

基于  $K$ -阶结构熵的网络异构性研究

黄丽亚 霍宥良 王青 成谢锋

Network heterogeneity based on  $K$ -order structure entropy

Huang Li-Ya Huo You-Liang Wang Qing Cheng Xie-Feng

引用信息 Citation: *Acta Physica Sinica*, 68, 018901 (2019) DOI: 10.7498/aps.68.20181388

在线阅读 View online: <http://dx.doi.org/10.7498/aps.68.20181388>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2019/V68/I1>

---

您可能感兴趣的其他文章

Articles you may be interested in

[相依网络上基于相连边的择优恢复算法](#)

Enhancing resilience of interdependent networks against cascading failures under preferential recovery strategies

物理学报.2018, 67(8): 088901 <http://dx.doi.org/10.7498/aps.67.20172526>

[加权网络中基于冗余边过滤的  \$k\$ -核分解排序算法](#)

A ranking approach based on  $k$ -shell decomposition method by filtering out redundant link in weighted networks

物理学报.2017, 66(18): 188902 <http://dx.doi.org/10.7498/aps.66.188902>

[面向成本-收益好的无标度耦合网络构建方法](#)

A toward cost-effective scale-free coupling network construction method

物理学报.2016, 65(9): 098901 <http://dx.doi.org/10.7498/aps.65.098901>

[基于度的正/负相关相依网络模型及其鲁棒性研究](#)

The model of interdependent network based on positive/negative correlation of the degree and its robustness study

物理学报.2015, 64(4): 048902 <http://dx.doi.org/10.7498/aps.64.048902>

[一种信息传播促进网络增长的网络演化模型](#)

A complex network evolution model for network growth promoted by information transmission

物理学报.2014, 63(15): 158902 <http://dx.doi.org/10.7498/aps.63.158902>

# 基于 $K$ -阶结构熵的网络异构性研究\*

黄丽亚<sup>†</sup> 霍宥良 王青 成谢锋

(南京邮电大学, 电子与光学工程学院, 微电子学院, 南京 210023)

(2018年7月19日收到; 2018年11月8日收到修改稿)

结构熵可以考察复杂网络的异构性. 为了弥补传统结构熵在综合刻画网络全局以及局部特性能力上的不足, 本文依据网络节点在  $K$  步内可达的节点总数定义了  $K$ -阶结构熵, 可从结构熵随  $K$  值的变化规律、最大  $K$  值下的结构熵以及网络能够达到的最小结构熵三个方面来评价网络的异构性. 利用  $K$ -阶结构熵对规则网络、随机网络、Watts-Strogatz 小世界网络、Barabási-Albert 无标度网络以及星型网络进行了理论与仿真实验, 结果表明上述网络的异构性依次增强. 其中  $K$ -阶结构熵能够较好地依据小世界属性来刻画小世界网络的异构性, 且对星型网络异构性随其规模演化规律的解释也更为合理. 此外,  $K$ -阶结构熵认为在规则结构外新增孤立节点的网络的异构性弱于未添加孤立节点的规则结构, 但强于同节点数的规则网络. 本文利用美国西部电网进一步论证了  $K$ -阶结构熵的有效性.

**关键词:** 复杂网络, 异构性, 熵

**PACS:** 89.75.-k, 89.75.Fb

**DOI:** 10.7498/aps.68.20181388

## 1 引言

大量系统都可以使用复杂网络进行抽象的表达<sup>[1]</sup>. 复杂网络各部分之间虽相互联系, 但通常存在着功能结构上的差异<sup>[2,3]</sup>. 因此, 对网络异构性的研究是分析节点重要性排序<sup>[4,5]</sup>、社区结构划分<sup>[6,7]</sup>以及网络传播效率、同步能力等<sup>[8,9]</sup>内容的重要基础. 通常, 网络异构性是从拓扑结构的角度出发, 通过定义结构熵予以定量评价. 结构熵越小, 意味着网络各部分间的差异越大, 异构性越强; 反之意味着网络结构越趋于均衡, 异构性越弱. 目前, 已有诸多研究<sup>[10-15]</sup>各自从不同的角度定义了网络结构熵.

Wu 熵<sup>[10]</sup>以节点为主体, 试图通过分析网络节点所拥有的边的条数, 即各节点度值之间的差异来反映网络的异构性. 而度分布结构熵<sup>[11]</sup> (下文简称 DD 熵) 则以度值为主体, 依据拥有不同度值的节点数量之间的差异, 对网络的异构性进行了刻画. 可见, 以上两种方法仅将邻居节点纳入评价体系, 未考虑非邻居节点的影响, 因而刻画网络全局

特性的能力较弱, 且对“桥”连接节点的重视程度不足. 此外, Wu 熵难以充分地解释网络中出现孤立节点的情况. 蔡萌等<sup>[12]</sup>提出了一种基于点和边差异性的网络结构熵 (下文简称 SD 熵), 虽然将“点差异性”与“边差异性”进行加权求和, 但其本质仍是从网络局部特性的角度出发, 且加权系数是人为设置的, 结果较为主观; 此外, SD 熵在计算节点相对重要性时, 人为地引入了  $\Delta$  项, 缺少理论解释. 为了充分利用网络的全局特性, 文献<sup>[13,14]</sup>以节点介数以及边介数为判据, 定量地评价了网络的异构性 (下文简称节点介数熵为 NB 熵; 边介数熵为 EB 熵). 介数虽然突出了桥节点的重要性, 但对网络拓扑中包含有环型或星型结构的部分解释不足. 为了综合评估网络的局部以及全局特性, 蔡萌等<sup>[15]</sup>又提出了一种基于最大流的网络结构熵, 该方法使用网络流量变化和节点度值的加权平均来计算网络的异构性, 但仍存在权重的选择较为主观、对  $\Delta$  项的解释不足等问题; 且当网络规模较大时, 最大流矩阵的计算也十分复杂. 此外, 若网络流量变化与节点度值量级不同, 最终结果易受量级较大部分的

\* 国家自然科学基金 (批准号: 61373065, 61271334) 资助的课题.

<sup>†</sup> 通信作者. E-mail: [huangly@njupt.edu.cn](mailto:huangly@njupt.edu.cn)

影响. 因此, 寻找一种能够综合考查网络全局以及局部特性的方便有效的异构性评价方法是本研究的目的所在.

文献 [16,17] 将网络视为节点间相互作用的系统, 每个节点均处于其他所有节点产生的势场之中. 该方法假设网络拓扑为短程场, 并基于高斯势函数定义了网络拓扑势函数, 通过比较各个节点所处位置势场值的大小研究了节点重要性、社区划分等网络特性. 但网络势程的长短以及势函数的形式应依网络所代表的实际意义而各有不同, 短程场以及高斯势函数的假设均具有一定的局限性.

本研究试图摒弃势程长短、势函数等额外假设, 仅利用网络的结构特性提出了一种  $K$ -阶结构熵模型, 并对规则网络、随机网络、WS (Watts-Strogatz) 小世界网络、BA (Barabási-Albert) 无标度网络、星型网络、在规则结构外新增孤立节点的网络以及美国西部电网进行了理论研究与仿真分析, 以证明该结构熵可以较好地反映不同网络的异构性.

## 2 $K$ -阶结构熵模型

首先给定无权无向网络图  $G(V, E)$ , 其中  $V = \{v_1, v_2, \dots, v_n\}$  为节点集, 共  $n$  个节点,  $E$  为边集, 并假设  $l_{ij}$  为节点  $v_i$  到  $v_j$  的最短路径长度. 为了避免网络有无自环在分析上造成的差异, 本文假设节点  $0$  步即可到达该节点本身, 定义节点的  $K$ -阶邻居数为

$$N_{v_i}^K = \sum_{j=1, j \neq i}^n I(l_{ij} \leq K) + 1, \quad (1)$$

其中  $I(\cdot)$  为指示函数, 即当节点  $v_i$  与  $v_j$  之间的最短路径长度  $l_{ij} \leq K$  时  $I(\cdot) = 1$ , 否则  $I(\cdot) = 0$ . 由 (1) 式可知, 对于任意节点有  $N_{v_i}^0 = 1$ .

$K$ -阶邻居数衡量了网络节点在  $K$  步内可达的节点总数, 本文将  $K$  值命名为影响力尺度. 由于本文研究的网络大小是有限的, 显然当  $K$  大于网络最大连通部分的直径  $d$  时, 各节点的  $K$ -阶邻居数均不再随  $K$  值的变化而变化. 因此  $K \in \{0, 1, \dots, d\}$ , 并将  $d$  命名为最大影响力尺度.

若某节点的  $K$ -阶邻居数越多, 可认为该节点在尺度  $K$  下所影响的节点越多, 节点则越重要. 为了衡量网络在尺度  $K$  下各节点影响力或重要性的差异, 即网络的异构性, 可将所有节点的  $K$ -阶邻居数加总代入到信息熵公式, 从而得到  $K$ -阶结

构熵为

$$H^K = - \sum_{i=1}^n \frac{N_{v_i}^K}{\sum_{j=1}^n N_{v_j}^K} \log \left( \frac{N_{v_i}^K}{\sum_{j=1}^n N_{v_j}^K} \right), \quad (2)$$

$K \in \{0, 1, \dots, d\}.$

由随机矩阵理论可知, 网络邻接矩阵  $\mathbf{A}$  的  $k$  次幂  $\mathbf{A}^k$  中的第  $i$  行、第  $j$  列元素表示从节点  $v_i$  到  $v_j$  通过  $k$  条边连接的路径数目. 因此, (1) 式可以改写为

$$N_{v_i}^K = \sum_{j=1}^n I \left( \sum_{k=0}^K a_{ij}^{(k)} > 0 \right), \quad (3)$$

其中  $a_{ij}^{(k)}$  为矩阵  $\mathbf{A}^k$  中的第  $i$  行、第  $j$  列元素. 此处  $\mathbf{A}^0 = \mathbf{E}$  为单位阵, 恰对应于前文节点零步可达该节点自身这一假设.

事实上, (3) 式表示的是矩阵多项式  $\mathbf{A}^0 + \mathbf{A}^1 + \dots + \mathbf{A}^K = \sum_{k=0}^K \mathbf{A}^k$  中第  $i$  行或列向量 (无向图的邻接矩阵  $\mathbf{A}$  为对称阵, 则其多项式亦对称) 中非零元素的个数, 即  $L_0$  范数. 于是 (3) 式可以改写为

$$N_{v_i}^K = \left\| \left( \sum_{k=0}^K \mathbf{A}^k \right) \cdot \mathbf{e}_i \right\|_0 = \left\| \sum_{k=0}^K \mathbf{A}^k \cdot \mathbf{e}_i \right\|_0, \quad (4)$$

其中  $\|\cdot\|_0$  表示向量的  $L_0$  范数,  $\mathbf{e}_i$  表示第  $i$  维为 1, 其余维度为 0 的  $n$  维向量. 因此, 第  $i$  个节点的  $K$ -阶邻居数可被认为是  $\mathbf{e}_i$  经矩阵多项式  $\sum_{k=0}^K \mathbf{A}^k$  映射后向量的  $L_0$  范数值, 或映射前后向量  $L_0$  范数的比值 ( $\|\mathbf{e}_i\|_0 = 1$ ). 根据 (4) 式最终可以得到  $K$ -阶结构熵的表达式为

$$H^K = - \sum_{i=1}^n \frac{\left\| \sum_{k=0}^K \mathbf{A}^k \cdot \mathbf{e}_i \right\|_0}{\sum_{j=1}^n \left\| \sum_{k=0}^K \mathbf{A}^k \cdot \mathbf{e}_j \right\|_0} \times \log \left( \frac{\left\| \sum_{k=0}^K \mathbf{A}^k \cdot \mathbf{e}_i \right\|_0}{\sum_{j=1}^n \left\| \sum_{k=0}^K \mathbf{A}^k \cdot \mathbf{e}_j \right\|_0} \right), \quad (5)$$

$K \in \{0, 1, \dots, d\},$

$H^K$  值越小, 证明网络在影响力尺度  $K$  下的异构性越强, 反之越弱.

由前文可知,  $K$  的取值是有限的, 因此在分析网络异构性时仅需研究  $K$ -阶结构熵序列

$\{H^0, H^1, \dots, H^d\}$  即可. 对任意网络而言, 当  $K = 0$  时, 各节点 0 步内只可到达其本身, 因此有  $H^0 = \log(n)$ ; 当  $K = 1$  时, 各节点 1 步内可达该节点本身及其邻居节点, 因此  $H^1$  按各节点“度值加一”代入信息熵公式进行计算, 与 Wu 熵相似却存在差异. 若网络是连通的, 各节点在  $d$  步内均可到达网络任意节点, 因此  $H^d = \log(n)$ ; 若网络是非连通的,  $H^d$  将依网络结构而各不相同. 此外, 由于结构熵序列是可列有限的, 其最小值  $\min H = \min\{H^0, H^1, \dots, H^d\}$  可被认为表征了网络能够达到的最强异构性. 综上所述, 在应用  $K$ -阶结构熵模型研究网络的异构性时, 应综合考虑结构熵序列随  $K$  值的变化规律、 $H^d$  以及  $\min H$ .

当网络为有向网络时, 可将上述  $K$ -阶结构熵演变为  $K$ -阶发送以及接收结构熵两部分进行讨论. 此时, 可以将由某节点出发、 $K$  步内可达的节点总数 (包含其自身) 记为该节点的  $K$ -阶发送节点数, 依此得到  $K$ -阶发送结构熵; 同理可以计算出  $K$ -阶接收结构熵. 显然,  $K$ -阶发送以及接收结构熵分别衡量了网络发、收能力的异构性, 分析时均应予以考虑. 当网络为有权网络时, 若网络边权表示节点间的路径长度, 在计算节点的  $K$ -阶邻居数时, 只需按有权网络路径长度的方法进行计算即可; 若网络权值表示节点间连接的强度时, 则可选定某一阈值或者采用多阈值的方法将网络进行二值化处理, 再利用上述方法进行研究.

### 3 典型网络的结构熵

为了验证  $K$ -阶结构熵评价网络异构性的能力, 本文对规则网络、随机网络、WS 小世界网络、BA 无标度网络、星型网络、在规则结构外新增孤立节点的网络以及美国西部电网进行了仿真分析. 以下所有仿真结果均取 500 次实验的平均值.

#### 3.1 WS 小世界网络

小世界网络是指具有较大平均聚类系数以及较短特征路径长度的一类网络结构. Watts 与 Strogatz<sup>[18]</sup> 最先提出了一种构造小世界网络的方法, 即将最近邻耦合网络每条边的一个端点保持不变, 另外一个端点依概率  $p$  随机改连至其他节点, 通常将生成的网络称为 WS 小世界网络. 当  $p = 0$  时, 网络为规则结构, 其异构性应当最弱; 当  $p = 1$  时, 网络中任意两节点依一定概率随机连接, 其网络异构

性应强于规则网络. 值得注意的是, WS 小世界网络虽为规则与随机网络之间的过渡形式, 其“主体”保持着最近邻耦合结构, 而个别节点却拥有“长程”连接. 但正是由于这些特殊节点的存在, 使得网络的特征路径长度大大减少, 起到了“桥”的作用. 由此可见, WS 小世界网络中个别节点存在着连接上的突出优势, 而随机网络各节点的连接状态是相似的, 因此可认为 WS 小世界网络的异构性强于随机网络.

下面基于  $K$ -阶结构熵模型从结构熵随  $K$  值的变化规律、网络最强异构性  $\min H$  等角度对 WS 小世界网络进行考察. 本文约定网络随机重连后不能有自环、重边且仍为连通图, 易证  $H^d = H^0 = \log(n)$ . 接下来, 图 1 和图 2 分别仿真了 50 和 200 节点最近邻耦合网络在不同重连概率  $p$  下所生成网络的结构熵随  $K$  值的变化规律, 其中该最近邻耦合网络每个节点与其左右各 2 个邻居节点相连.

显然, 最近邻耦合网络的结构是对称的, 节点数为  $n$  的最近邻耦合网络的  $K$ -阶结构熵恒等于  $\log(n)$ , 如图 1 (a) 和图 2 (a) 所示. 由于随机重连打破了网络结构的对称性, 在 50 及 200 节点网络中均可观察到, 在任意重连概率  $p$  下, 网络结构熵随  $K$  值的增加先下降再上升, 且结构熵最小值  $\min H$  所对应的  $K$  值随  $p$  的增大而逐渐减小. 但值得注意的是, 当节点个数为 50 时,  $\min H$  随  $p$  的增加而减小; 但当节点个数为 200 时,  $\min H$  随  $p$  的增加先减小, 如图 2 (a)—(f), 后又略有增大, 如图 2 (f)—(h). 有理由认为,  $\min H$  不仅与网络的拓扑结构有关, 更与网络的规模有着密切的联系. 因此, 图 3 仿真了网络  $\min H$  与小世界指数  $\sigma$  随网络节点数  $n$  以及重连概率  $p$  的变化曲线, 其中对  $\min H$  进行了归一化处理. 参照有关文献<sup>[19,20]</sup>, 本文所使用的小世界指数定义为

$$\sigma = \frac{C/C_n}{L/L_n}, \quad (6)$$

其中,  $C$  和  $L$  分别为网络的平均聚类系数与特征路径长度,  $C_n$  与  $L_n$  分别为同节点数最近邻耦合网络的平均聚类系数与特征路径长度. 显然, 当网络为最近邻耦合网络时,  $\sigma = 1$ ; 由于小世界网络拥有较高的平均聚类系数, 与最近邻耦合网络相近, 即  $C \approx C_n$ ; 但其特征路径长度较短, 远小于最近邻耦合网络, 与同规模随机网络相近, 即  $L \ll L_n$ . 因此可认为  $\sigma$  值越大, 网络的小世界性越显著; 由

图3所示, 从某 $p$ 值开始,  $\sigma$ 随 $p$ 的增加而下降. 这是由于随着重连概率的增加, 网络将同时具有较短的特征路径长度以及较小的平均聚类系数, 逐渐呈现出随机网络的特性.

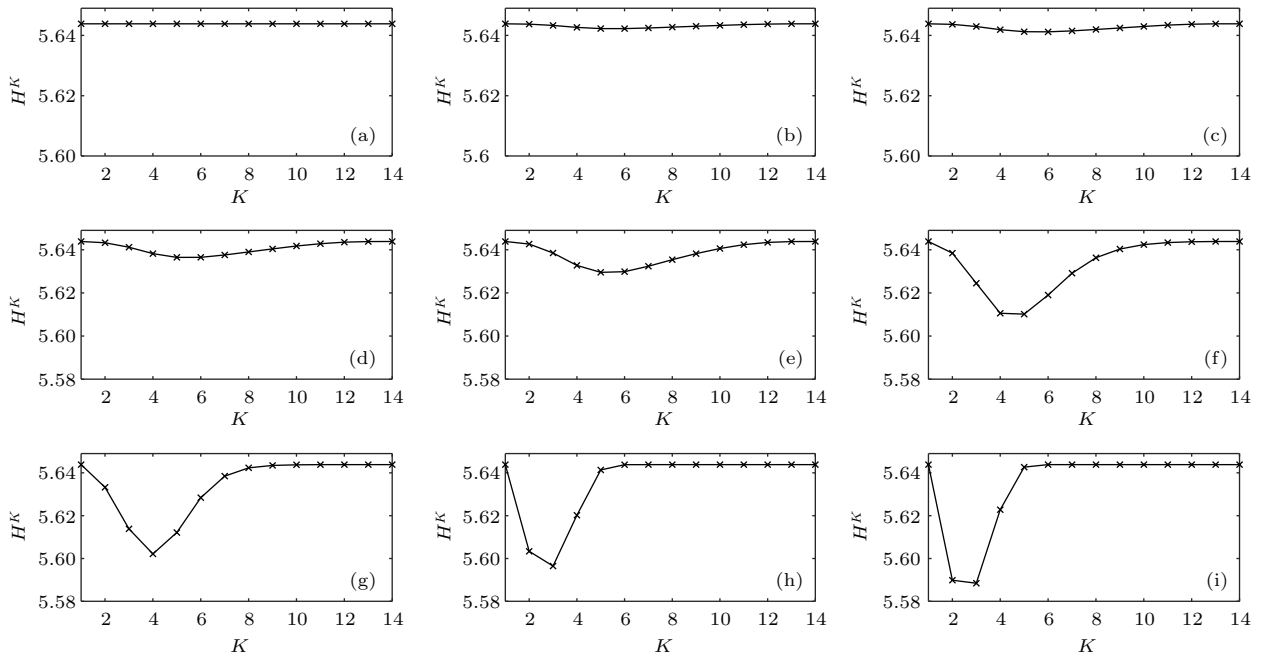


图1 50节点最近邻耦合网络在不同重连概率 $p$ 及影响力尺度 $K$ 下的结构熵 (a)  $p = 0$ ; (b)  $p = 0.001$ ; (c)  $p = 0.002$ ; (d)  $p = 0.005$ ; (e)  $p = 0.01$ ; (f)  $p = 0.05$ ; (g)  $p = 0.1$ ; (h)  $p = 0.5$ ; (i)  $p = 1$

Fig. 1.  $K$ -order structure entropy at influence scale  $K$  in graphs generated by randomly rewiring a 50 nodes nearest-neighbor coupled network based on different probability  $p$ : (a)  $p = 0$ ; (b)  $p = 0.001$ ; (c)  $p = 0.002$ ; (d)  $p = 0.005$ ; (e)  $p = 0.01$ ; (f)  $p = 0.05$ ; (g)  $p = 0.1$ ; (h)  $p = 0.5$ ; (i)  $p = 1$ .

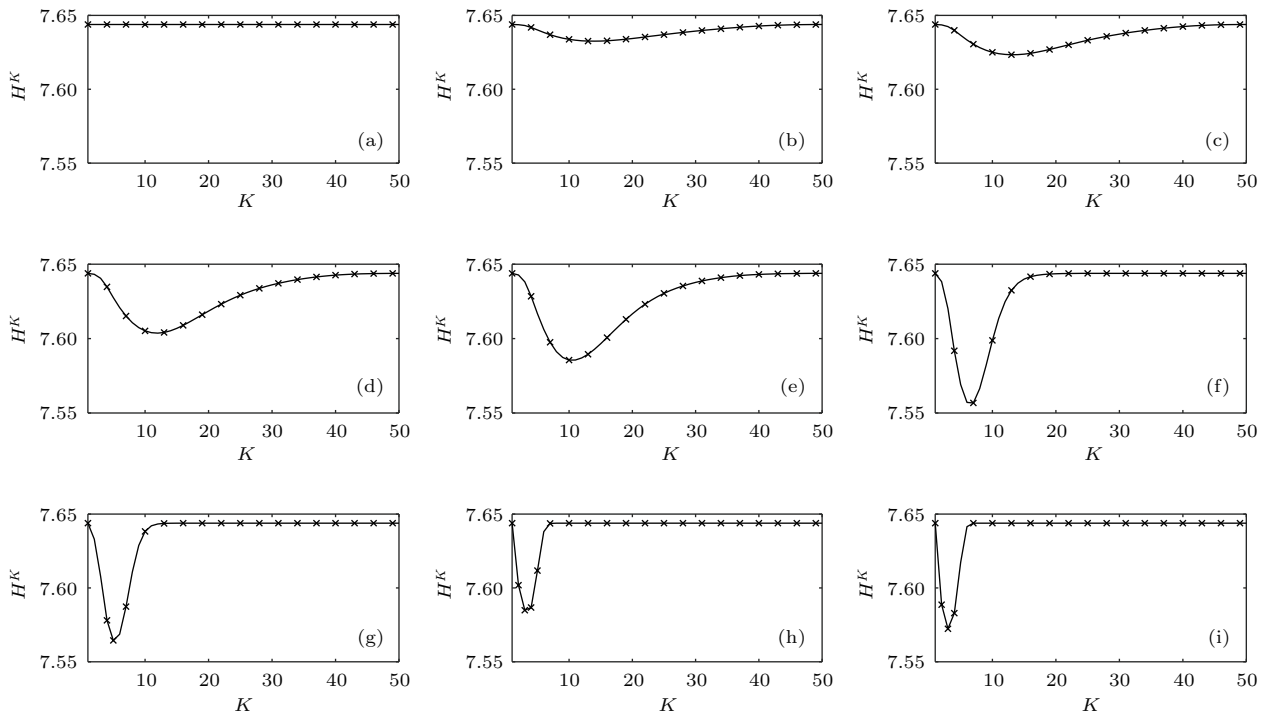


图2 200节点最近邻耦合网络在不同重连概率 $p$ 及影响力尺度 $K$ 下的结构熵 (a)  $p = 0$ ; (b)  $p = 0.001$ ; (c)  $p = 0.002$ ; (d)  $p = 0.005$ ; (e)  $p = 0.01$ ; (f)  $p = 0.05$ ; (g)  $p = 0.1$ ; (h)  $p = 0.5$ ; (i)  $p = 1$

Fig. 2.  $K$ -order structure entropy at influence scale  $K$  in graphs generated by randomly rewiring a 200 nodes nearest-neighbor coupled network based on different probability  $p$ : (a)  $p = 0$ ; (b)  $p = 0.001$ ; (c)  $p = 0.002$ ; (d)  $p = 0.005$ ; (e)  $p = 0.01$ ; (f)  $p = 0.05$ ; (g)  $p = 0.1$ ; (h)  $p = 0.5$ ; (i)  $p = 1$ .

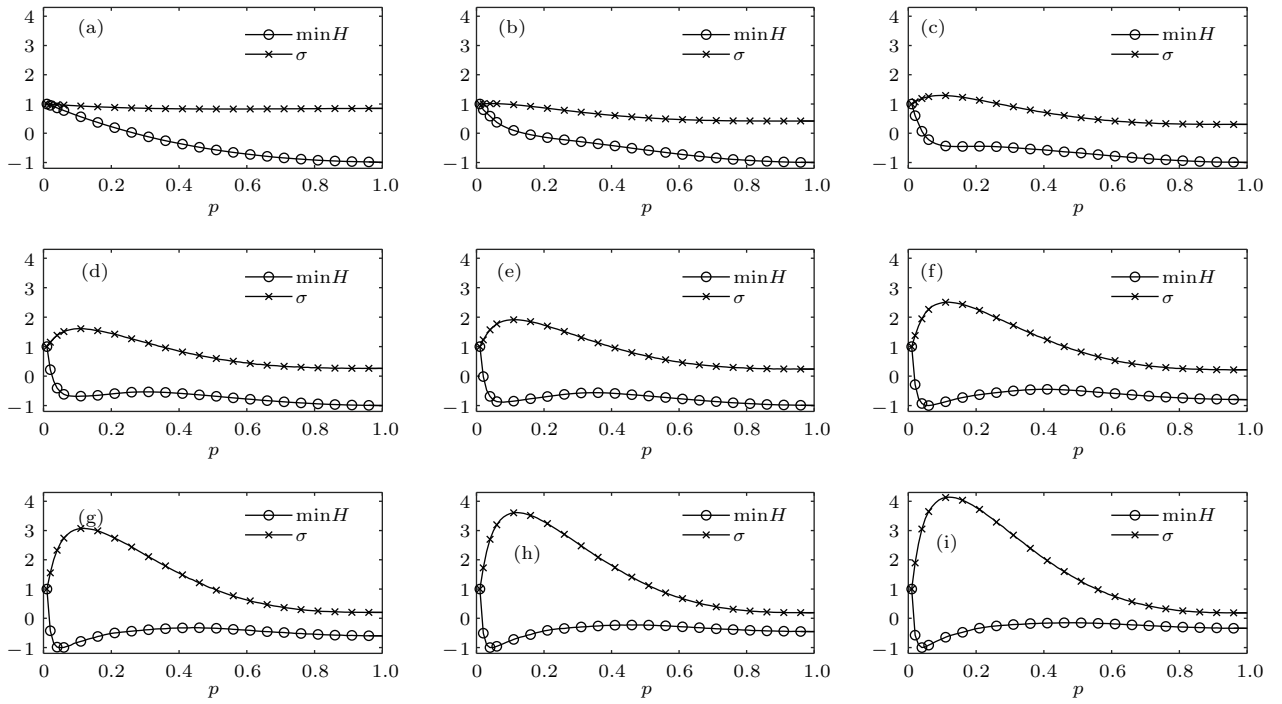


图3 不同网络规模  $n$  及重连概率  $p$  下的  $\min H$  和小世界指数  $\sigma$  (a)  $n = 10$ ; (b)  $n = 25$ ; (c)  $n = 50$ ; (d)  $n = 75$ ; (e)  $n = 100$ ; (f)  $n = 150$ ; (g)  $n = 200$ ; (h)  $n = 250$ ; (i)  $n = 300$

Fig. 3. The minimum structure entropy  $\min H$  and small-world coefficient  $\sigma$  at different network size  $n$  and rewiring probability  $p$ : (a)  $n = 10$ ; (b)  $n = 25$ ; (c)  $n = 50$ ; (d)  $n = 75$ ; (e)  $n = 100$ ; (f)  $n = 150$ ; (g)  $n = 200$ ; (h)  $n = 250$ ; (i)  $n = 300$ .

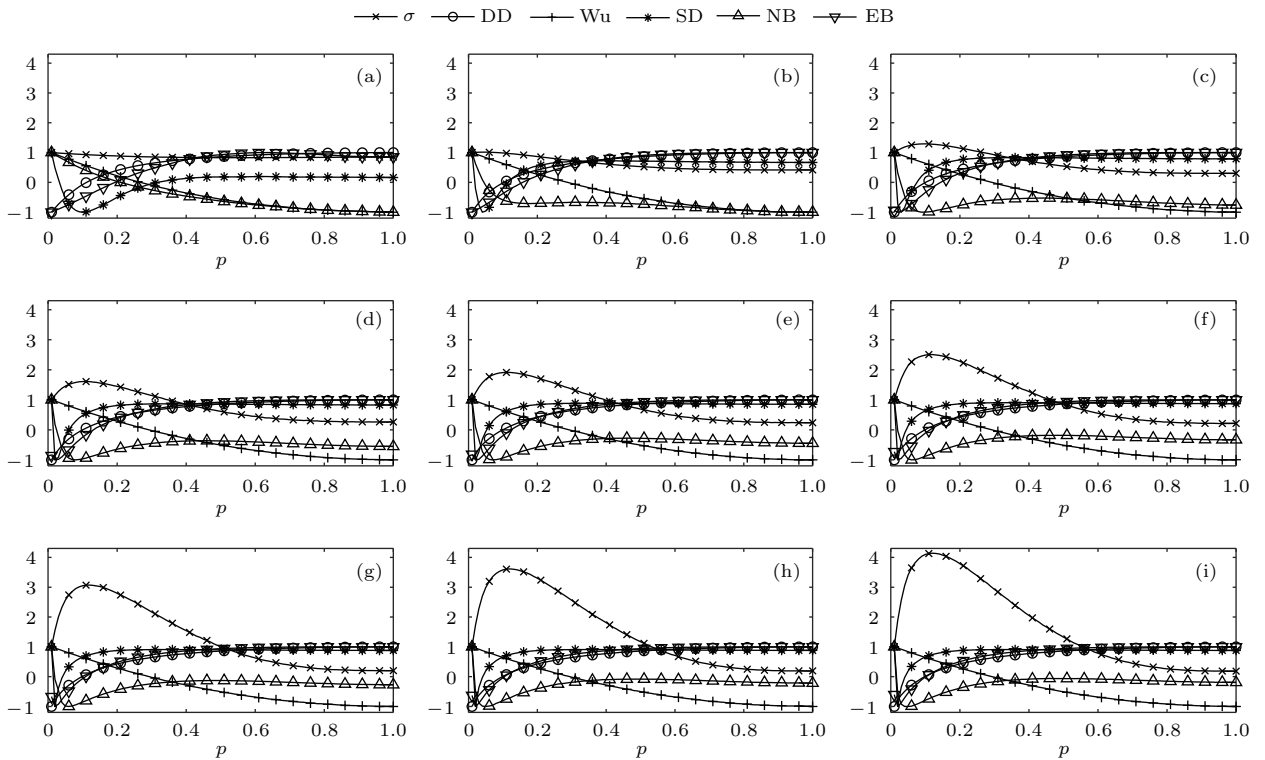


图4 不同网络规模  $n$  及重连概率  $p$  下的小世界指数  $\sigma$ , DD熵、Wu熵、SD熵、NB熵以及EB熵 (a)  $n = 10$ ; (b)  $n = 25$ ; (c)  $n = 50$ ; (d)  $n = 75$ ; (e)  $n = 100$ ; (f)  $n = 150$ ; (g)  $n = 200$ ; (h)  $n = 250$ ; (i)  $n = 300$

Fig. 4. Small-world coefficient  $\sigma$ , degree distribution entropy (DD entropy), Wu entropy (Wu entropy), node and edge difference entropy (SD entropy), node betweenness entropy (NB entropy), edge betweenness entropy (EB entropy) at different network size  $n$  and rewiring probability  $p$ : (a)  $n = 10$ ; (b)  $n = 25$ ; (c)  $n = 50$ ; (d)  $n = 75$ ; (e)  $n = 100$ ; (f)  $n = 150$ ; (g)  $n = 200$ ; (h)  $n = 250$ ; (i)  $n = 300$ .

由图 3 可知, 当网络规模较小时, 网络的小世界性不明显,  $\min H$  随  $p$  的增加单调下降; 当网络规模较大且  $p$  较小时, 网络具有显著的小世界性, 此时  $\min H$  随重连概率先下降再上升. 因此,  $K$ -阶结构熵认为最近邻耦合网络、随机网络以及 WS 小世界网络的异构性依次增强, 且网络的小世界性越强, 其异构性则越强. 可见,  $K$ -阶结构熵模型能够较好地依据小世界属性来反映网络的异构性.

为了进一步验证  $K$ -阶结构熵的性能, 图 4 绘制了小世界指数  $\sigma$  与 DD 熵、Wu 熵、SD 熵、NB 熵以及 EB 熵随网络节点数  $n$  以及重连概率  $p$  的变化曲线.

由图 4 可见, 在任意网络规模下, DD 熵随  $p$  单调递增, 而 Wu 熵随  $p$  单调递减, 可见 DD 熵与 Wu 熵均难以表征小世界网络的异构性; 此外, DD 熵认为随机网络的异构性弱于规则网络, 与通常理解存在偏差; 当网络规模较小时, SD 熵认为最近邻耦合网络、随机网络以及 WS 小世界网络的异构性依次增强, 与本文结论相似. 但当网络规模较大时, SD 熵随  $p$  的增加迅速上升, 随之保持在随机网络

的异构性水平不变, 但此时网络仍具有较强的小世界性; EB 熵在网络小世界性较强时认为 WS 小世界网络的异构性最强, 但却认为随机网络的异构性弱于最近邻耦合网络, 这一结论同样与普遍理解存在偏差. 只有 NB 熵与本文结论相似.

### 3.2 BA 无标度网络与星型网络

BA 无标度网络依据增长和择优机理而构建 [21], 所生成的网络度分布符合  $\gamma = 3$  的幂律形式, 网络中少数节点拥有极多的连接, 而大多数节点只有少量的连接, 通常认为无标度网络具有较强的异构性 [22,23]. 星型网络是指以中央节点为中心, 其余节点均只与中央节点相连的拓扑结构, 通常可被看作是无标度网络的极端形式, 其异构性要强于一般的无标度网络. 下面基于  $K$ -阶结构熵对二者的异构性进行考察. 文中规定 BA 无标度网络与星型网络均为连通图, 同有  $H^d = H^0 = \log(n)$ .

图 5 仿真了在不同网络规模下, BA 无标度网络与星型网络归一化后的  $K$ -阶结构熵随  $K$  的变化规律.

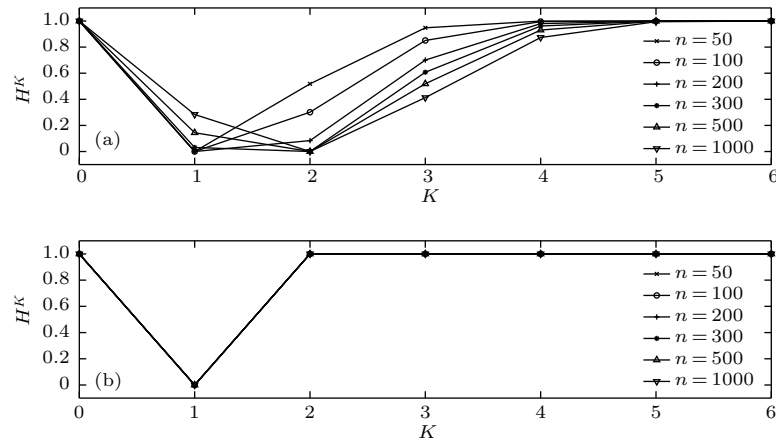


图 5 BA 无标度网络及星型网络在不同网络规模  $n$  及影响力尺度  $K$  下的结构熵 (a) BA 无标度网络; (b) 星型网络

Fig. 5.  $K$ -order structure entropy of BA scale-free networks and star networks at different influence scale  $K$  and the network size  $n$ : (a) BA scale-free network; (b) star network.

由图 5 可见, 在任意网络规模下, BA 无标度网络与星型网络的结构熵均随  $K$  的增加先下降再上升. 此外, BA 无标度网络  $\min H$  所对应的  $K$  值随网络规模的增加逐渐增大, 而星型网络  $\min H$  所对应的  $K$  值与网络规模无关, 即  $K = 1$  时达到最小. 这是由于在任意网络规模下, 星型网络的节点均可在二步内到达其余节点. 图 6 仿真了最近邻耦合网络、随机网络、BA 无标度网络以及星型网络等不同

结构的最小  $H$ , DD 熵、Wu 熵、SD 熵、NB 熵和 EB 熵随网络节点数  $n$  的变化情况.

由图 6 可见,  $\min H$  与 Wu 熵认为, 在任意网络规模下, 最近邻耦合网络、随机网络、BA 无标度网络以及星型网络的异构性依次增强; 且随着网络规模的增加, 以上拓扑结构的  $\min H$  与 Wu 熵均呈上升趋势. 在此观点下, 星型网络的异构性随网络规模的增加而减弱, 这可以解释为在任意网

络规模下, 星型网络只拥有一个中心节点, 其余节点均是度1节点. 随着网络规模的增加, 重要性相同的度1节点的数量显著增加, 这导致了网络的异构性有所减弱. 此结论比SD熵和DD熵认为的星型网络的异构性随其规模的增大略有增强后保持不变的解释更为合理; 此外, DD熵区分随机网络与BA无标度网络的能力不足, 且无法充分反

映最近邻耦合网络异构性随其网络规模的演化规律; NB熵对最近邻耦合网络、随机网络以及BA无标度网络异构性随其网络规模演化规律的解释与K-阶结构熵相同, 但却无法充分反映星型网络的异构性; EB熵认为随机网络与BA无标度网络的异构性弱于最近邻耦合网络, 与通常解释存在偏差.

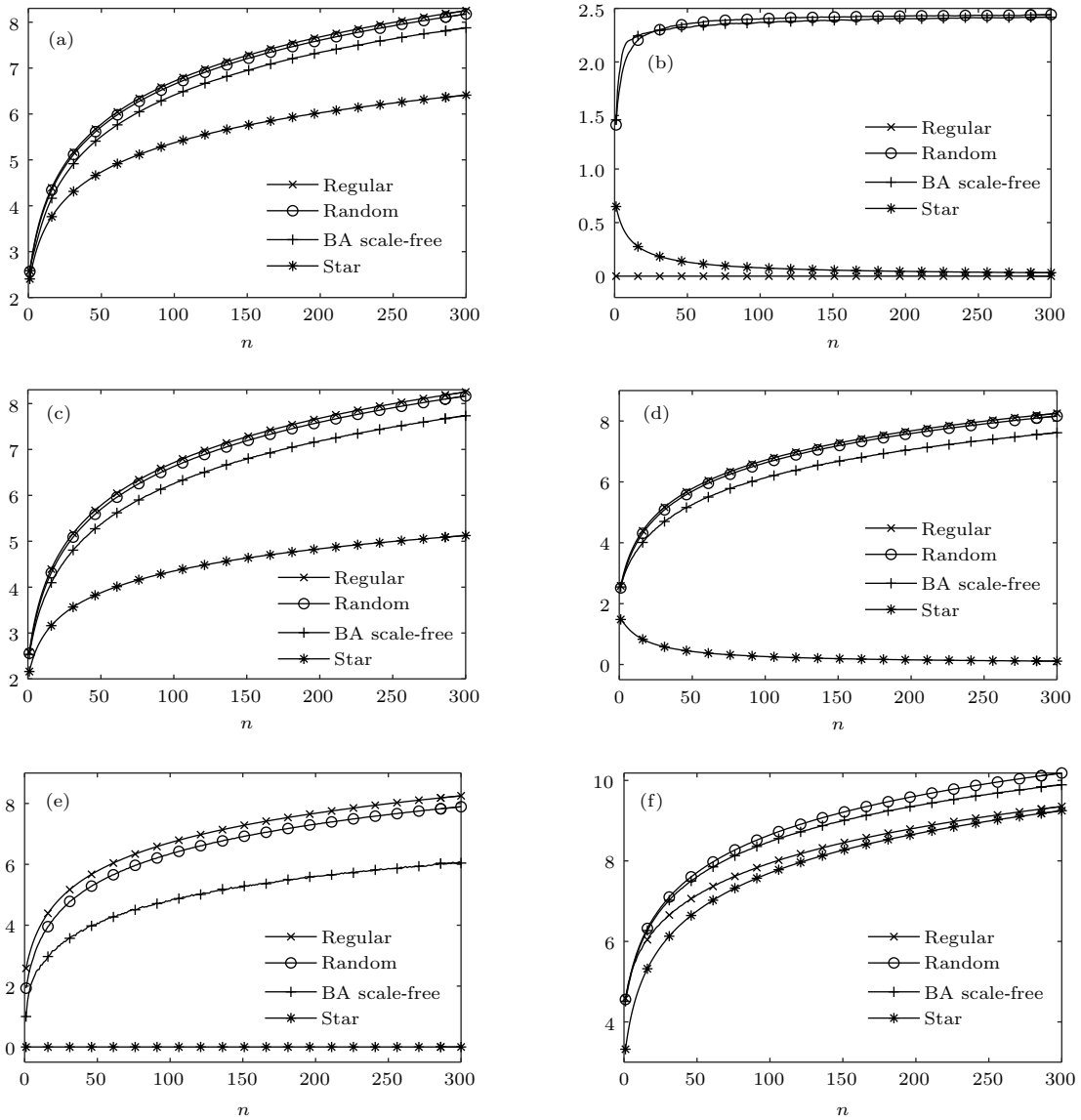


图6 最近邻耦合网络、随机网络、BA无标度网络、星型网络在不同网络规模  $n$  下的  $\min H$ 、DD熵、Wu熵、SD熵、NB熵以及EB熵 (a)  $\min H$ ; (b) DD熵; (c) Wu熵; (d) SD熵; (e) NB熵; (f) EB熵

Fig. 6. The  $\min H$ , DD entropy, Wu entropy, SD entropy, NB entropy, EB entropy of nearest-neighbor coupled networks, random networks, BA scale-free networks and star networks at different network size  $n$ : (a)  $\min H$ ; (b) DD entropy; (c) Wu entropy; (d) SD entropy; (e) NB entropy; (f) EB entropy.

### 3.3 在规则结构外新增孤立节点的网络

由于最近邻耦合网络、全连接网络和环型网络等规则网络的结构是对称的, 因此其异构性应当较

弱. 现假设在规则网络结构之外再新增若干个孤立节点, 由于孤立节点在网络中的结构与作用是固定不变的, 它们只影响其自身, 不与原规则部分中的任何节点发生作用, 即孤立节点的存在增加了网络

的无序程度, 因此该网络的异构性应弱于原规则结构; 但由于孤立节点与原规则部分的节点存在着结构上的差异, 因此在规则结构外新增孤立节点的网络的异构性应强于同节点数的规则网络.

不失一般性, 以下内容以最近邻耦合网络为例进行分析. 由前文可知, 当最近邻耦合网络未新增孤立节点时,  $K$ -阶结构熵在任意影响力尺度  $K$  下均为  $\log(n)$ ; 现假设最近邻耦合网络的节点数为  $n$ , 额外新增的孤立节点数为  $i$ , 在任意影响力尺度  $K$  下, 易证原最近邻耦合部分各节点的  $K$ -阶邻居数相同, 记为

$$N^K = \begin{cases} 4K + 1, & K < d, \\ n, & K \geq d; \end{cases} \quad (7)$$

而各孤立节点的  $K$  阶邻居数始终为 1. 因此, 得到该网络结构的  $K$ -阶结构熵为

$$H^K = -\frac{n \cdot N^K}{n \cdot N^K + i} \log\left(\frac{N^K}{n \cdot N^K + i}\right) - \frac{i}{n \cdot N^K + i} \log\left(\frac{1}{n \cdot N^K + i}\right). \quad (8)$$

若假设  $n$  和  $i$  不变, 由 (8) 式易证  $H^K$  将随着  $K$  的增加单调减少, 因此有  $H^d \leq H^0$ , 即  $H^d \leq \log(n+i)$ ; 且最大影响力尺度  $d$  下的结构熵  $H^d$  与最小结构熵  $\min H$  相等. 在最大影响力尺度  $d$  下, 对于原最近邻耦合网络部分的节点有  $N^d = n$ , 因此网络的  $H^d$  即  $\min H$  为

$$\begin{aligned} H^d &= \min H \\ &= -\frac{n^2}{n^2 + i} \log\left(\frac{n}{n^2 + i}\right) - \frac{i}{n^2 + i} \log\left(\frac{1}{n^2 + i}\right) \\ &= \log(n^2 + i) - \frac{n^2}{n^2 + i} \log(n). \end{aligned} \quad (9)$$

由 (9) 式易证  $\log(n) \leq H^d$ . 综上有  $\log(n) \leq H^d \leq \log(n+i)$ , 即可认为在  $n$  节点最近邻耦合网络外

新增  $i$  个孤立节点后, 其网络异构性介于  $n$  与  $n+i$  节点最近邻耦合网络之间, 与前文讨论相符. 事实上, 节点为  $n$  的任意连通图额外新增  $i$  个孤立节点后, 其最大影响力尺度  $d$  下的异构性  $H^d$  均符合 (9) 式的结果, 但网络最小结构熵  $\min H$  未必等于  $H^d$ . 此外, 随着孤立节点的增多, 当  $i \gg n^2$  时, (9) 式可近似为  $\log(i)$ , 这可认为当孤立节点的数量远超原网络 (任意连通图) 的规模时, 其网络异构性将更多地受到孤立节点的影响, 也是符合实际的. 值得注意的是, 任意连通图的  $H^d$  始终为  $\log(n)$ . 因此只有当网络非连通时,  $H^d$  的讨论才有意义.

图 7 仿真了 20 节点最近邻耦合网络新增  $i$  个孤立节点后, 其  $K$ -阶结构熵随  $K$  值的变化规律. 图 8 则绘制了 20 节点最近邻耦合网络新增  $i$  个孤立节点后, 其  $H^d(\min H)$ , DD 熵、Wu 熵、SD 熵、NB 熵以及 EB 熵随  $i$  的变化规律.

如图 8 所示, Wu 熵、SD 熵、NB 熵以及 EB 熵均认为  $n$  节点最近邻耦合网络新增  $i$  个孤立节点后, 其网络异构性要强于  $n+i$  节点最近邻耦合网络; 但 Wu 熵、NB 熵以及 EB 熵却认为  $n$  节点最近邻耦合网络增添孤立节点前后的网络异构性不发生变化, 这一结论忽略了孤立节点对网络结构的影响; 而 SD 熵认为与原  $n$  节点最近邻耦合网络相比, 网络异构性随增添孤立节点个数的增加先增大后减少, 目前对于这一结论缺乏合理的解释; DD 熵认为  $n$  节点最近邻耦合网络增添  $i$  个孤立节点后, 其网络异构性要比  $n$  和  $n+i$  节点最近邻耦合网络都弱, 与前文讨论不一致. 由此可见,  $K$ -阶结构熵对最近邻耦合网络新增孤立节点后的网络异构性的解释

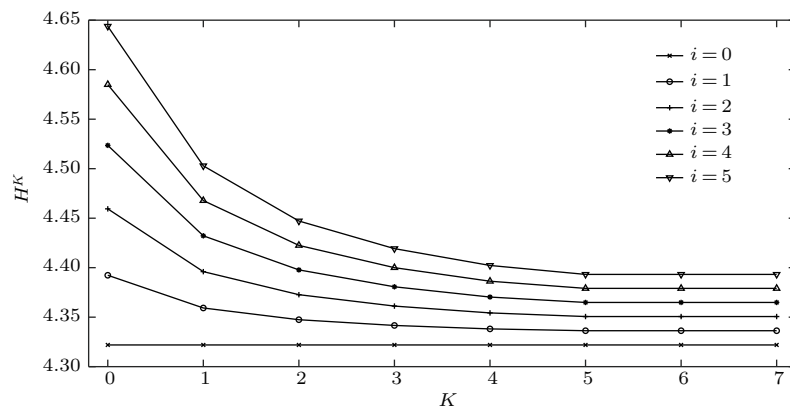


图 7 在 20 节点最近邻耦合网络外新增  $i$  个孤立节点后的网络在不同  $K$  值下的结构熵

Fig. 7.  $K$ -order structure entropy at different influence scale  $K$  of different networks generated by adding isolated nodes to a 20 nodes nearest-neighbor coupled network, where the number of isolated nodes is  $i$ .

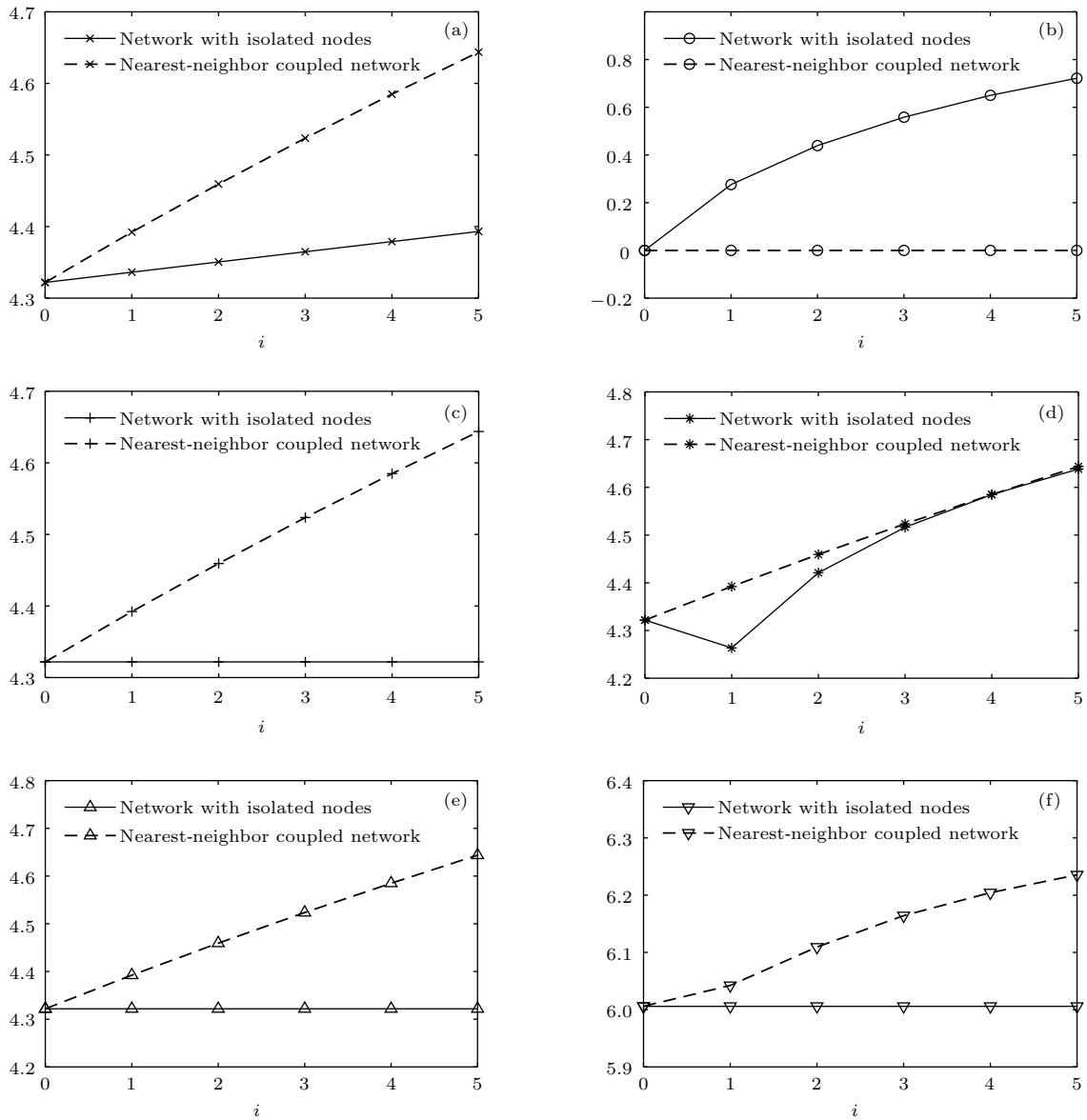


图8 在20节点最近邻耦合网络外新增*i*个孤立节点后的网络与20+*i*节点最近邻耦合网络的 $H^d(\min H)$ , DD熵、Wu熵、SD熵、NB熵以及EB熵 (a)  $H^d(\min H)$ ; (b) DD熵; (c) Wu熵; (d) SD熵; (e) NB熵; (f) EB熵

Fig. 8. The  $H^d(\min H)$ , DD entropy, Wu entropy, SD entropy, NB entropy, EB entropy of  $(20 + i)$  nodes nearest-neighbor coupled networks and graphs generated by adding  $i$  isolated nodes to 20 nodes nearest-neighbor coupled networks: (a)  $H^d(\min H)$ ; (b) DD entropy; (c) Wu entropy; (d) SD entropy; (e) NB entropy; (f) EB entropy.

更为合理。其实,  $K$ -阶结构熵与Wu熵类似, 均将节点作为基本单元来评价网络的异构性, 但本文将节点自身纳入影响力范围, 从而避免了Wu熵中由于孤立节点度值为0所造成的分析上的不便。

综上所述,  $K$ -阶结构熵能够较好地刻画并区分规则网络、随机网络、WS小世界网络、BA无标度网络以及星型网络的异构性, 认为以上网络的异构性依次增强, 与文献[24]的理论研究保持一致。其中,  $K$ -阶结构熵能够较好地依据小世界属性来刻画WS小世界网络的异构性, 且对星型网络异构性随其规模演化规律的解释也更为合理。此外, 本

文还分析了在规则结构外新增孤立节点后的网络的异构性, 认为 $n$ 节点规则网络增添*i*个孤立节点后, 其异构性介于 $n$ 和 $n + i$ 节点规则网络之间, 并初步探索了 $\min H$ 和 $H^d$ 之间的关系。

### 3.4 美国西部电网

为了进一步阐述 $K$ -阶结构熵评估网络异构性的能力, 本文以某一公用数据集为例进行说明。该数据集为美国西部电网[18], 描述了美国西部高压输电网的拓扑结构, 共有4941个节点, 代表发电厂、

变电站以及中间电气连接点等场所; 网络共有 6594 条边, 代表输电线和变压器支路. 该网络规定各节点无差别, 且忽略输电线的物理构造及电气参数差异, 即认为各边无差异, 且无向无权.

Watts 和 Strogatz 首次揭示了美国西部电网的小世界性, 认为该网络同时具有近似于同规模随机网络的较短的特征路径长度, 以及远大于随机网络的平均聚类系数<sup>[18]</sup>. 表 1 列出了美国西部电网与同规模随机网络的平均聚类系数、特征路径长度等参数.

表 1 美国西部电网及同规模随机网络的网络参数  
Table 1. Network features of the Western Power Grid of the United States and random networks.

网络类型	节点数	边数	平均度	平均聚类系数	特征路径长度
西部电网	4941	6594	2.67	0.16020	18.9892
随机网络	4941	6885	2.67	0.000314	8.1527

由表 1 可知, 美国西部电网的特征路径长度仅为随机网络的 2.33 倍, 而平均聚类系数是随机网络的 510 倍, 以文献<sup>[18]</sup>的观点可认为美国西部电网是具有小世界性的. 下面将基于  $K$ -阶结构熵对美国西部电网进行考察.

图 9 仿真了美国西部电网以及同规模最近邻耦合网络、WS 小世界网络、随机网络和 BA 无标度网络在不同影响力尺度  $K$  下的结构熵. 由图 9 可见, 美国西部电网的结构熵随  $K$  先减少后增加, 其  $\min H$  约为 11.95, 与  $p = 0.01$  下的 WS 小世界网络最为接近 ( $\min H = 11.997$ ), 远大于 BA 无标度网络 ( $\min H = 11.544$ ), 且小于随机网络 ( $\min H = 12.187$ ). 因此, 从最强异构性  $\min H$  的角度出发, 可认为美国西部电网的异构性与小世界网络相似, 与前文平均聚类系数、特征路径长度等特性的分析结果以及相关文献<sup>[18, 25-27]</sup>的研究均保持一致.

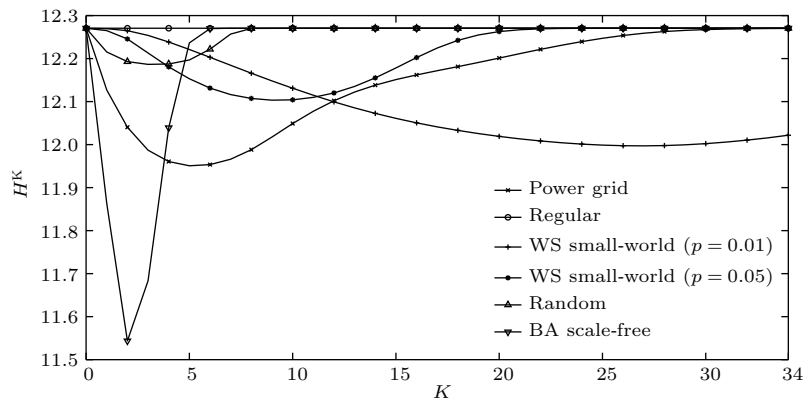


图 9 美国西部电网、最近邻耦合网络、WS 小世界网络、随机网络以及 BA 无标度网络在不同影响力尺度  $K$  下的结构熵

Fig. 9.  $K$ -order structure entropy at different influence scale  $K$  of the Western Power Grid of the United States, nearest-neighbor coupled networks, WS small-world networks, random networks and BA scale-free networks.

前文指出, 网络的异构性还应从结构熵随  $K$  值的变化规律等方面予以考察. 然而, 美国西部电网的结构熵  $H^K$  随  $K$  的变化曲线却与 WS 小世界网络、BA 无标度网络等模型存在较大的差异. 其中, 该电网  $\min H$  所对应的  $K$  值为 5, 远小于  $p = 0.01$  下的 WS 小世界网络 ( $K = 27$ ), 与随机网络 ( $K = 3$ ) 和 BA 无标度网络 ( $K = 2$ ) 更为接近. 事实上, Barabási 和 Albert 等以美国西部电网为研究对象时指出, 该电力网络的度分布符合  $\gamma \approx 4$  的幂率形式, 具有一定的无标度性<sup>[21]</sup>, 而 WS 小世界模型的度分布却为泊松形式. 此外, BA 无标度模

型的度分布虽为幂率形式, 但其幂指数  $\gamma = 3$ , 与美国西部电网比较也存在差异; 且 BA 无标度模型的聚类系数较小, 无法体现美国西部电网的小世界性.

可见, 若从  $\min H$  的角度出发, 可认为美国西部电网的异构性与小世界网络相似, 但若从  $K$ -阶结构熵随  $K$  值的变化规律的角度出发, 美国西部电网的异构性又与随机网络、WS 小世界网络、BA 无标度网络均不相同. 目前, 本文未能定量地给出  $H^K$  随  $K$  值的变化规律与网络特性之间的关系, 不能确定美国西部电网的全部性质, 且美国西部电网

也十分复杂,并不完全符合WS小世界、BA无标度等传统的理论模型.因此,深入探寻结构熵随 $K$ 值的变化规律与网络特性之间的关系、探寻幂指数与 $\min H$ 等指标的关系、提出更符合诸如美国西部电网等实际网络的模型,均为未来的工作提出了新的要求.

值得注意的是,本文是通过定义 $K$ -阶邻居数来计算 $K$ -阶结构熵的, $K$ -阶邻居数表征了某节点在 $K$ 步内可达的节点总数,在一定程度上反映了节点的重要性,而节点重要性排序是研究网络传播效率等网络功能的重要前提.显然,节点重要性的排序将依 $K$ 值的选择而有所不同.依前文所述,当网络具有最强的异构性时,各节点重要性分化达到最大限度.因此,选择 $\min H$ 所对应的 $K$ 值作为节点重要性排序的参数是最为合理的.此外, $K$ -阶邻居数还可以作为社区结构划分以及研究网络传播效率的依据.以社区结构划分为例,由于社区内部节点间的连接较为密集,而社区之间的连接则相对稀疏,因此社区中心节点的 $K$ -阶邻居数往往较多,而社区边界节点的 $K$ -阶邻居数则相对较少.由此,通过寻找被边缘节点所分割的中心节点,便可实现网络社区结构的自然划分.这里仍可将 $\min H$ 所对应的 $K$ 值作为社区结构划分的参数选择.而基于 $K$ -阶邻居数的节点重要性排序、社区结构划分及网络传播等其他研究的具体过程本文不再展开叙述.

## 4 结 论

本文依据网络节点在不同影响力尺度下的 $K$ -阶邻居数提出了一种 $K$ -阶结构熵模型来评价网络的异构性. $K$ -阶结构熵不是依据单一指标对网络进行分析,而是基于不同的影响力尺度 $K$ 对网络的异构性进行多角度的讨论.研究发现, $K$ -阶结构熵随 $K$ 的变化规律、最大影响力尺度 $d$ 下的异构性 $H^d$ 以及网络能够达到的最强异构性 $\min H$ 均可以从不同的角度反映网络的异构性.通过对规则网络、随机网络、WS小世界网络、BA无标度网络和星型网络的理论研究和仿真分析发现, $K$ -阶结构熵模型能够克服DD熵、Wu熵、SD熵、EB熵无法充分依据小世界属性来反映网络异构性的缺陷,并弥补了DD熵以及NB熵对星型网络解释的不足,且 $K$ -阶结构熵对在规则结构外新增孤立节点的网络异构性的解释也优于其他结构熵.此外,本文还基于 $K$ -阶结构熵对美国西部电网进行了讨论.由

于本文的主要工作集中于提出新的网络结构熵模型以及利用经典的网络结构进行实验分析,对于网络结构熵的定量计算、结构熵随 $K$ 值的变化规律与网络结构特性之间的关系、基于 $K$ -阶结构熵的节点重要性排序、社区结构划分等其他应用将会是后续研究的重点.值得注意的是,本研究仅探索了在规则结构外新添孤立节点的特殊情况,对其他拓扑结构添加孤立节点后,结构熵随 $K$ 的变化规律、 $H^d$ 与 $\min H$ 之间的关系等问题也仍需进一步探索.

## 参考文献

- [1] Wang X F, Li X, Chen G R 2006 *Complex Network Theory and Application* (Beijing: Tsinghua University Press) p8 (in Chinese) [汪小帆, 李翔, 陈关荣 2006 复杂网络理论及其应用 (北京: 清华大学出版社) 第8页]
- [2] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D U 2006 *Phys. Rep.* **424** 175
- [3] Vázquez A, Dobrin R, Sergi D, Eckmann J P, Oltvai Z N, Barabási A L 2004 *Proc. Natl. Acad. Sci. USA* **101** 17940
- [4] Pinto P C, Thiran P, Vetterli M 2012 *Phys. Rev. Lett.* **109** 068702
- [5] Yang Y Y, Xie G 2016 *Inform. Process. Manage.* **52** 911
- [6] Newman M E J 2004 *Eur. Phys. J. B* **38** 321
- [7] Lermansinkoff D B, Barch D M 2016 *Neuroimage-Clin.* **10** 96
- [8] Grabow C, Grosskinsky S, Timme M 2011 *Eur. Phys. J. B* **84** 613
- [9] Marceau V, Noël P A, Hébert-Dufresne L, Allard A, Dubé L J 2010 *Phys. Rev. E* **82** 036116
- [10] Wu J, Tan Y J, Zheng H Z, Zhu D Z 2007 *System Eng. Theor. Prac.* **27** 101 (in Chinese) [吴俊, 谭跃进, 郑宏钟, 朱大智 2007 系统工程理论与实践 **27** 101]
- [11] Solé R V, Valverde S 2004 *Lect. Notes Phys.* **650** 189
- [12] Cai M, Du H F, Ren Y K, Feldman M 2011 *Acta Phys. Sin.* **60** 110513 (in Chinese) [蔡萌, 杜海峰, 任义科, 费尔德曼 M 2011 物理学报 **60** 110513]
- [13] Yoon J, Blumer A, Lee K 2006 *Bioinformatics* **22** 3106
- [14] Zhang Q, Li M Z, Deng Y 2014 arXiv:1407.0097v1 [cs, SI]
- [15] Cai M, Du H F, Feldman M 2014 *Acta Phys. Sin.* **63** 060504 (in Chinese) [蔡萌, 杜海峰, 费尔德曼 M 2014 物理学报 **63** 060504]
- [16] Xiao L P, Meng H, Li D Y 2008 *Geomatics Inf. Sci. Wuhan Univ.* **33** 379 (in Chinese) [肖俐平, 孟晖, 李德毅 2008 武汉大学学报 (信息科学版) **33** 379]
- [17] Gan W Y, He N, Li D Y, Wang J M 2009 *J. Software* **20** 2241 (in Chinese) [淦文燕, 赫南, 李德毅, 王建民 2009 软件学报 **20** 2241]
- [18] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [19] Humphries M D, Gurney K, Prescott T J 2006 *Proc. R. Soc. B* **273** 503
- [20] Humphries M D, Gurney K 2008 *PLoS One* **3** e0002051
- [21] Barabási A L, Albert R 1999 *Science* **286** 509

- [22] Pan Z F, Wang X F 2006 *Acta Phys. Sin.* **55** 4508 (in Chinese) [潘灶烽, 汪小帆 2006 物理学报 **55** 4508]  
 [23] Wang B H, Zhou T, Wang W X, Li J, Jiang P Q 2006 *Acta Phys. Sin.* **55** 4051 (in Chinese) [汪秉宏, 周涛, 王文旭, 李季, 蒋品群 2006 物理学报 **55** 4051]  
 [24] Yang J M, Zhang N 2005 *J. Univ. Shanghai Sci. Technol.* **27** 413 (in Chinese) [杨建民, 张宁 2005 上海理工大学学报 **27** 413]  
 [25] Holmgren Å J 2006 *Risk Anal.* **26** 955  
 [26] Newman M E J 2003 *SIAM Rev.* **45** 167  
 [27] Chassin D P, Posse C 2005 *Physica A* **355** 667

## Network heterogeneity based on $K$ -order structure entropy\*

Huang Li-Ya<sup>†</sup> Huo You-Liang Wang Qing Cheng Xie-Feng

(College of Electronic and Optical Engineering and College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

( Received 19 July 2018; revised manuscript received 8 November 2018 )

### Abstract

Structure entropy can evaluate the heterogeneity of complex networks, but traditional structure entropy has deficiencies in comprehensively reflecting the global and local network features. In this paper, we define a new structure entropy based on the number of the  $K$ -order neighbor nodes which refer to those nodes which a node can reach within  $K$  steps. It can be supposed that the more  $K$ -order neighbors a node has, the more important role the node plays in the network structure. Combining the formula of Shannon entropy, the  $K$ -order structure entropy can be defined and figured out to explain the differences among the relative importance among nodes. Meanwhile, the new structure entropy can describe the network heterogeneity from the following three aspects. The first aspect is the change tendency of structure entropy with the value of  $K$ . The second aspect is the structure entropy under a maximum influence scale  $K$ . The last aspect is the minimum value of the  $K$ -order structure entropy. The simulation compares the heterogeneities of five classic networks from the above three aspects, and the result shows that the heterogeneity strengthens in the from-weak-to-strong sequence: regular network, random network, WS (Watts-Strogatz) small-world network, BA (Barabási-Albert) scale-free network and star network. This conclusion is consistent with the previous theoretical research result, but hard to obtain from the traditional structure entropy. It is remarkable that the  $K$ -order structure entropy can better evaluate the heterogeneity of WS small-world networks and suggests that the greater small-world coefficients a network has, the stronger heterogeneity the network has. Besides, the  $K$ -order structure entropy can fully reflect the heterogeneity variation of star networks with network size, and reasonably explain the heterogeneity of regular networks with additional isolated nodes. It suggests that when  $i$  additional isolated nodes are added to a regular network with  $n$  nodes, the new network has weaker heterogeneity than the old one, but has stronger heterogeneity than the regular network with  $n+i$  nodes. Finally, the validity of the  $K$ -order structure entropy is further confirmed by simulations for the western power grid of the United States. Based on the minimum value of the  $K$ -order structure entropy, the heterogeneity of the western power grid is the closest to that of WS small-world networks.

**Keywords:** complex network, heterogeneity, entropy

**PACS:** 89.75.-k, 89.75.Fb

**DOI:** 10.7498/aps.68.20181388

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 61373065, 61271334).

† Corresponding author. E-mail: [huangly@njupt.edu.cn](mailto:huangly@njupt.edu.cn)