具有良好安全性能的混沌映射二进制序列 [1]

何振亚 李 克 杨绿溪
(东南大学无线电系 南京 210096)

摘 要 有限精度问题降低了混沌映射序列的密码学和统计特性，且对那些与 Tent 映射拓扑共轭的映射产生的序列更可通过预测的方法精确重建。本文讨论了混沌映射的拓扑共轭变换及其性质，导出了 Tent 、 Logistic 和二阶 Chebyshev 映射的共轭关系，并针对这种攻击提出了一种混沌序列的产生方法可有效地抵抗这种攻击。

关键词 混沌映射，序列，拓扑共轭，安全性
中图号 TN914.4, TN918

1 引 言

目前扩频通信中最常用的线性码如 m-序列、 Gold 码等均有良好的自、互相关特性，但对传输系统尤其是保密通信系统而言它们是不安全的，只需知道序列中 $2n$ 个比特 (n 为寄存器级数) 的码元就可很容易地破译 [1,2] 。一维混沌映射方程简单、具有类似白噪声的统计特性，对初值的极端敏感又使得由它所产生的序列具有很好的密码学特性且数量众多，因此在扩频通信中有较好的应用前景。但在具体实现时由于有限精度问题影响了序列的各项性能。虽然采用较高的精度可使其安全性和相关性接近理想序列，但同时也提高了电路的硬件复杂度和运算代价。

Tent 映射 [3] 作为一种典型的一维混沌映射模型，它与其他的一些混沌映射，如二阶 Chebyshev 、 Logistic 映射等均有拓扑共轭 [4] 的性质，而它与 Bernoulli 映射的二进序列间又有简单的逻辑关系，这就使得一切针对 Bernoulli 映射的短序列预测 [5] 攻击对 Tent 映射以及与之拓扑共轭的映射也同样适用；反之，如果采用一种能够抵抗这种攻击的 Bernoulli 映射实现方法，则在此基础上实现的其他映射亦可抵抗这种攻击。

2 一维混沌映射和拓扑共轭性

定义 1 设 $z = h(x)$ 是单调连续函数，因而是可逆的，存在唯一的 $h^{-1}(x)$ 。对映射 $f(x)$ 作变换：

$$g(z) = h(f(h^{-1}(z))) \tag{1}$$

或简记为 $g = h \circ f \circ h^{-1}$, 则称 (1) 式为 $f(x)$ 到 $g(z)$ 的拓扑共轭变换。具有拓扑共轭关系的两个映射实质上是不同坐标表示下的同一种映射，因此拓扑共轭变换具有一些不变性质。

性质 1 如果映射 $f(x)$ 和 $g(z)$ 具有拓扑共轭关系，则

$$g^{(n)} = h \circ f^{(n)} \circ h^{-1}, \tag{2}$$

即 f 的迭代与 g 的迭代存在一一对应关系，或者说拓扑等价关系。(由定义 1 可立即证明该结论。)

[1] 1998-02-09 收到， 1998-12-11 定稿
国家自然科学基金 (No.69735101) 和江苏省自然科学基金 (No.BK97011) 资助课题

推论 1　设有一对满足拓扑共轭变换的映射 f 、 g ，如果 f 有 n 周期轨道，则 g 也有 n 周期轨道，且两者具有相同的稳定性，即 Lyapunov 指数。

Tent 映射是一种典型的一维混沌映射：

$$f_T(x) = 1 - |1 - 2x|, \quad x \in [0,1]. \tag{3}$$

它具有唯一的分布函数 $P_T(x) = 1$ ，因为与初值无关，所以必定是遍历的；又由于该映射在除 $x = 1/2$ 外各点处导数的绝对值恒为 2 ，其 Lyapunov 指数 $\lambda_T = \lim_{n\to\infty} \ln|\mathrm{d}x_n/\mathrm{d}x_0|/n = \ln 2$ 。

二阶 Chebyshev 映射 [6] 和 logistic 映射的方程分别为

$$f_C(x) = 2x^2 - 1, \qquad x \in [-1,1], \tag{4}$$

$$f_L(x) = 4x(1-x), \quad x \in [0,1]. \tag{5}$$

取变换函数

$$h(x) = \cos(\pi x), \quad x \in [0,1]; \tag{6}$$

则 $x = h^{-1}(z) = \arccos(z)$, $z \in [-1,1]$ 。由拓扑共轭的定义，有

$$g(z) = h(f(h^{-1}(z))) = h[f((1/\pi)\arccos(z))] = \cos(\pi - |\pi - 2\arccos(z)|) = 2z^2 - 1. \tag{7}$$

又由于 $h(x)$ 是单调的，所以说二阶 Chebyshev 映射是 Tent 映射的拓扑共轭变换。由拓扑共轭的性质可立即得到它的分布函数和 Lyapunov 指数分别为

$$P_C(z) = P_T(h^{-1}(z))\left|\frac{d}{dz}h^{-1}(z)\right| = \frac{1}{\pi\sqrt{1-z^2}}, \quad z \in [-1,1]; \tag{8}$$

$$\lambda_C = \lambda_T = \ln 2. \tag{9}$$

同样地，取变换函数：

$$h(x) = \sin^2(\pi x/2), \quad x \in [0,1]. \tag{10}$$

可以验证 logistic 映射与 Tent 映射也满足拓扑共轭关系，其分布函数为

$$P_L(z) = P_T(h^{-1}(z))\left|\frac{d}{dz}h^{-1}(z)\right| = \frac{1}{\pi\sqrt{z(1-z)}}, \quad z \in [0,1]. \tag{11}$$

3　混沌映射二进序列间的逻辑关系

一般地，从模拟序列向二进序列转换多采用如下两种方法：设定一个判决门限 ξ ，通过判决函数 $T(x) = [1 + \mathrm{sgn}(x - \xi)]/2$ 进行转换；或将模拟序列各点值表示成二进制数，取其中某一位作为输出序列的一个比特。对 Tent 映射可采用后者，即每次取其二进小数表示的最高位作为序列的一个比特，它也等价于取 $\xi = 1/2$ 时的第一种方法。对 logistic 、 Chebyshev 映射可分别对其相应的模拟序列取 $\xi = 1/2$ 和 0 按方法一进行转换。

对于满足拓扑共轭性质的映射，其二进序列之间亦满足一定的逻辑关系。对 Chebyshev 映射，根据它与 Tent 映射的关系 (6) 式可很容易地得到其二进序列间的逻辑关系：

$$c_n = \bar{t}_n, \tag{12}$$

其中 c_n 、 t_n 分别是 Chebyshev 和 Tent 映射的二进序列值。同样，对 logistic 映射，按照它与 Tent 映射的关系 (10) 式也可得到两者的二进序列间的逻辑关系：

$$l_n = t_n, \tag{13}$$

其中 l_n 是 logistic 映射的二进序列值。

另外，由 Bernoulli 映射的定义 $f_B(x) = 2x\mathrm{mod}1,\ x \in [0,1]$ 可知，它与 Tent 映射同属于移位映射，其区别仅在于前者是在每次迭代时将当前值左移一位并截去整数部分，而后者若首位为 0 是左移，若首位为 1 则取反后移位，由此可得到二者的二进序列间的逻辑关系为

$$t_0 = b_0, \quad t_n = b_n \oplus b_{n-1}. \tag{14}$$

结合 (12) 、 (13) 式即可得到 Chebyshev 、 logistic 映射与 Bernoulli 映射二进序列间的关系分别为

$$c_0 = \bar{b}_0, \quad c_n = \overline{b_n \oplus b_{n-1}}; \tag{15}$$

$$l_0 = b_0, \quad l_n = b_n \oplus b_{n-1}. \tag{16}$$

4 针对有限精度混沌映射序列的预测攻击和对策

在计算机上取任意初值对 Bernoulli 映射进行模拟时发现，经过很短的几次迭代后都迅速收敛至 0 。这是由于每次迭代都是将当前值乘 2 再折叠到 $[0,1]$ 区间上，用二进制小数表示时即不断地左移并截去整数部分，这就要求初值必须是能用无穷位不循环二进小数 (即无理数) 来表示才能确保映射处于混沌态，这在有限精度下是无法实现的。这样由二进序列中的一段即可很容易地恢复出映射的初始值。设序列初值为 $x_0 \in (0,1)$, 将其表示成二进小数的形式： $x_0 = \sum_{i=1}^{n} b_i \cdot 2^{-i},\ b_i \in \mathrm{GF}(2)$, 因为每次迭代输出当前 x 的最高位 b_1 ，经过 n 次迭代后即可得到 x_0 的值，若 x_0 用单精度数表示，最多只需进行 32 次迭代，即使是双精度数只需 64 次。这样就可从部分 Chebyshev 二进序列 $\{c_i, i = 1,2,\cdots,k\}$ 在任取 $b_1 = 0$ 或 1 后得到相应的部分 Bernoulli 二进序列 $\{b_i\}$ ，对不同的 b_1 会得到互反的 $\{b_i\}$ 序列即互反的初值 x_0 ，它同时也就是 Tent 序列的初值。然后再利用 Tent 映射与 Chebyshev 映射的关系 (6) 式对 x_0 进行变换即可得出 Chebyshev 映射的初值 y_0 ，进而重建出该映射完整的模拟和二进序列。对 logistic 映射序列重建的方法与此相同。

因此解决问题的关键是能否找到一种无法由部分序列完整重建的 Bernoulli 映射二进序列的实现。这里我们提出了一种基于素数的方法，即选取一大于等于所需序列长度的素数 P ，任意给出一个初始值 (小于 P 的正整数)z_0 ，按下式进行迭代：

$$\left.\begin{aligned} z_{k+1} &= 2z_k \mathrm{mod} P, \\ x_k &= z_k / P, \end{aligned}\right\} \quad z_k \in \{1,2,\cdots,P-1\}; \tag{17}$$

其中 $\{x_k\}$ 为输出模拟序列，当 $P \to \infty$ 时方程等价于 $x_{k+1} = 2x_k \mathrm{mod}1,\ x_k \in (0,1)$ 。其相应的二进序列仍可通过每次取模拟序列用二进制表示时的最高位得到。

类似地，可得到 Tent 映射的基于素数的实现方法为

$$\left.\begin{aligned} z_{k+1} &= P - |P - 2z_k|, \\ x_k &= z_k / P, \end{aligned}\right\} \quad z_k \in \{1,2,\cdots,P-1\}; \tag{18}$$

其中 $\{x_k\}$ 为输出模拟序列，向二进序列的转换与 Bernoulli 映射相同。

在由 (17) 式生成的 Bernoulli 二进序列的基础上分别用 (15) 式和 (16) 式进行转换即可得到相应的 Chebyshev 和 logistic 映射二进序列，或者对由 (18) 式生成的 Tent 模拟和二进序列分别用 (6) 、 (10) 式和 (12) 、 (13) 式转换亦可得到 Chebyshev 、 logistic 映射的模拟和二进序列。

因为除 2 以外的任一素数 P 的倒数均不能用有限位二进小数精确表示，而只能用周期为 $T(T<P)$ 的循环二进小数表示，若 $T=P-1$ 则由此产生的序列是类似于 m-序列的另一种意义上的最大长度序列。经验证发现大部分的素数均满足此条件。在扩频 CDMA 通信中，常使用超长码截短来传送信息流以提高系统的保密性和抗侦破能力，所以这里我们取较大的素数 P，选取不存在移位重叠的一组初值 $\{z_0^{(i)}\}$ 进行迭代，按给定的码周期 $T(T<<P)$ 截取得到相应的序列。

由于理论上混沌映射序列具有 δ-函数形式的自相关和零值互相关 [3,6]，而该实现方法是在有限精度下对理想混沌映射的最佳逼近，因而最大程度地保留了混沌映射的这种相关特性。以 Tent 映射为例，取一较大的素数 $P=1000003$，经验证此时对任意初值用该方法产生的序列是遍历的。取序列长度 $N=4096$，随机选取一组初值 $\{z_0^{(i)}, i=1,2,\cdots,10\}$ 按 (18) 式迭代得到 Tent 映射的一组模拟和二进序列，分别计算它们的自相关和互相关，其绝对值的统计结果列于表 1，由表可见在有限长度情况下其相关性能接近于理论值。良好的自相关特性使得序列的捕捉和同步与传统扩频序列相同。由于符合条件的素数很多，既可以用同一素数取不同的初值得到一组同族序列，也可以对不同的素数取任意初值产生一组不同族序列，而根据混沌映射对初值的极端敏感性特点，使得对无移位重叠的任意两个初值所产生的序列其互相关很小，因此大大增加了可用序列的数目，避免了 m-序列等数量有限的缺点，提高了多址通信用户容量。

表 1　Tent 映射序列相关特性的统计结果

	模拟序列最大自相关旁瓣	二进序列最大自相关旁瓣	模拟序列最大互相关	二进序列最大互相关
最大值	0.00512	0.0566	0.00589	0.0619
最小值	0.00394	0.0495	0.00430	0.0407
均值	0.00425	0.0524	0.00447	0.0481
标准差	0.00038	0.0037	0.00047	0.0052

根据 Bernoulli 映射的特点，其迭代方程可用简单的数字电路实现。素数 P 的二进制表示为 $P=\sum_{i=0}^{M-1} p_i\cdot 2^i=(p_{M-1}p_{M-2}\cdots p_1p_0)$, $p_i\in \mathrm{GF}(2)$，它的补为 $P'=2^M-P$，z 用 $M+1$ 位二进制数表示，则在每个时钟周期将 z 左移一位，判断是否超出 P，若是则将 z 的低 M 位和 P' 相加并置最高位为 0，每次输出 z 的第 M 位，对截短序列还需增加一个状态检测器对跳跃点状态进行监测，在到达跳跃点时强置移存器状态为初始态。在此基础上作相应调整即可得到 Tent 等映射的数字实现。

5　混沌映射序列的安全性能分析和仿真实验

序列的线性复杂度是其线性不可预测性的一个指标。定义 GF(q) 上序列 $a=a_0,a_1,\cdots,a_{N-1}$ 的线性复杂度 $C(a)$ 为产生该序列的 GF(q) 上级数最少的线性反馈移位寄存器的级数。对 m-序列、Gold 序列而言就是构成该序列的移存器阶数 n，由定理 1，只需 $2n$ 个相继的码元就可求出序列的线性递推公式中的系数，进而得到构成 m-序列的 n 阶本原多项式及寄存器的初始状态。

定理 1[2] 设 $a=\{a_i\}$ 是 GF(q) 上的周期序列，且序列的线性复杂度 $C(a)=L\geq 1$，则只要知道 $\{a_i\}$ 中任意相继的 $2L$ 位就可确定整个序列 $\{a_i\}$ 和产生该序列的极小多项式。

对有限长序列，线性复杂度充分大是必要的，但并非越大越好，通常把复杂度在序列长度一半附近作为序列密码安全随机性的标准之一。对于 N 位随机二进序列其线性复杂度的均值在 $N/2$ 左右[2]，而混沌扩频序列本质上是随机二进序列，因此它的线性复杂度的均值亦为序列长度的 1/2(即 $L=N/2$)，我们用 Berlekamp-Massey 算法[7] 对 Tent 和 Bernoulli 映射二进序列的线性复杂度进行计算并列于表 2，计算结果与此很好地吻合。由定理 1 可知需 N 位的码元才能确定整个序列，显然这是毫无意义的。因此，用破译 m-序列的方法无法对混沌映射序列进行预测。

表 2 混沌映射二进序列的线性复杂度

N	Tent 映射	Bernoulli 映射
128	61	60
256	129	126
512	256	255
1024	511	512
4096	2046	2048

由于本文中映射方程的特殊结构使得上述针对有限精度混沌映射序列的攻击方法亦无法对它所产生的序列进行预测重建。因为序列的初值是一个分数 z_0/P，相当于由 z_0 和 P 构成了一个密钥对 (z_0,P)，它对每个序列是唯一的。由短序列预测的方法只能得到一有限精度小数表示的初值，虽然随着运算精度的提高和截获序列长度的增加，该小数逐渐接近该分数值 z_0/P，但由于可供选用的素数数目众多，攻击者很难通过试错的方法来找到 P，因而就无法破译出该密钥对组合以及整个序列。

以 Chebyshev 映射为例比较常规方法和基于素数方法在该预测攻击下的性能。取素数 $P=1019$，经验证对任意初值用后者产生的序列是遍历的。取 $z_0=347$，相应的模拟序列初值 (取单精度) 为 $z_0/P\approx 0.3405299$，分别用这两种方法产生 Chebyshev 映射的模拟、二进序列，序列长度取为 100。图 1 中 (a)、(b) 是用常规方法产生的 Chebyshev 模拟、二进序列，(c) 是与 (b) 对应的 Bernoulli 二进序列，对单精度运算只需取该序列前 26 位即可精

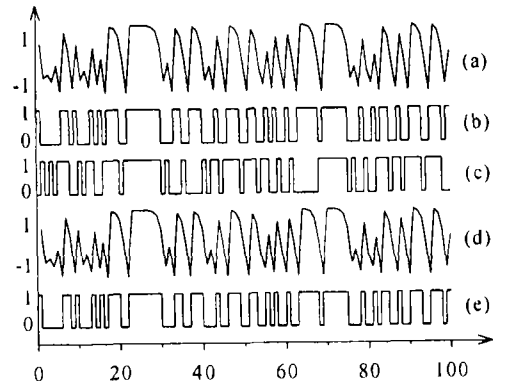

图 1 对常规的 Chebyshev 映射序列的重建

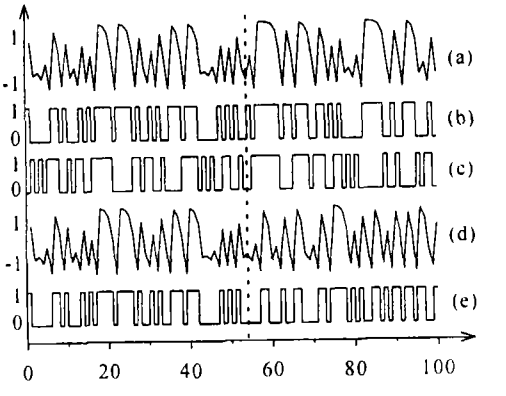

图 2 对基于素数的 Chebyshev 映射序列的重建

确恢复该映射的初值，从而重建出完整的 Chebyshev 序列 (d)、(e)。图 2 中 (a)、(b) 是用本文方法产生的 Chebyshev 序列，(c) 是与 (b) 对应的 Bernoulli 二进序列，仍用预测方法对 (a)、(b) 进行重建，运算精度为双精度，即取序列 (c) 的前 64 位，重建结果见图中 (d)、(e)，它与原序列 (a)、(b) 相比仅前 54 位相同，即不能完全重建。

6 结　论

本文针对常规混沌映射序列由于实现精度所限容易受到短序列预测攻击的缺点提出了一种新的基于素数的实现方法，理论分析和仿真实验表明，该方法所产生的映射序列具有较高的安全性能，可有效地防止短序列预测的攻击。而且由于它保留了混沌映射序列固有的特点，使得生成的序列具有较好的相关性能，大的线性复杂度，易于产生且序列数目众多，因此混沌映射序列在有较高保密性要求的扩频通信系统中有较好的应用前景。

参 考 文 献

[1] Dixon R C. Spread Spectrum Systems. New York: John Wiley & Sons, Inc. 1976, Chapter 3.

[2] 杨义先，林须端. 编码密码学. 北京：人民邮电出版社，1992, 第 15, 16 章.

[3] Nagai Y, *et al.* Guassian-like processes produced by fully developed chaos. Phys. Lett., 1985, A-112(6, 7): 259–264.

[4] 陈式刚. 映象与混沌. 北京：国防工业出版社，1992, 第 3, 9 章.

[5] 张申如，王庭昌. 混沌二进制序列构成的安全性研究. 通信保密. 1995, 95(4): 42–46.

[6] Geisel T, Fairen V. Statistical properties of chaos in Chebyshev maps. Phys. Lett., 1984, A-105(6): 263–266.

[7] Massey J L. Shift-register synthesis and BCH decoding. IEEE Trans. on IT, 1969, IT-15(1): 122–127.

CHAOTIC MAP BINARY SEQUENCES WITH GOOD SECURITY

He Zhenya　Li Ke　Yang Luxi

(*Department of Radio Engineering, Southeast University, Nanjing* 210096)

Abstract The problem of finite precision degrades the cyptologic and statistical properties of chaotic maps, and sequences generated from maps conjugated with Tent map can be reconstructed precisely by short-sequence-prediction. This paper discusses topological conjugation of chaotic maps and its properties, derives the conjugate relation between Tent, logistic and 2nd-order Chebyshev maps. A method is given to produce chaotic sequences which can sustain this predictive attack.

Key words Chaotic map, Sequence, Topological conjugation, Security

何振亚：男，1923 年生，教授，博士生导师，国家攀登计划首席科学家，主要研究领域为自适应信号处理、神经网络与智能信息处理.

李　克：男，1972 年生，博士生，主要研究方向为混沌信息处理与保密通信.

杨绿溪：男，1964 年生，博士，副教授，现从事神经网络信息处理和图像处理等的教学和研究工作.