一种基于混沌神经网络的序列密码算法*)

韦玉轩[1] 韦鹏程[2] 张 伟[2]

(广西建设职业技术学院 南宁 530003)[1] (重庆教育学院计算机与现代教育技术系 重庆 400067)[2]

摘 要 混沌序列具有带宽大、类噪声、难于预测和重构等特点,因而非常适用于网络通信和数据加密。本文结合神经网络和混沌映射的特点,提出了一种基于混沌神经网络和混沌映射序列密码的设计方法,该方法可以克服有限精度效应对混沌系统的影响,从而改善混沌序列特性,理论和实验结果表明:在有限精度实现下,该方法可以有效提高混沌系统的复杂性和随机性,并且算法安全性高、运算速度快,并且便于软硬件的实现。

关键词 序列密码,混沌序列,混沌神经网络,混沌映射

A Stream Cryptographic Algorithm Based on Chaotic Neural Network

WEI Yu-Xuan[1] WEI Peng-Cheng[2] ZHANG Wei[2]

(Guangxi Polytechnic College of Construction, Nanning 530003)[1]

(Department of Computer and Modern Education Technology, Chongqing Education College, Chongqing 400067)[2]

Abstract As we all know, chaotic system is sensitive to initial values and system parameters, so it is suitable for information encryption. In this paper, a stream cryptographic algorithm based on chaotic neural network is proposed, this scheme can improve the complexity and the period of the chaotic system under the finite-precision circumstances. The theoretic and simulation show that the new approach not only owns high complexity, good randomness properties, but also has high security and implemented easily in both software and hardware.

Keywords Stream cipher, Chaotic sequences, Chaotic neural network, Chaotic map

1 引言

混沌神经网络模型最早是根据生物神经元的混沌特性于20世纪90年代初由K. Aihara, T. Takabe和M. Toyoda等人首次提出来的[1,2],它具有非常丰富和复杂的非线性动力学特性,特别是其混沌动力学特性,不仅能产生无法预测的伪随机序列轨迹,而且是一个非常复杂难解的NP问题。与基于移位寄存器的序列加密法相比,混沌神经网络在序列周期、随机统计性以及线性复杂度方面均有优势,而且这种混沌加密算法的安全性更高。

利用混沌映射产生混沌序列的理论研究已经很成熟。但是,混沌序列发生器总是在有限精度下实现,混沌迭代过程必将退化为周期序列。因此,有限精度效应是混沌序列从理论走向应用的主要障碍[3,4]。本文结合混沌神经网络和混沌映射的特点,提出了一种基于混沌神经网络和混沌映射伪随机序列的设计方法,设计出混合混沌的序列密码算法,该方法可以克服有限精度效应对混沌系统的影响,改善混沌序列特性,并从理论和数字实验两方面对其该序列密码安全性进行了评估、分析。结果表明,这种基于混沌神经网络和混沌映射的序列密码具有随机性好、实现容易、周期长等优点,为在低成本下得到比较实用的序列密码提供了一种新的思路。

2 细胞神经网络和Chebyshev映射

细胞神经网络(CNN)理论及其应用是由Chua等于1988年首先提出的,其规则的结构和局部的连接性质使其易于实现超大规模集成电路(VLSI),故CNN具有广泛的应用前景,目前CNN作为一种灵活而有效的神经网络模型在保密通信、图像处理、模式识别和物理学等领域的应用受到很多学者的关注[5,6]。CNN的应用(如图像处理、模式识别和控制)在很大程度上取决于其动力学行为,往往需要网络收敛于稳定的平衡点,在保密通信和物理学中的应用往往需要网络具有混沌吸引子或极限环解[7]。本文研究如下三阶CNN动态模型:

$$\frac{dx_j}{dt}=-x_j+a_jy_j+\sum_{k=1,k\neq j}^{3}a_{jk}y_k+\sum_{k=1}^{3}S_{jk}x_k+i_j \quad j=1,2,3 \tag{1}$$

其中x_j是状态变量,y_j是相应的输出,满足如下公式:

$$y_j=0.5(|x_j+1|-|x_j-1|) \quad j=1,2,3 \tag{2}$$

令:$a_{12}=a_{13}=a_2=a_{23}=a_{32}=a_3=a_{21}=a_{31}=0$;$S_{13}=S_{31}=S_{22}=0$;$i_1=i_2=i_3=0$;$S_{21}=S_{23}=1$,则系统(1)变成:

$$\begin{cases}\dfrac{dx_1}{dt}=-x_1+a_1y_1+S_{11}x_1+S_{12}x\\ \dfrac{dx_2}{dt}=-x_2+x_1+x_3\\ \dfrac{dx_3}{dt}=-x_3+S_{32}x_2+x_3\end{cases} \tag{3}$$

从图1可以看出,只要调节非线性扰动参数a_1、S_{11}、S_{12}和S_{32},系统的运动轨迹变为不同的混沌吸引子。

Chebyshev映射式(4)是一个典型的混沌系统,满足绝对连续不变测度、等分布和对称特性的条件。

$$x_{n+1}=\cos(k\arccos(x_n)),-1\leqslant x_n\leqslant 1,n=1,2,3\cdots \tag{4}$$

*)基金项目:重庆市教委资助项目(No. kj051501),重庆教育学院重点项目。韦玉轩 硕士研究生,主要研究方向信息安全技术;韦鹏程 博士研究生,主要研究方向为信息安全技术,混沌理论和混沌密码学;张 伟 教授,博士后,主要研究方向为信息安全、计算智能与数据挖掘。

这里我们主要讨论 k≥2，k=4 的混沌特性。图 2 是两个初始值相差仅为 10^{-5} 的两个混沌时间序列，这表明 Chebyshev 映射具有良好的初始值敏感性，能产生良好性能的伪随机序列。

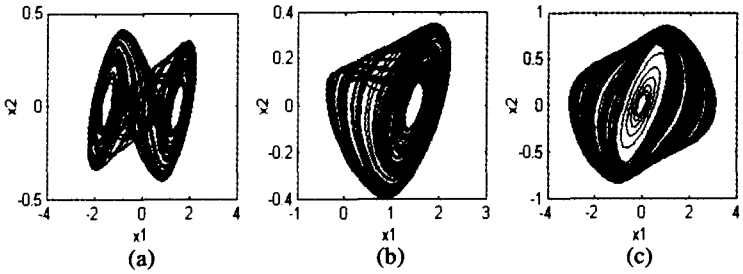

图 1　三阶细胞神经网络的混沌吸引子(a)：$a_1=3.86$，$s_{11}=-1.55$，$s_{12}=8.98$，$s_{32}=-14.25$；(b)$a_1=3.85$，$s_{11}=-1.55$，$s_{12}=8.76$，$s_{32}=-14.35$；(c) $a_1=-4.198$，$s_{11}=2.365$，$s_{12}=7.45$，$s_{32}=-10.98$

3　算法描述

从理论上讲，对任意给定系统(3)式和(4)式的初始值，通过迭代会分别产生一个非周期的无穷数值序列，从而对应一个无穷随机序列。该随机序列也是非周期的和类随机的，但计算机精度有限使得实际产生的序列必然具有周期性。事实上，我们无法消除周期性，而只能设法延长序列的周期。为此我们提出一种新的思想，用三阶细胞混沌神经网络系统和 Chebyshev 混沌映射来产生伪随机序列，系统的结构如图 3。

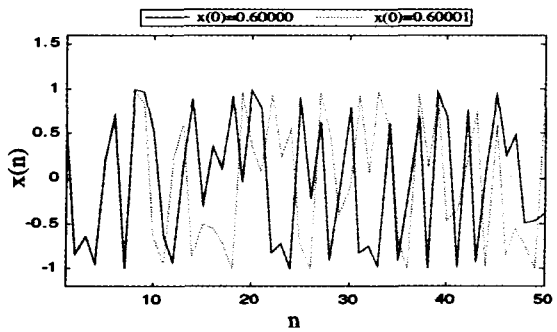

图 2　初始值相差 10^{-5} 的 Chebyshev 时间序列

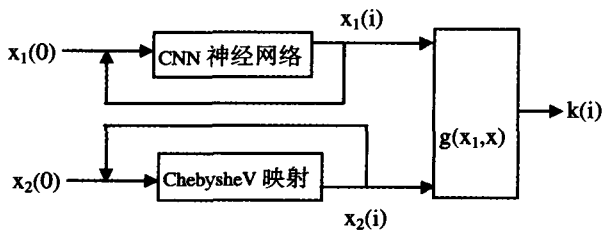

图 3　混沌伪随机序列发生器结构图

在混沌神经序列密码系统中，以系统(3)式和系统(4)式的初值作为序列密码系统的密钥，为获得较好的随机效果，混沌系统的暂态过程即初始的 N_0 次迭代不予使用。加密过程如下：

(1)输入：细胞神经网络(3)式的非线性扰动参数 a_1、S_{11}、S_{12} 和 S_{32}；Chebyshev 映射(4)式的初值 x_0；

(2)第 i 次迭代，细胞神经网络(3)式产生的值为 $x_1(i)$，Chebyshev 映射(4)式产生的值为 $x_2(i)$；

(3)若 $x_1(i)>x_2(i)$，置 $k_1=1$；若 $x_1(i)<x_2(i)$，置 $k_1=0$；若 $x_1(i)=x_2(i)$，不输出。

(4)输出：混合混沌序列$\{k_i|i=1,2,\cdots\}$。

(5)加密：得到的序列$\{k_i|i=1,2,\cdots\}$与明文$\{p_i|i=1,2,\cdots\}$进行异或运算得$\{c_i|i=1,2,\cdots\}$，即：$c_i=k_i\oplus p_i$。

细胞神经网络(3)式和 Chebyshev 映射(4)式通过时钟控制，使它们同时开始启动。解密过程与加密相同。

4　性能分析及对输出结果的数值模拟

4.1　抗干扰性能分析

首先我们分析系统的抗干扰性能。从混沌序列中任意截取一段序列记为 $S=\{s_{k1},\cdots,s_{km}\}$，设在传输中噪声 $E=\{e_1,\cdots,e_m\}$加入到输出序列中，则实际输出序列为：$R=S+E$。如果我们把混沌序列 S 和实际输出序列 R 看成 m 维向量，那么相关系数 C 等于 S 和 R 的数量积，即

$$C=S\cdot R=S\cdot S+S\cdot E$$

混沌序列的相关函数具有快速衰减特性，因此这些噪声与现有的序列是不相关的，即 $S\cdot E$ 很小，系统对于附加噪声和乘性干扰传输是鲁棒的。

4.2　输出序列性能分析

定义 1[8]　设 $a^{(1)}$，$a^{(2)}$ 分别表示长度为 N 的两个不同混沌伪随机序列。序列 $a^{(1)}$ 和 $a^{(2)}$ 的非周期互相关函数为：

$$C_{a^{(1)}a^{(2)}}(l)=\begin{cases}\dfrac{1}{N}\sum\limits_{i=0}^{N-1-l}a_i^{(1)}(a_{i+l}^{(2)})^*, & 0\leqslant l<N\\ \dfrac{1}{N}\sum\limits_{i=0}^{N-1+l}a_{i-l}^{(1)}(a_i^{(2)})^*, & 1-N\leqslant l<0\\ 0, & |l|\geqslant N\end{cases}\tag{5}$$

当 $a^{(1)}=a^{(2)}$ 时，上式表示的函数是扩频序列的非周期自相关函数。

我们对基于混沌神经网络和混沌映射的伪随机序列进行相关特性检测，其非周期自相关与互相关特性如图 4 和图 5。由实验结果看出，混沌序列具有尖锐的自相关特性和很小的互相关值。

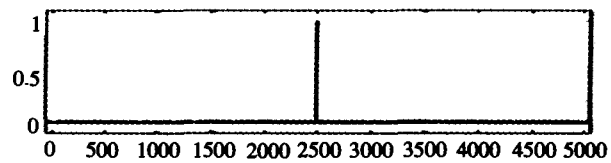

图 4　输出序列的自相关

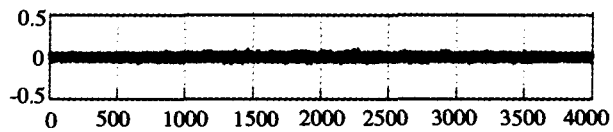

图 5　输出序列和互相关

(下转第 206 页)

吗？是，转(9)；

(3)否：$i=i+1$；读入候选结果集中的第 i 个歧义字段；

(4)利用从训练语料中得到的数据对第 i 个歧义字段，计算每个二字应成词的 $I(x:y)$、$t_{x,z}(y)$和 $\Delta t(x:y)$；

(5)$I(b_m:c_1)>I(a_i:b_1)$且 $\Delta t(b_m:c_1)>\Delta t(a_i:b_1)$或 $I(b_m:c_1)<I(a_i:b_1)$且 $\Delta t(b_m:c_1)<\Delta t(a_i:b_1)$吗？是，$I$ 和 Δt 的判断一致，按的判断结果切分，转(2)；

(6)否：$|I(b_m:c_1)-I(a_i:b_1)|\geqslant\alpha$ 吗？，是则由 I 的判断结果切分，，转(2)；

(7)$|\Delta t(b_m:c_1)-t(a_i:b_1)|\geqslant\beta$ 吗？是则由 Δt 的判断结果切分，转(2)；

(8)否：由 I 的判断结果切分，转(2)；

(9)算法结束。

4 实验与讨论

歧义字段切分查准率和查全率是评价切分歧义消除效果的两个重要指标，定义如下：

查准率＝

$$\frac{\text{正确切分的歧义字段的个数}}{\text{正确切分的歧义字段个数＋错误切分的歧义字段的个数}}\times 100\%$$

查全率＝

$$\frac{\text{正确切分的歧义字段的个数}}{\text{正确切分的歧义字段个数＋未发现的歧义字段的个数}}\times 100\%$$

用本文的方法对 200 句 MIS 检索用语进行了测试，从分词结果来看，歧义切分查准率达到 84%，查全率为 77%。说明应用互信息和 t-信息差这两个统计量对检索用语的歧义字段进行切分，提高了分词精度，本文的分词策略和歧义消除算法是可行的。

结束语 切分歧义问题是中文分词处理中的一个难点，本文提出了 MIS 汉语检索接口中分词系统的构建方法，给出了对应的分词策略。应用互信息和 t-信息差等统计消歧方法设计了切分歧义的消除算法并对算法进行了测试，取得了较好的分词效果。分词策略与消歧算法适用于 MIS 智能检索等专业性和领域性较强的应用环境。受词典及训练语料规模及领域的限制，分词系统在测试时发现了部分切分错误，说明本文方法对部分交集型歧义的发现还不够完善。同时系统需要进行大量的数据计算，时间开销较大，消歧效率还有待提高。下一步将进一步扩大训练语料的规模和覆盖范围，降低消歧算法的计算复杂度，提高分词效率。

参 考 文 献

1 Michael R G, Nils J N. Logical Foundation of Artificial Intelligence. Morgan Kaufmken Publishers, Inc, 1987
2 Nguyend T, Vindrow B. Nearal networks for Self-Learning Control Systems. IEEE CSM 1990, 10(3): 18～23
3 Christopher D. Manning, Hinrich Schutze. 统计自然语言处理基础[M]. 苑春法，等译. 北京：电子工业出版社，2005. 143～163
4 孙茂松，肖明，邹嘉彦. 基于无指导学习策略的无词表条件下的汉语自动分词[J]. 计算机学报，2004，27(6)：736～742
5 谈文蓉，杨宪泽. MIS 的智能处理的近似评判法及其算法研究[J]. 计算机科学，2005，32(3)：226～228
6 曹娟，周经野. 一种计算汉字串之间相关程度的新方法[J]. 中文信息学报，2005，18(4)：55～59
7 孙茂松，黄昌宁，等. 利用汉字二元语法关系解决汉语自动分词中的交集型歧义[J]. 计算机研究与发展，1997，34(5)：332～339
8 杨宪泽，谈文蓉，唐向阳，等. 一种混合式机器翻译方法及其算法[J]. 计算机应用与软件，2005，22(9)：142～146

(上接第 146 页)

4.3 扰乱与扩散性能分析

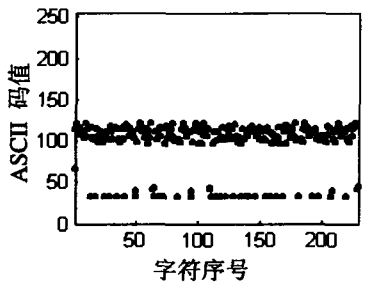

图 6 明文

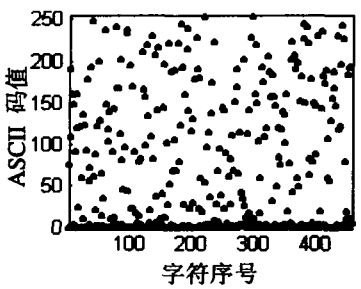

图 7 密文

扩散是将每一位明文的影响尽可能地作用到较多的输出密文位中去，同时，还要尽量使得每一位密钥的影响也尽可能迅速地扩展到较多的密文位中去。其目的是有效隐藏明文的统计特性，这也就是混沌系统的初始条件敏感依赖性。扰乱，是指密文和明文之间的统计特性的关系尽可能地复杂化，这也就是混沌映射通过迭代，将初始域扩散到整个相空间。为更清晰地描述这一特性，我们使用二维图形来表达。在图 6 和图 7 中，横轴代表信息中字符出现的序号，纵轴代表对应字符的 ASCII 码值(范围 0～255)。从明文和密文的图形来看，明文的码值比较集中，而根据本文算法所得到的密文在整个密文空间的分布都非常均匀。也就是说，通过扩散、扰乱等作用后，密文中不包含明文的任何信息(包括明文的统计概率信息)。这正是我们想要达到的加密效果。

结论 本文在分析三阶细胞神经网络和 Chebyshev 映射混沌特性的基础上，提出了基于神经网络和混沌映射的序列密码算法，给出了系统实现原理和算法描述，对混沌序列进行模拟实验和计算机仿真，同时对该系统产生的安全性能进行了分析。结果表明，这种方法设计的序列密码具有随机性好，在较低精度下序列的相关性能好，这大大降低了实现了成本。

参 考 文 献

1 van Schyndel R G, Tirkel A Z, Svalbe I D. Key independent watermark detection. IEEE International Conference on Multimedia Computing and Systems, Florence, Italy, 199: 580～585
2 Eggers J J, Su J K, Girod B. Public key watermarking by eigenvectors of linear transforms. European Signal Processing Conference, Tampere, Finland, 2000・428～435
3 丘水生，陈艳峰，吴敏，等. 混沌加密的若干问题与新的加密系统方案. 见：2002 中国非线性电路与系统学术会议论文集，中国，深圳，2002. 174～179
4 王育民. 混沌密码序列使用化问题. 西安电子科技大学学报，1997，24(4)：560～562
5 Chua L O, Roska T. The CNN paradigm. IEEE Trans. CAS-I, 1993, 40: 47～156
6 Civalleri P P, Gilli M. On dynamic behaviour of two-cell cellular neural networks. Int. J. Circ. Th. Appl., 1993, 21: 451～471
7 何振亚，张毅锋，卢宏涛. 细胞神经网络动态特性及其在保密. 通信中的应用. 通信学报，1999，20(3)：59～67
8 Tohur K, Akio T. Pseudonoise sequence by chaotic nonlinear and their correlation properties. IEICE Trans commun, 1993, E97-B (8): 855～862