第 26 卷第 11 期 电 子 与 信 息 学 报 Vol.26No.11
2004 年 11 月 Journal of Electronics & Information Technology Nov. 2004

周期序列线性复杂度与 k-错复杂度的数学期望 [1]

牛志华 白恩健 肖国镇

(西安电子科技大学综合业务网国家重点实验室 西安 710071)

摘 要: 密码学意义上强的序列不仅应该具有足够高的线性复杂度，而且当少量比特发生改变时不会引起线性复杂度的急剧下降，即具有高的 k-错复杂度. 该文以多项式的因式分解为主要工具研究了任意有限域 GF(q) 上，周期 N 与 p 互素以及 $N=p^v$ 这两种情况下，计数函数 $\mathcal{N}_{N,0}(c)$ 的值，并给出了线性复杂度的数学期望 $E_{N,0}$ 的值以及 k-错复杂度的数学期望 $E_{N,k}$ 的一个有用的下界，这里 p 是有限域 GF(q) 的特征.

关键词: 流密码，周期序列，线性复杂度， k-错复杂度

中图分类号: TN918.4 **文献标识码**: A **文章编号**: 1009-5896(2004)11-1787-05

On the Expected Value of the Linear Complexity and the k-Error Linear Complexity of Periodic Sequences

Niu Zhi-hua Bai En-jian Xiao Guo-zhen

(*National Key Lab of Integrated Service Networks, Xidian Univ., Xi'an* 710071, *China*)

Abstract Cryptographically strong sequences not only should have a large linear complexity, but also no a significant decrease of the linear complexity when a few terms are changed. This requirement leads to the concept of the k-error linear complexity of periodic sequences. In the following two cases: (1) $\gcd(N,p)=1$; (2) $N=p^v$, where p denotes the characteristic of the finite field GF(q), the counting function $\mathcal{N}_{N,0}(c)$, i.e., the number of N-periodic sequences with given linear complexity c, is showed, the expected value of the linear complexity $E_{N,0}$ is determined, and a useful lower bound on the expected value of the k-error linear complexity $E_{N,k}$ is established.

Key words Stream cipher, Periodic sequences, Linear complexity, k-Error linear complexity

1 引言

周期序列的线性复杂度是流密码强度的一个重要度量指标，序列 (S) 的线性复杂度定义为生成它的最小线性反馈移位寄存器的长度，记为 LC(S)，密钥序列的线性复杂度必须足够大，因为只要知道连续 2LC(S) 个比特，就可以通过解线性方程组或借助 BM 算法将整个序列完全确定.

但是，有的序列虽然具有很高的线性复杂度但其线性复杂度却极不稳定，即当改变这些序列周期的一位或几位时，其线性复杂度发生很大的变化，这样的序列用来作为密钥序列是不安全的. 因此，序列的线性复杂度的稳定性与序列的不可测性是密切相关的. 我国学者早就注意到这个问题并率先创立了流密码的稳定性理论，并引入球体复杂度，重量复杂度等流密码稳定

[1] 2003-05-15 收到， 2003-12-02 改回
国家自然科学基金项目 (60073051) 和国家自然科学基金重大项目 (90104005) 资助课题

性度量指标 [1]，国外学者 Stamp 和 Martin 也注意到这个问题，在文献 [2] 中提出了类似球体复杂度的 k-错复杂度的概念。

设 (S) 是有限域 $\mathrm{GF}(q)$ 上周期为 N 的序列，当改变 (S) 的周期中至多 $k(0 \le k \le N)$ 位后，得到的所有序列的线性复杂度中最小的线性复杂度，称为 (S) 的 k-错复杂度，记为 $\mathrm{LC}_k(S)$，即 $\mathrm{LC}_k(S) = \min_{0 \le W_H(E^N) \le k} \mathrm{LC}(S+E)$，这里 (E) 是周期为 N 的序列，$W_H(E^N)$ 表示 (E) 的周期中不为零的元素个数，$\mathrm{LC}(S)$ 表示 (S) 的线性复杂度。

在文献 [2] 中 Stamp 和 Martin 还设计了一个计算周期为 2^n 的二元序列 k-错复杂度的有效算法，之后，k-错复杂度及其算法的研究引起了人们的广泛兴趣，文献 [3] 将 Stamp-Martin 算法做了推广，给出了任意有限域 $\mathrm{GF}(p^m)$ 上周期为 p^n 的序列的 k-错复杂度的有效算法，但目前还没有一个有效的计算任意周期序列 k-错复杂度的通用算法。

研究周期序列线性复杂度和 k-错复杂度的随机性，如分布、数学期望、方差等，有助于我们进一步研究和设计流密码的生成器以及计算周期序列线性复杂度和 k-错复杂度的算法。Rueppel 对二元序列线性复杂度的随机性进行了研究 [4]，文献 [5] 以广义离散傅里叶变换作为主要的研究工具，给出了任意有限域 $\mathrm{GF}(q)$ 上周期序列线性复杂度和 k-错复杂度的数学期望的一些有用的下界。本文仅利用多项式的因式分解等简单工具，研究了任意有限域 $\mathrm{GF}(q)$ 上，周期 N 与 p 互素以及 $N = p^v$ 这两种情况下，计数函数 $\mathcal{N}_{N,0}(c)$ 的值，并给出了线性复杂度的数学期望 $E_{n,0}$ 的值以及 k-错复杂度的数学期望 $E_{N,k}$ 的一个有用的下界。其中 $N = p^v$ 的情况下的结果与文献 [5] 中完全相同，但我们使用的方法更加简单。本文中 p 均表示有限域 $\mathrm{GF}(q)$ 的特征。

本文内容安排如下：第 2 节研究周期 N 与 p 互素时线性复杂度和 k-错复杂度的数学期望；第 3 节研究 $N = p^v$ 时的情况；第 4 节是结束语。

2 周期 N 与 p 互素时线性复杂度和 k-错复杂度的数学期望

我们先介绍本文将要用到的一些记号。$\mathcal{N}_{N,0}(c)$ 表示 $\mathrm{GF}(q)$ 上周期为 N，线性复杂度为 c 的序列的个数；$\mathcal{N}_{N,k}(c)$ 表示 $\mathrm{GF}(q)$ 上周期为 N，k-错复杂度为 c 的序列的个数；$\mathcal{M}_{N,0}(c)$ 表示 $\mathrm{GF}(q)$ 上周期为 N，线性复杂度不超过 c 的序列的个数；$\mathcal{M}_{N,k}(c)$ 表示 $\mathrm{GF}(q)$ 上周期为 N，k-错复杂度不超过 c 的序列的个数；$E_{N,0}$ 表示 $\mathrm{GF}(q)$ 上周期为 N 的序列的线性复杂度的数学期望；$E_{N,k}$ 表示 $\mathrm{GF}(q)$ 上周期为 N 的序列的 k-错复杂度的数学期望。

引理 1[5] 对于任意的整数 $N \ge 1$ 和 $0 \le k \le N$，$\mathrm{GF}(q)$ 上周期为 N 的序列的 k-错复杂度的数学期望为

$$E_{N,k} = N - \frac{1}{q^N} \sum_{c=0}^{N-1} \mathcal{M}_{(N,k)}(c)$$

当周期 N 与有限域的特征 p 互素时，$1 - x^N$ 在 $\mathrm{GF}(q)$ 上无重根。

引理 2[6] 设 p 是有限域 $\mathrm{GF}(q)$ 的特征，N 是与 p 互素的正整数，$1 - x^N$ 在 $\mathrm{GF}(q)$ 上的不可约分解为 $1 - x^N = f_1(x) f_2(x) \cdots f_l(x)$，则任意周期为 N 的序列 (S) 的生成函数可表示为

$$S(x) = \frac{S^N(x)}{1 - x^N} = \frac{g_1(x)}{f_1(x)} + \frac{g_2(x)}{f_2(x)} + \cdots + \frac{g_l(x)}{f_l(x)} \tag{1}$$

这里 $\gcd(g_i(x), f_i(x)) = 1$，且 $\deg(g_i(x)) < \deg(f_i(x))$ 或 $g_i(x) = 0$, $i = 1, 2, \cdots, l$。

引理 3 设 p 是有限域 $\mathrm{GF}(q)$ 的特征，N 是与 p 互素的正整数，$1 - x^N$ 在 $\mathrm{GF}(q)$ 上可分解为一次因式的乘积，则

$$\mathcal{N}_{N,0}(c) = C_N^c (q-1)^c$$

C_N^i 表示 N 中取 i 的组合数.

证明　设 $1-x^N$ 在 GF(q) 上的不可约分解为 $1-x^N=f_1(x)f_2(x)\cdots f_N(x)$, $f_i(x)$ $(i=1,2,\cdots,N)$ 均为一次因式，则由引理 2，

$$S(x)=\frac{S^N(x)}{1-x^N}=\frac{g_1(x)}{f_1(x)}+\frac{g_2(x)}{f_2(x)}+\cdots+\frac{g_N(x)}{f_N(x)} \tag{2}$$

其中 $g_i(x)=0$，或 $g_i(x)$ 为常数，$i=1,2,\cdots,N$，则 $\mathrm{LC}(S)=c$ 当且仅当式 (2) 中恰有 c 个 $g_i(x)$ 不为 0，这样的序列 (S) 有 $C_N^c(q-1)^c$ 个，即 $\mathcal{N}_{N,0}(c)=C_N^c(q-1)^c$.

证毕

由引理 3 立即可得：

定理 1　设 p 是有限域 GF(q) 的特征，N 是与 p 互素的正整数，$1-x^N$ 在 GF(q) 上可分解为一次因式的乘积，则 GF(q) 上周期为 N 的序列的线性复杂度的数学期望为

$$E_{N,0}=\frac{1}{q^N}\sum_{c=1}^{N}cC_N^c(q-1)^c$$

设序列 (S) 是一个 GF(q) 上的周期为 N，线性复杂度不超过 c 的序列，则 $\mathcal{M}_{N,k}(c)$ 表示的就是与某个这样的 (S) 的 Hamming 距离不超过 k 的序列的个数. 所以 $\mathcal{M}_{N,k}(c)\le\mathcal{M}_{N,0}(c)\sum_{t=0}^{k}C_N^t(q-1)^t$，从而

$$\mathcal{M}_{N,k}(c)\le\min\left\{q^N,\mathcal{M}_{N,0}(c)\sum_{t=0}^{k}C_N^t(q-1)^t\right\} \tag{3}$$

定理 2　设 p 是有限域 GF(q) 的特征，N 是与 p 互素的正整数，$1-x^N$ 在 GF(q) 上可分解为一次因式的乘积，则 GF(q) 上周期为 N 的序列的 k-错复杂度的数学期望满足：

$$E_{N,k}\ge\lambda+1-\frac{1}{q^N}\sum_{c=0}^{\lambda}\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i$$

其中 $\lambda=\max\left\{c\left|\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i\le q^N\right.\right\}$.

证明　由引理 3，$\mathcal{M}_{N,0}(c)=\sum_{r=0}^{c}\mathcal{N}_{N,0}(r)=\sum_{r=0}^{c}C_N^r(q-1)^r$，

$$\begin{aligned}\mathcal{M}_{N,0}(c)\sum_{i=0}^{k}C_N^t(q-1)^t&=\sum_{r=0}^{c}C_N^r(q-1)^r\sum_{t=0}^{k}C_N^t(q-1)^t\\&=\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i\end{aligned}$$

令 $\lambda=\max\left\{c\left|\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i\le q^N\right.\right\}$，则由引理 1 和式 (3)，

$$q^NE_{N,k}=q^NN-\sum_{c=0}^{N-1}\mathcal{M}_{N,k}(c)$$

$$\ge q^NN-q^NN(N-\lambda-1)-\sum_{c=0}^{\lambda}\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i$$

$$=q^N(\lambda+1)-\sum_{c=0}^{\lambda}\sum_{i=0}^{c+k}\sum_{\substack{0\le j\le c\\0\le i-j\le k}}C_N^jC_N^{i-j}(q-1)^i$$

两边除以 q^N 即可得到定理的表达式. 证毕

3 周期 $N=p^v$ 时线性复杂度和 k-错复杂度的数学期望

引理 4 设 p 是有限域 GF(q) 的特征，$N=p^v$，则 $\mathcal{N}_{N,0}(0)=1$，当 $c\ge 1$ 时，$\mathcal{N}_{N,0}(c)=q^{c-1}(q-1)$.

证明 周期为 $N=p^v$ 的 q 元序列 (S) 的生成函数可表示为

$$S(x)=\frac{S^N(x)}{1-x^N}=\frac{S^N(x)}{(1-x)^{p^v}}=\frac{g(x)}{(1-x)^i} \tag{4}$$

这里 $1-x$ 不整除 $g(x)$，$\deg(g(x))<i$ 或 $g(x)=0$. $c=0$ 当且仅当 (S) 为全 0 序列，即 $\mathcal{N}_{N,0}(0)=1$，当 $c\ge 1$ 时，(S) 的线性复杂度为 c 当且仅当式 (4) 中 $i=c$，满足上述条件的 $g(x)$ 有 $q^c-1-(q^{c-1}-1)=q^{c-1}(q-1)$ 个，从而这样的序列 (S) 有 $q^{c-1}(q-1)$ 个，即 $\mathcal{N}_{N,0}(c)=q^{c-1}(q-1)$. 证毕

定理 3 设 p 是有限域 GF(q) 的特征，$N=p^v$，则 GF(q) 上周期为 N 的序列的线性复杂度的数学期望为：

$$E_{N,0}=N-[(1-q^{-N})/(q-1)]$$

证明

$$E_{N,0}=\frac{1}{q^N}\sum_{c=0}^{N}c\mathcal{N}_{N,0}(c)=\frac{1}{q^N}\sum_{c=1}^{N}cq^{c-1}(q-1)=\frac{q-1}{q^N}\left(\frac{Nq^N}{q-1}-\frac{q^N-1}{(q-1)^2}\right)$$

$$=N-\frac{1-q^{-N}}{q-1}$$

证毕

定理 4 设 p 是有限域 GF(q) 的特征，$N=p^v$，则 GF(q) 上周期为 N 的序列的 k-错复杂度的数学期望满足：

$$E_{N,k}\ge N-\log_q\sum_{t=0}^{k}C_N^t(q-1)^t-\frac{q}{q-1}$$

证明 当 $c=0$ 时，$\mathcal{M}_{N,0}(0)=\mathcal{N}_{N,0}(0)=1=q^0$；当 $c\ge 1$ 时，$\mathcal{M}_{N,0}(c)=\sum_{r=0}^{c}\mathcal{N}_{N,0}(r)=1+\sum_{r=1}^{c}q^{r-1}(q-1)=q^c$. $\mathcal{M}_{N,0}(c)\sum_{t=0}^{k}C_N^t(q-1)^t=q^c\sum_{t=0}^{k}C_N^t(q-1)^t<q^N$ 当且仅当

$c \le \lambda := \left\lfloor N - \log_q \sum_{t=0}^{k} C_N^t (q-1)^t \right\rfloor$，则由引理 1 和式 (3)，

$$
\begin{aligned}
q^N E_{N,k} &= q^N N - \sum_{c=0}^{N-1} \mathcal{M}_{N,k}(c) \\
&\ge q^N N - q^N (N-\lambda-1) - \sum_{c=0}^{\lambda} q^c \sum_{t=0}^{k} C_N^t (q-1)^t \\
&= q^N (\lambda+1) - \frac{q^{\lambda+1}-1}{q-1} \sum_{t=0}^{k} C_N^t (q-1)^t \\
&\ge q^N \left(N - \log_q \sum_{t=0}^{k} C_N^t (q-1)^t \right) - \frac{1}{q-1} q^{N-\log_q \sum_{t=0}^{k} C_N^t (q-1)^t + 1} \sum_{t=0}^{k} C_N^t (q-1)^t \\
&= q^N \left(N - \log_q \sum_{t=0}^{k} C_N^t (q-1)^t \right) - \frac{1}{q-1} q^{N-\log_q \sum_{t=0}^{k} C_N^t (q-1)^t + 1} q^{\log_q \sum_{t=0}^{k} C_N^t (q-1)^t} \\
&= q^N \left(N - \log_q \sum_{t=0}^{k} C_N^t (q-1)^t \right) - \frac{q^{N+1}}{q-1}
\end{aligned}
$$

两边除以 q^N 即可得到定理的表达式。　　证毕

在这一节中我们仅以多项式的因式分解为主要工具，得到了与文献 [5] 完全相同的结果，而文献 [5] 得到这样的结果是以广义离散傅里叶变换为其主要工具的。

4 结束语

研究周期序列线性复杂度和 k-错复杂度的数学期望，有助于我们进一步研究和设计流密码的生成器和计算周期序列线性复杂度和 k-错复杂度的算法，尤其是 k-错复杂度随机性的研究才刚刚起步，本文给出了 k-错复杂度数学期望的一个下界，进一步搞清周期序列 k-错复杂度的分布对于了解序列的随机性及其稳定性具有重要的意义。

参　考　文　献

[1] Ding C, Xiao G, Shan W. The stability theory of stream ciphers. Lecture Notes in Computer Science. Vol.561, Berlin: Springer-Verlag, 1991.

[2] Stamp M, Martin C F. An algorithm for the k-error linear complexity of binary sequences of period 2^n. *IEEE Trans. on Information Theory*, 1993, IT-39(4): 1398–1401.

[3] Kaida T, Uehara S, Imamura K. A new algorithm for the k-error linear complexity of sequences over GF(p^m) with period p^n. In Sequences and Their Applications. Ding C, Helleseth T, Niederreiter H. Eds. London, U. K. : Springer, 1999: 284–296.

[4] Rueppel R A. Analysis and Design of Stream Ciphers. Berlin: Springer-Verlag, 1986.

[5] Meidl W, Niederreiter H. On the expected value of the linear complexity and the k-error linear complexity of periodic sequences. *IEEE Trans. on Information Theory*, 2002, IT-48(11): 2817–2825.

[6] Wei S, Zhang Y, Xiao G. Distribution of linear complexity for periodic sequences. CrypTEC'99, Hong Kong, July 1999: 250–253.

牛志华：　女，1976 年生，博士生，主要研究方向为信息安全与密码学.

白恩健：　男，1977 年生，博士生，主要研究方向为信息安全与密码学.

肖国镇：　男，1934 年生，教授，博士生导师，主要研究方向为信息论、编码学与密码学.