第 53 卷 第 9 期 2004 年 9 月
1000-3290/2004/53(09)/2877-05

物 理 学 报
ACTA PHYSICA SINICA

Vol.53, No.9, September, 2004

混沌伪随机序列复杂度分析的符号动力学方法*

肖方红[1)†]　阎桂荣[1)]　韩宇航[2)]

[1)] (西安交通大学建筑工程与力学学院,西安　710049)

[2)] (中国工程物理研究院,绵阳　610003)

(2003 年 12 月 17 日收到;2004 年 1 月 18 日收到修改稿)

通过将混沌伪随机序列看成一个符号序列,提出了用符号动力学的方法来分析混沌伪随机序列的复杂度.以 Logistic 映射和耦合映射格子系统产生的混沌伪随机序列为例,说明了该方法的应用,并将计算结果与近似熵 ApEn 法的计算结果作了比较.结果表明,该方法可以有效地判断出不同的混沌伪随机序列的复杂程度,而且比近似熵法更为优越.

关键词:混沌, 伪随机序列, 符号动力学, 熵

PACC:0545

1. 引　　言

近几年来,混沌序列和同步混沌在信息通信中的应用研究引起了人们的极大兴趣,混沌在通信中的应用得到了广泛而深入的研究[1—7],其中用混沌伪随机序列代替一般的伪随机码,作为扩频通信系统的扩频序列,是利用混沌实现秘密通讯的一个重要方面.用混沌伪随机序列代替一般的伪随机码,其主要优点在于混沌序列具有宽带功率谱,有良好的相关特性和类随机性,同时混沌序列作为确定性非线性系统的产物,它的产生和复制都很方便[8].

为了保证扩频通信的最大通信容量,理想的伪随机码应具有尽可能大的序列复杂度,因此,混沌伪随机序列的复杂度分析是混沌伪随机序列用于扩频通信的一个重要研究内容.扩频序列复杂度分析的一个常用方法是线性复杂度分析算法,文献[9]的研究表明该算法不能有效地判断混沌伪随机序列的复杂度,并提出利用混沌运动产生的信息量的大小来度量混沌伪随机序列的复杂度,用近似熵(approximate entropy, ApEn)作为判断复杂度大小的准则.近似熵法虽然能较有效地判断混沌伪随机序列的复杂度,但计算涉及多个参数的选取,且参数选取会影响到 ApEn 的计算精确度,带有经验性,此外,ApEn 的计算显得有些繁复.

由于混沌伪随机序列是由混沌迭代产生的序列经过量化和判决得到的,将混沌序列经过量化和判决得到的每一个取值看作一个符号,这样混沌伪随机序列实际上又是一个符号序列.本文提出用符号动力学的方法来分析混沌伪随机序列的复杂性,用符号熵作为判断序列复杂度大小的准则,并以 Logistic 映射和耦合映射格子系统迭代产生的混沌伪随机序列为例来说明方法的有效性.结果表明,该方法不但可以有效地判断混沌伪随机序列的复杂度,而且该方法具有不涉及参数的选取、计算比 ApEn 方法简单、结果比 ApEn 方法更为可靠的特点.

2. 混沌序列符号动力学与符号熵

为了识别一个混沌时间序列中的不同的动力学演化形式,一般地可以通过粗粒化的方法,将其转化为一个相应的符号序列,这样混沌时间序列其轨道演化的信息就通过符号序列达到了编码[10—13].混沌序列的粗粒化可以通过对相空间的分割而实现.

给定一个由 m 个符号组成的符号集 $\{S_0, S_1, \cdots, S_{m-1}\}$和一个 $m+1$ 个临界点组成的集合$\{C_0, C_1, \cdots, C_m\}$,可以用以下分割规则将时间序列 $\{X_1, X_2, X_3, \cdots\}$ 转换为一个符号序列

* 国家自然科学基金(批准号:10276032)与中国工程物理院联合基金 NSAF 资助的课题.

† E-mail:xiaofh.xjtu@163.com;电话:029-82660978,82673456.

$\{S(1), S(2), S(3), \cdots\}$,

$$\text{若}\ C_k < X_j \leqslant C_{k+1}, \quad \text{则}\ S(j) \equiv S_k . \tag{1}$$

将所得符号序列 $\{S(1), S(2), S(3), \cdots\}$ 分割成长度为 L 的短序列,如图 1 所示是 $L=5$ 的分割.这些短序列可以通过下式来标记和辨别[10—13],

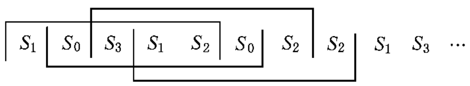

图 1 短序列分割示意图

$$l_X(L, i) = \sum_{p=1}^{L} m^{L-p} S(p+i), \tag{2}$$

其中 m 是符号集 $\{S_0, S_1, S_2, \cdots, S_{m-1}\}$ 不同符号的个数,L 表示短序列的长度,i 表示短序列沿着符号序列 $\{S(1), S(2), S(3), \cdots\}$ 从第 i 个符号开始.为了方便,通过将符号 S_k 用相应的整数 k 来代替,由(2)式可以看出每一个短序列都可以方便地用整数集 $\{0, 1, 2, \cdots, m^L - 1\}$ 中一个整数唯一地进行标记和辨别.用 P_{l_X} 表示特定短序列 l_X 出现的概率,它可以用短序列 l_X 在混沌符号序列 $\{S(1), S(2), S(3), \cdots\}$ 中出现的次数除以所有短序列的总数来计算.包含在符号序列中的信息量可以用下式引入的符号熵来定量表示[10,11]

$$E = -\frac{1}{L} \sum_{l_X} P_{l_X} \ln P_{l_X} . \tag{3}$$

一般地,由规则(1)定义的符号语言对混沌序列的信息编码能力和分割临界点 C_i 的数目及取值相关.在给定临界点 C_i 数目的情况下,可以通过使熵 E 最大化来寻找出最优临界点,而增加临界点的数目也使符号语言对混沌序列有更强的编码能力,因此,相应的熵随着临界点数目的增加而增大,当临界点数目达到一定值时,这时符号序列已包含了原混沌序列的充分多的信息量,熵不再随临界点数目的增多而增大,这时的符号语言可以称为达到了优化[10].

对于一个由混沌序列经过量化和判决而得到的伪随机序列,伪随机序列的每一个值又可以看作一个符号,因此,混沌伪随机序列可以直接看成一个符号序列.将混沌伪随机序列分割成长度为 L 的短序列,通过用(2)式统计出不同短序列出现的次数,求出各短序列在混沌伪随机序列中出现的概率,就可以用(3)式直接计算混沌伪随机序列的符号熵,根据符号熵的大小来判断序列的复杂度.

3. 复杂度分析的 ApEn 算法

为了便于分析和对比,这里简要地归纳一下复杂度分析的 ApEn 算法[9,14,15].对于一个长度为 N 的序列样本空间 $[u(1), u(2), \cdots, u(N)]$,定义 m 维向量组

$$x(1), x(2), \cdots, x(i), \cdots, x(N-m+1) \in R^m,$$

其中

$$x(i) = [u(i), u(i+1), \cdots, u(i+m-1)] \quad (1 \leqslant i \leqslant N-m+1).$$

定义以下几个中间量:m 维向量 $x(i)$ 和 $x(j)$ 的最大距离,

$$d[x(i), x(j)] = \max_{k=1,2,\cdots,m} (|u(i+k-1) - u(j+k-1)|), \tag{4}$$

满足与第 i 个 m 维向量 $x(i)$ 的最大距离小于 r 的向量数,

$$C_i^m(r) = (\text{满足}\ d[x(i), x(j)] \leqslant r\ \text{的}\ j\ \text{的个数}) / (N-m+1). \tag{5}$$

根据 $C_i^m(r)$,定义 $\phi^m(r)$,

$$\phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r), \tag{6}$$

则 ApEn 可以被定义为

$$\mathrm{ApEn}(m, r, N) = \phi^m(r) - \phi^{m+1}(r). \tag{7}$$

向量维数 m 的最大值由观察空间的长度 N 确定,当 m 越大时,ApEn 越接近序列的测度熵,距离参数 r 决定了算法的分辨率,r 越小,ApEn 的分辨率越高.对于混沌伪随机序列,可选择 $r=0$,这时(5)式可写为

$$C_i^m(r) = (\text{满足}\ x(i) = x(j)\ \text{的}\ j\ \text{的个数}) / (N-m+1). \tag{8}$$

这样根据 ApEn 的大小,就可以确定混沌序列和混沌伪随机序列复杂度的大小,ApEn 越大,复杂度越高.

4. 混沌伪随机序列复杂度分析实例

这里以 Logistic 映射和耦合映射格子系统迭代产生的伪随机序列为例,说明符号动力学用于混沌伪随机序列复杂度分析的应用和有效性.

考虑 Logistic 映射

$$x_{n+1} = 4x_n(1-x_n) \quad (0 \leqslant x_n \leqslant 1, n = 0, 1, 2, \cdots), \tag{9}$$

和耦合映射格子系统[1,9]

$$y_{n+1}(1) = (1-\varepsilon)f(y_n(1)) + \varepsilon g_n,$$
$$y_{n+1}(i) = (1-\varepsilon)f(y_n(i)) + \varepsilon f(y_n(i+1)),$$
$$i = 2,\cdots,6 \quad (n = 0,1,2,\cdots),$$
$$g_n = f(y_n(2)) \quad (0 \leqslant y_n \leqslant 1, n = 0,1,2,\cdots), \tag{10}$$

其中边缘条件满足 $y_n(7) = y_n(1)$，映射子函数 $f(x) = 4x(x-1)$，选择耦合信号 g_n 作为伪随机序列的生成函数.

在区间(0,1)上随机地选取迭代初始值进行迭代，产生混沌序列 $\{x_0, x_1, x_2, \cdots\}$ 和 $\{g_0, g_1, g_2, \cdots\}$，分别作出其相空间如图 2 所示，可以看出 $\{g_0, g_1, g_2, \cdots\}$ 比 $\{x_0, x_1, x_2, \cdots\}$ 具有更高的复杂度.

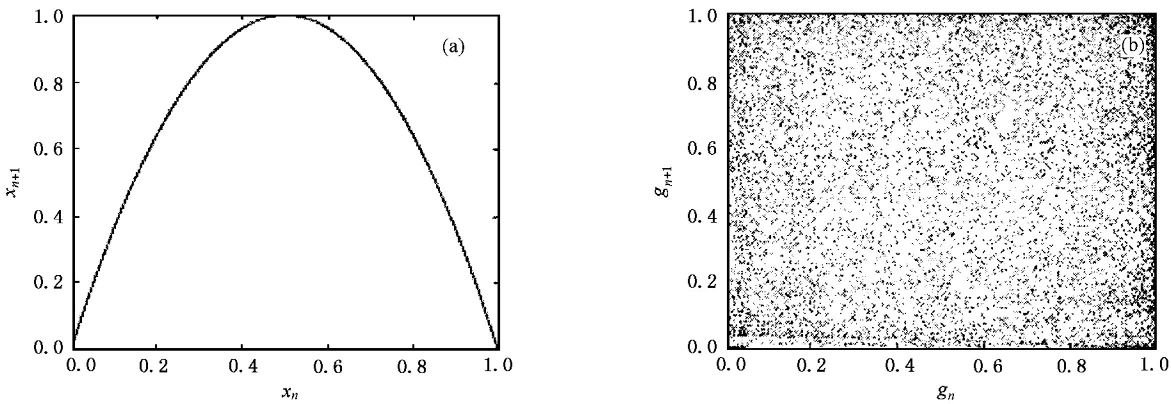

图 2　混沌序列的相空间结构　(a) Logistic 映射，(b)耦合映射格子系统

现定义判决公式[9]：

$$\sigma_c(x) = j, \tag{11}$$

若

$$\sin^2\left|\frac{j\pi}{2K}\right| < x \leqslant \sin^2\left|\frac{(j+1)\pi}{2K}\right|,$$
$$j = 0,1,2,\cdots,K-1.$$

对(9)和(10)式迭代产生的序列进行量化，可以得到 $K = 2^n$ 进制伪随机序列 $\{\sigma_c(f^n(x))\}_{n=0}^{\infty}$. 文献[9]根据 Berlekamp-Massey 线性复杂度计算法，计算了两种映射所产生的 8 进制伪随机序列的线性复杂度，如表 1 所示.

表 1　混沌伪随机序列的线性复杂度

序列长度	200	400	600	800	1000
Logistic 映射	100	199	299	400	498
耦合信号 $\varepsilon = 0.99$	100	200	300	398	499

可以看出两种映射所产生序列的线性复杂度近似相同，因此线性复杂度不能够有效地区分混沌伪随机序列的复杂度. 文献[9]的研究表明，由(9)和(10)式产生的混沌序列的近似熵 ApEn 与参数 m, r, N 的选取有关，对于两种映射所产生的伪随机序列，取 $r = 0$, $N = 10000$，本文用 ApEn 法计算了不同嵌入维数 m 下的近似熵，如表 2 所示.

表 2　8 进制混沌伪随机扩频序列的 ApEn

m	1	2	3	4	5
Logistic 映射	0.693	0.692	0.692	0.690	0.686
耦合信号 $\varepsilon = 0.99$	2.047	2.027	1.850	0.991	0.210

本文再用符号动力学方法根据(2)和(3)式计算了不同短序列长度 L 下的符号熵，如表 3 所示(序列长度 $N = 10000$).

表 3　8 进制混沌伪随机扩频序列的符号熵

L	1	2	3	4	5
Logistic 映射	2.079	1.386	1.155	1.039	0.969
耦合信号 $\varepsilon = 0.99$	2.051	2.049	2.042	1.994	1.793

从表 2 和表 3 可以看出，用符号动力学方法取 $L \geqslant 2$ 时，即可有效地判断伪随机序列的复杂度，当 $L = 1$ 时，不同短序列只有 8 个，即 8 个符号，短序列分布接近均匀分布，显然短序列太少，不能有效地表征混沌伪随机序列中的信息变化形式. 从表 3 可以看出，当 $L \geqslant 2$ 时，由耦合映射格子系统迭代产生的混沌伪随机序列的复杂度比 Logistic 映射产生的混沌伪随机序列的复杂度大，在不同短序列长度 L 下的情况相一致，而近似熵 ApEn 法在嵌入维数较大

时,结果不可靠.另外,本文的计算表明,当短序列长度 L 不是很大时,符号熵的计算时间比 ApEn 的计算时间大为缩短,因此符号动力学方法是分析混沌伪随机序列复杂度的一种简单而且很有效的方法.

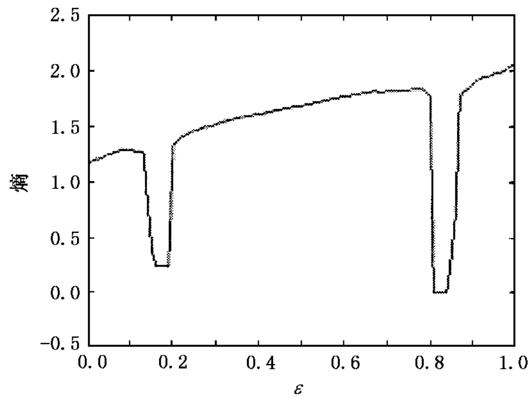

图3 不同耦合系数 ε 下的伪随机序列的复杂度

下面再应用符号动力学的方法对耦合映射格子系统(10)在不同耦合系数 ε 下所产生的混沌伪随机序列的符号熵进行计算.在区间[0,1]中均匀地选取 101 个点作为 ε 的值,在各 ε 值下,在区间(0,1)中随机地选取各变量的初始值进行迭代,去除前面过渡过程中的点,取序列长度 $N=10000$,用(11)式进行判决,得到长度 $N=10000$ 混沌伪随机序列,计算伪随机序列的符号熵,短序列长度 $L=3$,其结果如图 3 所示.从图 3 可以看到,由耦合映射格子系统产生的伪随机序列的复杂度总体上随着耦合系数 ε 的增大而增大,其中有两个特殊的 ε 区域,在这两个区域内,序列的复杂度大为减小.计算表明当 $\varepsilon\in[0.16,0.19]$ 时,这时系统输出的是一个周期为 2 序列,当 $\varepsilon\in[0.81,0.84]$ 时,系统输出的是一个常数序列,这表明可以通过调节耦合系数 ε 来控制耦合映射格子系统中的混沌运动.

5. 结 论

本文提出了用符号动力学的方法来分析混沌伪随机序列的复杂度,根据符号熵来判断混沌伪随机序列复杂度大小的准则,并以 Logistic 映射和耦合映射格子系统为例,分析了这两个混沌系统产生的伪随机序列的复杂度.通过与用近似熵 ApEn 法计算出的结果作比较,说明用符号动力学方法来分析混沌伪随机序列的复杂度是一种简单比近似熵法更为优越的方法.通过对耦合映射格子系统在不同耦合系数下产生的混沌伪随机序列的复杂度的分析,结果表明该方法可以有效地判断出不同的混沌伪随机序列的复杂程度.

[1] Xiao J H, Hu G and Qu Z L 1996 *Phys. Rev. Lett.* **77** 4162

[2] Li J F and Li N 2002 *Chin. Phys.* **11** 1124

[3] Wu L and Zhu S Q 2003 *Chin. Phys.* **12** 300

[4] Mu J, Tao C and Du G H 2003 *Chin. Phys.* **12** 381

[5] Zhang J S and Xiao X C 2001 *Acta Phys. Sin.* **50** 2121 (in Chinese)[张家树、肖先赐 2001 物理学报 **50** 2121]

[6] Liu J B, Ye C F and Zhang S J 2000 *Acta Phys. Sin.* **49** 20 (in Chinese)[刘剑波、叶春飞、张树京 2000 物理学报 **49** 20]

[7] Kuang J Y, Deng K and Huang R H 2001 *Acta Phys. Sin.* **50** 1856 (in Chinese)[匡锦瑜、邓 昆、黄荣怀 2001 物理学报 **50** 1856]

[8] Hu G, Xiao J H and Zheng Z G 2000 *Chaos control* (Shanghai: Shanghai Science and Technology Education Publishing House) p181 (in Chinese)[胡 岗、萧井华、郑志刚 2000 混沌控制(上海:上海科技教育出版社)第 181 页]

[9] Cai J P, Li Z and Song W T 2003 *Acta Phys. Sin.* **52** 1871 (in Chinese)[蔡觉平、李 赞、宋文涛 2003 物理学报 **52** 1871]

[10] Lehrman M and Rechester A B 1997 *Phys. Rev. Lett.* **78** 54

[11] Azad R K, Rao J S and Ramaswamy R 2002 *Chao, Solitons and Fractals* **14** 633

[12] Rechester A B and White R B 1991 *Phys. Lett.* A **156** 419

[13] Rechester A B and White R B 1991 *Phys. Lett.* A **158** 51

[14] Jorge A et al 2000 Physica A **276** 425

[15] Pincus S and Kalman R E 1997 *Pro. Nat. Acad. Sci.* USA **94** 35130

A symbolic dynamics approach for the complexity analysis of chaotic pseudo-random sequences *

Xiao Fang-Hong[1)]　Yan Gui-Rong[2)]　Han Yu-Hang[2)]

[1)] (*School of Civil Engineering and Mechanics*, *Xi 'an Jiaotong University*, *Xi 'an* 710049, *China*)

[2)] (*China Academy of Engineering Physics*, *Mianyang* 610003, *China*)

(Received 17 December 2003; revised manuscript received 18 January 2004)

Abstract

By considering a chaotic pseudo-random sequence as a symbolic sequence, we present a symbolic dynamics approach for the complexity analysis of chaotic pseudo-random sequences. The method is applied to the cases of Logistic map and one-way coupled map lattice to demonstrate how it works, and a comparison is made between it and the approximate entropy method. The results show that this method is applicable to distinguish the complexities of different chaotic pseudo-random sequences, and it is superior to the approximate entropy method.

Keywords: chaos, pseudo-random sequence, symbolic dynamics, entropy

PACC: 0545

* Project supported by the National Natural Science Foundation of China (Grant No 10276032) and the Science Foundation of China Academy of Engineering Physics NSAF.