

基于自规避随机游走的节点排序算法

段杰明 尚明生 蔡世民 张玉霞

A ranking method based on self-avoiding random walk in complex networks

Duan Jie-Ming Shang Ming-Sheng Cai Shi-Min Zhang Yu-Xia

引用信息 Citation: *Acta Physica Sinica*, 64, 200501 (2015) DOI: 10.7498/aps.64.200501

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.200501>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I20>

您可能感兴趣的其他文章

Articles you may be interested in

负荷作用下相依网络中的级联故障

Load-induced cascading failure in interdependent network

物理学报.2015, 64(4): 048901 <http://dx.doi.org/10.7498/aps.64.048901>

基于局部路由策略的复杂网络拥塞控制

Congestion control in complex network based on local routing strategy

物理学报.2014, 63(24): 248901 <http://dx.doi.org/10.7498/aps.63.248901>

在线社交网络中谣言的传播与抑制

The propagation and inhibition of rumors in online social network

物理学报.2012, 61(23): 238701 <http://dx.doi.org/10.7498/aps.61.238701>

无标度立体 Koch 网络上随机游走的平均吸收时间

Exact solution for mean trapping time of random walk on a scale-free Koch network

物理学报.2012, 61(20): 200503 <http://dx.doi.org/10.7498/aps.61.200503>

混沌半导体激光器的弛豫振荡频率对随机序列速率的影响

The effect of the relaxation oscillation frequency of chaotic semiconductor laser on the rate of random sequence

物理学报.2012, 61(15): 150502 <http://dx.doi.org/10.7498/aps.61.150502>

基于自规避随机游走的节点排序算法^{*}

段杰明¹⁾ 尚明生^{1)2)†} 蔡世民¹⁾²⁾ 张玉霞³⁾

1)(电子科技大学计算机科学与工程学院, 成都 611731)

2)(电子科技大学大数据研究中心, 成都 611731)

3)(华南理工大学物理与光电学院, 广州 510640)

(2015年4月8日收到; 2015年6月2日收到修改稿)

评估复杂网络系统的节点重要性有助于提升其系统抗毁性和结构稳定性. 目前, 定量节点重要性的排序算法通常基于网络结构的中心性指标如度数、介数、紧密度、特征向量等. 然而, 这些算法需要以知晓网络结构的全局信息为前提, 很难在大规模网络中实际应用. 基于自规避随机游走的思想, 提出一种结合网络结构局域信息和标签扩散的节点排序算法. 该算法综合考虑了节点的直接邻居数量及与其他节点之间的拓扑关系, 能够表征其在复杂网络系统中的结构影响力和重要性. 基于三个典型的实际网络, 通过对极大连通系数、网络谱距离数、节点连边数和脆弱系数等评估指标的实验对比, 结果表明提出的算法显著优于现有的依据局域信息的节点排序算法.

关键词: 复杂网络系统, 节点排序, 自规避随机游走, 局域信息

PACS: 05.40.Fb, 87.23.Ge, 89.75.Fb

DOI: 10.7498/aps.64.200501

1 引言

随着社会网络化和经济全球化, 关乎国计民生的复杂网络系统如电力网、世界经济网, 其数据规模与日俱增且内在结构日益复杂. 2003年, 美国俄亥俄州少数电站失效引发了北美大停电事故, 造成上千万人陷入黑暗, 经济损失估计高达百亿美元^[1]. 2008年, 美国次级贷款危机的爆发引发全球化的经济危机, 导致世界经济的大衰退. 这些事件的发生使得人们对复杂网络系统的稳定性和安全性越来越重视. 因此, 能够有效地发现这些大规模网络系统中具有结构影响力的重要节点, 对提高网络系统的鲁棒性与抗毁性有重要的现实意义.

复杂网络研究的蓬勃发展^[2-7], 使发现重要节点的相关研究得到了进一步的认识. 重要节点可以是指拓扑结构中的核心节点, 也可以是指动态传播过程中传播效率最高的节点^[8]. 通常来讲,

基于网络结构的中心性指标如度数、介数、紧密度、特征向量等能够有效地评估节点重要性或核心程度^[9-13]. Ghoshal和Barabási^[14]分析了网络结构对PageRank算法的影响, 并强调了其中超稳定节点所起的作用. Zhao等^[15]则从图信号处理的角度分析了网络中节点的重要性, 并使用传播动力学的方法进行了验证. Chen等^[16]依据网络中节点的直接邻居和次邻居的度数大小, 提出一种基于多级邻居信息进行节点排序的算法. Zhang等^[17]利用连边两端节点的度数大小, 提出了新的节点排序算法. Cheng等^[18]提出了基于最大团的Bridgeness排序算法, 但同样存在计算复杂等缺点. Tang等^[19,20]利用网络上节点的多源传播影响力来识别节点重要性.

然而, 不少研究成果^[13,21,22]指出, 若要发现网络中具有结构影响力的重要节点, 应该综合考虑节点的度数以及其与网络上其余节点间的结构关系. 尽管多数算法考虑了网络结构的因素, 却需要预知

^{*} 国家自然科学基金(批准号: 61370150, 61433014, 71490720)和中央高校基本科研业务费(批准号: 2014ZM0079)资助的课题.

[†] 通信作者. E-mail: msshang@uestc.edu.cn

网络结构的全局信息, 这对于当前结构复杂且规模巨大的网络来说, 未必是合适的有效方法. 于是, 有学者提出依据局域信息的节点排序算法, 其中最具有代表性的是熟人算法 (acquaintance algorithm), 实验结果证实其在无标度网络中表现优异^[23,24]. 熟人算法通过随机选点然后再次随机选择其邻居的方法, 避免了对全局信息的依赖, 但由于其只考虑了节点的直接邻居情况, 所以无法反映出节点在整个网络中的结构特性. 因此, 本文基于自规避随机游走, 提出一种结合局域信息和标签游走的节点排序算法 (tag walking algorithm, 简称标签算法). 该算法通过节点向直接邻居投放标签, 基于自规避随机游走实现标签转发过程, 利用标签的游走路径测量它与其他节点的结构关系, 评估节点在整体网络中的结构影响力. 基于三个典型的实际网络, 本文采用极大连通系数、网络谱距离数、节点连边数及脆弱系数等指标, 对节点的排序结果展开比较与分析. 实验结果验证了本文提出的算法可以更有效地发现网络上具有结构影响力的重要节点.

2 理论算法

表征网络拓扑结构的信息可分为局域信息和全局信息^[25]. 基于全局信息的节点排序算法, 由于需要知晓全部网络拓扑结构, 因而适用范围有限. 例如因特网, 其节点规模巨大且动态变化, 全局信息往往并不可知.

随机游走是指信息从一个或多个顶点出发向整个网络进行扩散, 也就是在每个节点以一定的概率随机地游走到它的邻接节点. 随机游走模型已经被广泛地应用在各个领域, 可用于刻画实际复杂网络系统中的动力学过程, 定量表征节点的重要性^[14]. 例如, 前述电力网络中, 某处电站作为电能中转站的职能失效, 无法实现电能流通, 造成了电力网络的大面积停电 (但不是整个电力网络), 因此我们可以认为该电站是电力网络中具有显著结构影响力的节点. 类似地, 把网络中流动的对象都抽象为标签, 在扩散过程中节点收到标签的数量, 能够用来表征其对网络结构的影响力, 从而判定节点在网络中的重要性.

为此, 本文提出的标签算法将依据网络局域信息, 兼顾与网络其余节点的结构关系, 利用基于自规避随机游走的标签扩散, 定量分析节点的重要性. 自规避随机游走是指在每一步游走过程中, 不

再选择那些已经到达过的节点, 相对于传统随机游走, 可以缩短扩散时间, 并且避免在扩散过程中形成区域范围内的标签扩散闭环. 标签算法通过随机选择节点, 并向其直接邻居转发标签, 根据自规避随机游走实现标签在网络其余节点间的转发过程, 综合考虑了节点的直接邻居数量以及和网络其余节点的拓扑关系, 由此测量节点在整个拓扑结构中所发挥的结构影响力. 图 1 给出标签传递的示意图, 其流程如下:

- (a) 选取网络节点 i , 以广播形式向直接邻居转发标签信息 l ;
- (b) 邻居节点如果首次收到该标签信息 l , 则接受并记录该标签, 接着产生一个随机概率 p , 如果 p 小于传递系数 t ($t \in (0, 1)$), 那么同样以广播形式转发该标签 l ;
- (c) 如果其邻居节点已收到过标签 l , 则丢弃;
- (d) 当网络中再无标签转发行为时, 一次试验结束.

依据上述流程, 独立运行 n 次试验, 计算节点接收到的标签数量并进行排序. 节点接收到的标签数量越多, 说明其结构影响力越大.

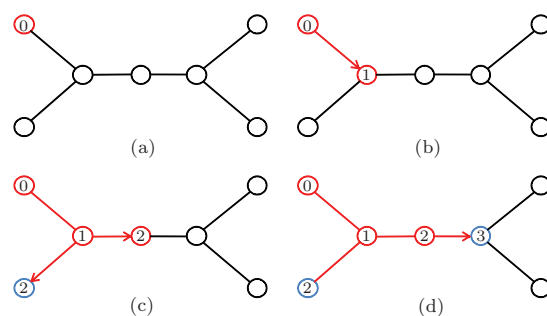


图 1 (网刊彩色) 标签算法示意图. 红色表示已经收到标签且选择转发的节点, 蓝色表示收到标签且选择不转发的节点, 箭头表示在每一步运行过程中标签的传播方向, 其中, 节点 0 为随机选择的初始节点.

Fig. 1. (color online) Sketch of the tag walking algorithm. The node with red color represent node that received and transmitted the tag, while blue means received but not transmitted, the arrow indicates the direction of propagation of the tag, node 0 is the source.

在上述算法中传递系数 t 作为可调参量, 对于每个网络而言在算法运行过程中是个确定值. 由于 t 值对标签在网络上的扩散程度有直接的影响, 过小的 t 值使得标签有可能只扩散到直接邻居, 而过大的 t 值又使得网络中每个节点都有可能接受到标签, 因此合理地选取 t 值有助于提高标签算法的有效性.

值得注意的是标签算法不同于基于中心性的排序算法. 首先, 本文算法是一种随机化算法, 没有确定性地利用所有网络结构信息, 不能够准确地描述算法的时间复杂度. 其次, 标签扩散与网络结构局域信息的结合, 主要体现在两个方面: 1) 由于网络结构的限制, 标签并不会扩散到整个网络, 更多的是在几层邻居或网络的局部区域之内; 2) 对于算法中的多次随机选点过程, 可能存在部分网络中的节点从未被选中或某些节点多次被选中, 但它们的结构重要性依然会在扩散过程中体现出来, 具体的表现就是收到标签的数目.

3 节点结构重要性评价方法

对于重要节点的定义, 不同文献不尽相同 [24, 26–28]. 本文采用四个常用的指标评估节点的结构重要性, 即极大连通系数、谱距离、节点连边数和脆弱性.

1) 极大连通系数: 可以用来量化节点在网络中的结构重要性, 即移除节点后网络的极大连通图规模与网络规模的比率 [29], 计算公式如下:

$$G = n_g/n, \quad (1)$$

其中, n 表示网络总节点数, n_g 表示移除部分节点后剩余网络中极大连通子图的节点数.

2) 谱距离: 对于移除节点后剩余的非连通网络则采用谱距离 L 来表示网络中节点间的距离, 也就是平均最短路径长度 [30]. 谱距离越大说明破碎后的网络节点间距离越远, 则破碎效果越好. 谱距离计算公式如下:

$$L = 1 / \left[\frac{1}{n(n-1)} \times \sum_{i \neq j} \frac{1}{d_{ij}} \right], \quad (2)$$

其中, n 表示网络节点数, i 和 j 表示网络中的节点, d_{ij} 表示节点 i 与 j 之间的最短路径长度.

3) 节点连边数: 对于被移除节点的度数大小则可以用网络中累积移除的边数来表示, 计算公式

如下:

$$\sum m = m - m_i, \quad (3)$$

其中, m 表示网络总边数, i 表示网络中的节点, m_i 表示移除 i 后网络中剩余的边数.

4) 脆弱性: 通常来讲, 可以用脆弱性来衡量网络中节点失效或者遭受攻击后引起的整个网络系统性能的下降 [30–32]. 脆弱性与连通效率密切相关 [33]. 节点脆弱系数 V 验证移除节点后网络连通效率发生变化的情况, 计算公式如下:

$$V = \frac{E - E_i}{E}, \quad (4)$$

其中, E 反映的是当前网络连通效率, i 表示网络中的节点, E_i 表示移除 i 后的网络连通效率, 它们都可通过计算谱距离 L 的倒数求取, 具体见 (2) 式.

4 实验结果与分析

4.1 实验数据

为了验证标签算法的效果, 本文选取三个典型的网络: 美国国家电力网 (Power)、线虫脑细胞网络 (C. elegans) 和科学家合著关系网 (Netcoauthor). 表 1 给出了这三个网络的拓扑结构统计特征. 其中, n 和 m 分别表示网络节点数和连边数; $\langle k \rangle$ 表示节点平均度; k_{\max} 表示节点最大度; $\langle d \rangle$ 表示网络节点之间的最短路径平均数; C 表示聚类系数 [3], 用来评估节点聚集成团的程度; R 表示同配系数 [33], 用于反映邻接节点间的度相关性; H 表示异质系数 [34], 反映的是网络节点度负相关性. 可以看出, C. elegans 和 Netcoauthor 网络拥有明显的小世界特性, 即高聚类系数与短的平均最短路径长度, 同时 C. elegans 网络拥有明显的大度节点且具有高异质性, 因此它同时也具有无标度网络的特性. 相反, Power 网络聚类系数较低, 并且它的平均度和最大度都比较小, 可以认为它更接近于随机网络. 所以, 三个典型的网络表征不同的网络结构, 能够更广泛地评估标签算法的有效性.

表 1 网络拓扑结构的统计特征
Table 1. Structural properties of three empirical networks.

网络	n	m	$\langle k \rangle$	k_{\max}	$\langle d \rangle$	C	R	H
Power	4941	6594	2.66	19	18.98	0.080	0.003	0.324
C. elegans	453	2040	9	239	2.66	0.646	-0.219	0.492
Netcoauthor	379	914	4.82	34	6.04	0.741	-0.081	0.373

4.2 实验结果

基于上述三个网络, 本文对标签算法与同样是采用局域信息的熟人算法进行了比较与分析. 在模拟实验过程中依据两种算法分别取得排序后的结果, 用参数 f 代表降序节点集所占网络规模的比例, 从而评估网络上重要节点的准确性. 这里, 由于 Power 和 Netcoauthor 的平均度较小 (表 1), 它们所选择的传播系数 t 应取比较大的值, 只有这样才可以使得标签在整个网络范围内传播, 相反 C. elegans 的平均度明显大于另外两个网络, 因此它的传播系数应该较小. 经实验分析传递系数 t 得到三个网络 Power, C. elegans 和 Netcoauthor 的最优取值

分别为 0.9, 0.3, 0.9.

依据确定的传递系数, 对标签算法和熟人算法分别进行 10^6 次独立重复实验. 然后, 基于四个评估节点结构重要性的指标, 我们对标签算法和熟人算法进行对比分析. 值得注意的是, 文献 [23,24] 已经验证熟人算法优于基于中心性指标的节点排序算法, 在此仅以熟人算法作为基准.

图 2 反映的是不同的网络中, 算法排序后移除的节点比例与极大连通系数的关系. 如图 2 所示, 横坐标 f 表示的是移除比例, 纵坐标 G 表示的是移除比例 f 节点后的网络极大连通系数, 其中 TAG 与 ACQ 分别表示标签算法和熟人算法 (下同). 数值反映了网络结构完整情形与极大连通图规模的

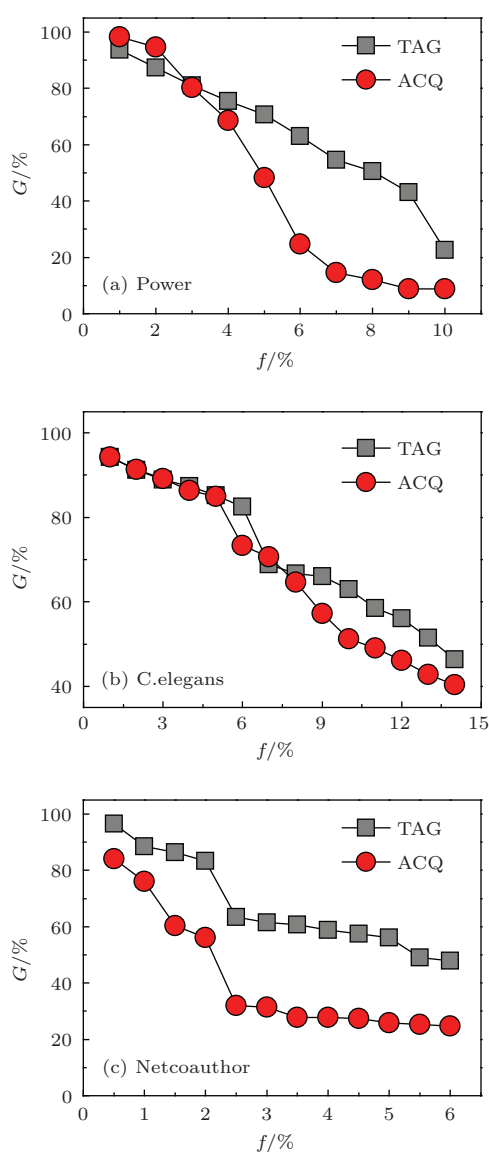


图 2 (网刊彩色) 极大连通系数随节点移除比例的变化
Fig. 2. (color online) Giant components as a function of node removed ratio f .

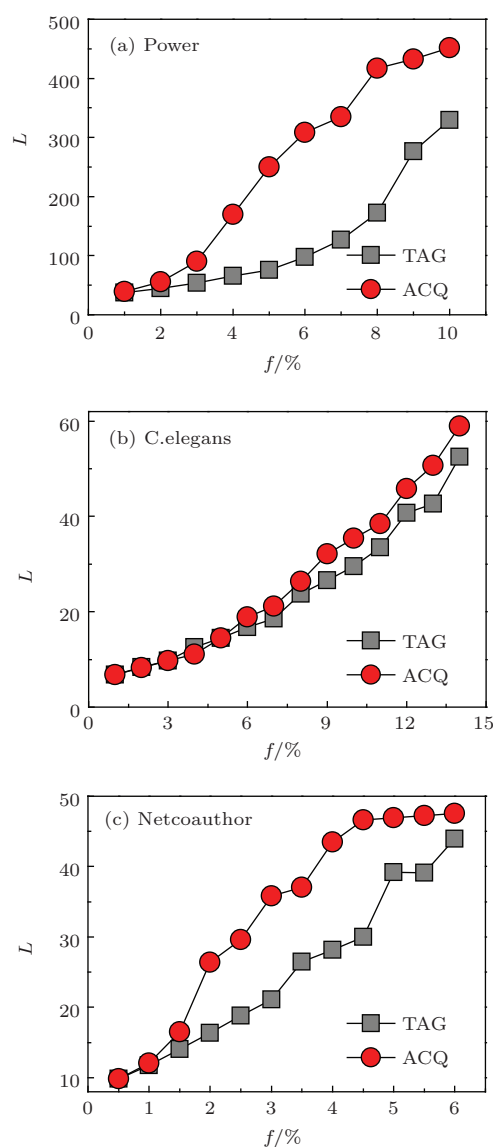


图 3 (网刊彩色) 网络谱距离数随节点移除比例的变化
Fig. 3. (color online) Spectral distance as a function of node removed ratio f .

正关联性, 即极大连通图规模越小, 网络结构完整性越差. 观察可见, 在 Power, C. elegans 和 Netcoauthor 中, 标签算法可以更好地发现网络上具有结构影响力的节点, 而此类节点起到了明显的全局影响作用, 可以更容易地破碎整个网络的拓扑结构, 使得移除少量节点即可产生结构性溃散的效果.

对于移除节点后造成网络结构发生的破碎效果, 同样反映的是网络中任意节点间的距离疏远变化. 换言之, 移除重要节点后的网络碎片化现象越明显, 网络连通性越差, 于是网络的谱距离越大. 如图 3 所示, 比较了标签算法与熟人算法对谱距离

的实验结果, 横坐标 f 表示的是移除比例, 纵坐标 L 表示的是网络的谱距离平均数随移除比例 f 的变化. 观察不难发现, 标签算法考虑了节点与网络其余节点的结构关系, 使得排序的综合结果更加合理, 其在 Power 和 Netcoauthor 中的效果更加显著.

对于熟人算法而言, 重要节点意味着度数最大的节点 [23,24], 即与度数中心性具有正相关性. 这类节点通常反映了局域结构, 刻画了节点对其直接邻居的影响. 图 4 描述的正两种采用局域信息的算法在度数特征方面的差异, 其中横坐标 f 表示的是节点移除比例, 纵坐标 $\sum m$ 表示的是节点连边数随移除比例 f 的变化. 观察结果可见, 熟人算法对

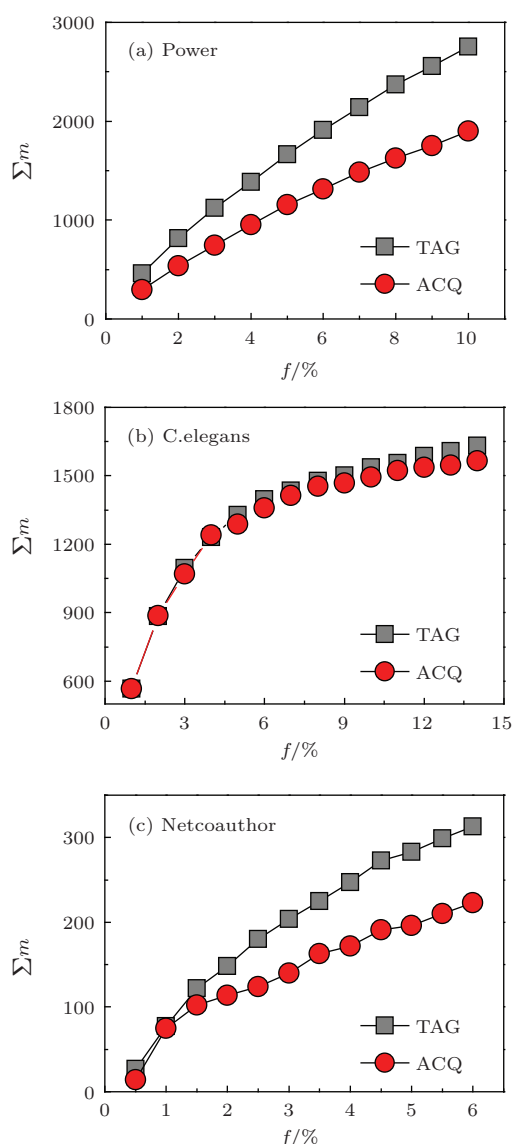


图 4 (网刊彩色) 连边累积数随节点移除比例的变化

Fig. 4. (color online) Cumulative edges among nodes as a function of node removed ratio f .

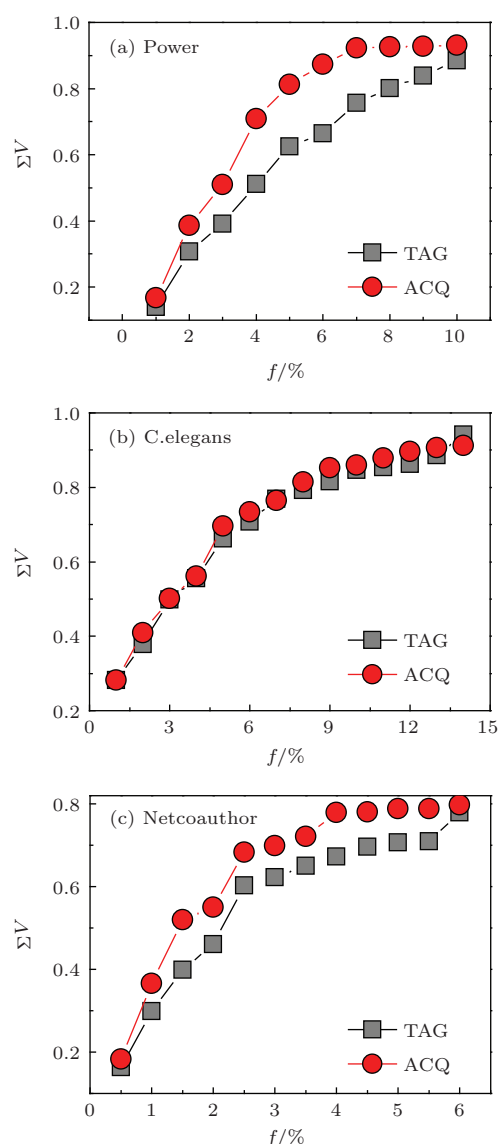


图 5 (网刊彩色) 累计脆弱系数随节点移除比例的变化

Fig. 5. (color online) Cumulative fragility among nodes as a function of node removed ratio f .

重要节点的发现更倾向于网络中的大度节点, 而标签算法在三个网络中发现的重要节点在度数方面并未表现出鲜明的特征. 这里, 标签算法兼顾了节点度数的局域影响力以及对网络上其余节点的结构关系, 因此能够更加合理地反映出节点对整体结构所产生的影响作用.

重要节点之所以能够在网络结构方面发挥关键性作用, 不仅是节点度数具有的连通优势, 更关键的是与网络中其余节点间存在的拓扑关系. 那么, 移除重要节点必定给网络结构造成更大的伤害, 这点可以通过脆弱系数来表现. 图5比较了标签算法与熟人算法的节点脆弱系数的实验效果, 横坐标 f 表示的是移除比例, 纵坐标 $\sum V$ 表示的是脆弱系数的累积值随移除比例 f 的变化. 在三个网络中, 实验结果都证实了标签算法发现的节点在网络结构方面显得更加“脆弱”. 在此, 极大连通系数、谱距离数以及脆弱系数等都表现出标签算法整体优于熟人算法的结果, 同时也验证了我们所提出算法的有效性.

4.3 相关性分析

图6反映的是实际网络中节点的中心性指标、脆弱系数以及通过实验计算得到的标签次数三者间所表现的关联. 这里的横坐标代表节点的中心性指标, 纵坐标代表节点通过算法计算得出的标签次数, 圆点颜色代表节点的脆弱系数大小. K壳分解算法曾提出节点重要性应该依赖于整个网络结构或者位置的思想, 同时指出度数和介数在不少场合无法有效地刻画节点重要性^[13]. 从图6结果中不难发现度数或者介数在衡量节点的结构脆弱性方面不太准确, 多数节点的度数或介数与标签次数的差异偏小(密集且无法区分), 所以曲线未能表现出近似规律. 另一方面, 结果也显示了标签算法与紧密度存在较明显的关联现象(近似规律), 说明该算法所利用的自规避随机游走过程, 考虑了与网络上其余节点的拓扑关系, 所以能够较好地发现对网络结构具有影响作用的节点.

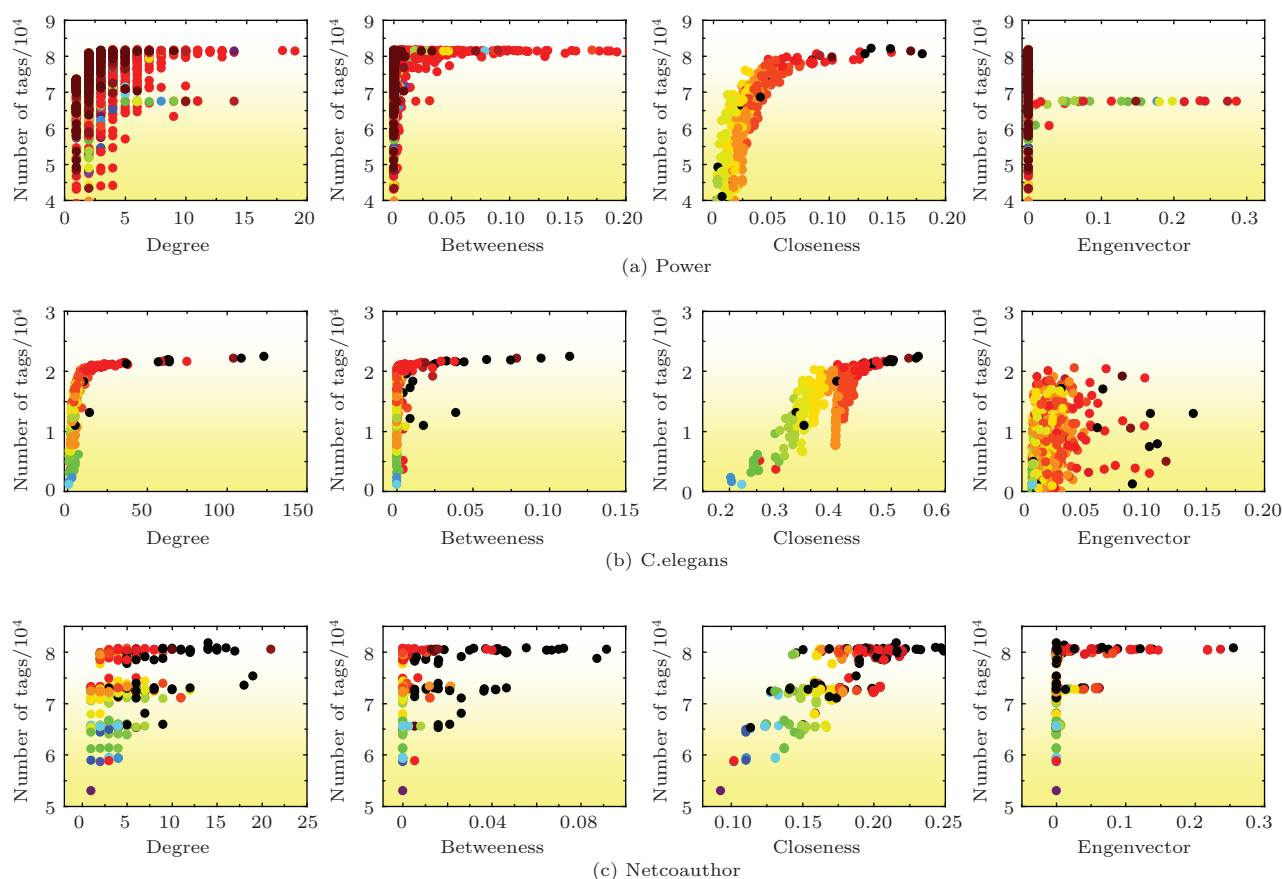


图6 (网刊彩色) 节点中心性、标签次数以及节点脆弱系数三者间的关联

Fig. 6. (color online) Correlations among centralities, count obtained from tag avoid-walking algorithm and fragility in empirical networks.

5 结 论

发现复杂网络系统中那些具有结构影响力的重要节点, 对于提升其系统抗毁性和结构稳定性有重要作用, 例如, 应对电力网故障或流行病大爆发等. 面对规模不断增大的复杂网络系统, 怎样利用网络局域信息来有效地测量节点在系统中所发挥的结构重要性, 是摆在科学研究与工程应用方面所面临的难题.

为此, 本文基于自规避随机游走的思路, 综合考虑节点的直接邻居数量以及与网络上其余节点间的结构关系, 提出一种基于自规避随机游走, 结合局域信息和标签扩散的节点排序算法. 以 Power, C. elegans 和 Netcoauthor 三个典型网络为基础, 通过极大连通系数、网络谱距离数、节点度数与脆弱系数等指标验证了算法的有效性和适用性. 与熟人算法的实验结果比较可知, 标签算法能够更好地发现网络系统中具有结构影响力的节点.

参考文献

- [1] Kinney R, Crucitti P, Albert R, Latora V 2005 *Eur. Phys. J. B* **46** 101
- [2] Strogatz S H 2001 *Nature* **410** 268
- [3] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [4] Barabasi A L, Albert R 1999 *Science* **286** 509
- [5] Lü L, Medo M, Yeung C H, Zhang Y C, Zhang Z K, Zhou T 2012 *Phys. Rep.* **519** 1
- [6] Albert R, Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
- [7] Liang Z W, Li J P, Yang F, Petropulu A 2014 *Chin. Phys. B* **23** 098902
- [8] Liu J G, Ren Z M, Guo Q, Wang B H 2013 *Acta Phys. Sin.* **62** 178901 (in Chinese) [刘建国, 任卓明, 郭强, 汪秉宏 2013 物理学报 **62** 178901]
- [9] Newman M 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
- [10] Freeman L 1977 *Sociometry* **40** 35
- [11] Sabidussi G 1966 *Psychometrika* **31** 581
- [12] Stephenson K, Zelen M 1989 *Soc. Networks* **11** 1
- [13] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, Makse H A 2010 *Nat. Phys.* **6** 888
- [14] Ghoshal G, Barabási A L 2010 *Nat. Commun.* **2** 394
- [15] Zhao J, Yu L, Li J R, Zhou P 2015 *Chin. Phys. B* **24** 058904
- [16] Chen D, Lü L, Shang M S, Zhang Y C, Zhou T 2012 *Physica A* **391** 1777
- [17] Zhang H F, Li K Z, Fu X C, Wang B H 2009 *Chin. Phys. Lett.* **26** 068901
- [18] Cheng X Q, Ren F X, Shen H W, Zhang Z K, Zhou T 2010 *J. Stat. Mech.* **20** 595
- [19] Zhao Y X, Huang B, Tang M, Zhang H F, Chen D B 2014 *EPL* **108** 68005
- [20] Liu Y, Tang M, Zhou T, Do Y 2015 *Sci. Rep.* **5** 9602
- [21] Hu Q, Gao Y, Ma P, Yin Y, Zhang Y, Xing C 2013 *Web-Age Information* (Berlin: Springer Berlin Heidelberg) pp99–104
- [22] Ren X L, Lü L Y 2014 *Chin. Sci. Bull.* **59** 1175 (in Chinese) [任晓龙, 吕琳媛 2014 科学通报 **59** 1175]
- [23] Cohen R, Havlin S, Ben-Avraham D 2003 *Phys. Rev. Lett.* **91** 247901
- [24] Cohen R, Erez K, Ben-Avraham D, Havlin S 2001 *Phys. Rev. Lett.* **86** 3682
- [25] Salathé M, Jones J H 2010 *PLoS Comput. Biol.* **4** e1000736
- [26] Albert R, Jeong H, Barabasi A L 2000 *Nature* **406** 378
- [27] Muff S, Rao F, Cafilisch A 2005 *Phys. Rev. E* **72** 056107
- [28] Dorogovtsev S N, Mendes J F 2001 *Phys. Rev. Lett.* **87** 219801
- [29] Holme P, Kim B J, Yoon C N, Han S K 2002 *Phys. Rev. E* **65** 056109
- [30] Costa L F, Rodrigues F A, Traviesso G, Villas Boas P R 2007 *Adv. Phys.* **56** 167
- [31] Latora V, Marchiori M 2001 *Phys. Rev. Lett.* **87** 198701
- [32] Latora V, Marchiori M 2003 *Eur. Phys. J. B* **32** 249
- [33] Moreno Y, Nekovee M, Vespignani A 2004 *Phys. Rev. E* **69** 055101
- [34] Newman M E 2002 *Phys. Rev. Lett.* **89** 208701

A ranking method based on self-avoiding random walk in complex networks*

Duan Jie-Ming¹⁾ Shang Ming-Sheng^{1)2)†} Cai Shi-Min¹⁾²⁾ Zhang Yu-Xia³⁾

1) (School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

2) (Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China)

3) (Physics and Photoelectricity School, South China University of Technology, Guangzhou 510640, China)

(Received 8 April 2015; revised manuscript received 2 June 2015)

Abstract

Evaluation of node importance is helpful to improve the invulnerability and robustness of complex networked systems. At present, the classic ranking methods of quantitatively analyzing node importance are based on the centrality measurements of network topology, such as degree, betweenness, closeness, eigenvector, etc. Therefore, they often restrict the unknown topological information and are not convenient to use in large-scale real networked systems. In this paper, according to the idea of self-avoiding random walking, we propose a novel and simplified ranking method integrated with label propagation and local topological information, in which the number of labels that node collects from propagating process quantitatively denotes the ranking order. Moreover, the proposed method is able to characterize the structural influence and importance of node in complex networked system because it comprehensively considers both the direct neighbors of node and the topological relation of node to other ones. Through performing the experiments on three benchmark networks, we obtain interesting results derived from four common evaluating indices, i. e. , the coefficient of giant component, the spectral distance, the links of node, and the fragility, which indicate that the proposed method is much more efficient and effective for ranking influential nodes than the acquaintance algorithm.

Keywords: complex networks, node ranking, self-avoiding random walk, local information

PACS: 05.40.Fb, 87.23.Ge, 89.75.Fb

DOI: 10.7498/aps.64.200501

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61370150, 61433014, 71490720) and the Fundamental Research Funds for the Central Universities, China (Grant No. 2014ZM0079).

† Corresponding author. E-mail: msshang@uestc.edu.cn