混沌序列产生方法及其在图象加密中的应用

孙秀花, 戴跃伟, 王执铨

(南京理工大学自动化系, 210094, 南京)

[摘要]　介绍了一个混沌二值序列产生和实验的软件平台, 它利用三种混沌动力学模型, 即一阶时延模型、Lorenz 模型、Logistic 模型产生加密序列, 并通过大量实验对产生的序列的伪随机性、敏感性、重复性进行了测试. 测试表明混沌伪随机序列具有良好的加密性能. 本系统还可以提供对多种对象的加/解密实验和分析. 同时也给出了对. bmp 格式图像加/解密的实验结果.

[关键词]　混沌序列, 性能测试, 序列密码, 位图图像

[中图分类号] TP273^{+}. 5; TP311. 56,　[文献标识码] B,　[文章编号] 1672- 1292- (2004) 01- 0056- 04

0　引言

在信息化社会中, 信息(数据) 安全问题已成为与人们生存和发展休戚相关的重要问题.

20 世纪 60 年代人们发现了一种特殊的自然现象——混沌(英文为 chaos). 混沌是确定性系统的伪随机性. 混沌系统具有伪随机性, 它对系统的参数和初始条件极端敏感. 同时, 它又具有确定性, 其输出值由非线性系统的方程、参数和初始条件完全决定. 只要系统参数及初始条件相同, 就可以重构混沌信号. 以上这些特征可被用于保密通信, 混沌学和密码术相结合, 形成了所谓的"混沌密码体制"[1]. 分析表明[2,3]: 采用混沌序列不比两种最常用的序列 m- 序列或 Gold- 序列差.

为了利用混沌动力学模型产生混沌二值序列, 并验证其伪随机性、敏感性、重复性, 我们开发了实验的软件平台. 该平台还可以将产生的混沌序列用在序列密码体制中, 实现对多种对象的加/解密, 本文给出了对. bmp 格式图像的加/解密结果.

1　混沌序列的生成

我们在 windows 操作系统下, 利用 C+ + Builder 6. 0 开发环境编制了相应的实验程序. 图 1 是我们设计的实验系统的总体结构框图.

1.1　三种典型混沌动力学方程

本实验用到的三种典型混沌动力学模型的方程如下:

(1) 虫口模型 ——Logistic 映射:

$$x_{n+1} = f(x_n) = ux_n(1 - x_n) \qquad (1)$$

这是最简单的非线性函数, 当 3. 5699456… < $u \leqslant 4$ 时, Logistic 映射进入混沌(chaos) 区域.

(2) 大气对流模型 ——Lorenz 方程

$$\begin{cases} dx/dt = -\sigma(x - y) \\ dy/dt = -xz + rx - y \\ dz/dt = xy - bz \end{cases} \qquad (2)$$

这是一个三维的自治系统, 反映了大气的对流运动. 1963 年 Lorenz 在标准情形(即参数 $\sigma = 10$, $b = 8/3$, $r = 28$) 下, 发现了奇怪吸引子.

(3) 一阶时延模型: $\dot{x} = -ax + b\sin(wx(t - \tau))$ (3)

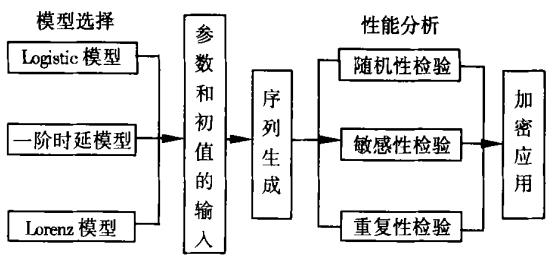


图 1　系统框图

1.2　混沌伪随机二值序列产生过程

以 Lorenz 方程(如(2) 式所示) 描述的三维混沌连续系统为例(利用另外两种模型产生混沌二值序列的步骤类似):

收稿日期: 2003- 08- 29.

基金项目: 江苏省自然科学基金[1. 混沌控制理论及其用于多媒体信息安全的研究(BK2001054); 2. 信息隐藏技术的信息论模型研究(BK2002101)].

作者简介: 孙秀花, 女, 1980- , 硕士研究生, 主要从事信息安全技术等方面的学习和研究. E- mail: sxh1001@tom. com.

通讯联系人: 戴跃伟, 1962- , 工学博士, 教授, 主要从事信息安全等方面的研究.

(1) 设定(2) 式中的参数和初始条件(包括 n、σ、r、b、x_0、y_0、z_0、h 等); n 为序列长度; n 可为任意正整数; h 为步长.

(2) 用数值算法(如四阶龙格-库塔算法) 得到上述系统的离散数值解: 由于该映射是用常微分方程组描述的, 我们可以得到 3 个离散序列: $x_1, x_2, x_3, \cdots x_n$; $y_1, y_2, y_3, \cdots y_n$; $z_1, z_2, z_3, \cdots z_n$. 我们可以利用其中的某一个序列, 例如 y 序列.

(3) 把该离散数值解 y 序列转化为混沌二值序列, 定义如下:

$$C_i = \begin{cases} 1 & y_i > u \\ 0 & y_i \leqslant u \end{cases} \quad (i = 1, 2, 3, \cdots n),\ u \text{ 为量化阈值} \tag{4}$$

u 的选取对该序列的性能有着极为重要的影响. 实验中, 我们取 $u = \frac{1}{n}\sum_{i=1}^{n} y_i$.

2 混沌伪随机二值序列性能分析

2.1 混沌二值序列的随机性能分析方法

根据 Golomb 的 3 个随机性公设[4], 为确保伪随机序列在其周期内的任一子序列尽可能具有更好的随机性能, 我们对混沌序列的随机性能进行以下几方面测试[5,6]:

①频数检验　用以确保大致有等量的 0和 1, 计算

$$\chi^2 = \frac{(n_0 - n_1)^2}{n} \tag{5}$$

其中: n_0 为 0 的个数; n_1 为 1 的个数. 与 1 自由度的 χ^2 分布比较, 对应 5% 的显著性水平, χ^2 的值为 3.84, 只要得到的值不大于 3.84, 则序列通过检验.

②序列检验　用以判定转移概率是否合理. 已经证明:

$$\chi^2 = \frac{4}{n-1}\sum_{i=0}^{1}\sum_{j=0}^{1}(n_{ij})^2 - \frac{2}{n}\sum_{i=0}^{1}[(n_i)^2 + 1] \tag{6}$$

其中: n_{ij} 代表"ij" 的个数, 对 2 自由度的 χ^2 分布, 对应 5% 的显著性水平, χ^2 的值为 5.99. 只要得到的值不大于 5.99, 则序列通过检验.

③自相关检验　设 n 比特序列为 $a_1, a_2, \cdots a_n$, 则自相关为:

$$A(d) = \sum_{i=1}^{n} a_i a_{i+d},\ 0 \leqslant d \leqslant n-1 \tag{7}$$

为了考察并演示混沌二值序列作为加密序列的安全性能, 根据前述分析方法, 我们分别对这 3 种模型进行了仿真实验, 就其随机性、敏感性、密钥的重复性进行检验, 并以图表形式给出了实验结果. 为得到性能较理想的随机序列, 对 Logistic 映射和一阶时延模型我们每 150 点取一次数值解(一阶时延模型步长 h 取 0.3), 构成混沌实值序列, 再进行量化得到混沌二值序列; Lorenz 映射是三维的, 我们对 y 进行取值, 每 900 点取一次数值解(步长 h 取 0.005), 构成混沌实值序列, 再量化得到二值序列.

除下列表格中所列举的参数取值之外, 我们还就参数的取值范围做了大量实验, 用本算法产生的混沌序列, Logistic 映射中 x_0 在 0 与 1 之间的取值; 一阶时延模型各参数和初始值均有较广的取值范围, 其中 b、w、x_{t_0} 甚至可以取负数; Lorenz 映射中 r 也有较大的取值范围. 在上述参数范围内实验, 序列性能均较好.

2.2 随机性检验结果

Lorenz 映射随机性分析结果如表 1 所示. 其它两种模型也具有比较理想的测试结果, 限于篇幅, 不在此列举.

实验表明, 这 3 种混沌动力学模型均通过序列检验和频数检验, 且有较好的游程特性.

图 2 是一阶时延模型在 $a = 3.00$, $b = 8.00$, $w = 5.00$, $\tau = 8.75$, $x_{t_0} = 0.5$, $t_0 = 0$, $h = 0.3$ 时, 取序列长度 $n = 100\,000$ 得到的序列自相关图. 显然, 满足 Golomb 第三个随机性公设.

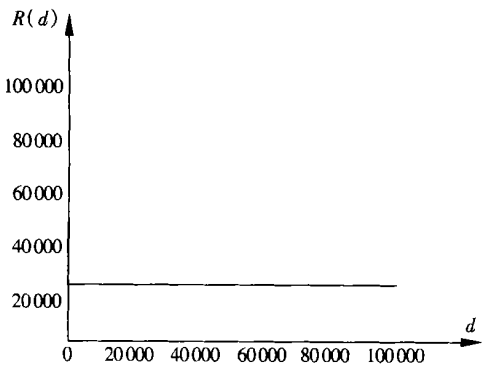


图 2　序列自相关图

表1　Lorenz映射随机性分析结果

参数及初值分别为: $\sigma = 10.00$, $r = 28.00$, $b = 2.67$, $x_{t_0} = 2.0$, $y_{t_0} = 12.0$, $z_{t_0} = 7.03$					
序列长度 n	10 000	20 000	50 000	80 000	100 000
0的个数 n_0	4 968	9 965	24 929	39 916	49 963
1的个数 n_1	5 032	10 035	25 071	40 084	50 037
00的个数 n_{00}	2 482	4 983	12 526	20 039	25 122
01的个数 n_{01}	2 486	4 982	12 403	19 876	24 841
10的个数 n_{10}	2 485	4 981	12 402	19 876	24 840
11的个数 n_{11}	2 546	5 053	12 668	20 208	25 196
频数检验 χ_f^2	0.409 600	0.245 000	0.403 280	0.352 800	0.054 760 0
序列检验 χ_s^2	−0.265 586	−0.488 612	0.711 715	1.752 94	2.253 46
游程总数	4 972	9 964	24 806	39 573	49 682
1游程数	2 489	5 020	12 364	19 746	24 678
2游程数	1 202	2 383	6 070	9 932	12 414
3游程数	639	1 280	3 225	5 052	6 280
4游程数	331	642	1 556	2 495	3 134
5游程数	142	311	775	1 243	1 577
6游程数	85	169	414	646	802
7游程数	45	89	195	309	390
8游程数	24	38	99	155	189
9游程数	12	20	55	80	104
10以上游程数	3	12	53	95	114

2.3　敏感性检验结果

好的加密算法应具备对初始条件的敏感性,这是所有混沌系统的内在性质. 对每种模型,让某些参数或初始值发生微小变化,比较所产生的混沌序列之间的差异.

表2、表3、表4表明,三种混沌动力学模型均具有对参数和初始值的极端敏感性,即使某一或某些参数发生极小的变化,例如Logistic映射中 x_0 仅发生了 $1/10^{16}$ 的变化,所产生的二进制序列中仍有半数左右的位发生改变.

2.4　重复性检验结果

考察当初始条件不同时,产生的加密序列是否会有两个或多个是重复的. 为此,让某些参数或初值在某个区间内变化来产生多个等长序列,对每一个序列我们从中间某个确定位置开始连续取128位,两两比较这些序列中有无重复序列. 考虑到本实验处理的数据量非常之大,为节省运行时间,我们产生的每个序列都相对较短,其中Logistic映射和一阶时延模型序列长度 n 取500,Lorenz映射序列长度 n 取400. 表5所示实验结果说明三种混沌伪随机序列均通过重复性检验.

表2　Logistic映射敏感性分析结果

序列长度 n	20 000		
固定参数	$u = 3.970$		无
变化的参数	$x_0 = 0.300\,0$ $x_0' = 0.300\,1$	$x_0 = 0.700\,000\,000\,000\,000\,0$ $x_0' = 0.700\,000\,000\,000\,000\,1$	$x_0 = 0.450\,0, u = 3.970\,1$ $x_0' = 0.450\,01, u = 3.970\,0$
位变化率	49.77%	50.63%	49.68%

表3　一阶时延模型敏感性分析结果

序列长度 n	20 000				
固定参数	$t_0 = 0$, $h = 0.3$				
变化的参数	$a = 3.0$ $a' = 3.1$	$b = 8.00$ $b' = 8.01$	$w = 5.000$ $w' = 4.999\,999\,999\,999\,99$	$\tau = 8.750$ $\tau' = 8.751$	$x_{t0} = 0.500\,0$ $x_{t0}' = 0.500\,1$
位变化率	49.58%	50.395%	49.635%	50.09%	49.96%

表4　Lorenz映射敏感性分析结果

序列长度 n	10 000			
固定参数	$\sigma = 10.00, x_{t_0} = 2.0,$			
变化的参数	$r = 28.00$ $r' = 28.01$	$b = 2.67$ $b' = 2.671$	$y_{t_0} = 12.0$ $y_{t_0}' = 13.0$	$z_{t_0} = 7.03$ $z_{t_0}' = 7.030\,000\,000\,000\,01$
位变化率	49.49%	50.12%	49.28%	49.71%

表5　重复性检验结果

模型类型	Logistic 映射	一阶时延模型	Lorenz 映射 x_{t0}
固定参数取值	$u = 3.970$	$a = 3.00$, $b = 8.00$, $\tau = 8.75$, $x_{t0} = 0.5$, $t_0 = 0$, $h = 0.3$	$\sigma = 10.00$, $b = 2.67$, $x_{t0} = 2.0$, $y_{t0} = 12.0$, $z_{t0} = 7.03$
序列长度 n	500	500	400
变化的参数	x_0	w	r
参数变化范围	0.2000 ~ 0.9999	5.000 ~ 9.999	25.000 ~ 27.999
参数变化步长	0.0001	0.001	0.001
序列总个数	8000	5000	3000
重复对数	0	0	0

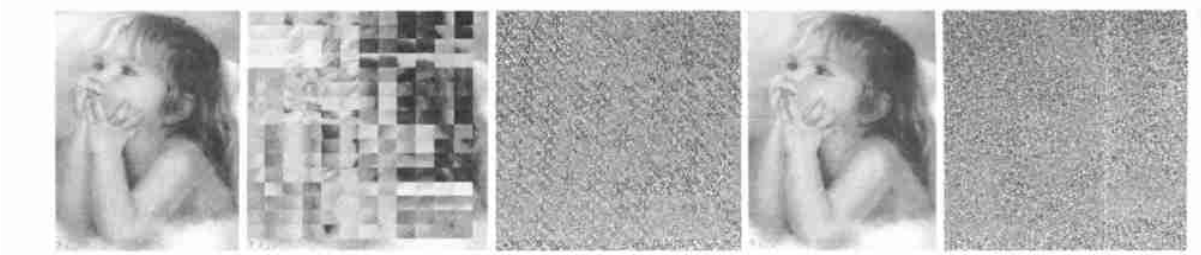
随机性、敏感性、重复性检验结果表明, 混沌序列随机性好、密钥量大、产生方法多样(包括非线性映射的多样性和初值选择的多样性). 另外, 量化阈值的选取、二值序列的抽取间隔、步长 h 的选取等对二值序列的取值均有较大影响. 攻击者要破译一段或数段明文或密文对, 既要破译混沌动力学系统的初始条件, 又要破译系统参数, 这项工作几乎是不可能完成的.

3　混沌二值序列在位图图像加密中的应用

目前多媒体通信已成为人们交流信息的重要手段, 多媒体信息的保密变得尤为重要. 由于混沌映射是确定性的, 可根据模型方程及参数初值复制混沌信号, 故可以将其用在对称密码体制中. 我们利用混沌序列对. bmp 格式的图像进行加密和解密, 以便直观的验证混沌序列的安全性能.

我们的加密在时域进行, 先通过标准幻方矩阵对图像置乱[7], 再用序列加密体制. 图 3 是加密、解密的效果图. 为突出混沌密码的加密效果, 加密部分我们分两步演示: (b) 图是仅经过 8 次幻方置乱后的图像, (c) 图是混沌伪随机序列进行序列加密的效果图. 加密、解密采用的是 Lorenz 映射, 密钥为 $\sigma = 10.00$, $r = 28.00$, $b = 2.67$, $x_{t_0} = 2.0$, $y_{t_0} = 12.0$, $z_{t_0} = 7.03$, $n = 20000$, 置乱次数为 8 次. (d) 图是解密密钥正确时的解密图像, (e) 图是解密密钥中除了参数 $r = 28.01$ 是错误的以外, 其余密钥(包括置乱次数) 全正确的解密效果图.

可见, 即使密钥值有微小的差异也会得到完全不同的解密结果(如图 e 所示), 从而无法对图像正确解密.



(a) baby 原图　(b) 8 次置换图像　(c) 加密图像　(d) 解密图像　(e) 密钥错解密图像

图3　图像加密与解密结果

4　结束语

本文通过大量实验表明了混沌伪随机序列的随机性、敏感性和重复性. 并以位图图像加密为例, 直观演示了混沌伪随机序列的加密效果. 这都预示着混沌理论在密码学领域有较广阔的发展前景. 但对混沌序列的统计分析还较为困难, 对它们的密码学验证基本还停留在数值模拟阶段, 缺乏严密的理论证明. 由于计算机实际运算精度有限, 因此, 所谓"混沌序列"的周期性是不可避免的, 算法应考虑尽量避免短周期现象的出现[1].

[参考文献]

[1] 黄月江. 信息安全与保密: 现代战争的信息卫士[M]. 北京: 国防工业出版社, 1999.

[2] Ljupco Kocarev, Goce Jakimoski. Chaos and Cryptography-PART Ⅰ: From Chaotic Maps to Encryption Algorithms[J/OL]. http://rfic.ucsd.edu/chaos/papers.html, 2000.

[3] Goce Jakimoski, Ljupco Kocare. Chaos and Cryptography-PART Ⅱ: Block Encryption Based on Chaotic Maps[J/OL]. http://rfic.ucsd.edu/chaos/papers.html, 2000.

[4] 吴伟陵. 信息处理与编码[M]. 北京: 人民邮电出版社, 1999, 7.

(下转第 78 页)

[参考文献]

[1] GB/T17986.1-2000,房产测量规范[S].国家质量技术监督局.2000.

[2] 张桥平,石强,陈晓东.房产公用建筑面积分摊计算模型研究[J].测绘工程,2000,9(4):40~42.

[3] 徐爱俊,罗学年,黄全义.房产测绘管理系统中公用建筑面积分摊模型的设计与实现[J].测绘通报.2001,11:31~33.

[4] 许捍卫,张桥平.房屋分层分户图信息管理软件的开发及相关技术[J].微型计算机应用,2002,18(9):33~41.

The Calculating Model of Sharing Floor Area for Public Use

Zhang Xin[1], Li Ruolin[2], Zhang Hong[1], Jiang Wenming[1]

(1. Key Laboratory of Jiangsu Province in Geographical Information Science, Nanjing Normal University, Nnajing 210097, PRC)
(2. Nanjing Real Estate Bureau, Nanjing 210097, PRC)

Abstract: This paper discusses the calculating model for sharing the floor area for public use, which not only conforms to the measuring standard of real estate but also is easy to calculate with the computer.

Key words: house property survey and mapping, floor area sharing, calculating model

[责任编辑:刘健]

(上接第59页)

[5] 戴跃伟,卓成春,王执铨.一种二值混沌加密序列的产生及性能分析[J].南京理工大学学报,2001,25(5).

[6] 邓浩.混沌伪随机序列和数字语音保密通信[J].通信学报,1999,20(4):29~35.

[7] 丁玮,齐东旭.数字图像变换及信息隐藏与伪装技术[J].计算机学报,1998,21(9):838~843.

The Generation of Chaotic Sequence and Its Application to Image Encryption

Sun Xiuhua, Dai Yuewei, Wang Zhiquan

(Department of Automation, Nanjing University of Science and Technology, Nanjing 210094, PRC)

Abstract: Chaos systems are characterized by its critical sensitivity to initial conditions and parameters, pseudo-randomness, and determinability. In this paper a software platform is brought forward, which can be used to generate and evaluate chaotic binary sequences. It utilizes three typical kinds of chaotic dynamics model to generate encryption sequences. Massive experiments on randomness, sensitivity and repeatability verify that chaotic sequence has higher performance in encryption. Besides, it has the capability to encrypt and decrypt many kinds of objects. The experimental results have been given on bitmap image encryption as an example.

Key words: chaotic sequence, performance evaluation, stream cipher, bitmap

[责任编辑:刘健]