细胞自动机组合伪随机序列发生器

张传武

(西南民族大学电气信息工程学院 成都 610041)

【摘要】结合细胞自动机所特有的单元结构的简单性、单元之间作用的局部性和信息处理的高度并行性等特点，利用细胞自动机产生高速序列；分析了比特与、或、异或运算周期特性，其周期等于各自周期的最小公倍数；证明了比特异或运算的频率特性优于原有的频率特性。研究了比特组合运算的线性复杂独特性，比特与、或运算的线性复杂度等于各自线性复杂度的乘积，异或运算的线性复杂度等于各自线性复杂度之和。利用伪随机特性检测方法和线性复杂度的测试方法的计算机模拟表明细胞自动机组合伪随机序列发生器实现简单、速度高、能有效增加序列周期长，改善序列伪随机统计特性，并能有效增加伪随机序列的线性复杂度。

关 键 词 细胞自动机； 组合伪随机序列发生器； 密码学； 线性复杂度； 统计特性

中图分类号 TN918.91 **文献标识码** A

Combined Pseudorandom Sequence Generator Based on Cellular Automata

ZHANG Chuan-wu

(College of Electrical and Information Engineering, Southwest University for Nationalities Chengdu 610041)

Abstract By analyzing the period of the bit computation of AND, OR, and XOR, it is proved that the frequency characters of the bit computation of XOR is better than the original frequency characters. The study of the linear complexity of the combined bits demonstrates that: the linear complexity of the bit computation of AND and OR is equal to the product of the linear complexity of the originals; and the linear complexity of the bit computation of XOR equal to the addition of the linear complexity of the originals. Computer simulation demonstrates that the combined pseudorandom sequence generator has simple architecture and high speed information processing characters, and can efficiently increase the period and linear complexity efficiently of pseudorandom sequence.

Key words cellular automata (CA); combined pseudorandom sequence generator (CPRSG); cryptography; linear complexity; statistical properties

伪随机序列发生器是利用较短的初始随机序列产生较长的伪随机序列的一种确定性方法[1]。目前广泛使用的是基于同余和线性反馈移位寄存器、以及基于它们组合的伪随机序列发生器，其中组合伪随机序列发生器具有增加序列周期、改善序列随机统计特性和提高序列线性复杂度等优点[2]。由于细胞自动机具有组成单元的简单规则性、单元之间作用的局部互连性和信息处理的高度并行性，并表现出复杂的全局特性等特点[3-4]，使其成为移位寄存器的替代品并广泛应用于密码学、通信和测试等领域[5]。所以，基于细胞自动机的组合伪随机序列发生器的研究具有重要意义。

1 细胞自动机理论

基本细胞自动机是一组如图1所示的具有一定状态$s_i \in \{0,1\}$，i=0,1,…,N–1的细胞单元组成的阵列。其每个单元的转移状态s_i^{t+1}由其相应的邻域规则f和该单元的邻域状态$(s^t_{i-1}, s^t_i, s^t_{i+1})$确定，称由$f$确定的邻域状态$(s^t_{i-1}, s^t_i, s^t_{i+1})$与转移状态$s_i^{t+1}$的映射为基本细胞自动机的规则表，并称邻域状态$(s^t_{i-1}, s^t_i, s^t_{i+1})$的映射 $\{l_7 = f(111),\cdots,l_1 = f(001),l_0 = f(000)\}$ 的组合 $I_f = \sum_{i=0}^{7} l_i 2^i$ 为基本细胞自动机的规则号[6]。如{l_7

收稿日期：2007 – 03 – 31；修回日期：2007 – 09 – 10
基金项目：国家自然科学基金(60603009)
作者简介：张传武(1971 –)，男，博士，教授，主要从事细胞自动机、信息安全和通信网等方面的研究.

$l_6\ l_5\ l_4\ l_3\ l_2\ l_1\ l_0$=01011010}对应的规则号为90，其逻辑函数表达式为 $s_i^{t+1}=s_{i-1}^t+s_{i+1}^t$，其中“+”表示模二加运算。

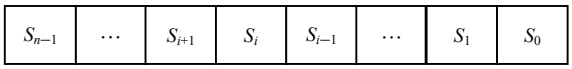

图1　细胞自动机的结构图

在细胞自动机中，不同的细胞单元可以采用不同的规则*f*，所以细胞自动机采用其细胞单元使用的规则序列来表示，如(90,90,150)表示90、90、150规则的3单元细胞自动机。

定义 1　称状态转移的映射相补的两个规则为互补规则，如规则90和由 $\{\overline{l_7 l_6 l_5 l_4 l_3 l_2 l_1 l_0}=01011010\}$ 表示的规则165。

*N*单元加性细胞自动机状态转移方程可表示为：

$$s_i^{t+1}=a_{i,-1}s_{i-1}^t+a_{i,0}s_i^t+a_{i,1}s_{i+1}^t+h_i \qquad 0<i<N \quad (1)$$

式中　边界单元 s_0 和 s_{N-1} 的缺失邻域单元的状态恒为0；$h_i=1$ 时表示第*i*个细胞单元使用的邻域函数规则为补规则，称 $h_i=0, i=0,1,\cdots,N-1$ 的加性细胞自动机为线性细胞自动机。将式(1)改为矩阵方程的形式为：

$$\boldsymbol{S}^{t+1}=\boldsymbol{T}\boldsymbol{S}^t+\boldsymbol{H} \qquad (2)$$

式中　$\boldsymbol{T}$为线性细胞自动机的状态转移矩阵；$\boldsymbol{S}^t=(s_0^t,s_1^t,\cdots,s_{N-1}^t)^{\mathrm{T}}$ 为细胞自动机在*t*时刻的全局状态配置；$\boldsymbol{H}=(h_0,h_1,\cdots,h_{N-1})^{\mathrm{T}}$ 是根据线性细胞自动机构造加性细胞自动机的补规则指示位，简称为加性细胞自动机的补规则向量。所以可以采用线性细胞自动机的规则序列构成的规则向量和补规则向量确定加性细胞自动机。如对于(60,90,150)线性细胞自动机，由它和补规则向量(0,0,1)确定了(60,90,105)加性细胞自动机，并称由(60,90,150)线性细胞自动机和其补规则向量确定的8个加性细胞自动机为(60,90,150)线性细胞自动机对应的加性细胞自动机簇。

2　发生方法

设有一组共*M*个细胞自动机，每个细胞自动机的单元数为$N_i \geqslant N_{i+1}$，*i*=0,1,…,*M*–2，状态转移方程为：

$$\boldsymbol{S}_l^t=\boldsymbol{T}_l\boldsymbol{S}_l^{t-1}+\boldsymbol{H}_l \qquad l=0,1,\cdots,M-1 \qquad (3)$$

其*t*次迭代的状态转移方程为：

$$\boldsymbol{S}_l^t=(\boldsymbol{T}_l)^t\boldsymbol{S}_l^0+\sum_{i=0}^{t-1}(\boldsymbol{T}_l)^i\boldsymbol{H}_l \qquad l=0,1,\cdots,M-1 \quad (4)$$

式中　$\boldsymbol{T}_l$为第*l*个细胞自动机的状态转移矩阵；$\boldsymbol{S}_l^t=(s_{l,0}^t,s_{l,1}^t,\cdots,s_{l,N_l-1}^t)^{\mathrm{T}}$ 为第*l*个细胞自动机在*t*时刻的状态配置；$\boldsymbol{H}_l=(h_{l,0},h_{l,1},\cdots,h_{l,N_l-1})^{\mathrm{T}}$ 为第*l*个线性细胞自动机对应的加性细胞自动机的补规则向量。组合细胞自动机在*t*时刻的全局状态为：

$$\boldsymbol{S}^t=\boldsymbol{S}_0^t\otimes\hat{\boldsymbol{S}}_1^t\otimes\cdots\otimes\hat{\boldsymbol{S}}_{M-1}^t \qquad (5)$$

式中　“$\otimes$”运算符可以分别表示对应位的与、或、异或等运算；$\hat{\boldsymbol{S}}_l^t$ 为第*l*个细胞自动机的从N_l维状态到N_0维状态的扩展，它只简单地在 $\boldsymbol{S}_l^t$ 状态的后面添加N_0–N_l个零，即 $\boldsymbol{S}_l^t=(s_{l,0}^t,s_{l,1}^t,\cdots,s_{l,N_l-1}^t,0,\cdots,0)^{\mathrm{T}}$。

将式(5)中的第 $i_0<N_{M-1},i_1<N_{M-1},\cdots,i_{k-1}<N_{M-1}$ 位的组合 $\left\{y^t\middle|y^t=\sum_{j=0}^{k-1}(s_{i_j}^t2^j),i_j\in Z_{N_{M-1}}\right\}$ 称为系统产生的伪随机数。并将其第 $i<N_{M-1}$ 位的状态 s_i^t 称为系统产生的伪随机序列：

$$\left\{y^t\middle|y^t=s_i^t=\prod_{l=0}^{M-1}\left(\sum_{n=0}^{N_l-1}([(\boldsymbol{T}_l)^t]_{i,n}s_{l,n}^0)+\sum_{n=0}^{N_l-1}\left(\sum_{j=0}^{t-1}[(\boldsymbol{T}_l)^j]_{i,n}h_{l,n}\right)\right)\right\} \qquad i\in Z_{N_{M-1}} \qquad (6)$$

式(6)便为组合细胞自动机产生伪随机序列的模型。

结论 1　对于周期为$P_0,P_1,\cdots,P_{M-1}$的*M*个独立的细胞自动机伪随机序列发生器，通过比特与、或、异或运算后的组合细胞自动机伪随机序列的周期为这*M*个细胞自动机伪随机序列发生器周期的最小公倍数LCD($P_0,P_1,\cdots,P_{M-1}$)。

证明：假设序列$S_0,S_1,\cdots,S_{M-1}$的周期分别为$P_0,P_1,\cdots,P_{M-1}$，且*S*为$S_0,S_1,\cdots,S_{M-1}$的比特与运算，其周期为*P*，则有：$\begin{cases}S^i=S_0^i\times S_1^i\times\cdots\times S_{M-1}^i\\ S^i=S^{i+P}\end{cases}$。即 $S_0^i\times S_1^i\times\cdots\times S_{M-1}^i=S_0^{i+P}\times S_1^{i+P}\times\cdots\times S_{M-1}^{i+P}$。因为$S_0,S_1,\cdots,S_{M-1}$相互独立，所以上式成立，当且仅当：$S_j^i=S_j^{i+P}, j\in(0,1,\cdots,M-1)$。而由假设有：$S_j^i=S_j^{i+l_jP_j}$，$(l_j\in Z, j=0,1,\cdots,M-1)$，所以：$S_j^{i+P}=S_j^{i+l_jP_j}$，$(l_j\in Z, j=0,1,\cdots,M-1)$。即有：$P=l_jP_j,(j=0,1,\cdots,M-1)$，即*P*为$P_0,P_1,\cdots,P_j,\cdots,P_{M-1}$的最小公倍数LCD($P_0,P_1,\cdots,P_{M-1}$)。

同理可证或、异或组合伪随机序列发生方法的周期特性。

结论 2　*M*个独立的细胞自动机伪随机序列发生器通过模二加运算组合而成的细胞自动机组合伪随机序列发生器产生的伪随机序列的频率特性(0/1出现的概率)好于等于其中任何一个细胞自动机伪随机序列发生器产生的比特序列的频率特性(0/1出现的概率)。

证明：对伪随机序列，其频率特性越接近0.5越好，由于比特序列中比特0和1的频率特性关于0.5对

称，所以使用比特1的频率特性与0.5的差的绝对值表示频率特性好坏的度量，其值越小越好。

假设CA_1和CA_2的比特序列中比特0的频率为$p_0^i=0.5+a_i, i=1,2, a_i\in[0,0.5]$，那么比特1的频率为$p_1^i=0.5-a_i, i=1,2$，由CA1和CA2模2加运算得到的序列中1的频率为(CA1与CA2独立)：

$$p_1=(0.5+a_1)(0.5-a_2)+(0.5-a_1)(0.5+a_2)=0.5-2a_1a_2$$

对CA1、CA2伪随机序列发生器和细胞自动机组合伪随机序列发生器的频率特性的量度为：

$$|p_1^i-0.5|=a_i, i=1,2，\ |p_1-0.5|=2a_1a_2$$

因为$a_1,a_2\in[0,0.5]$，所以有$2a_1a_2\leqslant a_i, i=1,2$，即两个细胞自动机组合而成的伪随机序列发生器的平衡特性好于等于其组成的任一细胞自动机伪随机序列发生器。再由两个细胞自动机组合而成的伪随机序列发生器与另一个细胞自动机伪随机发生器组合而成另外一个组合伪随机序列发生器，上述结论同样成立，以此类推到M个组合上式均成立，即定理成立。

引理 1[7] 线性复杂度为$L_0,L_1,\cdots,L_{M-1}$的M个m序列的与、或序列的线性复杂度为$L=L_0\times L_1\times\cdots\times L_{M-1}$，其中，$L_i\neq L_j, i\neq j, i\in Z_M, j\in Z_M$。

引理 2[8] 两个序列的模二加组合序列的最小多项式为两个序列的最小多项式的乘积。

从引理2可知，M个细胞自动机产生的伪随机序列的模二加序列的最小多项式是它们的乘积，即异或运算组合伪随机序列的线性复杂度为其各伪随机序列的线性复杂度之和。

结论 3 采用先比特与/或、后比特异或的组合方法可产生以周期为组合的周期的最小公倍数、随机统计特性好、线性复杂度等于各与/或序列的线性复杂度乘积再加上异或序列的线性复杂度的序列。

证明：由定理1有周期为组合序列的周期的最小公倍数；由定理2可知先比特与/或、后比特异或的组合方法产生的伪随机序列的伪随机统计特性由于其中的任何一个伪随机序列；而由引理1、引理2可知先比特与/或、后比特异或的组合方法产生伪随机序列的线性复杂度等于与/或序列的线性复杂度的乘积，再加上异或序列的线性复杂度。

以上证明了细胞自动机组合伪随机序列发生器具有增加序列周期、改善序列的频率特性等统计特性等。下面是基于细胞自动机的组合伪随机序列的产生方法：(1) 利用一组单个细胞自动机产生伪随机序列；(2) 将其中一部分采用与/或运算进行组合；(3) 将与/或运算组合后的组合序列与其他序列进行异或组合即可产生组合伪随机序列。

3 算法仿真实验

在构造细胞自动机组合伪随机序列发生器时，应采用具有最大阶群特性的细胞自动机，否则细胞自动机组合伪随机序列发生器会受初始状态的约束。单一细胞自动机一般不会有规律的出现最大阶群特性的性质，而混合细胞自动机有时会规律性的出现最大阶群特性。在Werner Pries和Peter D. Hortensius等人实验的基础上，文献[9]提出了根据特征多项式综合90和150混合细胞自动机的高效方法。

细胞自动机与/异或组合伪随机序列发生方法实验用的细胞自动机的补规则向量为全0的90/150加性细胞自动机。其规则向量分别为：30单元的0x3a26ca3d细胞自动机CA_0、31单元的0x2656fb98细胞自动机CA_1、以及32单元的0x4609bd5细胞自动机CA_2来实现组合伪随机序列发生，初始状态均为除第15位为1外其余位均为0。输出先使用CA_0和CA_2的第15位进行与运算，以提高组合序列的线性复杂度，然后将与组合序列与CA_1的第15位产生的伪随机序列进行异或运算，以提高组合序列的随机统计特性。

实验仿真的平台是Windows2000和Turbo C 2.0，并使用Matlab绘制。序列伪随机性的统计测试采用符合美国联邦信息处理标准的经典五项测试准则[10]，其每次产生的n=20 000 bit的组合伪随机序列的统计特性和线性复杂度特性测试如图2～图7所示。

图2～图7的仿真测试说明组合伪随机序列的频率测试、Poker测试、游程测试和自相关测试均满足伪随机序列的特性，同时线性复杂度测试说明组合方法可以有效增加组合伪随机序列的线性复杂度。

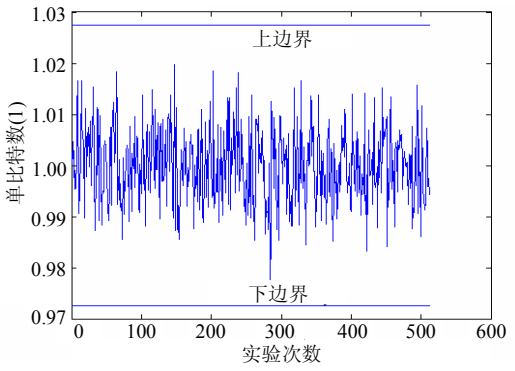

图2 频率测试(比特1)

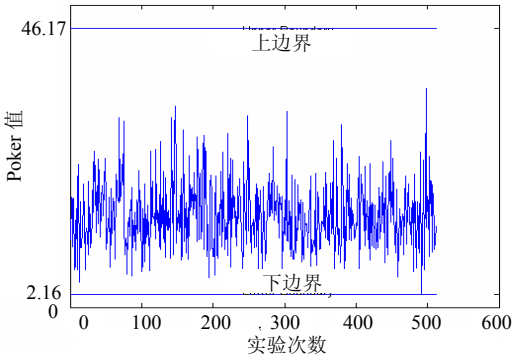

图3　Poker测试

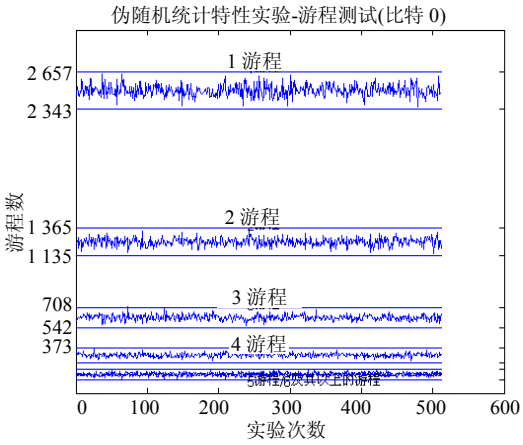

图4　游程测试(比特0)

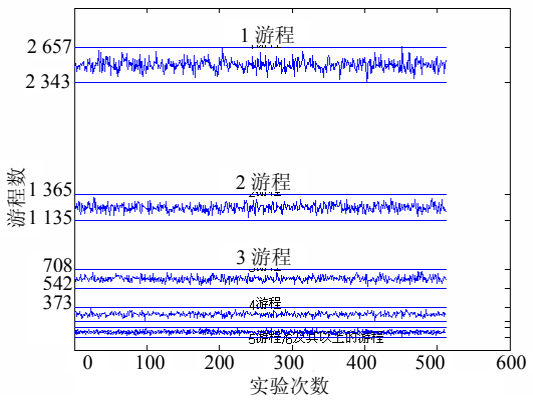

图5　游程测试(比特1)

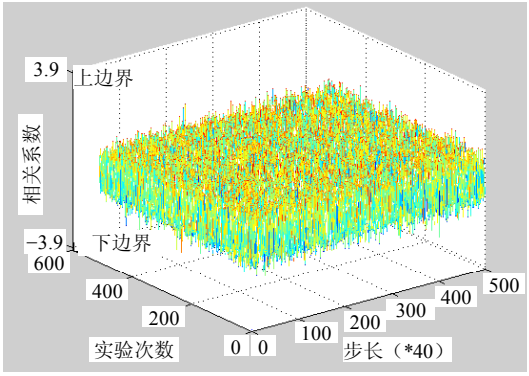

图6　自相关测试

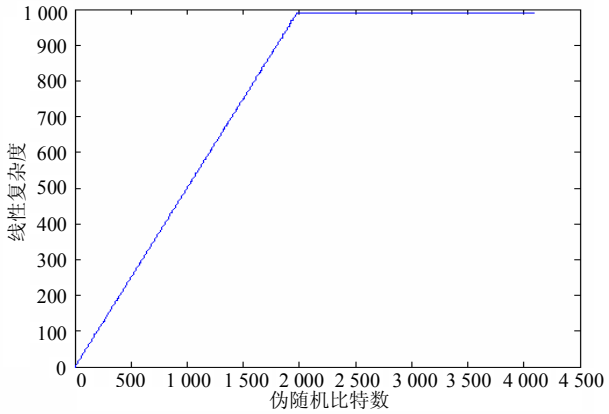

图7　线性复杂度测试

4　结 束 语

本文的研究表明：基于细胞自动机的组合伪随机序列发生器不仅具有产生序列周期长、统计特性好等优点，而且具有成本低、实现简单和处理速度快等特点。细胞自动机组合伪随机序列发生器技术将是取代基于反馈移位寄存器的伪随机序列发生器技术。

参 考 文 献

[1] GUAN S U, TAN S K. Pseudorandom number generation with self-programmable cellular automata[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2004, 23(7): 1095-1101.

[2] PIERRE L E. Efficient and portable combined random number generators[J]. Communications of the ACM, 1988, 31(6): 742-774.

[3] ROSIN P L. Training cellular automata for image processing [J]. IEEE Transactions on Image Processing, 2006, 15(7): 2076-2087.

[4] TOMASSINI M, SIPPER M, PERRENOUD M. On the generation of high-quality random numbers by two-dimensional cellular automata[J]. IEEE Transactions on Computers, 2000, 49(10): 1146-1151.

[5] 张传武, 沈野樵, 彭启琮. 细胞自动机反向迭代加密技术研究[J]. 计算机学报, 2004, 27(1): 125-129.
ZHANG Chuan-wu, SHEN Ye-qiao, PENG Qi-cong. Encryption based on cellular automata inverse iteration[J]. Chinese Journal of Computers, 2004, 27(1): 125-129.

[6] 张传武. 加性细胞自动机同构性分析[J]. 电子科技大学学报, 2006, 35(5): 774-777.
ZHANG Chuan-wu. The homogeneous characteristic of additive cellular automata[J]. Journal of University of Electronic Science and Technology of China, 2006, 35(5): 774-777.

[7] RAINER A R, OTHMAR J S. Products of linear recurring sequences with maximum complexity[J]. IEEE Transactions on Information Theory, 1987, IT-33(1): 124-131.

[8] 万哲先. 代数和编码[M]. 北京: 科学出版社, 1980: 488-499.
WAN Zhe-xian. Algebra and coding[M]. Beijing: Science Press, 1980: 488-499.

[9] CHO S J, CHOI U S, KIM H D, et al. New synthesis of one-dimensional 90/150 linear hybrid group cellular automata[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2007, l26(9): 1720-1724.

[10] NIST. FIPS PUB 140-2 security requirements for cryptographic modules[S]. 1999.

编　辑　漆　蓉