第 57 卷 第 11 期 2008 年 11 月
1000-3290/2008/57(11)/7380-10

物 理 学 报
ACTA PHYSICA SINICA

Vol.57,No.11,November,2008
©2008 Chin. Phys. Soc.

利用复杂网络研究中国温度序列的拓扑性质*

周 磊[1)2)] 龚志强[1)3)] 支 蓉[1)3)] 封国林[1)2)3)†]

1)(扬州大学物理科学与技术学院,扬州 225002)

2)(中国科学院大气物理研究所东亚区域气候-环境重点实验室,北京 100029)

3)(国家气候中心,中国气象局气候研究开放实验室,北京 100081)

(2008 年 1 月 28 日收到;2008 年 4 月 22 日收到修改稿)

依据粗粒化方法,将中国 1961—2002 年逐日平均温度序列转化为由 5 个特征字符 {R,r,e,d,D} 构成的温度符号序列.以符号序列中的 125 种 3 字串组成的温度波动模态为网络的节点(即连续 4 d 的温度波动组合),并按照时间顺序连边,构建有向加权的温度波动网络,进而将温度波动模态间的相互作用等综合信息蕴含于网络的拓扑结构之中.对随机序列和 Lorenz 系统的混沌序列分别构建随机和混沌波动网络.计算三种网络的度与度分布、聚类系数、最短路径长度等动力学统计量,从网络的角度对比研究三种序列内秉性质的差异.结果表明:温度波动网络的动力学拓扑性质与混沌网络相似,与随机网络差异显著,体现了温度变化的复杂性.此外,温度波动网络的重要节点中大都包含了 R,r 和 e 三种符号,可能对应了全球增暖的背景下,温度的波动主要以上升为主;温度波动网络部分节点的中介中心性能力较强,4% 的节点承担了网络 12.93% 的中介中心性功能,这些具有拓扑统计重要性的节点对于理解温度波动的内在规律和波动信息的传递等有一定指导意义.Lorenz 系统对应的混沌波动网络的中介中心性与温度网络类似,而随机波动网络中节点的中介中心性几乎是均等的.因此,有别于传统的 Lyapunov 指数、分维数等方法,从网络结构拓扑性质这一新的角度验证了温度变化具有类混沌属性,而不是随机变化过程.

关键词:气候变化,气候复杂网络,拓扑结构

PACC:9260X

1. 引 言

气候系统是一个十分庞大而又复杂的系统[1,2],对其复杂性进行研究具有重要的理论意义和实用价值[3].目前,讨论气候系统复杂性的方法很多,已有的研究大致分为三种:对描述气候系统动力过程的非线性微分方程的研究,对连续的气候观测资料的研究及对离散的历史气候等级序列的研究[3,4].温度作为气候系统的一个重要组成部分,在气候变化研究中有着重要的地位.随着全球变暖,人们对于温度的观测以及代用资料的研究越来越密切关注[5,6],以期揭示全球变暖的成因.20 世纪 90 年代初以来,科学家对近 50 年气候变化观测事实进行了系统分析,并取得了卓越成果.例如分区建立了 1951—1996 年平均温度序列,进而得到中国年平均温度序列[7];对中国地区年和季节平均温度、降水变化特点及其与全球变暖的关系进行了分析[8],对黄淮海地区近 40 年温度变化特征进行了研究[9];利用西藏地区 1961—2000 年月平均温度、最高气温和最低气温资料,分析了近 40 年高原气温的变化趋势[10].这些研究表明,20 世纪 60 年代以来中国气候总体趋向变暖,80 年代以后变暖尤为明显.但温度序列本身是非线性、非平稳的,传统的方法大都分析温度变化的趋势,很难揭示气象要素的动力学特征,不能突出其本质行为.

郝柏林[11]提出可以通过符号动力学建立运动轨道和形式语言的联系,然后借助语法复杂性理论来刻画复杂性,其核心内容就是符号动力学与时间序列的粗粒化.不同程度的粗粒化,舍去更小层次上的细节,使它们在所关注的层次中表现为某些特征量,有利于突出本质的特征.所以,利用复杂网络来研究经济系统、气候系统等复杂系统受到了广泛的关注.例如,Li 等[12]根据香港证券市场恒生指数构

* 国家自然科学基金(批准号:40675044),国家重点发展基础研究项目(批准号:2006CB400503)和国家科技支撑计划(批准号:2007BAC03A01)共同资助的课题.

† 通讯联系人. E-mail:fenggl@cma.gov.cn

建证券指数网络，模拟作为复杂网络的证券市场，获取香港证券市场的结构和变化特征.

本文针对中国气象局公布的全国 720 个常规观测站的观测数据，选取其中无缺测的 435 个站点 1961—2002 年日平均温度资料讨论温度序列的复杂性，旨在从复杂网络的角度揭示温度变化的动力学特征.依据粗粒化同质划分方法，按照等概率原理将温度涨落转换成符号序列，构建一个有向加权的温度波动网络，将相关的温度数值编码在网络的拓扑结构之中，并对网络的动力学统计特征量和拓扑参数等进行分析，以获取温度波动网络的内在规律性.在此基础上，对比研究随机和混沌网络，分析这 3 种网络的基本统计性质，进而更好地认识温度系统的复杂性特征.

2. 资料及其粗粒化

2.1. 资料和理论模型

我们所选用的温度资料是中国气象局公布的全国 720 个常规站观测数据中无缺测的 435 个站点 1961—2002 年的逐日平均温度序列 $T_i(t)$，($t=1,2,3,4,5,\cdots,N$，$i=1,2,\cdots,435$)，$N=14976$，并对所有站点温度序列计算平均值得到平均温度序列 $\bar{T}(t)$，即

$$\bar{T}(t)=\frac{1}{435}\sum_{i=1}^{435}T_i(t). \tag{1}$$

对 Lorenz 模型，取积分步长 0.01，相当于实际大气演化 10 min，按照相应一天的时间长度计算对应每一天的平均值以得到混沌序列 $C(t)$，从而保证 $C(t)$ 与 $\bar{T}(t)$ 具有相同的长度.应用随机函数，构建随机序列 $R(t)$，$N=14976$.

2.2. 三种序列的粗粒化方法

对复杂性的描述经常伴随着对系统的粗粒化和符号化.例如，构成 DNA 核苷酸分子只用 A，T，G，C 四种符号来代表，而所有生物的 DNA 都由这四种核苷酸编码而成.粗粒化常用的方法是对系统区间进行同质划分[13]，它是将整个区间平均分成有限个子区间.如果对划分后的每个子区间赋予一个符号，则整个系统区间就转换为一个符号序列.对粗粒化符号序列的研究也就是对相应时间序列的研究，由于粗粒化过程舍去了小层次上的细节，粗粒化形式的符号序列又具有有限性，这就使得粗粒化方法能突出系统的本质行为，得出更合理的结论[14].因此，我们对 $\bar{T}(t)$，$C(t)$ 和 $R(t)$ 分别做粗粒化处理，使其转换为字符形式表示的符号序列.本文的工作都将在这些离散的符号序列上进行，结论的有效性将在很大程度上依赖于粗粒化的精度.这就要求在进行符号转换时，确定的符号种类不能太多，每一个符号尽可能代表系统变化中的某一种基本的、相对独立的模态.

下面以温度序列为例扼要介绍粗粒化方法.首先，我们计算温度数值的波动 $k(t)$：

$$k(t)=\frac{\bar{T}(t+\Delta t)-\bar{T}(t)}{\Delta t}, \tag{2}$$

其中，Δt 为时间间隔，本文主要研究 $\Delta t=1$ 时任意连续两天之间的温度波动情况，Δt 取其他值的情况在下文中也做了简单讨论，并给出了标度分布.气候系统中温度波动是天气变化过程的一种重要体现方式.因此，我们将温度序列转换为符号序列就是先考虑温度系统中的数值波动.根据

$$P(k)=\int_{-\infty}^{k}\frac{Num(x)}{N}\mathrm{d}x \tag{3}$$

计算不同温度波动值可能出现的概率 $P(k)$.我们用等概率的思想将温度波动 $k(t)$ 划分为 5 个区间，把落在这 5 个区间 $k(t)$ 分别用符号 R,r,e,d,D 表示为

$$S_i=\begin{cases}R, & 0>P(k)>0.2,\\ r, & 0.2\geqslant P(k)>0.4,\\ e, & 0.4\geqslant P(k)>0.6,\\ d, & 0.6\geqslant P(k)>0.8,\\ D, & 0.8\geqslant P(k)>1.0.\end{cases} \tag{4}$$

其中 R 代表温度的快速上升，r 代表缓慢上升，e 代表相对平稳，d 代表缓慢下降，D 为快速下降.依据这一思想，我们将 42 年平均温度序列 $\bar{T}(t)$ 转化为相应的符号序列

$$s_{\mathrm{T}}=(s_1s_2s_3\cdots),s_i\in(R,r,e,d,D), \tag{5}$$

对 $C(t)$ 和 $R(t)$ 也进行类似处理，可以得到其符号序列

$$\begin{aligned}s_{\mathrm{C}}&=(s_1s_2s_3\cdots),s_i\in(R,r,e,d,D),\\ s_{\mathrm{R}}&=(s_1s_2s_3\cdots),s_i\in(R,r,e,d,D).\end{aligned} \tag{6}$$

天气系统中，人们关注的往往都是天气过程的变化情况，我们工作所要研究的重点是代表着连续几天温度涨落的组合模态，所以对这些温度数值波动进行分类时，应保证这些元模态在数量上是相等的，这样才能确保这些温度波动模态在理论上可能出现的概率也是均等的.

在温度数值序列转换为符号序列过程中时间间隔 Δt 是一个重要的参数,Δt 选取的大小不同,对时间序列的分辨率也就不同;对同一个时间序列,其对应的符号序列长度,各字符的占有量 $N(R)$,$N(r)$,$N(e)$,$N(d)$,$N(D)$以及字符之间的关联性也就不同.为此我们对 1961—2002 年平均温度序列 $\bar{T}(t)$,$C(t)$和 $R(t)$分别统计了不同 Δt 下所构建的字符序列中 $N(R)$,$N(r)$,$N(e)$,$N(d)$,$N(D)$的发生次数,发现 $N(R)$,$N(r)$,$N(e)$,$N(d)$,$N(D)$与时间间隔 Δt 均满足幂律关系,它们之间的关系如图 1 所示.

$$N \sim \Delta t^{\tau}. \tag{7}$$

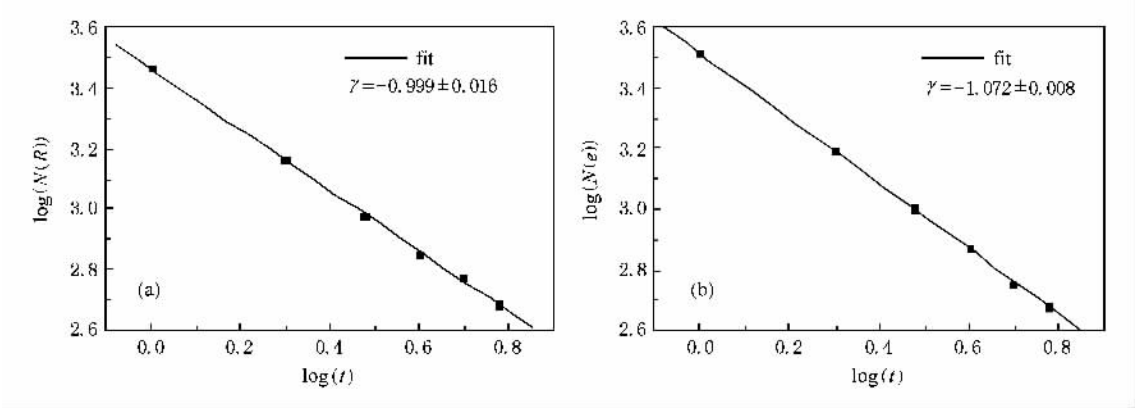

图 1 字符统计 $\log(N)$与 $\log(\Delta t)$关系图 (a),(b)分别为温度序列 $N(R)$-Δt 和 $N(e)$-Δt 关系图

显然,幂律关系的简单性反映了复杂系统的动力学特征的规律性;幂律指数 τ 的波动范围较小反映了动力学规律的稳定性;而关于各种字符 R,r,e,d,D 的动力学规律具有很好的一致性,反映了动力学规律的普遍性;字符幂律关系的双对数图为一很好的线性关系,反映了温度波动本身具有较好的无标度性[15,16].温度波动的无标度性反映了优选的网络演化过程[17,18],说明对符号进行粗粒化处理更倾向于取小的时间尺度,因此,本文将温度序列向符号序列有效转化,研究连续两天之间的温度波动规律[19].

3. 温度波动网络的构建及其拓扑性质

3.1. 三种波动网络

通过粗粒化处理得到的符号序列中,称有 n 个特征字符构成的字串为 n 字串.对给定的 n,可以有 5^n 种不同的 n 字串,这些 n 字串即为各种波动模态,即连续 n 个不同波动的组合[12].本文取 $n=3$ 时,共有 125 种波动模态,它们是 {*RRR*,*RRr*,*RRe*,*RRd*,*RRD*,*RrR*,*Rre*,*Rrd*,*RrD*,…}.图 2 给出了 *RRR*,*rrr* 和 *drD* 温度波动的示意图.

温度符号序列 S_T 中,表现为一种温度波动模

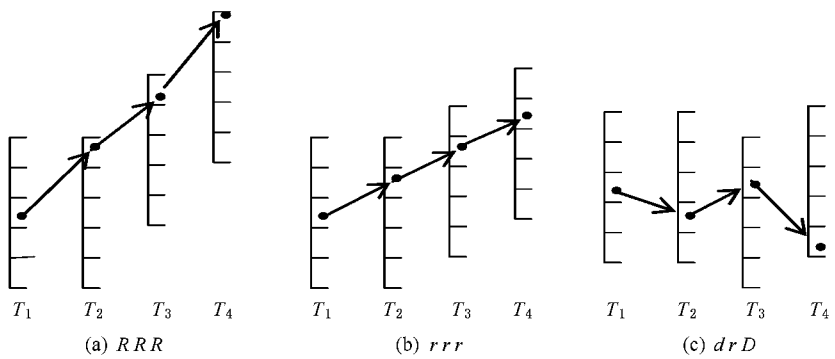

图 2 三种温度波动模态示意图 (a),(b)和(c)分别表示连续四天内温度均为快速上升的波动 *RRR*,温度均为缓慢上升的波动 *rrr*,温度为缓慢下降,缓慢上升,快速下降的波动 *drD*(黑点为当天温度值,5 个区间分别代表了 {*R*,*r*,*e*,*d*,*D*}5 种温度变化)

态向另一种模态的转换，转换过程中，模态之间可能存在着信息传递或支配等关系，具有多元化的特征；后一个模态的出现以前一个或多个模态为基础，体现了一定的记忆性，因此模态之间的联系形式必然是‘有向’的，联系的强弱程度也互不相同，各种不同的联系便构成了复杂系统. 显然，温度符号序列 S_T 中元结构相互作用的关系在符号序列 S_C 和 S_R 中也存在.

复杂网络是大量复杂系统得以存在的拓扑基础，它舍弃了小尺度上的细节作用而更关心其本质行为，为处理复杂系统提供了途径[20—22]. 近年来复杂网络在电力，社会，交通，Internet，生物学等许多领域得到了广泛应用[23—25]. 因为各种波动模态之间存在着多元化的作用关系，我们尝试从符号动力学的角度应用复杂网络理论来描述温度、混沌和随机符号序列内部各种相互作用. 为此，我们引入一个加权网络来描述温度序列中各波动模态之间的作用和联系，其中网络的节点就是 125 个 3 字串的波动模态；边为前一个节点指向它的下一个节点，即一种模态向下一个模态转换，表征了一种天气过程向另一种天气过程的转变. 如果网络中的两个节点之间有多条互不相交的并联连接通路，则并联连接的通路数就是这两个节点连边的权重. 例如，在我们所构建的网络中，温度符号序列为

RRdDeRRrdRDDDrRDDDrRedDDDdRrrRRRRerreRRd
rDDdrRdDrDDRdDRRRDDDDRRedeRRRdrdRDDD…

以 3 字串的元结构 {*RRR*，*RRr*，*RRe*，*RRd*，*RRD*，*RrR*，*Rre*，*Rrd*，*RrD*，…} 作为网络的节点，则网络节点的有向连接形式为

RRd → *DeR* → *Rrd* → *RDD* → *DrR* →
DDD → *rRe* → *dDD* → *DdR* → *rrR* →
RRR → *err* → *eRR* → *drD* → *Ddr* →
RdD → *rDD* → *RdD* → *RRR* → *DDD* →
RRe → *deR* → *RRd* → *rdR* → *DDD* → …

由此我们可以构建体现三种序列内部各波动模态间相互作用的有向含权网络图（图 3），其权重就是两个节点之间的直接连边次数，反映了两个节点之间的关联程度. 为了体现网络的结构特征，反映网络的一些基本性质，我们仅给出了部分节点的关联图像.

图 3(a)是温度序列部分节点构建的复杂网络，节点与节点连边线条粗细反映了节点之间关联程度的强弱，例如节点 *DDD* 与节点 *rRR* 之间的连线最粗，表示两种温度波动模态之间的关联程度最为密集，这两种变化模态在长期的温度变化中具有很好的联系. 图 3(b)是随机序列部分节点构建的复杂网络，随机网络中节点之间连线的粗细基本相同，即节点间的关联强度比较平均，没有关联特别密集的节点对. 图 3(c)是混沌序列部分节点构建的复杂网络，由于 Lorenz 系统本身具有的复杂性，混沌波动网络中节点之间连边也存在粗细差异，即节点之间的关联程度也有一定的差异. 因此就网络的外部特征而言，温度和混沌的波动网络具有类似特征，与随机波动网络差异较显著.

温度波动网络的有向性可以直观地反映表征各种温度波动模态的节点的变化方向；节点之间连边关系的权重反映着长期以来温度资料中一种变化模态向另一种变化模态的转换频数，研究这些不同变化模态之间转换的方向和频数，可以为温度的预测提供一种可能的新途径.

3.2. 波动网络的动力学拓扑性质

3.2.1. 度与度分布

度是节点属性中简单而又重要的概念，网络中所有节点度的平均值，即为平均顶点度〈*k*〉. 我们所构建的温度波动网络是一个有向网络，其节点度分为出度和入度，出度是指从某一节点指向其他节点

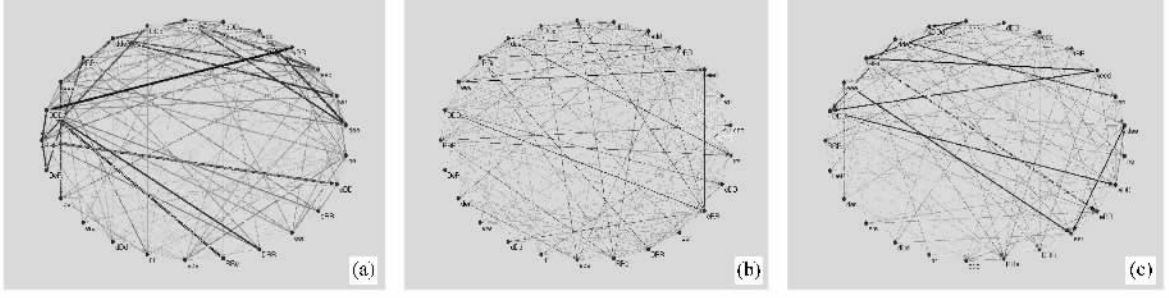

图 3　三种序列的部分节点构建的复杂网络　(a)，(b)和(c)分别为温度波动网络、随机波动网络和混沌波动网络

的边数，入度是指从其他节点指向该节点的边数[26]．以温度波动网络为例，节点的入度表示其他温度波动模态向某一模态的直接转换，节点的出度代表着这种波动模态向其他模态的直接转换．因为节点之间的连边是通过时间顺序来连接的，所以在温度波动网络中，除了第一个节点和最后一个节点，其他节点的出度和入度必定是相等的，所以只要研究其中的一种就可以来研究温度波动网络中的度与度分布了．我们选择顶点的出度作为研究对象，即一种波动模态向另一种模态间的转换，中间不需要其他节点的中转，表现为各种波动模态之间的短程相关．度的大小表示它与其他温度波动模态之间短程相关的强弱，一个节点的度越大就意味着这个节点在某种意义上越重要[26]．

表 1 为温度波动网络各种波动模态的节点度大小的排序，发现节点 *RRR*，*rRr*，*RRe*，…的节点度数比较大，这说明在温度波动网络中，这些节点所代表的温度波动模态在天气尺度的温度变化过程中起到重要的直接关联作用，各种温度波动模态向这几个重要模态转换，或被这几个重要模态转换的频率较高．另外，我们还对网络的节点度进行了字频统计，发现在度数较大的前 15 个节点中，字符 *R* 出现的次数为 20，而代表温度值急剧下滑的字符 *D* 出现的次数仅为 3，说明在温度变换中，带有急剧上升的温度波动模态与其他波动模态之间的直接转换或被转换的概率较高．从这一侧面可以反映出在全球变暖的大背景之下[5,6]，急剧上升的温度波动在温度变换中出现的次数越来越多．另外表中还显示着代表连续 3 天都急剧上升的极端温度波动模态（*RRR*）在与其他变换模态之间的关联程度越来越强，这一极端温度波动模态的强关联性质与全球气温变暖有着直接的关系．

表 1 温度波动网络各种波动模态的节点度大小的排序

节点	*RRR*	*rrd*	*rRr*	*RRe*	*eee*	*erR*	*Rdd*	*RRr*	*Rre*
度	330	180	172	164	156	140	130	128	116
等级	1	2	3	4	5	6	7	8	9
节点	*RRd*	*Rrr*	*rRR*	*dDD*	*eRR*	*RRD*	…	*DRd*	
度	114	112	110	106	104	102	…	2	
等级	10	11	12	13	14	15	…	125	

图 4 给出三种网络的累计节点度分布．一方面，半对数坐标系下，温度和混沌波动网络均服从衰减的指数分布，且指数值较接近，但随机波动网络则不服从衰减指数分布；另一方面，随机网络的度分布范围相对较狭窄，度显著较大的节点数较温度波动网络和混沌波动网络要少许多．由此可知温度序列的波动特征有别于随机序列的随机波动特征，具有和 Lorenz 系统变量相似的混沌特征．此外，衰减的指数分布意味着随机挑选的网络演化过程[17,18]，这说明各种温度波动模态的发生带有一定的随机性（即何时发生何种波动模态是不确定的），但各种温度波动模态出现的次数在一个较长的时间段内是遵循一定规律的．温度涨落变化的这种规律类似于地球系统中各种等级地震发生的次数分布，即服从指数分布，其原因可能在于气候系统、地球系统等各种复杂系统均遵循系统最复杂原理（最大熵原理）[27—29]，而指

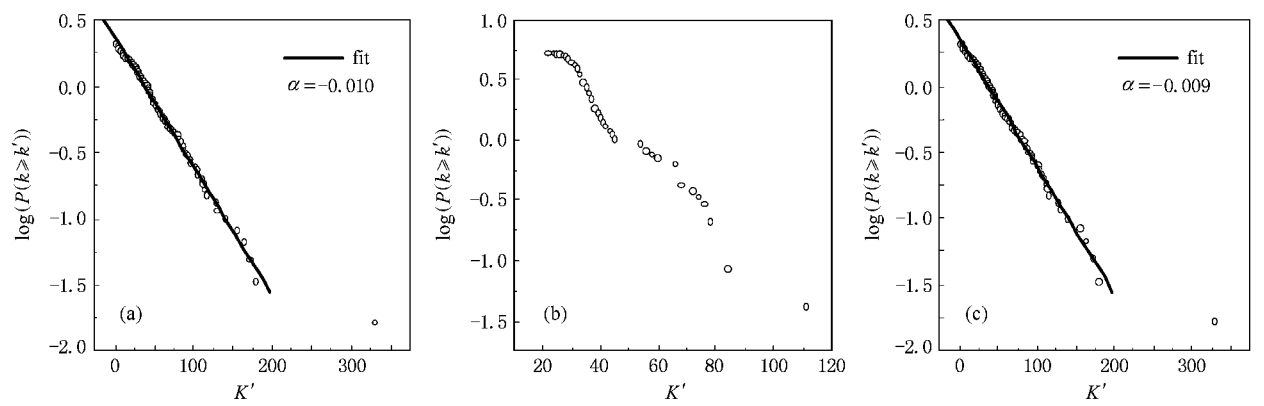

图 4 累计节点度分布 (a)，(b)和(c)分别为温度波动网络、随机波动网络和混沌波动网络

数分布则是其重要的表现形式之一，因此这一分布在一定程度上体现了复杂系统的内在动力学本质特征.

虽然温度和混沌波动网络都服从衰减的指数分布，但是由于 Lorenz 系统序列是在理论意义上产生的，在外界强迫参数不变的情况下，它是一个平稳的混沌系统，温度波动网络的指数衰减较混沌波动网络快.

3.2.2. 聚类系数

聚类系数反映网络的紧密程度，整个网络的聚类系数 C 就是所有节点 i 的聚类系数 c_i 的平均值. 温度波动网络是含权网络，考虑边权，则网络中某一节点的聚类系数可以定义为

$$C_i^w = \frac{1}{s_i(k_i-1)}\sum_{(j,k)}\frac{w_{ij}+w_{ik}}{2}a_{ij}a_{jk}a_{ik},$$

其中 w_{ij} 表示两个节点 (i,j) 连边的边权，s_i 表示节点 i 的点强度，定义为 $s_i=\sum_j w_{ij}$，即节点 i 的邻边边权之和，k_i 为节点 i 的度，$\sum_{k>j}a_{ij}a_{jk}a_{ik}$，表示网络中包含节点 i 的三角形的总数. 很明显，$0\leqslant C\leqslant 1$，$C=0$ 当且仅当所有的节点均为孤立节点时，没有任何连接边；$C=1$ 当且仅当网络是全局耦合的，即网络中任意两个节点都直接相连[26].

图 5 给出了三种网络的聚类系数随节点度的分布图，其中 $C(k)$ 是度为 k 的所有节点聚类系数的平均值. 随机网络中的聚类系数随度分布并没有特性，不同的度对应的聚类系数相差不大，而在温度波动网络中聚类系数的大小随节点度的变化幅度较大，这意味着温度波动网络并不是完全随机的，而是在某种程度上具有类似于社会关系网络中“物以类聚，人以群分”的特性[26]. 它又不像社会网络表现出来的那样，度数小的节点有着高的聚类系数，度大的节点有着低的聚类系数，即温度网络比较明显的群聚特性既可能会发生在小的集团中，也可能会发生在大的集团中. 在随时间变化的实际温度资料中，温度变化的群发性有时会反映在小的时间尺度上，也有时会反映在大的时间尺度上. 通过对温度波动网络聚类系数的研究，可以为我们以后研究温度变化的群发性提供一定的参考和帮助.

在图 5 中，温度系统和 Lorenz 系统相似，这两种系统下网络的聚类系数随度分布都表现出一定的特性，体现出温度系统的复杂性. 同时，因为温度序列是非线性、非平稳的复杂系统，而 Lorenz 系统在外界强迫参数不变的情况下是一个平稳的混沌系统，所以温度网络中的差异波动要更加明显，所反映出来的特性要比 Lorenz 系统更加强烈，表现出更高的复杂性.

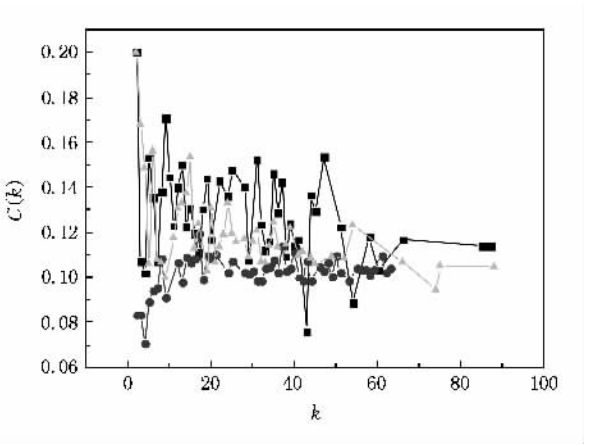

图 5　聚类系数随节点度的分布(■，●，▲分别表示温度，Lorenz 和随机网络)

3.2.3. 平均最短路径长度

两个节点 i 和 j 之间的距离为 d_{ij}(即任意两个节点相互连接所要经过的连边数)，网络的平均路径长度 L 定义为任意两个节点之间距离的平均值. 在有向网络中，网络的平均路径长度定义为 $L=\frac{1}{N(N-1)}\sum_{i\neq j}d_{ij}$，$N$ 为网络的节点数. 计算温度波动网络的平均路径长度，并和随机和混沌波动网络进行比较(表 2)，发现温度波动网络的平均路径长度较混沌网络要大，比随机网络要大许多. 随机波动网络中的 125 种波动模态之间随机关联的比较紧密，各种模态之间的转换不需要花太长的时间；Lorenz 系统由于其本身的复杂性，各种波动模态之间的转换表现出多样性，所以网络的平均最短路径长度要比随机网络大；而温度系统要比 Lorenz 系统更加复杂，其各种波动模态之间的转换更具多样性，所以温度网络表现出来的平均最短路径长度要比混沌网络大，且温度波动网络中某种温度波动模态向另一种模态转换要经过较长的时间，即表现出长程关联的温度波动模态相对较多.

表 2 给出了三种网络的平均路径长度、聚类系数及平均顶点度. 温度和混沌波动网络类似，均表现出较高的聚类系数，大的平均路径长度和大的平均顶点度. 温度网络中各种温度波动模态之间多表现出长程相关性，但其平均顶点度比随机网络大，这说明各种温度波动模态可能存在一系列中小规模的群簇，群簇内部节点之间的关联较好，而群簇之间的关联则较差，因此有长程关联和短程关联并重的情况.

如春夏秋冬不同的季节内,可能是不同部分波动模态起主要作用,这些波动模态就构成了一个群簇,不同的季节则对应了不同的群簇.对于混沌序列而言,也存在一定的群簇性,但群簇间的差异没有季节差异那样显著,因此平均路径长度相较温度网络要小.

表 2 三种网络的聚类系数、平均路径长度和平均顶点度

	C	L	$< k >$
温度	0.13	22.06	49.11
Lorenz	0.12	14.80	49.54
随机	0.09	3.83	40.08

4. 网络的拓扑参数——中介中心性

温度波动网络包含了许多不同模态,反映了各种模态的特定关系,每一种波动模态都受其所处网络的拓扑结构影响.一个好的网络拓扑结构的测量必须结合网络上更多的全局信息,如一个节点在网络中任意两个节点的连接路径上所起的作用等.节点的拓扑重要性等价于该节点与其他节点的连接性,节点的中介中心性(BC)则可以提供关于网络拓扑重要性的信息[30].在网络理论中,用中心性来衡量一个节点所代表的对象在网络中能力的大小,即通过网络的拓扑属性可了解该对象获得控制信息(资源)的能力.中介中心性衡量网络的中介性,可用来考察网络中各个节点在特定的网络拓扑结构中所处位置的枢纽程度[31,32].对于温度波动网络而言,了解中枢节点可以更好地认识温度变化的中间过程,为温度预测提供理论依据.

网络节点中介中心性测度 g_k 的定义:设 $c(i,j)$ 为节点对 (i,j) 之间所有最短路径的总数,路径的长度就是路径通过的各条边的权重的总和[33—35].在这些路径中通过中间节点 k 的路径数为 c_k,并记 $g_k(i,j)$ 为

$$g_k(i,j) = c_k(i,j)/c(i,j), \tag{8}$$

则节点 k 的中介中心性的测度 g_k 为所有节点对的 $g_k(i,j)$ 和,即

$$g_k = \sum_{i,j} c_k(i,j)/c(i,j) = \sum_{i,j} g_k(i,j). \tag{9}$$

计算了网络中各个节点的中心参数 BC,并把它们按数值百分比进行排序,结果如表 3.就节点的 BC 值而言,在温度波动网络中,各节点的 BC 值之间的差异是比较明显的,其中节点 *RRR*,*DDD*,*RRr*,*eee*,*dDD* 中介中心性的能力均在 2% 以上,这 5 个节点对网络的中介中心性的贡献达到了 12.93%,即 4% 的节点承担了网络 12.93% 的中介中心性功能,所以从网络的中介中心性角度来看,这意味着在温度波动网络,有些节点所代表的温度波动模态有着重要的意义,任意两个温度波动模态之间的转换经过这几个温度波动模态中转的概率较大.尤其是 *RRR* 和 *DDD* 这两个模态中介中心性的能力均在 3% 以上,充分说明了在全球增暖的背景下,极端高温与极端低温剧烈化程度.这几种温度波动模态对控制其他模态的转换以及信息传输过程具有重要的意义,它们在一定程度上可以作为各种温度波动模态之间转换的前兆,对它们的研究将有助于更好地把握温度变化的规律性.

图 6 为三种网络节点的中介中心性,在随机波动网络中,各节点的中介中心性 BC 大小都在 0.59—0.98 之间.由于 Lorenz 混沌系统本身的复杂性,混沌波动网络中节点的中介中心性也有一定差别(表 3).温度波动网络类似于混沌波动网络,表现出一定的复杂性,体现了温度系统本身的复杂性.同样,温度波动网络中节点中介中心性 BC 值的差异涨落要较 Lorenz 系统更为明显,相对于 Lorenz 系统来说,温度系统在温度波动网络的节点中介中心性上表现出更高的复杂性,节点的特性要更加强烈些,即温度波动网络中,高频出现的节点和低频出现的节点之间的中介差异性更显著,这可能有意义于天气尺度过程的温度预测.

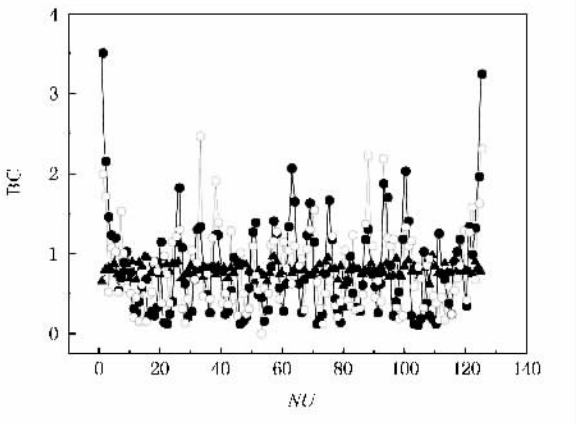

图 6 三种网络节点的中介中心性(▲,●,○分别表示温度,Lorenz 和随机网络)

5. 结果和讨论

复杂网络舍弃了小尺度上的细节作用而更关心

表 3　三种网络中节点的 BC

温度波动网络	节点	*RRR*	*DDD*	*RRr*	*eee*	*dDD*	*DDd*	*dde*	*rRR*	*Ddd*	*eDD*	*eed*	*edd*	…	*DRd*
	BC/%	3.51	3.25	2.16	2.07	2.04	1.97	1.89	1.83	1.70	1 .67	1.66	1.63	…	0.12
	等级	1	2	3	4	5	6	7	8	9	10	11	12	…	125
随机波动网络	节点	*Drd*	*Err*	*ReD*	*RDR*	*eDD*	*RdR*	*dDr*	*rdD*	*DdR*	*DRR*	*ReR*	*eDd*	…	*dRD*
	BC/%	0.98	0.97	0.96	0.95	0.95	0.95	0.94	0.93	0.91	0.91	0.90	0.90	…	0.59
	等级	1	2	3	4	5	6	7	8	9	10	11	12	…	125
混沌波动网络	节点	*Rre*	*DDD*	*dee*	*dde*	*RRR*	*ree*	*RRr*	*DDd*	*DDr*	*edD*	*Rrr*	*red*	…	*eRe*
	BC/%	2.48	2.32	2.23	2.20	2.01	1.91	1.72	1.63	1.59	1.55	1.54	1.40	…	0.00
	等级	1	2	3	4	5	6	7	8	9	10	11	12	…	125

注：“等级”行给出了网络节点按照其中介中心性 BC 大小的排列.

系统的本质行为，可以方便而又直观的追溯温度资料中数值波动的动力学行为.对温度资料做等概率粗粒化处理，各字符占有量满足幂律关系 $N\text{-}\Delta t^{\tau}$.这种简单关系反映了温度序列动力学特征规律及其普遍性；温度波动本身的无标度性；幂律指数 τ 的波动范围有限性反映了动力学规律的稳定性.在此基础上，构建了有向加权温度波动网络，实现了将温度波动模态的相互关系和相互作用等信息编码于网络的拓扑结构之中.

对随机序列构建随机波动网络、Lorenz 系统变量构建混沌波动网络，就三种网络的动力学统计量——顶点度分布、聚类系数和平均最短路径长度等进行比较分析发现：1）节点 *RRR*，*rRr*，*RRe*，…的节点度数比较大，这些节点所代表的温度波动模态在温度变化过程中起到重要的直接关联作用，且这些温度波动模态中大都带有上升或急剧上升的温度变化，体现了全球变暖的大背景.2）温度波动网络和混沌波动网络的顶点度分布均满足衰减指数分布，且指数值较接近，但随机波动网络则不服从衰减指数分布，这说明各种温度波动模态的发生既带有一定的随机性，但同时各种温度波动模态出现的次数在一个较长的时间段内是遵循系统最复杂原理.3）随机网络聚类系数随节点度的变化不大，而在温度波动网络中聚类系数的大小随节点度的变化则较大，其群聚特性可能会发生在小的集团中，也可能会发生在大的集团中.换而言之，温度变化的群发性有时会反映在小的时间尺度上，也有时会反映在大的时间尺度上.4）温度波动网络和混沌波动网络类似，表现出较高的聚类系数，大的平均路径长度和大的平均顶点度，说明各种温度波动模态可能存在一系列中小规模的群簇，群簇内部节点之间的关联较好，而群簇之间的关联则较差，因此有长程关联和短程关联并重的情况.如春夏秋冬不同的季节内，可能是不同部分波动模态起主要作用，这些波动模态就构成了一个群簇，不同的季节则对应了不同的群簇.

温度波动网络中节点的中介中心性具有明显的差异性，节点 *RRR*，*DDD*，*RRr*，*eee*，*dDD* 中介中心性的能力均在 2% 以上，而这 5 个节点对网络的中介中心性的贡献达到了 12.93%，即 4% 的节点承担了网络 12.93% 的中介中心性功能，所以从网络的中介中心性角度来看，这意味着在温度系统中，这些节点所代表的温度波动模态在控制信息的变化与传输过程中具有重要的意义.随机网络中不存在高的中介中心性节点，即不存在拓扑重要性节点，这表明温度系统在统计意义下并不是随机的，而是具有一定特性的；它与 Lorenz 系统具有一定的相似性，表明了温度系统的复杂性，但同样，温度系统在网络节点的中介中心性上表现出来的特性要比 Lorenz 系统更加复杂.

某些连续的特征波动是如此频繁的出现在温度资料中，我们很难看出它们的特性，然而当我们从网络的中介中心性角度来看待这些频繁的变化模态时，它们则是温度转变的前兆，温度变化其前奏是可以从这些变化模态身上发觉的.从这一观点来看，温度的波动并不是无关联的随机过程，识别这些具有拓扑统计重要性的节点模态对于我们理解温度变化的规律性，以及信息的传输是有积极意义的.

[1] Liu S D, Liu S K 1999 *Atmosphere dynamics* (Beijing: Peking University Press) p536 (in Chinese) [刘式达、刘式适 1999 大气动力学(北京:北京大学出版社)第 536 页]

[2] Shi N 2005 *Chin. Phys.* **14** 844

[3] Dai X G, Fu Z B, Wang P 2005 *Chin. Phys.* **14** 850

[4] Chen L X 1995 *U. S. country studies management team and department of Science and Technology for Social Development the State Science and Technology Commission of China* 112

[5] IPCC 2001 *Climate Change* 2001 (UK: Cambridge University Press) p24

[6] Damtoft J S, Lukasik J, Herfort D, Sorrentino D, Gartner E M 2008 *Cement and Concrete Research* **38** 115

[7] Tu Q P, Deng Z W, Zhou X L 2000 *Acta. Meteorologica Sinica* **58** 288 (in Chinese) [屠其璞、邓自旺 周晓兰 2000 气象学报 **58** 288]

[8] Hu Z, Song Y, Wu R 2003 *Journal of Geophysical Research* **108** 4614

[9] Qu J H, Sun A J 1991 *Journal of Applied Meteorology* **2** 423 (in Chinese) [曲建和、孙安健 1991 应用气象学报 **2** 423]

[10] Du J 2001 *Acta. Geographical Sinica* **56** 682 (in Chinese) [杜 军 2001 地理学报 **56** 682]

[11] Hao B L 1999 *Science* **51** 3(in Chinese) [郝柏林 1999 科学 **51** 3]

[12] Li P, Wang B H 2006 *Chinese Science Bulletin* **5** 624

[13] Wackerbauer R, Witt A, Atmanspacher H 1994 *Chaos, Solutions & Fractals* **4** 133

[14] Hao B L 1999 *Science* **51** 3 (in Chinese) [郝柏林 1999 科学 **51** 3]

[15] Feng G L, Dong W J, Gong Z Q, Hou W, Wan S Q, Zhi R 2006 *Nonlinear theories and methods on spatial-temporal distribution of the observational data* (Beijing: Metrological press) p227 (in Chinese) [封国林、董文杰、龚志强、侯 威、万仕全、支 蓉 2006 观测数据非线性时空分布理论和方法(气象出版社) 第 227 页]

[16] Cao H X, Zheng Z F, Zhang X L, Xie X L, Pan J H 2007 *Chinese Journal o f Geophysics* **50** 420

[17] Zhang P P, Chen K, He Y, Zhou T, Su B B, Jin Y, Chang H, Zhou Y P, Sun L C, Wang B H, He D R 2006 *Physica* A **360** 599

[18] Barabasi A L, Albert R 1999 *Science* **286** 509

[19] Li P, Wang B H 2003 *Journal of the Graduate School of the Chinese Academy of Sciences* **20** 200(in Chinese) [李 平、汪秉宏 2003 中国科学院研究生院学报 **20** 200]

[20] Barabási A L, Albert R 1999 *Science* **286** 509

[21] Barabási A L, Albert R, Jeong H 1999 *Physica* A **272** 173

[22] Newman M E J, Watts D J 1999 *Phys. Rev.* E **60** 7332

[23] Liljeros F, Edling C R, Amaral L A N 2001 *Nature* **411** 907

[24] Sen P, Dasgupta S, Chatterjee A 2003 *Phys. Rev.* E **67** 036106

[25] Barabási A L, Oltvai Z N 2004 *Nature Reviews Genetics* **5** 101

[26] Wang X F, Li X, Chen G R 2006 *The theories and application of complex network* (Beijing: Tsinghua Unversity Press) p10 (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 10 页]

[27] Bak P, Tang C, Wiesenfel K J 1990 *J. Appl. Meteorol.* **29** 1167

[28] Gutenberg B, Richer C F 1994 *Bull. Seismol. Soc. Am.* **34** 185

[29] Frette V, Christensen K, Malthe-Sorenssen A, Feder J, Jossang T 1996 *Nature* (London) **379** 49

[30] Freeman L C 1978 *Social Networks* **1** 215

[31] Freeman L C, Borgatti S P, White D R 1991 *Social Networks* **13** 141

[32] Goh K I, Oh E, Kahng B 2003 *Phys. Rev.* E **67** 017101

[33] Newman M E J 2001 *Phys. Rev.* E **64** 016132

[34] Gong Z Q, Zhou L, Zhi R, Feng G L 2008 *Acta Phys. Sin.* **57** 5351(in Chinese) [龚志强、周 磊、支 蓉、封国林 2008 物理学报 **57** 5351]

[35] Feng G L, Gong Z Q, Dong W J 2005 *Acta Phys. Sin.* **54** 5494 (in Chinese) [封国林、龚志强、董文杰 2005 物理学报 **54** 5494]

An approach to research the topology of Chinese temperature sequence based on complex network*

Zhou Lei[1)2)]　Gong Zhi-Qiang[1)3)]　Zhi Rong[1)3)]　Feng Guo-Lin[1)2)3)†]

1)(*Department of Physics*, *Yangzhou University*, *Yangzhou* 225002, *China*)

2)(*Key Laboratory of Regional Climate-Environment for Temperate East Asia*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing* 100029, *China*)

3)(*Laboratory for Climate Studies*, *National Climate Center*, *China Meteorological Administration*, *Beijing* 100081, *China*)

(Received 28 January 2008; revised manuscript received 22 April 2008)

Abstract

To analyze the dynamics of the temperature data, using homogenous partition of coarse graining process, the series of Chinese daily main temperature from 1961 to 2002 is transformed into symbolic sequences consisting of 5 characters $\{R, r, e, d, D\}$. The vertices of the temperature fluctuation network is 125 3-symbol strings (i.e., 125 fluctuation patterns in durations of 4 days), linked in the network's topology by time sequence. It contains Integrated information about interconnections and interactions between fluctuation patterns of temperature in network topology. Random fluctuant network and chaos fluctuant network by using random sequence and chaos sequence of Lorenz system were built. We calculated the dynamical statistics of the degrees and the distribution of degrees, clustering coefficient and the shortest path length, and compared the difference of these three sequences from the view point of network. The result is that. The main vertices of temperature fluctuant network generally contain the 3 characters R, r and e, and corresponding the background of global warming, the feature of temperature fluctuation is mainly ascending. On all accounts, the temperature fluctuant network is similar with chaos network in its dynamical statistics, but is markedly different from random network, which reflects the complexity of temperature fluctuation. And besides, some vertices in temperature fluctuant network have high betweenness centrality (BC), 4% of vertices bear 71.9% of betweenness centrality of networks, these vertices of importance in topological statistics are helpful to understanding the inherent law and information transmission. The vertices' BC in chaos network is similar with that of the temperature network, but vertices in random network almost have identical betweenness centrality. As a result, the patterns of temperature fluctuation corresponding to the process of weather change have similar property with chaos rather than the random fluctuation process.

Keywords: temperature fluctuant, complex network, topological structure

PACC: 9260X

* Project supported by the National Natural Science Foundation of China (Grant No. 40675044), the State Key Development Program for Basic Research (Grant No. 2006CB400503) and the Ministry of Science and Technology Project (Grant No. 2007BAC03A01).

† Corresponding author. E-mail: fenggl@cma.gov.cn