Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Theoretical Computer Science 352 (2006) 288–292

Theoretical Computer Science

www.elsevier.com/locate/tcs

On linear complexity of sequences over $GF(2^n)$

A.M. Youssef[a,*], G. Gong[b]

[a]Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Que., Canada H3G 1M8
[b]Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ont., Canada N2L 3G1

Received 29 June 2005; accepted 19 November 2005

Communicated by V. Pan

Abstract

In this paper, we consider some aspects related to determining the linear complexity of sequences over $GF(2^n)$. In particular, we study the effect of changing the finite field basis on the minimal polynomials, and thus on the linear complexity, of sequences defined over $GF(2^n)$ but given in their binary representation. Let $\mathbf{a} = \{a_i\}$ be a sequence over $GF(2^n)$. Then a_i can be represented by $a_i = \sum_{j=0}^{n-1} a_{ij}\alpha^j$, $a_{ij} \in GF(2)$, where α is the root of the irreducible polynomial defining the field. Consider the sequence $\mathbf{b} = \{b_i\}$ whose elements are obtained from the same binary representation of $\mathbf{a}$ but assuming a different set of basis (say $\{\gamma_0, \gamma_1, \ldots, \gamma_{n-1}\}$), i.e., $b_i = \sum_{j=0}^{r-1} a_{ij}\gamma_j$. We study the relation between the minimal polynomial of $\mathbf{a}$ and $\mathbf{b}$.

Keywords: Linear complexity; Berlekamp–Massey algorithm; Cryptanalysis; Finite fields

1. Introduction

The linear complexity of a given sequence is defined as the length of the shortest linear feedback shift register (LFSR) [2] which can produce this sequence [11]. The sequence linear complexity is one measure of its predictability. If the linear complexity of a sequence $\mathbf{s}$ is l, then the cryptanalyst can recover the entire sequence by observing $2l$ consecutive elements of $\mathbf{s}$ [7]. Thus, stream cipher designers should ensure that sequences produced by their ciphers have large linear complexity

While many classical stream ciphers were based on bit-wise operations, most of modern stream cipher [1] tend to operate on words (e.g. 8–32 bit) in order to provide efficient software implementations. Feedback shift registers over $GF(2^n)$ are becoming popular building blocks in many of these ciphers (e.g., the NESSIE submissions SOBER-t16 and SOBER-t32 [8]).

In this paper, we consider a practical issue that faces the cryptanalyst when trying to determine the linear complexity of a given $GF(2^n)$ sequence. Consider a cryptanalyst who observes a key stream, naturally represented as a binary stream of bits. Assuming that n as well as the boundary between the n-tuples are known, an interesting question is which basis should the cryptanalyst use to evaluate the linear complexity of the given sequence. The following example illustrate this point.

* Corresponding author. Tel.: +1 514 848 2424; fax: +1 514 848 3171.
E-mail addresses: youssef@ciise.concordia.ca (A.M. Youssef), G.Gong@ece.uwaterloo.ca (G. Gong).

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2005.11.044

Example 1. Consider the following binary representation of a periodic sequence over $GF(2^3)$:

{001, 011, 000, 100, 111, 000, 110,
001, 000, 101, 100, 000, 010, 110,
000, 011, 101, 000, 111, 010, 000}.

If we use the irreducible polynomial $x^3 + x + 1$ and assume that binary representation above is using the polynomial basis $(\alpha^2, \alpha, 1)$ where α is a root of the irreducible polynomial above, then the above sequence can be described as

$$\mathbf{a} = \{1, \alpha^3, 0, \alpha^2, \alpha^5, 0, \alpha^4, \alpha^0, 0, \alpha^6, \alpha^2, 0, \alpha, \alpha^4, 0, \alpha^3, \alpha^6, 0, \alpha^5, \alpha, 0\}.$$

The minimal polynomial, calculated using the Berlekamp–Massey algorithm [5], is given by $x^2 + \alpha^3 x + \alpha^6 = 0$, and hence the linear complexity is 2. Thus it suffices to know 4 consecutive tuples of this sequence to fully recover it.

On the other hand, if we use the basis $(\beta^2, \beta, 1)$ where β is a root of the irreducible polynomial $x^3 + x^2 + 1$, then the sequence above is given by

$$\mathbf{b} = \{1, \beta^5, 0, \beta^2, \beta^4, 0, \beta^6, 1, 0, \beta^3, \beta^2, 0, \beta, \beta^6, 0, \beta^5, \beta^3, 0, \beta^4, \beta, 0\}.$$

Again, the Berlekamp–Massey algorithm yields the following minimal polynomial $x^6 + x^5 + x^4 + x^2 + 1 = 0$, and hence the linear complexity is 6 which means that the sequence designer might think that at least 12 consecutive tuples of this sequence are needed to fully recover it.

One can think of a $GF(2^n)$ sequence given in its binary representation as a set of n sequences over $GF(2)$. For example, if $\mathbf{a}$ is an m-sequence sequences over $GF(2^n)$ with minimal polynomial m of degree d, then $\mathbf{a}$ can be decomposed into n shift equivalent m-sequences over $GF(2)$ with minimal polynomials of degree $d \times n$ [3,9]. However, as illustrated by the following example, this may not always be the preferred choice for the cryptanalyst.

Example 2. Consider the periodic sequence $\mathbf{a}$ as defined in Example 1. Then $\mathbf{a}$ can be decomposed into the following three shift equivalent binary sequences by setting $\mathbf{a}_i = (\mathbf{a}_{3i}, \mathbf{a}_{2i}, \mathbf{a}_{1i})$

$$\begin{aligned}
\mathbf{a}_1 &= \{1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0\},\\
\mathbf{a}_2 &= \{0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0\},\\
\mathbf{a}_3 &= \{0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0\}
\end{aligned}$$

with $m_{\mathbf{a}_1}(x) = m_{\mathbf{a}_2}(x) = m_{\mathbf{a}_3}(x) = x^6 + x^5 + x^4 + x^2 + 1$. Thus the cryptanalyst needs to observe 12 (required by Berlekamp–Massey algorithm to determine both the connection polynomial and initial state) + 6 (to determine the initial state) + 6 (to determine the initial state) = 24 bits in order to recover the entire sequence. Note that from Example 1, only 4 tuples = 12 bits were required to recover the sequence $\mathbf{a}$.

Moreover, for a general $GF(2^n)$ sequence, the corresponding binary sequences are not necessarily shift equivalent, which implies that the cryptanalyst needs to run the Berlekamp–Massey algorithm for every individual bit separately as illustrated by the following example.

Example 3. Let $\mathbf{a} = \{1, \alpha^3, \alpha, \alpha^5, 0, \alpha^4, \alpha^6, 1, \alpha^3, \alpha, \alpha^5, 0, \alpha^4, \alpha^6\}$ where α is a root of $x^3 + x + 1$ over $GF(2^3)$. Then we have

$$m_{\mathbf{a}}(x) = x^2 + \alpha^2 x + \alpha^6 = (x + 1)(x + \alpha^6).$$

If we assume that the binary representation of $\mathbf{a}$ is obtained using the polynomial basis, then we get

$$m_{\mathbf{a}_1}(x) = m_{\mathbf{a}_2}(x) = x^3 + x^2 + 1,$$

$$m_{\mathbf{a}_3}(x) = x^4 + x^2 + x + 1.$$

2. Mathematical background and definitions

For a background about the general theory of finite fields, the reader is referred to [4] and for a background about finite fields of characteristic 2, the reader is referred to [6]. Throughout the rest of this paper, we will only consider sequences with minimal polynomials that have no multiple roots.

Definition 1. A polynomial having the special form

$$L(x) = \sum_{i=0}^{t} \beta_i x^{2^i}$$

with coefficients β_i from $GF(2^n)$ is called a linearized polynomial over $GF(2^n)$.

Lemma 1. *Let A be a linear mapping over $GF(2)^n$, then this mapping can be expressed in terms of a linearized polynomial over $GF(2^n)$*

Lemma 2. *Let $\alpha_1, \alpha_2, \ldots, \alpha_t$ be elements in $GF(2^n)$. Then*

$$(\alpha_1 + \alpha_2 + \cdots + \alpha_t)^{2^k} = \alpha_1^{2^k} + \alpha_2^{2^k} + \cdots + \alpha_t^{2^k}.$$

Lemma 3. *For $i = 1, 2, \ldots, h$ let $\mathbf{s}_i$ be a homogeneous linear recurring sequence in $GF(q)$ with minimal polynomial $m_i(x)$. Then, the minimal polynomial of the sequence $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2 + \cdots + \mathbf{s}_h$, $\mathbf{s}_i \neq \mathbf{s}_j$ for $i \neq j$, is given by the least common multiple of m_i's, i.e., $m(x) = lcm(m_1, \ldots, m_h)$* [4].

3. Main results

Let $\mathbf{a} = \{a_i\}$ be a sequence over $GF(2^n)$. Then a_i can be represented by

$$a_i = \sum_{j=0}^{n-1} a_{ij}\alpha^j, \quad a_{ij} \in GF(2), \quad i = 0, 1, \ldots, \tag{1}$$

where α is the root of the irreducible polynomial defining the field. Consider the sequence $\mathbf{b} = \{b_i\}$ whose elements are obtained from the same binary representation of $\mathbf{a}$ but assuming a different set of basis (say $\{\gamma_0, \gamma_1, \ldots, \gamma_{n-1}\}$), i.e.,

$$b_i = \sum_{j=0}^{n-1} a_{ij}\gamma_j, \quad i = 0, 1, \ldots . \tag{2}$$

Let σ_j, $j = 0, \ldots, n-1$ denotes the transform that maps a polynomial $u(x) = \sum_i u_i x^i$, $u_i \in GF(2^n)$ to

$$\sigma_j(u(x)) = \sum_{i=0}^{r-1} u_i^{2^j} x^i. \tag{3}$$

In what follows, we study the relation between the minimal polynomial of $\mathbf{b}$ and $\mathbf{a}$.

Besides its cryptanalytic significance, the results presented in this paper can be used to efficiently generate of linear streams over $GF(2^n)$ without performing any multiplication operations. Moreover, most of the results in this paper can also be extended in a straightforward way from $GF(2^n)$ to $GF(p^n)$.

Lemma 4. *Let $\mathbf{s}^{(2)} = \{s_i^{(2)}\}$ be the sequence obtained from $\mathbf{s} = \{s_i\}$ by squaring every element of $\mathbf{s}$, i.e., $s_i^{(2)} = s_i^2$. Let $f_1(x) = \sum_{i=1}^{n} a_i x^i$ be the minimal polynomial of $\mathbf{s}$, then the minimal polynomial of $\mathbf{s}^{(2)}$ divides $\sigma_1(f_1(x))$.*

Proof. Since f_1 is the characteristic polynomial of $\mathbf{s}$, then the sequence $\mathbf{s}$ satisfies the recurrent relation

$$a_n s(i) + a_{n-1} s(i-1) + \cdots + a_0 s(i-n) = 0.$$

By squaring the above equation, and using Lemma 2, we have

$$a_n{}^2(s(i))^2 + a_{n-1}^2(s(i-1))^2 + \cdots + a_0^2(s(i-n))^2 = 0.$$

Thus we get

$$a_n^2 s^{(2)}(i) + a_{n-1}^2 s^{(2)}(i-1) + \cdots + a_0^2 s^{(2)}(i-n) = 0,$$

which proves the lemma.

Example 4. Consider the same periodic sequence given in Example 1 above

$$\mathbf{s} = \left\{1, \alpha^3, 0, \alpha^2, \alpha^5, 0, \alpha^4, \alpha^0, 0, \alpha^6, \alpha^2, 0, \alpha, \alpha^4, 0, \alpha^3, \alpha^6, 0, \alpha^5, \alpha, 0\right\}$$

defined over $GF(2^3)$ where α is a root of the irreducible polynomial $x^3 + x + 1$. The minimal polynomial of $\mathbf{s}$ is given by $f_1(x) = x^2 + \alpha^3 x + \alpha^6$.

The minimal polynomial of the sequence

$$\mathbf{s}^{(2)} = \left\{1, \alpha^6, 0, \alpha^4, \alpha^3, 0, \alpha, 1, 0, \alpha^5, \alpha^4, 0, \alpha^2, \alpha, 0, \alpha^6, \alpha^5, 0, \alpha^3, \alpha^2, 0\right\}$$

is given by $x^2 + (\alpha^3)^2 x + (\alpha^6)^2 = x^2 + \alpha^6 x + \alpha^5 = \sigma_1(f(x))$. Consequently, the minimal polynomial of the sequence

$$\mathbf{s}^{(4)} = \left\{1, \alpha^5, 0, \alpha, \alpha^6, 0, \alpha^2, 1, 0, \alpha^3, \alpha, 0, \alpha^4, \alpha^2, 0, \alpha^5, \alpha^3, 0, \alpha^6, \alpha^4, 0\right\}$$

is given by $x^2 + (\alpha^3)^4 x + (\alpha^6)^4 = x^2 + \alpha^5 x + \alpha^3 = \sigma_2(f(x))$.

Lemma 5. *Let* **a** *and* **b** *be the sequences described by* (1), (2) *respectively. Then* **b** *can be expressed as*

$$b_i = \sum_{j=0}^{n-1} c_j a_i^{2^j}, \quad c_j \in GF(2^n). \tag{4}$$

Proof. By noting that changing the basis is equivalent to applying a linear transformation to the co-ordinates of **a**, the proof follows by applying Lemma 1.

Theorem 1. *Let $f(x)$ and $g(x)$ denote the minimal polynomials corresponding to* **a** *and* **b** *described by* (1), (2) (*and consequently* (4)), *respectively. Then $g(x)$ is given by*

$$lcm(\sigma_{i_0}(f(x)), \sigma_{i_1}(f(x)) \cdots \sigma_{i_r}(f(x))), \tag{5}$$

where σ_{i_j} is given by (3), *and $i_j \in \{j | c_j \neq 0\}$.*

Proof. The result follows by applying Lemma 3.

Example 5. Consider the sequence **a** defined over $GF(2^4)$ with $m_a(x) = x^5 + \alpha x^4 + \alpha^9$ where α is a root of the irreducible polynomial $x^4 + x^3 + 1$. Let **b** denote the sequence obtained from **a** assuming the basis $\{\alpha^6, \alpha^{14}, \alpha, \alpha^2\}$ which corresponds to applying a linear transformation whose linearized polynomial $L(x) = \alpha^{12}x + \alpha^8 x^2 + \alpha^4 x^4$ to the coordinates of **a**. Thus we have $m_b(x) = lcm(x^5 + \alpha x^4 + \alpha^9, x^5 + \alpha^2 x^4 + \alpha^3, x^5 + \alpha^4 x^4 + \alpha^6) = x^{14} + \alpha^2 x^{13} + \alpha^5 x^{12} + \alpha^{11} x^{11} + \alpha x^{10} + \alpha^{11} x^9 + \alpha^8 x^8 + \alpha^{14} x^7 + \alpha^4 x^6 + \alpha^9 x^5 + \alpha^5 x^4 + \alpha^{13} x^3 + \alpha^3 x^2 + \alpha^8 x + \alpha^{13}$.

It can be shown that applying a linear transformation for which all the coefficients of the corresponding linearized polynomial are no-zero results in obtaining a characteristic polynomial with 0–1 coefficients. This observation can be used to efficiently generate of linear streams over $GF(2^n)$ without performing any multiplication operations.

Example 6. Consider the sequence defined in Example 1. If we use the basis $(1, \alpha, \alpha^2)$ instead of $(\alpha^2, \alpha, 1)$ then the minimal polynomial of the sequence is given by $x^6 + x^5 + x^4 + x^2 + 1 = lcm(\sigma_0(f(x)), \sigma_1(f(x)), \sigma_2(f(x)))$ where $\sigma_0(f(x)), \sigma_1(f(x)), \sigma_2(f(x))$ are as given in Example 4.

It is interesting to extend the results in this paper to sequences defined over Galois rings for which the linear complexity can be calculated using the extended version of Berlekamp–Massey algorithm given in [10].

References

[1] A. Biryukov, Block Ciphers and Stream Ciphers: The State of the Art, Lecture Notes in Computer Science, Proc. COSIC Summer Course 2003, Springer, Berlin, to appear.

[2] S.W. Golomb, Shift Register Sequences, Aegean Park Press, Laguna Hills, CA, 1982.

[3] G. Gong, G.Z. Xiao, Synthesis and uniqueness of m-sequences over $GF(q^n)$ as n-phase sequence over $GF(q)$, IEEE Trans. Commun. 42 (8) (1994) 2501–2505.

[4] R. Lidl, H. Niederreiter, Finite Fields (Encyclopedia of Mathematics and Its Applications), Addison-Wesley, Reading, MA, 1983.

[5] J.L. Massey, Shift-register synthesis and BCH decoding, IEEE Trans. Inform. Theory IT-15 (1) (1969) 122–127.

[6] R.J. McEliece, Finite Fields For Computer Scientists and Engineers, Kluwer Academic Publishers, Dordrecht, 1987.

[7] A.J. Menezes, P.C. van Oorschot, S.A. Vanstone, Handbook of Applied Cryptography, CRC Press, Boca Raton, FL, 1996.

[8] NESSIE: New European Schemes for Signatures, Integrity, and Encryption. See ⟨http://www.cryptonessie.org⟩.

[9] W.J. Park, J. Komo, Relationships between m-sequences over $GF(q)$ and $GF(q^m)$, IEEE Trans. Inform. Theory 35 (1) (1989) 183–186.

[10] J.A. Reeds, N.A.J. Sloane, Shift-register synthesis modulo M, Siam. J. Comput. 14 (3) (1985) 505–513.

[11] R.A. Rueppel, Analysis and Design of Stream Ciphers, Springer, Berlin, 1986.