On the Linear Span of a Binary Sequence Family with Large Size and Low Correlation

Jun CHEN [1,2], Yun CHEN [2]

[1]*Lab of Information Coding & Transmission, Southwest Jiaotong University, Chengdu, China*

[2]*Information Security Institute, Chengdu University of Information Technology, Chengdu, China*

Email: chenjun@cuit.edu.cn, chy@cuit.edu.cns

Abstract: For positive number $n=2m$, a binary sequence family with period 2^n-1, family size 2^{3n} and maximum out-of-phase correlation value $6\cdot 2^{n/2}-1$ is proposed, and then an upper bound of the linear spans of these sequences is given. The linear spans of the new sequences are proved to be larger than those of the previously known large families of sequences, and the exact value of the linear span of each sequence are determined, when $n\equiv 2 \bmod 6$, $r=(2^{n/2-1}-1)/7$.

Keywords: pseudorandom sequence; linear span; low cross-correlation; d-form sequence

一类大集合二元低相关序列集的线性复杂度

陈　俊 [1,2]，陈　运 [2]

1. 西南交通大学信息编码与传输实验室，成都，中国，610031

2. 成都信息工程学院信息安全研究所，成都，中国，610225

Email: chenjun@cuit..edu.cn, chy@cuit.edu.cn

摘　要: 对正整数 $n=2m$，该文构造了一类周期为 2^n-1，序列数目为 2^{3n}，相关函数最大边峰值为 $6\cdot 2^{m/2}-1$ 的二元伪随机序列集并给出了其线性复杂度的上界。当 $n\equiv 2 \bmod 6$，$r=(2^{n/2-1}-1)/7$ 时，证明了该序列集中的序列的线性复杂度大于已知的拥有大集合容量(family size)的二元序列集的线性复杂度，并给出了每条序列的线性复杂度的精确值。

关键词: 伪随机序列；线性复杂度；低相关性；d-型序列

1 引言

伪随机序列被广泛应用于通信领域和密码学领域。在码分多址通信中，伪随机序列的数目越多，通信系统的容量就越大，从而可以降低通信的拥塞率和掉线率；当序列间的相关函数值较小时，可降低来自同一信道中其它用户的干扰。另外，当伪随机序列具有较大的线性复杂度时，可以有效的抵抗基于Berlekamp-Massey 算法的攻击，提高系统的安全性[11]。特别是在流密码系统中，伪随机序列的线性复杂度是衡量系统安全的重要标准之一。因此，设计出具有大线性复杂度和大集合容量(family size)的低相关序列集具有重要的意义。

文献[1],[2],[3],[4],[5],[6],[7],[8]中分别构造出了具有大集合容量和较好相关特性的二元伪随机序列集，但它们的线性复杂度都很低。

该文构造了一类具有大集合容量的伪随机序列集，证明了当 r 为一般值时，序列线性复杂度的理论上界。当参数 $r=(2^{n/2}-1)/7$，$n\equiv 2 \bmod 6$ 时，证明了该序列集中的序列都拥有较大的线性复杂度，且给出了线性复杂度的精确值，其最大值远大于几类已知序列[1]-[8]的线性复杂度。

资助信息：国家自然科学基金资助项目(60873216)

2 序列集的构造

设 $S=\{s_i \mid 0\le i\le M-1\}$ 是由 M 条周期为 N 的二元序列组成的序列集，其中

$$s_i=\{s_i(t)\}_{t=0}^{N-1}, s_i(t)\in\{0,1\}.$$

序列 s_i 和 s_j 的周期相关函数定义为：

$$R_{i,j}(\tau)=\sum_{t=0}^{N-1}(-1)^{s_i(t)+s_j(t+\tau)}.$$

其中 $0\le i,j\le M-1$， $0\le\tau<N$， $t+\tau$ 是模 N 加。

序列集 S 的周期相关函数的最大边峰值 $R_{\max}$ 定义为 $R_{\max}=\max\{|R_{i,j}(\tau)|\,i\ne j$ 或 $\tau\ne 0\}$。

令 $GF(2^n)$ 表示含有 2^n 个元素的有限域。

设正整数 n,m,e 满足 $n=me$，定义从 $GF(2^n)$ 到 $GF(2^m)$ 的迹函数为

$$tr_m^n(x)=\sum_{k=0}^{e-1}x^{2^{mk}}$$

其中 $x\in GF(2^n)$。

迹函数的性质参见文献[9]。

引理 1 如果正整数 $m\equiv 1 \bmod 3$，那么 $\gcd(2^m-1,(2^{m-1}-1)/7)=1$。

证明：因为当 $m\equiv 1 \bmod 3$ 时，$(2^{m-1}-1)/7$ 是整数，而 $\gcd(2^m-1,2^{m-1}-1)=1$，所以引理 1 成立。证毕。

以下总令 $n=2m$ 是正偶数，α 为有限域 $GF(2^n)$ 的一个本原元，$\xi,\eta,\gamma\in GF(2^n)$，$\gcd(r,2^m-1)=1$。

定义序列集

$$\Xi=\{s_{\xi,\gamma,\eta}\,|\,\xi,\gamma,\eta\in GF(2^n)\},$$

其中

$$s_{\xi,\gamma,\eta}=\{s_{\xi,\gamma,\eta}(t)=tr_1^m\{[tr_m^n(\alpha^t+\xi\alpha^{(2^{m+1}-1)t}+\gamma\alpha^{(3\cdot 2^m-2)t})+\eta\alpha^{(2^{m+2}-3)t}]^r\}\}_{t=0}^{2^n-2}.$$

定理 1[6] 序列集 Ξ 的周期相关函数的最大边峰值为 $6\cdot 2^{n/2}-1$。

定理 2[6] 序列集 Ξ 中共有 2^{3n} 个不等价的序列。

3 序列的线性复杂度

令 $LS(s_{\xi,\gamma,\eta})$ 为序列 $s_{\xi,\gamma,\eta}$ 的线性复杂度。根据文献[10]知，若将 $s_{\xi,\gamma,\eta}(t)$ 表示成 α^t 的多项式，则 $s_{\xi,\gamma,\eta}$ 的线性复杂度就等于该多项式中包含的 α^t 的单项式的数目。若令 $x=\alpha^t$，则 $s_{\xi,\gamma,\eta}(x)=tr_1^m\{[tr_m^n(x+\xi x^{2^{m+1}-1}+\gamma x^{3\cdot 2^m-2}+\eta x^{2^{m+2}-3}]^r\}$ 的展开式中关于 x 的单项式的个数就等于 $LS(s_{\xi,\gamma,\eta})$。

令 $y=x^{2^m-1}$，则

$$\begin{aligned}s_{\xi,\gamma,\eta}(x)&=tr_1^m\{[tr_m^{2m}(x+\xi x^{2^{m+1}-1}+\gamma x^{3\cdot 2^m-2}+\eta x^{2^{m+2}-3}]^r\}\\&=tr_1^m\{[xy^{-3}(\eta^{2^m}+\gamma^{2^m}y+\xi^{2^m}y^2+y^3+y^4+\xi y^5+\gamma y^6+\eta y^7)]^r\}\\&=\sum_{k=0}^{m-1}[xy^{-3}(\eta^{2^m}+\gamma^{2^m}y+\xi^{2^m}y^2+y^3+y^4+\xi y^5+\gamma y^6+\eta y^7)]^{2^k r}.\end{aligned}$$

若令 $\Delta_k(x)=[xy^{-3}(\eta^{2^m}+\gamma^{2^m}y+\xi^{2^m}y^2+y^3+y^4+\xi y^5+\gamma y^6+\eta y^7)]^{2^k r}$。

则有如下的引理。

引理 2 若 $k\ne k'$，则 $\Delta_k(x)$ 的展开式中关于变量 x 的单项式的指数与 $\Delta_{k'}(x)$ 的展开式中关于变量 x 的单项式的指数互不相同。

证明：因为 $y=x^{2^m-1}$，所以 $\Delta_k(x)$ 的展开式中，关于变量 x 的指数模 2^m-1 同余 $2^k r$。如果存在 k,k' 使得 $2^k r\equiv 2^{k'}r \bmod 2^m-1$，那么因为 $\gcd(r,2^m-1)=1$，所以有 $2^k\equiv 2^{k'} \bmod 2^m-1$，从而 $k=k'$。证毕。

若用 $|\Delta_k(x)|$ 表示 $\Delta_k(x)$ 中关于 x 的单项式的数目，则由引理 2 可知，

$$LS(s_{\xi,\gamma,\eta})=m\cdot|\Delta_0(x)|. \tag{1}$$

令 $\Gamma_{\xi,\gamma,\eta}(y)=(\eta^{2^m}+\gamma^{2^m}y+\xi^{2^m}y^2+y^3+y^4+\xi y^5+\gamma y^6+\eta y^7)^r$ (2)

则有

$$LS(s_{\xi,\gamma,\eta})=m\cdot|\Gamma_{\xi,\gamma,\eta}(y)| \tag{3}$$

因此只需计算 $\Gamma_{\xi,\gamma,\eta}(y)$ 的展开式中，关于 y 的单项式的数目。

引理 3 当 $r=(2^{m-1}-1)/7$ 时，$\Gamma_{\xi,\gamma,\eta}(y)$ 的展开式中，关于 y 的指数互不相同，且互不相同的指数的数目分别为 $2^{(m-1)/3}$，$4^{(m-1)/3}$，$6^{(m-1)/3}$ 和 2^{m-1}。

证明：令 b 表示 $\Gamma_{\xi,\gamma,\eta}(y)$ 的展开式中 y 的指数。

因为 $r=(2^{m-1}-1)/7=1+2^3+\cdots+2^{m-4}$，所以指数 b 的形式为

$$b=\sum_{j=0}^{(m-4)/3}t_j 2^{3j}=\sum_{j=0}^{(m-4)/3}t_j 8^j,\ t_j\in\{0,1,\cdots,7\}.$$

容易知道 b 为某个正整数的 8 进制表示，并且有 $7b\le 2^{m-1}-1$，从而可知 y 的指数互不相同。因此当 $\xi\cdot\gamma\cdot\eta\ne 0$ 时，(2)式的展开式中共有 $8^{(m-1)/3}=2^{m-1}$ 个指数互不相同的 y 的单项式；而当 $\xi=0,\gamma=0,\eta=0$ 时，共有 $2^{(m-1)/3}$ 个指数互不相同的 y 的单项式；当 ξ,γ,η 三者中只有一个为零时，共有 $6^{(m-1)/3}$ 个指数互不相同的 y 的单项式；其它情况下，关于 y 的单项式的数目为 $4^{(m-1)/3}$。证毕。

由引理 2、引理 3 以及(1)式和(3)式，可得定理 3。

定理 3 当 $n=2m,m\equiv 1 \bmod 3,r=(2^{m-1}-1)/7$ 时，序列 $s_{\xi,\gamma,\eta}$ 的线性复杂度分别为 $n\cdot 2^{(n-8)/6}$，$n\cdot 2^{(n-5)/3}$ 以及 $n\cdot 3^{(n-2)/6}\cdot 2^{(n-8)/6}$ 和 $n\cdot 2^{n/2-2}$。

定理 4 当 r 为一般值时，集合 Ξ 中序列的线性复杂度最大为 $n(2^{n/2}+1)/2$。

证明：因为在(2)式中，$y^{2^m+1}=1$，所以在(2)式的

978-1-935068-10-5 © 2010 SciRes.

展开式中，y 的指数属于集合 $\{0,1,2,\cdots,2^m\}$，故最多有 2^m+1 个互不相同的指数。根据(2)和(3)式，Ξ 中序列的最大线性复杂度为 $m(2^m+1)=n(2^{n/2}+1)/2$。证毕

集。从中可以看出本文构造的序列的具有很大的线性复杂度，并且由定理3和4可知，当 $r=(2^{n/2-1}-1)/7$ 时，其最大线性复杂度约为理论上界的 1/2。

Table 1 several binary sequences set of large collection volume
表 1. 几类具有大集合容量的二元序列集

序列集	n	周期	容量	最大边峰值	线性复杂度(最小,最大)
Kasami(Large Set)[5]	偶数	2^n-1	$2^{\frac{n}{2}}(2^n+1)$	$2^{\frac{n+1}{2}}+1$	$(n,\frac{5n}{2})$
Chang 等[1]	奇数	2^n-1	2^{2n}	$2^{\frac{n+3}{2}}-1$	$(n,3n)$
Rothaus[2]	奇数	2^n-1	$2^{2n}+2^n+1$	$2^{\frac{n+3}{2}}-1$	$(n,3n)$
Shanbhag 等[3]	奇数	$2(2^n-1)$	2^{2n-1}	$2^{\frac{n+3}{2}}+2$	文献中未给出
Yu 和 Gong $S_o(3)$ [4]	奇数	2^n-1	2^{3n}	$2^{\frac{n+5}{2}}-1$	$(\frac{n(n-5)}{2},\frac{n(n+1)}{2})$
Yu 和 Gong $S_e(3)$ [4]	偶数	2^n-1	2^{3n}	$2^{\frac{n}{2}+3}-1$	$(\frac{n(n-5)}{2},\frac{n(n+1)}{2})$
本文的序列集	模 6 余 2	2^n-1	2^{3n}	$6\cdot 2^{\frac{n}{2}}-1$	$(n2^{\frac{n-8}{6}},n2^{\frac{n}{2}-2})$

4 结论

该文构造了一类具有大线性复杂度和大集合容量的二元低相关序列集，是首类同时具有上述三个性质的序列集。其突出的优点是具有很大的线性复杂度和集合容量。若将这类序列用于码分多址通信系统，可以增大系统的容量并提高系统的安全性。

References (参考文献)

[1] A. Chang, et al. “On a conjectured ideal autocorrelation sequence and a related triple-error correcting cyclic code”, IEEE Trans. Inf. Theory, vol.46, no.2, pp.680-687, Mar. 2000.
[2] S. Rothaus. “Modified Gold codes”, IEEE Trans. Inf. Theory, vol.39, no.2, pp.654-656, Mar. 1993.
[3] A. G. Shanbhag, el al, “Improved binary codes and seque- nces families from Z4-linear codes”, IEEE Trans. Inf. The- ory, vol.42,no.5, pp1582-1587, Sep. 1996.
[4] N. Y. Yu and G. Gong, “A new binary sequence family with low correlation and large size”, IEEE Trans. Inf. Theory, vol.52, no.4, pp.1624-1636, April. 2006.
[5] T. Kasmi, “Weight enumerators for several classes of sub- codes of the 2nd order Reed-Muller codes”, Information and control, vol. 18, pp.369-394, 1971.
[6] Fanxin Zeng. “Two classes of large families of sequences with low correlation”, IEEE Proceedings of IWSDA’97, pp.56-60, 2007.
[7] Niho Y; Multivalued Cross-Correlation Functions between Two Maximal Linear Recursive Sequences [D], Ph. D. dissertation, Univ. Southern Calif., Los Angeles, 1972.
[8] T. Helleseth. Some results about the cross-correlation function between two maximal linear sequences. Discrete Mathematics, 16(3):209-232, 1976.
[9] Lidl R, Niederreiter H. Introduction to Finite Fields and their Applications. Cambridge: Cambridge University Press, 1994.
[10] Key E L. An analysis of the structure and complexity of nonlin- ear binary sequence generators [J]. IEEE Trans. Inform. Theory, 1976, 22(6): 732-736.
[11] S. W. Golomb and G. Gong, Signal Designs With Good Correla- tion: For Wireless Communications, Cryptography and Radar Applications. Cambridge, U.K.: Cambridge University Press, 2005.