

Nonlinear time series analysis by means of complex networks

邹勇, [DONNER Reik V.](#), [Norbert MARWAN](#), [DONGES Jonathan F.](#) and [Jürgen KURTHS](#)

Citation: [中国科学: 物理学 力学 天文学](#) **50**, 010509 (2020); doi: 10.1360/SSPMA-2019-0136

View online: <http://engine.scichina.com/doi/10.1360/SSPMA-2019-0136>

View Table of Contents: <http://engine.scichina.com/publisher/scp/journal/SSPMA/50/1>

Published by the [《中国科学》杂志社](#)

Articles you may be interested in

[Space-time means and solutions to a class of nonlinear parabolic equations](#)

Science in China Series A-Mathematics **41**, 682 (1998);

[Multifractal mass exponent spectrum of complex physiological time series](#)

Chinese Science Bulletin **55**, 1996 (2010);

[Identifying topologies and system parameters of uncertain time-varying delayed complex networks](#)

SCIENCE CHINA Technological Sciences **62**, 94 (2019);

[Hilger-type impulsive differential inequality and its application to impulsive synchronization of delayed complex networks on time scales](#)

SCIENCE CHINA Information Sciences **61**, 078201 (2018);

[SOME NOTES ON APPLICATION OF MAXIMUM ENTROPY SPECTRAL ANALYSIS TO SEASONAL ADJUSTMENT OF TIME SERIES](#)

Chinese Science Bulletin **28**, 289 (1983);



非线性时间序列的复杂网络分析

邹勇^{1*}, DONNER Reik V.^{2,3}, MARWAN Norbert³, DONGES Jonathan F.^{3,4},
KURTHS Jürgen^{3,5}

1. 华东师范大学物理与电子科学学院, 上海 200241;

2. Department of Water, Environment, Construction and Safety, Magdeburg-Stendal University of Applied Sciences, Breitscheidstraße 2, Magdeburg 39114, Germany;

3. Potsdam Institute for Climate Impact Research (PIK) — Member of the Leibniz Association, Telegrafenberg A31, Potsdam 14473, Germany;

4. Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, Stockholm 114 19, Sweden;

5. Department of Physics, Humboldt University Berlin, Newtonstraße 15, Berlin 12489, Germany

*联系人, E-mail: yzou@phy.ecnu.edu.cn

收稿日期: 2019-04-17; 接受日期: 2019-05-24; 网络出版日期: 2019-09-25

国家自然科学基金(编号: 11872182, 11835003)和上海市自然科学基金(编号: 17ZR1444800)资助项目

摘要 近十多年, 运用复杂网络的方法进行时间序列分析, 是非线性动力学发展的最新方向, 为传统的以混沌理论为基础的非线性时间序列分析注入了新生力量, 在刻画动力系统非线性特征方面, 提供了很多崭新的见解, 在很多不同领域内都取得了广泛应用. 本文回顾三种主要的方法: 递归网络、可视图以及转换网络. 着重介绍三种网络方法的理论基础和各自的最新进展. 本文将指明今后时间序列网络方法的发展方向, 并为实际数据分析提供指导作用.

关键词 复杂网络, 非线性动力学, 回归性, 可视图, 传递网络

PACS: 05.45.Ac, 05.45.Tp, 89.75.Fb

1 引言

人工智能正在产生更多复杂形式的数据, 进入了大数据的新时代^[1], 数据越来越具有高容量、高速度和高品种等复杂性, 这给来自各个领域的研究人员带来了巨大挑战. 开发先进的跨学科数据分析技术有助于在混乱的数据点中捕获隐藏的结构, 这包括来自机器学习、数据挖掘、统计、自然语言和文本处理等方法^[2], 快速学习数据特征, 并做出更

好更快的决策. 在动态系统和统计物理学等背景下, 这些新方法通常与复杂网络^[3]和复杂性理论等概念紧密关联.

本文关注一些在过去十多年中引起极大研究兴趣的特定领域——应用复杂网络理论进行非线性时间序列分析^[4,5]. 时间序列是由观察时间得到的数据点序列, 时间点之间通常都是等间隔且离散性的. 时间序列在日常生活中到处可见, 比如, (1) 天气条件, 如从气象站或者气象卫星获得的地面气

引用格式: 邹勇, Donner R V, Marwan N, 等. 非线性时间序列的复杂网络分析. 中国科学: 物理学 力学 天文学, 2020, 50: 010509

Zou Y, Donner R V, Marwan N, et al. Nonlinear time series analysis by means of complex networks (in Chinese). Sci Sin-Phys Mech Astron, 2020, 50: 010509, doi: [10.1360/SSPMA-2019-0136](https://doi.org/10.1360/SSPMA-2019-0136)

温、海平面气压和风速等; (2) 经济活动数据, 如琼斯工业平均指数每日收盘价等股票市场或汇率变化等金融数据; (3) 人的生物医学状况, 如通过脑电图(Electroencephalogram, EEG)或高分辨率脑成像技术(如磁共振成像MRI和断层扫描CT)收集的生理和临床数据.

本文回顾的时间序列分析方法是基于动力系统理论, 结合复杂网络理论而进行的非线性时间序列分析^[4-7]. 从复杂网络研究的角度来看, 所讨论的课题可以看作是网络理论在多个应用领域内不同实验数据的成功应用. 本文以综述文献^[8]为基础, 将原文中的内容进一步凝练而成, 供国内读者参考.

2 非线性时间序列分析

时间序列分析实质上就是数据压缩^[7]. 通过系统测量的样本时间序列, 计算出的几个特征指标来解释其潜在的基本动力学性质, 因此这些特征量必须突出系统的某些特定性质. 大多数时间序列分析的方法都依赖于线性假设, 例如自回归和滑动平均等模型, 且这些模型中的自相关函数几乎都是呈指数衰减. 现在普遍认为, 支配自然或人类活动的动力学规律大部分是非线性的, 非线性现象无处不在. 但是非线性时间序列分析没有像线性时间序列分析那样成熟, 也远没有线性时间序列分析那样容易^[4,7]. 非线性时间序列分析的思想和方法, 源于动力系统理论或所谓的“混沌理论”, 该理论通过一组非线性差分方程或非线性常微分方程研究系统动力学. 混沌理论描述非线性导致的系统特性的复杂演化规律, 例如, 对初始条件的敏感依赖和有限时间尺度下的可预测性. 这些非线性演化规律通常情况下不能用经典线性分析技术来获得.

自20世纪80年代初期以来^[9], 非线性时间序列分析取得了巨大进展, 其数学美在于能够用许多不同的方法来描述相空间的不变测度性质. 比如基于相空间重构的吸引子, 可从两个方面揭示存在的混沌行为: (1) 动态方面, 如通过计算李雅普诺夫指数或熵(如Kolmogorov-Sinai 熵); (2) 几何方面, 如通过分形维数或关联维数等指标刻画(或广义分形维数谱 D_q ^[10]).

这些都是基于单变量时间序列, 进行系统的某

些相空间不变特性的描述^[7]. 相比之下, 双变量法是通过分析相互耦合的成对时间序列来分析两个动态系统之间的动力学性质. 在许多物理和生物系统中都观察到了耦合混沌系统的同步行为, 双变量时间序列分析方法可以检测和区分从非同步状态到同步的过渡形式(例如, 相位同步、滞后同步、完全同步和广义同步等). 双变量分析就是要获取耦合强度和方向等相互作用信息. 更具有挑战的任务是解开复杂网络子系统之间的复杂拓扑连接结构, 这包括各种统计关联算法, 如皮尔逊相关性、相互信息(包括时滞信息)^[11]、格兰杰因果关系^[12]、传递熵^[13]、耦合方向性指标等^[14-16]. 一般来说, 系统间的耦合可以有多种形式, 感兴趣的读者可以参考综述文献^[17, 18].

非线性时间序列分析为许多工程数据分析提供了强大的工具, 但必须注意实际应用时方法的局限性, 实验测量数据中常见的一些问题对非线性计算分析提出了巨大挑战. 例如, 现有的大多数非线性方法仅适用于低维动力系统, 但实际上, 很少有真实的数据集是在低维系统上运行并且被传感器完全测量的. 数据平稳性、相空间嵌入参数的合理选择、数据长度的依赖性(在许多实际情况下, 时间序列可能相当短)、噪声的影响或不规则采样等, 都是实际处理数据时必须解决的重要问题^[8]. 算法本身同样需要正确选择参数, 这在很大程度上取决于研究者的经验, 例如, 分形维数和李雅普诺夫指数等动态不变量的计算时应采用对数坐标下线性拟合算法, 线性区间的合理选择往往会对结果产生显著的影响. 此外, 算法本身的复杂性也必须得到很好的评估.

本文将展现利用复杂网络方法进行非线性时间序列分析的基本思想. 更重要的是, 复杂网络方法可以部分解决时间序列分析中其他传统方法未能解决的一些问题, 例如, 网络分析方法计算分形维数更加便捷, 或和李雅普诺夫指数等相媲美^[19,20].

3 复杂网络基本概念

随着计算能力和各研究领域数据量的飞速增长, 复杂网络已经成为描述系统单元和系统之间相互作用和依赖性的强有力通用工具^[3,21], 同时它又

是一种空间和时间数据分析的强有力工具 [8].

本节简要介绍复杂网络的基本概念, 更一般的网络统计分析及其最新进展可参考文献[3, 21]. 一个复杂网络包含一个节点集合 V 和一个表示节点间连边情况的连边集合 E , 用图 $G = (V, E)$ 来表示 [22]. 每一个节点用一个整型编号表示 $p = 1, \dots, N$, 每一条连边用一对节点 (p, q) 表示. 一般情形, 连边可以存在不同强度, 用权重 W_{pq} 表示. 时间序列网络可以是无向无权重的简单网络, 也可以是有向有权重的网络. 一个无权重网络可以在权重矩阵 $\mathbf{W}$ 中设定合适的阈值 T 获得, 通常用二进制邻接矩阵 $\mathbf{A}$ 表示. 也就是若 $W_{pq} > T$, 则 $A_{pq} = 1$, 反之则 $A_{pq} = 0$, 邻接矩阵 $\mathbf{A}$ 中的每一个非零元素即组成连边集合 E . 且无向网络的矩阵 $\mathbf{A}$ 存在对称性, 即 $A_{pq} = A_{qp}$. 网络路径是指从节点 p 到 q 需要经过的路径, 由路径长度 l_{pq} 表示. 根据不同的时间序列网络转换算法, 邻接矩阵 $\mathbf{A}$ 通常依赖于算法的参数选择.

基于邻接矩阵 $\mathbf{A}$, 描述网络节点和连边的统计性质, 也就构成了时间序列的网络结构分析. 常用统计量包括度 k 及其分布 $p(k)$ 、局部类聚系数 C_p 、局部近邻系数 c_p 、节点介数 b_p 、连边匹配系数 m_{pq} 、连边介数 b_{pq} 、网络连接密度 ρ 、全局类聚系数 C 、传递系数 $\mathcal{T}$ 、平均路径长度 $\mathcal{L}$ 、全局效率 $\mathcal{E}$. 网络子图或模体的分布同样为网络结构的刻画提供了很好的手段, 这里不展开网络统计量的具体表达式, 可参考文献[8]. 有关复杂网络的诸如小世界、无标度和社团结构等有趣的性质, 也可参考文献[23].

许多复杂系统包括多个不同层次的子系统. 由于子系统内部成员构成的局域连接和子系统之间构成的外部全局连接的共同作用, 系统动力学演化呈现更多不同层次的交互作用而错综复杂. 构建多层或多路网络可以更好地模拟复杂系统相互作用的多层次性质 [24]. 对于相互作用的一对双变量(或多变量)时间序列, 可以通过构造多层多路网络或者相互耦合的作用网络来分析, 由此获得系统间相互作用强度和作用方向等信息.

在单个无向无权重的网络 $G = (V, E)$ 基础上, Donges等人 [25]提出了适用于时间序列相互作用网络的通用表达式. 设 G 由 G_α 和 G_β 两层组成, 则 $E_{\alpha\alpha}$ 表示 G_α 内部的连接, 而集合 $E_{\alpha\beta}$ 表示子网络 G_α 和 G_β 之

间的连接. 在这样的描述框架下, 单个时间序列网络的结构统计量可以直接套用在子网络 G_α 和 G_β 之间相互作用的刻画上, 包含节点和连边两大类特征量. 此时更加关心刻画子网络 G_α 每一个节点 p 在影响子网络 G_β 的交叉作用, 主要包括交叉度 $k_p^{\alpha\beta}$ 、局部交叉类聚系数 $C_p^{\alpha\beta}$ 、交叉近邻系数 $c_p^{\alpha\beta}$ 、交叉介数 $b_p^{\alpha\beta}$ 、子网络交叉连接密度 $\rho^{\alpha\beta}$ 、全局子网络交叉类聚系数 $C^{\alpha\beta}$ 、交叉平均路径长度 $\mathcal{L}^{\alpha\beta}$ 和交叉效率 $\mathcal{E}^{\alpha\beta}$ 等. 这些统计量的具体表达式可参考文献[25].

4 时间序列的网络转换方法

将时间序列转换为复杂网络, 首先需要提出网络节点和连边的定义. 根据不同的定义, 同一条时间序列构建的网络节点数和连边数都有可能不同, 这些网络大致上可分成三大类(见表 1).

第一类近似网络把时间序列观测点或者片段定义为节点, 连边定义为节点之间的相似性或邻近关系 [26–28]. 这一类方法又细分为周期网络 [29]、关联网络 [30]和递归网络等 [26, 27, 31].

第二类方法根据时间序列采样点之间的可视性条件而构成的可视图方法 [32–34].

第三类方法是转换网络, 它主要利用了符号动力学和随机过程的思想, 首先将时间序列转换为符号表示, 即定义为网络的节点, 动力学的信息则转换为符号之间演化, 不同网络节点间的连接强度由不同符号之间的转移概率来描述, 由此所获得的转换网络很大程度上是把原有动力学转换为马尔可夫链的描述形式 [35].

需要指出的是上述近似网络和转换网络在某种程度上都或多或少和动力系统回归性质有关联 [28],

表 1 网络节点和连边的定义

Table 1 Definitions of vertices and the edges

	方法	节点	连边	方向性
近似性网络	周期网	周期	关联系数或距离	无向
	关联网	状态矢量	关联系数	无向
	k -近邻网	状态矢量	固定邻居数	有向
	递归网	状态矢量	固定连边数	无向
	自适应近邻网 ε -递归网	状态矢量	固定领域大小	无向
可视图	可视图基本型	观测值	可视条件	无向
	水平可视图	观测值	水平可视条件	无向
转换网络	阈值粗粒化	相空间划分	时间延续	有向
	排序图	排序模体	时间延续	有向

这两类时间序列网络方法都运用了非线性时间序列分析时常用的相空间划分和序列的符号化等技巧. 也正因为如此, 这些网络方法可以用来估计系统的动态不变量, 如熵和互信息 [19].

网络方法包含两点共性. 首先, 网络具有节点重标不变性, 所以近似网络的结构拓扑特征同样具有节点重排不变性, 这和时间序列分析中关注观测值之间的时间先后信息具有明显不同. 其次, 近似网络是嵌入动力学相空间的一类空间网络 [36], 也被称为是随机几何图 [19]. 可视图及其相关概念可以看作是以时间轴为距离的空间网络.

为了更好地理解时间序列的网络分析方法, 首先数值求解混沌洛伦兹系统, 获得 x 时间序列 ($\dot{x} = 10(y - x); \dot{y} = x(28 - y); \dot{z} = xy - 8/3z$, 采样间隔 $\Delta t = 0.02$), 图 1 展现了几种常见的时间序列网络图(不同方法得到的网络节点数不同, 具体的参数选择可参考文献 [8]). 下面展开讨论构建网络的算法以及相应网络特征的描述, 并且介绍已经被广泛应用的常见方法, 即递归网络、可视图和转换网络.

4.1 相空间递归网络

设动力系统的采样点为 $x_i = x(t_i)$, 组成序列 $\{x_i\}_{i=1}^N$. 在连续系统中, 时间序列通常是将其相空间轨迹根据给定的采样频率进行时间离散化得到的有限长度序列. 相空间吸引子可以通过时间滞后法进行重构 [4, 37], 即选择合适的重构维数 m 和时间滞后 τ , 得到相空间点矢量 $\mathbf{x}_i = (x_i, x_{i-\tau}, \dots, x_{i-(m-1)\tau})$. 已经证明当 $m > 2D_F$ 时, 重构吸引子和原系统拓扑等价 [37]. 通常情况下, m 和 τ 都是未知的, 需要合理设置. 在处理实际数据时, 首先必须确保重构是否可行, 否则极易引入很大的计算偏差 [4].

非线性混沌系统的一个重要性质是递归性, 即随着时间趋于无穷, 相空间轨迹将无限多次回到初始点所在的无限小领域内 [6]. Eckmann 等人 [38, 39] 引入了递归图方法, 从重构序列生成一个递归矩阵 $\mathbf{R}(\varepsilon)$ 来刻画系统的递归特性, 其定义为

$$R_{ij}(\varepsilon) = \Theta(\varepsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (1)$$

其中, $\|\cdot\|$ 表示相空间的两点距离(如曼哈顿距离或

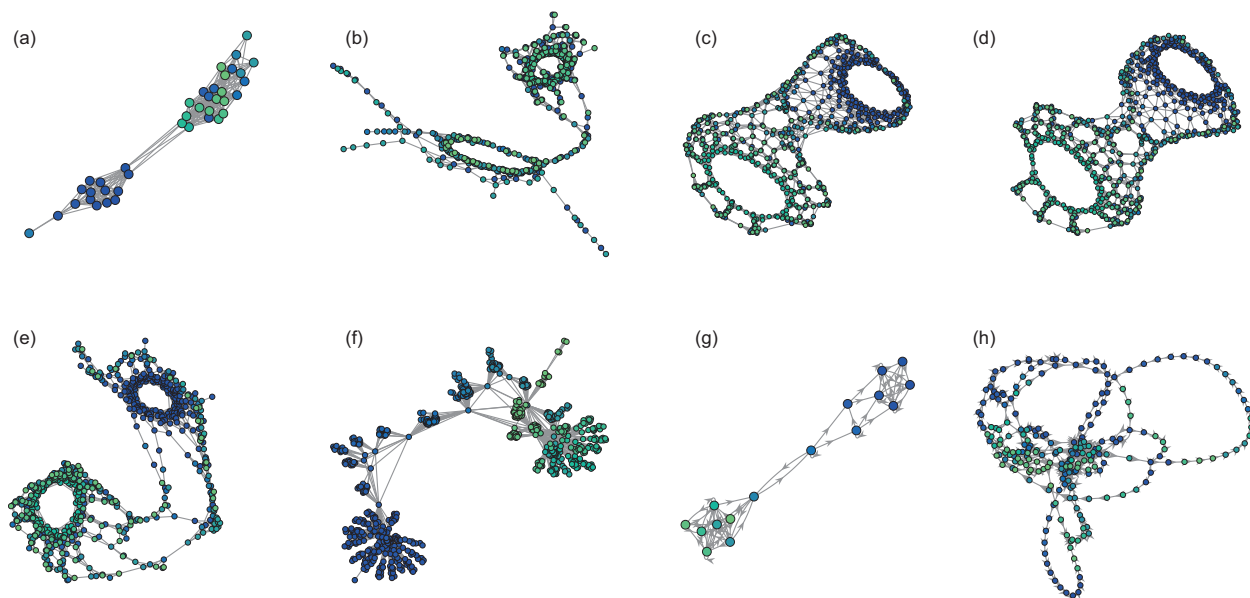


图 1 (网络版彩图)洛伦兹系统 x 序列构建的网络. (a) 周期网络; (b) 关联网络; (c) k -近邻网络; (d) 自适应近邻网络; (e) ε -递归网络; (f) 可视图; (g) 粗粒化转换网络; (h) 排序转换网络. 网络孤立节点已经排除在外. 引自文献 [8]

Figure 1 (Color online) Networks constructed from the x -coordinate of the Lorenz system. (a) Cycle network; (b) correlation network; (c) k -nearest neighbor network; (d) adaptive nearest neighbor network; (e) ε -recurrence network; (f) visibility graph; (g) coarse-graining based transition network; (h) ordinal partition transition network. Disconnected vertices have been removed from the representations. Modified from ref. [8].

欧氏距离等), 当空间两点间 $\mathbf{x}_i$ 和 $\mathbf{x}_j$ 的距离小于给定阈值 ε , 相应的 R_{ij} 设为1, 否则设为0. 显见, ε 的选取决定了空间点之间靠近的程度, 并且会影响网络统计量. 单个系统递归图的思想已经被推广为能够处理双变量或者多变量的交叉递归图^[40]、联合递归图^[41]或者多路递归图^[42]. 一些更好捕捉递归区域的最新算法同样值得关注^[43–45].

递归网络将每一个相空间点 $\mathbf{x}_i$ 当作网络节点, 且网络连边表示节点在相空间内靠近的程度. 排除递归图中的自身连接, 即得到时间序列的递归网络邻接矩阵 $\mathbf{A}$. 参数 ε 直接影响邻接矩阵, 因此被称为 ε -递归网络^[28]. 二进制邻接矩阵 $\mathbf{A}(\varepsilon)$ 具有对称性. 图2中展示了混沌洛伦兹系统 x 时间序列的递归网络构建的基本思想. 关于递归网络构建的更详细内容可以参考文献^[27, 28].

递归网络中, 路径是指从节点 i 开始到 j 结束跨过的连边数目 l_{ij} , 如图2(b)所示. 具体而言, 一条路径是用尺度为 ε 的 m 维小球 $B_\varepsilon(x_i)$ 去覆盖 i 和 j 之间的吸引子距离, $B_\varepsilon(x_i), B_\varepsilon(x_{k_1}), \dots, B_\varepsilon(x_{k_{l_{ij}-1}}), B_\varepsilon(x_j)$, 且每个小球都符合 $B_\varepsilon(x) = \{y \mid \|x - y\| < \varepsilon\}$ 且它们之间具有重叠区域, 即 $B_\varepsilon(x_i) \cap B_\varepsilon(x_{k_1}) \neq \emptyset, \dots, B_\varepsilon(x_{k_{l_{ij}-1}}) \cap B_\varepsilon(x_j) \neq \emptyset$. 因此, 最短距离也就是相空间点 i 和 j 之间需要的 ε 小球的最少数目, 表示为 l_{ij} . 需要注意网络的路径和时间序列在相空间的轨迹是完全不一样的概念^[27].

网络的其他结构统计量同样可以从邻接矩阵

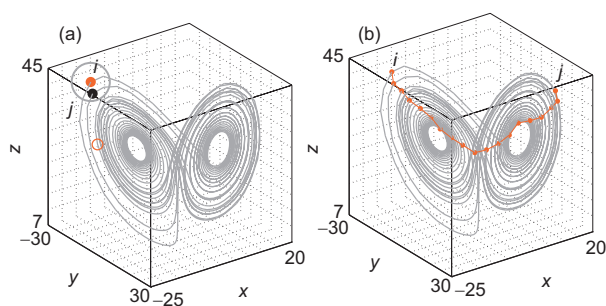


图2 (网络版彩图)递归网络构建基本思想. (a) 两个实心时间点之间距离小于阈值 ε ; (b) 一条网络路径 l_{ij} . 引自文献^[28]

Figure 2 (Color online) Basic concepts of recurrence networks. (a) The distance between two time (filled) points is less than the threshold ε ; (b) a network path l_{ij} . Reproduced from ref. ^[28] with permission by World Scientific Publishing Co.

$\mathbf{A}(\varepsilon)$ 得到, 但是需要注意邻接矩阵的网络展示可以运用很多不同的方法. 图1(e)采用了图像软件GUESS中的受力导向放置算法, 当然还可以运用弹簧电气模型算法等获得和吸引子相空间更加接近的几何结构^[46].

除了采取固定阈值的 ε -递归网络, 还有稍微不同的变型, 比如以 k -近邻递归网络^[38]和 k -近邻自适应递归网络^[31, 47]. 递归网络的其他形式可以参考文献^[8], 值得指出的是这些变化形式和基本型相比较, 到底获得了多少额外的网络结构信息, 这还值得进一步研究.

4.1.1 递归网络的刻画

递归网络忽略了时间序列的时间演化信息, 这和传统的递归图分析显著不同^[39]. 递归网络结构分析更多关注吸引子在相空间内的几何特性^[27].

基于网络邻接矩阵 $\mathbf{A}$, 前面提到的常用结构统计量都可以显式写下来^[8, 28]. 注意这些量都是从有限长度 N 的序列中计算得到的, 并且依赖于阈值 ε 的选取. 另外, 如果 $N \rightarrow \infty$ 和任意给定阈值 ε , 网络统计量都可以在随机几何图的理论下找到各自的理论解^[48], 前面讨论的统计量都是这些解析理论的离散化估计, 这就使得混沌吸引子在相空间的递归网络分析更具理论基础.

一个有趣的问题是递归网络是否同样存在复杂网络的无标度和小世界性质. 运用随机几何图理论, 递归网络的度分布 $p(k)$ 可以从吸引子的密度函数 $p(x)$ 获得, 并且在一维Logistic映射系统中, 严格证明了 $p(k)$ 是符合幂律分布的^[49]. 但是对于更高维或者一般的连续系统, 无标度性质产生的条件尚且未知. 另外, 无标度特性和相空间分形结构是两个不同的概念, 通常情况下两者之间没有直接联系. 网络小世界性质是指网络同时具有高的类聚系数和小的平均路径长度, 而递归网络的平均路径长度是由阈值 ε 确定的, 即 $\mathcal{L} \sim \varepsilon^{-1}$ ^[27], 和网络大小 N 关系不大, 因此递归网络不可能具有小世界性质^[50]. 如果时间序列的噪声强度和吸引子尺度可比的话, 递归网络的小世界性质有可能会存在^[50, 51].

网络传递系数 $\mathcal{T}$ 可以作为吸引子分维数的另外

一种补充定义^[19], 并且解析得到了 $\mathcal{T}$ 和维数 m 之间符合指数级下降的关系, 这个结论可以扩展到具有分形维数的吸引子流形. 具体而言, 分为上和下网络传递维数, 即^[19]

$$D_{\mathcal{T}}^u = \limsup_{\varepsilon} \frac{\log(\mathcal{T}(\varepsilon))}{\log(3/4)}, \quad D_{\mathcal{T}}^l = \liminf_{\varepsilon} \frac{\log(\mathcal{T}(\varepsilon))}{\log(3/4)}, \quad (2)$$

此处定义的上和下传递维数主要是部分混沌映射吸引子形成了康拓集后, 维数值会随着 ε 的变化产生振荡现象. 网络类聚系数同样可以定义为类聚维数, 作为吸引子分维数的补充^[19].

4.1.2 构建递归网络的挑战

阈值 ε 、重构参数、采样频率、甚至多变量时间序列的选取等, 都有可能影响递归网络的分析, 这些讨论可以参考文献[8, 52]:

- 阈值选取^[39, 52]: 过大或过小的 ε 都不合适, 可采用类似关联维数 D_2 算法中的思想, 选取的阈值应该使得网络连接密度在双对数坐标下处在线性区域^[10, 52]. 另外一种经验方式是选取的 ε 使得网络的连接密度为 $\rho \leq 0.05$ ^[26, 27], 这在文献[52]做了详细讨论. 确保网络连通性时所对应的临界阈值同样是一个不错的选择^[48, 53]. 最近有根据递归网络的计算复杂性而提出的阈值选取方法, 同样取得有意义的结果^[54].

- 重构维数 m 和时间滞后 τ : 几乎所有的非线性分析都是以相空间重构为基础的^[4], 这对于平稳过程通常是可行的. 但对一些非平稳过程, 必须注意引入合适的的数据预处理, 尽量消除非平稳带来的影响, 否则容易得到不合理结果^[55, 56].

- 网络统计量的阈值依赖性: 如前所述, 阈值的选取决定了网络连接的密度, 并且影响了网络结构统计量的数值, 为了更全面反映阈值的影响, 通常可以把阈值依赖性直接展示^[19].

- 噪声对于结果的影响: 在递归图中, 根据噪声的标准差 σ , 选取 $\varepsilon \sim 5\sigma$ 可以有效减小噪声的影响^[57]. 噪声对网络统计量的影响, 可参考文献[58]. 诚然在实际数据分析时, 噪声只是采集数据不确定性的一个方面, 还包括数据测试误差. 不规则采样等, 此时通常需要合理地进行统计推断^[59].

4.1.3 多层递归网络

在多变量递归网络的构建方面, 存在几个不同的工作. 文献[42]从一个 M 维多变量时间序列 $\{\mathbf{x}_i\}_{i=1}^N$ ($\mathbf{x}_i = (x_i^{[1]}, x_i^{[2]}, \dots, x_i^{[M]}) \in \mathbb{R}^M$), 提出计算层间互信息 $I^{\alpha\beta}$ 或者层间连接重合度 ω 两种方法, 构建了多路递归网络. 在耦合的周期系统中, 可以得到高数值 $I^{\alpha\beta}$ 和 ω , 而混沌系统则只能得到较小值^[42].

交叉递归图^[40]和联合递归图^[41]是双变量递归图的两种典型构造方法. 进一步可以构建跨系统递归网络^[60]. 首先任取两层 X_α 和 X_β , 交叉递归网络考虑 (α, β) 共享的相空间邻域, 即 $CR_{ij}^{[\alpha\beta]}(\varepsilon_{\alpha\beta}) = \Theta(\varepsilon_{\alpha\beta} - \|\mathbf{x}_i^{[\alpha]} - \mathbf{x}_j^{[\beta]}\|)$ ^[40], 由此, 两两分析得到的交叉递归连接密度矩阵 $\mathbf{CR}^{[\alpha\beta]}(\varepsilon_{\alpha\beta})$ 即组成跨系统递归网络^[60]. 在这样的框架下, 网络结构统计量同样具有简洁的理论表达形式. 跨系统递归网络方法得到了很好地应用, 包括油水两相流多通道测量数据分析^[61-63].

联合递归网络就是多个递归网络的乘积, 即联合邻接矩阵 $\mathbf{JR}$ 的元素是 $JR_{ij}(\varepsilon_1, \dots, \varepsilon_M) = \prod_{\alpha=1}^M R_{ij}^{[\alpha]}(\varepsilon_\alpha)$, 其中 $R_{ij}^{[\alpha]}(\varepsilon_\alpha) = \Theta(\varepsilon_\alpha - \|\mathbf{x}_i^{[\alpha]} - \mathbf{x}_j^{[\alpha]}\|)$ 为 α 层的递归矩阵, $\alpha \in [1, M]$. 联合网络连边 (i, j) 表示所有 M 个时间序列同一时刻递归到相应的领域^[60].

4.2 可视图

最初, 可视图是用来刻画二维景观障碍物之间的相互可见关系, 其应用范围包括机器人运动规划、建筑设计和地形空间描述^[64, 65]. 自从Lacasa等人^[32]把这一思想引入到时间序列分析开始, 可视图方法得到了深入和广泛地应用^[66, 67], 尤其在地球物理过程的时间序列分析方面^[34]. 此处讨论一下可视图的基本思想, 并指出在应用时可能会碰到的问题.

4.2.1 构建方法

可视图^[32]把一维时间序列 $\{x_i\}_{i=1}^N$ (其中 $x_i = x(t_i)$, $i = 1, \dots, N$)的每一个采样点作为网络节点, 两个时间点 t_i 和 t_j 之间的所有其他时间点 t_k ($t_i < t_k < t_j$)满足下列可视化条件, 那么 t_i 和 t_j 在网络中就相

连, 即

$$\frac{x_i - x_k}{t_k - t_i} > \frac{x_i - x_j}{t_j - t_i}. \quad (3)$$

因此网络连边直接考虑了时间信息. 图3(a)中以太阳黑子序列为例说明了可视图的构建方法. 根据定义, 相邻两个时间点总是相连的, 整个可视图是联通的, 不存在孤立点, 但是时间序列首尾部分存在边界效应.

如果时间点 t_i 和 t_j 连接与否由下边的水平可视化条件决定, 即构成水平可视图[33]:

$$x_k < \min\{x_i, x_j\}, \quad (4)$$

其中, $t_i < t_k < t_j$. 由此可知, 水平可视图的条件式(4)比一般的条件式(3)更为严格, 两者算法差异如图3(b)所示. 水平可视图中的连接是一般可视图连接的子集, 因此网络的连接密度会降低, 两个网络中连接的差异性可以提供补充信息.

基于一般可视和水平可视图, 还有很多推广[69-71], 并且还需关注其算法复杂性[72].

4.2.2 可视图基本性质

周期信号由于规则性, 可视图的度分布 $p(k)$ 通常只有有限个不一样的 k 值. 经验结果表明, 在分形随机过程中, $p(k)$ 为幂律分布, 即 $p(k) \sim k^{-\gamma}$ [32], 且 γ

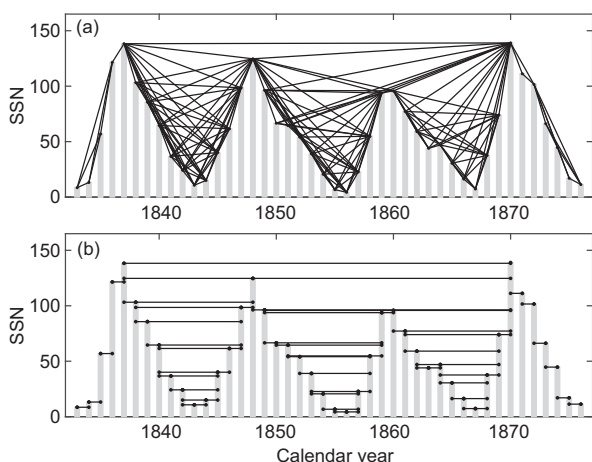


图3 可视图(a)和水平可视图(b)的构建方法, 引自文献[68]

Figure 3 Schematic illustration of the algorithm for constructing (a) natural visibility graphs and (b) horizontal VG. Reproduced from ref. [68].

和分数布朗运动的Hurst指数 H 之间存在线性关系 $\gamma = 3 - 2H$, 而对于分数高斯噪声, 则 $\gamma = 5 - 2H$ [73,74]. 当然, 数值求解 γ 需要验证统计可靠性[75].

对于某些简单的随机过程序列, 水平可视图的 $p(k)$ 可以得到解析解[33]. 比如在无关联随机序列中(设其概率分布为 $p(x)$, $x \in [0, 1]$), 则 $p(k) = \frac{1}{3}(\frac{2}{3})^{k-2}$. 数值结果显示, 当 $p(x)$ 为均匀、高斯或者幂律分布时(如 $p(x) \sim x^{-2}$), $p(k)$ 和理论解非常一致[33]. 对于更一般具有不变测度 $p(x)$ 情形, 如马尔可夫过程和平稳Ornstein-Uhlenbeck等过程, $p(k)$ 同样可以解析获得[76].

网络大值 k 节点影响了 $p(k)$ 的尾部特征, 也通常反映了时间序列局部采样大值点在网络中的作用, 但这会受到周期性变化趋势的影响[77], 因此有时候大值 k 点和大数值 x 不能够一一对应[32]. 为了刻画时间序列小数值 x 在网络中的作用, 可以把时间序列预先做一个负号变换[77]. 刻画 $p(k)$ 的复杂性, 可计算熵 $S = -\sum_k p(k) \log p(k)$, 这可以作为时间序列香农熵的近似[78].

对于非关联噪声, 水平可视图 $p(k) \sim e^{-\lambda k}$, 尤其当白噪声时, $\lambda_c = \ln(3/2)$ [76,79]. 但是运用临界指数 λ_c 来区分随机过程和混沌系统时, 必须小心验证其可靠性[80,81].

局部类聚系数 C 在非关联随机序列中的分布具有简洁的表达式, 即 $p(C) = \frac{1}{3}(\frac{2}{3})^{2/C-2}$, 这得到了很好的数值验证[33]. 关于其他局部统计量(如介数 b_i 和近邻系数 c_i), 目前都仅停留在数值研究, 还不清楚是否存在理论解[34]. 在全局统计量方面, 更关注类聚系数 C 和平均路径长度 L , 比如分数布朗运动中, C 随着Hurst指数增加而减小[82]. 最后, 可视图的模体分布可提供额外的信息[83].

4.2.3 构建可视图的挑战

可视图处理实际时间序列时, 还会碰到一些特殊问题, 比如经常会存在漏采样或者采样间隔不均匀等问题, 此时通常需要进行插值等数据预处理获得等间隔数据. 更为普遍的例子是在地球物理中的点过程序列(如地震强度的序列), 表现为时间轴上不均匀的事件. 可视化条件直接考虑了采样时间的

作用, 因此这对于可视图的构建同样重要^[34].

4.2.4 多层可视图

目前提出了多路可视图、交叉可视图和联合可视图, 这可以用来研究两个时间序列之间的耦合和信息传递. 假设 M 维实值序列 $\{\mathbf{x}_i\} = (\{x_i^{[1]}\}, \{x_i^{[2]}\}, \dots, \{x_i^{[M]}\})$, 且 $\mathbf{x}_i = \mathbf{x}(t_i) \in \mathbb{R}^M, t_i \in [1, N]$. 每一维序列标号为 $X^{[\alpha]} = \{x_i^{[\alpha]}\}_{i=1}^N, \alpha \in [1, M]$.

- 多路可视图^[84]: 整个全体 M 是邻接矩阵的集合, 即 $\mathcal{M} = \mathbf{A} = \{\mathbf{A}^{[1]}, \mathbf{A}^{[2]}, \dots, \mathbf{A}^{[M]}\}$, 其中 $\mathbf{A}^{[\alpha]} = \{A_{ij}^{[\alpha]}\}_{ij}$ 是 α 层的邻接矩阵. 多路可视图采用了和一般多层网络一致的描述方法^[24]. 具体说来, 每两个不同层 α 和 β 之间的相似度可以通过计算层间互信息 $I^{\alpha\beta}$ 或层间连接重合密度 ω ^[84] 来刻画. 多路可视图方法在很多实际数据中都得到了很好地应用^[85].

- 联合度和余度法^[68]: 考虑一对时间序列 (α, β) 及其邻接矩阵 $\mathbf{A}^{[\alpha]}$ 和 $\mathbf{A}^{[\beta]}$. 每一个时间点 i 在 α 和 β 两个层内的度 $k_i^{[\alpha, \beta]}$ 及邻居集 $\mathcal{N}_i^{[\alpha, \beta]} = \{A_{ij}^{[\alpha, \beta]} \equiv 1, j \in \{1, \dots, N\} \setminus \{i\}\}$. 时间点 i 的联合度 k_i^{joint} 即考虑这两个邻居集的重合度. 如果仅考虑在 α 层的连接而不考虑 β 层的情况, 就可以刻画 α 层多余的度信息(交换 α 和 β 则为 β 层的多余度). 余度法即刻刻画了 α 和 β 层间的多余度信息的非对称性, 这一方法很好地刻画了太阳黑子活动序列在南北半球分布的非对称性^[68].

4.2.5 可视图分解

假设可视图邻接矩阵 $\mathbf{A}$, 把时间信息作为连接的方向性, 很容易得到有向可视图. 在节点 i 的邻居中, 将 $t_j > t_i$ 的部分邻居定义为 i 节点的出度 k_i^{out} , 而 $t_j < t_i$ 的部分邻居定义为 i 节点的入度 k_i^{in} , 由此, i 节点的度分解成 $k_i = k_i^{\text{in}} + k_i^{\text{out}}$, 相应地, 网络度分布同样分成两类 $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$. 另外, 局域类聚系数 C_i 同样可以分解为时间向前和向后两类^[86]. 以网络路径为主的量(介数、平均路径等), 时间信息分解相对较为困难^[87].

运用可视图分解可以测试时间序列的时间反演性质, 这是系统非线性的重要标志. 以可视图提出的评判标准是 $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$ 之间的差

异性强弱, 比如用 Kullback Leibler 散度 $D_{\text{KLD}}(\text{in}||\text{out})$ 来刻画^[87, 88]. 如果 $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$ 相同, 散度值 $D_{\text{KLD}}(\text{in}||\text{out})$ 为零, 则说明时间序列具有时间反演性; 反之, 大于零的散度值则表明 $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$ 是不同的, 时间序列不具有时间反演性. 但是需要注意, KLD 散度只能一阶近似反映时间序列的时间反演性质, 具体的 D_{KLD} 数值可能会受到序列的长程关联性的影响, 通常还需要进行假设检验^[89].

Donges 等人^[86] 提出时间反演性质的 KS 检验. 在时间反演的零假设下, $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$ 应该来自同一个分布. 而备择假设成立只需要验证 $p(k^{\text{out}})$ 和 $p(k^{\text{in}})$ 之间存在足够的统计偏差. KS 检验同样可以对局域类聚系数的两个分布 $p(C^{\text{out}})$ 和 $p(C^{\text{in}})$ 进行.

4.3 转换网络

转换网络源自时间序列符号化分析, 网络节点对应于某些离散状态或模式, 连边对应于两个离散状态之间在时间轴上具有可观测的非零概率. 由此可见, 连边定义和描述马尔可夫链中用到的转移概率一致, 因此, 转换网络与马尔可夫链等数学理论息息相关^[35].

4.3.1 符号编码和马尔科夫链

从连续数据中定义转换网络的节点有多种方法, 这通常被称为符号编码^[4]. 符号编码将时间序列转换成一组 K 个离散状态或模式(“符号”) $\{\pi_1, \dots, \pi_K\}$, 进一步运用复杂性指标刻画 K 个离散状态的非均匀性^[4]. 一定程度上, 符号编码的定义具有很大的自由. 在实践中, 常见方法有 3 个^[90]. (1) 相空间静态粗粒化, 即根据一组预先定义的阈值 $\{\xi_1, \dots, \xi_{K-1}\}$, 将数据分为 K 个不同的组. 在这种情况下, 符号 π_p ($p = 1, \dots, K$) 根据各自组中数据值(即 $\pi_p = q$) 的关系进行分配. (2) 根据 p 阶差分滤波序列进行动态粗粒化, 如 $p = 1$, 则差分序列为 $\Delta x_i = x_{i+1} - x_i$, 然后以 Δx_i 为分析对象进行粗粒化, 此时的粗粒化反映了时间序列增加还是减小的趋势. (3) 排序编码法^[91] 考虑 m 个连续数值, 用其数值大小的排序关系来编码(最多有 $m!$ 种排序), 这和置换熵的分析一致. 其他方法可参考文献^[90], 基于符

合编码以后的其他分析可以参考[91].

每一个符号也就成了描述动力系统的状态, 也就是转换网络的节点, 节点连接用两个符号之间转移概率来描述, 即构成有向有权重的网络 $\mathbf{W} = \{w_{pq}\}$, $p, q \in [1, \dots, K]$, 其中 $w_{pq} = p(\xi_{i+1} = \pi_q | \xi_i = \pi_p)$. 通常情况下, 矩阵需要按行归一, 即有 $\sum_{q=1}^K w_{pq} = 1$ 对所有 $p \in [1, K]$. 显见, 转换网络即是用马尔可夫链描述了经过符号变换的动力系统. 如果时间序列存在时间关联, 在计算 $\mathbf{W}$ 时可以直接引入时间信息, 比如从节点 p 到 q 在时间 t_i 时存在连边可以记为 $(p \rightarrow q; i)$, 由此可以构建时变转换网络[92]. 转换网络节点定义的其他形式请参考文献[8], 下面介绍排序转换网络[20,93-96].

4.3.2 一维序列排序转换图

从一维时间序列 $\{x_i\}$ 构建排序转换网络, 第一步需要重构出合适的相空间矢量, 即选取重构维数 m 和时间滞后 τ , 这是构建转换网络的两个重要的参数[94,97,98]. 本文建议采用传统的相空间重构的标准(见4.1节中的讨论). 每一个重构相空间点用其对应的大小排序关系来替代, 即为网络节点. 网络节点 p 和 q 之间的连接用转换频率 w_{pq} 来刻画. 图4展示了构建的具体方法[20].

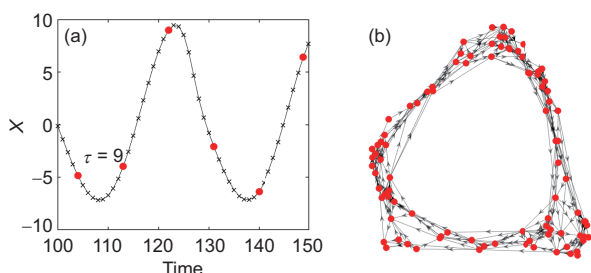


图 4 (网络版彩图)一维序列排序转换图的构建方法. (a) 时间序列排序关系的定义, 假设 $\tau = 9$ 和 $m = 6$. 重构相空间点 $\mathbf{x}_{104} = \{x_{104}, x_{113}, x_{122}, x_{131}, x_{140}, x_{149}\}$ 及其排序关系记为 $\pi_{104} = \{5, 1, 2, 4, 6, 3\}$. (b) 相应的转换网络. 引自文献[8].

Figure 4 (Color online) The construction method of one-dimensional sequence ordering conversion graph. (a) Illustration of permutations from an example time series. Assume $\tau = 9$ and $m = 6$. One embedded state vector $\mathbf{x}_{104} = \{x_{104}, x_{113}, x_{122}, x_{131}, x_{140}, x_{149}\}$ is highlighted, and its corresponding pattern is defined by the rank ordering $\pi_{104} = \{5, 1, 2, 4, 6, 3\}$. (b) Resulting OPTN. Reproduced from ref. [8].

进一步从局部和全局两个层面刻画转换网络节点之间转移复杂性[95]. 从局部来看, 由按行归一化可知, 从节点 π_p 到 π_q 的出边转移概率可以写为

$$p_{pq} = \begin{cases} 0, & \text{if } p = q, \\ \frac{w_{pq}}{\sum_{q:q \neq p} w_{pq}}, & \text{if } p \neq q. \end{cases} \quad (5)$$

由此, 节点 π_p 的出边连接复杂性香农熵为 $s_p^L = -\sum_q p_{pq} \log p_{pq}$. 根据每一个节点在网络中占有不同的比例, 全局出边转移复杂性熵为 $S^G = \sum_p p_p s_p^L$. 若熵值 $S^G = 0$, 则说明系统是确定性的, 且独立同均匀分布噪声对应最大熵值 S^G [95].

McCullough等人[20]以混沌Rössler系统为例讨论了排序转换网络的构建, 结果表明周期和混沌序列分别具有不同的网络结构, 并且和传统的李雅普诺夫指数比较发现, 网络的平均出度及其方差可以很好地区分不同的动力学行为[20].

4.3.3 多维序列的排序转换图

以三维序列 (x_i, y_i, z_i) 为例, 说明多变量排序转换网络的构建[99]. 首先得到差分序列 $(\Delta x_i, \Delta y_i, \Delta z_i)$. 网络节点定义为三维差分序列在时刻 i 的正负排列组合, 即根据 $\Delta x_i, \Delta y_i$ 和 Δz_i 的正负号共有8个网络节点 $(\pi_1, \dots, \pi_8)$, 如表2所示.

更一般, 对 n 维序列, 共有 2^n 种不同的排序关系. 通过这样定义的排序转换网络获取了原有时间序列在差分空间的性质, 反映了轨迹在相空间穿越零线的变化情况.

网络节点之间的连边根据时间上的转移频率而定. 系统确定性成分的高低决定了每一个排序子图 π_p 的频率高低 $p(\pi_p)$ 及其不同子图之间的转换概率 w_{pq} , 这可以由下面两个香农熵分别刻画:

表 2 三维序列 (x_i, y_i, z_i) 排序子图定义

Table 2 Order patterns in three-dimensional time series (x_i, y_i, z_i)

Π	π_1	π_2	π_3	π_4	π_5	π_6	π_7	π_8
Δx	+	+	+	+	-	-	-	-
Δy	+	+	-	-	+	+	-	-
Δz	+	-	+	-	+	-	+	-

$$S_O = - \sum_{p=1}^{2^n} p(\pi_p) \log_2 p(\pi_p), \quad (6)$$

$$S_T = - \sum_{p,q=1}^{2^n} w_{pq} \log_2 w_{pq}. \quad (7)$$

这两个复杂性统计量可以用来研究耦合振子产生相同步的动力学过程^[99], 发现在同步产生之前, 网络排序子图及其相互转换的概率分布较为均匀. 随着耦合强度的增大, 转换概率更趋异质, 并且伴随禁止排序子图的出现.

4.3.4 交叉和联合排序转换网络

交叉和联合排序转换网络是上述方法^[99,100]的进一步拓展. 以混沌Rössler为例说明: 设两个耦合系统的时间序列分别为 $(x_{1,i}, y_{1,i}, z_{1,i})$ 和 $(x_{2,i}, y_{2,i}, z_{2,i})$. 由此, 首先分别得到两个系统的差分序列为 $(\Delta x_{1,i}, \Delta y_{1,i}, \Delta z_{1,i})$ 和 $(\Delta x_{2,i}, \Delta y_{2,i}, \Delta z_{2,i})$.

交叉排序转换网络节点根据两个系统间差分序列差值的排序确定, 即根据 $(\Delta x_{1,i} - \Delta x_{2,i})$, $(\Delta y_{1,i} - \Delta y_{2,i})$ 和 $(\Delta z_{1,i} - \Delta z_{2,i})$ 这三个差值的正负号排列关系, 共得到8个网络节点, 其定义如表3所示. 为了更好说明, 图5(a)展示了两个耦合混沌Rössler系统在非

表 3 交叉转换网络节点定义

Table 3 Pattern definitions of a COPTN

Π	π_1	π_2	π_3	π_4	π_5	π_6	π_7	π_8
$\Delta x_1 - \Delta x_2$	+	+	+	+	-	-	-	-
$\Delta y_1 - \Delta y_2$	+	+	-	-	+	+	-	-
$\Delta z_1 - \Delta z_2$	+	-	+	-	+	-	+	-

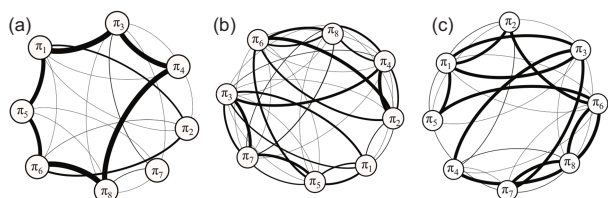


图 5 耦合Rössler系统非同步情形的交叉和联合排序转换网络图, 引自文献^[100]. (a) 交叉转换图的基本形式; (b) 交叉转换图的归一化形式; (c) 联合转换图. 修改自文献^[100]

Figure 5 Cross and joint OPTNs, which are from two coupled Rössler systems in the non-synchronized regime. (a) Normal cross OPTN; (b) alternative version; (c) joint OPTN. Adapted from ref. ^[100].

同步情形下的交叉排序转换网络.

很明显, 交叉网络节点定义受到时间序列数值的直接影响, 如果两个序列幅值相差很大(在多维序列中还有可能物理量单位不同), 那么上述网络节点定义的适用性就会降低, 因此可以引入适当的归一化^[100], 此时不同节点之间的转换在非同步时更加随机, 如图5(b)所示.

联合排序转换网络的节点是根据两个系统差分序列乘积正负号而定义的, 即 $\Delta x_1 \cdot \Delta x_2$, $\Delta y_1 \cdot \Delta y_2$ 和 $\Delta z_1 \cdot \Delta z_2$, 根据正负号的组合, 同样可得到 $K = 8$ 个不同的排序关系, 即网络节点的定义和表3类似, 联合转换网络如图5(c)所示. 和交叉转换图对比, 联合转换网络节点定义考虑了两个系统变化趋势的一致性, 和具体的数值大小和单位无关, 因此很容易推广为 n 个相互作用的子系统的情形. 最后指出, 在多维时间序列分析角度看, 交叉和联合转换网络提供了两种不一样的方法, 可以提供互为补充的信息.

5 结论和展望

复杂网络为非线性时间序列分析提供了崭新的方法, 并且取得了大量的应用^[8]. 根据网络节点和连边的不同定义, 本文仅回顾了三种主要方法, 即递归网络、可视图和转换网络. 这些不同方法为原有非线性时间序列分析提供了包括相空间几何结构和动力学特性两个方面的补充知识.

时间序列网络方法方兴未艾, 还有很大的拓展空间, 尤其需要关注网络方法和传统方法的比较研究, 这将极大推动非线性时间序列的分析. 下面探讨本文作者认为重要的几个可能方向.

- 时序网络方法: 目前大部分网络方法都是揭示了时间序列的静态网络结构特性, 建立时间序列的演化网络, 通过研究网络节点和连边的时间依赖性来刻画序列动态演化特性. 因此网络的构建与分析需要引入时间信息, 比如考虑连边的时序性, 构建时序网络可以提供新思路, 提供诸如时间序列记忆性的刻画^[92,101].

- 多层多路网络的构建: 本文回顾了构建多层多路递归网络和可视图的主要方法. 在实际中, 构建的网络需要考虑多变量时间序列在不同的时间和

空间尺度上的相互作用, 并且能够反映不同层次的作用斑图.

- 时间序列网络法的反问题研究, 即从网络到时间序列的重构: 目前大多数研究关注的是时间序列的网络结构特性, 对其反问题的研究具有很大的挑战, 通用的结果不多^[102–106]. 重构时间序列的反问题, 主要难点在于网络拓扑结构不受节点重标号的影响, 会丢失序列的时间信息(如递归网络), 而这些恰是时间序列重构的关键. 重构序列可以进一步展开统计推断和假设检验等研究工作.

- 数据挖掘方法的结合: 由于大数据和云计算的出现, 可用数据集的规模和复杂性不断增加, 在计算机科学研究领域提出的数据挖掘工具得到了相当

大的快速发展, 这在时间序列挖掘中同样得到了许多应用. 时间序列挖掘的部分研究问题和时间序列网络分析具有相似性^[107], 比如, 挖掘序列的模体、分块和类聚等, 这些方法可以和时间序列网络分析相结合, 分析和展示更复杂的数据特性.

- 时间序列的预测: 运用复杂网络方法进行时间序列的建模与预测方面的工作鲜有报道, 这可以说是非线性时间序列分析的最终目标^[4,7], 应该成为今后的主要方向.

最后, 为了实现时间序列网络的展现和统计分析, 我们推荐计算软件包pyunicorn和Matlab环境下的CRP工具箱, 两者可以从<http://tocsy.pik-potsdam.de>获得.

参考文献

- 1 Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work, and Think. New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013
- 2 Hurwitz J, Kirsch D. Machine Learning For Dummies, IBM Limited Edition. New York: John Wiley & Sons, Inc., 2018
- 3 Albert R, Barabási A L. Statistical mechanics of complex networks. *Rev Mod Phys*, 2002, 74: 47–97
- 4 Kantz H, Schreiber T. Nonlinear Time Series Analysis. 2nd ed. Cambridge: Cambridge University Press, 2004
- 5 Abarbanel H D I, Brown R, Sidorowich J J, et al. The analysis of observed chaotic data in physical systems. *Rev Mod Phys*, 1993, 65: 1331–1392
- 6 Ott E. Chaos in Dynamical Systems. 2nd ed. Cambridge: Cambridge University Press, 2002
- 7 Bradley E, Kantz H. Nonlinear time-series analysis revisited. *Chaos*, 2015, 25: 097610
- 8 Zou Y, Donner R V, Marwan N, et al. Complex network approaches to nonlinear time series analysis. *Phys Rep*, 2019, 787: 1–97
- 9 Packard N H, Crutchfield J P, Farmer J D, et al. Geometry from a time series. *Phys Rev Lett*, 1980, 45: 712–716
- 10 Grassberger P, Procaccia I. Characterization of strange attractors. *Phys Rev Lett*, 1983, 50: 346–349
- 11 Vejmelka M, Paluš M. Inferring the directionality of coupling with conditional mutual information. *Phys Rev E*, 2008, 77: 026214
- 12 Granger C W J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 1969, 37: 424–438
- 13 Schreiber T. Measuring information transfer. *Phys Rev Lett*, 2000, 85: 461–464
- 14 Rosenblum M G, Pikovsky A S. Detecting direction of coupling in interacting oscillators. *Phys Rev E*, 2001, 64: 045202
- 15 Paluš M, Vejmelka M. Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Phys Rev E*, 2007, 75: 056211
- 16 Romano M C, Thiel M, Kurths J, et al. Estimation of the direction of the coupling by conditional probabilities of recurrence. *Phys Rev E*, 2007, 76: 036211
- 17 Hlavackovaschindler K, Palus M, Vejmelka M, et al. Causality detection based on information-theoretic approaches in time series analysis. *Phys Rep*, 2007, 441: 1–46
- 18 Stankovski T, Pereira T, McClintock P V E, et al. Coupling functions: Universal insights into dynamical interaction mechanisms. *Rev Mod Phys*, 2017, 89: 045001
- 19 Donner R V, Heitzig J, Donges J F, et al. The geometry of chaotic dynamics—A complex network perspective. *Eur Phys J B*, 2011, 84: 653–672
- 20 McCullough M, Small M, Stemler T, et al. Time lagged ordinal partition networks for capturing dynamics of continuous dynamical systems. *Chaos*, 2015, 25: 053101
- 21 Boccaletti S, Latora V, Moreno Y, et al. Complex networks: Structure and dynamics. *Phys Rep*, 2006, 424: 175–308
- 22 Costa L F, Rodrigues F A, Travieso G, et al. Characterization of complex networks: A survey of measurements. *Adv Phys*, 2007, 56: 167–242
- 23 Barabási A-L. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512

- 24 Boccaletti S, Bianconi G, Criado R, et al. The structure and dynamics of multilayer networks. *Phys Rep*, 2014, 544: 1–122
- 25 Donges J F, Schultz H C H, Marwan N, et al. Investigating the topology of interacting networks. Theory and application to coupled climate subnetworks. *Eur Phys J B*, 2011, 84: 635–651
- 26 Marwan N, Donges J F, Zou Y, et al. Complex network approach for recurrence analysis of time series. *Phys Lett A*, 2009, 373: 4246–4254
- 27 Donner R V, Zou Y, Donges J F, et al. Recurrence networks—A novel paradigm for nonlinear time series analysis. *New J Phys*, 2010, 12: 033025
- 28 Donner R V, Small M, Donges J F, et al. Recurrence-based time series analysis by means of complex network methods. *Int J Bifurcat Chaos*, 2011, 21: 1019–1046
- 29 Zhang J, Small M. Complex network from pseudoperiodic time series: Topology versus dynamics. *Phys Rev Lett*, 2006, 96: 238701
- 30 Yang Y, Yang H. Complex network-based time series analysis. *Phys A-Statist Mech Appl*, 2008, 387: 1381–1386
- 31 Xu X, Zhang J, Small M. Superfamily phenomena and motifs of networks induced from time series. *Proc Natl Acad Sci*, 2008, 105: 19601–19605
- 32 Lacasa L, Luque B, Ballesteros F, et al. From time series to complex networks: The visibility graph. *Proc Natl Acad Sci*, 2008, 105: 4972–4975
- 33 Luque B, Lacasa L, Ballesteros F, et al. Horizontal visibility graphs: Exact results for random time series. *Phys Rev E*, 2009, 80: 046103
- 34 Donner R V, Donges J F. Visibility graph analysis of geophysical time series: Potentials and possible pitfalls. *Acta Geophys*, 2012, 60: 589–623
- 35 Nicolis G, CantúA G A, Nicolis C. Dynamical aspects of interaction networks. *Int J Bifurcat Chaos*, 2005, 15: 3467–3480
- 36 Barthélemy M. Spatial networks. *Phys Rep*, 2011, 499: 1–101
- 37 Takens F. Detecting strange attractors in turbulence. In: Rand D A, Young L S, eds. *Dynamical Systems and Turbulence*, Warwick 1980, Volume 898 of Lecture Notes in Mathematics. New York: Springer, 1981. 366–381
- 38 Eckmann J P, Kamphorst S O, Ruelle D. Recurrence plots of dynamical systems. *Europhys Lett*, 1987, 4: 973–977
- 39 Marwan N, Romano M C, Thiel M, et al. Recurrence plots for the analysis of complex systems. *Phys Rep*, 2007, 438: 237–329
- 40 Marwan N, Thiel M, Nowaczyk N R. Cross recurrence plot based synchronization of time series. *Nonlin Process Geophys*, 2002, 9: 325–331
- 41 Romano M C, Thiel M, Kurths J, et al. Multivariate recurrence plots. *Phys Lett A*, 2004, 330: 214–223
- 42 Eroglu D, Marwan N, Stebich M, et al. Multiplex recurrence networks. *Phys Rev E*, 2018, 97: 012312
- 43 Pham T D. Fuzzy recurrence plots. *Europhys Lett*, 2016, 116: 50008
- 44 Graben P, Hutt A. Detecting recurrence domains of dynamical systems by symbolic dynamics. *Phys Rev Lett*, 2013, 110: 154101
- 45 Costa D G B, Reis B M F, Zou Y, et al. Recurrence density enhanced complex networks for nonlinear time series analysis. *Int J Bifurcat Chaos*, 2018, 28: 1850008
- 46 Yang H, Liu G. Self-organized topology of recurrence-based complex networks. *Chaos*, 2013, 23: 043116
- 47 Liu C, Zhou W X. Superfamily classification of nonstationary time series based on DFA scaling exponents. *J Phys A-Math Theor*, 2010, 43: 495005
- 48 Donges J F, Heitzig J, Donner R V, et al. Analytical framework for recurrence network analysis of time series. *Phys Rev E*, 2012, 85: 046105
- 49 Zou Y, Heitzig J, Donner R V, et al. Power-laws in recurrence networks from dynamical systems. *Europhys Lett*, 2012, 98: 48001
- 50 Jacob R, Hari Krishnan K P, Misra R, et al. Cross over of recurrence networks to random graphs and random geometric graphs. *Pramana-J Phys*, 2017, 88: 37
- 51 Jacob R, Hari Krishnan K P, Misra R, et al. Characterization of chaotic attractors under noise: A recurrence network perspective. *Commun Nonlinear Sci Numer Simul*, 2016, 41: 32–47
- 52 Donner R V, Zou Y, Donges J F, et al. Ambiguities in recurrence-based complex network representations of time series. *Phys Rev E*, 2010, 81: 015101
- 53 Kraemer K H, Donner R V, Heitzig J, et al. Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions. *Chaos*, 2018, 28: 085720
- 54 Wiedermann M, Donges J F, Kurths J, et al. Mapping and discrimination of networks in the complexity-entropy plane. *Phys Rev E*, 2017, 96: 042304
- 55 Liu J L, Yu Z G, Anh V. Topological properties and fractal analysis of a recurrence network constructed from fractional Brownian motions. *Phys Rev E*, 2014, 89: 032814
- 56 Zou Y, Donner R V, Kurths J. Analyzing long-term correlated stochastic processes by means of recurrence networks: Potentials and pitfalls. *Phys Rev E*, 2015, 91: 022926
- 57 Thiel M, Romano M C, Kurths J, et al. Influence of observational noise on the recurrence quantification analysis. *Phys D-Nonlinear Phenom*, 2002, 171: 138–152
- 58 Subramaniyam N P, Hyttinen J. Characterization of dynamical systems under noise using recurrence networks: Application to simulated and

- EEG data. *Phys Lett A*, 2014, 378: 3464–3474
- 59 Goswami B, Boers N, Rheinwalt A, et al. Abrupt transitions in time series with uncertainties. *Nat Commun*, 2018, 9: 48
- 60 Feldhoff J H, Donner R V, Donges J F, et al. Geometric detection of coupling directions by means of inter-system recurrence networks. *Phys Lett A*, 2012, 376: 3504–3513
- 61 Gao Z K, Zhang X W, Jin N D, et al. Multivariate recurrence network analysis for characterizing horizontal oil-water two-phase flow. *Phys Rev E*, 2013, 88: 032910
- 62 Gao Z K, Yang Y X, Cai Q, et al. Multivariate weighted recurrence network inference for uncovering oil-water transitional flow behavior in a vertical pipe. *Chaos*, 2016, 26: 063117
- 63 Gao Z K, Yang Y X, Zhai L S, et al. Multivariate multiscale complex network analysis of vertical upward oil-water two-phase flow in a small diameter pipe. *Sci Rep*, 2016, 6: 20052
- 64 Nagy G. Terrain visibility. *Comput Graphics*, 1994, 18: 763–773
- 65 De Florian L, Marzano P, Puppo E. Line-of-sight communication on terrain models. *Int J Geograph Inf Syst*, 1994, 8: 329–342
- 66 Lacasa L, Just W. Visibility graphs and symbolic dynamics. *Phys D-Nonlinear Phenom*, 2018, 374-375: 35–44
- 67 Nuñez A, Lacasa L, Patricio J, et al. Visibility algorithms: A short review. In: *New Front. Graph Theory*. InTech, 2012. 119–152
- 68 Zou Y, Donner R V, Marwan N, et al. Long-term changes in the north-south asymmetry of solar activity: A nonlinear dynamics characterization using visibility graphs. *Nonlin Process Geophys*, 2014, 21: 1113–1126
- 69 Gao Z K, Cai Q, Yang Y X, et al. Multiscale limited penetrable horizontal visibility graph for analyzing nonlinear time series. *Sci Rep*, 2016, 6: 35622
- 70 Mutua S, Gu C, Yang H. Visibility graphlet approach to chaotic time series. *Chaos*, 2016, 26: 053107
- 71 Lacasa L, Iacovacci J. Visibility graphs of random scalar fields and spatial data. *Phys Rev E*, 2017, 96: 012318
- 72 Lan X, Mo H, Chen S, et al. Fast transformation from time series to visibility graphs. *Chaos*, 2015, 25: 083105
- 73 Lacasa L, Luque B, Luque J, et al. The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion. *Europhys Lett*, 2009, 86: 30001
- 74 Ni X H, Jiang Z Q, Zhou W X. Degree distributions of the visibility graphs mapped from fractional Brownian motions and multifractal random walks. *Phys Lett A*, 2009, 373: 3822–3826
- 75 Ahmadi M, Adeli H, Adeli A. Improved visibility graph fractality with application for the diagnosis of Autism Spectrum Disorder. *Phys A-Stat Mech Appl*, 2012, 391: 4720–4726
- 76 Lacasa L. On the degree distribution of horizontal visibility graphs associated with Markov processes and dynamical systems: Diagrammatic and variational approaches. *Nonlinearity*, 2014, 27: 2063–2093
- 77 Zou Y, Small M, Liu Z, et al. Complex network approach to characterize the statistical features of the sunspot series. *New J Phys*, 2014, 16: 013051
- 78 Gonçalves B A, Carpi L, Rosso O A, et al. Time series characterization via horizontal visibility graph and Information Theory. *Phys A-Stat Mech Appl*, 2016, 464: 93–102
- 79 Lacasa L, Toral R. Description of stochastic and chaotic series using visibility graphs. *Phys Rev E*, 2010, 82: 036120
- 80 Manshour P. Complex network approach to fractional time series. *Chaos*, 2015, 25: 103105
- 81 Zhang R, Zou Y, Zhou J, et al. Visibility graph analysis for re-sampled time series from auto-regressive stochastic processes. *Commun Nonlinear Sci Numer Simul*, 2017, 42: 396–403
- 82 Xie W J, Zhou W X. Horizontal visibility graphs transformed from fractional Brownian motions: Topological properties versus the Hurst index. *Phys A-Stat Mech Appl*, 2011, 390: 3592–3601
- 83 Iacovacci J, Lacasa L. Sequential visibility-graph motifs. *Phys Rev E*, 2016, 93: 042309
- 84 Lacasa L, Nicosia V, Latora V. Network structure of multivariate time series. *Sci Rep*, 2015, 5: 15508
- 85 Sannino S, Stramaglia S, Lacasa L, et al. Visibility graphs for fMRI data: Multiplex temporal graphs and their modulations across resting-state networks. *Network Neurosci*, 2017, 1: 208–221
- 86 Donges J F, Donner R V, Kurths J. Testing time series irreversibility using complex network methods. *Europhys Lett*, 2013, 102: 10004
- 87 Lacasa L, Nuñez A, Roldán E, et al. Time series irreversibility: A visibility graph approach. *Eur Phys J B*, 2012, 85: 217
- 88 Lacasa L, Flanagan R. Time reversibility from visibility graphs of nonstationary processes. *Phys Rev E*, 2015, 92: 022817
- 89 Telesca L, Czechowski Z. Relation between HVG-irreversibility and persistence in the modified Langevin equation. *Chaos*, 2018, 28: 073107
- 90 Donner R, Hinrichs U, Scholz-Reiter B. Symbolic recurrence plots: A new quantitative framework for performance analysis of manufacturing networks. *Eur Phys J Spec Top*, 2008, 164: 85–104

- 91 Bandt C, Pompe B. Permutation entropy: A natural complexity measure for time series. *Phys Rev Lett*, 2002, 88: 174102
- 92 Weng T, Zhang J, Small M, et al. Memory and betweenness preference in temporal networks induced from time series. *Sci Rep*, 2017, 7: 41951
- 93 Small M. Complex networks from time series: Capturing dynamics. In: 2013 IEEE International Symposium on Circuits and Systems (IS-CAS2013). Beijing, 2013. 2509–2512
- 94 Kulp C W, Chobot J M, Freitas H R, et al. Using ordinal partition transition networks to analyze ECG data. *Chaos*, 2016, 26: 073114
- 95 McCullough M, Small M, Iu H H C, et al. Multiscale ordinal network analysis of human cardiac dynamics. *Phil Trans R Soc A*, 2017, 375: 20160292
- 96 Small M, McCullough M, and Sakellariou K. Ordinal network measures—Quantifying determinism in data. In: 2018 IEEE International Symposium on Circuits and Systems (ISCAS). Florence, 2018. 1–5
- 97 McCullough M, Sakellariou K, Stemler T, et al. Counting forbidden patterns in irregularly sampled time series. I. The effects of under-sampling, random depletion, and timing jitter. *Chaos*, 2016, 26: 123103
- 98 Sakellariou K, McCullough M, Stemler T, et al. Counting forbidden patterns in irregularly sampled time series. II. Reliability in the presence of highly irregular sampling. *Chaos*, 2016, 26: 123104
- 99 Zhang J, Zhou J, Tang M, et al. Constructing ordinal partition transition networks from multivariate time series. *Sci Rep*, 2017, 7: 7795
- 100 Guo H, Zhang J Y, Zou Y, et al. Cross and joint ordinal partition transition networks for multivariate time series analysis. *Front Phys*, 2018, 13: 130508
- 101 Holme P, Saramäki J. Temporal networks. *Phys Rep*, 2012, 519: 97–125
- 102 Thiel M, Romano M C, Kurths J. How much information is contained in a recurrence plot? *Phys Lett A*, 2004, 330: 343–349
- 103 Hirata Y, Oda A, Ohta K, et al. Three-dimensional reconstruction of single-cell chromosome structure using recurrence plots. *Sci Rep*, 2016, 6: 34982
- 104 Robinson G, Thiel M. Recurrences determine the dynamics. *Chaos*, 2009, 19: 023104
- 105 Zhao Y, Weng T, Ye S. Geometrical invariability of transformation between a time series and a complex network. *Phys Rev E*, 2014, 90: 012804
- 106 McCullough M, Sakellariou K, Stemler T, et al. Regenerating time series from ordinal networks. *Chaos*, 2017, 27: 035814
- 107 Fu T. A review on time series data mining. *Eng Appl Artif Intell*, 2011, 24: 164–181

Nonlinear time series analysis by means of complex networks

ZOU Yong^{1*}, DONNER Reik V.^{2,3}, MARWAN Norbert³,
DONGES Jonathan F.^{3,4} & KURTHS Jürgen^{3,5}

¹*School of Physics and Electronic Sciences, East China Normal University, Shanghai 200241, China;*

²*Department of Water, Environment, Construction and Safety, Magdeburg-Stendal University of Applied Sciences, Breitscheidstraße 2, Magdeburg 39114, Germany;*

³*Potsdam Institute for Climate Impact Research (PIK) — Member of the Leibniz Association, Telegrafenberg A31, Potsdam 14473, Germany;*

⁴*Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, Stockholm 114 19, Sweden;*

⁵*Department of Physics, Humboldt University Berlin, Newtonstraße 15, Berlin 12489, Germany*

In the last decade, there has been a growing body of literatures addressing the utilization of complex network methods for the characterization of dynamical systems based on time series, which has allowed addressing fundamental questions regarding the structural organization of nonlinear dynamics as well as the successful treatment of a variety of applications from a broad range of disciplines. In this report, we provide an in-depth review of three existing approaches of recurrence networks, visibility graphs and transition networks, covering their methodological foundations, interpretation and the recent developments. The overall aim of this report is to provide the Chinese readers with the future directions of time series network approaches and how the complex network approaches can be applied to their own field of real-world time series analysis.

complex networks, nonlinear dynamics, recurrences, visibility, transition networks

PACS: 05.45.Ac, 05.45.Tp, 89.75.Fb

doi: [10.1360/SSPMA-2019-0136](https://doi.org/10.1360/SSPMA-2019-0136)