a function of time in Email, Groad, Yeast and USair97 with the initially infected nodes being those that either appear in the top-10 list by the proposed method or W-TOPSIS.

forms W-TOPSIS, because the average number of infected nodes in each step by the proposed method is larger than W-TOPSIS at the beginning of infecting process shown in Fig. 3(c). In USair97, Fig. 3(d) presents the cumulative number of the infected number of the proposed method and W-TOPSIS. Obviously, the proposed method outperforms W-TOPSIS.

Finally, Kendall's tau coefficient τ [69] are used to measure the correlation between the node's influence measured by two methods and the result simulated by the SI model. It assumes that there are two random

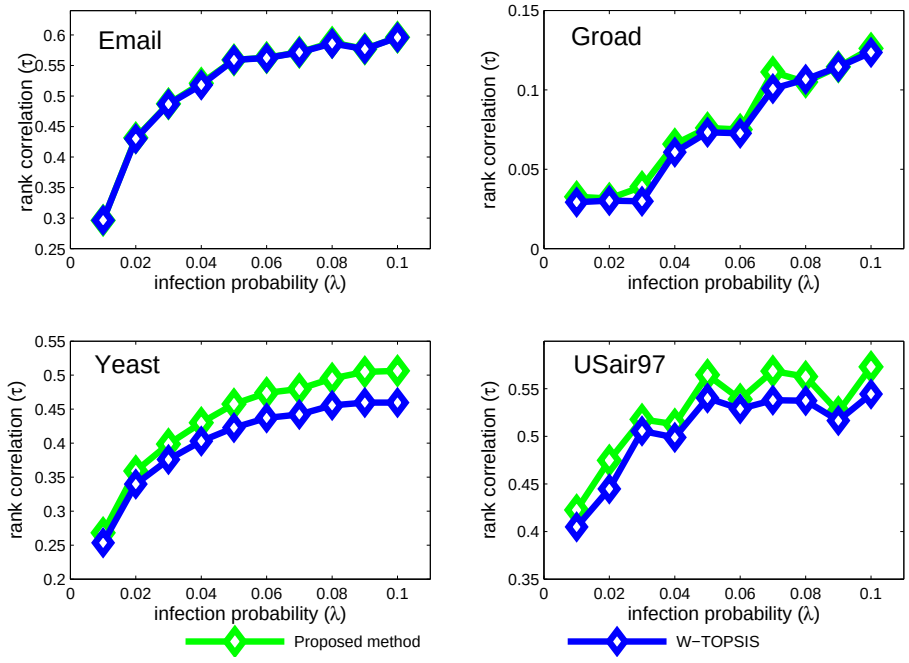

Figure 4: The Kendall's tau coefficient τ obtained by comparing the ranking list generated by the two methods and the ranking list generated by SI model on four real networks. The results are obtained by over 100 independent runs where the spreading probability λ is ranging from 0.01 to 0.1.

variables X and Y, the ith value of two random variables are respectively considered as X_i and Y_i. The corresponding elements in X and Y form an element pair of the set XY, which contains the elements (X_i, Y_i). If both $X_i > X_j$ and $Y_i > Y_j$ or if both $X_i < X_j$ and $Y_i < Y_j$, these two elements (X_i, Y_i), (X_j, Y_j) are considered coincident. In contrast, if $X_i > X_j$ and $Y_i < Y_j$ or if $X_i < X_j$ and $Y_i > Y_j$, they are said to be discordant. If $X_i = X_j$ or $Y_i = Y_j$, these two elements are neither concordant nor discordant. The Kendall's tau coefficient is defined as[70, 71]:

$$\tau = \frac{n_c - n_d}{0.5n(n-1)} \tag{18}$$

where n_c and n_d represent the number of concordant and discordant pairs respectively. Kendall's tau coefficient can clearly show the matching extent between two kinds of sorts, it can be more intuitive to reflect the correlation between the two methods in the numerical results. So we set the spreading probability of SI model are gradually increasing from 0.01 to 0.1, then calculate the Kendall's tau coefficient of different sorting methods and SI models with the change of spreading probability, the results are shown in Fig. 4. It is clear that the higher the value of the Kendall coefficient τ, the higher the correlation between the method and the SI model. From Fig. 4, we can see that the proposed method outperforms the W-TOPSIS on strongly positive correlation with real spreading process in Groad, Yeast, and USair97. In Email, our method and W-TOPSIS almost have the same performance. But our method is slightly better than W-TOPSIS around $\lambda = 0.09$.

5.3. The proposed method applied in informative networks

In recent years, the concepts of complex networks have been widely used in textual analysis[72, 73, 74, 75, 76, 77]. In this section, an application of keywords detection is employed to illustrate the validity and practicability of the proposed method[78]. We use the word adjacencies network to represent adjacency of words in English text, in this case the novel *DavidCopperfield*[79] by Charles Dickens is used to carry out the analysis. The vertices in the informative network represent words and if one word adjoins one another at any point in the text then an edge will generate to connect these two words. There are two traditional approach-

Table 7: Top ten keywords ranked using entropic approach (E), clustering approach (C), DC, CC, BC, EC, SLC, PR, LR and the proposed method from the novel *DavidCopperfield*.

Rank	E	C	DC	CC	BC	EC	SLC	PR	LR	Proposed method
1	other	little	little	little	little	half	little	little	little	little
2	good	old	old	old	old	glad	face	old	old	old
3	young	good	other	good	other	fancy	friend	other	other	other
4	dear	other	good	other	first	lost	way	good	good	good
5	pretty	same	same	way	good	family	place	first	first	first
6	great	room	first	same	whole	world	right	whole	same	same
7	best	way	way	first	man	full	other	same	whole	man
8	eye	small	dear	time	better	greater	good	man	man	whole
9	short	poor	man	room	way	anything	pretty	nothing	better	way
10	beautiful	black	young	thing	same	true	old	heart	dear	better

es used to copy with the problem of keyword detection have been proved to have good effect: the entropic and clustering approaches. The entrop-

ic approach (E) was proposed by Herrera and Pury[80], at first we need to divide the text composed of N words into P parts. Then a probability measure over the partition $p_i(w)$ for every word type w can be defined as follows[78]:

$$p_i(w) = \frac{f_i(w)}{\sum_{j=1}^{P} f_j(w)} (i = 1, ..., P), \tag{19}$$

where $f_i(w)$ is the relative frequency of occurrence of the word type w in the ith part.

Shannons entropy is employed by the expression as follows:

$$S(w) = -\frac{1}{ln(P)} \sum_{i=1}^{P} P_i(w) ln(p_i(w)). \tag{20}$$

Finally, Herrera and Pury[80] normalized the measure as follows:

$$E_{nor}(w) = \frac{1 - S(w)}{1 - S_{ran}(w)} \tag{21}$$

where $S_{ran} \approx 1 - \frac{P-1}{2nln(P)}$ means a word with absolute frequency n in a random text. The larger the value of $E_{nor}(w)$, the larger its relevance.

The clustering approach (C) was developed by Ortuño et al.[81], and then modified by Carpena et al.[82]. Considering a text has N words and a word with frequency of occurrence n in that text. We denote the inter-occurrences distances as $d_i (i = 1, 2, ..., n - 1)$. Ortuño et al.[81] proposed $\sigma = s/ < d >$ as a measure of the relevance of the word, where $s = (< d^2 > - < d >^2)^{1/2}$ is the standard deviation and $< d >$ is the average distance. Then Carpena et al.[82] normalized the σ as follows:

$$\sigma_{nor} = \frac{\sigma}{\sqrt{1 - p}} \tag{22}$$

where $\sqrt{1-p}$ is the expected value of σ for a word with probability p randomly distributed, $p = n/N$.

Finally, the clustering approach C, that uses the clustering σ_{nor} and the frequency n, is defined as follows:

$$C(\sigma_{nor}, n) = \frac{\sigma_{nor} - < \sigma_{nor} > (n)}{sd(\sigma_{nor})(n)}, \tag{23}$$

where the mean $< \sigma_{nor} >$ and the standard deviation $sd(\sigma_{nor})(n)$ is well fitted by the functions:

$$< \sigma_{nor} >= \frac{2n-1}{2n+2}, sd(\sigma_{nor}) = \frac{1}{\sqrt{n}(1+2.8n^{-0.865})} \tag{24}$$

The larger the value of C, the larger its relevance.

The top-10 keywords ranked using entropic approach, clustering approach, DC, CC, BC, EC, SLC, PR, LR and the proposed method from the novel $DavidCopperfield$ are shown in Table 7, from the results we can observe that the proposed method works better than other centrality measures, because the proposed method has most of the same words with other measures in the top-10 lists, where the number of the same words in the top-10 lists between the proposed method and other measures are 2, 6, 8, 7, 10, 5, 8 and 9 respectively in entropic approach, clustering approach, DC, CC, BC, SLC, PR and LR.

6. Conclusion

In this paper, a new method is proposed to identify the influential nodes in complex network based on the AHP. In our method, we consider several different centrality measures as the multi-attribute of complex

network in AHP application and give the corresponding weights to each attribute according to the matching degree with $F(t)$. AHP is used to aggregate the multi-attribute to evaluate the importance of each node, which can comprehensively consider different centrality measures. To evaluate the performance, we used the SI model to estimate the spreading influence of nodes by different methods. The Kendall's tau coefficient τ between different methods and F(t) is calculated to demonstrate the effectiveness of the method. The experimental results on four real networks and an informative network show that our method can successfully identify the influential nodes in networks.

Acknowledgment

We greatly appreciate the Editor's encouragement and the anonymous reviewer's valuable comments and suggestions to improve this work. The work is partially supported by National High Technology Research and Development Program of China (863 Program)(Grant No. 2013AA013801), National Natural Science Foundation of China(Grant Nos. 61174022, 61573290, 61503237), China State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No. BUAA-VR-14KF-02).

Reference

[1] D.-H. Kim, J. D. Noh, H. Jeong, Scale-free trees: The skeletons of complex networks, Physical Review E 70 (4) (2004) 046126.

[2] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Reviews of modern physics 74 (1) (2002) 47.

[3] H. Zhang, D. Wei, Y. Hu, X. Lan, Y. Deng, Modeling the self-similarity in complex networks based on coulombs law, Communications in Nonlinear Science and Numerical Simulation 35 (2016) 97–104.

[4] A. P. Masucci, K. Stanilov, M. Batty, Limited urban growth: London's street network dynamics since the 18th century, PLoS One 8 (8) (2013) e69469.

[5] D. R. Amancio, O. N. Oliveira, L. da Fontoura Costa, Three-feature model to reproduce the topology of citation networks and the effects from authors visibility on their h-index, Journal of informetrics 6 (3) (2012) 427–434.

[6] A. P. Masucci, A. Kalampokis, V. M. Eguíluz, E. Hernández-García, Wikipedia information flow analysis reveals the scale-free architecture of the semantic space, PloS one 6 (2) (2011) e17333.

[7] H. Jeong, S. P. Mason, A.-L. Barabási, Z. N. Oltvai, Lethality and centrality in protein networks, Nature 411 (6833) (2001) 41–42.

[8] H. Yu, P. M. Kim, E. Sprecher, V. Trifonov, M. Gerstein, The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics, PLoS Comput Biol 3 (4) (2007) e59.

[9] X. Zhang, J. Zhu, Q. Wang, H. Zhao, Identifying influential nodes in complex networks with community structure, Knowledge-Based Systems 42 (2013) 74–84.

[10] S. Wang, Y. Du, Y. Deng, A new measure of identifying influential nodes: Efficiency centrality, Communications in Nonlinear Science and Numerical Simulation 47 (2017) 151–163.

[11] Y. A. Kim, H. S. Song, Strategies for predicting local trust based on trust propagation in social networks, Knowledge-Based Systems 24 (8) (2011) 1360–1371.

[12] H. Mo, C. Gao, Y. Deng, Evidential method to identify influential nodes in complex networks, Journal of Systems Engineering and Electronics 26 (2) (2015) 381–387.

[13] A.-L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nature Reviews Genetics 12 (1) (2011) 56–68.

[14] R. Yang, B.-H. Wang, J. Ren, W.-J. Bai, Z.-W. Shi, W.-X. Wang, T. Zhou, Epidemic spreading on heterogeneous networks with identical infectivity, Physics Letters A 364 (3) (2007) 189–193.

[15] J. Borge-Holthoefer, Y. Moreno, Absence of influential spreaders in rumor dynamics, Physical Review E 85 (2) (2012) 026116.

[16] S. Suo, Y. Chen, The dynamics of public opinion in complex networks, Journal of Artificial Societies and Social Simulation 11 (4) (2008) 2.

[17] Z. Wang, C.-Y. Xia, S. Meloni, C.-S. Zhou, Y. Moreno, Impact of social punishment on cooperative behavior in complex networks, Scientific reports 3.

[18] C. Gao, X. Lan, X. Zhang, Y. Deng, A bio-inspired methodology of identifying influential nodes in complex networks, PloS one 8 (6) (2013) e66732.

[19] S. H. Strogatz, Exploring complex networks, Nature 410 (6825) (2001) 268–276.

[20] M. E. Newman, The structure and function of complex networks, SIAM review 45 (2) (2003) 167–256.

[21] L. Lü, T. Zhou, Link prediction in complex networks: A survey, Physica A: Statistical Mechanics and its Applications 390 (6) (2011) 1150–1170.

[22] Z. Cheng, H.-T. Zhang, M. Z. Chen, T. Zhou, N. V. Valeyev, Aggregation pattern transitions by slightly varying the attractive/repulsive function, PloS one 6 (7) (2011) e22123.

[23] E. E. Schadt, Molecular networks as sensors and drivers of common human diseases, Nature 461 (7261) (2009) 218–223.

[24] D. Amancio, M. Nunes, O. Oliveira, T. Pardo, L. Antiqueira, L. d. F. Costa, Using metrics from complex networks to evaluate machine translation, Physica A: Statistical Mechanics and its Applications 390 (1) (2011) 131–142.

[25] M. Yang, G. Chen, X. Fu, A modified sis model with an infective medium on complex networks and its global stability, Physica A: Statistical Mechanics and its Applications 390 (12) (2011) 2408–2413.

[26] B. Hou, Y. Yao, D. Liao, Identifying all-around nodes for spreading dynamics in complex networks, Physica A: Statistical Mechanics and its Applications 391 (15) (2012) 4012–4017.

[27] M. Jalili, Social power and opinion formation in complex networks, Physica A: Statistical Mechanics and its Applications 392 (4) (2013) 959–966.

[28] L. C. Freeman, Centrality in social networks conceptual clarification, Social networks 1 (3) (1978) 215–239.

[29] U. Brandes, A faster algorithm for betweenness centrality*, Journal of mathematical sociology 25 (2) (2001) 163–177.

[30] M. E. Newman, A measure of betweenness centrality based on random walks, Social networks 27 (1) (2005) 39–54.

[31] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, T. Zhou, Identifying influ-

ential nodes in complex networks, Physica a: Statistical mechanics and its applications 391 (4) (2012) 1777–1787.

[32] P. Bonacich, P. Lloyd, Eigenvector-like measures of centrality for asymmetric relations, Social networks 23 (3) (2001) 191–201.

[33] U. Sehgal, K. Kaur, P. Kumar, The anatomy of a large-scale hyper textual web search engine, in: 2009 Second International Conference on Computer and Electrical Engineering, IEEE, 2009, pp. 491–495.

[34] L. Lü, Y.-C. Zhang, C. H. Yeung, T. Zhou, Leaders in social networks, the delicious case, PloS one 6 (6) (2011) e21202.

[35] 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: Statistical Mechanics and its Applications 403 (2014) 130–147.

[36] T. Gwo-Hshiung, Multiple attribute decision making: methods and applications, Multiple Attribute Decision Making: Methods and Applications.

[37] M. Behzadian, S. K. Otaghsara, M. Yazdani, J. Ignatius, A state-of the-art survey of topsis applications, Expert Systems with Applications 39 (17) (2012) 13051–13069.

[38] X. Zhang, F. T. Chan, A. Adamatzky, S. Mahadevan, H. Yang, Z. Zhang, Y. Deng, An intelligent physarum solver for supply chain

network design under profit maximization and oligopolistic competition, International Journal of Production Research 55 (1) (2017) 244–263.

[39] X. Zhou, Y. Shi, X. Deng, Y. Deng, D-DEMATEL: A new method to identify critical success factors in emergency management, Safety Science 91 (2017) 93–104.

[40] X. Guo, Y. Zhou, J. Qian, Y. Deng, Using evidence credibility decay model for dependence assessment in human reliability analysis, Annals of Nuclear Energy 100 (2017) 107–118.

[41] S. Galam, J.-D. Zucker, From individual choice to group decision-making, Physica A: Statistical Mechanics and its Applications 287 (3) (2000) 644–659.

[42] X. Zhang, Y. Deng, F. T. Chan, P. Xu, S. Mahadevan, Y. Hu, Ifsjsp: a novel methodology for the job-shop scheduling problem based on intuitionistic fuzzy sets, International Journal of Production Research 51 (17) (2013) 5100–5119.

[43] R. Zhang, X. Ran, C. Wang, Y. Deng, Fuzzy Evaluation of Network Vulnerability, Quality and Reliability Engineering International 32 (5) (2016) 1715–1730.

[44] J. Wang, Y. Hu, F. Xiao, X. Deng, Y. Deng, A novel method to use fuzzy soft sets in decision making based on ambiguity measure and

Dempster-Shafer theory of evidence: An application in medical diagnosis, Artificial intelligence in medicine 69 (2016) 1–11.

[45] Y. Du, X. Lu, X. Su, Y. Hu, Y. Deng, New Failure Mode and Effects Analysis: An Evidential Downscaling Method, Quality and Reliability Engineering International 32 (2) (2016) 737–746.

[46] K. Yuan, F. Xiao, L. Fei, B. Kang, Y. Deng, Modeling Sensor Reliability in Fault Diagnosis Based on Evidence Theory, SENSORS 16 (1). doi:10.3390/s16010113.

[47] H. Mo, Y. Deng, A new aggregating operator in linguistic decision making based on d numbers, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 24 (6) (2016) 831–846.

[48] B. Kang, Y. Hu, Y. Deng, D. Zhou, A New Methodology of Multicriteria Decision-Making in Supplier Selection Based on Z-Numbers, Mathematical Problems in Engineeringdoi:10.1155/2016/8475987.

[49] 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: Statistical Mechanics and its Applications 399 (2014) 57–69.

[50] J. Hu, Y. Du, H. Mo, D. Wei, Y. Deng, A modified weighted topsis to identify influential nodes in complex networks, Physica A: Statistical Mechanics and its Applications 444 (2016) 73–85.

[51] T. L. Saaty, The analytic hierarchy process: planning, priority setting, resources allocation, New York: McGraw.

[52] T. L. Saaty, Fundamentals of decision making and priority theory with the analytic hierarchy process, Vol. 6, Rws Publications, 2000.

[53] X. Zhang, Y. Deng, F. T. Chan, S. Mahadevan, A fuzzy extended analytic network process-based approach for global supplier selection, Applied Intelligence 43 (4) (2015) 760–772.

[54] X. Zhang, Y. Deng, F. T. S. Chan, A. Adamatzky, S. Mahadevan, Supplier selection based on evidence theory and analytic network process, Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 230 (3) (2016) 562–573. doi:10.1177/0954405414551105.

[55] X. Zhou, X. Deng, Y. Deng, S. Mahadevan, Dependence assessment in human reliability analysis based on d numbers and ahp, Nuclear Engineering and Design 313 (2017) 243–252.

[56] J. Nixon, P. Dey, P. Davies, Which is the best solar thermal collection technology for electricity generation in north-west india? evaluation of options using the analytical hierarchy process, Energy 35 (12) (2010) 5230–5240.

[57] T. Kurka, Application of the analytic hierarchy process to evaluate the regional sustainability of bioenergy developments, Energy 62 (2013) 393–402.

[58] T. Zhou, J.-G. Liu, W.-J. Bai, G. Chen, B.-H. Wang, Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity, Physical Review E 74 (5) (2006) 056109.

[59] S. Brin, L. Page, Reprint of: The anatomy of a large-scale hypertextual web search engine, Computer networks 56 (18) (2012) 3825–3833.

[60] A. N. Langville, C. D. Meyer, Google's PageRank and beyond: The science of search engine rankings, Princeton University Press, 2011.

[61] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, T. Zhou, Vital nodes identification in complex networks, Physics Reports 650 (2016) 1–63.

[62] T. L. Saaty, The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making, in: Multiple criteria decision analysis: state of the art surveys, Springer, 2005, pp. 345–405.

[63] P. Aragonés-Beltrán, F. Chaparro-González, J.-P. Pastor-Ferrando, A. Pla-Rubio, An ahp (analytic hierarchy process)/anp (analytic network process)-based multi-criteria decision approach for the selection of solar-thermal power plant investment projects, Energy 66 (2014) 222–238.

[64] T. Opsahl, F. Agneessens, J. Skvoretz, Node centrality in weighted networks: Generalizing degree and shortest paths, Social Networks 32 (3) (2010) 245–251.

[65] D. Krackhardt, Assessing the political landscape: Structure, cognition, and power in organizations, Administrative science quarterly (1990) 342–369.

[66] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, A. Arenas, Self-similar community structure in a network of human interactions, Physical review E 68 (6) (2003) 065103.

[67] M. Kaiser, C. C. Hilgetag, Spatial growth of real-world networks, Physical Review E 69 (3) (2004) 036103.

[68] D. J. Watts, S. H. Strogatz, Collective dynamics of small-worldnetworks, nature 393 (6684) (1998) 440–442.

[69] M. G. Kendall, The treatment of ties in ranking problems, Biometrika 33 (3) (1945) 239–251.

[70] W. Hoeffding, A non-parametric test of independence, The annals of mathematical statistics (1948) 546–557.

[71] 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, Springer, 2007, pp. 149–172.

[72] D. R. Amancio, Authorship recognition via fluctuation analysis of network topology and word intermittency, Journal of Statistical Mechanics: Theory and Experiment 2015 (3) (2015) P03005.

[73] L. Sheng, C. Li, English and chinese languages as weighted complex networks, Physica A: Statistical Mechanics and its Applications 388 (12) (2009) 2561–2570.

[74] D. R. Amancio, O. N. Oliveira Jr, L. d. F. Costa, Structure–semantics interplay in complex networks and its effects on the predictability of similarity in texts, Physica A: Statistical Mechanics and its Applications 391 (18) (2012) 4406–4419.

[75] D. R. Amancio, S. M. Aluisio, O. N. Oliveira Jr, L. d. F. Costa, Complex networks analysis of language complexity, EPL (Europhysics Letters) 100 (5) (2012) 58002.

[76] D. R. Amancio, E. G. Altmann, D. Rybski, O. N. Oliveira Jr, L. d. F. Costa, Probing the statistical properties of unknown texts: application to the voynich manuscript, PLoS One 8 (7) (2013) e67310.

[77] D. R. Amancio, E. G. Altmann, O. N. Oliveira Jr, L. da Fontoura Costa, Comparing intermittency and network measurements of words and their dependence on authorship, New Journal of Physics 13 (12) (2011) 123024.

[78] C. Carretero-Campos, P. Bernaola-Galván, A. Coronado, P. Carpena, Improving statistical keyword detection in short texts: Entropic and clustering approaches, Physica A: Statistical Mechanics and its Applications 392 (6) (2013) 1481–1492.

[79] M. E. Newman, Finding community structure in networks using the eigenvectors of matrices, Physical review E 74 (3) (2006) 036104.

[80] J. Herrera, P. Pury, Statistical keyword detection in literary corpora, The European Physical Journal B 63 (1) (2008) 135.

[81] M. Ortuño, P. Carpena, P. Bernaola-Galvn, E. Muñoz, A. Somoza, Keyword detection in natural languages and dna, EPL (Europhysics Letters) 57 (5) (2002) 759–764.

[82] P. Carpena, P. Bernaola-Galvan, M. Hackenberg, A. Coronado, J. Oliver, Level statistics of words: finding keywords in literary texts and symbolic sequences, Physical Review E 79 (3) (2009) 035102.