

复杂网络谱粗粒化方法的改进算法

周建 贾贞 李科赞

Improved algorithm of spectral coarse graining method of complex network

Zhou Jian Jia Zhen Li Ke-Zan

引用信息 Citation: *Acta Physica Sinica*, 66, 060502 (2017) DOI: 10.7498/aps.66.060502

在线阅读 View online: <http://dx.doi.org/10.7498/aps.66.060502>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2017/V66/I6>

您可能感兴趣的其他文章

Articles you may be interested in

强迫 Lorenz 系统的可预报性研究

Predictability of forced Lorenz system

物理学报.2017, 66(6): 060503 <http://dx.doi.org/10.7498/aps.66.060503>

基于广义似然比判决的混沌信号重构方法

Reconstruction algorithm of chaotic signal based on generalized likelihood ratio threshold-decision

物理学报.2017, 66(4): 040503 <http://dx.doi.org/10.7498/aps.66.040503>

含三个忆阻器的六阶混沌电路研究

Research on a six-order chaotic circuit with three memristors

物理学报.2017, 66(4): 040502 <http://dx.doi.org/10.7498/aps.66.040502>

忆感器文氏电桥振荡器

Meminductive Wein-bridge chaotic oscillator

物理学报.2017, 66(2): 020502 <http://dx.doi.org/10.7498/aps.66.020502>

余维-1 非光滑分岔下的簇发振荡及其机理

Bursting oscillations as well as the mechanism with codimension-1 non-smooth bifurcation

物理学报.2017, 66(2): 020501 <http://dx.doi.org/10.7498/aps.66.020501>

复杂网络谱粗粒化方法的改进算法*

周建¹⁾ 贾贞^{1)†} 李科赞²⁾

1)(桂林理工大学理学院, 桂林 541004)

2)(桂林电子科技大学数学与计算科学学院, 桂林 541004)

(2016年10月2日收到; 2016年11月21日收到修改稿)

大规模网络的同步问题是网络科学的重要研究课题之一. 粗粒化方法提供了一种将大规模网络转化为小规模网络, 同时又能较好地保持原始网络的拓扑性质或动态特性的研究途径, 其中比较有代表性的谱粗粒化方法能较好地保持初始网络的同步能力. 然而, 谱粗粒化方法在实际计算中计算量大、对实际大规模网络可执行性差. 本文提出一种改进的谱粗粒化算法, 能大幅减少计算量, 同时获得更好的谱粗粒化效果. 通过理论分析和大量的数值仿真实验验证了所提改进算法的粗粒化效果和计算量都明显优于原谱粗粒化方法.

关键词: 复杂网络, 同步, 谱粗粒化, 改进算法

PACS: 05.45.-a, 05.45.Xt

DOI: 10.7498/aps.66.060502

1 引言

同步是自然界、人类社会和工程技术领域中一种常见的集合运动现象, 如蛙声齐鸣、有节律的鼓掌等. 近二十多年来, 研究人员对复杂动态网络的同步进行了大量研究, 特别是对于中尺度网络的同步研究取得了丰富的研究成果^[1-11]. 然而, 许多实际网络是拥有成千上万甚至上亿节点的大规模网络, 研究大规模耦合复杂动态网络同步常常会产生大量的耦合微分方程, 给计算和仿真实验都带来巨大困难, 许多中尺度网络同步算法在大规模网络研究中难以实现, 因此人们提出了一些粗粒化方法试图将大规模网络转化为中尺度网络来研究^[12-17]. 粗粒化方法通过合并相似节点将大规模网络缩减为中尺度网络, 同时尽量保持原始网络的某些重要性质, 如拓扑特性或动力学性质等. 近年来, 在粗粒化的研究方面也取得了一些重要进展. 例如, Kim^[18]提出了一种粗粒化方法能保持原始网络的度分布、聚类系数、同配系数等性质; Chen等^[19]提出了一种基于网络度的粗粒化方法, 能保持原始

网络平衡态统计分布的一致性. 在这些粗粒化方法中, 比较典型的是由Gfeller和Rios^[20,21]在2008年提出的谱粗粒化(spectral coarse graining, SCG)方法, 通过分析网络结构矩阵的特征值谱, 合并网络中结构相同和相似的节点, 从而有效地将大规模网络缩减为小规模网络同时又能较好地保持初始网络的同步能力. Chen等^[15]进一步研究了SCG方法在聚类网络中的应用效果, Zeng和Lü^[16]进一步提出了基于有向网络的SCG方法, 能很好地保持粗粒化网络的同步能力.

SCG方法在保持网络同步能力方面独占优势, 然而在确定网络中的哪些节点被合并时需要反复判断, 这样会产生非常巨大的计算量, 其计算复杂度为 $O(N^3)$ ^[16]. 对于现实中少则成千上万节点的网络, 多则上百万甚至上亿节点的大规模网络, 其巨大的计算量致使SCG方法在实际中难以执行. 为了克服SCG方法计算量大的缺陷, 本文提出了一种改进的SCG方法, 记为ISCG(improved spectral coarse graining)算法, 将计算复杂度从 $O(N^3)$ 降到 $O(N^2)$, 同时还改善了粗粒化效果, 所得粗粒化网络能够更好地保持初始网络的同步能力.

* 国家自然科学基金(批准号: 61563013, 61663006)资助的课题.

† 通信作者. E-mail: jjjzzz0@163.com

2 网络同步能力的刻画

考虑 N 个节点的复杂网络, 其动力学方程为

$$\dot{x}_i(t) = \mathbf{F}(x_i(t)) - \delta \sum_{j=1}^N L_{ij} \mathbf{H}(x_j(t))$$

$$(i = 1, 2, \dots, N), \quad (1)$$

其中 $x_i \in \mathbf{R}^m$ 表示第 i 个节点的 m 维状态变量; $\dot{x}_i(t) = \mathbf{F}(x_i(t))$ 是第 i 个节点的动态方程; $\delta > 0$ 为耦合强度; $\mathbf{H}(\cdot): \mathbf{R}^m \rightarrow \mathbf{R}^m$ 是节点内部耦合函数; Laplacian 矩阵 $\mathbf{L} = (L_{ij})_{N \times N} \in \mathbf{R}^{N \times N}$ 为外部耦合矩阵, 刻画了网络的耦合拓扑结构. 若网络节点 j 和 i ($i \neq j$) 相连, 则 $L_{ij} = 1$, 否则 $L_{ij} = 0$; Laplacian 矩阵 $\mathbf{L}$ 的对角线元素满足 $L_{ii} = - \sum_{j=1, j \neq i}^N L_{ij}$,

即矩阵 $\mathbf{L}$ 满足耗散耦合条件: $\sum_{j=1}^N L_{ij} = 0$. 当网络是无向无权的连通网络时, 对应的矩阵 $\mathbf{L}$ 是对称不可约的正半定矩阵, 具有非负特征值且满足 $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_N$. 网络 (1) 同步状态 $x_1(t) = x_2(t) = \dots = x_N(t) = s(t)$ 的同步流形 $s(t)$ 满足 $\dot{s}(t) = \mathbf{F}(s(t))$, 它可以是孤立节点的平衡点、周期轨道或混沌轨道. 将方程 (1) 线性化, 令 ξ_i 为第 i 个节点状态的变分, 得到变分方程

$$\dot{\xi}_i = \mathbf{D}\mathbf{F}(s)\xi_i - \delta \sum_{j=1}^N L_{ij} \mathbf{D}\mathbf{H}(s)\xi_j, \quad (2)$$

其中 $\mathbf{D}\mathbf{F}(s)$ 和 $\mathbf{D}\mathbf{H}(s)$ 分别是 $\mathbf{F}(s)$ 和 $\mathbf{H}(s)$ 关于 $s(t)$ 的 Jacobi 矩阵. 将方程 (2) 对角化可得到

$$\dot{\eta}_k = [\mathbf{D}\mathbf{F}(s) - \delta \lambda_k \mathbf{D}\mathbf{H}(s)]\eta_k$$

$$(k = 2, \dots, N), \quad (3)$$

η_k 是矩阵 $\mathbf{L}$ 关于特征值 λ_k 的特征模式. 将方程 (3) 一般化, 得到主稳定方程:

$$\dot{y} = [\mathbf{D}\mathbf{F}(s) - \alpha \mathbf{D}\mathbf{H}(s)]y. \quad (4)$$

该方程的最大 Lyapunov 指数 $L_{\max}$ 是实变量 α 的函数, 称之为网络 (1) 的主稳定函数, 记为 $L_{\max}(\alpha)$. 网络 (1) 的同步化区域是指使 $L_{\max}(\alpha) < 0$ 的实数 α 的取值范围, 记为 SR , 它是由孤立节点动力学函数 $\mathbf{F}(x_i(t))$ 和内连函数 $\mathbf{H}(\cdot)$ 确定的, 只要网络的耦合强度 δ 与耦合矩阵 $\mathbf{L}$ 的特征值之积全部落入同步区域即 $\delta \lambda_k \in SR$ ($k = 2, 3, \dots, N$),

网络 (1) 就能达到完全同步 [22]. 网络的同步区域主要分为以下 4 种类型: 1) 有界区域 (α_1, α_2) ; 2) 无界区域 $(\alpha_1, +\infty)$; 3) 若干个有界区域的并集 $\cup(\alpha_{1i}, \alpha_{2i})$; 4) 空集. 一般地, 3) 和 4) 的情形下网络很难或完全无法达到同步, 幸运的是大多数网络都是 1) 和 2) 两种情形, 此时只有当 $\mathbf{L}$ 的特征值满足 $\lambda_N/\lambda_2 < \alpha_2/\alpha_1$ 或者 $\lambda_2 > \alpha_1/\delta$ 时网络才能达到同步稳定状态 (其中 λ_2 和 λ_N 分别为 Laplacian 矩阵的最小非零特征值和最大特征值). 这意味着当 λ_N/λ_2 越小或者 λ_2 越大时网络更容易达到同步, 因此在同步研究中, 常常用特征值比 λ_N/λ_2 和最小非零特征值 λ_2 两个指标来刻画网络的同步能力 [23].

3 基于网络同步的 SCG 方法

3.1 SCG 方法

由于网络的同步能力可用 λ_2 或 λ_N/λ_2 来刻画, 因此要使粗粒化后的网络仍保持原始网络的同步能力, 就应使粗粒化前后网络的 λ_2 或 λ_N/λ_2 尽可能接近或不变. 2008 年, Gfeller 和 Rios [20, 21] 提出的 SCG 方法就是依此设计的, 在缩减网络规模的同时, 尽可能地保持网络的同步能力. 该方法主要解决了两个关键问题: 一是如何合并节点和更新边; 二是确定哪些节点可以被合并.

首先, 对于如何合并节点和更新边, SCG 方法将初始网络的 N 个节点标记为 $i = 1, 2, \dots, N$, 粗粒化后网络节点标记为 $C = 1, 2, \dots, \tilde{N}$, $\tilde{N}$ 为粗粒化后的网络规模, C 也表示初始网络中的 $\tilde{N}$ 个团, 每个团的节点将被合并为粗粒化网络中的一个节点. 粗粒化后网络的边 (对应新的 Laplacian 矩阵 $\tilde{\mathbf{L}}$) 可通过下面的矩阵乘积进行更新:

$$\tilde{\mathbf{L}} = \mathbf{K}\mathbf{L}\mathbf{Q}, \quad (5)$$

其中矩阵 $\mathbf{K} \in \mathbf{R}^{\tilde{N} \times N}$ 和 $\mathbf{Q} \in \mathbf{R}^{N \times \tilde{N}}$ 分别为

$$\mathbf{K}_{Ci} = \psi_{C,C_i}/|C|; \quad \mathbf{Q}_{iC} = \psi_{C,C_i}. \quad (6)$$

这里, $|C|$ 是团 C 中包含节点的个数; C_i 是节点 i 所在团 C 的标号, ψ 是常用的 Kronecker 函数.

其次, 对于哪些节点被合并, 分别考虑 λ_2 和 λ_N/λ_2 两个指标. 若考虑指标 λ_2 (网络 (1) 的 Laplacian 矩阵 $\mathbf{L}$ 的最小非零特征值), 记 λ_2 对应的特征向量为 $\mathbf{p}^2$, 将 $\mathbf{p}^2$ 中相同或相近的分量对应的节点进行合并, 此时 $\tilde{\mathbf{L}}$ 的最小非零特征值

$\tilde{\lambda}_2$ 与 λ_2 相同或相近. 理论上, 特征向量 $\mathbf{p}^2$ 的两个分量 p_i^2 和 p_j^2 相同或相近, 可表示为 $d_{i,j}^2 \equiv |p_i^2 - p_j^2| / |p_{\max}^2 - p_{\min}^2| \ll 1$, 其中 $p_{\max}^2$ 和 $p_{\min}^2$ 分别为 $\mathbf{p}^2$ 的 N 个分量中的最大值和最小值. 实际计算时, 把区间 $[p_{\min}^2, p_{\max}^2]$ 划分为 I 个等分区间, 对落入同一等分区间的 $\mathbf{p}^2$ 分量对应的节点进行合并. 一般地, I 越小对应的 $\tilde{N}$ 越小, 因此可通过控制 I 值的大小来大致控制粗粒化网络的规模. 相同或相似节点的合并过程及 λ_2 的保持情况如图 1 所示.

同理, 当考虑指标 λ_N/λ_2 (网络 (1) 的 Laplacian 矩阵 $\mathbf{L}$ 中最大与最小非零特征值之比) 时, 合并节点需同时满足条件 $d_{i,j}^2 \ll 1$ 和 $d_{i,j}^N \ll 1$. 实际计算时分别把区间 $[p_{\min}^2, p_{\max}^2]$, $[p_{\min}^N, p_{\max}^N]$ 划分为 I 个等分区间 (其中 $\mathbf{p}^N$ 为最大特征值 λ_N 对应的特征向量), 找出 $\mathbf{p}^2$ 和 $\mathbf{p}^N$ 中同时落入等分区间内的分量, 将其对应的节点进行合并. 此时 $\tilde{\mathbf{L}}$ 的特征值比 $\tilde{\lambda}_N/\tilde{\lambda}_2$ 与 λ_N/λ_2 相同或相近, 即能保持粗粒化网络的同步能力.

图 1 显示, 把结构相同或相似的节点进行合并, 网络的最小非零特征值 λ_2 变化不大, 即网络的同步能力能够很好地被保持.

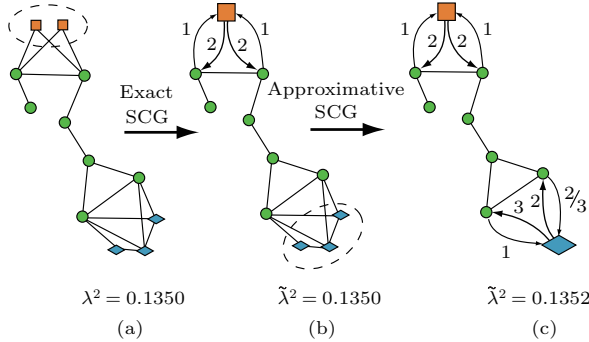


图 1 (网刊彩色) 合并节点过程 [21]

Fig. 1. (color online) Processing of merging nodes [21].

3.2 SCG 方法的局限

我们通过大量仿真实验发现, 很多网络的特征向量分量分布极不均匀, 大多数分量分布相对集中、间距较小, 极少数分量分布特别分散且间距很大. 以节点 $N = 100$ 的 NW 小世界网络为例, 根据特征值 λ_2 对应的特征向量 $\mathbf{p}^2$, 将区间 $[p_{\min}^2, p_{\max}^2]$ 划分为 $I = 15$ 的等分区间, $\mathbf{p}^2$ 的分量落入等分区间内的分布情况如图 2 所示, 图 2 中菱形对应 $\mathbf{p}^2$ 的所有分量, 16 条竖线将区间 $[p_{\min}^2, p_{\max}^2]$ 划分为 15 个等分区间. 从图 2 可看出, $\mathbf{p}^2$ 的分量分布极不均匀. 因此等区间划分节点时可能造成有些距离特别近的两个分量 (对应极相似的节点, 如图 2 中蓝色菱形对应分量 p_{80}^2 和 锰紫色菱形对应分量 p_{65}^2) 被分割到两个相邻区间中, 而与间距更大的分量 (对应不太相似的节点) 划分到同一区间而合并相应节点, 这会影响粗粒化效果. 另一方面, 计算中需要将每一个 $\mathbf{p}^2$ 分量与等分区间所有的端点值进行比较, 依此判断落在哪个等分区间内, 这会造成较大的计算量, 特别是对规模巨大的网络更为明显, 其计算复杂度为 $O(N^3)$ [16]. 例如, 当网络节点 $N = 1000$ 时, SCG 方法的计算量达到 10 亿次. 此外, 由于 $\mathbf{p}^2$ 的分量分布极不均匀, 一些等分区间内没有分量落入, 导致合并后的网络规模 $\tilde{N}$ 随 I 变化不大. 例如, 节点 $N = 1000$ 的 NW 小世界网络, 当划分等区间数 I 为 $10^2, 10^3, 10^4, 10^5, 10^6$ 时, $\tilde{N}$ 分别为 14, 69, 381, 871, 991, 因此 I 值对网络规模 $\tilde{N}$ 的控制效果并不理想, 也在一定程度上会影响粗粒化效果. 综上所述, SCG 化方法存在两方面的缺陷. 一是计算量过大, 对于大规模网络实际计算难以实现; 二是对某些网络粗粒化效果并不十分理想.

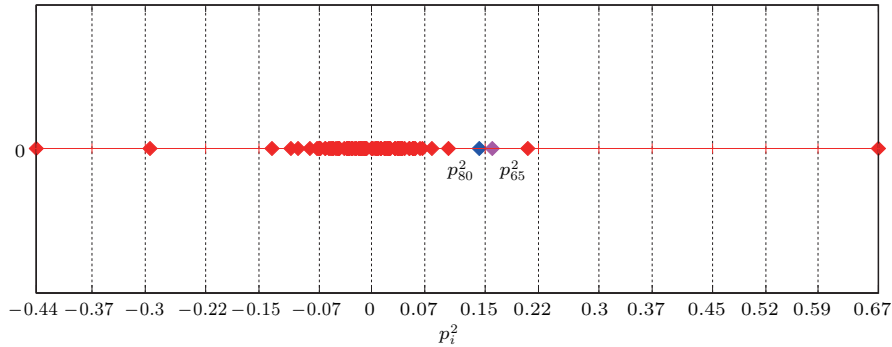


图 2 (网刊彩色) 特征向量 $\mathbf{p}^2$ 的分量 p_i^2 分布情况

Fig. 2. (color online) Distribution of components p_i^2 of eigenvector $\mathbf{p}^2$.

4 改进的SCG算法

本节我们在原方法的基础上提出一种改进的SCG算法, 不仅降低了计算复杂度, 还改善了粗粒化效果, 使粗粒化后的网络更好地保持初始网络的同步能力.

与SCG方法采取等分区间确定合并节点的做法不同, ISCG算法依特征向量分量的分布情况采取分裂聚类方法来确定合并节点. 具体做法是: 先将特征向量分量由小到大排序, 计算两两相邻分量之间的距离, 根据相邻分量之间的间距大小将分量分裂成若干个聚类, 然后将每个聚类中对应的节点合并为一个节点. 选择间距的大小将决定同一聚类中节点的相似程度和粗粒化网络的规模. 例如, 选择较大的前 $\tilde{N} - 1$ 个间距将所有分量分裂成 $\tilde{N}$ 个聚类, 这 $\tilde{N}$ 个聚类对应节点合并为 $\tilde{N}$ 个节点, 从而可以精确控制网络的规模 $\tilde{N}$, $\tilde{N}$ 越大, 相邻分量的间距越小, 合并节点的相似程度越高, 粗粒化效果越好.

下面我们以保持 λ_2 为例进一步说明ISCG算法的步骤. 首先假设 λ_2 的特征向量 $\mathbf{p}^2$ 的分量 $p_1^2, \dots, p_N^2$ 由小到大排序为 $\tilde{p}_1^2, \dots, \tilde{p}_N^2$, 相邻分量之间的间距为 $d_1^2, \dots, d_{N-1}^2$ ($d_i^2 = \tilde{p}_{i+1}^2 - \tilde{p}_i^2$, $i = 1, \dots, N - 1$), 假设选择最大的前3个间距, 分别记为 d_a^2, d_b^2, d_c^2 ($a < b < c$), 这3个间距将 $\tilde{p}_1^2, \dots, \tilde{p}_N^2$ 分裂成4个聚类, 分别为 $[\tilde{p}_1^2, \tilde{p}_a^2]$, $[\tilde{p}_{a+1}^2, \tilde{p}_b^2]$, $[\tilde{p}_{b+1}^2, \tilde{p}_c^2]$, $[\tilde{p}_{c+1}^2, \tilde{p}_N^2]$, 其中间距 d_a^2, d_b^2, d_c^2 分裂区间时对应的左右端点分别为 $\tilde{p}_a^2$ 和 $\tilde{p}_{a+1}^2$, $\tilde{p}_b^2$ 和 $\tilde{p}_{b+1}^2$, $\tilde{p}_c^2$ 和 $\tilde{p}_{c+1}^2$, 再将这4个聚类中对应的节点进行合并, 即得到 $\tilde{N} = 4$ 的粗粒化网络. 由此可见, ISCG算法可以根据需要, 选择一定的间距将原始网络 N 个节点合并到任何指定规模 $\tilde{N}$ 的粗粒化网络, 并且合理区分不同节点间的相似程度, 从而能确保被合并节点之间的相似程度高于不同聚类之间节点的相似程度, 因此粗粒化效果较SCG方法更好. 在计算量方面, 由于ISCG算法只需计算分量之间的间距, 而不需判断比较每一个特征向量分量落入等区间的情况, 只要确定了粗粒化网络模型, 便可一次完成所有节点的合并, 因此大幅减少了计算量, 其计算复杂度为 $O(N^2)$. 此外, ISCG算法能精确控制粗粒化网络规模, 这点优于SCG算法对粗粒化网络规模的控制.

类似地, 对于保持特征值比 λ_N/λ_2 的情况, 可

根据 λ_2 和 λ_N 的特征向量 $\mathbf{p}^2$ 和 $\mathbf{p}^N$ 的分量之间的间距将各分量分裂成不同的聚类, 将同时分裂在同一聚类中对应节点合并为一个节点, 计算过程与上述保持 λ_2 的情形类似, 这里不再赘述.

5 数值仿真

本节将分别对连续时间动态网络和耦合相振子网络模型进行数值仿真, 分别应用ISCG算法和SCG方法对网络进行粗粒化, 比较两种不同粗粒化算法对几种典型网络同步能力的保持效果.

5.1 连续时间动态网络的同步效果

下面对连续时间动态网络模型(1)进行数值仿真. 分别选取四种典型的复杂网络, 即NW小世界网络、ER随机网络、BA无标度网络 and 一类复杂聚类网络^[15], 分别应用ISCG和SCG算法粗粒化网络, 观察并比较粗粒化效果. 原始网络规模均为 $N = 1000$.

图3和图4展示了分别采用ISCG和SCG两种算法对四种典型网络进行粗粒化的结果和网络同步能力的保持情况. 无论对哪种网络, 采取ISCG算法获得的粗粒化网络与原始网络的 λ_2 和 λ_N/λ_2 的近似程度都明显优于SCG方法获得的粗粒化网络, 说明ISCG算法的粗粒化效果和同步能力保持情况优于SCG方法. 从运算时间上看, 在实验中获得相同规模的粗粒化网络, ISCG算法比SCG方法所用的时间大大减少. 以上述 $N = 1000$ 的NW小世界网络为例, 在保持 λ_2 的情形, 获得 $\tilde{N} = 600$ 的粗粒化网络, 采用ISCG算法和SCG方法在MATLAB软件试验中的运行时间分别是2.94 s和18.4 s. 实际上, 由于后者需要经过多次反复实验才能获得规模 $\tilde{N} = 600$ 的粗粒化网络, 故实际获得结果时间远不止18.4 s, 而ISCG算法可以精确控制粗粒化网络规模, 一次完成实验, 故运行时间只需要2.94 s, 可见ISCG算法的运算速度大幅提高. 此外, ISCG算法可以精确获得任意规模的粗粒化网络, 从而能够更精细地展示网络的粗粒化演变过程, 而SCG方法却做不到这点. 当然, 粗粒化网络规模越小, 合并的节点就越多, 对原始网络同步能力的保持效果越差, 但对于同等规模的粗粒化网络, 显然ISCG算法的同步能力的保持效果更好.

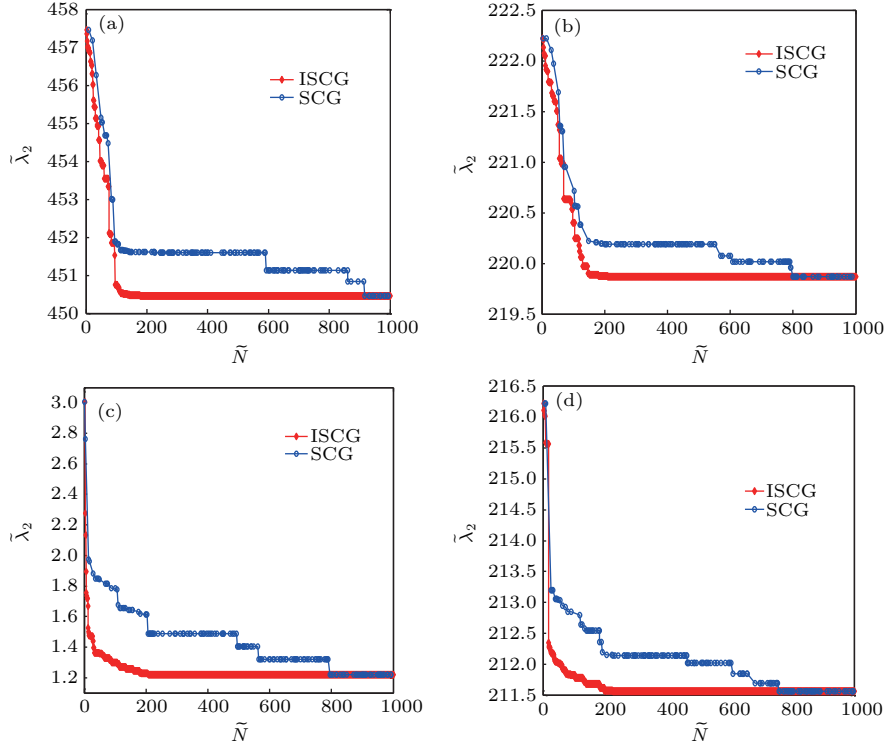


图3 (网刊彩色) 分别采用 ISCG 和 SCG 算法获得粗粒化网络对 λ_2 的保持情况 (a) NW 小世界网络; (b) ER 随机网络; (c) BA 无标度网络; (d) 聚类网络

Fig. 3. (color online) The maintaining of λ_2 obtained by using ISCG and SCG algorithms in coarse graining network: (a) NW network; (b) ER network; (c) BA network; (d) clustered network.

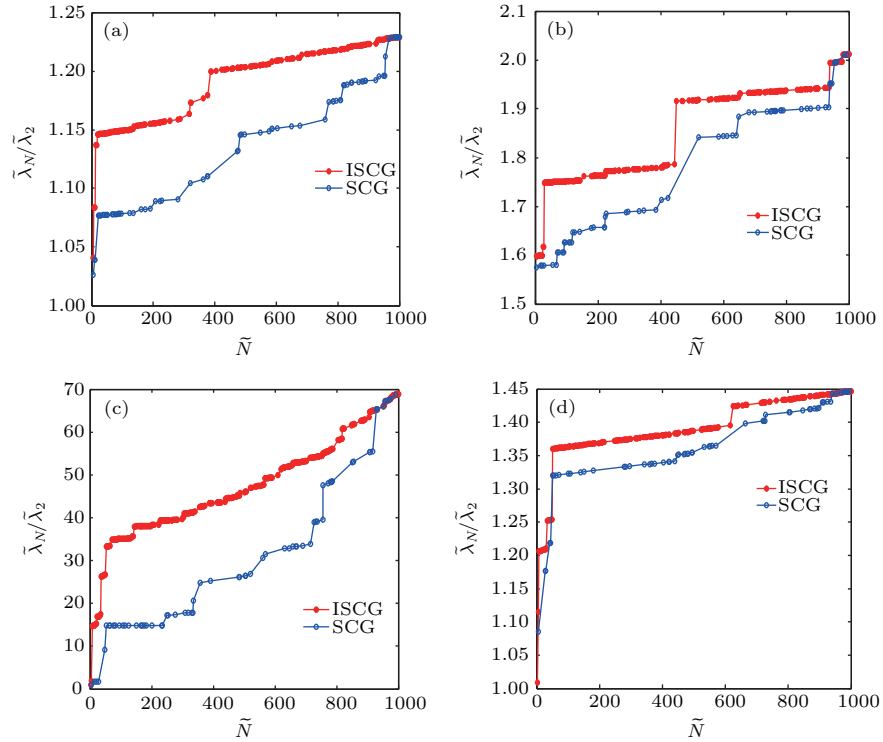


图4 (网刊彩色) 分别采用 ISCG 和 SCG 算法获得粗粒化网络对 λ_N/λ_2 的保持情况 (a) NW 小世界网络; (b) ER 随机网络; (c) BA 无标度网络; (d) 聚类网络

Fig. 4. (color online) The maintaining of λ_N/λ_2 obtained by using ISCG and SCG algorithms in coarse graining network: (a) NW network; (b) ER network; (c) BA network; (d) clustered network.

5.2 耦合Kuramoto相振子网络的同步效果

下面对耦合Kuramoto网络^[24,25]模型进行数值仿真, 进一步比较ISCG和SCG算法对保持相振子同步的效果. 耦合Kuramoto网络模型方程为

$$\begin{aligned} \dot{\theta}_i &= \omega_i + \sigma \sum_{j=1}^N \mathbf{A}_{ij} \sin(\theta_j - \theta_i) \\ (i &= 1, 2, \dots, N), \end{aligned} \quad (7)$$

其中, ω_i 表示第 i 个振子的固有频率, σ 表示全局振子耦合强度, $\mathbf{A}_{ij}$ 为邻接矩阵, θ_i 表示第 i 个振子的相位, N 表示振子的个数. 全局振子的相同步程度可用序参量 $r(t)$ 刻画:

$$r(t) e^{i\phi(t)} = \frac{1}{N} \sum_{j=1}^N e^{i\theta_j(t)}, \quad (8)$$

其中序参量 $r(t) (\in [0, 1])$ 为0时表示全局振子以各自的固有频率运动, 即完全不同步状态, 为1时表示全局振子以相同频率运动, 即达到完全同步状态; $\phi(t)$ 表示全局振子的平均相位. 因此可通过序参量 $r(t)$ 来研究粗粒化网络的同步保持情况.

以上文 $N = 1000$ 的NW小世界网络、ER随机网络、BA无标度网络、聚类网络为例, 采用ISCG和SCG算法粗粒化网络, 得到网络规模 $\hat{N}$ 均为200的ISCG网络和SCG网络, ω_i 在 $(-0.5, 0.5)$ 中随机选择, 初始相位 θ_i 在 $(-\pi, \pi)$ 中随机选择, σ 分别为0.2, 0.2, 1.5, 0.5. 其序参量 $r(t)$ 保持情况如图5所示.

图5展示了分别采用ISCG和SCG两种算法对振子相同步的保持情况, 可见ISCG获得网络的序参量 $r(t)$ 收敛性与原始网络基本一致, 且ISCG算法获得的粗粒化网络同步能力保持情况优于SCG方法.

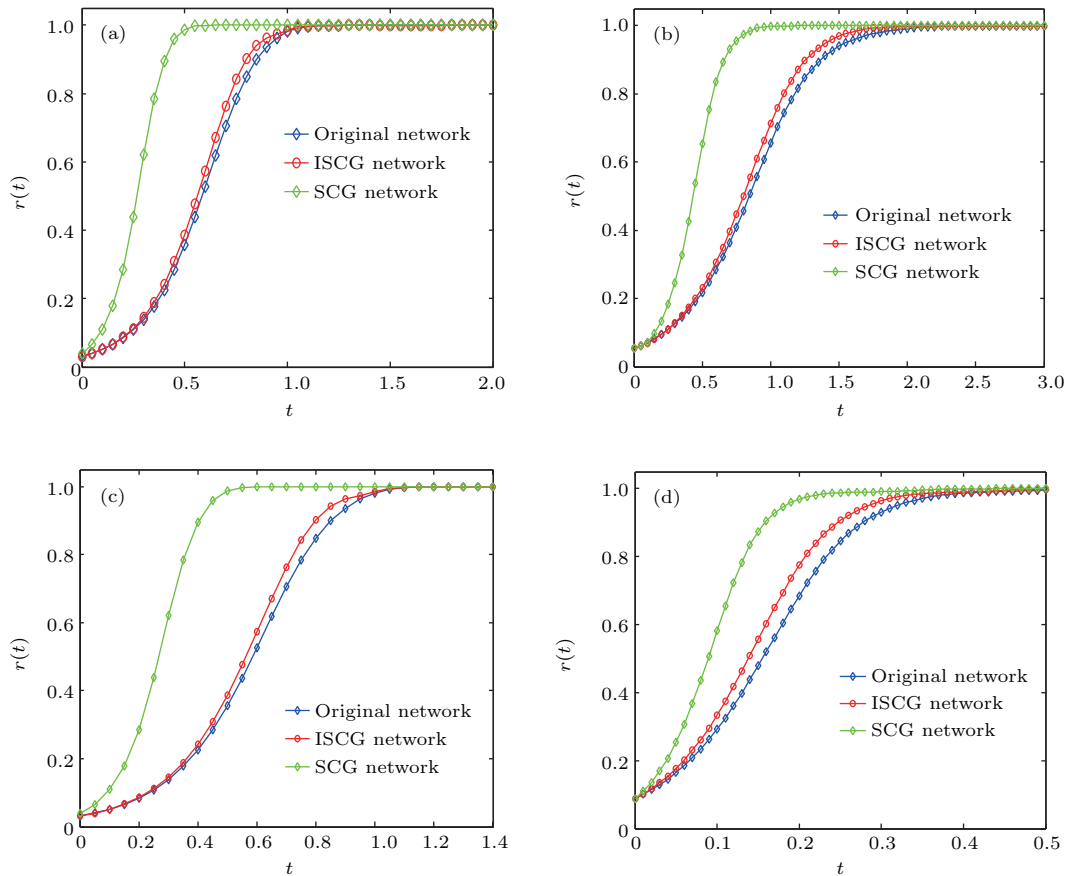


图5 (网刊彩色) 分别采用ISCG和SCG算法保持序参量 $r(t)$ 的情况 (a) NW小世界网络; (b) ER随机网络; (c) BA无标度网络; (d) 聚类网络

Fig. 5. (color online) The maintaining of the order parameter $r(t)$ obtained by using ISCG and SCG algorithms in coarse graining network: (a) NW network; (b) ER network; (c) BA network; (d) clustered network.

6 结 论

复杂动态网络的耦合拓扑结构直接影响网络的动力学性质. 一般来说, 两个节点的外部连接相同或相近, 意味着它们接收外部的信息和作用也相同或相近, 因而具有相近的动力学性质, SCG方法的基本思想正是基于这一点, 把网络中外部链接相同或相近似的节点进行合并, 从而把大规模网络缩减为中尺度网络, 这是研究大规模网络的重要途径之一. 本文提出的ISCG算法, 以特征向量分量之间的间距大小来划分相似节点, 从而自适应地对聚类节点进行分组, 能够确保网络中越相近的节点越被优先合并, 避免了SCG方法对节点进行硬性分组而导致某些情况下相似节点被拆分而与不相似节点合并的缺陷. 仿真实验也验证了ISCG算法既能很好地改善粗粒化效果, 又大幅降低了运算的时间复杂度.

从算法分析和仿真结果看, 在粗粒化过程中应该存在一个最优的聚类个数, 即存在一个既能很好地保持原始网络同步能力又能最大程度地缩减原始网络规模的聚类个数, 例如, 用 λ_2 来刻画网络的同步能力, 可以满足 $|\lambda_2 - \tilde{\lambda}_2| < \varepsilon$ (ε 比较小, 可设定)的 $\tilde{\lambda}_2$ 对应的最小 $\tilde{N}$ 即为最优聚类个数. 以上文中所列举的四种典型即NW小世界网络、ER随机网络、BA无标度网络和聚类网络为例, 可求得最优聚类个数 $\tilde{N}$ 分别为915, 805, 795, 768. 然而, 对于最优的聚类个数问题, 还有更多的未知需要我们去探索, 例如衡量最优粗粒化聚类个数的指标函数如何确定? 各种不同结构的网络的最优指标有何差异? 且有待我们进行更深入的探索研究, 也是我们今后的努力方向.

本文提出的ISCG算法具有误差更小、效果更优、运算量大幅降低等诸多优点, 能极大地改善原粗粒化效果, 提高了网络同步能力的保持状况, 同时为实现大规模网络的同步研究提供了一种更简单、高效和可操作的有效方法.

参考文献

- [1] Pecora L M, Carroll T L 1998 *Phys. Rev. Lett.* **80** 2109
- [2] Jost J, Joy M P 2001 *Phys. Rev. E* **65** 016201
- [3] Wang X F, Chen G R 2002 *IEEE Trans. Circuits-I* **49** 54
- [4] Barahona M, Pecora L M 2002 *Phys. Rev. Lett.* **89** 054101
- [5] Wang X F, Chen G R 2002 *Int. J. Bifurcat. Chaos* **12** 187
- [6] Motter A E, Zhou C S, Kurths J 2005 *Phys. Rev. E* **71** 016116
- [7] Nishikawa T, Motter A E 2006 *Physica D* **224** 77
- [8] Zhou J, Lu J A, Lu J H 2006 *IEEE Trans. Auto. Control* **51** 652
- [9] Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C S 2008 *Phys. Rep.* **469** 93
- [10] Zhu T X, Wu Y, Xiao J H 2012 *Acta Phys. Sin.* **61** 040502 (in Chinese) [朱廷祥, 吴晔, 肖井华 2012 物理学报 **61** 040502]
- [11] Xu M M, Lu J A, Zhou J 2016 *Acta Phys. Sin.* **65** 028902 (in Chinese) [徐明明, 陆君安, 周进 2016 物理学报 **65** 028902]
- [12] Kurkcuglu O, Jernigan R L, Doruker P 2004 *Polymer* **45** 649
- [13] Marrink S J, Vries A H D, Mark A E 2004 *J. Phys. Chem. B* **108** 750
- [14] Bornholdt S 2005 *Science* **310** 449
- [15] Chen J, Lu J A, Lu X F, Wu X Q, Chen G R 2013 *Commun. Nonlinear Sci.* **18** 3036
- [16] Zeng A, Lü L Y 2011 *Phys. Rev. E* **83** 056123
- [17] Saunders M G, Voth G A 2013 *Annu. Rev. Biophys.* **42** 73
- [18] Kim B J 2004 *Phys. Rev. Lett.* **93** 168701
- [19] Chen H S, Hou Z H, Xin H W, Yan Y J 2010 *Phys. Rev. E* **82** 011107
- [20] Gfeller D, Rios P D L 2007 *Phys. Rev. Lett.* **99** 038701
- [21] Gfeller D, Rios P D L 2008 *Phys. Rev. Lett.* **100** 174104
- [22] Chen G R, Wang X F, Li X, Lü J H 2009 *Some Recent Advances in Complex Networks Synchronization* (Heidelberg: Springer) pp3–16
- [23] Lu J A, Liu H, Chen J 2016 *Synchronization in Complex Dynamical Networks* (Beijing: Higher Education Press) pp120–125 (in Chinese) [陆君安, 刘慧, 陈娟 2016 复杂动态网络的同步 (北京: 高等教育出版社) 第120—125页]
- [24] Kuramoto Y 1975 *Lect. Notes Phys.* **39** 420
- [25] Acebrón J A, Bonilla L L, Pérez Vicente C J, Ritort F, Spigler R 2005 *Rev. Mod. Phys.* **77** 137

Improved algorithm of spectral coarse graining method of complex network*

Zhou Jian¹⁾ Jia Zhen^{1)†} Li Ke-Zan²⁾

1) (College of Science, Guilin University of Technology, Guilin 541004, China)

2) (School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin 541004, China)

(Received 2 October 2016; revised manuscript received 21 November 2016)

Abstract

Complex network as a key approach to understanding many complex systems, such as biological, chemical, physical, technological and social systems, is ubiquitous in nature and society. Synchronization of large-scale complex networks is one of the most important issues in network science. In the last two decades, much attention has been paid to the synchronization of complex dynamic networks, especially the meso-scale networks. However, many real networks consist of even hundreds of millions of nodes. Analyzing the synchronization of such large-scale coupled complex dynamic networks often generate a large number of coupled differential equations, which may make many synchronization algorithms inapplicable for meso-scale networks due to the complexities of simulation experiments. Coarse graining method can map the large-scale networks into meso-scale networks while preserving some of topological properties or dynamic characteristics of the original network. Especially, the spectral coarse-graining scheme, as a typical coarse graining method, is proposed to reduce the network size while preserving the synchronization capacity of the initial network. Nevertheless, plenty of studies demonstrate that the components of eigenvectors for the eigenvalue of the coupling matrix, which can depict the ability to synchronizing networks, distribute unevenly. Most of the components distribute concentrically and the intervals are small, while some other components distribute dispersedly and the intervals are large, which renders the applications of original spectral coarse graining method unsatisfactory. Inspired by the adaptive clustering, we propose an improved spectral coarse graining algorithm, which clusters the same or the similar nodes in the network according to the distance between the components of eigenvectors for the eigenvalue of network coupling matrices, so that the nodes with the same or the similar dynamic properties can be effectively clustered together. Compared with the original spectral coarse graining algorithm, this method can improve the accuracy of the result of clustering. Meanwhile, our method can greatly reduce algorithm complexity, and obtain better spectral coarse graining result. Finally, numerical simulation experiments are implemented in four typical complex networks: NW network, ER network, BA scale-free network and clustering network. The comparison of results demonstrate that our method outperforms the original spectral coarse graining approach under various criteria, and improves the effect of coarse graining and the ability to synchronize networks.

Keywords: complex network, synchronization, spectral coarse graining, improved algorithm

PACS: 05.45.-a, 05.45.Xt

DOI: 10.7498/aps.66.060502

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61563013, 61663006).

† Corresponding author. E-mail: jjjzzz0@163.com