

基于主成分分析的网络节点重要性指标贡献评价

胡钢¹,徐翔²,张维明²,周璠²

(1. 安徽工业大学管理科学与工程学院, 安徽马鞍山 243032; 2. 国防科技大学信息系统工程重点实验室, 湖南长沙 410073)

摘要: 为研究不同网络节点重要性指标对网络中重要节点的影响程度,进而优选出较能体现网络重要节点性质的指标. 本文基于主成分分析(Principal Component Analysis, 简记 PCA), 选取七个节点重要性指标对网络重要性节点贡献率进行计算分析, 同时选取了七种不同的网络进行实验, 得到指标贡献率大小顺序, 利用肯德尔系数对重要指标与其余指标进行相关性分析, 得到不同指标之间的相关系数及相关系数大小的影响因素. 本文为研究网络重要节点选择指标提供了一种思路, 同时为研究不同节点间的相互关系提供了研究方法.

关键词: 主成分分析; 节点重要性; 贡献率; 肯德尔系数

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112(2019)02-0358-08

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2019.02.015

Contribution Analysis for Assessing Node Importance Indices with Principal Component Analysis

HU Gang¹, XU Xiang², ZHANG Wei-ming², ZHOU Yun²

(1. School of Management Science and Engineering, Anhui University of Technology, Maanshan, Anhui 243032, China;

2. Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha, Hunan 410073, China)

Abstract: In network theory, it is interest to study the influences of different nodes on the key nodes in the network, and build or select the proper node importance index to model it. This paper selects seven node importance indices to calculate and analyze their contributions in nodes' importance evaluation with Principal Component Analysis. Seven empirical networks are used for experiments. Moreover, the order of different contributions of indices is obtained, and the correlation analysis between the most important index and the other indices is carried out using the Kendall coefficient, and factors affecting the correlation coefficient are also discussed. This paper provides a way to select the node importance index in the network, and the results could also be used for studying the relationships between different nodes.

Key words: principal component analysis; node importance index; contribution analysis; Kendall coefficient

1 引言

随着复杂网络研究的兴起,网络中重要节点的识别也成为研究热点.从不同的研究角度,节点识别的方法也不相同,例如:基于节点度值的度中心性^[1],从传播路径角度考虑的介数中心性^[2],考虑网络中节点之间距离的接近中心性^[3],对每个节点同时考虑权威值(authorities)和枢纽值(hubs)的 HITS 算法^[4],删除节点后,考虑网络拉普拉斯算子能量大小的拉普拉斯中心性^[5],考虑节点邻接节点间相互连接关系的结构洞指标^[6].还有一些其他类型的指标与方法,如:半局部中心性^[7],特征向量中心性^[8],残余接近中心性^[9],k-壳分解法^[10],PageRank 算法^[11],累计提名^[12],LeaderRank 算法^[13]等一系列指标和算法.

为了较全面的研究不同节点重要性指标对网络重要节点的影响程度,进而选择较能体现网络重要节点性质的指标.本文应用主成分分析方法^[14]对选取的七个指标进行研究,同时,本文选取了七种不同类型的网络进行实验分析,七种网络分别为:风筝网络^[15]、ARPA 网络、艾滋病患者性关系网络^[16]、水网络^[17]、换热网络^[18]、无标度网络、小世界网络.

2 理论基础

2.1 主成分分析的数学模型

(1) 主成分的定义

如果将描述系统的 n 个指标看作是 n 维空间的 n 个随机变量,则主成分的定义如下:

设 $a = (a_1, a_2, \dots, a_n)$ 为 n 维空间 R_n 的单位向量, 并记所有单位向量集合为:

$$R_0 = \{a \mid aa^T = 1\} \quad (1)$$

即 n 个线性相关的随机变量为:

$$X = (X_1, X_2, \dots, X_n)^T \quad (2)$$

记 $D(X_i)$ 为 X_i 的方差, $z_i = a_i X$, $a_i \in R_0$, 则有如下定义:

若 $D(z_1) = \max_{a_i \in R_0} \{a_i X\}$, 称 z_1 为 X 的第一主成分, 记为: $z_1 = \beta_1 X$, $\beta_1 \in R_0$.

(2) 主成分的计算

设 X 为 n 维空间的随机变量, 且 $E(X) = 0$, $\sigma = E(XX^T)$, 则

$$\begin{aligned} \sigma &= E(XX^T) = E(X)E(X^T) + \text{cov}(XX^T) \\ &= \text{cov}(XX^T) \end{aligned} \quad (3)$$

即 σ 为一实对称的 n 阶协方差矩阵, 可以证明 σ 具有 n 个大于零的特征根, 记为: $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$, 则 X 的第 k 个主成分 $z_k = \beta_k X$ 的线性系数 β_k 为 σ 的第 k 个特征根 λ_k 的特征向量, 则可求得 n 个主成分.

2.2 主成分计算步骤

Step 1: 对样本进行标准化处理, 标准化消除了原来各指标的量纲, 使各指标之间具有可比性. 计算公式为:

$$X_{ij} = \frac{Y_{ij} - \bar{Y}_j}{s_j} \quad (4)$$

其中:

$$\bar{Y}_j = \frac{1}{m} \sum_{i=1}^m Y_{ij} \quad (j=1, 2, \dots, n) \quad (5)$$

$$s_j^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_{ij} - \bar{Y}_j)^2 \quad (j=1, 2, \dots, n) \quad (6)$$

Step 2: 利用标准化后的样本估计 σ . 由 $\sigma = E(XX^T) = \text{cov}(XX^T)$ 可知, 目的为通过样本估计总体的协方差矩阵. 计算公式如下:

$$\sigma_{ij} = \frac{1}{m-1} \sum_{i=1}^m X_{ki} X_{kj} \quad (i, j=1, 2, \dots, n) \quad (7)$$

通过计算可以得到一个实对称的协方差矩阵 σ .

Step 3: 计算各主成分. 根据得到的协方差矩阵 σ 即可得到 n 个非负特征根 $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$, 从而得到 n 个单位化特征向量, 构成一个正交矩阵, 记为 a , 则

$$a = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (8)$$

a_{ij} 中的 i 为第 i 个主分量, j 为第 j 个主分量.

3 实例验证

3.1 实验指标及网络的选取

(1) 指标的选取

为了使指标较为全面地从不同角度反映网络节点

的重要性, 并综合国内外研究, 本文选取了 7 个节点重要性指标: 度中心性 (DC)、介数中心性 (BC)、接近中心性 (CC)、聚类系数 (C)、HITS 值、拉普拉斯中心性 (LP)、结构洞 (SH). 7 个指标的具体含义见表 1.

表 1 七种不同指标及其定义

指标	含义
度中心性	节点 i 的实际连边数与其最大可能的连边数的比值
介数中心性	经过节点 i 的最短路径数占所有最短路径数的比例
接近中心性	节点 i 到网络其他所有节点的距离平均值的倒数
聚类系数	节点 i 相邻节点间的连边数与其最大可能连边数之比
HITS 值	一个节点的权威性和枢纽性
拉普拉斯中心性	节点 i 被删除后, 相应减少的拉普拉斯算子能量大小
结构洞	节点 i 与其他节点间的约束系数

(2) 指标选取说明

节点的度中心性指标、聚类系数和介数中心性指标分别从局部、中观和全局视角研究了网络节点重要性. 接近中心性指标从一个新的角度考虑, 指标值越大的节点在网络中所处的位置越靠近网络中心. HITS 算法将网络节点的属性分为两个相互影响的指标: 权威值和枢纽值, 其中权威值描述节点对网络中信息的原创性, 枢纽值则衡量节点在网络中传播的作用. 拉普拉斯中心性通过研究节点递阶删除, 相应损失拉普拉斯算子能量越大, 则该节点越重要. 处于结构洞位置的相关节点具有获取网络中信息的结构优势, 且对网络中信息的传播影响较强.

综上, 七个不同指标分别从不同角度、不同层次对网络中的节点重要性进行了研究, 不同指标间相互补充, 进而能够全面地体现节点在网络中的重要程度.

(3) 网络的选取

本文选取了 7 种不同类型的网络, 其中 5 个实际网络: 风筝网络、ARPA 网络、艾滋病患者性关系网络、水网络和换热网络; 2 个通过 Pajek 仿真软件模拟生成的网络: 无标度网络和小世界网络.

3.2 指标说明

(1) 度中心性

$$DC_i = \frac{k_i}{N-1} \quad (9)$$

k_i 为节点 i 的度值, N 为网络节点总数. 度中心性是一种简单的节点重要性衡量指标, 度中心性越大节点越重要.

(2) 介数中心性

$$BC_i = \sum_{s \neq t \neq i} \frac{n_{st}^i}{g_{st}} \quad (10)$$

g_{st} 为节点 s 到节点 t 的最短路径数, n_{st}^i 为从节点 s 到节点 t 且经过节点 i 的最短路径数. 节点的介数值描述了

节点对网络中其他节点对之间沿着最短路径传输信息的控制能力大小。

(3) 接近中心性

$$CC_i = \frac{1}{d_i} = \frac{N}{\sum_{j=1}^N d_{ij}} \quad (11)$$

$$d_i = \frac{1}{N} \sum_{j=1}^N d_{ij} \quad (12)$$

d_i 表示节点 i 到网络中其他所有节点距离的平均值, N 为网络节点总数, d_{ij} 为节点 i 到节点 j 的距离. 接近中心性能够反映节点在网络中的相对位置, 节点的接近数越大, 则节点越靠近网络的中心.

(4) 聚类系数

$$C_i = \frac{E_i}{(k_i(k_i - 1))/2} = \frac{2E_i}{k_i(k_i - 1)} \quad (13)$$

E_i 为节点 i 的邻居之间实际存在的边数, k_i 为节点 i 的度值, 即其邻居数. 聚类系数反映了节点之间的紧密程度, 聚类系数越大, 则节点越聚集, 反之节点越松散.

(5) HITS 算法

$$x'_i(k) = \sum_{j=1}^N a_{ji} y_j(k-1), i=1, 2, \dots, N \quad (14)$$

$$y'_i(k) = \sum_{j=1}^N a_{ij} y'_j(k), i=1, 2, \dots, N \quad (15)$$

$$x_i(k) = \frac{x'_i(k)}{\|x'(k)\|}, i=1, 2, \dots, N \quad (16)$$

$$y_i(k) = \frac{y'_i(k)}{\|y'(k)\|}, i=1, 2, \dots, N \quad (17)$$

$x'_i(k)$, $y'_i(k)$ 分别为时刻 k 时, 节点 i 的权威值和枢组值. a_{ji} 为邻接矩阵中节点 j 指向节点 i 的值, a_{ij} 为邻接矩阵中节点 i 指向节点 j 的值, N 为网络节点总数. $x_i(k)$, $y_i(k)$ 分别为归一化后节点 i 在时刻 k 时的权威值和枢

组值.

(6) 拉普拉斯中心性

$$C_i^L = (\Delta E)_i = E_L(G) - E_L(H) \quad (18)$$

$$E_L(G) = \sum_{i=1}^N \lambda_i^2 \quad (19)$$

$E_L(G)$ 为网络 G 的拉普拉斯能量, $E_L(H)$ 为删除节点 i 后网络的拉普拉斯能量, N 为网络节点总数, λ_i 为网络对应的拉普拉斯矩阵的特征值.

(7) 结构洞约束系数

$$C_i = \sum_{j \in \Gamma(i)} (p_{ij} + \sum_q p_{iq} p_{qj})^2, q \neq i, j \quad (20)$$

$$p_{ij} = \frac{z_{ij}}{\sum_{j \in \Gamma(i)} z_{ij}} \quad (21)$$

$$z_{ij} = \begin{cases} 1, & i \text{ 到 } j \text{ 有连边} \\ 0, & i \text{ 到 } j \text{ 无连边} \end{cases} \quad (22)$$

$\Gamma(i)$ 为节点 i 的邻居节点集合, p_{ij} 表示节点 i 为维持与节点 j 的邻居关系所耗费的成本占总成本的比例. 节点的结构洞约束系数刻画了节点形成结构洞可能性的大小, 系数越小节点越易形成结构洞, 则节点越重要, 反之亦然.

3.3 网络实验分析

为了研究在某种网络中不同节点重要性指标对节点重要性影响程度, 以网络中的节点数据为样本对 7 种网络单独进行主成分分析, 以风筝网络为例, 具体实验步骤如下:

Step 1: 通过 Pajek 仿真软件得到网络中各个节点 7 种指标的具体数值见表 2.

Step 2: 对样本数据进行标准化处理, 见公式 (4) ~ (6), 通过标准化处理得到如下数据, 见表 3.

表 2 风筝网络中七种指标具体数值

节点序号	DC	BC	CC	C	HITS 值	LP	SH
1	0.111	0.000	0.310	0.000	0.0112	6	1.000
2	0.222	0.222	0.429	0.000	0.048	14	0.500
3	0.333	0.389	0.600	0.333	0.196	32	0.431
4	0.556	0.231	0.643	0.500	0.398	52	0.445
5	0.556	0.231	0.643	0.500	0.3978	52	0.445
6	0.333	0.000	0.500	1.000	0.286	40	0.664
7	0.667	0.102	0.600	0.533	0.481	86	0.470
8	0.667	0.000	0.500	1.000	0.286	40	0.664
9	0.444	0.023	0.529	0.667	0.352	54	0.541
10	0.444	0.023	0.529	0.667	0.352	54	0.541

表 3 标准化后指标数值

节点序号	DC	BC	CC	C	HITS 值	LP	SH
1	-1.742	-0.889	-2.117	-1.501	-1.758	-1.636	2.483
2	-1.142	0.728	-0.963	-1.501	-1.518	-1.282	-0.405
3	-0.542	1.944	0.695	-0.540	-0.553	-0.486	-0.804
4	0.663	0.793	1.112	-0.058	0.764	0.398	-0.723
5	0.663	0.793	1.112	-0.058	0.763	0.398	-0.723
6	-0.542	-0.889	-0.274	1.385	0.034	-0.133	0.542
7	1.263	-0.146	0.695	0.038	1.305	1.902	-0.578
8	1.263	-0.889	-0.274	1.385	0.034	-0.133	0.542
9	0.058	-0.722	0.007	0.424	0.464	0.486	-0.168
10	0.058	-0.722	0.007	0.424	0.464	0.486	-0.168

Step 3: 求出 7 个指标的相关系数矩阵表, 见表 4. 可通过 EXCEL 中自带的 CORREL 函数求解.

表 4 指标间的相关系数矩阵表

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1	-0.021	0.729	0.626	0.870	0.830	-0.527
X_2	-0.021	1	0.519	-0.442	-0.036	-0.060	-0.650
X_3	0.729	0.519	1	0.377	0.821	0.739	-0.859
X_4	0.626	-0.442	0.377	1	0.629	0.524	-0.140
X_5	0.870	-0.036	0.821	0.629	1	0.958	-0.587
X_6	0.830	-0.060	0.739	0.524	0.958	1	-0.571
X_7	-0.527	-0.650	-0.859	-0.140	-0.587	-0.571	1

Step 4: 求解相关系数矩阵的特征值和特征向量, 见表 5. 可通过 MATLAB 中的 eig 函数求解.

表 5 相关系数矩阵的特征值与特征向量

特征值	特征向量						
	c_{i1}	c_{i2}	c_{i3}	c_{i4}	c_{i5}	c_{i6}	c_{i7}
0.0007	-0.0156	-0.0241	-0.3273	-0.8243	0.0652	0.1506	-0.4308
0.0317	0.3787	0.4296	0.2791	-0.2309	-0.2075	-0.7004	-0.0844
0.1291	-0.6421	-0.3030	0.4654	0.0073	-0.1531	-0.2518	-0.4386
0.1932	0.1023	0.2803	0.0472	0.1444	-0.7700	0.4612	-0.2870
0.3919	0.6371	-0.4314	0.3152	0.1589	0.2314	0.1398	-0.4586
1.8868	-0.1651	0.6549	0.0230	0.2685	0.5117	0.1342	-0.4375
4.3666	-0.0196	0.1716	0.7045	-0.3859	0.1451	0.4160	0.3617

Step 5: 计算各个主成分的贡献率和累积贡献率, 前 r 个主成分的累积贡献率, 见表 6.

$\lambda_k / \sum_{i=1}^n \lambda_i$ 为第 k 个主成分的贡献率; $\sum_{i=1}^r \lambda_i / \sum_{i=1}^n \lambda_i$ 为

表 6 各个主成分的贡献率和累积贡献率

特征值	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
贡献率	0.0001	0.0045	0.018	0.028	0.056	0.270	0.624
累积贡献率	0.0001	0.005	0.023	0.051	0.107	0.376	1.000

同理,可以得到其他 6 个网络的特征值、特征向量、主成分贡献率和累积贡献率.

4 实验结果分析

4.1 实验指标与指标贡献率

为了直观地反映不同指标对节点重要性的贡献率,绘制七种不同指标对七种不同网络节点的贡献率曲线图,见图 1.

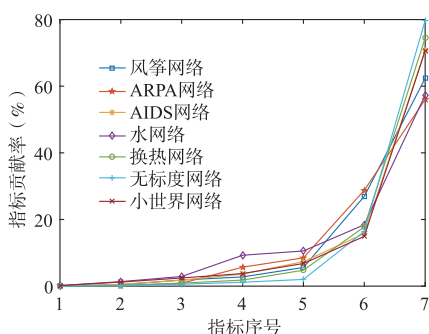


图 1 指标贡献率分布图

为了进一步研究七种指标对七种网络节点重要性的贡献率,绘制如下七种网络与指标贡献率关系图,见图 2.

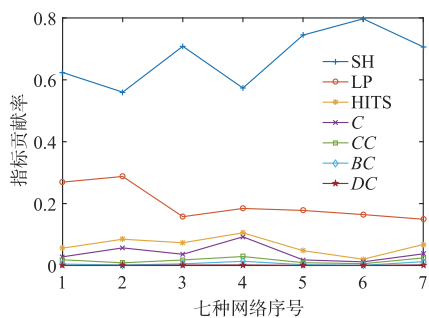


图 2 网络与指标贡献率分布图

由图 2 可以看出:结构洞指标的贡献率最大,拉普拉斯中心性指标次之,其余指标对节点重要性贡献率较小且相差不大.本文对 7 种经典的节点重要性指标进行了属性贡献率研究.其他未被选入参与本文实验的指标对节点重要性的影响有两方面,一方面可能来自指标本身具有的特殊属性;另一方面可能来自节点所处网络的固有属性产生的内生性影响.

4.2 相关文献对比分析

文献[19]使用结构洞方法对节点重要性进行研究,得到结构洞指标优于其他指标,其中接近中心性指标优于介数中心性指标;文献[20]同样使用结构洞理论,并对 Netscience 和 E-mail 网络节点影响力进行研

究,得到基于结构洞的指标要优于介数中心性与度值中心性指标;文献[21]基于结构洞,在 ARPA、AIDS 等网络中研究不同指标对节点重要性的影响,实验结果表明结构洞指标要优于其它所选指标,上述参考文献结果均与本文研究结果一致.

文献[22]基于多属性决策对节点重要性进行综合评价,其中使用主成分分析确定指标权重,该文献只使用了四种不同的指标,本文在其基础上增加了拉普拉斯中心性、HITS 值、聚类系数三个指标,能够更加全面的描述节点的重要性.文献[23]利用主成分方法得到不同指标的权重,没有进一步研究不同指标对网络节点重要性的影响,以及不同指标在不同网络中对节点重要性的影响.

4.3 实验指标相关性分析

为了探究结构洞指标与其他指标之间的相关性,本文采用肯德尔相关系数^[24]对指标进行相关性分析.

假设两个随机变量分别为 X, Y , 它们的元素个数均为 N , X_i, Y_i 分别为两个随机变量的第 i 个元素值. 当 $X_i > X_j$ 且 $Y_i > Y_j$ 或者 $X_i < X_j$ 且 $Y_i < Y_j$ 时,两元素是一致的;当 $X_i > X_j$ 且 $Y_i < Y_j$ 或者 $X_i < X_j$ 且 $Y_i > Y_j$ 时,两元素是不一致的;当 $X_i = X_j$ 或者 $Y_i = Y_j$ 时,两元素间无关系.肯德尔相关系数定义如下:

$$\tau = \frac{N_c - N_d}{N(N-1)/2} \quad (23)$$

$N_c - N_d$ 分别表示元素对集合 XY 中拥有一致性的元素对数和不一致性的元素对数, N 为网络节点总数.

图 3 显示了七种网络中结构洞指标与其它四种指标之间的相关性关系:

从图 3 可以看出,结构洞指标与其它四种指标呈负相关,这是因为结构洞指标值越小表明节点越重要,而其它四种指标相反.结构洞指标与度中心性指标、介数中心性指标的相关性较大, τ 平均在 -0.63 左右;而与另外两个指标的相关性较小, τ 平均在 -0.54 左右,可能的原因是结构洞指标和度中心性指标都利用了节点的局部信息;从网络结构上看,结构洞节点也是一个“桥接节点”和介数中心性具有相似的性质.

从不同网络对指标相关性影响方面可以看出:不同类型的网络对指标之间的相关性程度会产生影响.在 ARPA 网络中,结构洞指标与四种指标的相关性都比较小, τ 平均在 -0.40 左右;而在无标度网络中,结构洞指标与四种指标的相关性较大, τ 平均在 -0.76 左右,这可能是由于网络的不同拓扑结构产生的结果.

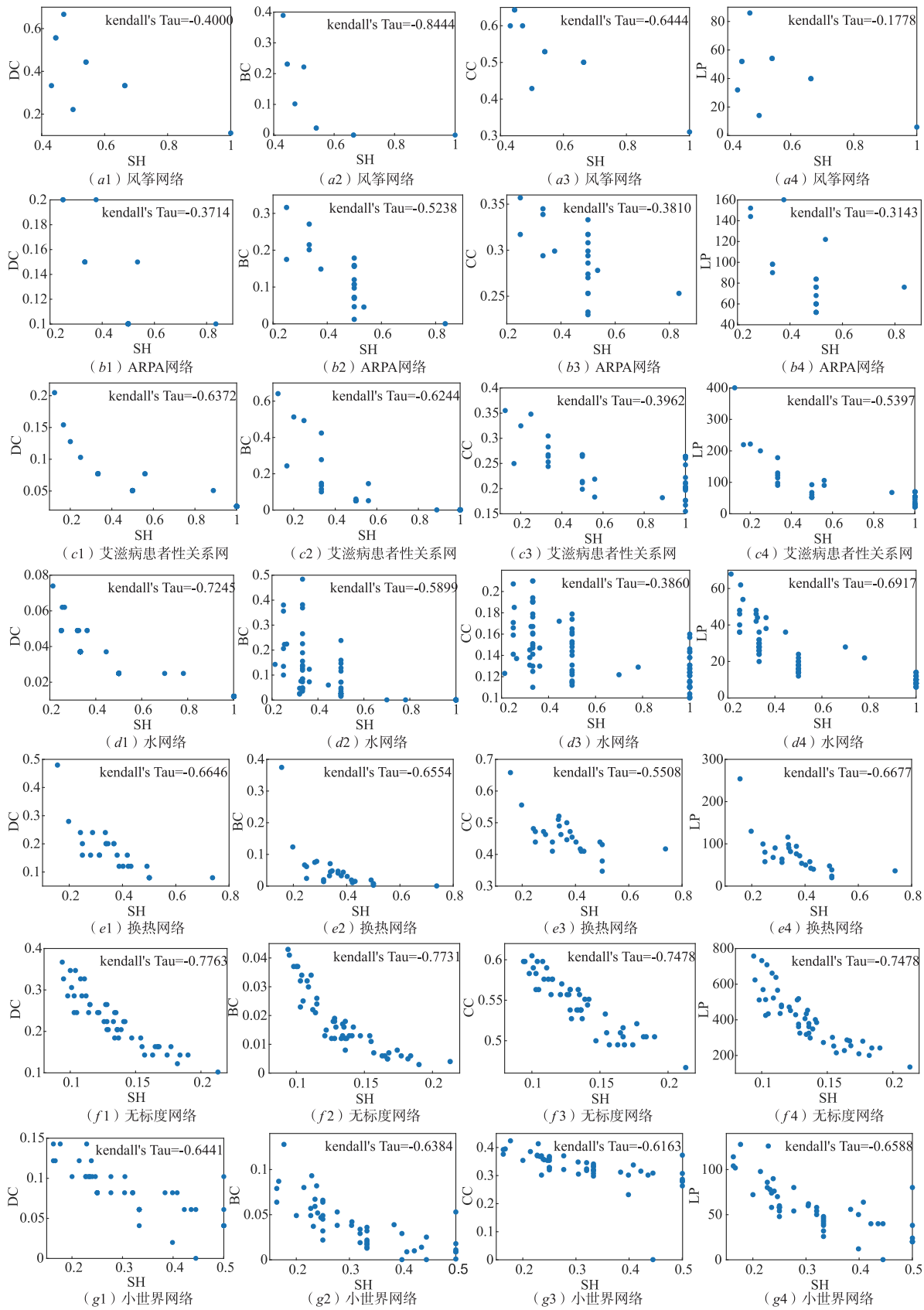


图 3 指标相关性分析图

为了更加直观地研究结构洞与其它指标在不同网络中的相关性变化,以七种网络的序号为横坐标,以结构洞与其它指标之间的相关系数为纵坐标,得到下图,见图 4.

由图 4 可以看出,结构洞指标与其他指标之间的相关系数在换热网络、无标度网络、小世界网络中的差异较小;而在风筝网络、ARPA 网络、艾滋病患者性关系网络和水网络中差异较大,反映了不同指标在不同网络中对重要节点的贡献率的差异.

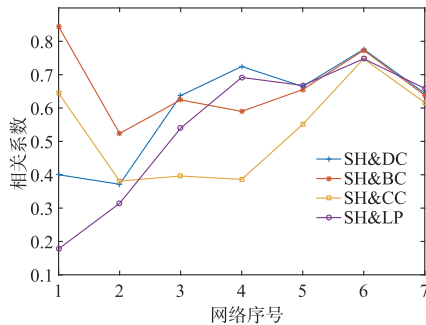


图 4 网络与相关系数变化关系图

5 结论

基于主成分分析,本文研究了不同类型的节点重要性指标对节点重要性的贡献率,进而能够选择更能够反映节点性质的指标.同时,探究了各指标之间的相关性以及不同类型网络对指标贡献率产生影响的原因.本文为多属性指标选择问题提供了研究方法,为研究不同指标间的相互关系以及外界环境对指标产生影响的原因提供了一种思路.

参考文献

- [1] Bonacich P. Factoring and weighting approaches to status scores and clique identification[J]. *Journal of Mathematical Sociology*, 1972, 2(1): 113 - 120.
- [2] Freeman L C. A set of measures of centrality based on betweenness[J]. *Sociometry*, 1977, 40(1): 35 - 41.
- [3] Freeman L C. Centrality in social networks conceptual clarification[J]. *Social Networks*, 1979, 1(3): 215 - 239.
- [4] Kleinberg J M. Authoritative sources in a hyperlinked environment[J]. *Journal of Acm*, 1999, 46(5): 604 - 632.
- [5] Qi X, Duval R D, Christensen K, et al. Terrorist networks, network energy and node removal: a new measure of centrality based on laplacian energy[J]. *Social Networking*, 2013, 2(1): 19 - 31.
- [6] Burt R S. *Structural Holes: The Social Structure of Competition*[M]. Harvard University Press, 2010.
- [7] Chen D B, Lü L, Shang M S, et al. Identifying influential nodes in complex networks[J]. *Physica A Statistical Mechanics & Its Applications*, 2012, 391(1): 1777 - 1787.
- [8] Phillip Bonacich. Factoring and weighting approaches to status scores and clique identification[J]. *Journal of Mathematical Sociology*, 1972, 2(1): 113 - 120.
- [9] Dangelchev C. Residual closeness in networks[J]. *Physica A Statistical Mechanics & Its Applications*, 2006, 365(2): 556 - 564.
- [10] Kitsak M, Gallos L K, Havlin S, et al. Identification of influential spreaders in complex networks[J]. *Nature Physics*, 2011, 6(11): 888 - 893.
- [11] Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine [A]. *International Conference on World Wide Web*[C]. Elsevier Science Publishers B. V. 1998. 107 - 117.
- [12] Poulin R, Boily M C, Msse B R. Dynamical systems to define centrality in social networks[J]. *Social Networks*, 2000, 22(3): 187 - 220.
- [13] Linyuan L, Zhang Y C, Ho Y C, et al. Leaders in social networks, the delicious case [J]. *Plos One*, 2011, 6(6): e21202.
- [14] 汪应洛. 系统工程. 第 5 版[M]. 北京: 机械工业出版社, 2015. 52 - 55.
Wang Ying-luo. *Systems Engineering. Fifth Edition*[M]. Beijing: China Machine Press, 2015. 52 - 55. (in Chinese)
- [15] 汪小帆, 李翔, 陈关荣. 网络科学导论[M]. 北京: 高等教育出版社, 2012. 161 - 162.
Wang Xiao-fan, Li Xiang, Chen Guan-rong. *An Introduction of Network Science*[M]. Beijing: Higher Education Press, 2012. 161 - 162. (in Chinese)
- [16] Klovdahl A S. Social networks and the spread of infectious diseases: the AIDS example[J]. *Social Science & Medicine*, 1985, 21(11): 1203 - 1216.
- [17] 刘忠华, 于华, 杨方廷. 基于复杂网络理论的水网节点重要性评价研究[J]. *中国科学: 技术科学*, 2014(12): 1280 - 1294.
Liu Zhong-hua, Yu Hua, Yang Fang-ting. Evaluate the node importance for water network based on complex network theory[J]. *Chinese Science: Technical Science*, 2014(12): 1280 - 1294. (in Chinese)
- [18] 王政, 孙锦程, 刘晓强, 等. 基于复杂网络理论的大型换热网络节点重要性评价[J]. *化工进展*, 2017, 36(5): 1581 - 1588.
Wang Zheng, Sun Jin-cheng, Liu Xiao-qiang, Jiang Ying, Jia Xiaoping, Wang Fang. Evaluation of the node importance for large heat exchanger network based on complex network theory [J]. *Chemical Industry and Engineering Progress*, 2017, 36(5): 1581 - 1588. (in Chinese)
- [19] 韩忠明, 吴杨, 谭旭升, 等. 面向结构洞的复杂网络关键

- 节点排序[J]. 物理学报, 2015, 64(5): 421 - 429.
Han Zhong-ming, Wu Yang, Tan Xu-sheng, Duan Da-gao, Yang Wei-jie. Ranking key nodes in complex networks by considering structural holes [J]. Acta Physica Sinica, 2015, 64(5): 421 - 429. (in Chinese)
- [20] 苏晓萍, 宋玉蓉. 利用邻域“结构洞”寻找社会网络中最具影响力节点[J]. 物理学报, 2015, 64(2): 1 - 11.
Su Xiao-ping, Song Yu-rong. Leveraging neighborhood “structural holes” to identifying key spreaders in social networks[J]. Acta Physica Sinica, 2015, 64(2): 1 - 11. (in Chinese)
- [21] Yu H, Cao X, Liu Z, et al. Identifying key nodes based on improved structural holes in complex networks[J]. Physica A Statistical Mechanics & Its Applications, 2017, 486.
- [22] 于会, 刘尊, 李勇军. 基于多属性决策的复杂网络节点重要性综合评价方法 [J]. 物理学报, 2013, 62(2): 020204-1 - 020204-9.
Yu Hui, Liu Zun, Li Yong-jun. Key nodes in complex networks identified by multi-attribute decision-making method [J]. Acta Physica Sinica, 2013, 62(2): 020204-1 - 020204-9. (in Chinese)
- [23] Zhang K, Zhang H, Wu Y D, et al. Evaluating the importance of nodes in complex networks based on principal component analysis and grey relational analysis [A]. IEEE International Conference on Networks [C]. 2011. 6955 (2): 231 - 235.
- [24] Kendall M G. The treatment of ties in ranking problems [J]. Biometrika, 1945, 33(3): 239 - 251.

作者简介



胡 钢 男, 1970 年出生, 甘肃天水人. 博士后, 副教授, 主要研究方向为复杂网络系统建模仿真与均衡分析、多属性决策、供应链金融.
E-mail: hug_2004@126.com



徐 翔 男, 1993 年出生, 安徽滁州人. 现为在读博士研究生, 主要研究方向为复杂网络系统建模仿真与均衡分析、供应链网络仿真与应用.
E-mail: 18726008700@163.com