第 16 卷 第 5 期　　电 子 科 学 学 刊　　Vol. 16, No. 5
1994 年 9 月　　JOURNAL OF ELECTRONICS　　Sept. 1994

纠错码用于局部随机序列发生器*

杨 义 先
(北京邮电大学信息工程系 北京 100088)

摘要　本文证明了好的线性分组码的编码器可作为好的线性局部随机序列发生器（具体含义见定理 2 之后的说明）. 此结论再一次揭示了纠错编码理论与现代密码学之间的有机联系.

关键词　纠错编码,密码,随机序列

1 引　　言

设 S_n 表示由长度为 n 的所有二进制向量所组成的集合，即 $S_n=\{0,1\}^n$. 所谓 (k,n) 序列发生器[1]就是一个从 S_k 到 S_n 的映射,它将每个长度为 k 的二进制向量扩张成长度为 n 的二进制向量 ($k\leqslant n$). 比如一般分组纠错码[2,3]的编码器便可看作一种特殊的 (k,n) 序列发生器，因为一个长度为 k 的信息向量经过编码之后便被扩张成为一个长度为 $n(n>k)$ 的码字向量.在现代密码学中,特别是在密钥分散管理系统中,比较有用的 (k,n) 序列发生器是这样的: k 维随机向量经过它扩张之后所得到的 n 维向量输出仍然具有良好的随机性.此类序列发生器可严格地用如下定义来描述(见文献[1]中的定义5).

定义 1　(k,n) 序列发生器 G 称为一个 (k,n,e,δ) 局部随机序列发生器,如果当输入序列是 k 维二进独立均匀分布随机向量时，它所对应的 n 维输出向量中任意 e 个分量所组成的随机子向量 S 满足 $H(S)\geqslant e-\delta>0$. 这里 $H(S)$ 是熟知的信息熵函数[4],即 $H(S)=-\sum_a P(S=a)\log P(S=a)$; 前述的二进均匀分布随机变量 x 意指 $P(x=0)=P(x=1)=1/2$，由相互独立的 k 个此种随机变量组成的 k 维向量便是上述的 k 维二进独立均匀分布随机向量.

在定义 1 中,当 $\delta=0$ 时，$(k,n,e,0)$ 局部随机序列发生器又简称为 (k,n,e) 最佳局部随机序列发生器，此时输出序列中任意 e 个分量所组成的向量 S 具有最好的随机特性,因为其熵函数达到了最大值,即 $H(S)=e$. Maurer 和 Massey 已在文献[1]中对最佳局部随机序列发生器进行了研究. 本文将研究一般的 (k,n,e,δ) 局部随机序列发生器的设计问题. 单从随机性要求来看，$\delta=0$ 是最好的情况. 但是当 $\delta=0$ 时,若 $e>k$ 或 $1<e=k<n-1$，那么就根本不存在 (k,n,e) 最佳局部随机序列发生器[1]，所以

1993-05-14 收到,1993-12-17 定稿
* 国家教委跨世纪优秀人才基金资助课题.
杨义先　男,1961 年生，教授、博士生导师，现从事信号论、编码、密码、现代通信与人工神经网络等方面的研究.

$\delta > 0$ 的情况也值得研究. 本文主要以熟知的纠错码为研究工具，证明了好的纠错码可以设计出好的随机序列发生器(具体含义见定理 2 的说明)，从而再一次揭示了纠错码理论与现代密码学之间的有机联系. 关于纠错码方法在密码学中的其它应用实例可参见文献[3,5,6].

2 用纠错码设计随机序列发生器

本文仅限于考虑线性随机序列发生器，非线性情况不在此考虑.

定义 2 (k,n) 序列发生器 G 称为一个线性序列发生器当且仅当对任意 $a,b \in S_k$ 都成立 $G(a \oplus b) = G(a) \oplus G(b)$，否则称 G 为一个非线性序列发生器. 这里"$\oplus$"表示向量对应分量的模 2 加，也称为并元和.

显然从定义 2 知线性序列发生器实际上就是一个线性分组码的编码器，它由一个 $k \times n$ 阶二进矩阵(在纠错编码中称为生成矩阵)来描述. 长度为 k 的序列 $a = (a_1, \cdots, a_k)$ 按照线性分组码的编码方式扩张成为一个长度为 n 的序列 $c = (c_1, \cdots, c_n)$，即 $c = aA$. 下面的定理 1 给出了线性 (k,n,e,δ) 局部随机序列发生器的一个等价定义.

定理 1 由 $k \times n$ 阶矩阵 A 确定的 (k,n) 线性序列发生器是一个线性 (k,n,e,δ) 局部随机序列发生器的充要条件是由矩阵 A 中任意 e 个列组成的 $k \times e$ 阶子矩阵 A^* 的秩不小于 $e - \delta$，即 $\mathrm{Rank}(A^*) \geqslant e - \delta$.

证明 设 $x = (x_1, \cdots, x_k)$ 是 k 维二进独立均匀分布随机向量，

$$y = (y_1, \cdots, y_n) = xA$$

是对应于输入序列 x 的输出序列. 任取 e 个正整数. $1 \leqslant i_1 < i_2 < \cdots < i_e \leqslant n$，将 y 中由第 $i_1, i_2, \cdots, i_e$ 个分量所组成的 e 维向量记为 $y^* = (y_1^*, \cdots, y_e^*)$；由矩阵 A 中第 $i_1, i_2, \cdots, i_e$ 列所组成的 $k \times e$ 阶子矩阵记为 A^*. 为证明此定理，只需证明成立等式 $\mathrm{Rank}(A^*) = H(y^*)$.

不失一般性，假设 $\mathrm{Rank}(A^*) = r$ 并且矩阵 A^* 的第 $1,2,\cdots,r$ 行线性独立.

为求熵函数 $H(y^*)$ 的值，先求随机向量 y^* 的概率分布. 任取一个 e 维二进向量 $a = (a_1, \cdots, a_e)$，当向量 a 不包含在由 A^* 的行向量所张成的 r 维线性空间中，即 $a \notin L(A^*)$，这时显然有 $P(y^* = a) = 0$. 以下假定 $a \in L(A^*)$，即向量 a 位于由 A^* 的行向量所张成的线性空间中. 由于已假定 A^* 的前面 r 个行线性独立，所以

$$\begin{aligned}
P(y^* = a) &= P(xA^* = a) \\
&= \sum_b P((x_{r+1}, \cdots, x_k) = (b_1, \cdots, b_{k-r})) \\
&\quad \times P[xA^* = a \mid (x_{r+1}, \cdots, x_k) = (b_1, \cdots, b_{k-r})] \\
&= \sum_b \frac{1}{2^{k-r}} P[xA^* = a \mid (x_{r+1}, \cdots, x_k) = (b_1, \cdots, b_{k-r})] \\
&= \sum_b \frac{1}{2^{k-r}} P[(x_1, \cdots, x_r) = (d_1, \cdots, d_r)] \\
&= \sum_b \frac{1}{2^{k-r}} \cdot \frac{1}{2^r}
\end{aligned}$$

$$= \frac{1}{2^r}. \tag{1}$$

(1) 式中倒数第 3 个等号右边二进向量 $(d_1,\cdots,d_r)$ 是由向量 α 和 $(b_1,\cdots,b_{k-r})$ 共同确定的. 于是

$$\begin{aligned} H(y^*) &= -\sum_{\alpha} P(y^*=\alpha)\log P(y^*=\alpha) \\ &= -\sum_{\alpha\in L(A^*)} P(y^*=\alpha)\log P(y^*=\alpha) \\ &= \sum_{\alpha\in L(A^*)} r/2^r = r \end{aligned} \tag{2}$$

(2) 式中最后一个等号是由于 $L(A^*)$ 是 r 维线性空间，它含有 2^r 个二进向量.

证毕

众所周知，在纠错编码理论中一个 (n,k) 线性分组码 C 既可以由 $k\times n$ 阶生成矩阵 G 确定又可以由 $(n-k)\times n$ 阶一致校验矩阵 H 确定. 一个 n 维二进向量

$$\mathbf{c}=(c_1,\cdots,c_n)$$

是一个码字 (即 $\mathbf{c}\in C$) 的充要条件是 $\mathbf{c}H^T=\mathbf{0}$. 由此可知，若记 d 为码 C 的码间最小Hamming 距离，那么当 $e<d$ 时矩阵 H 中的任意 e 列所组成的子矩阵的秩为 e；当 $e\geqslant d$ 时矩阵 H 中的任意 e 列所组成的子矩阵的秩均不小于 $d-1$. 结合定理 1 立即可得如下定理.

定理 2　设 H 是某个 (n,k,d) 线性分组码的一致校验矩阵. 那么由 H 所确定的线性 $(n-k,n)$ 序列发生器既可以用作一个线性 $(n-k,n,e,0)$ (若 $e<d$) 局部随机序列发生器，又可以用作一个线性 $(n-k,n,e,e-d+1)$ (若 $e\geqslant d$) 局部随机序列发生器.

需要指出的是在某些特例中，上述定理 2 中 $(n-k,n,e,\delta)$ 局部随机序列发生器的 δ 值还可以进一步减小，但是目前还未找到其一般性规律，此问题有待进一步研究.

例　(7, 4, 3) Hamming 码是最常见的线性分组码，它的一致校验矩阵为如下 3×7 阶矩阵 H，即

$$H=\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \tag{3}$$

若直接利用定理 2，那么我们只能断定由此 H 矩阵确定的(3,7)线性序列发生器可以用作一个(3,7,4,2)线性局部随机序列发生器，即 $e=4$ 时可取 $\delta=2$. 但是经过直接验证后可以将上述 δ 值降低为 1，即由 (7,4,3) Hamming 码可以设计出一个线性(3,7,4,1)局部随机序列发生器，因为由 H 矩阵的任意 4 列所组成的 3×4 阶子矩阵的秩均为 3(>2)，再结合定理 1 就知上述结论正确.

由定理 2 立即可知由码间最小距离大的纠错码可以设计出 δ 值小的 (k,n,e,δ) 线性局部随机序列发生器，即由纠错能力越强的线性分组码可以设计出随机性能越好的线性局部随机序列发生器. 此结论再一次揭示了纠错编码理论与现代密码学之间的密切联系. 相信随着今后研究的继续深入将会发现彼此间的更多联系.

定理 3 设 $h(x)$ 是二值熵函数,即 $h(x)=-x\log x-(1-x)\log(1-x)$,这里 $\log(\cdot)$ 是以 2 为底的对数函数,那么以下结论成立:

(1) 当 $e\leqslant k/\log(n-\delta)$ 时,一定存在 (k,n,e,δ) 线性局部随机序列发生器;

(2) 当 $e\leqslant(n+1-\delta)/2$ 并且 $h[(e-1)/(n-\delta-1)]<k/(n-\delta-1)$ 时,也一定存在线性 (k,n,e,δ) 局部随机序列发生器;

(3) 当 $(n-\delta\lfloor e/(\delta+1)\rfloor)h[\lfloor e/(2(\delta+1))\rfloor/(n-\delta\lfloor e/(\delta+1)\rfloor)]\geqslant k+1/2+\log(n-\delta\lfloor e/(\delta+1)\rfloor)/2$ 时,不存在任何线性 (k,n,e,δ) 局部随机序列发生器. 此处 $\lfloor x\rfloor$ 意指不超过 x 的最大整数.

证明 由于当 $n>\delta+1$ 时,恒成立

$$\sum_{i=0}^{e-1}\binom{n-\delta-1}{i}<(n-\delta)^e, \tag{4}$$

所以由 $e\leqslant k/\log(n-\delta)$ 推出 $\sum_{i=0}^{e-1}\binom{n-1-\delta}{i}<2^k$.

又由于下面这个熟知的不等式[7]

$$2^{nh(t/n)}/\sqrt{2n}\leqslant\binom{n}{t}\leqslant\sum_{i=0}^{t}\binom{n}{i}\leqslant 2^{nh(t/n)},\quad(t\leqslant n/2). \tag{5}$$

所以由条件 $e\leqslant(n+1-\delta)/2$ 和 $h[(e-1)/(n-1-\delta)]<k/(n-1-\delta)$ 也可以推出 $\sum_{i=0}^{e-1}\binom{n-1-\delta}{i}<2^k$.

因此下面我们可以在不等式 $\sum_{i=0}^{e-1}\binom{n-1-\delta}{i}<2^k$ 成立的条件下证明结论(1)和结论(2).

由定理 1 可知,若能找到一个 $k\times n$ 阶二进矩阵 H 使其任意 e 个列所组成的子矩阵的秩均不小于 $e-\delta$,那么就存在线性 (k,n,e,δ) 局部随机序列发生器. 下面就来构造此种 $k\times n$ 阶二进矩阵 H.

首先取一个非零 k 维二进列向量 $\boldsymbol{a}$,矩阵 H 的前 $\delta+1$ 列均取为 $\boldsymbol{a}$. 现在假设矩阵 H 的前面 $i(i\leqslant n-1)$列均已被选定并且满足由其中任意 $r(r\leqslant e)$列所组成的 $k\times r$ 阶子矩阵的秩均不小于 $r-\delta$. 由于矩阵 H 的前面 i 列中不超过 $e-1$ 个列的任意线性组合最多能得到 $\sum_{j=0}^{e-1}\binom{i-\delta}{j}$ 个相异的 k 维列向量(这是因为以下两个事实:(a)两个相同的二进向量之并元和为零向量,(b)矩阵 H 的前面 $\delta+1$ 列都取同一个向量),所以根据如下不等式

$$\sum_{j=1}^{e-1}\binom{i-\delta}{j}\leqslant\sum_{j=0}^{e-1}\binom{n-1-\delta}{j}<2^k$$

可知,我们能找到一个 k 维列向量 $\boldsymbol{b}$ 使其不是 H 的前面 i 列中某组不超过 $e-1$ 个列的线性组合. 将列向量 $\boldsymbol{b}$ 放入矩阵 H 中作为它的第 $i+1$ 列.显然在矩阵 H 的前面 $i+1$

列中仍然满足任意 $r(r \leqslant e)$ 个列所组成的 $k \times r$ 子矩阵的秩不小于 $r-\delta$. 重复上述过程，便可以在条件 $\sum_{i=0}^{e-1}\binom{n-1-\delta}{i} < 2^k$ 之下最终找到一个 $k \times n$ 阶矩阵使得其中任意 e 个列所组成的 $k \times e$ 阶子矩阵的秩不小于 $e-\delta$. 从而结论 (1) 和结论 (2) 被证明.

现在来证明结论(3). 首先由不等式(5)式可以从已知条件

$$\left(n-\delta\left\lfloor\frac{e}{\delta+1}\right\rfloor\right) h\left(\left\lfloor\frac{e}{2(\delta+1)}\right\rfloor\Big/\left(n-\delta\left\lfloor\frac{e}{\delta+1}\right\rfloor\right)\right)$$

$$\geqslant k+1/2+\log\left(n-\delta\left\lfloor\frac{e}{\delta+1}\right\rfloor\right)\Big/2$$

推出如下不等式

$$\sum_{i=1}^{\lfloor e/(2(\delta+1))\rfloor}\binom{n-\delta\lfloor e/(\delta+1)\rfloor}{i} \geqslant 2^k. \tag{6}$$

设 A 是任意一个 $k \times n$ 阶二进矩阵. 下面我们将从矩阵 A 中找出某 e 个列使得由它们组成的子矩阵的秩不超过 $e-\delta-1$，从而证明结论(3)正确.

第 1 步　由不等式(6)式可推出

$$\sum_{i=1}^{\lfloor e/(2(\delta+1))\rfloor}\binom{n}{i} \geqslant \sum_{i=1}^{\lfloor e/(2(\delta+1))\rfloor}\binom{n-\delta\lfloor e/(\delta+1)\rfloor}{i} \geqslant 2^k,$$

所以在矩阵 A 中或者可以找到一组个数不超过 $\lfloor e/(2(\delta+1))\rfloor$ 列的列向量使其并元和为零向量；或者可以找到某两组列向量使得每组中向量个数都不超过 $\lfloor e/(2(\delta+1))\rfloor$ 并且这两组向量的并元和相同. 一句话，我们一定可以在矩阵 A 中找到 $\lfloor e/(\delta+1)\rfloor$ 个列向量使得它们是线性相关的向量组.

第 2 步　记矩阵为 A_1 为从矩阵 A 中去掉第 1 步中所选出的那 $\lfloor e/(\delta+1)\rfloor$ 个列向量之后得到的 $k \times (n-\lfloor e/(\delta+1)\rfloor$ 阶矩阵.

由不等式(6)式可知

$$\sum_{i=1}^{\lfloor e/(2(\delta+1))\rfloor}\binom{n-\lfloor e/(\delta+1)\rfloor}{i} \geqslant \sum_{i=1}^{\lfloor e/(2(\delta+1))\rfloor}\binom{n-\delta\lfloor e/(\delta+1)\rfloor}{i} \geqslant 2^k$$

与第 1 步相似，我们又可以从矩阵 A_1 中找到 $\lfloor e/(\delta+1)\rfloor$ 个列向量使得它们是线性相关向量组.

与此相同的步骤反复进行，一直到如下的第 $\delta+1$ 步.

第 $\delta+1$ 步　记矩阵 A_δ 为从矩阵 $A_{\delta-1}$ 中去掉第 δ 步中所选出的 $\lfloor e/(\delta+1)\rfloor$ 个线性相关列向量之后得到的 $k \times (n-\delta\lfloor e/(\delta+1)\rfloor)$ 阶矩阵. 由不等式(6)式，仿第 1 步的过程，我们可以从矩阵 A_δ 中选出 $\lfloor e/(\delta+1)\rfloor$ 个线性相关的列向量.

现在将第 1 步至第 $\delta+1$ 步中所选出的全部 $(\delta+1)\lfloor e/(\delta+1)\rfloor$ 个列向量和矩阵 A 中的任意其它 $e-(\delta+1)\lfloor e/(\delta+1)\rfloor$ 个列向量组成的 $k \times e$ 阶子矩阵，此子矩阵的秩显然不超过 $e-\delta-1$.　　证毕

参 考 文 献

[1] Maurer U, Massey J.J. Cryptology, 1991, 4(2): 135—149.
[2] 王新梅. 纠错码与差错控制. 北京: 人民邮电出版社,1989,第2章.
[3] 杨义先,林须端. 编码密码学. 北京: 人民邮电出版社,1992,第2章.
[4] 周炯槃. 信息论基础. 北京: 人民邮电出版社,1983,第4章.
[5] Yang Yixian. Electron. Lett., 1988, 24(3): 154—156.
[6] 王新梅. 通信学报,1986,9(5): 1—6.
[7] Wozencraft J, Reiffen B. Sequential decoding. Cambridge, MA: MIT Press, 1961, Chapter 5.

APPLICATION OF ERROR-CORRECTING CODES TO DESIGNING LOCAL-RANDOM SEQUENCE GENERATORS

Yang Yixian

(Beijing University of Posts and Telecommunications, Beijing 100088)

Abstract This paper proves that a good linear block error-correcting code implements a good linear local-random sequence generator. Therefore the close relation between error-correcting coding theory and modern cryptography is discovered once again.

Key words Error-correcting coding, Cryptography, Random sequences