Nonlinear Filterings of m-Sequences for Cryptographic Purposes

A. Fúster Sabater
Instituto de Física Aplicada (C.S.I.C.)
Serrano 144, 28006 Madrid (SPAIN)
E-mail: amparo@iec.csic.es

August 27, 1999

Keywords: Cryptography, linear complexity, keystream generators.

1 Introduction

The nonlinear filtering of binary m-sequences is one of the most common keystream generators in secret key Cryptography (stream ciphers). The nonlinear filtering preserves the good cryptographic characteristics of the binary m-sequences but increases the low value of their linear complexity [9]. In this work, we present a new representation of the nonlinear filterings, the so-called Sequential Decomposition in Cosets, which identifies a kth-order nonlinear filtering with the sum of the sequences associated with every cyclotomic coset of weight $\leq k$. This new vision of the nonlinear filterings lets us study the probability of choosing nonlinear filterings of m-sequences with optimal cryptographic properties as well as derive a method to design nonlinear filtering functions with a guaranteed large linear complexity.

2 Basic Concepts and Notation

A kth-order nonlinear filtering F is a kth-order nonlinear function applied to the L stages of a maximal-length Linear Feedback Shift Register. Let $\{a_n\}$

be the LFSR output sequence; then the generic element a_n can be written as

$$a_n = \alpha^n + \alpha^{2n} + \cdots + \alpha^{2^{(L-1)}n} \tag{1}$$

$\alpha \in GF(2^L)$ being a root of the LFSR primitive feedback polynomial and $\{z_n\}$ the filtered sequence.

Definition 1 *Let Z_{2^L-1} denote the set of integers $[1, \cdots, 2^L - 1]$. We consider the following equivalence relation R defined on its elements: q_1 R q_2 with q_1, $q_2 \in Z_{2^L-1}$ if there exists an integer j, $0 \leq j \leq L-1$, such that*

$$2^j \cdot q_1 \;=\; q_2 \mod 2^L - 1 \tag{2}$$

The resultant equivalence classes into which Z_{2^L-1} is partitioned are called the cyclotomic cosets modulo $2^L - 1$. The leader element of the cyclotomic coset E, denoted by E, is the smallest integer in such an equivalence class.

Definition 2 *The characteristic polynomial of a cyclotomic coset E is a polynomial P_E in the indeterminate D defined by*

$$P_E(D) \;=\; (D + \alpha^E)(D + \alpha^{2E}) \cdots (D + \alpha^{2^{(r-1)}E}) \tag{3}$$

where the degree r $(r \leq L)$ of $P_E(D)$ equals the cardinal of the cyclotomic coset E.

Definition 3 *The characteristic sequence of a cyclotomic coset E is a binary sequence, denoted by $\{S_n^E\}$, that obeys the expression*

$$\{S_n^E\} = \{\alpha^{En} + \alpha^{2En} + \cdots + \alpha^{2^{(r-1)}En}\},\ n \geq 0 \tag{4}$$

Note that $\{S_n^E\}$ is a decimation of $\{a_n\}$ formed from this sequence by taking one out of E terms.

In brief, every cyclotomic *coset* E can be characterized by its *leader element* E or its *characteristic polynomial* $P_E(D)$ or its *characteristic sequence* $\{S_n^E\}$.

3 Sequential Decomposition in Cosets

According to equation (1), the generic element z_n of the filtered sequence $\{z_n\}$ can be written as

$$\begin{aligned} z_n \;=\; F(a_n, a_{n+1}, ..., a_{n+L-1}) \;= \\ C_1\alpha^{E_1 n} + (C_1\alpha^{E_1 n})^2 + ... + (C_1\alpha^{E_1 n})^{2^{(r_1-1)}} + \\ C_2\alpha^{E_2 n} + (C_2\alpha^{E_2 n})^2 + ... + (C_2\alpha^{E_2 n})^{2^{(r_2-1)}} + \\ \vdots \\ C_i\alpha^{E_i n} + (C_i\alpha^{E_i n})^2 + ... + (C_i\alpha^{E_i n})^{2^{(r_i-1)}} + \\ \vdots \end{aligned} \tag{5}$$

r_i being the cardinal of *coset* E_i and $1 \leq i \leq N$ (N being the number of cosets of weight $\leq k$). At this point, different features can be pointed out:

- Note that every row of (5) corresponds to the *nth*-element of the *characteristic sequence* $\{C_i\alpha^{E_i n} + (C_i\alpha^{E_i n})^2 + \cdots + (C_i\alpha^{E_i n})^{2^{(r_i-1)}}\}$. The coefficient C_i determines the phase shift of this sequence referred to its characteristic phase in equation (4).

- It can be proved [6] that every coefficient $C_i \in GF(2^{r_i})$, so that as long as C_i is within its corresponding field, we shift along the sequence $\{S_n^{E_i}\}$.

- If $C_i = 0$, then $\{S_n^{E_i}\}$ does not contribute to the filtered sequence $\{z_n\}$. In that case the cyclotomic *coset* E_i is degenerate.

Now equation (5) can be rewritten as

$$\{z_n\} = \{F(a_n, a_{n+1}, \ldots, a_{n+L-1})\} = \sum_i C_i\{S_n^{E_i}\} \tag{6}$$

where $C_i\{S_n^{E_i}\}$ denotes the *characteristic sequence* $\{C_i\alpha^{E_i n} + (C_i\alpha^{E_i n})^2 + \cdots\}$ and the sum is extended to the N *cosets* of weight $\leq k$.

Thus, the output sequence of an arbitrary kth-order nonlinear filtering can be decomposed as the termwise sum of the *characteristic sequences* associated with the cyclotomic cosets of weight $\leq k$. Such a decomposition is called the

Sequential Decomposition in Cosets. It can be proved that there is a one-to-one correspondence between each nonlinear filtering F and each S.D.C.

Remark that if the degree of the *minimal polynomial* of the filtered sequence is $\sum \deg(P_{E_i}(D))$, then the fact of adding new *characteristic sequences* $\{S_n^{E_i}\}$ to its sequential decomposition is equivalent to the fact of introducing new factors $P_{E_i}(D)$ into its *minimal polynomial*. In this way, we can construct filtered sequences with maximum period and linear complexity as large as desired.

Two different applications can be obtained from this model as it will be seen in the following sections.

4 Nonlinear Filterings of m-Sequences with Optimal Cryptographic Properties

Let nfk be the number of kth-order nonlinear filterings F defined as before and nfm the number of the previous filterings whose output sequences $\{z_n\}$ have maximum linear complexity ($C_i \neq 0, \forall i$). The probability of choosing a nonlinear filtering whose output sequence has optimal cryptographic properties in which period and linear complexity is concerned is:

$$\Pr = \frac{nfm}{nfk} = \frac{(2^{r_1}-1)\,(2^{r_2}-1)\,\ldots\,(2^{r_N}-1)}{(2^{\binom{L}{k}}-1)\cdot 2^{\binom{L}{k-1}}\,\ldots\,2^{\binom{L}{1}}} \tag{7}$$

If L is a prime number (the most common cryptographic case), then nfm can be rewritten as

$$nfm = (2^L-1)^N = (2^L-1)^{\frac{1}{L}\sum_{i=1}^{k}\binom{L}{k}} = (2^L-1)^{\frac{N_k}{L}} \tag{8}$$

$$\Pr > \left(\frac{2^L-1}{2^L}\right)^{\frac{N_k}{L}} = \left(1-\frac{1}{2^L}\right)^{2^L\frac{N_k}{2^L L}} \tag{9}$$

As $k \simeq L/2$ then $N_k \simeq 2^{L-1}$. Thus,

$$\Pr > e^{-\frac{N_k}{2^L L}} \simeq e^{-\frac{1}{2L}} \tag{10}$$

For $L = 257$ a typical value in Cryptography, $\Pr > 0.998$

In addition, this kind of nonlinear filtering also has maximum period. Indeed, as those filterings contain the *characteristic sequences* $\{S_n^{E_i}\}$ associated with all the *cosets* E_i, they also contain that of *coset* E_1 whose period is $2^L - 1$.

5 A Method of Constructing Filtering Functions with Large Linear Complexity

Let A be the set of the kth-order nonlinear filterings applied to the L stages of a maximal-length LFSR. We are going to group the elements of A producing the filtered sequence $\{z_n\}$ or a shifted version $\{z_n\}^*$. From equation (6), it is clear that if we substitute C_i for $C_i\alpha^{E_i} \quad \forall i$, we will obtain $\{z_{n+1}\}$. In general,

$$C_i \to C_i\alpha^{j\,E_i} \qquad \forall i \quad \Longrightarrow \qquad \{z_n\} \to \{z_{n+j}\}$$

This fact enables us to define an equivalence relation $\sim$ on the set A such as follows: $F_1 \sim F_2$ with $F_1, F_2 \in A$ if

$$\{F_1(a_n, a_{n+1}, \ldots, a_{n+L-1})\} = \{F_2(a_n, a_{n+1}, \ldots, a_{n+L-1})\}^* \tag{11}$$

It is easy to see that the relation defined above is an equivalence relation and that the coefficients $(C_i)_1, (C_i)_2$ of two different nonlinear filterings F_1, F_2 in the same equivalence class verify

$$(C_i)_2 = (C_i)_1 \cdot \alpha^{j\,E_i} \quad \forall i \tag{12}$$

The number of elements in every equivalence class equals the period of the filtered sequence, T, so that in equation (12) the index j verifies $0 \leq j \leq T-1$.

According to the previous expression, we know that the sum of nonlinear filterings in an equivalence class turns out to be a new nonlinear filtering in a different class. The new filtering keeps the same *cosets* E_i as before. Thus, the constructing method can be stated as follows:

1. Start from a nonlinear filtering whose number of nondegenerate cosets is known to be large, for instance all the cosets of weight k [9].

2. Sum two consecutive nonlinear filterings in this class in order to jump into a new equivalence class.

3. Repeat step 2 in order to range in the different equivalence classes.

In this way, we can generate the basic terms of order k which preserve the initial linear complexity. Next, we can obtain different filtering functions with the same linear complexity by adding new terms of lower order .

Remark that:

- The procedure is general and can be applied to the generation of filtering functions with (L, k) in the range of cryptographic interest.
- As we guarantee the presence of all the cosets of weight k among which there are proper cosets, the period of the filtered sequence will be maximum.

6 Conclusions

A new representation, the so-called Sequential Decomposition in Cosets, of the kth-order nonlinear filterings of m-sequences has been developed. Emphasis is on the structural properties of the filtered sequences (period, linear complexity, minimal polynomial) that can be studied in terms of this representation. According to the sequential decomposition, likelihood that a keystream generator has optimal cryptographic properties has been considered. The work concludes with a method that allows us to generate filtering functions with maximum period and a large guaranteed linear complexity.

References

[1] A. Fúster-Sabater and L.J. García-Villalba, 'Likelihood that a pseudo-random sequence generator has optimal properties'. Electronics Letters, Vol. 34, No. 7, 1998.

[2] S.W. Golomb, 'Shift Register-Sequences', Holden-Day, San Francisco, 1967.

[3] E.J. Groth, 'Generation of Binary Sequences with Controllable Complexity'. IEEE Trans. on Information Theory, Vol. IT-17, May 1971.

[4] E.L. Key, 'An Analysis of the Structure and Complexity of Non-Linear Binary Sequence Generators'. IEEE Trans. on Information Theory, Vol. IT-22, No. 6, pp. 732-736, Nov. 1976.

[5] P.V. Kumar and R.A. Scholtz, 'Bounds on the Linear Span of Bent Sequences'. IEEE Trans. on Information Theory, Vol. IT-29, pp. 854-862, Nov. 1983.

[6] R. Lidl and H. Niederreiter, Introduction to Finite Fields and Their Applications. Cambridge University Press, 1986.

[7] J.L. Massey and S. Serconek, 'A Fourier Transform Approach to the Linear Complexity of Nonlinearly Filtered Sequences'. Advances in Cryptology-CRYPTO'94. Lecture Notes in Computer Science Vol. 839, pp. 332-340, Springer-Verlag, 1994.

[8] K.G. Peterson, 'Root Counting, the DFT and the Linear Complexity of Nonliear Filtering'. Designs, Codes and Crytography, Vol. 14, No. 3, September 1998.

[9] R.A. Rueppel, 'Analysis and Design of Stream Ciphers'. Springer-Verlag, New York, 1986.

[10] G.J. Simmons (ed.). 'Contemporary Cryptology: The Science of Information Integrity'. IEEE Press, 1991.