基于相对熵的节点影响力测量方法

陈俊华 边宅安 李慧嘉 关闰丹
(中央财经大学管理科学与工程学院 北京 100081)

摘 要 识别中心节点是复杂网络分析的一个关键问题。文中结合现有的中心性测度方法,提出了一种利用 TOPSIS 法的相对熵,以此识别网络中的中心节点。现有的中心性测度方法可以被看作在复杂网络中确定各节点属性的排名。因此,提出的方法可以利用各种中心性测度方法的优点,获得一个更优的排名结果。最后用数值实验验证了所提方法的有效性。

关键词 复杂网络,中心节点,中心性测度,相对熵,TOPSIS 法
中图法分类号 TP393 **文献标识码** A

Measuring Method of Node Influence Based on Relative Entropy

CEHN Jun-hua BIAN Zhai-an LI Hui-jia GUAN Run-dan
(School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100081, China)

Abstract Recognizing central nodes is a key problem in complex network analysis, this paper proposed a method of relative entropy using TOPSIS (Technique for Order Performance by Similarity to Ideal Solution) method to identify the influential node in the network. The existing central measure methods can be considered as determining the rank of each node attribute in a complex network. Therefore, the method proposed in this paper can use the advantages of various central measure methods to obtain a better ranking result. Finally, the validity of the proposed method was verified by numerical experiments.

Keywords Complex networks, Influential nodes, Centrality measure, Relative entropy, TOPSIS method

1 引言

近年来,复杂网络理论在许多领域变得越来越重要[1-8],如自然科学、经济学、管理学、生物科学和计算机科学与技术[9-14]。同时,如级联机制[15-16]、传播效应[17]和同步效应等一系列机制和效应,会因为一小部分重要节点的变化而受到较大影响[18-19]。评估节点的影响力(重要性)在网络分析中具有非常重要的意义[20-21],如对疾病传播的控制[22-23]、对民意的研究和谣言传播等[24]。

1.1 经典中心性测度方法

中心节点在网络结构和功能方面起着关键作用。虽然重要的中心节点在网络中并不多,但它们的影响很快就会蔓延到整个网络[25]。因此,研究人员提出了许多方法来评估复杂网络中的中心节点,如节点度(Degree Centrality, DC)、介数中心性(Betweeness Centrality, BC)、接近中心性(Closeness Centrality, CC)[26]、特征向量中心(EC)[27]、pagerank 指标(pr)[28],以及许多其他方法[29-35]。这些中心性测度方法由研究人员开发,并适用于许多实际工作。文献[36]提出了一个基于生物启发的中心性测度模型来寻找网络中的重要节点。Bonacich 等[27]开发了特征向量中心性测度方法,旨在处理非对称图的情况。文献[37]扩展了不同维度的介数中心性测度方法。现有的中心性测度方法都分别从各个方面进行了改进,然而单一的中心性测度方法在识别有影响的节点时,仍存在一些不足。

1.2 现有方法的不足

中心性测度方法被广泛地用于复杂网络中的中心节点识别。然而,不可否认的是,它们存在一些局限性。例如,DC 虽然具有低复杂性和高实用性的优点,但它只关注节点的局部信息,而忽略了全局信息。而对于 CC,如果两个节点有不同的组件,它将获得不合理的结果,因此它的优势在包含断开组件的网络中并不能得到很好的体现。BC 虽然弥补了 DC 和 CC 的缺点,但它仍然存在一些问题。例如,若许多节点未包括在其他节点的最短路径中,那么 BC 的结果将是 0[38]。通过以上分析发现,当用不同的中心性测度方法来识别中心节点时,会产生不同的结果。图 1 给出了一个人际关系网络的例子[39]。

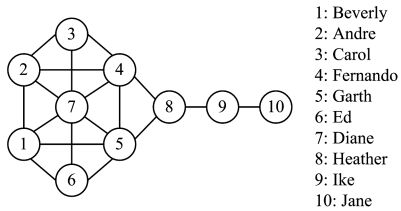


图 1 经典的人际关系网络

每个节点的 DC,CC 和 BC 如表 1 所列。可以发现,如果

陈俊华(1975-),男,博士,副教授,主要研究方向为社会网络;**边宅安**(1995-),男,硕士生,主要研究方向为复杂网络;**李慧嘉**(1985-),男,博士,副教授,主要研究方向为复杂网络、数据挖掘、运筹学,E-mail:Hjli@amss.ac.cn;**关闰丹**(1997-),女,硕士生,主要研究方向为复杂网络。

使用 DC，那么最有影响力的节点是 Diane，但如果使用 BC，Heather 有最大的介数中心性，虽然其标准化后的度中心性只有 0.333。此外，如果使用 CC，Fernando 和 Garth 的节点是最重要的。因此，在这种网络中将从不同的中心性测度方法中得到不同的结果。

表 1　经典人际网络中各节点的 DC，CC，BC 及排名

Degree(DC)		Closeness(CC)		Betweenness(BC)	
Nodes	Value	Nodes	Value	Nodes	Value
Diane	6	Fernando	0.0714	Heather	0.5778
Fernando	5	Garth	0.07114	Fernando	0.3556
Garth	5	Diane	0.0667	Ike	0.3333
Andre	4	Heather	0.0667	Garth	0.2889
Beverly	4	Andre	0.0588	Diane	0.0444
Carol	3	Beverly	0.0588	Andre	0.0370
Ed	3	Carol	0.0556	Beverly	0.0370
Heather	3	Ed	0.0556	Jane	0.0000
Ike	2	Ike	0.0486	Ed	0.0000
Jane	1	Jane	0.0345	Carol	0.0000

本文提出了一种利用 TOPSIS 法中的相对熵来确定有中心节点的新方法。本文第 2 节简要介绍了图形的基本特征和一些相关概念；第 3 节提出了利用 TOPSIS 法中的相对熵来确定有中心节点的新方法；第 4 节使用多个实际网络研究了方法的有效性；最后总结全文。

2　理论方法

2.1　经典中心性测度方法

图形可以被表示为 $G=(V,E)$，其中 V 和 E 分别是节点集和边集。在加权和非加权网络中，每对节点之间的距离 $d(i,j)$ 的定义是不同的。在非加权网络中，最短路径被定义为最小的中间节点数。DC，BC 和 CC 的中心性测度方法的定义如下。

定义 1(DC)　节点 i 的 DC 表示为 k_i，定义为[40]：

$$k_i=\sum_j^N x_{ij} \tag{1}$$

其中，i 是中心节点，j 表示连接到 i 的其他节点，x_{ij} 表示节点 i 和节点 j 之间的边。如果节点 i 连接到节点 j，则 x_{ij} 的值为 1，否则为 0。

定义 2(BC)　节点 i 的 BC 表示为 b_i，定义为[26]：

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

其中，g_{jk} 是节点 j 和节点 k 之间的最短路径数，是通过节点 i 的路径的数目。

定义 3(CC)　节点 i 的 CC 表示为 c_i，定义为[26]：

$$c_i=\frac{1}{\sum_j^N d_{ij}} \tag{3}$$

其中，d_{ij} 表示节点 i 和节点 j 之间的距离。

上述措施已扩展到加权网络，在加权网络中两个节点之间的距离表示为节点之间的权重之和。

定义 4(一个加权网络中的 DC)　节点 i 的 DC 表示为 k_i^w，定义为[38]：

$$k_i^w=\sum_j^N w_{ij} \tag{4}$$

其中，w_{ij} 是节点 i 和 j 之间的边的加权，当节点 i 连接到节点 j 时大于 0。

定义 5(一个加权网络中的 BC)　节点 i 的 BC 表示为 b_i^w，定义为[38]：

$$b_i^w=\sum_{j,k\neq i}\frac{g_{jk}^w(i)}{g_{jk}^w} \tag{5}$$

其中，g_{jk}^w 是节点 j 和节点 k 之间最短路径数的二进制，$g_{jk}^w(i)$ 是 g_{jk}^w 中通过节点 i 的路径的数目。

定义 6(一个加权网络中的 CC)　节点 i 的 CC 表示为 c_i^w，定义为[38]：

$$c_i^w=[\sum_j^N d^w(i,j)]^{-1} \tag{6}$$

其中，$d^w(i,j)$ 是节点 i 和 j 之间的最短距离。

2.2　相对熵

相对熵（又称 KL 散度）是概率论和信息论中的一个基本概念，由 Kullback 等[41]提出。相对熵是两个概率分布之间差值的非对称测度。两个概率分布 P 和 Q 的相对熵的定义如下：

$$D_{KL}(P\parallel Q)=\sum_{i=1}^n P(i)\log\frac{P(i)}{Q(i)} \tag{7}$$

其中，P 和 Q 有相同数量的组件，这两个概率中的组件数为 n。

系统 A 和 B 的状态 A_i 和 $B_i(i=1,2,\cdots,N)$ 的差异度可以由信息论中的相对熵测度。

$$C=\sum_{i=1}^N\{A_i\log\frac{A_i}{B_i}+(1-A_i)\log\frac{1-A_i}{1-B_i}\} \tag{8}$$

其中，C 被称为两个系统的相对熵。由于相对熵是不对称的，因此它不是两个系统的实际距离，但把它作为两个系统的距离却具有实际意义。

2.3　TOPSIS 法

TOPSIS 法是由 Huang 等于 1981 年首次提出的[43]一种多属性决策方法。它试图选择与最优解有着最短距离，同时与最劣解有着最远距离的集合。最优解方案最大化了效益标准，并最大限度地降低了成本标准，而最劣解方案最大化了成本标准，并最大限度地减少了利益标准[32]。

假设有一个决策矩阵 $D=\{x_{ij}\}_{m\times n}$，将决策矩阵标准化为：

$$r_{ij}=\frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}},i\in 1,2,\cdots,m,j\in 1,2,\cdots,n \tag{9}$$

其中，x_{ij} 表示第 i 个节点的第 j 个属性的标准化值。

将标准化的决策矩阵的列乘以相关权重以获得加权决策矩阵：

$$v_{ij}=w_{ij}\times r_{ij},i\in 1,2,\cdots,m,j\in 1,2,\cdots,n \tag{10}$$

其中，w_{ij} 是第 j 个属性的权重。

标准化加权矩阵的每个属性，为相对熵的计算做准备。

$$u_{ij}=\frac{x_{ij}}{\sum_{i=1}^m x_{ij}},i\in 1,2,\cdots,m,j\in 1,2,\cdots,n \tag{11}$$

其中，x_{ij} 表示第 i 个节点的第 j 个属性的标准化值。

确定最优解方案和最劣解方案，最优解方案表示为 A^+，最劣解方案表示为 A^-，定义如下：

$$\begin{aligned}A^+&=\{v_1^+,v_2^+,\cdots,v_n^+\}\\&=\{(\max_i v_{ij}\mid j\in T_1)(\min_i v_{ij}\mid j\in T_2)\}\end{aligned} \tag{12}$$

$$A^-=\{v_1^-,v_2^-,\cdots,v_n^-\}$$
$$=\{(\max_i v_{ij}\mid j\in T_2)(\min_i v_{ij}\mid j\in T_1)\} \quad (13)$$

其中，T_1 和 T_2 分别为福利标准和成本标准。

用最优解和最劣解选择的集合计算相对熵，分别表示为 S_i^+ 和 S_-^-：

$$S_i^+=\sum_{i=1}^{N}\{v_i^+\log\frac{v_i^+}{v_i}+(1-v_i^+)\log\frac{(1-v_i^+)}{1-v_i}\},i\in1,2,\cdots,m,j\in1,2,\cdots,n \quad (14)$$

$$S_i^-=\sum_{i=1}^{N}\{v_i^-\log\frac{v_i^-}{v_i}+(1-v_i^-)\log\frac{(1-v_i^-)}{1-v_i}\},i\in1,2,\cdots,m,j\in1,2,\cdots,n \quad (15)$$

计算与理想解的相对贴近度：

$$C_i=\frac{S_i^-}{S_i^++S_i^-},i\in1,2,\cdots,m \quad (16)$$

在复杂网络中使用相对熵来识别中心节点时有一个关键问题需要解决：当分母为 0 时，分数值趋于无穷大，其对数值也趋于无穷大。如果出现这种情况，本文方法将失去意义，因此用非常小的数字 1×10^{-12} 在上述现象发生时替换 0 值[44]。大量的实验证明，这不会影响计算结果。

2.4 示例说明

本文构建了一个小的网络，以展示本文方法的具体步骤，如图 2 所示。

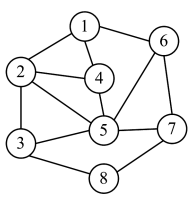


图 2 示例网络

首先计算每个节点的 BC，CC 和 DC，结果如表 2 所列。

表 2 示例网络的 BC，CC 和 DC 值

node	d	c	b
1	3	0.0833	0.0714
2	4	0.1000	0.1905
3	3	0.0909	0.1667
4	3	0.0833	0
5	5	0.1111	0.2500
6	3	0.0909	0.0714
7	3	0.0909	0.0714
8	2	0.0714	0

其次，得到的替代矩阵如下：

$$D=\begin{bmatrix}3&0.0833&0.0714\\4&0.1000&0.1905\\3&0.0909&0.1667\\3&0.0833&0\\5&0.1111&0.2500\\3&0.0909&0.0714\\3&0.0909&0.0714\\2&0.0714&0\end{bmatrix}$$

再次，通过式(9)将矩阵 D 标准化，对其中一个元素的标准化方式如下：

$$r_{11}=\frac{x_{11}}{\sqrt{\sum_{1=1}^{8}x_{i1}^2}}$$
$$=\frac{3}{\sqrt{3^2+4^2+3^2+3^2+5^2+3^2+3^2+2^2}}$$
$$=0.3162$$

其余元素的标准化可以用同样的方式完成，从而得到标准化的决策矩阵如下：

$$R=\begin{bmatrix}0.3162&0.3241&0.1896\\0.4216&0.3889&0.5057\\0.3162&0.3535&0.4425\\0.3162&0.3241&0\\0.5270&0.4321&0.6637\\0.3162&0.3535&0.1896\\0.3162&0.3535&0.1896\\0.2108&0.2778&0\end{bmatrix}$$

我们假设每个属性的权重相同，为 1/3。加权决策矩阵如表 3 所列，标准化的决策矩阵如表 4 所列。

表 3 加权后的数据

node	d	c	b
1	0.1054	0.1080	0.0632
2	0.1405	0.1296	0.1686
3	0.1054	0.1178	0.1475
4	0.1054	0.1080	0
5	0.1757	0.1440	0.2212
6	0.1054	0.1178	0.0632
7	0.1054	0.1178	0.0632
8	0.0703	0.0926	0

表 4 标准化后的数据

node	d	c	b
1	0.1154	0.1154	0.0870
2	0.1538	0.1385	0.2319
3	0.1154	0.1259	0.2029
4	0.1154	0.1154	1×10^{-12}
5	0.1923	0.1539	0.3043
6	0.1154	0.1259	0.0870
7	0.1154	0.1259	0.0870
8	0.0769	0.0989	1×10^{-12}

然后，用式(12)和式(13)计算最优解方案和最劣解方案：

$A^+=\{0.1923,0.1539,0.3043\}$

$A^-=\{0.0769,0.0989,1\times10^{-12}\}$

使用相对熵计算第 1 个节点的 S^+ 和 S^-：

$$S^+=0.1923\times\log\frac{0.1923}{0.1154}+(1-0.1923)\times\log\frac{1-0.1923}{1-0.1154}+0.1539\times\log\frac{0.1539}{0.1154}+(1-0.1539)\times\log\frac{1-0.1539}{1-0.1154}+0.1539\times\log\frac{0.3043}{0.0807}+(1-0.3043)\times\log\frac{1-0.3043}{1-0.0807}$$
$$=0.7557$$

同理，$S^-=0.1449$。

最后，通过式(16)可以得到第 1 个节点的相对贴近度：

$$C_1=\frac{S_1^-}{S_1^++S_1^-}=\frac{0.1449}{0.7557+0.1449}=0.1609$$

重复以上过程，可以计算出每一个节点的相对贴近度，其结果以及节点的排名如表 5 所列。表 5 中，第 6 个节点和第 7

个节点有相同的排名，这是因为它们具有相同的 BC，CC 和 DC 的值。

表 5　节点的相对贴近度及排名

node	S_i^+	S_i^-	C_i	Ranking
1	0.7557	0.1449	0.1609	6
2	0.2047	0.4299	0.6774	2
3	0.3643	0.3439	0.4856	3
4	11.8155	0.0137	0.0012	7
5	0	0.6185	1	1
6	0.7363	0.1480	0.1674	4
7	0.7363	0.1480	0.1674	4
8	11.9623	0	0	8

3　提出的方法

为了进一步研究如何确定中心节点，目前已经提出了许多中心性测度方法，然而这些方法只在它们各自关注的方面优势明显，存在着局限性。为了解决上述问题，本文引入了多属性决策方法（TOPSIS 法）来识别中心节点，将 TOPSIS 法的思想与相对熵相结合，选择不同的中心性测度方法作为多属性。标准的 TOPSIS 法旨在找到距离最优解最近且距离最劣解最远的结果。它可以从不同的方面聚合这些属性从而做出决策或进行评估。本文方法可以结合现有中心测度的优点，为实际网络提供更好的排序结果，具体流程如图 3 所示。

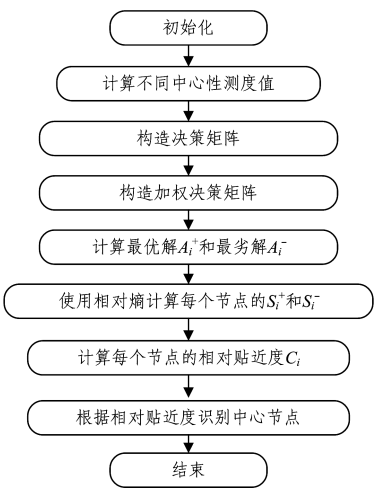


图 3　本文方法的流程图

本文方法的具体步骤如下。

步骤 1　计算不同中心性测度值。

根据网络中的信息传播方式，选取不同的中心性测度方法作为 TOPSIS 法的多属性。在某些运输网络中，若信息通过测地线路径（最短路径）进行传播，则需要选择度中心性（DC）测度方法、接近中心性（CC）测度方法和介数中心性（BC）测度方法。同样，若信息在一些不健全的网络中传播，则需要选择度中心性（DC）测度方法、接近中心性（CC）测度方法和特征向量中心性（EC）测度方法。

步骤 2　使用每个节点的相对熵计算 S_i^+ 和 S_i^-。

首先利用式（9），用不同中心性测度值构造决策矩阵；其次，用式（10）构造加权后的决策矩阵；然后，用式（12）和式（13）计算最优解 A_i^+ 和最劣解 A_i^-；最后，利用式（14）和式（15）得到 S_i^+ 和 S_i^-。

步骤 3　通过式（16）计算节点的相对贴近度 C_i。

步骤 4　根据 C_i 识别中心节点。

C_i 值越高，选择的方法越好，相应的节点影响力就越高。

4　实验分析

4.1　实验数据来源

本节引入 4 个不同的网络实验来说明本文方法的有效性：1）电子邮件（Email）网络[45]，此网络描述不同用户之间的邮件交换关系，节点表示用户，如果两个用户相互通信，则在相对应的节点之间产生一条相连接的边；2）美国航空（UsAir97）网络，这个网络记录了美国航空公司在 1997 年的状况，美国的机场代表节点，在机场之间有一个航空公司，代表节点之间的边，机场之间的坐标距离代表权重；3）空手道俱乐部（Zachary's Karate club）社交网络[46]，数据信息来自 20 世纪 70 年代美国大学空手道俱乐部 34 位成员之间的社交网络；4）《大卫·科波菲尔》书中名词和形容词的网络[47]，这是一个由英国作家查尔斯·狄更斯的小说《大卫·科波菲尔》（David Copperfield）中的普通名词和形容词组成的网络，一个名词或一个形容词代表一个节点，边连接在相邻位置出现的两个单词。4 个网络拓扑特征的基本统计数据如表 6 所列。

表 6　4 个实际网络拓扑特征的基本统计信息

Network	n	m	$\langle k\rangle$	$\langle d\rangle$	C
Karate club	34	78	4.5882	2.34	0.5706
Email	1133	10903	9.6240	3.71	0.2202
UsAir97	332	2126	12.8072	2.78	0.3126
David Copperfield	112	425	7.5893	2.98	0.1574

表 6 中，n 和 m 分别表示节点和边的数目，$\langle k\rangle$ 和 $\langle d\rangle$ 表示每个网络的平均度和平均最短距离，C 代表聚类系数。

4.2　易感染模型

在实验部分，本文利用易感染模型（SI 模型）评估网络节点的传输容量，从而反映出节点[48]的影响力和重要性。在疾病动态传播的模拟领域[34,49-50]中，SI 模型的研究越来越受关注。在 SI 模型中，每个节点都有两个离散状态：易感性状态和已感染状态。易感染节点有 β 的概率被已感染的节点感染。β 表示已感染的节点在疾病传输网络中的影响或感染范围。在加权网络中，节点 j 被节点 i 感染的概率可以表示为：$\beta=(\frac{w_{ij}}{w_M})^{\alpha}$，其中 w_{ij} 表示节点 i 和 j 的权重，w_M 表示整个网络中所有权重的最大值。在非加权网络中，概率 $\beta=(\frac{1}{2})^{\alpha}$。在下文的实验中，用 $I(t)$ 表示 t 时刻已感染节点的数量。

4.3　实验分析

实验 1　比较各中心性测度方法和所提方法影响力前 10 的节点

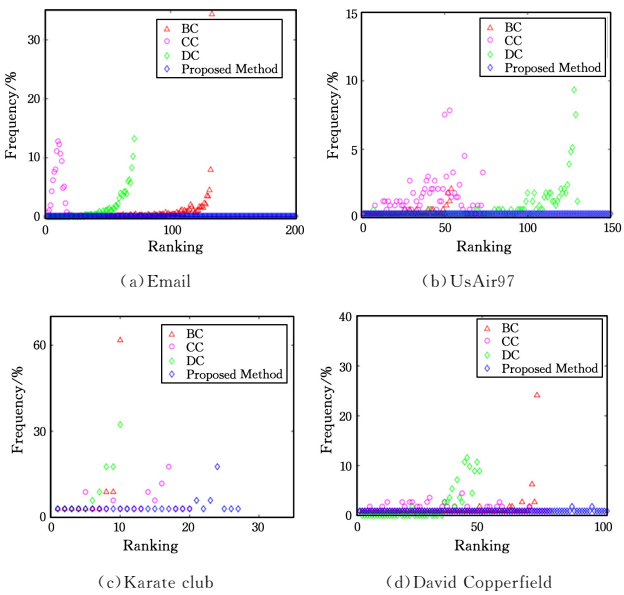
表 7 列出了不同的中心性测度方法中的前 10 个最具影响力的节点。在电子邮件网络中，用 DC，CC 和 BC 测度方法测出的前 10 个最具影响力的节点与所提出的方法前 10 个最具影响力的节点中，分别有 7 个、9 个和 9 个相同。在 UsAir97 网络中，分别有 7 个、7 个和 9 个相同。在空手道俱乐部成员关系网络中，分别有 8 个、7 个和 10 个相同。在《大卫·科波菲尔》的网络中，分别有 7 个、7 个和 10 个相同。实验表明，本文提出的方法与现有的中心性测度方法有高度的相似性。

表 7 中心性测度方法与所提方法测出的排名前 10 的节点

Rand	Email				UsAir97				Karate club				David Copperfield			
	DC	CC	BC	Proposed Method	DC	CC	BC	Proposed Method	DC	CC	BC	Proposed Method	DC	CC	BC	Proposed Method
1	105	333	23	23	118	118	118	118	34	1	1	1	18	18	18	18
2	333	23	105	105	261	261	8	261	1	3	3	34	3	3	3	3
3	42	105	333	333	248	67	261	8	33	34	34	3	52	52	44	44
4	23	42	76	42	67	255	47	47	3	32	33	33	44	44	9	52
5	16	41	42	76	255	201	201	201	2	33	32	32	105	105	52	9
6	41	76	578	41	47	182	67	67	32	14	6	2	9	28	80	80
7	196	233	135	135	166	47	313	313	4	9	2	6	51	9	1	1
8	233	52	41	578	201	248	13	182	24	20	28	24	28	27	29	105
9	76	135	52	52	182	166	182	152	14	2	24	28	25	25	28	28
10	21	378	355	233	147	112	152	255	9	4	9	9	26	26	105	29

实验 2 比较同一排名的节点的发生频率

毫无疑问，如果使用某些方法显示出相同的计算值，则在确定中心节点的过程中，节点排名将相同。因此，具有相同排名的节点数量可以被看作是评价排序方法有效性的指标。如果某一名次具有太多节点，那么用来排名的方法与标准就不能很好地识别中心节点。图 4 给出了 4 个不同网络中具有相同排名节点的出现频率。



(a) Email (b) UsAir97

(c) Karate club (d) David Copperfield

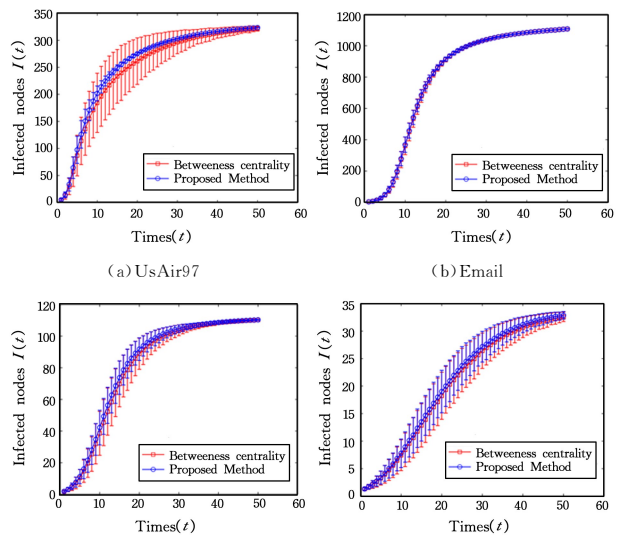
图 4 4 个实际网络中具有相同排名节点的发生频率

由图 4 可以明显看出，本文所提方法的测度结果中具有相同排名的节点是最少的，而其他中心性测度方法中具有相同排名节点数的多少程度不一。从这一角度可以得出结论：本文提出的方法在识别中心节点时比其他算法更有效。

实验 3 比较前 10 个节点的平均感染容量

众所周知，如果节点在网络中是重要的，那么它在 SI 模型中的传染能力会很强。因此，本文通过比较前 10 个节点的感染节点的平均值来评估排序方法的有效性。根据不同的感染时间，分别计算各种方法中受前 10 个节点感染节点数的平均值。然后将各网络的结果及其误差绘制出来，结果如图 5 所示。

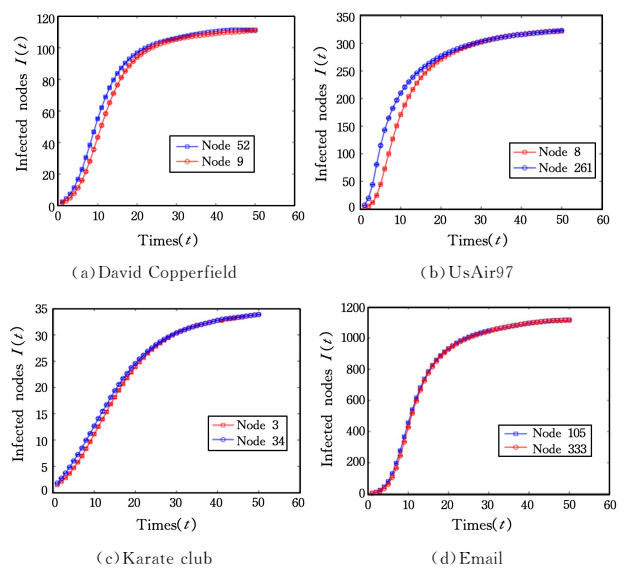
显然在 4 个不同网络中，本文提出的方法在传播率和已感染节点的数量方面均优于 BC 测度方法，与 DC 和 CC 测度方法的结果较相近，为了进一步比较 BC 测度方法和本文提出的方法，我们需要区分单个节点的传播能力，结果如图 6 所示。



(a) UsAir97 (b) Email

(c) David Copperfield (d) Karate club

图 5 前 10 个节点的感染节点的平均值及误差



(a) David Copperfield (b) UsAir97

(c) Karate club (d) Email

图 6 4 个网络中单个节点的传播能力

从图 6 可以发现，《大卫·科波菲尔》网络中的节点 52、空手道俱乐部社交网络中的节点 34、电子邮件网络中的节点 105 和 UsAir97 网络中的节点 261，具有相对较强的感染能力，都是本文提出的方法所确定的中心节点。

实验 4 比较各中心性测度方法与本文方法的相关性

利用各中心测度和所提方法计算每个网络的节点排名，然后分别得到本文方法与 DC、CC 和 BC 测度方法之间的相关性，相关关系如图 7—图 10 所示。

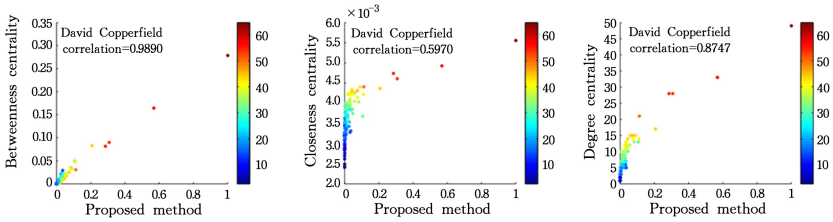


图 7　《大卫·科波菲尔》网络中 4 种方法的相关性

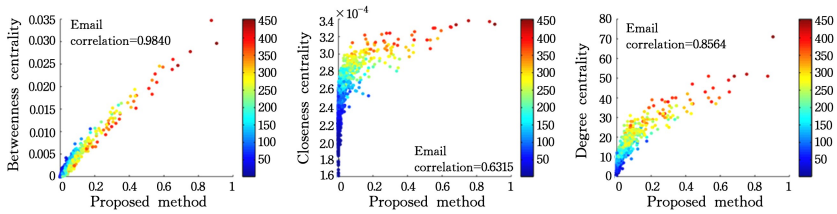


图 8　电子邮件网络中 4 种方法的相关性

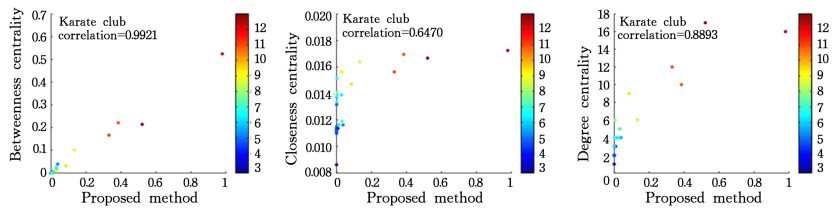


图 9　空手道俱乐部网络中 4 种方法的相关性

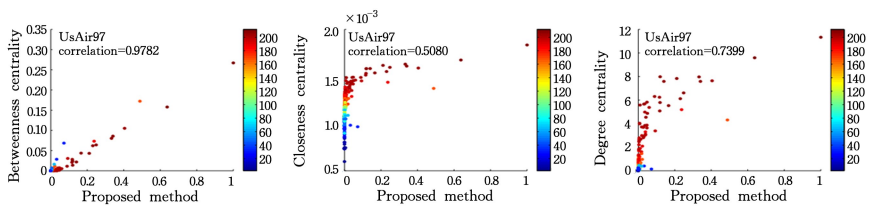


图 10　美国航空网络中 4 种方法的相关性

从图中可以发现，所提方法与 BC 测度方法和 DC 测度方法有较强的相关性，而与 CC 测度方法的相关程度较弱。在图 7—图 10 中，每个点表示网络中的一个节点，其颜色显示的是十倍感染影响。在 4 个网络中，3 种中心性测度方法都与所提方法呈正相关，尤其是 BC 和 DC 测度方法。此外，中心性测度值较大的节点具有较强的影响力。

总之，本文方法能够在不同的现实网络中与各中心性测度方法保持较稳定的相关性。

结束语　本文提出了一种新的基于相对熵的中心节点的识别方法。该方法将不同的中心性测度方法作为多属性，以确定实际网络中有影响的节点。此外，本文方法结合现有的中心性测度方法的优势，可以获得更好的排名结果。本文通过 4 个现实网络来说明该方法的有效性。实验结果表明，与其他中心性测度方法相比，本文方法具有优越性。

参 考 文 献

[1] TAN S，LÜ J. Characterizing the effect of population heterogeneity on evolutionary dynamics on complex networks[J]. Scientific Reports，2014，4(4)：5034.

[2] LU J，CHEN G. A time-varying complex dynamical network model and its controlled synchronization criteria[J]. IEEE Transactions on Automatic Control，2005，50(6)：841-846.

[3] LI X，YANG D，LIU X，et al. Bridging Time Series Dynamics and Complex Network Theory with Application to Electrocardiogram Analysis[J]. Circuits & Systems Magazine IEEE，2012，12(4)：33-46.

[4] LU W，LI X，RONG Z. Global stabilization of complex networks with digraph topologies via a local pinning algorithm[J]. Automatica，2010，46(1)：116-121.

[5] JIN Q，WANG L，XIA C Y，et al. Spontaneous Symmetry Breaking in Interdependent Networked Game[J]. Sci. Rep.，2012，4(6172)：4095.

[6] BOZZO E，FRANCESCHET M. Resistance distance，closeness，and betweenness[J]. Social Networks，2013，35(3)：460-469.

[7] WANG W，TANG M，YANG H，et al. Asymmetrically interacting spreading dynamics on complex layered networks[J]. Scientific Reports，2014，4(7502)：5097.

[8] WANG Z，KOKUBO S，JUSUP M，et al. Universal scaling for the dilemma strength in evolutionary games[J]. Physics of Life Reviews，2015，14：47-48.

[9] STROGATZ S H. Exploring Complex Networks. Nature 410，268[J]. Nature，2001，410(6825)：268-276.

[10] NEWMAN M E. The structure and function of complex networks[J]. SIAM Review，2003，45(2)：167-256.

[11] LÜ L，ZHOU T. Link prediction in complex networks：A survey[J]. Physica A Statistical Mechanics & Its Applications，2011，390(6)：1150-1170.

[12] SCHADT E E. Molecular networks as sensors and drivers of

common human diseases[J]. Nature,2009,461(7261):218.
[13] AMANCIO D R,NUNES M G V,JR O N O,et al. Using metrics from complex networks to evaluate machine translation[J]. Physica A Statistical Mechanics & Its Applications,2011,390(1):131-142.
[14] YANG M,CHEN G,FU X. A modified SIS model with an infective medium on complex networks and its global stability[J]. Physica A Statistical Mechanics & Its Applications,2011,390(12):2408-2413.
[15] MOTTER A E,LAI Y C. Cascade-based attacks on complex networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics,2002,66(2):065102.
[16] TAO Z,BING-HONG W. Catastrophes in scale-free networks [J]. Chinese Physics Letters,2005,22(5):1072-1075.
[17] PASTORSATORRAS R,VESPIGNANI A. Immunization of complex networks[J]. Phys Rev E Stat Nonlin Soft Matter Phys,2002,65(3 Pt 2A):036104.
[18] ZHAO M,ZHOU T,WANG B H,et al. Enhanced synchronizability by structural perturbations[J]. Phys. Rev. E,2005,72(5 Pt 2):057102.
[19] ZEMANOVÁ L,ZHOU C,KURTHS J. Structural and functional clusters of complex brain networks[J]. Physica D Nonlinear Phenomena,2006,224(1):202-212.
[20] CHEN D B,XIAO R,ZENG A,et al. Path diversity improvesthe identification of influential spreaders[J]. Europhysics Letters,2013,104(6):68006.
[21] MOTTER A E,ZHOU C,KURTHS J. Enhancing complex-network synchronization[J]. Europhysics Letters,2005,69(3):334.
[22] BARABÁSI A L,GULBAHCE N,LOSCALZO J. Network medicine: a network-based approach to human disease[J]. Nature Reviews Genetics,2011,12(1):56.
[23] YANG R,WANG B H,REN J,et al. Epidemic spreading on heterogeneous networks with identical infectivity[J]. Physics Letters A,2007,364(3):189-193.
[24] BORGE-HOLTHOEFER J,MORENO Y. Absence of influential spreaders in rumor dynamics[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics,2012,85(2 Pt 2):026116.
[25] GLATTFELDER J B. The Network of Global Corporate Control[J]. Business & Politics,2011,15(3):357-379.
[26] FREEMAN L C. Centrality in social networks conceptual clarification[J]. Social Networks,1978,1(3):215-239.
[27] BONACICH P,LLOYD P. Eigenvector-like measures of centrality for asymmetric relations[J]. Social Networks,2001,23(3):191-201.
[28] BRIN S,PAGE L. Reprint of: The anatomy of a large-scale hypertextual web search engine[J]. Computer Networks,2014,56(18):3825-3833.
[29] LÜ L,ZHANG Y C,CHI H Y,et al. Leaders in Social Networks,the Delicious Case[J]. Plos One,2011,6(6):e21202.
[30] KITSAK M,GALLOS L K,HAVLIN S,et al. Identification of influential spreaders in complex networks[J]. Nature Physics,2010,6(11):888-893.
[31] LIU J G,REN Z M,GUO Q. Ranking the spreading influence in complex networks[J]. Physica A Statistical Mechanics & Its Applications,2013,392(18):4154-4159.
[32] DU Y,GAO C,HU Y,et al. A new method of identifying influential nodes in complex networks based on TOPSIS[J]. Physica A:Statistical Mechanics and its Applications,2014,399:57-69.
[33] BAUER F,LIZIER J T. Identifying influential spreaders and efficiently estimating infection numbers in epidemic models: a walk counting approach[J]. Epl,2012,99(6):367-372.
[34] WEI D,DENG X,ZHANG X,et al. Identifying influential nodes in weighted networks based on evidence theory[J]. Physica A Statistical Mechanics & Its Applications,2013,392(10):2564-2575.
[35] GAO C,WEI D,HU Y,et al. A modified evidential methodology of identifying influential nodes in weighted networks[J]. Physica A Statistical Mechanics & Its Applications,2013,392(21):5490-5500.
[36] GAO C,LAN X,ZHANG X,et al. A Bio-Inspired Methodology of Identifying Influential Nodes in Complex Networks[J]. Plos One,2013,8(6):e66732.
[37] GÓMEZABC D. Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems [J]. European Journal of Operational Research,2013,226(2):354-365.
[38] OPSAHL T,AGNEESSENS F,SKVORETZ J. Node centrality in weighted networks: Generalizing degree and shortest paths [J]. Social Networks,2010,32(3):245-251.
[39] KRACKHARDT D. Assessing the Political Landscape: Structure,Cognition,and Power in Organizations[J]. Administrative Science Quarterly,1990,35(2):342-369.
[40] BONACICH P. Factoring and weighting approaches to status scores and clique identification[J]. Journal of Mathematical Sociology,1972,2(1):113-120.
[41] KULLBACK S,LEIBLER R A. On Information and Sufficiency [J]. Annals of Mathematical Statistics,1951,22(1):79-86.
[42] COVER T M,THOMAS J A. Elements of Information Theory Wiley[J]. New York,1991.
[43] HWANG C L,YOON K. Methods for Multiple Attribute Decision Making[M]// Multiple Attribute Decision Making. Springer Berlin Heidelberg,1981:58-191.
[44] GUO X Z,XIN X L. Partial Entropy and Relative Entropy of Fuzzy Sets[J]. Fuzzy Systems & Mathematics,2005(2):97-102.
[45] GUIMERÀ R,DANON L,DÀAZ-GUILERA A,et al. Self-similar community structure in a network of human interactions[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics,2003,68(6 Pt 2):065103.
[46] ZACHARY W W. An Information Flow Model for Conflict and Fission in Small Groups [J]. Journal of Anthropological Research,1977,33(4):452-473.
[47] NEWMAN M E J. Finding community strcuture in networks using the eigenvectors of matrics. Phys. Rev. E 74,036104[J]. Physical Review E,2006,74(3 Pt 2):036104.
[48] ZHOU T,LIU J G,BAI W J,et al. Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity [J]. Physical Review E Statistical Nonlinear & Soft Matter Physics,2006,74(2):056109.
[49] BAI W J,ZHOU T,WANG B H. Immunization of susceptible-infected model on scale-free networks[J]. Physica A Statistical Mechanics & Its Applications,2007,384(2):656-662.
[50] KARSAI M,KIVELÄ M,PAN R K,et al. Small but slow world: how network topology and burstiness slow down spreading[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics,2011,83(2):025102.