

文章编号: 1007-6735(2008)03-0227-04

复杂网络节点中心性

荣莉莉, 郭天柱, 王建伟

(大连理工大学 系统工程研究所, 大连 116023)

摘要: 将网络中心性方法按照理论特征划分为节点关联性、网络最短路和模拟流问题, 并对现实网络的局域性、信息完备性和动态性进行了深入分析, 在此基础上建立了中心性方法与实际网络之间的匹配关系。

关键词: 复杂网络; 中心性; 网络结构; 应用; 适应性

中图分类号: N 94 **文献标识码:** A

Centralities of nodes in complex networks

RONG Li-li, GUO Tian-zhu, WANG Jian-wei

(Institute of Sysyem Engineering, Dalian University of Technology, Dalian 116023, China)

Abstract: The centralities suggested in literature are classified into three groups as nodes' correlations, networks' shortest paths and network stimulating flow problem. A match network is constructed between the centralities and real networks based on the analysis of node's domain, information maturity and dynamics property.

Key words: complex networks; centrality; network structure; application; adaptation

网络节点的重要性可以通过节点的中心性来衡量。在复杂网络的研究中, 采用了不同的中心性定义来描述网络的动态特性, 如网络在攻击状态下的鲁棒性、交通网络特性等。本文分析了节点的关联性、网络的最短路径和模拟流问题, 从这些角度对比了各种方法之间的异同。在总结了现实网络的局域性、信息完备性和动态性的基础上, 讨论了中心性方法与实际网络应用之间的适应性, 提出了针对不同网络结构特征和实际应用要求的中心性判断方法选择原则, 并探讨了将不同的中心性判断方法互补应用于同一网络模型。

1 中心性方法介绍和分类

当前很多网络研究都基于 Barabási^[1~3] 提出的择

优连接机制, 寻求合适模型对现实现象的合理解释, 这些模型都定义网络中心节点是那些在一定条件下拥有大量连接边的节点^[4], 但并不是所有的现实网络的特性都能通过这种简单的机制得到合理的解释。因此, 出现了其他一些不同的中心性判断方法。针对现实网络中出现节点拥有较小的连接度, 却对网络的联通性起到关键作用的介数中心性(BC)^[4]; 引入网络流特征, 根据网络中节点到达整个网络其他所有节点的难易程度, 提出了凝聚中心性(CC)^[5]; 特征值中心性(EC)^[6]是考虑节点已建立连接节点的重要性对该节点的影响而提出的; 网络流中心性(FC)^[7]按照流通的方式确定网络的几何中心; 基于节点对网络全局信息未知条件下, 提出了随机行走方法(RDW)^[8]; 根据网络节点在构造不同网络子图中的参与程度, 提出了子

收稿日期: 2008-03-25

基金项目: 国家自然科学基金资助项目(70771016)

作者简介: 荣莉莉(1964-), 女, 教授。

©1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

图中心性(SC)^[9].

a. 节点度

节点度^[4], 即与节点连接的边数, 也是与该节点建立连接的节点数. 设网络具有 n 个节点, k 为节点度, 则节点 i 的度中心性

$$C_d(i) = k(i) \tag{1}$$

b. 介数

节点介数^[10, 11] 定义为网络中节点对最短路径中经过节点 i 的个数占所有最短路径数的比例. 用 $g_{st,i}$ 表示节点对 s 和 t 最短路径经过 i 点的路径数, n_{st} 表示节点 s 和节点 t 之间存在所有最短路径的路径数, 则节点 i 的介数中心性

$$C_b(i) = \frac{2 \sum_{s < t} g_{st,i} \setminus n_{st}}{n(n-1)} \tag{2}$$

其中, $n(n-1)/2 = \sum_{s < t} 1$, 用来将节点介数标准化, 使得 $C_b \in [0, 1]$.

c. 凝聚度

与介数中心性方法类似, 提出了另外一种基于最短路的方法——凝聚中心性^[5], 它是基于网络节点达到网络其他所有节点所需路径综合的差异性, 利用信息在网络中的广播时间长短来确定网络节点的重要性, 定义为节点 i 到网络其他所有节点最短距离的总和, 凝聚度

$$C_c(i) = \sum_{s < t} d_{st} \tag{3}$$

d. 网络流

设网络具有 n 个节点 则节点 i 的网络流的定义为

$$C_f(i) = \sum_{s < t} \frac{g_{st,i}}{g_{st}} \tag{4}$$

式中, g_{st} 为网络中节点对 (s, t) 之间的所有路径数, 不包含回路; $g_{st,i}$ 为节点对 (s, t) 之间经过节点 i 的路径数.

e. 随机行走

随机行走模型的提出基于一个多数网络的事实, 网络节点对网络的整体特性是未知的, 这样就使得对整体网络选择最优不可能. New man 提出的随机行走算法(RDW)^[8] 为:

(a) 构建关系矩阵 $L = D - A$. 其中, A 为目标网络的邻接矩阵, D 为节点度组成的对角矩阵.

(b) 变换矩阵, 把 L 矩阵去掉最后一行和最后一列, 变成可逆矩阵.

(c) 求 L 矩阵的逆矩阵 L^{-1} , 在 L^{-1} 基础上添加元素全为 0 的一行一列, 构成新矩阵 T . T_{ij} 为 T

中的元素.

随机行走算法中网络节点的介数为

$$C_r(i) = \frac{2 \sum_{s < t} I_{st,i}}{n(n-1)} \tag{5}$$

其中

$$I_{st,i} = \frac{1}{2} \sum_j A_{ij} |T_{is} - T_{it} - T_{js} + T_{jt}| \quad i \neq s, i \neq t$$
$$I_{st,i} = 1 \quad i = s \text{ 或 } i = t$$

f. 子图

子图中心性方法^[9] 是对节点度中心性的改进, 基于节点对所在网络局部子图的参与程度来确定节点的重要性, 节点 i 的子图中心性

$$C_s(i) = \sum_{k=0}^{\infty} \frac{\mu_k(i)}{k!} \tag{6}$$

$$\mu_k(i) = (A^k)_{ii} \tag{7}$$

式中, A 为邻接矩阵; $\mu_k(i)$ 表示以节点 i 为起点经 k 个连边回到节点 i 的路径数目, 路径包含回路.

g. 特征向量

设网络具有 n 个节点, A 表示网络的邻接矩阵. 其中, 节点对 (i, j) 之间存在连接, 则 $a_{ij} = 1$; 否则, $a_{ij} = 0$; $\lambda_1, \lambda_2, \dots, \lambda_n$ 表示 A 的特征值, 且每个特征值 λ_i 对应的特征向量为 $a = (e_1, e_1, \dots, e_n)$, 其关系可表示为 $\lambda e_i = \sum_{j=1}^n a_{ij} e_j$, 则特征向量中心性定义为

$$C_e(i) = \lambda^{-1} \sum_{j=1}^n a_{ij} e_j \tag{8}$$

前面介绍的 7 种网络中心性判断方法都可以应用到实际网络中发现网络的中心节点, 但由于方法的原理和侧重不同, 对同一网络可能会出现不同的结果. 按照各判断方法的定义可以划分为三类: 关联性、最短路径和模拟流问题, 如图 1 所示. 关联性强调个体之间的直接连接, 最短路径问题表示节点传播途径始终选择最优方式, 而模拟流则侧重于对现实的一个模拟, 如引入物理分压的概念模拟网络流问题.

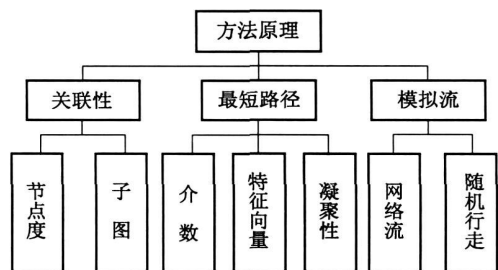


图 1 中心性判断方法原理分类图

Fig. 1 Centrality measures classifications

2 现实网络特征分析

网络中心性判断方法的差异主要源于研究者对网络本身特征和应用要求的不同侧重。前者主要关注网络结构特征,如网络的连通性;后者则更注重固定网络结构上产生的效率问题,如网络信息传播的深度和广度以及时间效率。因此,从衡量标准上存在节点在网络局部和全局贡献的差异,Newman 提出随机行走算法改变了以往对网络整体性质已知的假设,引发基于网络节点对全局信息缺失条件下的研究。另外,对于复杂网络特别是实际大型网络中心性研究,算法的时间复杂度也是十分重要的问题,现从这些角度对比分析这些方法的异同。

2.1 局域性

通常按照网络中心性方法依照考察节点连接方式不同划分为局部和全局方法^[12,13]。以节点度为主,考察节点直接连接,在社会网络中体现个体的影响力。以最短路为基础的方法,如介数、凝聚性以及网络流,都是考察网络全局性质。子图中心性方法是将局部信息扩大,体现整体的性质。在子图算法中如果把回路为1的权值设为1,其他回路的权值为0,那么子图中心性方法就同节点度是一致的。

科研合作网络、论文引用网络及语言网络等强调个体之间的连接,并通过个体的连接来确定节点的重要性,都适合使用局域性方法;而对病毒传播网络、通讯网络,交通网络及电力网络等,节点的特性能在多个连接之间传播,需要采用全局分析法。

2.2 网络信息

中心性方法中无论是节点度、凝聚性,还是介数,都是基于整个网络结构特征信息完全已知,从节点自身的连接到整个网络分布对决策者都是已知的,并且研究网络结构对实际网络的影响力、信息流及交通流等如何控制,属于静态的。而随机行走是针对节点拥有网络有限信息的研究,是对现实网络更真实的模拟,其算法思想来源于电路分压,即电路节点只对直接连接点已知。

社会学研究中的很多问题对网络结构特征获取是十分困难的,如社交网络,还有大规模网络上的局部问题,采用随机行走能有效地利用有限信息模拟网络性质。

2.3 动态性

网络的动态性主要体现在网络自身节点和边的

变化,主要形式有节点的增长、老化,边的添加、删除和改变连接。这些变化都会对网络的节点中心性产生直接的影响。这里讨论的动态性还涉及信息在网络节点之间传播的方式,节点度代表信息只存在与两个节点之间的传播,没有扩散性,而介数、网络流和随机行走都是典型的广播式传播。

3 中心性方法应用匹配体系和互补

3.1 应用匹配体系

对网络节点重要性的应用研究主要可以分为两类。一类是社会网络分析法,该方法假设网络的重要性等于节点的显著性,通过直接连接数来体现;另一类通过节点删除法来判断节点的重要性,其基础为若删除一个节点给网络带来的连通性破坏大,则该节点重要性高。在实际应用中,由于不同方法对节点的重要性排序结果不同,关键节点的发现效率存在差异,因此,对方法的选择是十分重要的。结合前面的分析,将网络结构特征和应用要求相结合,选取了三组关键的指标对实际网络进行划分,与方法之间建立一个匹配关系,如图2(见下页)所示。其中,网络指标分为局域性、动态性和信息透明度,结合三者可以对实际网络选取合理的方法。通常考虑节点局部连接、静态网络,并且网络结构已知,会选择度分布来衡量节点的重要性。

在对网络的实证研究中,重要性的衡量标准与网络实际应用是密切相关的,很多实证网络中用节点度和度分布来统计网络的性质,没有考虑应用的要求。如人际关系网络中,个体的大量连接可以凝聚较强的影响力,但是那些与影响力大的个体建立连接的人,因此也提升了自身的影响力,更典型的例子就是桥连接的重要性如何衡量。普通情况下,考虑算法实现的复杂度,如果只需要对网络的节点属性进行统计,可以采用节点度。介数、凝聚性是建立在最短路径的基础上,对于信息传播、交通网络及电力网络等,它们都是按照最优方案执行,并且在出现堵塞的情况下,会出现流量的扩散,导致整个网络受影响,这些网络中心性可以采用介数和凝聚性,他们之间的区别就在于应用中是强调个体在信息传播中承担的信息量,还是个体对信息广播的速度。如果实际网络不需要建立在最优方案的基础上,考虑全局信息,可以选择网络流和特征向量,强调网络中的可达性。子图和随机行走用于特别的实际需要,子图将局部信息有权值扩展到全局网络中,可以应用到

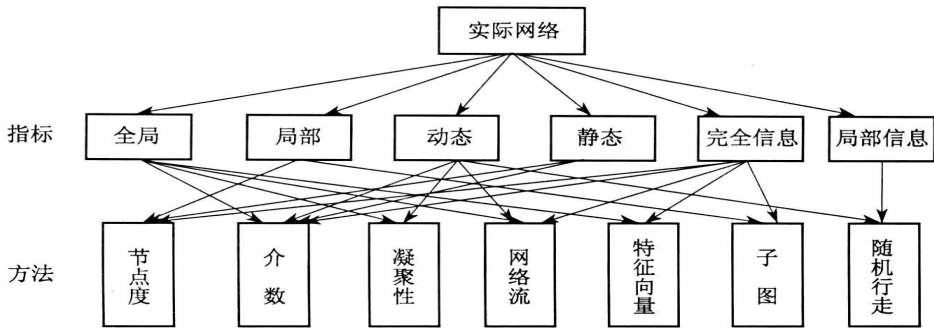


图 2 中心性方法应用匹配图

Fig. 2 Match of centrality measures and applications

表 1 算法性质和现实应用背景

Tab. 1 Property of algorithms and Application fields

名 称	局域性	网络信息	时间复杂度	现实意义	典型网络
节点度	局部	已知	$o(1)$	节点直接影响力	社交网络
介数	全局	已知	$o(n^3)$	节点对网络信息传播的控制能力	交通网络
凝聚性	全局	已知	$o(n^2)$	节点为源信息在网络中广播时间	通讯网络
网络流	全局	已知	$o(n^3)$	节点在网络中的拓扑中心性	航空网络
随机行走	全局	未知	$o((m+n)n^2)$	网络特征不全的信息传播	计算机网络
特征向量	局部	已知	$o(n^2)$	网络节点间接影响力	传染病网络
子图	局部	已知	$o(n^2)$	节点在网络局部参与程度	蛋白质网络

社团和蛋白质网络中衡量元素的参与程度, 如表 1 所示.

3.2 互补性

尽管不同方法在对同一网络进行分析的时候可能出现不同的结果, 但是, 各种方法之间并不是完全独立的, 它们之间存在不同强度的关联性. 在经典网络模型中, 节点度和介数有很强的关联性, 除少数桥节点, 其他节点序相对一致; 在现实网络的实证研究中, 对蛋白质网络、科研合作网络和演员合作网络, 各种方法之间显示了一定的关联性, 但较经典网络模型要弱. 因此, 实际网络考察中需要综合利用不同方法, 或者提出一些相互之间转换的机制, 如将网络桥节点按照所连接区域最大度数转换成为其连接度, 将介数统一到节点度中, 既能突出特殊节点, 又能保证计算简单.

4 结 论

复杂网络的中心性判断方法由于定义中对网络结构关注的侧重不同, 因此会出现对同一现实网络出现不同的分析结果, 并且现实网络对节点的重要性判断是根据应用要求而定, 本文就判断方法和实际网络之间建立了一个匹配体系, 抽取实际网络中

的局域性、动态性和信息透明度为具体指标, 与中心性判断方法原理中的节点关联、最短路径和网络模拟相对应, 并且探讨了如何将不同方法结合应用, 满足实际需求. 由于中心性判断方法不是完全独立的, 所以, 需要进一步研究不同方法之间的关联性, 探讨方法之间的整合.

参考文献:

[1] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439): 509—512.

[2] ALBERT R, JEONG, H, BARABÁSI A L. Error and attack tolerance of complex networks[J]. Nature, 2000, 406(6794): 378—382.

[3] ALBERT R, JEONG H, BARABASI A L. Diameter of the world-wide web[J]. Nature, 1999, 401(6749): 130—131.

[4] FREEMAN L C. A set of measures of centrality based upon betweenness[J]. Sociometry, 1977, 40(1): 35—41.

[5] FREEMAN L C. Centrality in social networks: ii. Experimental results[J]. Social Networks, 1979, 2(2): 119—141.

(下转第 236 页)

扩张可以用来研究子系统相互之间的拓扑关系, 腐蚀可以用来研究系统内部要素的层次关系, 开启可以忽略一些小系统的影响, 闭合可以连接小的系统进入大的系统. 通过不同的扩张、腐蚀、开启和闭合形态操作, 可以得到不同尺度下的系统拓扑结构, 便于人们从微观到宏观全角度来研究复杂网络.

参考文献:

[1] COSTA L D F, ROEHA L E. A generalized approach to complex networks [EB/ OL]. <http://arxiv.org/PS>

[6] FREEMAN L C. Centrality in social networks: Conceptual clarification[J]. *Social Networks*, 1979, 1(3): 215—239.

[7] BRANDES U. A fast algorithm for betweenness centrality[J]. *Journal of Mathematical Sociology*, 2001, 25(2): 163—177.

[8] NEWMAN M E J. A measure of betweenness centrality based on random walk[J]. *Social Networks*, 2005, 27(1): 39—54.

[9] ESTRADA E, RODRIGUEZ-VELAZQUEZ J A. Subgraph centrality in complex networks[J]. *Physical Review E*, 2005, 71(5): 1—9.

[10] GOH K, OH E, KAHNG B, et al. Betweenness centrality

cache/ cond-mat/pdf/0408/0408076v2. pdf, 2004—08—03.

[2] COSTA L D F, SILVA F N. Hierarchical characterization of complex networks.[EB/ OL]. <http://arxiv.org/PS> cache/ cond-mat/pdf/0401/0412761v4. pdf, 2004—12—30.

[3] COSTA L D F, RODRIGUES F A, TRAVIESO G, et al. Characterization of complex networks: A survey of measurements. [EB/ OL]. <http://arxiv.org/PS> cache/ cond-mat/pdf/0505/0505185v5. pdf, 2005—05—07.

(上接第 230 页)

correlation In social networks[J]. *Physical Review E*, 2003, 67(1): 017101, 1—4.

[11] BARTHELEMY M. Betweenness centrality in large complex networks[J]. *European Physical Journal B*, 2004, 38(1 434): 163—168.

[12] JUNGSLUTH M, BURGHARDT B, HARTMANN A K. What do we learn from correlation of local and global network properties[J]. *Physics Soc*, 2006, 7(10): 150—158.

[13] ALBERT R, BARABÁSI A L. Topology of evolving networks: Local events and universality[J]. *Phys Rev Lett*, 2000, 85(24): 5 234—5 237.