

基于分形特性的复杂网络全局效率估计方法

张胜, 戴维凯, 吴锋, 蓝文祥

(南昌航空大学信息工程学院, 江西 南昌 330063)

摘要: 针对大型网络中效率计算时间复杂度高、计算耗时长的问題, 提出一种基于分形特性的网络效率估计方法。利用复杂网络拓扑结构的分形特性, 分析网络效率与节点关联和的关系, 用部分节点关联和来估计网络全局效率。此外, 为了快速判断复杂网络的分形特性, 提出基于节点关联和的分形特性判别方法。在构造网络和真实网络中进行实验分析, 结果表明, 所提方法能准确有效地估算网络全局效率, 比原始的网络效率计算方法可缩减不低于90%的计算时间。

关键词: 复杂网络; 网络效率; 分形; 关联和

中图分类号: O157.5

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2020118

Global efficiency estimation method of complex network based on fractal property

ZHANG Sheng, DAI Weikai, WU Feng, LAN Wenxiang

School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China

Abstract: A method of network efficiency estimation based on fractal property was proposed for solving the problems of high complexity and time-consuming calculation of efficiency in large scale networks. Considering the fractal properties of complex network topology, the relationship between network efficiency and node correlation sum was analyzed, and the global network efficiency was estimated by partial nodes. Besides, to rapidly find the fractal properties of complex networks, a fractal property discrimination method based on node correlation sum was proposed. The experimental analysis in the construction network and the real-world network show that the proposed method can accurately and effectively estimate the global efficiency of the network, which reduces the calculation time by at least 90% compared with the original method.

Key words: complex network, network efficiency, fractal, correlation sum

1 引言

复杂网络是具有非平凡拓扑特征的图, 能够表征各种各样的复杂系统。真实复杂系统的图模型中常常表现出小世界、无标度等拓扑特征^[1]。网络科学中全局效率是衡量信息交换效率的指标^[2]。最初提出全局效率是用来度量复杂网络中的小世界特

征, 使“小世界性”有了明确的物理意义和定量分析的标准。后来这一概念逐渐应用到多个领域, 如社会网络中寻找疫情传播过程中的重要节点^[3]; 生物网络中发现高智商青年人的脑网络具有更高的全局和局部网络效率, 以此评估脑网络的完整性^[4-5]; 通信网络中用于评估网络拓扑的稳健性^[6]; 交通网络中利用网络全局效率分析地铁系统的传输效率和

收稿日期: 2020-01-15; 修回日期: 2020-05-13

基金项目: 国家自然科学基金资助项目 (No.61661037); 江西省教育厅科技基金资助项目 (No.GJJ170575); 南昌航空大学研究生创新专项资金资助项目 (No.YC2018018)

Foundation Items: The National Natural Science Foundation of China (No.61661037), Science and Technology Project of Jiangxi Province Education Department (No.GJJ170575), Nanchang Hangkong University Graduate Innovation Special Foundation (No.YC2018018)

规划维护道路网络^[7-8]; Web 网络中使用网络效率解释信息传输的方式^[9]; 电力网络中用来分析系统结构的脆弱性^[10]。在无权和加权的复杂网络中, 有研究者利用网络效率计算节点中心性^[11-12]。

尽管网络全局效率作为信息交换效率的衡量指标应用广泛、作用巨大, 但其计算十分复杂、耗时, 且计算过程中最短路径矩阵的理论计算时间复杂度达到 $O(N^3)$ ^[13]。

大量研究发现, 许多复杂网络在结构上具有分形特性, 其直观地表现为网络的整体与局部间存在自相似性。Song 等^[14-15]分析了各种真实的复杂网络, 并发现它们在所有长度尺度上都由自我重复的模式组成, 揭示了复杂网络的分形与自相似特性。在求得一定长度尺度信息后, 通过其自相似性可以推测出其他尺度上的模式。此后, 研究者们提出多种分形维数, 用来定量分析网络分形特性。例如通过 Rényi 熵^[16]、结构熵^[17]、最大熵^[18]、相关熵^[19]等从信息论角度研究复杂网络的分形特性; 通过提出节点体积和网络的黎曼 ζ 函数等概念完善体积维数, 从而进行网络分形特性分析^[20-22]。

利用网络的分形特性, 不仅揭示了网络全局的拓扑结构特征, 还反映了节点的局部信息^[23], 一些分形维数也逐渐得到应用。Pu 等^[24]通过局部维数识别网络中具有影响力的节点。同样的信息维也对网络中重要节点识别起到了关键的作用^[25-26]。文献[27]引入非广延参数 q , 利用多重局部维度的方法提供了一个更通用的识别重要节点的方法。文献[28]利用分形维数分析了互联网中演化的拓扑特征。

本文针对网络全局效率计算复杂度高、耗时长的问题, 利用分形网络具有整体与局部的自相似性特征, 深入分析网络关联维数与网络效率的关系, 将网络全局效率的计算由部分局部节点效率的计算来近似估计, 大大降低了网络全局效率计算的复杂度。具体实现过程如下: 1) 判断复杂网络是否具有分形特性; 2) 利用关联维数与尺度距离的关系估计节点关联和; 3) 通过节点关联和与网络效率的关系估算网络全局效率。

2 理论基础

设网络 $G=(N, E)$, 其中, N 表示节点的集合, E 表示边的集合。

2.1 距离

2 个节点间的距离 d 定义为一个节点到另一个

节点的最短路径的边权和, 即

$$d_{ij} = \min \{w_{ik} + w_{kl} + \dots + w_{mj}\} \quad (1)$$

其中, k, l, m 为节点序号; d_{ij} 表示节点 i 与节点 j 的距离; w_{ik} 为节点 i 与节点 k 的边权值; \min 值为从节点 i 到节点 j 所经过的所有可能中最小的边权和, 即最短路径长度, 可以通过 Floyd^[13] 或 Dijkstra 算法^[29] 求得。当所有边的权值为 1 时即为无权网络, 此时距离表现为 2 个节点间的最少跳数。

2.2 关联维数

分形维数是定量分析复杂网络分形特性的主要工具^[14,30]。关联维数是分形维数的一种, 在复杂网络中表示随机节点所占空间维数的度量^[31]。在平面网络关联维数^[32]计算中, 距离小于或等于 r 的节点对数量所占总数的比例 $C(r)$ 称为关联和, 即

$$C(r) = \frac{2 \sum_{i=1}^n \sum_{j=i+1}^n H(r - d_{ij})}{n(n-1)} \quad (2)$$

其中, n 为节点总数, $H(x)$ 为 Heaviside 阶跃函数^[33], 如式(3)所示。

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad (3)$$

r 与 $C(r)$ 呈幂律关系, 随着半径 r 的扩大, $C(r)$ 也相应增加, 并且 $C(r)$ 随 r 以幂律的形式增长, 即 $C(r) \sim r^\beta$, 则认为该网络存在分形特性, 幂指数 β 为关联维数。通常 $C(r) \sim r^\beta$ 并不是在所有尺度上都存在幂律增长的形式。在该幂律关系区间内, 对 r 与 $C(r)$ 取对数, 拟合直线 $\log(r)$ 与 $\log(C(r))$, 则该拟合直线斜率即为网络的关联维数值。

2.3 网络全局效率

网络的全局效率简称为网络效率, 表示整个网络的平均效率, 效率的度量允许对信息流进行精确的定量分析。数值上等于所有节点对距离倒数之和的平均值, 定义为^[34]

$$E = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \quad (4)$$

最初提出网络效率的概念是用于定量分析复杂网络中小世界的拓扑结构特征^[2], 效率越高代表节点间的交互将消耗更少能量^[35]。网络效率的计算有助于识别网络中有影响力的个体和改进网络结果等功能^[11]。

2.4 节点效率

节点效率 I_k 是指节点 k 到达其他所有节点的难易程度, 数值上等于该节点与其他所有节点距离倒数之和的平均值^[6], 如式(5)所示。

$$I_k = \frac{1}{n-1} \sum_{i=1, i \neq k}^n \frac{1}{d_{ik}} \quad (5)$$

网络效率刻画的是整个网络的平均接近程度, 而节点效率则表示一个节点与其他所有节点信息交换的效率。节点效率越高, 则该节点与其他节点的信息交换越容易, 消耗的资源越少, 同时也包含了网络局部拓扑结构特征。从节点效率的定义中可看出, 所有节点效率的平均值等于网络效率 E , 因此网络效率 E 也可通过式(6)求解^[6]。

$$E = \frac{1}{n} \sum_{k=1}^n I_k \quad (6)$$

3 基于关联维数的网络效率估计方法

3.1 关联维数与网络效率分析

从式(6)可知, 网络效率可以转换为求所有节点效率的均值。在计算所有节点效率时, 会发现一些节点的效率实际上相差不大, 如邻居节点之间仅差一跳, 那么它们具有近似的效率值, 在规模较大的网络中, 邻居节点之间效率值差距将更小。在关联维数的计算中, 需要计算一定距离下的节点对数量, 并且关联和 $C(r)$ 也可以通过每个节点距离小于或等于 r 的数量再取平均得到。从式(2)可得, 节点对距离小于或等于 r 的数量为

$$A(r) = \frac{1}{2} C(r) n(n-1) \quad (7)$$

则距离等于 r 的节点对数量为

$$S(r) = A(r) - A(r-1) \quad (8)$$

$A(r)$ 和 $S(r)$ 是反映总体数量的。考虑单个节点的情况, 以图 1 的网络为例, 以节点 i 为中心、 $r=1$ 为半径的范围包含 8 个邻居节点, 2 跳距离以内的节点数为 21 个, 与节点 i 在 r 跳以内的节点数记为 $A_i(r)$ 。随着 r 增大, $A_i(r)$ 也不断增大, 直到 r 达到网络的半径时, $A_i(r)$ 将等于除节点 i 外的总节点数。对所有节点求和, 就可以得到 $A(r) = \sum_{k=1}^n A_k(r)$ 。以节点 i 为中心、 r 为半径的节点的分布情况如表 1 所示。

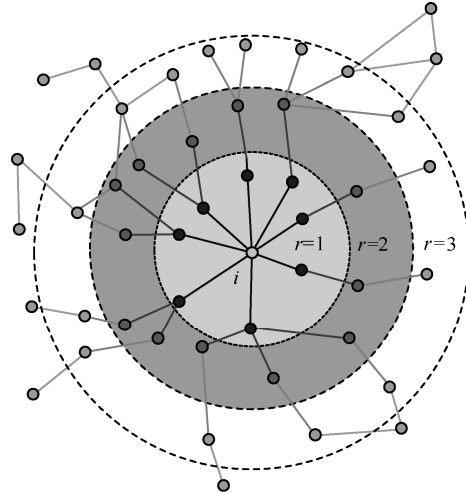


图 1 以节点 i 为中心、 r 为半径的节点分布

表 1 以节点 i 为中心、 r 为半径的节点的分布情况

统计量	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$
$A_i(r)$	8	21	36	44	46
$S_i(r)$	8	13	15	8	2

表 1 中, $S_i(r)$ 代表节点 i 的 r 阶邻居数, $S_i(1)=8$ 表示直接(一阶)与节点 i 相连的节点数量, 当 $r=2$ 时, 2 阶邻居即邻居的邻居, 与 i 的距离为 2, 因此 $S_i(2)=13=A_i(2)-A_i(1)$ 。同样, 总体的 r 阶邻居数量 $S(r) = \sum_{k=1}^n S_k(r)$ 。此时通过式(5)计算 i 的节点效率是节点 i 到其他节点的距离倒数的平均值, 共有 $S_i(1)$ 个一阶邻居, 其距离倒数为 1; 共有 $S_i(2)$ 个二阶邻居, 其距离倒数为 $\frac{1}{2}$; 依次类推, 可得

$$\sum_{k=1, k \neq i}^n \frac{1}{d_{ik}} = \sum_{r=1}^L S_i(r) \frac{1}{r} \quad (9)$$

其中, L 表示网络的直径, d_{ik} 表示节点 i 与节点 k 的距离。式(9)表达的是以节点 i 为中心, i 与其他节点距离倒数之和可等价于求节点 i 的 r 阶邻居数。所以网络总的距离倒数之和为

$$\sum_{j \neq k \in G} \frac{1}{d_{jk}} = \sum_{r=1}^L S(r) \frac{1}{r} \quad (10)$$

将式(10)代入式(4), 可得网络效率的另一种计算式, 即

$$E = \frac{1}{n(n-1)} \sum_{r=1}^L \frac{S(r)}{r} \quad (11)$$

考虑一个网络具有强分形特性, 如构造一个环

形的最近邻耦合网络，即为 Newman-watts 和 Watts-Strogatz 小世界网络模型的初始规则网络。在双对数坐标轴上，画出 $C(r)$ 与 r 的关系如图 2 所示， $\log(r)$ 与 $\log(C(r))$ 能完美地拟合成一条直线。

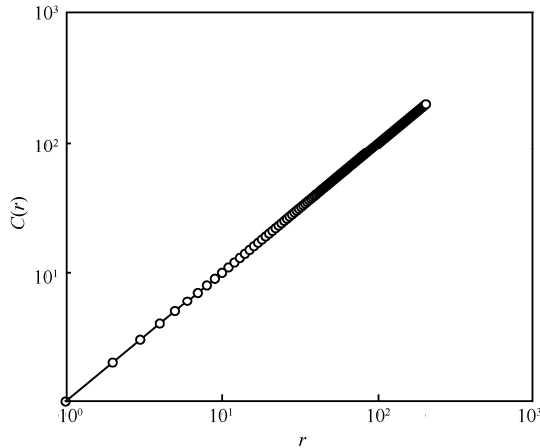


图 2 环形最近邻耦合网络中 $C(r)$ 与 r 的关系

r 与 $C(r)$ 的关系为

$$C(r) = r^\beta \tag{12}$$

根据式(7)与式(8)可得

$$S(r) = \frac{n(n-1)}{2} [C(r) - C(r-1)] \tag{13}$$

将式(12)代入式(13)可得

$$S(r) = \frac{n(n-1)}{2} [r^\beta - (r-1)^\beta] \tag{14}$$

将式(14)代入式(11)可得

$$E = \frac{1}{2} \sum_{r=1}^L \frac{r^\beta - (r-1)^\beta}{r} \tag{15}$$

此时，计算网络效率仅需要网络的直径 L 和关联维数 β 这 2 个参数，便可以求出准确答案。对于关联维数，可通过 $\beta = \frac{\log(C(r))}{\log(r)}$ 得出。

3.2 确定无标度区间

虽然环形的最近邻网络中， $C(r)$ 能在所有尺度 r 中都存在幂律关系，但实际上大部分网络中 $C(r)$ 与 r 只能在部分尺度范围存在幂律关系。如图 3 所示，整个线性区间 (r_1, r_m) 称为无标度区间。式(14)只能在区间 (r_1, r_m) 成立，将其代入式(10)后变为

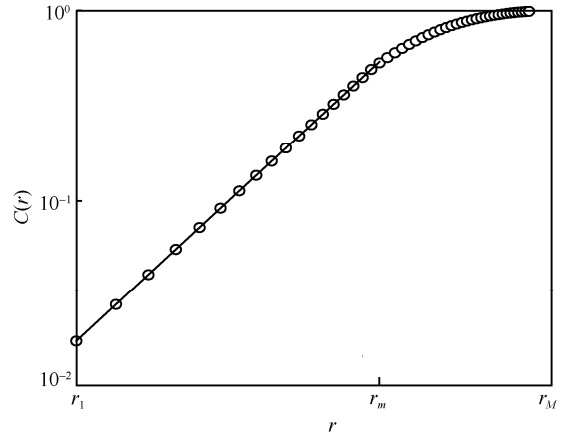


图 3 分形网络 $C(r)$ 与 r 的关系

$$E = \frac{1}{2} \sum_{r=r_1}^{r_m} \frac{r^\beta - (r-1)^\beta}{r} + \sum_{r=r_m+1}^L \frac{S(r)}{r} \tag{16}$$

当 r 数值较小时， $\frac{1}{r}$ 较大，这会导致节点对数

量 $S(r)$ 的误差对整体造成更大的偏差，因此具有分形特性的网络在区间 (r_1, r_m) 使用关联维数与 r 的幂律关系来代替 $S(r)$ ，可以降低在无标度区间网络效率估计的敏感性，使误差值减小。

由上述分析可知，网络效率并不能单纯地通过网络直径和关联维数求得。需要先确定网络是否存在分形特性和无标度区间范围。因此本文提出一种节点关联和的方法来判断复杂网络是否存在分形特性。定义单个节点的关联和为

$$V_k(r) = \frac{2}{n-1} \sum_{i=1, i \neq k}^n H(r - d_{ik}) \tag{17}$$

则 $C(r)$ 与 $V_k(r)$ 的关系为

$$C(r) = \frac{1}{n} \sum_{i=1}^n V_k(r) \tag{18}$$

总体关联和等于各节点关联和的平均值。对于 $C(r)$ 的求解，并不需要计算所有的节点平均，只需要在每求得一个节点关联和时计算累积当前关联和平均值，如果出现了无标度区间，则可认为不必继续计算更多的节点关联和。

3.3 确定决定系数阈值

无标度区间经常由人工判断，后来 Clauset 等^[36] 提出幂律分布的参数估计方法，同时也给出了可以用于估计无标度区间的方法。其中区间确认使用的是调整权重 KS (Kolmogorov Smirnov) 统计量的方法，通过 2 个经验分布之间的最大差值达到最小来确定区间范围。当区间 (r_1, r_m) 的 r_m 值越接

近 r_1 时, KS 统计量的差值通常就会越小, 导致无标度区间的估计偏小。为了避免无标度区间过小, 本文使用区间缩减的决定系数来确定无标度区间范围。在区间缩减过程中, 当决定系数值达到阈值上限, 则表明有较好的拟合优度, 当前的区间即为无标度区间。设共有 w 个观测值 x_1, \dots, x_w , 当区间缩减至 $(1, q)$ 时, 通过数据 (x_1, \dots, x_w) 中的前 q 个数据来估计数值 $(\hat{x}_1, \dots, \hat{x}_q)$, 则前 q 个数的决定系数值为

$$R^2(q) = \frac{\sum_{i=1}^q (x_i - \hat{x}_i)^2}{\sum_{i=1}^q (x_i - \bar{x})^2} \quad (19)$$

q 值从 w 递减到 $q = r_m$, 当 $R^2(r_m)$ 首次大于阈值 R_y^2 且 r_m 足够大时, $(1, r_m)$ 为无标度区间, 该网络具有分形特性。由式(16)可知, 网络效率由关联维数估计部分与区间外通过式(19)的 r 阶邻居 2 个部分组成。关联维数估计的部分所占网络效率整体的比重越大, 则该网络的分形特性越强, 估计方法将更准确。无标度区间关联维数估计部分占整体的比重为

$$P_f = \frac{\sum_{r=1}^{r_q} \frac{r^\beta - (r-1)^\beta}{r}}{\sum_{r=1}^{r_q} \frac{r^\beta - (r-1)^\beta}{r} + 2 \sum_{r=r_q+1}^L \frac{S(r)}{r}} \quad (20)$$

3.4 网络效率的估计算法

本文提出的基于关联维的网络效率估计 (CDEE, correlation dimension for efficiency estimation) 算法主要包括 3 个步骤: 选取部分节点通过广度优先遍历的方法求取关联和; 确定无标度区间范围; 当无标度区间的权重达到阈值时, 进行网络效率的估计, 否则重复上述步骤。算法的具体流程如算法 1 所示。

算法 1 CDEE

输入 网络图 $G = (N, E)$ 的邻接矩阵 $A \in \mathbb{R}^{n \times n}$, 计算关联维所需的最小节点数 t_{\min} , 决定系数上限值 R_y^2 , 无标度区间权重阈值 P_y

输出 网络效率估计值 E

- 1) function cal_scale_region(N_i, List_V)
- 2) for t in N_i do
- 3) 遍历集合 N_i 的所有节点;
- 4) 初始半径 $r=1$

- 5) while $V_t(r+1) - V_t(r)$ do
- 6) 节点 t 以 r 为半径统计遍历 r 范围内的节点总数 $V_t(r)$ 和 $V_t(r+1)$
- 7) 半径 r 自增 1
- 8) end while
- 9) $V_t(r)$ 加入列表 List_V 中
- 10) end for
- 11) 循环停止后记录下 r 的最大值记为 r_{\max} ;
- 12) 计算 List_V 内所有节点关联和平均值 $V_{\text{avg}}(r)$;
- 13) 区间缩减变量 $r_v = r_{\max}$;
- 14) while $r_v > 3$ & $R^2(r_v) < R_y^2$
- 15) 计算区间 $(1, r_v)$ 的缩减决定系数值 $R^2(r_v)$, 直到达到预设上限值 R_y^2 ;
- 16) $r_v = r_v - 1$;
- 17) end while
- 18) 达到决定系数上限值时, 区间 $(1, r_v)$ 即为无标度区间;
- 19) if $r_v < 3$ then
- 20) return None
- 21) else
- 22) return r_v
- 23) end if
- 24) end function
- 25) 将节点集合 N 分成 $\frac{|N|}{t_{\min}}$ 份;
- 26) for $i = 1$ to $\frac{|N|}{t_{\min}}$ do
- 27) 计算每份集合 N_i 的关联和列表 List_V , 并获取无标度区间;
- 28) $r_v = \text{cal_scale_region}(N_i, \text{List}_V)$
- 29) if r_v then
- 30) 通过式(20)计算无标度区间权重 P_f
- 31) if $P_f > P_y$ or $P_f < 0.01$ then
- 32) 当 P_f 大于阈值或者收敛时跳出循环
- 33) break
- 34) end if
- 35) end if
- 36) end for
- 37) 通过式(13)与式(16)求得网络效率的估计 E

38) return E

CDEE 输入 4 个部分：图 $G = (N, E)$ ；计算关联维数所需的最少节点数 t_{\min} ，即每次最少选择 t_{\min} 个节点来求解关联和的平均值；决定系数上限值 R_y^2 ；无标度区间权重阈值 P_f 。

算法 1 的步骤 1)~步骤 24)为求取无标度区间函数，对象为列表 $List_V$ 中所有节点的平均关联和 $C(r)$ 与 r 的关系。在每次调用该函数时，列表 $List_V$ 都会存储 t_{\min} 个节点的关联和。

步骤 25)为代码的第一个步骤，将集合 N 分成 $\frac{|N|}{t_{\min}}$ 份。第一个循环求取第一份的无标度区间，第二个循环求取前两份的节点集合的无标度区间，以此类推。

步骤 26)~步骤 36)开始的循环计算无标度区间权重，循环的终止条件为无标度区间权重值 P_f 超过阈值或者 P_f 收敛 (P_f 的变化小于 0.01)。停止循环后开始进行最终的计算网络效率的估计值。

4 实验结果及分析

实验将从构造生成的网络（分形网络和小世界网络）和真实网络（道路网络）来验证所提方法的有效性和准确性。

4.1 构造网络的效率估计

4.1.1 分形网络效率估计

迭代生成一个分形增长的复杂网络，文献[37]通过逆重整化的方法迭代生成分形网络，生长过程由 n 、 m 、 e 这 3 个参数控制，满足

$$N(t) = nN(t-1) \tag{21}$$

$$k_i(t) = mk_i(t-1) \tag{22}$$

其中， $N(t)$ 和 $k_i(t)$ 分别表示 t 次迭代时的节点数量和

节点 i 的度，节点 i 在逆重整化中将重新加入 $n-1$ 个节点数，其中 $m-1$ 个节点将连接到 i 上，剩余的 $n-m-2$ 个节点随机连接到节点 i 的邻居上。 $e \in [0, 1]$ 表示分形程度， e 越接近 1，则网络的分形特性强度越低。在这里先构造 $n=6$ 、 $m=3$ ，并迭代 4 次得到 1 296 个节点数和 1 295 条边的分形网络。对于不同 e 下生成的网络进行验证，这里取 $R_y^2=0.998$ 。

图 4 为构造的分形网络各参数 e 下 $C(r)$ 与 r 的关系。对于 $e=0$ 时，无标度区间范围最大；当 e 从 0 变化到 0.6 时，可以看出有明显的分形特性。

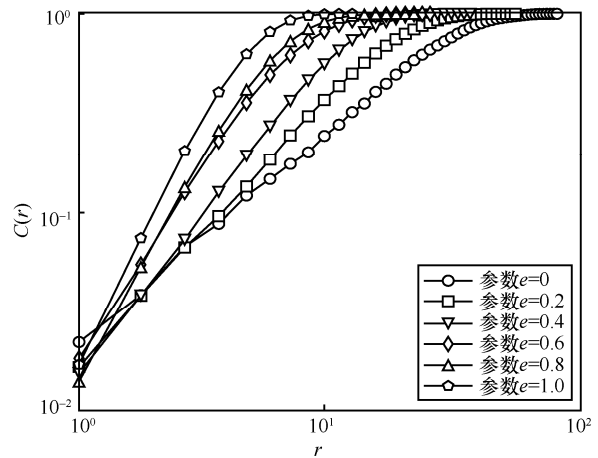


图 4 构造的分形网络各参数 e 下 $C(r)$ 与 r 的关系

不同分形程度的构造网络效率估计如表 2 所示。由表 2 可知，所有效率的估计误差不超过 1.2%，并且用于计算节点关联和的节点数 n_k 最大值为 155，占总节点 12%，其中参数 $e=0.2$ 和 $e=0.4$ 时 CDEE 仅需要计算 8%的节点关联和即可估计出误差小于 1.2%的全局效率，这意味着本文方法比传统方法最高节省了 92%的时间。 e 越小，表示该生成网络的分形性越强，则无标度区间占比

表 2 不同分形程度的构造网络效率估计

分形程度 e	r_k	决定系数 R^2	关联维数	P_f	n_k	T_{save}	估计值 \hat{E}	理论值 E	相对误差
0	24	0.998	1.46	92.20%	142	89%	0.084 9	0.085 6	0.82%
0.2	16	0.998	1.84	87.80%	104	92%	0.105 2	0.106 3	1.04%
0.4	11	0.998 6	2.23	80.50%	104	92%	0.124 3	0.125 8	1.19%
0.6	7	0.998 4	2.66	66.20%	155	88%	0.152 9	0.152 3	0.39%
0.8	5	0.998 1	3.09	57.10%	129	90%	0.169 7	0.168 3	0.83%
1.0	5	0.998 4	3.46	54.30%	142	89%	0.188 4	0.189 9	0.79%

注： r_k 为无标度区间最大值， P_f 为效率估计中无标度区间的比重， n_k 为出现无标度区间时选取的节点数量， T_{save} 为比原方法节省时间的百分比， \hat{E} 为网络效率估计值， E 为网络效率理论值。

P_f 也越大。关联维数随着网络效率的增大而线性增长，这是因为网络效率越高，则节点间信息交换效率越高，反映在局部中为小于 r 的节点对数量是增加的，导致 $C(r)$ 随着 r 更快增长，因此关联维数也增大。在构造分形网络中， β 与 E 的关系如图 5 所示，它们有着明显的线性关系。在这类分形网络中，关联维数与网络效率的关系为

$$E \approx 0.05\beta + 0.01 \quad (23)$$

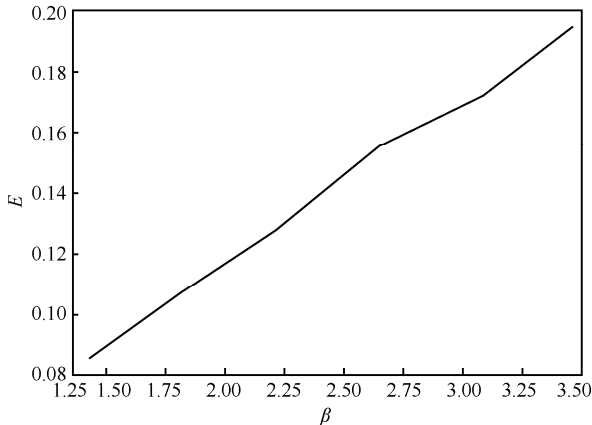


图 5 β 与 E 的关系

4.1.2 小世界网络效率估计

生成构造小世界网络模型为 Newman-Watts 小世界模型，简称为 NW 网络。网络模型由 n 、 K 和 p 控制，分别表示网络节点数、初始节点的邻居数量和节点间加边的概率。同样规模的网络， K 和 p 越大，理论上网络效率越高。生成的 NW 小世界网络节点数为 5 000， $K=8$ 即初始网络节点的邻居数为 8， $R_y^2=0.995$ 。小世界网络模型各参数 p 下 $C(r)$ 与 r 的关系如图 6 所示。从图 6 可以看到，当 $p=0$ 时，分形特性最明显。随着 p 增加，关联维数也增大，因此关联维数在定量分析小世界网络的性质上与网络效率同样有效。表 3 也表明了关联维数与网络效率呈现出 $p \propto \beta$ 的正相关。

由表 3 可知，所有效率估计误差不超过 0.3%，仅需不到 10% 的节点关联和来进行效率的估计，比传统方法至少节省 90% 的时间。

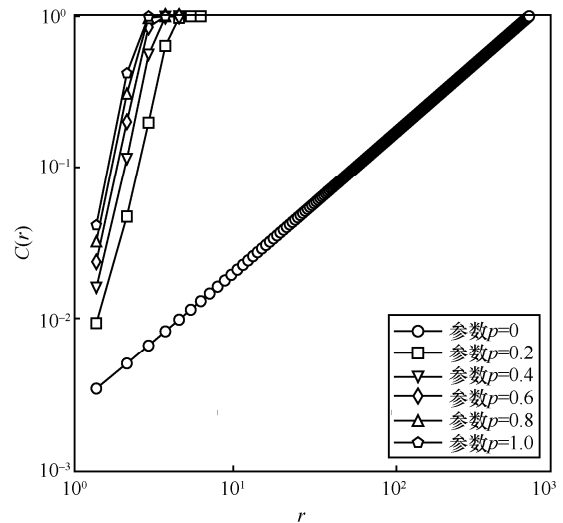


图 6 小世界网络模型各参数 p 下 $C(r)$ 与 r 的关系

4.2 真实世界网络效率估计

上述通过 2 种构造网络验证了所提方法的有效性，在真实世界的网络中，本文选取在 ICON 中公开的美国各州道路网络数据集。随机选取了 3 个公路网络，分别为犹他州、路易斯安那州和威斯康星州。数据中的每个道路网络都存在多个连通子网，本文分析时取网络中最大的连通子网（节点数最多）代表该州道路网络进行效率估计，3 个最大连通子网的平均节点数为 390 231 个，边为 469 217 条。图 7 为真实道路网络中 $C(r)$ 与 r 的关系。真实道路网络效率估计如表 4 所示。效率的估计只通过了不到 1% 的节点关联和，根据 T_{save} 显示，最高能节约 99.6% 的时间，同时估计出的绝对误差值低于 0.65% 的网络效率。估计的网络效率显示，犹他州为 0.004 027，高于威斯康星州的 0.003 625 和路易斯安那州的

表 3 不同连边概率 p 的小世界网络效率估计

连边概率 p	r_k	决定系数 R^2	关联维数	P_f	n_k	T_{save}	估计值 \hat{E}	理论值 E	相对误差
0	626	1	0.998	100%	100	98%	0.011 2	0.011 2	0%
0.2	5	0.995 9	4.6	70.70%	450	91%	0.203 9	0.203 3	0.30%
0.4	4	0.998 8	5.09	64.30%	300	94%	0.241 7	0.241 2	0.21%
0.6	4	0.999 8	5.18	88.30%	100	98%	0.264	0.264 2	0.08%
0.8	3	0.997 3	5.58	38.60%	400	92%	0.280 7	0.281 2	0.18%
1.0	3	0.995 6	5.7	51.30%	150	97%	0.294 1	0.293 8	0.10%

表 4 真实道路网络效率估计

道路网络	节点数	边数	r_k	决定系数 R^2	关联维数	P_f	n_k	T_{save}	估计值 \hat{E}	理论值 E	相对误差
犹他州	242 432	288 982	578	0.995	2.02	0.954	1 815	99.30%	0.004 027	0.004 026	0.03%
威斯康星州	514 687	623 277	503	0.995	1.91	0.868	3 088	99.40%	0.003 625	0.003 602	0.64%
路易斯安那州	413 574	495 392	660	0.995	1.85	0.916	1 836	99.60%	0.003 171	0.003 173	0.06%

0.003 171；同样地，犹他州的关联维数 2.02 也大于威斯康星州和路易斯安那州。结果基本符合 U.S. News&World Report 公布的美国 50 州交通排行，犹他州排第 2 名，威斯康星州排第 32 名，路易斯安那州排第 45 名。本文所提方法允许并行计算节点的关联和，能加速网络效率的估计，并进一步缩减计算时间。

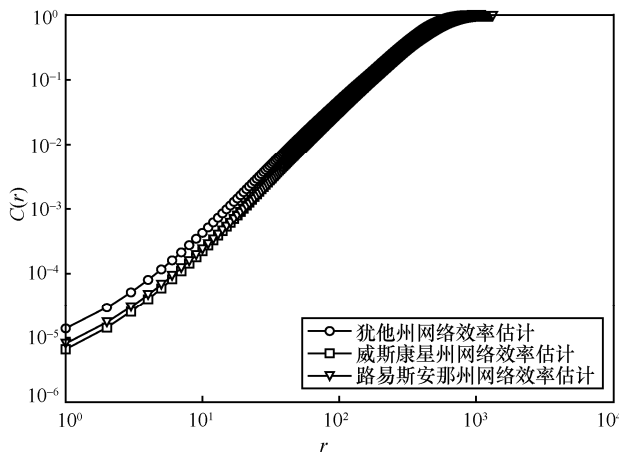


图 7 真实道路网络中 $C(r)$ 与 r 的关系

5 结束语

本文针对复杂网络中网络效率计算复杂度高、时间长的问题提出一种效率的估计方法。该方法考虑复杂网络中存在着分形的拓扑结构特性，从分形维的角度推导与分析了其中的关联维数与网络效率的关系，并给出了一种检测网络分形特性的新方法。实验分析发现，在 2 种构造生成的网络中，关联维数随网络效率线性增长，这说明关联维数也能在一定程度上衡量网络中节点信息交换的效率和复杂网络的小世界性。本文方法不仅对强分形网络有效，在弱分形网络中也同样有效，只需要知道无标度区间占比的变化情况也能得到较准确的效率估计。在节点数小于 5 000 的复杂网络中，需要约 10% 的节点关联和，与传统计算方法相比，可节省 90% 的时间，并得到了误差值小于

1.2% 的网络效率估计值。在规模更大的网络（节点数大于 200 000）中，3 个道路网络效率的估计最多用到了 0.75% 的节点关联和就可以获得误差不超过 0.65% 的网络效率估计值。本文方法更适用于大规模网络，大大降低了计算网络效率所需的时间，平均时间复杂度将缩减至 $O(N)$ 。由于节点关联和计算是独立的，因此本文方法将允许并行化操作，能进一步缩短效率估计所需的时间。

参考文献：

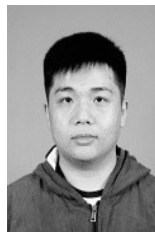
- [1] BARABÁSI A L, PÓSFAL M. Network science[M]. Cambridge: Cambridge University Press, 2016.
- [2] LATORA V, MARCHIORI M. Efficient behavior of small-world networks[J]. Physical Review Letters, 2001, 87(19): 3-6.
- [3] HÉBERT-DUFRESNE L, ALLARD A, YOUNG J, et al. Global efficiency of local immunization on complex networks[J]. Scientific Reports, 2013(3): 2171.
- [4] AVENA-KOENIGSBERGER A, MISIC B, SPORNS O. Communication dynamics in complex brain networks[J]. Nature Reviews Neuroscience, 2018, 19(1): 17-33.
- [5] BEATY R E, KAUFMAN S B, BENEDEK M, et al. Personality and complex brain networks: the role of openness to experience in default network efficiency[J]. Human brain mapping, 2016, 37(2): 773-779.
- [6] 周漩, 张凤鸣, 周卫平, 等. 利用节点效率评估复杂网络功能鲁棒性[J]. 物理学报, 2012, 61(19): 1-7.
- [7] ZHOU X, ZHANG F P, ZHOU W P, et al. Evaluating complex network functional robustness by node efficiency[J]. Acta Physica Sinica, 2012, 61(19): 1-7.
- [8] EK B, VERSCHNEIDER C, NARAYAN D A. Efficiency of star-like graphs and the Atlanta subway network[J]. Physica A: Statistical Mechanics and Its Applications, 2013, 392(21): 5481-5489.
- [9] ZHANG L, QIN J, HE Y, et al. Network-level optimization method for road network maintenance programming based on network efficiency[J]. Journal of Central South University, 2015, 22(12): 4882-4889.
- [10] LATORA V, MARCHIORI M. Economic small-world behavior in weighted networks[J]. The European Physical Journal B-Condensed Matter and Complex Systems, 2003, 32(2): 249-263.
- [11] BOMPARD E, WU D, XUE F. Structural vulnerability of power systems: a topological approach[J]. Electric Power Systems Research, 2011, 81(7): 1334-1340.
- [12] WANG S, DU Y, DENG Y. A new measure of identifying influential nodes: efficiency centrality[J]. Communications in Nonlinear Science and Numerical Simulation, 2017(47): 151-163.

- [12] WANG Y, WANG S, DENG Y. A modified efficiency centrality to identify influential nodes in weighted networks[J]. *Pramana*, 2019, 92(4): 68.
- [13] HOUGARDY S. The Floyd-Warshall algorithm on graphs with negative cycles[J]. *Information Processing Letters*, 2010, 110(8-9): 279-281.
- [14] SONG C, HAVLIN S, MAKSE H A. Self-similarity of complex networks[J]. *Nature*, 2005, 433(7024): 392-395.
- [15] SONG C, HAVLIN S, MAKSE H A. Origins of fractality in the growth of complex networks[J]. *Nature Physics*, 2006, 2(4): 275-281.
- [16] DUAN S, WEN T, JIANG W. A new information dimension of complex network based on Rényi entropy[J]. *Physica A: Statistical Mechanics and Its Applications*, 2019(516): 529-542.
- [17] LEI M, LIU L, WEI D. An improved method for measuring the complexity in complex networks based on structure entropy[J]. *IEEE Access*, 2019(7): 159190-159198.
- [18] ROSENBERG E. Maximal entropy coverings and the information dimension of a complex network[J]. *Physics Letters A*, 2017, 381(6): 574-580.
- [19] ZHANG Q, LI M, DENG Y. Measure the structure similarity of nodes in complex networks based on relative entropy[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018(491): 749-763.
- [20] SHANKER O. Defining dimension of a complex network[J]. *Modern Physics Letters B*, 2007, 21(6): 321-326.
- [21] SHANKER O. Graph zeta function and dimension of complex network[J]. *Modern Physics Letters B*, 2007, 21(11): 639-644.
- [22] WEI D, WEI B, ZHANG H, et al. A generalized volume dimension of complex networks[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2014, 2014(10): P10039.
- [23] WEN T, JIANG W. An information dimension of weighted complex networks[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018(501): 388-399.
- [24] PU J, CHEN X, WEI D, et al. Identifying influential nodes based on local dimension[J]. *Europhysics Letters*, 2014, 107(1): 10010.
- [25] BIAN T, DENG Y. Identifying influential nodes in complex networks: A node information dimension approach[J]. *Chaos: an Interdisciplinary Journal of Nonlinear Science*, 2018, 28(4): 43109.
- [26] WEN T, DENG Y. Identification of influencers in complex networks by local information dimensionality[J]. *Information Sciences*, 2020(512): 549-562.
- [27] WEN T, PELUSI D, DENG Y. Vital spreaders identification in complex networks with multi-local dimension[J]. *Knowledge-Based Systems*, 2020: 105717.
- [28] 关世杰, 赵海. 互联网中路由级和IP级拓扑分形特征分析[J]. *通信学报*, 2013, 34(11): 162-170.
GUAN S J, ZHAO H. Analysis of fractal characteristic of internet router-level and IP-level topology[J]. *Journal on Communications*, 2013, 34(11): 162-170.
- [29] DIJKSTRA E W. A note on two problems in connexion with graphs[J]. *Numerische Mathematik*, 1959, 1(1): 269-271.
- [30] SONG C, GALLOS L K, HAVLIN S, et al. How to calculate the fractal dimension of a complex network: the box covering algorithm[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2007, 2007(3): P03006.
- [31] LACASA L, GÓMEZ-GARDEÑES J. Correlation dimension of complex networks[J]. *Physical Review Letters*, 2013, 110(16): 1-5.
- [32] WANG X, LIU Z, WANG M. The correlation fractal dimension of complex networks[J]. *International Journal of Modern Physics C*, 2013, 24(5): 1350033.
- [33] 姜建国, 曹建中, 高玉明. 信号与系统分析基础[M]. 北京: 清华大学出版社, 2006.
JIANG J G, CAO J Z, GAO Y M. *Fundamentals of signal and system analysis*[M]. Beijing: Tsinghua University Press, 2006.
- [34] EK B, VERSCHNEIDER C, NARAYAN D A. Global efficiency of graphs[J]. *AKCE International Journal of Graphs and Combinatorics*, 2015, 12(1): 1-13.
- [35] ACHARD S, BULLMORE E. Efficiency and cost of economical brain functional networks[J]. *PLOS Computational Biology*, 2007, 3(2): e17.
- [36] CLAUSET A, SHALIZI C R, NEWMAN M E J. Power-law distributions in empirical data[J]. *SIAM Review*, 2009, 51(4): 661-703.
- [37] KITSACK M, HAVLIN S, PAUL G, et al. Betweenness centrality of fractal and nonfractal scale-free model networks and tests on real networks Maksim[J]. *Physical Review E*, 2007, 75(5): 056115.

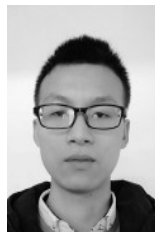
[作者简介]



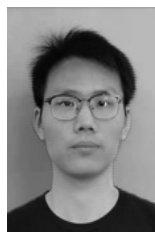
张胜 (1968-), 男, 湖北黄冈人, 博士, 南昌航空大学副教授, 主要研究方向为复杂系统、复杂网络、无线传感器网络等。



戴维凯 (1994-), 男, 广东惠州人, 南昌航空大学硕士生, 主要研究方向为复杂网络、分形理论等。



吴锋 (1996-), 男, 湖北黄冈人, 南昌航空大学硕士生, 主要研究方向为复杂网络、无线传感器网络、分形理论等。



蓝文祥 (1996-), 男, 江西高安人, 南昌航空大学硕士生, 主要研究方向为复杂网络、分形理论等。