文章编号 1004-924X(2011)09-2222-06

基于匹配搜索的伪随机序列生成多项式估计

柴先明[1*],彭 耿[2],师栋锋[1],吕守业[1],詹 明[1]
(1.北京遥感信息研究所,北京 100192;
2.海军装备研究院 信息工程技术研究所,北京 102249)

摘要:对截短和误码条件下伪随机序列生成多项式估计算法的误码适应性进行了研究。根据伪随机序列的线性约束原理,通过构造关于生成多项式系数的校验方程组,提出了基于匹配搜索方法的生成多项式估计算法,并在搜索过程中利用生成多项式在二元域的基本特性缩小搜索范围来优化算法,减少计算量。最后,对算法在不同门限值和误码率时的估计性能进行了仿真,且以 m 序列为例,对本算法与典型算法在截短序列和误码适应性方面进行性能对比。仿真结果表明,该算法能较好完成对 15 阶 m 序列生成多项式的估计,可以适应 20%以上的误码率,与现有算法相比,本算法具有更好的误码和截短序列适应性,能较好地满足工程实践中扩频序列分析的需要。

关 键 词:信息处理;伪随机序列;生成多项式;优化估计;匹配搜索
中图分类号:TN911.7 **文献标识码**:A **doi**:10.3788/OPE.20111909.2222

Generator polynomial estimation of pseudo-random sequence based on match-searching

CHAI Xian-ming[1*], PENG Geng[2], SHI Dong-feng[1], LÜ Shou-ye[1], ZHAN Ming[1]

(1. *Beijing Institute of Remote Sensing Information*, *Beijing* 100192, *China*;
2. *Institute of Information Engineering and Technology*,
Naval Academy of Armament, *Beijing* 102249, *China*)
* *Corresponding author*, *E-mail*: *elevant*110@163.*com*

Abstract: The generator polynomial estimation of a pseudo-random sequence in truncated codes or error codes is studied in the paper. An optimum estimation algorithm for the generator polynomial via a match-searching is proposed by constructing verification equations based on the linear principle of the pseudo-random sequence. The important characteristic of reducible polynomial in $GF(2)$ is used by reducing the computational amount in algorithm optimization. Finally, the estimation performance of the algorithm at different thresholds and bit errors is simulated, and the performance of the algorithm for the truncated codes or error codes is compared with that of the typical algorithm by taking m-sequences as an example. Simulation results indicate that the algorithm has good adaptability for the error codes and truncated codes and it completes a estimation for the m-sequences with rank of 15 and

收稿日期:2010-11-29;**修订日期**:2011-03-15.
基金项目:湖南省优秀博士论文基金资助项目

can adapte to an error rate more than 20%. The algorithm can preferably meets engineering application requirements.

Key words: information processing; pseudo-random sequence; generator polynomial; optimum estimation; match-searching

1 引　言

伪随机序列在扩频通信、伪码测距以及密码学等领域有着广泛的应用[1-2]。生成多项式是伪随机序列的重要参数，是完成扩频通信侦察、信息解密等信号处理后续工作的基础。伪随机序列包括线性和非线性移位寄存器序列。扩频通信中多采用 m 序列、Gold 序列等线性反馈移位寄存器(LFSR)序列。LFSR 序列的生成多项式估计是序列分析的重要内容，特别是在仅知道部分码序列和存在误码的情况下，如何估计其生成多项式是该领域的研究热点[3]。

针对 LFSR 序列的生成多项式估计，目前的方法主要有 BM 算法[4]、欧几里德算法[5]、格基约化算法[6]、基于高阶累积量的方法[7]和有限域傅里叶变换方法[8]等。BM 算法、欧几里德算法和格基约化算法效率很高，计算复杂度仅为 $O(N^2)$，但是这些算法不能适应存在误码的情况；当 PN 序列周期为 2^l-1 时，与上述方法相比，基于有限域傅里叶变换的方法能进一步降低计算量，但不足之处是仍缺乏误码适应能力[8]；基于高阶累积量的方法尽管能适应一定的误码，但当误码率超过 10%时不再适用，并且算法在序列截短时性能急剧下降，同时由于该算法利用的是 m 序列尖锐的自相关特性，所以它只能适用于 m 序列生成多项式的估计，在估计 Gold 序列等其它伪随机序列生成多项式时存在很大困难。总而言之，现有的 LFSR 序列生成多项式估计方法在截短和误码条件下还存在使用范围受限、效率不高、精度不够等问题，值得进一步深入研究。

本文针对截短和误码条件下 LFSR 序列的生成多项式估计问题进行研究，利用 LFSR 序列的线性约束关系，通过构造关于生成多项式系数的校验方程组，搜索最优解来完成对生成多项式的估计，以适应较高的误码率，并在搜索过程中利用生成多项式在二元域的特性缩小搜索范围来减少计算量，结果满足工程实践需要。

2 算法基本原理推导

LFSR 序列由移位寄存器及初始状态唯一确定。由于寄存器的阶数和系数构成了序列的生成多项式，因此可先利用信号处理得到的 LFSR 序列估计其生成多项式，进而得到相应的参数。本文以扩频通信[9]中常用的 m 序列、Gold 序列等 LFSR 序列为例展开研究，首先分析 LFSR 序列的线性递归关系，然后推导基于齐次方程检验的匹配搜索算法，最后对算法进行优化设计。

2.1 LFSR 序列的线性递归关系

m 序列的生成多项式是二元域 $GF(2)$ 上的本原多项式，而 Gold 序列是由两个同阶的 m 序列优选对组合产生的，其生成多项式可以表示成 $GF(2)$ 上两个相同阶次的本原多项式的乘积[8]（文中的运算不作特殊说明，均在 $GF(2)$ 上进行）。

已知 m 序列片段为 $a_i, i=1,\cdots,n$，其生成多项式为：

$$g(x)=c_Lx^L+c_{L-1}x^{L-1}\cdots+c_1x+1\ . \tag{1}$$

根据 m 序列的产生过程有[8]：

$$a_n=\sum_{i=1}^{l}{}_ia_{n-i}\ , \tag{2}$$

式中 L 为移位寄存器阶数，$c_i, i=1,\cdots,L$ 为移位寄存器系数。

根据(2)式，m 序列的线性递归关系可以表示为齐次线性方程组：

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{L-1} & a_L \\ a_1 & a_2 & \cdots & a_L & a_{L+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_L & a_{L+1} & \cdots & a_{L+2} & a_{2L} \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} c_L \\ c_{L-1} \\ \vdots \\ c_1 \\ 1 \end{pmatrix} = 0\ , \tag{3}$$

同理，L 阶的 Gold 序列的线性递归关系可表示为：

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{2L-1} & a_{2L} \\ a_2 & a_3 & \cdots & a_{2L+1} & a_{2L+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{2L} & a_{2L+1} & \cdots & a_{4L-1} & a_{4L} \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} c_{2L} \\ c_{2L-1} \\ \vdots \\ c_1 \\ 1 \end{pmatrix} = 0 \ , \tag{4}$$

此时 Gold 序列的生成多项式为：

$$g(x)=c_{2L}x^{2L}+c_{2L-1}x^{2L-1}+\cdots+c_1x+1=g_1(x)\cdot g_2(x) \ , \tag{5}$$

式中 $g_1(x)$，$g_2(x)$ 分别为对应的两个 L 阶 m 序列的生成多项式。

2.2 基于齐次方程检验的匹配搜索算法

信号处理可得到序列的估计值 $\hat{a}_i$，$i=1,\cdots,n$，可能存在误码，此时的线性递归关系可表示为：

$$\begin{pmatrix} \hat{a}_0 & \hat{a}_1 & \cdots & \hat{a}_{L-1} & \hat{a}_L \\ \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_L & \hat{a}_{L+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{a}_L & \hat{a}_{L+1} & \cdots & \hat{a}_{L+1} & \hat{a}_{2L} \\ \vdots & \vdots & \ddots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} c_L \\ c_{L-1} \\ \vdots \\ c_1 \\ 1 \end{pmatrix} = \boldsymbol{R} = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_L \\ \vdots \end{pmatrix} , \tag{6}$$

显然，当估计序列 $\hat{a}_i$，$i=1,\cdots,n$ 中没有误码时，$\boldsymbol{R}=0$。假设待估计的 LFSR 序列生成多项式为：

$$g(x)=c_px^p+c_{p-1}x^{p-1}\cdots+c_1x+1, p\geqslant 2 \ , \tag{7}$$

则可设计待定生成多项式阶数 p 和系数 c_i，$i=1,\cdots,p$，逐步提高阶数 p 并遍历所有可能系数取值的方法，构造较大数量（通常远大于 p）的齐次方程代入(6)式进行验证，并根据信号环境和粗估误码率设置门限值，当待定多项式使得齐次方程成立的比例达到门限值时，确定匹配解，进而选取更多的齐次方程来检验匹配解的正确性，最终确定生成多项式。

根据接收序列 $\hat{a}_i$，$i=1,2,\cdots,n$，按照式(6)构造包含 N（N 可根据精度要求进行设置）个方程的方程组，将 $g(x)$ 代入方程组后，计算 N 个方程的值，可得到：

$$\begin{pmatrix} \hat{a}_0 & \hat{a}_1 & \cdots & \hat{a}_{p-1} & \hat{a}_p \\ \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_p & \hat{a}_{p+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{a}_{N-2} & \hat{a}_{N-1} & \cdots & \hat{a}_{N+p-3} & \hat{a}_{N+p-2} \\ \hat{a}_{N-1} & \hat{a}_N & \cdots & \hat{a}_{N+p-2} & \hat{a}_{N+p-1} \end{pmatrix} \cdot \begin{pmatrix} c_p \\ c_{p-1} \\ \vdots \\ c_1 \\ 1 \end{pmatrix} = \boldsymbol{R} = \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_{N-2} \\ r_{N-1} \end{pmatrix} , \tag{8}$$

然后统计结果向量 $\boldsymbol{R}$ 中为 0 和 1 的个数，分别记为 N_0 和 N_1。

显然有 $N_1=\sum_{i=0}^{N-1} r_i$，$N_0=N-N_1$。在无误码的情况下，当遍历的 $g(x)$ 恰为生成多项式时，满足 $N_0=N$。根据信号环境设置门限值 T_0，并计算优势值 $T=N_0/N$。T 值越大，即满足齐次方程组的比例越大，此时 $g(x)$ 越有可能是正确解。当 $T\geqslant T_0$ 时，则认为搜索到齐次线性方程组的匹配解，即生成多项式。

由于算法采用了正向匹配验证的思路，所以通过设置不同的门限值 T_0，可以适应高误码率的情况；同时，因为算法是基于 LFSR 序列生成原理的，所以识别范围广，包括常用的 m 序列、Gold 序列等。

2.3 算法的优化设计

由于 m 序列的生成多项式都是本原多项式（特殊的不可约多项式），因此可利用二元域上不可约多项式以及本原多项式的相关特性[10]缩小遍历范围、提高匹配搜索效率，实现算法的优化设计。

目前，在判断不可约多项式和本原多项式方面有许多较为成熟的方法：基于不可约充分条件的判断算法[11]和基于素性检验思想的判断算法[12]等，但这些方法都是用来寻找、验证不可约或本原多项式的，在实现上计算量较大。本文算法在搜索 LFSR 序列生成多项式的过程中，需要对 2 的几十阶个可能多项式进行匹配，若对每个多项式按上述判断算法进行检测，则计算量相当大、耗时太久，不能满足快速去除可约多项式的工程实践需求。

为此，本文利用二元域中可约多项式的几种简单特性，快速去除不满足 m 序列或 Gold 序列

生成多项式条件的可约多项式，缩小搜索匹配范围，使 2.2 节中匹配检验过程的计算耗时大幅度减小。

引理[13]：设 $g(x)=\sum_{i=0}^{n}c_i x^i$ 是 $GF(2)$ 上的次多项式，其中 $c_i\in\{0,1\}$，$c_n=c_0=1$，则有：

(1) $\sum_{i=0}^{n}c_i \bmod 2=0$；

(2) $\forall c_i, i \bmod 2=1$ 时 $c_i=0$；

(3) 令 $I_0=\{i;c_i=1,i\bmod 3=0\}$，$I_1=\{i;c_i=1,i\bmod 3=1\}$，$I_2=\{i,c_i=1,i\bmod 3=2\}$，$S_0=\sum_{i\in I_0}c_i$，$S_1=\sum_{i\in I_1}c_i$，$S_2=\sum_{i\in I_2}c_i$，则 $S_0 \bmod 2=0$，$S_1=S_2$，$S_1 \bmod 2=1$。

若 $g(x)$ 满足(1)、(2)或(3)中的任一条件，则称 $g(x)$ 为可约多项式。

利用引理中的性质，在匹配搜索过程中可以快速去除可约多项式，跳过齐次方程组的检验过程。通过分析可知：对于 L 阶的序列进行估计时，2.2 节提出的原始算法需要 $2^{L+2}\cdot L$ 次乘法运算和 $2^{L+2}\cdot(L-1)$次加法运算；利用性质(1)，计算量减少一半，需要 $2^{L+1}\cdot L$ 次乘法运算和 $2^{L+1}\cdot(L-1)$次加法运算；利用性质(2)，计算量减少到只需 $2^{L+1}\cdot L-2^{\frac{L}{2}}\cdot\frac{L}{2}$次乘法和 $2^{L+1}\cdot(L-1)-2^{\frac{L}{2}}\cdot(\frac{L}{2}-1)$次加法；利用性质(3)，仍能减少一定的计算量。由此可见，设计的优化算法可极大地缩小搜索范围，大大减少计算量。

在利用上述特性快速去除可约多项式的同时，还可利用性质“本原多项式的互反多项式也必定是本原多项式”[14]，得出“非本原多项式的互反多项式也必定是非本原多项式”的结论，从而每次检验均可剔除 2 个非本原多项式，减小计算量，提高算法效率。

3　算法仿真分析

本节结合仿真实验，分析算法在不同门限值和误码率时的估计性能，且以 m 序列为例，将本算法与典型算法在截短序列和误码适应性方面进行性能对比。

3.1　仿真实验 1

以阶数为 15、生成多项式 $g(x)=x^{15}+x^{14}+1$ 的 m 序列为估计对象，本算法在验证方程个数 N 为 500、所需序列长度为 515、门限值分别为 0.8、0.7 和 0.6 时，误码率与识别正确率的关系曲线如图 1 所示。

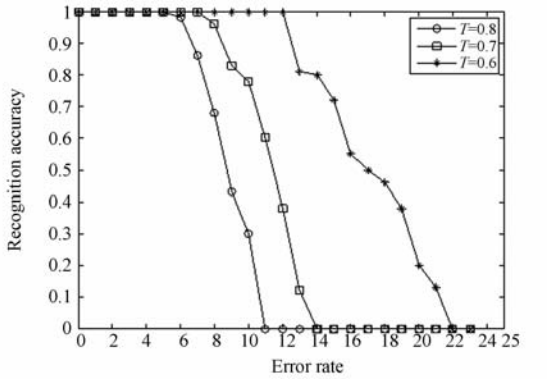

图 1　不同门限值时误码率与识别正确率的关系

Fig. 1　Relationship between bit error rate and recognition accuracy in different thresholds

3.2　仿真实验 2

由于现有的 BM 算法、欧几里得算法和格基约化算法不能适应存在误码的情况，而基于高阶累积量的方法仅适用于 m 序列，所以将本文算法与基于高阶累积量的方法在误码适应性和截短序列适应性等方面进行性能对比。以阶数为 10、生成多项式为 $g(x)=x^{10}+x^7+1$ 的 m 序列为识别对象，对序列分别为完整周期、800 位截短和 512 位截短时的算法性能进行仿真实验。本算法的验证方程个数取该序列长度下的最大值，门限值取 0.7，得仿真结果如图 2 所示。

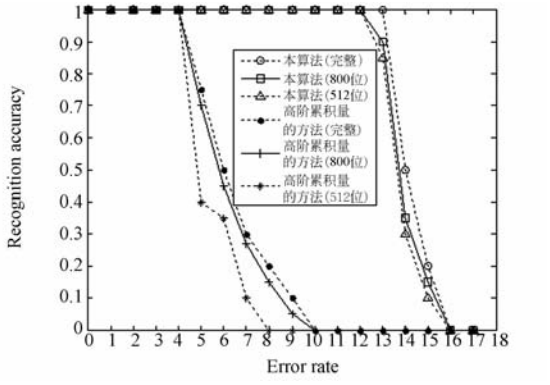

图 2　本文算法与基于高阶累积量的方法的性能比较

Fig. 2　Performance comparison of algortithm in this paper and using higher order statistics

3.3　仿真结果分析

从图 1 可以看出：本算法能完成对该 15 阶 m

序列生成多项式的估计，并且具有较好的误码适应性。在门限值为0.6时，可以适应20%以上的误码率。由于误码使得方程正确率下降，适当降低门限值可以适应一定的误码，从而提高对生成多项式的正确识别率。

从图2可以看出：本文算法在误码适应性能上要优于基于高阶累积量的方法，并且在对于截短序列进行估计时，基于高阶累积量的方法随着截短序列长度的减小，性能下降较大，而本文算法不需要完整的周期序列，只需序列长度能够取足一定数量的齐次方程，估计性能受到的影响很小。

对比图1、2可知：对低阶生成多项式的估计性能优于高阶的情况，这是因为随着阶数的增加，齐次校验方程中所用到的序列增长，所以要使得方程成立，要求越长的一段序列中没有误码，从而降低了对误码的适应性。

综上所述，在特殊的信号环境和工程应用背景中，通过选取足够的验证方程个数和设置合适的门限值，利用本文算法能快速有效地识别出LFSR序列，并且具有良好的误码适应性。

4 结　论

本文通过对LFSR序列的线性关系进行分析，构建关于生成多项式系数的齐次线性方程组，提出了基于校验思想的生成多项式匹配搜索算法，并利用GF(2)上可约多项式的特性对算法进行优化设计。仿真结果表明，算法有较强的误码适应能力和良好的识别性能，同时运算速度能较好地满足通信信号处理中常用20阶以内的扩频序列分析的需求，具有一定的军事意义和经济价值。

参考文献：

[1] WADE TRAPPE. *Introduction to Cryptography with Coding Theory*[M]. Beijing: Posts & Telecom Press, 2008.

[2] 彭耿，黄知涛，陆凤波，等. 双通道卫星通信信号快速盲检测[J]. 光学 精密工程，2009，17(10)：2535-2541.
PENG G, HUANG ZH T, LU F B, *et al*.. Double-channel fast blind detection of satellite communication signals[J]. *Opt. Precision Eng.*, 2009, 17(10): 2535-2541. (in Chinese)

[3] 吴迪. 直扩信号的快速同步技术研究[D]. 南京：南京理工大学，2009.
WU D. *Fast Synchronization technology of DS signal*[D]. NanJing: Nanjing University of Science and Technology, 2009. (in Chinese)

[4] BERLEKAMP E R. *Algebraic Coding Theory*[M]. NY USA: McGraw-Hill Book Company, 1968.

[5] HEYDTMANN A E, JENSEN J M. On the equivalence of the Berlekamp Massey and the Euclidean algorithms for decoding[J]. *IEEE Transactions on Information Theory*, 2000, 46(7): 2614-2624.

[6] 王丽萍，祝跃飞. F[x]-格基约化算法和多条序列综合[J]. 中国科学 E辑，2003，33(2)：8-12.
WANG L P, ZHU Y F. F[x]- Geikie reduction algorithm and multiple sequence synthesis[J]. *Science in China*, *Ser. E*, 2003, 33(2): 8-12. (in Chinese)

[7] SAID E E. Efficient detection of truncated m-sequence using higher order statistics[C]. *Proceeding of the 20th National Radio Science Conference*, *Cairo*, *Egypt*, 2003, 8: 1-9.

[8] WANG F H, HUANG ZH T, ZHOU Y Y. A new method for m-sequence and Gold-sequence generator polynomial estimation[C]. *Proceeding of IEEE International Symposium on Microwave Antenna*, *Propagation and EMC Technologies for Wireless Communications*, 2007.

[9] 史进. 自编码扩频技术的研究[D]. 西安：西安电子科技大学，2010.
SHI J. *Research on Self-encoded Spread Spectrum Technology*[D]. XiAn: XiDian University, 2010. (in Chinese)

[10] NINA D, LI C. LDPC Encoding based on the primitive polynomial[C]. *Wireless Communications Networking and Mobile Computing* (*WiCOM*), 2010: 1-2. (in Chinese)

[11] 王鑫，王新梅，韦宝典. 判定有限域上不可约多项式及本原多项式的一种高效算法[J]. 中山大学学报，2009，48(1)：6-9.
WANG X, WANG X M, WEI B D. An efficient and deterministic algorithm to determine irreducible and primitive polynomials over finite fields[J]. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2009, 48(1): 6-9. (in Chinese)

[12] 曹涵，陈恭亮. 基于素性检验思想的不可约多项

式判断[J]. 信息安全与通信保密，2006，3：73-74.
CAO H，CHEN G L. Test of irreducible polynomials based on primality-test[J]. *Information Security and Communications Privacy*，2006，3：73-74.（in Chinese）

[13] SHU L，DANIEL J C. *Error Control Coding*[M]. BeiJing：China Machine Press，2007.

[14] 郭鑫. 伪随机序列构造及其随机性分析研究[D]. 上海：上海交通大学，2008.
GUO X. *Study on construction and randomness analysis of pseudorandom sequences*[D]. ShangHai：Shanghai Jiao Tong University，2008.（in Chinese）

作者简介：

柴先明（1984－），男，安徽宿松人，助理工程师，2006 年于浙江大学获得学士学位，2008 于国防科技大学获得硕士学位，研究方向为信道编码盲识别、伪随机序列盲估计、电子侦察信号处理等。E-mail：elevant110@163.com

彭　耿（1980－），男，湖南平江人，博士，工程师，2010 年于国防科技大学获得博士学位，研究方向为通信侦察、空间信息对抗和系统仿真等。E-mail：hjhy-penggeng@163.com

师栋锋（1981－），男，山西永济人，硕士，工程师，2005 年于装备指挥技术学院获得硕士学位，研究方向为电子侦察信号处理、空间信息对抗等。Email：xinghelion@163.com

吕守业（1979－），男，山东寿光人，博士，副研究员，2005 年于北京理工大学获得博士学位，研究方向为电子侦察信号处理、空间信息对抗等。E-mail：Lvsy200310@163.com

詹　明（1965－），男，北京人，副研究员，1987 年于航空工程大学获得学士学位，研究方向为图像判读、红外探测等。E-mail：Zhm_2000@163.com