第 56 卷 第 3 期 2007 年 3 月
1000-3290/2007/56(03)/1297-08

物 理 学 报
ACTA PHYSICA SINICA

Vol.56,No.3,March,2007

基于可变参数双向耦合映像系统的时空混沌 Hash 函数设计

刘建东　余有明

(北京石油化工学院信息工程学院,北京　102617)

(2006 年 8 月 8 日收到;2006 年 8 月 20 日收到修改稿)

在分析单向与双向耦合映像格子系统的初值与参数敏感性的基础上,提出了一种基于可变参数双向耦合映像系统的时空混沌单向 Hash 函数构造方案.该方案以耦合映像系统的部分初态作为密钥,在迭代过程中,通过上一次的迭代值和线性变换后的不同位置的明文消息比特动态确定双向耦合映像系统模型参数,将明文消息多格点并行注入时空混沌轨迹中;取迭代序列中最后一轮迭代结果的适当空间项,线性映射为 Hash 值要求的 128 bit 值.由于耦合映像系统的双向扩散机理与混乱作用,迭代过程具有极强的不可逆性及初值与参数敏感性,Hash 结果的每位都与明文及密钥有着敏感、复杂的非线性强耦合关系.仿真实验与分析结果表明,该算法达到了 Hash 函数的各项性能要求,安全性好,执行效率高.

关键词:Hash 函数,时空混沌,耦合映像格子

PACC:0545

1. 引　　言

Hash 函数是现代密码学的一个基本模块,它将任意长的消息映像为定长的 Hash 值,其目的是产生文件、消息或其他数据块的"指纹".Hash 函数主要用于数据完整性检验、鉴别协议、零知识证明和随机数发生器等,在数字签名和消息验证码中有着尤其重要的应用.传统的单向 Hash 方法,如 MD2,MD4,MD5,SHA 等,大多是基于复杂度假设,需要进行大量复杂的异或等逻辑运算或是分组多次迭代得到 Hash 结果[1].自王小云发现 MD5 等算法的碰撞问题以来[2],Hash 函数的研究又成为一个热点.由于混沌具有对初始条件敏感、伪随机和遍历等特性,近年来,人们将混沌应用到 Hash 算法的研究中,基于混沌映像模型构造出一些单向 Hash 算法[3—5].但应注意到,混沌动力学有自身的几何结构,由低维混沌系统产生的截尾混沌序列的相邻值之间具有很强的相互制约性,可以通过非线性动力学预测和重构等方法破译低维混沌系统[6,7].更为重要的是,在有限精度实现的情况下,数字化混沌系统的动力学特性相对连续系统而言存在严重的退化[8,9],这将直接影响混沌系统的安全性.

时空混沌系统具有非常多的正李雅普诺夫指数,系统在时间及空间方向上都是混沌的,其动力学行为非常丰富而复杂,空间上的任何一点的微小变化都会随时间的增加而扩散开去,产生很大的变化.利用时空混沌可以大大提高系统的复杂性,从而有利于提高系统的安全性.时空混沌研究中的典型例子是耦合映像格子(CML)模型[10].从常规密码学的角度来看 CML 模型,是相当新颖而令人兴奋的.在常规密码学的领域里,有两种广为使用的手段:混乱和扩散.在 CML 的迭代模型中,格点间的耦合起了扩散的作用,它能将一个格点的变化扩散开去,影响其他所有格点;而每个格点的非线性动力学函数就起了混乱的作用.混乱和扩散在 CML 迭代模型中被很好地结合在一起,这种结合会使经过多次迭代后的输出与初始条件及系统参数的关系变得极为复杂,而复杂的依赖关系可以很好地被用来设计单向 Hash 函数.文献[11]给出一种基于时空混沌系统的单向 Hash 函数构造方案.但是,该方案存在以下缺陷:1)采用了单向耦合映像格子系统模型(OCML),用该模型构造单向 Hash 函数存在初始扰动传播放大速度较慢、扩散机理不强的问题,初始条件或系统参数的微小差异,需要通过约 $4L$ 次迭代(L 为 OCML 的格点数),才能使混沌轨道有明显分离,变

成互不相关的两条轨道(在以往的研究中,由于OCML易同步,且计算效率高,因此在人们将时空混沌应用到密码学领域时,比较多地应用了该模型[12]).2)明文消息作为初值来驱动OCML,一个格点只能调制1个字节的明文消息,当明文文件较大时,需要的OCML格点规模太大,计算效率极低.

针对上述问题,本文提出一种基于双向耦合映像格子系统(TCML)构造时空混沌Hash函数的新方法.通过实验发现TCML与OCML的初值及参数敏感性存在较明显的差异,使用TCML构造的单向Hash函数更加快速可靠.为了进一步提高Hash函数的初值敏感性、不可逆性及防伪造性,将传统的密码设计中采用的取模操作及S-盒代数变换与双向耦合映像格子系统相结合,通过参数调制方式将明文信息并行注入双向耦合映像系统的混沌轨迹中,而它的迭代的初始点可作为Hash函数的密钥.算法保证了带密钥的Hash函数的安全性完全由密钥的安全性决定.实验结果表明,该方案有效地提高了算法的性能,并且具有较快的运算速度.

2. 耦合映像时空混沌系统的初值及参数敏感性分析

2.1. 单向耦合映像格子模型的初值及参数敏感性

格子映射为logistic映射的单向耦合单峰格子模型的形式为

$$x_{n+1}(i) = (1-\varepsilon)f(x_n(i)) + \varepsilon f[x_n(i-1)], \tag{1}$$

其中,n为离散时间步数;$i=1,2,\cdots,L$为离散格点坐标,L为系统尺寸;ε为耦合系数,且满足$0<\varepsilon<1$.非线性函数f为logistic映像,即$f(x)=\mu x(1-x)$.边界条件满足$x_n(L)=x_n(0)$,初始条件为$[0,1]$内的随机数.当参数$\mu=4$时,单个格子处于混沌状态.这时取$\varepsilon=0.99$,则耦合单峰格子的时空混沌行为进入完全发展湍流模式.在这种运动模式下,可以认为存在一个时空变换下不变的连续状态分布.

单向耦合单峰格子模型的物理背景是具有方向性的"开流",它对"扩散"的方向进行了限定,在空间的某个点上施加扰动,只能沿单一的方向传播开来,所以在i格点上的行为只对i以后的格点状态有影响.

OCML映射的初值向量为$X=[x_0(1),x_0(2),\cdots,x_0(L)]$,假设在第一个格点施加扰动$\delta$,即$X'=[x_0(1)+\delta,x_0(2),\cdots,x_0(L)]$,图1为扰动$\delta$取不同的值时,得到的第$L$个格点$x_n(L)$的迭代序列所产生的误差的测试结果.其中图1(a)为格子数$L=10$,初值误差$\delta=10^{-2}$的测试结果;图1(b)为格子数$L=50$,初值误差$\delta=10^{-2}$的测试结果;图1(c)为格子数$L=10$,初值误差$\delta=10^{-5}$的测试结果;图1(d)为格子数$L=50$,初值误差$\delta=10^{-15}$的测试结果.仔细观察测试结果,发现这样一个规律:迭代到$L,2L,3L$次时,$x_n(L)$与$x'_n(L)$出现间歇性分离,随后又回到相同的轨道,迭代到$4L$次时,$x_n(L)$与$x'_n(L)$才出现持续稳定的明显分离.扰动δ的强弱,会影响间歇性分离所持续的时间及分离的程度,但出现持续稳定分离的迭代步数(图中标注的m的位置)却几乎不受δ强弱的影响,m的位置标志$x_n(L)$与$x'_n(L)$出现持续稳定的明显分离的开始.图2给出$\bar{m}$随耦合格子长度L变化的情况.$\bar{m}$是从1000个不同初值向量X得到的m的算术平均值.由图2可看出,初值误差只在格点数较小($L<60$)时对$\bar{m}$的值有轻微的影响,$\bar{m}$与耦合格子长度L呈近似线性的比例关系,即$\bar{m}\approx 4\times L$.

实验结果所揭示的这个有趣的规律可以作如下解释:由于选择的耦合系数ε比较大($\varepsilon=0.99$),迭代时施加在某个格点的扰动单方向扩散的强度较大,也就是说,扰动对单一方向上相邻的格点的影响要比对自身的影响大很多(原因是$\varepsilon\gg 1-\varepsilon$).随着扩散的进行,初始扰动被逐渐扩大,经过$L-1$次迭代,施加在第一个格点的扰动$\delta$扩散到格点$L$,并使$x_n(L)$的值有了较明显的差别,也就是说,$L-1$次迭代后,在第$L$个格点上产生了较强的扰动,但在随后的迭代中,增强了的扰动继续向前扩散(在周期性边界条件$x_n(L)=x_n(0)$下,又扩散到第1个格点,随后开始下一个周期的扩散),由于$1-\varepsilon$的值非常小,扰动对自身的影响却在随后的迭代中逐渐得到抑制,这样就出现了间歇性分离现象.这种间歇性分离现象周期性地出现3次,当初始扰动经过$4\times L$次迭代放大后,扰动扩散产生的"能量"足以抵制$1-\varepsilon$的抑制作用,因而出现迭代序列持续稳定的明显分离现象.总之,OCML映射对初值十分敏感,但在迭代过程中会出现间歇分离现象,使映射轨迹产生持续明显分离的迭代次数约为$4L$次.

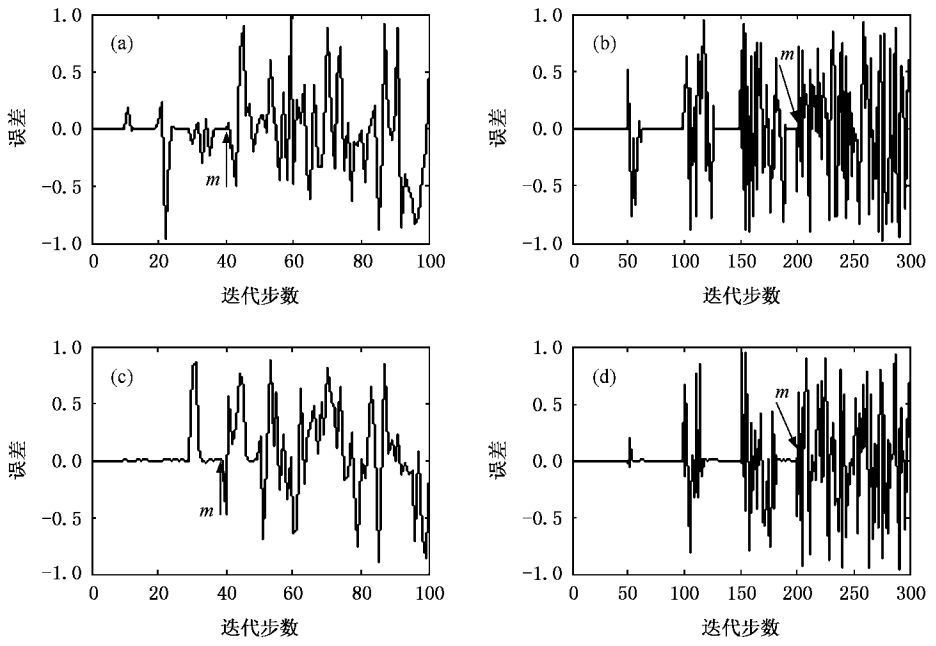

图 1　单向耦合映像格子初值敏感性　(a) $L=10$ 初值误差为 10^{-2} 数量级；(b) $L=50$ 初值误差为 10^{-2} 数量级；(c) $L=10$ 初值误差为 10^{-5} 数量级；(d) $L=50$ 初值误差为 10^{-15} 数量级

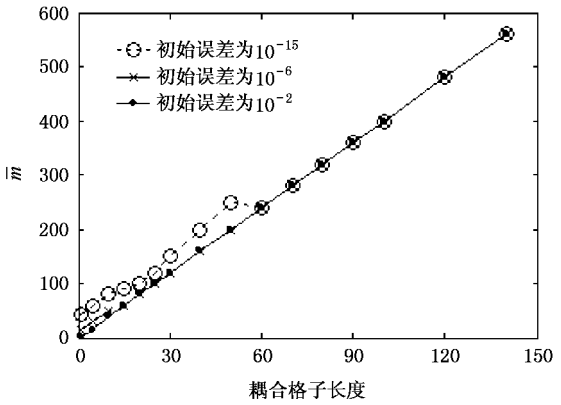

图 2　OCML 模型中 $\bar{m}$ 与耦合格子长度的关系

2.2. 双向耦合映像格子模型的初值及参数敏感性

格子映射为 logistic 映射的双向耦合单峰格子模型的形式为

$$x_{n+1}(i) = (1-\varepsilon)f(x_n(i) + \frac{\varepsilon}{2}[f(x_n(i-1)) + f(x_n(i+1))], \tag{2}$$

系统(2)中，各参数的取值与系统(1)相同.很明显，该模型中某个格点上的误差会沿着两个方向向其他格点扩散(注意使用的是周期边界条件)，因此格点上的扰动在空间及时间方向的扩散将更为复杂，速度会更快.仍假设在第一个格点上施加扰动 δ，图 3 为扰动 δ 取不同的值时，得到的第 $(L/2)$ 个格点的迭代序列所产生的误差的测试结果.其中图 3(a)为格子数 $L=50$，初值误差 $\delta=10^{-2}$ 的测试结果；图 3(b)为格子数 $L=70$，初值误差 $\delta=10^{-15}$ 的测试结果.未出现间歇性分离现象.图 4 给出初值误差 δ 取不同的数量级时，$\bar{m}$ 随耦合格子长度 L 变化的情况.$\bar{m}$ 仍是从 1000 个不同初值向量 X 得到的 m 的算术平均值.由图 4 可看出，$\bar{m}$ 的值较图 2 相应的值明显减小，随着耦合格子长度 L 的增加，在迭代步数还小于耦合格子长度 L 的情况下，迭代产生的序列就有持续的较明显的差别.施加在初始格点的扰动 δ 对 $\bar{m}$ 有一定影响，δ 越小，$\bar{m}$ 越大.

对于系统(1)及系统(2)，在[3.75,4]区间内，分别对 logistic 映射中的参数 μ 施加扰动，进行敏感性实验.扰动方法是在第 n 次迭代时，对第 1 个格点中的 μ 施加扰动 δ，第 $n+1$ 次迭代时 μ 恢复原来的值，实验得出与系统对初值敏感性相类似的结果.

实验结果表明，耦合单峰映像格子映射对初始状态和系统参数极度敏感，而且双向耦合单峰映像格子映射与单向耦合单峰映像格子映射的初值及参数敏感性存在较明显的差异，为了使迭代序列出现可靠的明显分离，当格子数 L 较大时，TCML 所需要

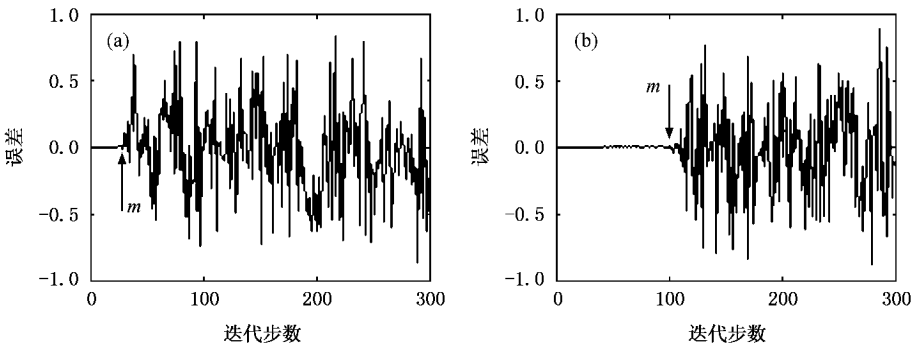

图 3 双向耦合映像格子初值敏感性 (a) $L=50$ 初值误差为 10^{-2} 数量级；(b) $L=70$ 初值误差为 10^{-15} 数量级

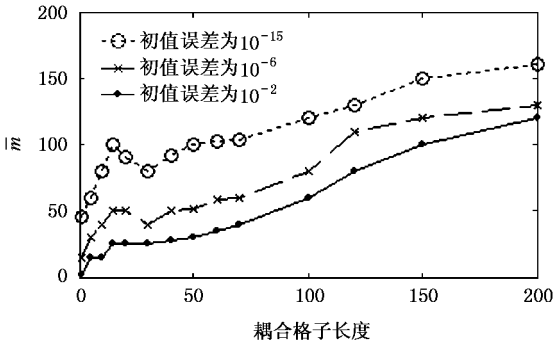

图 4 TCML 模型中 $\bar{m}$ 与耦合格子长度的关系

的迭代步数不足 OCML 的 $\frac{1}{4}$. 因此，虽然在每个格点的迭代中，TCML 比 OCML 要多计算一个非线性函数 $f(x)$，但用 TCML 构造的单向 Hash 函数要比用 OCML 构造的单向 Hash 函数快速可靠，用 TCML 构造单向 Hash 函数会有更高的执行效率.

3. 基于双向耦合映像系统的单向 Hash 函数构造

混沌系统的最本质特征就是初值及参数敏感性. 混沌系统中，初始条件的微小扰动在非线性机理的作用下，只需经历一个有限过程，就可被放大成宏观层次上的不确定性. 由于混沌轨道的局部不稳定性，相邻轨道以指数速度分离，初始条件包含的信息会在运动过程中逐步消失，因而满足了 Hash 函数设计的单向性及抗碰撞性要求. 然而，在有限精度条件下实现的混沌系统中，由于混沌特性的退化，敏感性会受到一定程度的影响. 因此，提高初值及参数敏感性应成为构造单向 Hash 函数时设计混沌系统所要考虑的主要问题. 我们用 TCML 模型并结合代数系统[13]来构造单向 Hash 函数，动力学方程为

$$
\begin{aligned}
&x_{n+1}(j) = (1-\varepsilon) f_j[x_n(j)] + \frac{\varepsilon}{2}[f_j(x_n(j-1)) \\
&\qquad + f_j(x_n(j+1))], \\
&f_j(x) = \mu x(1-x),\ j = 1,2,\cdots,p, \\
&\mu = 3.75 + (C_{(j,n)} + x_n(j))/8, \qquad (3) \\
&x_{n+1}(p+1) = (1-\varepsilon) f_j[x_n(p+1)] \\
&\qquad + \frac{\varepsilon}{2}[f_j(x_n(p)) + f_j(x_n(p+2))], \\
&Q'_n = [\mathrm{int}(x_n(p+1)) \times 2^{52}] \bmod 2^{\lambda}, \\
&Q_n = Sbox(Q'_n), \\
&x_n(p+1) = Q_n/2^{\lambda},\ f(x) = 4x(1-x), \qquad (4) \\
&x_{n+1}(j) = (1-\varepsilon) f_j[x_n(j)] \\
&\qquad + \frac{\varepsilon}{2}[f_j(x_n(j-1)) + f_j(x_n(j+1))], \\
&f(x) = 4x(1-x),\ j = p+2,\cdots,L, \qquad (5) \\
&K_n = [\mathrm{int}(x_n(L)) \times 2^{52}] \bmod 2^{\lambda}, \\
&x_n(0) = K_n/2^{\lambda}, \\
&D_n = [\mathrm{int}(x_n(1)) \times 2^{52}] \bmod 2^{\lambda}, \\
&x_n(L+1) = D_n/2^{\lambda}. \qquad (6)
\end{aligned}
$$

方程(4)中的 S-盒定义如下：

$$
\begin{aligned}
&A_1 = [(Q'_n \gg 24) \& 255], \\
&A_2 = [(Q'_n \gg 16) \& 255], \\
&A_3 = [(Q'_n \gg 8) \& 255], \\
&A_4 = [(Q'_n) \& 255], \\
&A_0 = A_1 \oplus A_2 \oplus A_3 \oplus A_4, \qquad (7) \\
&Q_n = [A_0 \ll 24] + [A_4 \ll 16] + [A_3 \ll 8] + A_2.
\end{aligned}
$$

方程(7)中的符号“$\gg$”(“$\ll$”)表示右移(左移)循环操作，“&”表示按位与操作，“$\oplus$”表示异或操

作.在方程(3)—(6)中，$C_{(j,n)}$ 为整个待处理文本按对应字节的 ASCII 码转换为数字，线性变换为 0—1 之间的数所得到的二维大数组，p 是注入文本信息空间的维数，在迭代过程中，根据上一次的迭代值和不同位置的原始文本字节来动态确定参数 μ，每轮迭代注入 p 字节信息，由于参数扰动能够在两个方向上扩散，因而增加了明文数组对整个迭代过程的影响程度.ε 是耦合系数，取 $\varepsilon = 0.99$；L 是时空混沌的系统尺寸.在方程(6)中，应用了取整(int)和取模(mod)的代数操作.取整及取模操作能够增加初值及系统参数的敏感性并能有效地改进迭代序列的随机性能.在方程(4)中，应用了 S-盒代数变换，极大地增加了初值及系统参数的灵敏性，增强了系统的扩散机理.

本文算法取 $L = 37$，$p = 32$，$\lambda = 32$.从图 4 可看到，在 $L = 30$ 左右取值时，$\bar{m}$的值出现一个低谷，因此取 $L = 37$ 有利于改善单向 Hash 函数的性能.

构造带密钥的单向 Hash 算法的基本思想是将密钥经线性变换作为双向耦合映像格子映射初值向量的部分分量，初值向量的其他分量为随机选择的[0,1]区间中的数，通过多轮迭代，产生离散的时空混沌序列.迭代过程中，将原始消息以字节为单位，线性变换后对双向耦合映像格子模型中 logistic 映像的参数进行调制.最后将空间格点为 11，31，37 的迭代结果 $x_R(11)$，$x_R(31)$，$x_R(37)$(R 为迭代次数)线性映像为 40 bit，40 bit，48 bit 的 3 个二进制数，从而形成 128 bit 的 Hash 值.算法的一般过程具体描述如下：

1) 将明文消息 x 分割成 p 字节的消息块 x_1，x_2，…，x_t，最后一个块填充为：x_t = * … * 10…0length(x)，其中 length(x)表示 x 的长度的二进制形式，长度为 64 bit，不足 64 bit 时高位添一个介符 1 再补 0.

2) 将每个消息块 x_i 按字节线性变换为 0—1 之间的数，由整个消息得到一个二维大数组 $C_{(p,t)}$，数组的行数为 p，列数为 t.

3) 将密钥分解为三部分，分别线性变换为 0—1 之间的数，作为时空混沌映射初值向量 $X = [x_0(1), x_0(2), \cdots, x_0(L)]$的第 1，3，5 个分量的值，其他分量取 0—1 之间的随机数(要求取 6 位以上有效数字，且最末位不为 0).

4) 应用方程(3)—(7)迭代初值向量 X，在迭代过程中，根据上一轮的迭代值和数组 $C_{(p,t)}$ 不同位置的元素来动态确定 logistic 映像的参数(方程(3))，每一轮迭代完成对数组 $C_{(p,t)}$ 一列元素的处理，因此，完成对数组 $C_{(p,t)}$ 所有元素的处理需 t 轮迭代，在此基础上，保持方程(3)中的参数 μ 不变，再进行 r 轮迭代，总的迭代次数 $R = t + r$.迭代生成时空混沌序列 L 组：$x_n(1)$，$x_n(2)$，$x_n(3)$，…，$x_n(L)$.

根据图 4 的实验结果，$L = 37$ 时，敏感度要达到 10^{-15}数量级，r 的取值应不小于 80.事实上，由于时空混沌模型中采取了取模及 S-盒代数变换，加速了扰动的扩散放大速度，可使 r 的取值大大减小($r \geqslant 19$ 时，初值及系统参数敏感度就能达到 10^{-15}数量级)，但为了增大时空混沌系统在迭代中的信息损失，保证算法的高安全性及高可靠性，我们取 $r = 80$.

5) 由于非相邻格子迭代序列的不相关性[14]，从迭代序列中取出最后一轮迭代的结果 $x_R(11)$，$x_R(31)$，$x_R(37)$，将它们经线性变换和取整运算映像为 2 个 40 bit，1 个 48 bit 的二进制数，合起来作为最后 128 bit 的 Hash 值.

4. 仿真实验

4.1. 文本 Hash 结果

初始文本 1 为‘ In recent years，a number of digital chaotic cryotographic approaches have been proposed for realizing private key cryptography with chaos. A typical one was proposed by Baptista that the message text is encrypted as the number of iterations applied in the chaotic map in order to reach the region correspondent to that text. The resultantciphertexts are integers and are suitable to be transmitted through public data communication networks. ”，文本 2 将文本 1 中首字母的 I 改为小写，文本 3 将 1 中的 approaches 写成 approachet，文本 4 将文本 1 中最后的句号改成逗号，文本 5 将文本 1 中最后加一个空格. Hash 结果用十六进制数表示，对于上述几种情况得到的 Hash 结果分别为

文本 1：D9AD8D3A1EA6C9A33725D35FE4836BDB；

文本 2：F87E953F40FD76b7D22A71E7DFA374BD；

文本 3：79A15217FCB0F6DE8231B012630CB4B2；

文本 4：1591B54656C15C4492A46547E8CB35F9；

文本 5：4AB0DE7CB3AD71E687F8AF7528C7579A.

可见输入的任何微小的改变都会引起输出的很大差异,算法具有高度的初值敏感性.

4.2. 混乱与扩散性质统计分析

Hash 结果的二进制表示中每 bit 只取 0 或 1,因此理想的 Hash 函数应该是初值的扰动将导致 Hash 结果的每 bit 都以 50% 的概率变化.参照文献[15]中测试方法和 $\bar{B}$,P,ΔB,ΔP 统计量定义,考察明文变化 1 bit 时,Hash 结果变化比特数 B 的分布.

在 $N=1024$ 次测试下,所得变化比特数 B 分布情况如图 5 所示,B 介于 50 和 78 之间,平均 bit 变化数为63.835个,非常接近理想状况下的 64 bit 变化数,其上下平均波动幅度很小,表明算法具备很强的明文置乱能力.

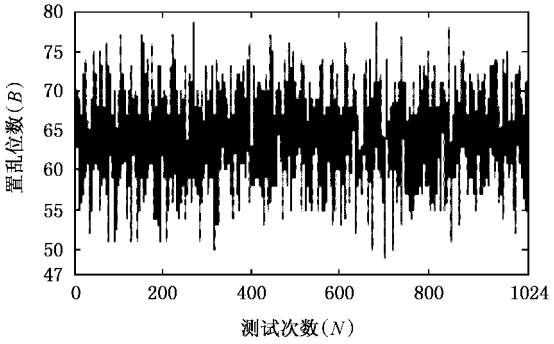

图 5 变化比特数 B 分布图

另外,分别经 $N=256,512,1024,2048$ 次测试,得到的 $\bar{B}$,P,ΔB,ΔP 值如表 1 所示.从表 1 中的数据可看出,算法的平均变化比特数和每比特平均变化概率都趋近于理想状况下 64 bit 和 50%,相当充分和均匀地利用了密文空间,明文的任何扰动,使得密文在统计上产生接近等概率的均匀分布,从统计效果上保证了攻击者无法在已知一些明文密文对的情况下伪造其他的明文密文.同时,ΔB 与 ΔP 都很小,表明算法对明文的混乱与扩散能力强而稳定.

表 1 Hash 结果统计表

	$N=256$	$N=512$	$N=1024$	$N=2048$	总平均
$\bar{B}$	64.311	63.798	63.835	63.957	63.975
ΔB	5.372	5.358	5.725	5.443	5.475
P/%	50.24	49.84	49.87	49.96	49.978
ΔP/%	4.196	4.180	4.47	4.252	4.275

4.3. 算法的碰撞分析

碰撞指给定不同的初值,而 Hash 结果却相同,即发生了多对一映射.我们通过以下的实验来定量地测试本算法的抗碰撞能力[3]:在明文空间中随机地选取一段明文求出其 Hash 值,并以十六进制值的方式来表示,然后随机地选择并改变明文中 1bit 的值得到另一新的 Hash 结果.比较两个 Hash 结果,记录在相同位置出现相同十六进制值的情况,并作统计分析.经过 2048 次测试,相同位置出现相同十六进制值的概率是 6.16%(图 6).图中可以看出每次变换一个明文 bit 位,Hash 值的 32 个十六进制值中,平均只有 1.9716 个十六进制值在相同位置取值相同,碰撞的程度很低.

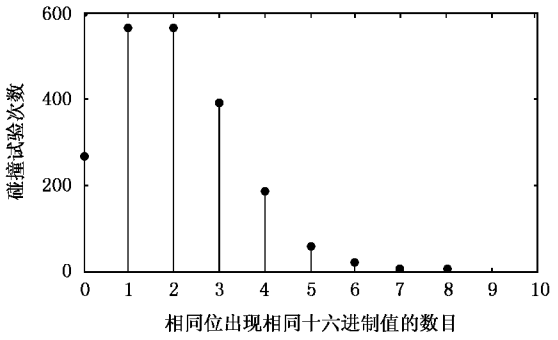

图 6 碰撞分析

4.4. 密钥敏感性分析

为了考察时空混沌映射初值向量 $X=[x_0(1),x_0(2),\cdots,x_0(L)]$ 的第 1,3,5 个分量(密钥)对 Hash 结果的影响,在(0,1)实数范围内,随机选取密钥分量($x_0(1)$,$x_0(3)$,$x_0(5)$)的 100 组值,定义 δ 为密钥各分量的变化量,$\bar{B}$ 为 δ 对应的 Hash 平均变化比特数.δ 取不同值时,测试相应变化量下 $\bar{B}$ 的大小.测试的 δ-$\bar{B}$ 关系如图 7 所示.

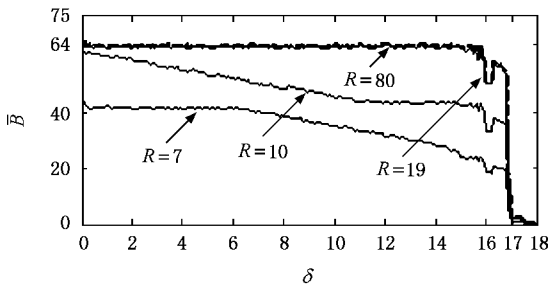

图 7 $\delta_{(x0(1))}$-$\bar{B}$ 曲线图

图 7 中,横坐标为密钥分量变化量的负对数表示,纵坐标为相应的 $\bar{B}$ 值.当迭代轮数 $R=80$,$\delta_{(x_0(1))}=10^{-16}$ 时,$\bar{B}=64$,当 $R=80$,$\delta_{(x_0(1))}=10^{-17}$

时，$\bar{B}=0$，因此算法对分量 $\delta_{(x_0(1))}$ 的敏感度为 10^{-16} 数量级．可见算法对密钥高度敏感．当 R 减小到 19 时，算法对分量 $\delta_{(x_0(1))}$ 的敏感度几乎没有受到影响．而当 $R<19$ 时，随着 R 的减小，密钥敏感度会明显下降．图中分别给出 R 取 10 和 7 的试验结果．经类似试验测得，算法对 $x_0(3)$，$x_0(5)$ 的敏感度也为 10^{-16} 数量级．按照 IEEE-754 标准，这实际上是双精度条件下所能达到的最高敏感度．此时，密钥空间的大小为 $\#(K)=10^{48}\approx 2^{158}$，达到 158 bit．根据加密强度的要求，可以选择更多的初值作为密钥，这样就可以实现更大的密钥空间．大量的数值实验还表明，在计算精度范围内，密钥分量在密钥集合(0,1)中具有连续性及均匀分布性．作为密钥的初值与其他初值向量分量呈相间分布，迭代时相互间的扩散作用很大程度上克服了传统混沌密码系统中的弱密钥问题[16]．

与文献[11]相比，本文算法有如下优点：

1) 双向耦合映像格子的双向扩散机理，加速了雪崩效应，具有更高的复杂性，加上取模运算及 S-盒代数变换对初值及参数扰动的高敏感性，提高了 Hash 函数的安全性．

2) 通过参数调制方式将明文信息逐轮并行注入 32 个格点中，这样为达到整体复杂性所需付出的计算代价由消息块的并行处理所分担，计算量大为降低．对于字节数为 N 的明文信息，文献[11]需要 N 个格点的规模，迭代轮数 $R\gg 3\times N$，每轮迭代需计算 $2\times N$ 个 logistic 映像，而本文算法中每轮注入 32 个字节，模型的格点数是 37，迭代轮数 $R=N/32+80$，每轮迭代需计算 111(37×3)个 logistic 映像．当明文较大时，本文算法的迭代轮数小于文献[11]的 1/96，需计算的 logistic 映像个数约为文献[11]的 $1/(1.73\times N)$．

3) 在前 t 轮迭代中，每次迭代的方程(3)的参数 μ 都由上一次的迭代值和某个消息字节共同确定，这将确保最终 Hash 值的任一比特都与消息的所有比特相关，同时，也使方程(3)的参数 μ 与初值(密钥)相关，增加了系统对密钥的敏感度．这里需指出，由于时空混沌模型中所用的非线性函数(logistic 映像)是一个关于直线 $x=0.5$ 左右对称的抛物线，即初始值 $x_0=a$ 与 $x_0=1-a$ 将会形成相同的迭代值，因此文献[11]中，作为初值的消息变换值 $x_0=a$ 与 $x_0=1-a$ 将会产生相同的 Hash 值，即发生了碰撞．为了避免这种情况发生，应将消息线性变换到区间(0,0.5)内．本文算法由于实现了密钥对参数 μ 的调制，破坏了 logistic 映像关于 $x=0.5$ 的对称性，进而将密钥分量由集合(0,0.5)拓展到(0,1)．

5. 结　论

本文提出了一种基于可变参数双向耦合映像系统的时空混沌单向 Hash 函数构造方案．时空混沌中大量正的 lyapunov 指数的存在和大量格点的并行非线性映像及邻近点状态的相互耦合发展使 Shannon 理论中的混合和扩散得到非常有效的实现．双向耦合映像系统强化了扩散机理，其时空混沌行为的复杂动力学特性足以保证算法的安全性．通过不同位置的明文消息与上一次的迭代值对 TCML 模型参数的共同调制，使所获得的迭代序列是 TCML 映射振荡参数下产生的迭代序列，因而更加复杂．取模运算、S-盒代数变换及明文消息的多格点并行参数调制，有效地减小了运算量，使算法既满足高安全度的要求，又具有高的执行效率．Hash 结果的每位都与明文及密钥有着敏感、复杂的非线性强耦合关系，可有效抵抗线性分析．由于算法具有很大的密钥空间，可以抵抗密钥的强力攻击．各格点计算过程相同，整个计算过程的并行程度很好，可以直接并行化．理论分析与计算机仿真实验表明，该方案很好地达到了 Hash 函数的各项性能要求，高效快速，可靠性好，有实际推广潜力．

[1] Pieprzyh J, Sadeghiyan B 1993 *Design of Hashing Algorithm* (Berlin: Springer)

[2] Wang X Y, Feng D G, Yu X Y 2005 *Science in China Series* E **35** 1 (in Chinese) [王小云、冯登国、于秀源 2005 中国科学 E 辑 **35** 1]

[3] Wong K W 2003 *Phys. Lett.* A **307** 292

[4] Liu J N, Xie J C, Wang P 2000 *Journal of Tsinghua University* (Natural Science Edition) **40** 55 (in Chinese) [刘军宁、谢杰成、王　普 2000 清华大学学报(自然科学版) **40** 55]

[5] Xiao D, Liao X F, Deng S J 2005 *Chaos, Solitons and Fractals* **24** 65

[6] Wang S H, Kuang J Y, Li J H 2002 *Phys. Rev.* **66** 1

[7] Ni W S, Hua Y M, Deng H et al 1996 *Progress in Physics* **16** 634 (in Chinese) [倪皖孙、华一满、邓 浩等 1996 物理学进展 **16** 634]

[8] Li S J, Mou X Q, Cai Y L et al 2003 *Computer Physics Communications* **153** 52

[9] Goce J, Ljupco K 2001 *Physics Letters* A **291** 381

[10] Yang W M 1994 *Spatiotemporal Chaos and coupled Map Lattice* (Shanghai: Shanghai Scientific and Technological Education Publishing House) P12 (in Chinese) [杨维明 1994 时空混沌和耦合映像格子(上海科学技术教育出版社)第 12 页]

[11] Zhang H, Wang X F, Li Z H et al 2005 *Acta Phys. Sin.* **54** 4006 (in Chinese) [张 瀚、王秀峰、李朝晖等 2005 物理学报 **54** 4006]

[12] Kuang J Y, Deng K, 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]

[13] Lü H P 2004 *Journal of Qujing Teachers College* **23** 1 (in Chinese) [吕华平 2004 曲靖师院学报 **23** 1]

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

[15] Wang X M, Zhang J S, Zhang W F 2003 *Acta Phys. Sin.* **52** 2737 (in Chinese) [王小敏、张家树、张文芳 2003 物理学报 **52** 2737]

[16] Wang L, Wang F P, Wang Z J 2006 *Acta Phys. Sin.* **55** 3964 (in Chinese) [王 蕾、汪芙平、王赞基 2006 物理学报 **55** 3964]

A TCML-based spatiotemporal chaotic one-way Hash function with changeable-parameter

Liu Jian-Dong Yu You-Ming

(*Information Engineering College, Beijing Institute of Petrochemical Technology, Beijing* 102617, *China*)

(Received 8 August 2006; revised manuscript received 20 August 2006)

Abstract

A TCML-based spatiotemporal chaotic one-way Hash function with changeable parameter was constructed based on the analysis of sensitivity to initial value and parameters of one-way and two-way coupled map Lattice (TCML) systems. The approach is implemented by employing part of the initial values of coupled map system as the secret key, and the parameters of two-way coupled map system in each iteration is dynamically determined by the value of the last iteration and the corresponding message bit in different positions, and then making message with multigrid embedded in spatiotemporal chaos track in parallel. Choosing some suitable spatial items of the result of the final iteration, the Hash value was obtained by means of linear transform limited with 128 bits. Iteration process has very strong irreversibility and sensitivity to initial values and parameters. Each bit of Hash value has very sensitive, complex and strongly nonlinear coupling relation with the corresponding message and secret key because of the bidirectional diffusion and confusion characteristics. Simulation and analysis demonstrate that the algorithm satisfies all the performance requirements of Hash function and is reliable, secure and efficient.

Keywords: Hash function, spatiotemporal chaos, coupled map lattice

PACC: 0545