

一种基于资源分配策略的复杂网络中心性测度

陈国强¹ 陈亮²

(河南大学计算机与信息工程学院 开封 475004)¹ (河南工业大学信息科学与工程学院 郑州 450001)²

摘要 针对标准中心性测度不适用于非联通网络这一问题,提出了基于资源分配策略的复杂网络中心性测度。节点的资源分配中心性测度定义为节点从其它节点接受的资源量,如果一个节点从其他节点接受的资源量越多,则该节点越重要。通过人工网络和现实网络实验表明,该度量不仅适用于联通网络,也适用于非联通网络,相较于标准测度,可以检测桥节点,而且具有良好的稳定性。

关键词 中心性测度,资源分配,复杂网络

中图法分类号 TP393 **文献标识码** A

New Measure of Complex Network Centrality Based on Resource Allocation

CHEN Guo-qiang¹ CHEN Liang²

(School of Computer and Information Engineering, Henan University, Kaifeng 475004, China)¹

(School of Information and Engineering, Henan University of Technology, Zhengzhou 450001, China)²

Abstract A new centrality measure for complex networks, called resource allocation centrality measure, was proposed in this paper. It can overcome some disadvantages of several often used centrality measures that can not be applicable to the disconnect networks. The resource allocation centrality of a node is defined as its amount of resource received from other nodes. If a node receives more resources from other nodes, the node is more important than others. Simulation tests on artificial networks and real networks show that the resource allocation centrality measure has good performances in detecting bridge node and has good stability.

Keywords Centrality measures, Resource allocation, Complex networks

1 引言

中心性是复杂网络分析中被研究最多的概念之一。对于一个规模巨大的网络拓扑而言,很难快速地获得全局信息,但是由于网络中节点的异质性,使得每个节点在网络中的重要程度不同^[1]。中心性测度用来量化网络中某些顶点或者边比其它顶点或者边更加处于中心地位这一直观感受。关于中心性一般就是通过为点或者边指定一个实数值来表示顶点或者边的重要性。历史上,对于中心性测度没有一个普遍接受的定义,几乎每个人在介绍自己的中心性测度时都没有给出严格的中心性定义。Bavelas 对小集团中人们的交流特征比较感兴趣,首次在人类交流研究中运用中心性的思想^[2],并且假定结构中心性和小组中有影响的人物具有一定的关系。此后,许多测度被设计出来,包括度数中心性、接近度、介数、特征向量中心性等,用来量化社会网络中个体的重要性^[3-5]。结构中心性不仅吸引了物理学家的注意,而且其应用也扩展到了生物、技术网络和地理网络等领域^[6,7]。

然而,在这些不同测度公式中都隐含着事物在网络中以某种方式流动的假设^[8]。例如,某些测度,比如 Freeman 提出的接近度和介数中心度^[9],仅计算测地线路径,显然假定事物仅沿着最短可能路径在网络中流动。其它测度,比如流介

数^[10],不是假设最短路径,而是假设适当的路径,即节点被访问不超过一次。不考虑流动轨迹,某些测度(比如介数)假设从一个节点到另一个节点的流是不可分割的(像一个包裹)并且必须选择一条路径;然而,其他的测度(比如特征向量)假设多条路径可以同时进行(像信息或者传染病)。

但上述的各类指标都是侧重于单一的拓扑信息,挖掘的结果存在不确定性。为了更加合理地找到复杂网络的拓扑中心,本文假设事物在两点之间通过公共邻居流动,并且可有 multiple 路径同时存在,提出一种新复杂网络的中心性度量指标,称为资源分配中心度。节点的资源分配中心度定义为节点从其它节点接受的资源量,一个节点从其他节点接受的资源量越多,该节点越重要。另外,该定义可以自然地扩展应用到组的中心性。通过人工网络和现实网络实验表明,该度量具有良好的稳定性。

2 标准中心性测度

下面列出了社会网络中一些标准的中心性测度,最近提出的模块度中心性也被列了出来。假定 G 表示一个有 N 个顶点 K 条边无向无权的图,图 G 的邻接矩阵用 A 表示,其中如果 i 和 j 之间有边连接,则 A_{ij} 等于 1,否则 A_{ij} 等于 0。

度中心性(C^D)基于重要的顶点应该是图中与其他顶点

到稿日期:2011-03-15 返修日期:2011-05-28 本文受国家自然科学基金课题(60873099)资助。

陈国强(1977—),男,博士生,主要研究方向为复杂网络、计算智能等,E-mail:chengq08@163.com;陈亮(1978—),男,硕士,讲师,主要研究方向为人工智能、复杂网络等。

联系最多的顶点这一思想。顶点 i 的度中心性定义为

$$C_i^D = \frac{k_i}{N-1} = \frac{\sum_{j \in G} A_{ij}}{N-1} \quad (1)$$

式中, k 表示顶点 i 的度。

接近度中心性(C^A)度量节点 i 沿着最短路径距离所有其他节点远近的程度^[3], 其定义为

$$C_i^A = (L_i)^{-1} = \frac{N-1}{\sum_{j \in G} d_{ij}} \quad (2)$$

介数中心度(C^B)基于如果一个顶点位于越多的顶点之间, 从感觉上它位于更多的连接两点的 shortest 路径上, 那么它就越居于中心位置^[9], 其定义为

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j \in G, j \neq i} \sum_{k \neq i, k \neq j} \frac{n_{jk}(i)}{n_{jk}} \quad (3)$$

信心中心度(C^I)是顶点的中心度与网络对顶点去掉后的反应能力建立关联。顶点 i 的信息中心度定义为从图 G 中去掉与顶点 i 相连的边后网络效率的相对变化值^[12]。其定义为

$$C_i^I = \frac{\Delta E}{E} = \frac{E(G) - E(G')}{E(G)} \quad (4)$$

其中网络的效率定义为

$$E(G) = \frac{1}{N(N-1)} \sum_{i, j \in G, i \neq j} \frac{1}{d_{ij}} \quad (5)$$

模块度中心性(C^M)基于模块矩阵的最大特征值对应的特征向量的元素的大小给出了顶点属于指派社区的强度的测度^[13, 14]。模块度 Q 定义为

$$Q = \sum_{j=1}^N \sum_{k=1}^p \beta_j (u_j^T x_k) \quad (6)$$

式中, p 表示非重叠的社区数, β_j 表示模块矩阵的第 j 个特征值, u_j 表示对应于 β_j 的特征向量, x_k 表示指派矩阵的第 k 行向量。定义 $\beta = \max(|\beta_j|)$, 如果 μ_i 表示对应于 β 的特征向量 u_j 的第 i 个元素的大小, 则顶点 i 的模块度中心性定义为

$$C_i^M = |\mu_i| \quad (7)$$

3 资源分配中心性测度

在本节中, 受网络中资源分配过程^[15, 16]的启发, 提出一个新的中心性测度, 该测度对于关键节点具有较好的结果和稳定性。

资源分配中心性基于以下思想: 在一个有资源流动的网络中, 节点(一组节点)的重要性与它从其它节点接受的资源量有关。如果一个节点从其它节点接受的资源越多, 则该节点越重要。节点 i 的资源分配中心性 C_i^R 定义为

$$C_i^R = \sum_{j \in G, j \neq i} T_i(j) \quad (8)$$

式中, $T_i(j)$ 表示顶点 i 从顶点 j 接受的资源量。

式(8)中 $T_i(j)$ 的精确定义依赖于具体应用。本文中, 考虑一对顶点 i 和 j , 假设顶点 j 拥有的资源量定义为 $d(j)$ 。如果顶点 i 和 j 直接相连, 那么 j 发给 i 的资源量定义为 $d(j)$; 如果顶点 i 和 j 不直接相连, 在顶点 j 给 i 发资源的过程中, 它们的公共邻居 k 扮演传递者的角色, 公共邻居 k 在收到资源后, 将平均转发给其直接相连的顶点。顶点 i 从 j 收到的资源量定义为

$$T_i(j) = \begin{cases} d(j), & \text{如果 } i \text{ 和 } j \text{ 直接相连} \\ \sum_{k \in N(i) \cap N(j)} \frac{d(k)}{d(k)}, & \text{否则} \end{cases} \quad (9)$$

式中, $d(j)$ 表示顶点 j 的度, $N(i)$ 表示顶点 i 的邻居。

为了定义组资源分配中心性(相对于单个顶点而言, 组资

源分配中心性就是一组顶点的中心性), 一个最直观的办法就是用超点 S 来代替组中的所有顶点, 超点的邻居就是组中所有顶点邻居的联合。即如果在组中至少有一个顶点和组外某顶点有连接关系, 那么超点 S 就与该组外顶点有连接关系。如果用 G 表示原始图, 则组点变成超点后的图用 G' 表示。如果图 G 有 n 个顶点, 组 C 有 c 个顶点, 那么 G' 就包含 $n-c+1$ 个顶点。通过计算 G' 中单个顶点的资源中心性, 就可以获得组中心性。超点 S 的组中心定义为

$$C_S^R = \sum_{j \in G', j \neq S, i \in S} T_i(j) \quad (10)$$

4 实验与结果

本节中给出人工网络和真实网络的例子来展示资源分配中心性测度的使用。

4.1 应用到人工网络

在极端条件下, 资源分配中心性可以和其它标准测度达到一致效果。例如: 它将星型网络的中央节点赋予最大中心性, 对完全图中的顶点赋予同等中心性。但是, 在下面的极端情况下, 比如在有桥节点连接两个主要部分的网络中, 本文所提测度和其他测度表现不同。桥节点在网络中通常起重要作用。在网络中, 并不是所有顶点地位都是相等的。去掉不同的顶点, 对网络的影响是不同的。如果一个节点是在末端, 则将其删除将不会有任何影响, 其它节点的相应等级也不会变化; 如果是桥节点则情况正好相反, 将其删除将会产生新的不连接部分。资源分配中心性可以检测桥接点, 其比标准测度有更好的结果和稳定性。下面, 将通过一个简单例子来展示。

考虑 $N1$ 和 $N2$ 两个无环图, 如图 1 所示。图 $N1$ 的 5 种中心性值如表 1 所列, 节点按 C^R 降序排列。尽管 C^R, C^M 和 C^C 都将顶点分为 5 个等级, 但有一些不同值需注意。

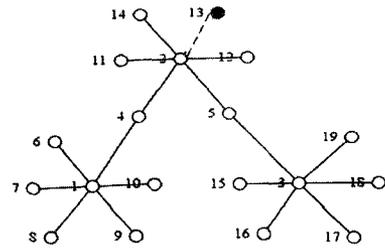


图 1 网络 $N1$ 有 19 个顶点, 网络 $N2$ 由网络 N 去掉顶点 13 得到

表 1 网络 $N1$ 的中心性值

顶点	C^R	C^M	C^C	C^B	C^D
2	14	0.5507	0.4500	0.7255	0.3333
4	13.8333	0.1787	0.4000	0.4706	0.1111
5	13.8333	0.1787	0.4000	0.4706	0.1111
1	10	0.5130	0.3462	0.4902	0.3333
3	10	0.5130	0.3462	0.4902	0.3333
11	7.1667	0.0946	0.3158	0	0.0556
12	7.1667	0.0946	0.3158	0	0.0556
13	7.1667	0.0946	0.3158	0	0.0556
14	7.1667	0.0946	0.3158	0	0.0556
6	7.0000	0.0841	0.2609	0	0.0556
7	7.0000	0.0841	0.2609	0	0.0556
8	7.0000	0.0841	0.2609	0	0.0556
9	7.0000	0.0841	0.2609	0	0.0556
10	7.0000	0.0841	0.2609	0	0.0556
15	7.0000	0.0841	0.2609	0	0.0556
16	7.0000	0.0841	0.2609	0	0.0556
17	7.0000	0.0841	0.2609	0	0.0556
18	7.0000	0.0841	0.2609	0	0.0556
19	7.0000	0.0841	0.2609	0	0.0556

如表 2 所列, C^R 将网络 $N1$ 中的最高值赋给了顶点 2, 第

二等级赋予顶点 4,5,第三等级赋予顶点 1,3,并且能将顶点 11,12,13,14(第 4 等级)从剩余顶点区别开。其它测度只有 C^M 和 C^C 能做此区分。 C^B 和 C^D 都没有 C^R, C^M 和 C^C 所取得的效果。实际上, C^D 将最高值赋给了顶点 1,2 和 3,这 3 个顶点都有 6 个邻居;第二级赋给了顶点 4,5,这两个顶点都有两个邻居。 C^B 和 C^D 都没有将顶点 11,12,13,14 从剩余顶点区别开。

表 2 网络 N1 的中心性等级

等级	C^R	C^M	C^C	C^B	C^D
1	2	2	2	2	2,1,3
2	4,5	1,3	4,5	1,3	4,5
3	1,3	4,5	1,3	4,5	others
4	11,12,13,14	11,12,13,14	11,12,13,14	others	
5	others	others	others		

网络 N2 的顶点等级如表 3 所列。在网络 N1 中,顶点 1,2,3 具有相同的邻居数;在网络 N2 中,顶点 2 比顶点 1 和 3 的邻居数少。这影响了 C^D 的顶点级数,但是没有影响基于 C^R, C^M, C^C 和 C^B 的顶点级数。 C^M 和 C^D 将顶点 2 从最高等级变成了第二等级。

表 3 网络 N2 的中心性等级

等级	C^R	C^M	C^C	C^B	C^D
1	2	1,3	2	2	1,3
2	4,5	2	4,5	1,3	2
3	1,3	4,5	1,3	4,5	4,5
4	others	others	11,12,14	others	others
5	11,12,14	11,12,14	others		

从上可以看出 C^R 比其他测度具有较好的解和稳定性。

4.2 应用到真实的网络

下面研究由 Linda Wolfe 收集的真实数据,该数据在 UCINET 中被收录为一个标准测试数据集^[17]。该数据记录了 20 只猴子在 3 个月之间的相互作用关系,该相互作用定义为两只猴子在河边同时出现。结果如图 2 所示,包含 6 个孤立点和一个由 14 个顶点相互连接的部分。该图已在文献^[12,18]中被研究过,标准顶点中心性测度 C^D, C^B, C^C 以及 C^I 已经推广到组中心性。

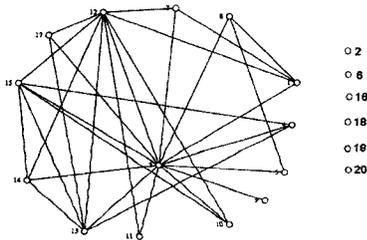


图 2 猴子交互图^[18]

通过计算不同测度的中心值,可以比较本文所提测度与标准测度的中心性。表 4 中报告了获得的顶点中心性值。

除了 C^M ,所有其它测度为第 3 只猴子赋予最高值,第 2 高值赋给了第 12 只猴子,第 3 高值赋给了第 13 和第 15 只猴子。根据 C^R, C^I, C^D, C^M ,6 个孤立的猴子被赋予最低的中心性。 C^B 赋予 14 个节点中心值为 0,包括 6 个孤立的猴子以及顶点 4,5,7,9,10,11,14,17。从数据集考虑,由 C^R 产生的 20 个顶点的等级和由 C^I 和 C^D 产生的等级相似。根据测度 C^I 和 C^D ,顶点 1 和 14,顶点 4,7,8,10 和 17,顶点 5 和 11 分别具有相同的等级。但是,根据测度 C^R ,除了顶点 10 和 17,它们

具有不同的等级。

表 4 猴子的不同测度的顶点中心性值

顶点	C^R	C^I	C^D	C^B	C^M
1	33.6410	0.1393	0.2105	0.0058	0.0816
2	0	0	0	0	0
3	49	0.3751	0.6842	0.2602	0.7818
4	32.9487	0.1311	0.1579	0	0.0921
5	20.7179	0.1230	0.1053	0	0.1030
6	0	0	0	0	0.0000
7	31.9551	0.1311	0.1579	0	0.1469
8	25.0769	0.1311	0.1579	0.0029	0.0154
9	16.6923	0.1148	0.0526	0	0.1519
10	34.9957	0.1311	0.1579	0	0.1826
11	28.1452	0.1230	0.1053	0	0.2125
12	46.4423	0.1803	0.4737	0.0604	0.4193
13	41.2179	0.1557	0.3158	0.0107	0.0600
14	39.5128	0.1393	0.2105	0	0.1527
15	41.2179	0.1557	0.3158	0.0107	0.0600
16	0	0	0.0	0	0.0000
17	34.9957	0.1311	0.1579	0	0.1826
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0

在图 3 中,每一个顶点的中心度值基于最高值进行了标准化,并且基于 C^R 的标准值进行排序。尽管具有不同的值,但 C^R 的表现与 C^I, C^D 以及 C^B 相似。 C^M 与其他测度表现不同:在等级 6,7,11 和 14 出现 4 个峰值,相应的顶点对应为 10,17,11,9,表明根据测度 C^M ,这 4 个猴子被赋予较高的等级。

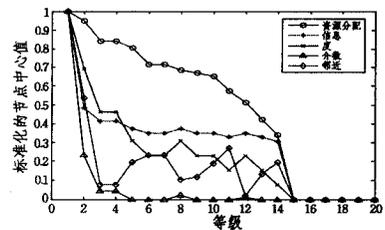


图 3 猴子的标准化中心值

结束语 本文介绍了一种新的复杂网络中心性测度,即资源分配中心性,并且扩展到组的中心性。通过标准测试集进行实验,所提测度比其它标准测度对桥节点可以得到更好的结果,并且具有良好的稳定性。下一步需要继续研究在何种情况下,所提测度比其它测度更适用。

参考文献

- [1] 韩言妮,李德毅,陈桂生. 软件网络的多粒度拓扑特性分析及其应用[J]. 计算机学报,2009,32(9):1711-1721
- [2] Scott J. Social network analysis: a handbook(2nd ed)[M]. London: Sage Publications, 2000
- [3] 李清泉,曾喆,杨必胜,等. 城市道路网络的中介中心性分析[J]. 武汉大学学报,2010,35(1):37-41
- [4] 江健,淦文燕,赵东杰,等. 基于拓扑势的社会通信网局域中心性分析[J]. 系统工程学报,2010,25(6):861-866
- [5] 肖俐平,孟晖,李德毅. 基于拓扑势的网络节点重要性排序及评价方法[J]. 武汉大学学报,2008,33(4):379-383
- [6] Pastor-Satorras R, Vázquez A, Vespignani A. Dynamical and correlation properties of the internet[J]. Physical Review Letter, 2001,87(25):258701

(下转第 52 页)

发进程竞争资源。因此,系统进行死锁检测时,先检测各站点局部是否出现死锁,若局部存在死锁,则整个系统肯定处于死锁状态,否则继续在全局范围内检测死锁。

4.1 死锁检测过程

(1)在单站点范围内,建立每个进程加锁请求的 Petri 网模型 $\Sigma_1=(S_1, T_1, F_1, W_1, M_{01})$;

(2)以资源作为连接点,构造站点的局部共享合成 Petri 网模型 $\Sigma_2=(S_2, T_2, F_2, W_2, M_{02})$;

(3)根据并发进程的请求给出 Petri 网的始态 M_0 ,画出可达标识图 $RG(\Sigma_2)=(R(M_{0P}), E, P)$;

(4)根据 $RG(\Sigma_2)$ 化简 $\Sigma_2=(S_2, T_2, F_2, W_2, M_{02})$ 得 $\Sigma'=(S', T', F', M_{0'})$;

(5)利用资源解锁的变迁作为连接点,构造分布式系统的全局 Petri 网模型 $\Sigma=(S, T, F, M_0)$;

(6)画出 $\Sigma=(S, T, F, M_0)$ 的可达标识图;

(7)找出死锁的进程,杀死其中一个进程;

(8)从 $\Sigma=(S, T, F, M_0)$ 中删除被杀死的进程的状态及相关变迁,重复(6)、(7)直至系统中不存在死锁。

4.2 死锁检测原理

分布式系统的全局 Petri 网模型 $\Sigma=(S, T, F, M_0)$ 在经过一系列的变迁之后,将得到一个最终状态 $M_T=\{M_T(P_1), M_T(P_2), \dots, M_T(P_m)\}$,即 $M_0[\sigma > M_T, \sigma \in T, \text{且} \rightarrow \exists t \in T \text{使} M_T[t >]$ 。系统死锁的充分必要条件是 $M_T(P_1), M_T(P_2), \dots, M_T(P_m)$ 中至少有两项大于 0。

证明:充分性:若 $M_T(P_1), M_T(P_2), \dots, M_T(P_m)$ 中至少有两项不等于 0,则说明至少有两个进程的加锁请求没有得到授权,且 M_T 是终态,这些加锁请求永远不会得到授权,那么可以肯定这些进程在相互等待对方解锁、释放资源。因此系统存在死锁。

必要性:若系统出现死锁,则系统中的每个进程的状态都将稳定不变,且有多于一个的进程 P_1, \dots, P_j 在相互等待其它进程释放资源, P_i, \dots, P_j 的加锁请求不能被授权,那么 $M(P_i), \dots, M(P_j) > 0$,即 $M_T(P_1), M_T(P_2), \dots, M_T(P_m)$ 中至

少有两项大于 0。

可以通过杀死进程的方法来解决死锁。设杀死的进程 P_k 有 n 个子进程 p_1, p_2, \dots, p_n ,记

$$M_0(P) = \sum_{i=1}^n M_0(P_i)$$

为减少重复加锁, P_k 应满足条件 $M_0(P_k) - M_T(P_k) = \max\{M_0(P_j) - M_T(P_j) | j \in \{1, 2, \dots, m\}\}$ 。

结束语 本文给出了分布式锁管理 DLM 中的 6 种锁的 Petri 网模型,并通过共享合成、同步合成及 Petri 网化简的方法建立分布式系统锁的 Petri 网模型,借助 Petri 网来检测系统的死锁状态。应用等待图或资源分配图检测死锁时,系统只会在进程迟迟不能得到资源的情况下发起死锁检测;基于 Petri 网的分布式死锁检测具有动态模拟性的优点,它能跟踪进程申请资源的动态变化过程,实时发现死锁,并且能选择杀死最合适的进程。

参考文献

- [1] 徐高潮,胡亮,鞠九滨.分布计算系统[M].北京:高等教育出版社,2004
- [2] Oracle. A42522_1;Oracle7 Parallel Server Concepts and Administration, Release 7. 3 [EB]. http://download.oracle.com/docs/pdf/A42522_1.pdf
- [3] 吴哲辉. Petri 网导论[M].北京:机械工业出版社,2004
- [4] 向学哲.分布式锁管理在集群系统中的应用研究[J]. MODERN COMPUTER, 2007(8): 20-22
- [5] 钱迎进,肖依,金士尧. Lustre 分布式锁管理器的分析与改进[J]. 计算机工程与科学, 2009, 31(A1): 146-149
- [6] 李春强,杨遂,陈春玲.分布式系统死锁的 Petri 网分析[J]. 微机发展, 2003, 13(5): 30-35
- [7] 刘荣峰,李引珍,吴桂华.基于 Petri 网的分布式系统并发进程死锁检测[J]. 计算机工程与设计, 2007, 28(22): 5353-5355
- [8] 韩耀军,蒋昌俊,罗雪梅.基于 Petri 网合成与化简的分布式数据库系统并发控制的死锁检测[J]. 小型微型计算机系统, 2004, 25(5): 821-826

(上接第 44 页)

- [7] Crucitti P, Latora V, Porta S. Centrality in networks of urban streets[J]. Chaos, an interdisciplinary journal of nonlinear science, 2006, 16(1): 015113
- [8] Borgatti S P. Centrality and network flow[J]. Social Networks, 2005, 27(1): 55-71
- [9] Freeman L C. Centrality in social networks conceptual clarification[J]. Social Networks, 1979, 1(3): 215-239
- [10] Freeman L C, Borgatti S P, White D R. Centrality in valued graphs: a measure of betweenness based on network flow[J]. Social Networks, 1991, 13(2): 141-154
- [11] Nieminen J. On the centrality in a graph[J]. Scandinavian Journal of Psychology, 1974, 15(1): 332-336
- [12] Latora V, Marchiori M. A measure of centrality based on network efficiency[J]. New Journal of Physics, 2007, 9(6): 188-198
- [13] Newman M E J. Finding community structure in networks using

the eigenvectors of matrices[J]. Physical Review E, 2006, 74(3): 036104

- [14] Wang Gao-xia, Shen Yi, Luan Ee. A measure of centrality based on modularity matrix[J]. Progress in Natural Science, 2008, 18: 1043-1047
- [15] Qing Ou, Jin Ying-di, Zhou Tao, et al. Power-law strength-degree correlation from a resource-allocation dynamics on weighted networks[J]. Physical Review E, 2007, 75(2): 021102
- [16] Zhou Tao, Lü Lin-yuan, Zhang Yi-chen. Predicting missing links via local information[J]. The European Physical Journal B, 2009, 71(4): 623-630
- [17] Borgatti S P, Everett M G, Freeman L C. UCINET, Version 1. 0 [Z]. Analytic Technologies, Columbia, SC, 1992
- [18] Everett M G, Borgatti S P. The centrality of groups and classed [J]. Journal of Mathematical Sociology, 1999, 23(3): 181-2011