

混沌伪随机序列的复杂度的稳定性研究^{*}

罗松江[†] 丘水生 骆开庆

(华南理工大学电子与信息学院 广州 510640)

(2009 年 1 月 19 日收到, 2009 年 2 月 17 日收到修改稿)

增强统计复杂度能反映混沌伪随机序列的随机本质, 在此基础上提出了 k 错增强统计复杂度的定义, 用来衡量混沌伪随机序列复杂度的稳定性, 并证明了其两个基本特性. 以 Logistic, Henon, Cubic, Chebyshev 和 Tent 映射产生的混沌伪随机序列为例, 说明了该方法的应用. 仿真结果表明, 该方法能区分不同混沌伪随机序列的稳定性, 是一种衡量混沌序列稳定性的有效方法.

关键词: 稳定性, k 错增强统计复杂度, 混沌, 伪随机序列

PACC: 0545

1. 引言

混沌系统由于类随机性、对初值敏感性和不可预测性的特点, 已引起人们的广泛关注. 混沌伪随机序列是由混沌系统迭代产生的序列经量化和判决后得到的, 其主要优点是具有宽带功率谱、类噪声、良好的相关特性、对初始条件和控制参数的敏感性, 同时便于产生和复制, 因而可取代传统的伪随机序列, 应用于密码学、随机模拟、扩频通信等领域. 各种不同的混沌映射, 例如 Logistic, Henon, Cubic, Chebyshev 等已被广泛用于产生混沌伪随机序列^[1-3].

混沌伪随机序列的统计特性是决定混沌保密通信安全性的关键因素之一, 因此衡量混沌伪随机序列的性能好坏是其获得广泛应用的前提. 最近有许多文献致力于这方面的研究, 如文献[2]中研究了用于产生流密码的几种常见混沌伪随机序列的密码学特性, 比较了它们的伪随机特性; 文献[3]研究了利用 Logistic 映射作为伪随机数产生器的可能性, 对由 Logistic 映射获得的序列做了一些统计测试. 以上文献分析的都是混沌伪随机序列的统计特性, 没有涉及到其稳定性的评价.

从密码学的观点来看, 好的伪随机序列不但应有很好的统计分布, 长周期和大的线性复杂度等特性, 而且应有很好的稳定性. 文献[4]中深入讨论了

流密码稳定性的重要性, 定义了重量复杂度, 即 k 错线性复杂度来量度二进制序列的稳定性. 由于线性复杂度表征的是能够产生该序列的最短线性反馈寄存器的长度, 而混沌伪随机序列是通过混沌系统的迭代, 并由其演化轨迹得到的, 文献[5]中证明了用线性复杂度来衡量混沌伪随机序列的复杂度是不合适的, 并提出用近似熵描述混沌伪随机序列的复杂度. 同样, 文献[6]表明 k 错线性复杂度并不能有效区分不同混沌伪随机序列的稳定性, 提出用 k 错近似熵衡量混沌伪随机序列的稳定性.

文献[7]指出熵是在信息论的框架内衡量伪随机数的统计特性, 然而测度熵不能保证一个时间序列的随机本质; 文献[8]指出增强统计复杂度测度不仅能量化序列的随机性, 且能反映序列的相关结构, 保证序列的随机本质. 基于增强统计复杂度在度量序列随机性方面的优越性, 本文首先用它对几种混沌伪随机序列进行了分析, 在此基础上提出了 k 错增强统计复杂度的概念, 用以测量不同混沌伪随机序列稳定性, 并以几种常见混沌映射产生的混沌伪随机序列为例, 说明了该方法的应用. 仿真结果表明, 该方法能区分不同混沌伪随机序列的稳定性, 是一种衡量混沌序列稳定性的有效方法.

2. 增强统计复杂度

在文献[9]中, López-Ruiz, Mancini 和 Calbet

^{*} 国家自然科学基金(批准号: 60372004)资助的课题.

[†] E-mail: luosongjiang@126.com

(LMC)认为仅熵的概念不能反映动态系统产生的时间序列的内部结构,然后基于“不平衡”的概念,提出了用统计复杂度测度作为时间序列内物理结构程度的量化器.对给定的系统状态,统计复杂度测度是测度熵(H)乘以到平衡状态(Q)距离的积.文献[7]中, Q 用 Wootters 统计距离建立, H 是归一化 Shannon 熵.文献[8]指出统计复杂度测度既不是内权量,也不是外延量,并提出了一种考虑内权量的统计复杂度测度,称为增强统计复杂度测度.增强统计复杂度测度能描述序列的相关结构,反映序列的随机本质,同时更方便应用于序列随机性的量度.

增强统计复杂度测度 $C_J[P]$ 是一个和动态系统产生的时间序列的概率分布 P 相联系的物理量,可通过下式计算:

$$C_J[P] = Q_J[P, P_e] \cdot H_J[P], \tag{1}$$

其中

$$H_J[P] = S[P] S_{\max} = \left[-\sum_{j=1}^N p_j \ln(p_j)\right] / S_{\max}, \tag{2}$$

且有 $S_{\max} = S[P_e] = \ln M (0 \leq H_S \leq 1)$, M 代表系统在相空间中总的状态数目; P_e 表示均匀分布,即 $P_e = \{1/M, \dots, 1/M\}$, S 为 Shannon 熵. Q_J 是根据 Jensen-Shannon 分歧定义的不均衡,有

$$Q_J[P, P_e] = Q_0 \cdot \{S[(P + P_e)/2] - S[P]/2 - S[P_e]/2\}, \tag{3}$$

Q_0 是归一化常数,

$$Q_0 = -2 \left\{ \left(\frac{M+1}{M} \right) \ln(M+1) - 2 \ln(2M) + \ln M \right\}^{-1}. \tag{4}$$

可见, $0 \leq Q_J \leq 1$, 且不平衡 Q_J 为能反映序列结构的内权量,同样 $C_J[P]$ 也能反映序列的相关结构. $C_J[P]$ 值越小,序列随机性越强,反之则序列规律性更强.对完全随机序列(3)式中 $P = P_e$, 故 $Q_J[P, P_e] = 0$, $H_J[P] = 1$, 从而 $C_J[P] = 0$, 即序列没有任何相关结构,序列中各个元素彼此无关.

下面介绍二进制伪随机序列概率分布 P 的计算.对于混沌伪随机序列 $\{x_t, t = 1, \dots, T\}$ 和嵌入维 $d > 1$, 对每一时刻 t , 其 d 阶顺次模式是由时刻 $t, t-1, \dots, t-(d-1)$ 的值组成的 d 维向量

$$(t) \rightarrow (x_{t-(d-1)}, x_{t-(d-2)}, \dots, x_{t-1}, x_t). \tag{5}$$

显然, d 越大, 向量提供的以前的信息越多.对 d 阶顺次模式, 共有 $M = 2^d$ 种可能排列.以 $d = 2$

为例,对 $\{x_{t-1}, x_t\}$, 有 $\{0, 0\}$, $\{0, 1\}$, $\{1, 0\}$ 和 $\{1, 1\}$ 四种排列.对每一种 d 阶排列 π , 统计其出现次数, 概率分布 $P = \{p(\pi)\}$ 定义为

$$p(\pi) = \frac{\#\{t | t \leq T-d+1(t), \text{有类型 } \pi\}}{T-d+1}, \tag{6}$$

符号 $\#$ 代表“数目”.增强统计复杂度通过这种“排列”概率分布来计算.

对 Logistic, Henon, Cubic, Chebyshev 和 Tent 映射, 方程和参数见表 1. 取五个混沌映射的初值 $x_0 = 0.01$, 为消除过渡态的影响, 截取 10000 次迭代后的 100 个迭代值经量化和判决后得到的二进制伪随机序列.设 $I = [a, b]$ 为混沌序列 $\{x_t\}$ 的值域, 量化函数 $T_n(\cdot)$ 定义为

$$\{s_n(t)\}_{t=1}^\infty = T_n(x_t) = \begin{cases} 0, & x \in U_{d=0}^{2^{n-1}-1} I_{2d}^n, \\ 1, & x \in U_{d=0}^{2^{n-1}-1} I_{2d+1}^n, \end{cases} \tag{7}$$

式中, $n > 0$ 为任意正整数; $I_0^n, I_1^n, \dots, I_{2^n-1}^n$ 代表 $[a, b]$ 的 2^n 个连续的等分区间.根据前述算法计算出相应二进制混沌伪随机序列的增强统计复杂度, 结果见表 2.

表 1 文中的五个混沌映射

混沌映射	映射方程	参数
Logistic	$x_{n+1} = \mu x_n(1 - x_n)$	$\mu = 4$
Henon	$x_n = 1 + bx_{n-2} - 1.4x_{n-1}^2$	$b = 0.3$
Cubic	$x_{n+1} = \lambda x_n(1 - x_n^2)$	$\lambda = 2.59$
Chebyshev	$x_{n+1} = \cos(k \arccos(x_n))$	$k = 4$
Tent	$x_{n+1} = \begin{cases} x_n/a, & x_n \leq a \\ (1-x_n)(1-a), & x_n \geq a \end{cases}$	$a = 0.4$

表 2 混沌伪随机序列的增强统计复杂度

d	2	3	4
Logistic	0.003461	0.0063973	0.009806
Henon	0.001741	0.0043608	0.007186
Cubic	0.011664	0.0138269	0.015201
Chebyshev	0.000874	0.0024841	0.005932
Tent	0.010745	0.0117719	0.013791

3. k 错增强统计复杂度

序列复杂度的稳定性一直是密码学者们关注的热点, 序列增强统计复杂度的稳定性与序列的可测

性也是密切相关的 ;度量有限混沌序列的伪随机性和增强统计复杂度的稳定性 ,引入如下定义 :

定义 设 s^N, t^N 是有限域 $GF(2)$ 上的 N 长二进制序列. s^N 的重量增强统计复杂度 (weight intensive statistical complexity) 定义为

$$ISC_k(d \text{ } s^N) = \max_{w_H(t^N)=k} ISC(d \text{ } s^N + t^N), \quad (8)$$

其中 $w_H(t^N)$ 表示 t^N 的 Hamming 重量 ,即 t^N 的非零分量个数 , $ISC(d \text{ } s^N)$ 表示序列 s^N 向量维数为 d 的增强统计复杂度 .

下面描述重量增强统计复杂度的几何意义 .考虑空间 $GF(2)^N$,以 Hamming 距离 l_H 作为它的距离 ,并记 $\mathcal{S}(s^N, k) = \{t^N : l_H(s^N, t^N) = k\}$.由定义有

$$ISC_k(d \text{ } s^N) = \max_{t^N \in \mathcal{S}(s^N, k)} ISC(d \text{ } t^N).$$

表明重量增强统计复杂度 $ISC_k(d \text{ } s^N)$ 是空间 $(GF(2)^N, l_H)$ 中球心在 s^N 以 k 为半径的所有 N 长序列当向量维数为 d 时的增强统计复杂度的上确界 ,所以它也可被称为球面增强统计复杂度 (spherical intensive statistical complexity) .也可这样来理解 :重量增强统计复杂度是原序列改变 k 位后得到的序列的增强统计复杂度的最大值 .所以 ,它也可被称为 k 错增强统计复杂度 (k -error intensive statistical complexity) .

设 $s^N, t^N \in GF(2)^N$, k 错增强统计复杂度有下述两个性质 :

性质 1

$$ISC_N(d \text{ } s^N) = ISC_0(d \text{ } s^N) = ISC(d \text{ } s^N).$$

由增强统计复杂度和重量增强统计复杂度的定义 ,性质 1 是显然的 .

性质 2

$$\begin{aligned} ISC_{N-w_H(s^N)}(d \text{ } s^N) &= ISC(\underbrace{[1, 1, \dots, 1]}_{N \uparrow}) \\ &= ISC(\underbrace{[0, 0, \dots, 0]}_{N \uparrow}). \end{aligned}$$

证明 由 (8) 式 ,

$$\begin{aligned} ISC_{N-w_H(s^N)}(d \text{ } s^N) &= \max_{w_H(t^N)=N-w_H(s^N)} ISC(s^N + t^N) \\ &= \max_{w_H(t^N)=s^N \text{ 中零分量的个数}} ISC(s^N + t^N). \end{aligned}$$

t^N 中非零分量的个数 = s^N 中零分量的个数 ,则 $s^N + t^N$ 后可能得到全为 0 或全为 1 的长度为 N 的序列 ,故性质 2 成立 .

取序列长度 $N = 100$,模式维数 $d \leq 3$,采用 k 错增强统计复杂度对表 1 混沌映射产生的混沌伪随机序列进行增强统计复杂度稳定性测试 ,测试结果见表 3 .

表 3 混沌伪随机序列的 k 错增强统计复杂度

	ISC _k (2)			ISC _k (3)			ISC _k (4)		
	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$	$k = 1$	$k = 2$	$k = 3$
Logistic	0.006587	0.010504	0.032256	0.025717	0.051946	0.100053	0.071289	0.090509	0.189451
Henon	0.004412	0.007188	0.017827	0.021237	0.031572	0.057606	0.049511	0.088675	0.184464
Cubic	0.052370	0.094361	0.156647	0.103148	0.141172	0.267412	0.125104	0.204149	0.344303
Chebyshev	0.001285	0.002094	0.004649	0.009738	0.013464	0.031151	0.027169	0.037608	0.087616
Tent	0.018227	0.030870	0.080265	0.035197	0.047441	0.086406	0.074885	0.114604	0.223552

由表 3 可看出 ,当 k 不变时 ,增强统计复杂度随着 d 的增大而增大 ;当 d 不变时 , k 错增强统计复杂度随着 k 的增大而增大 .

4. 基于 k 错增强统计复杂度的混沌伪随机序列复杂度稳定性

被测伪随机序列的稳定性优劣通过原序列增强统计复杂度与 k 错增强统计复杂度之间的偏离程度来衡量 ,即定义

$$R_{ISC} = \frac{ISC_k(d \text{ } s^N) - ISC(d \text{ } s^N)}{ISC(d \text{ } s^N)}. \quad (9)$$

根据 R_{ISC} 的大小量度增强统计复杂度的稳定性 .由表 2 和表 3 结果计算 R_{ISC} 值 ,作出增强统计复杂度的变化程度随序列改变比特个数的关系如图 1 .

从图 1 可看出 ,当 $d = 2$ 时 ,随着 k 的增加 ,Cubic 映射伪随机序列的 ISC 变化最快 ,Chebyshev 伪随机序列变化最慢 ,说明其增强统计复杂度最稳定 ;当 $d = 3$ 时 ,Cubic 映射的伪随机序列 ISC 仍然变化最快 ,Tent 伪随机序列最稳定 ; $d = 4$ 时 ,Chebyshev 伪随机序列的增强统计复杂度又是最稳定的 .

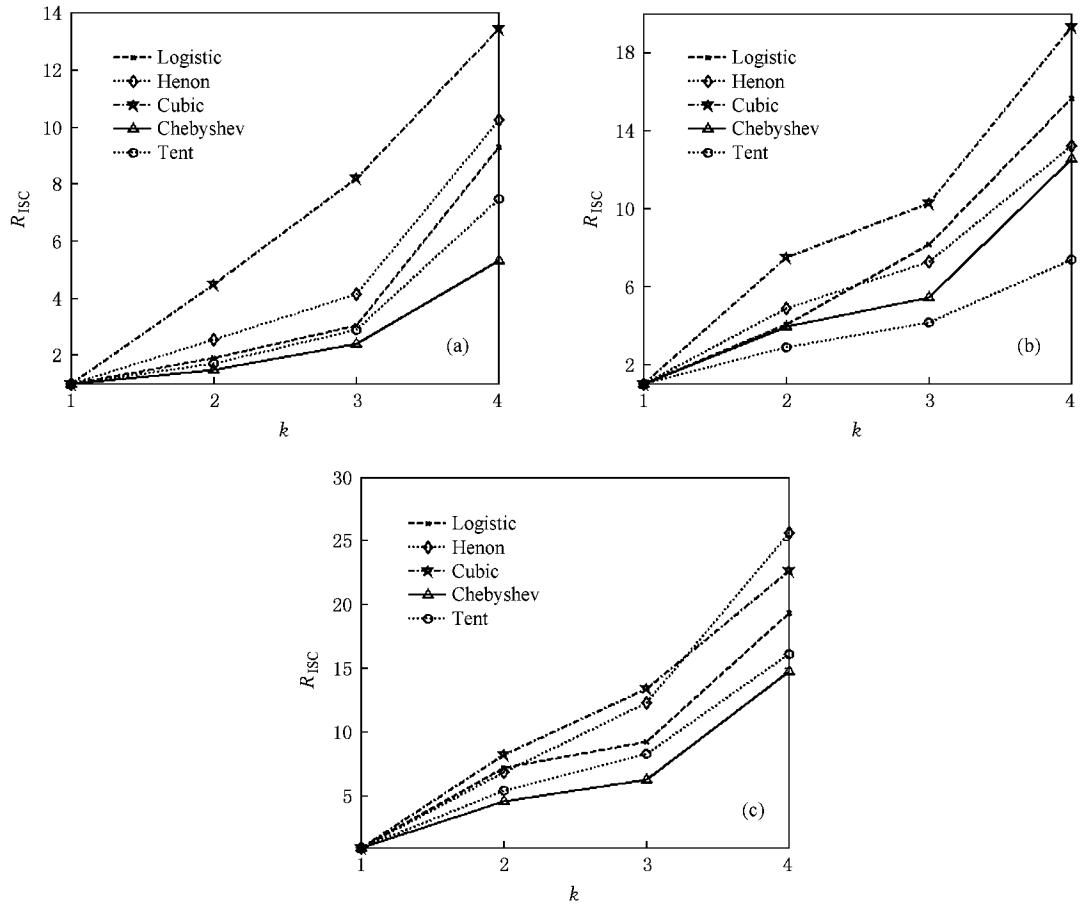


图 1 混沌伪随机序列的 ISC 随 k 的变化关系图 (a) $d=2$ (b) $d=3$ (c) $d=4$

5. 结 论

由于增强统计复杂度不仅能量化序列的随机性,且能呈现序列的相关结构,反映序列的随机本质,相对于测度熵,更适合用来度量混沌伪随机序列的随机性.本文首先用它对几种混沌伪随机序列进

行了分析,在此基础上提出了 k 错增强统计复杂度的概念,并证明了它的两个特性,用以测量不同混沌伪随机序列稳定性.以 Logistic, Henon, Cubic, Chebyshev 和 Tent 混沌映射产生的伪随机序列为例,说明了该方法的应用.仿真结果表明, k 错增强统计复杂度能区分不同混沌伪随机序列的稳定性,是一种衡量混沌伪随机序列稳定性的有效方法.

[1] Liao X F , Li X M , Peng J *et al* 2004 *Int. J. Commun. Syst.* **17** 437

[2] Khan M K , Zhang J S 2006 *Proc. 2006 IEEE Int. Conference on Engineering of Intelligent Systems* 1

[3] Phatak S C , Rao S S 1995 *Phys. Rev. E* **5** 3670

[4] Ding C S , Xiao G Z , Shan W J 1991 *The Stability Theory of Stream Ciphers* (Berlin ; New York : Springer-Verlag)

[5] Cai J P , Li Z , 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]

[6] Xiang F , Qiu S S 2008 *IEEE Commun. Lett.* **12** 337

[7] Martin M T , Plastino A , Rosso O A 2003 *Phys. Lett. A* **311** 126

[8] Larrondo H A , González C M , Martin M T , *et al.* 2005 *Phys. A* **356** 133

[9] López R R , Mancini H L , Calbet X 1995 *Phys. Lett. A* **209** 321

[10] Sheng L Y , Xiao Y Y , Sheng Z 2008 *Acta Phys. Sin.* **57** 4007 (in Chinese) [盛利元、肖燕予、盛 2008 物理学报 **57** 4007]

[11] Xiao F H , Yan G R , Han Y H 2004 *Acta Phys. Sin.* **53** 2877 (in Chinese) [肖方红、阎桂荣、韩宇航 2004 物理学报 **53** 2877]

Research on the stability of complexity of chaos-based pseudorandom sequence^{*}

Luo Song-Jiang[†] Qiu Shui-Sheng Luo Kai-Qing

(School of Electronic and Information Engineering , South China University of Technology , Guangzhou 510640 , China)

(Received 19 January 2009 ; revised manuscript received 17 February 2009)

Abstract

Intensive statistical complexity can reflect the random nature of chaos-based pseudorandom sequence. Based on this property, the definition of k -error intensive statistical complexity is presented and two basic properties of it are proved in this paper, which can be used to measure the stability of complexity of chaos-based pseudorandom sequences. Based on chaos-based pseudorandom sequences produced via Logistic, Henon, Cubic, Chebyshev and Tent maps, an example is given to demonstrate how it works. Simulation results indicate that the approach is effective, it can distinguish the stability of diverse chaos-based pseudorandom sequences, and is an effective way for evaluating the stability of chaos-based sequences.

Keywords : stability, k -error intensive statistical complexity, chaos, pseudorandom sequence

PACC : 0545

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 60372004).

[†] E-mail : luosongjiang@126.com