REVISTA MATEMÁTICA COMPLUTENSE
Volumen 11, número 1: 1998
http://dx.doi.org/10.5209/rev_REMA.1998.v11.n1.17307

Estudio de algunas secuencias pseudoaleatorias de aplicación criptográfica.

P. CABALLERO GIL and A. FÚSTER SABATER

Abstract

Pseudorandom binary sequences are required in stream ciphers and other applications of modern communication systems. In the first case it is essential that the sequences be unpredictable. The linear complexity of a sequence is the amount of it required to define the remainder. This work addresses the problem of the analysis and computation of the linear complexity of certain pseudorandom binary sequences. Finally we conclude some characteristics of the nonlinear function that produces the sequences to guarantee a minimum linear complexity.

1 Introducción y preliminares

En este trabajo estudiaremos determinadas secuencias binarias útiles en algunos de los sistemas de cifrado más difundidos hoy en día.

Muchos de los procedimientos de los sistemas de comunicación modernos requieren secuencias binarias que puedan ser generadas de una forma determinística y que a la vez parezcan aleatorias. Este tipo de secuencias reciben el nombre de pseudoaleatorias.

Concretamente los sistemas de cifrado en flujo basan su seguridad exclusivamente en las llamadas secuencias de clave cifrante, que deben tener, entre otras, la característica de pseudoaleatoriedad.

Para generar secuencias pseudoaleatorias, una de las herramientas

---

Clasificación de la A.M.S. de 1991: 11K45, 94A55.
Servicio Publicaciones Univ. Complutense. Madrid, 1998.

más conocidas es el llamado Registro de Desplazamiento con Realimentación Lineal (RDRL) de longitud L, cuyo funcionamiento se muestra en la figura 1, [3].

A intervalos de tiempo periódicos determinados por un reloj, los contenidos de la etapa s_i son transferidos a la etapa s_{i-1} obteniéndose un nuevo estado $(s_{j-1}, s_{j-2}, ..., s_{j-L})$. Para obtener cada nuevo valor de s_j se calcula el valor binario dado por la relación de recurrencia lineal $\sum_{k=1}^{L} c_k s_{j-k}$. Los coeficientes binarios c_k de esta expresión suelen definirse mediante el polinomio de grado L, $C(x) = 1 + c_1 x + c_2 x^2 + \cdots + c_L x^L$ (con c_L=1) llamado polinomio de realimentación del RDRL.

Las secuencias producidas por generadores lineales de este tipo verifican muchas de las propiedades ideales exigibles a las secuencias que se utilizan como clave cifrante, a saber: pseudoaleatoriedad, facilidad de implementación y gran período. Con respecto a esta última, el período de las secuencias producidas por determinados RDRLs de longitud L es $2^L - 1$ [7].

Ahora bien, la linealidad de los RDRLs hace de ellos generadores muy predecibles. Concretamente $2L$ términos seguidos de la secuencia producida se pueden utilizar fácilmente para describir completamente el generador [11]. A la cantidad de secuencia necesaria para definir el resto se le conoce con el nombre de complejidad lineal de la secuencia, y es la medida de impredecibilidad de las secuencias más utilizada habitualmente.

Con la intención de mantener el cumplimiento de las demás propiedades y al mismo tiempo resolver el inconveniente indicado en el párrafo anterior, se puede definir un generador no lineal conocido como filtrado no lineal y descrito en la figura 2.

El filtrado no lineal consiste básicamente en una función no lineal f que a cada golpe de reloj se aplica sobre los contenidos de las L etapas de un RDRL produciendo de esta forma un dígito de la secuencia filtrada $z_j = f(s_{j-1}, s_{j-2}, ..., s_{j-L})\ \forall j = 0, 1, ...$

Dada una secuencia binaria cualquiera siempre se pueden encontrar varios RDRLs de polinomios de realimentación distintos y varios filtrados no lineales diferentes que permiten generarla a partir de estados iniciales adecuados. En particular siempre puede utilizarse como generador de una secuencia de período p el RDRL de polinomio de realimentación

$1 + x^p$ al que se conoce como RDRL cíclico puro.

La complejidad lineal de una secuencia coincide exactamente con la longitud del menor RDRL que permite generar la secuencia, generador que recibe el nombre de equivalente lineal de la secuencia.

Se conocen varias publicaciones sobre el estudio de la complejidad lineal. Entre ellas destaca el algoritmo de Berlekamp-Massey [8], que se basa en el análisis de los dígitos de la secuencia. El caso de las secuencias filtradas, aunque resulta bastante difícil según se indica en [11, pag. 57], ha sido estudiado por diversos autores entre los que se encuentran las referencias [5], [6], [2] y [10].

Dada una secuencia binaria cualquiera $\{z_n\}$, se define su función generatriz *Z(x)* como la serie infinita $Z(x) = \sum_{j=0}^{\infty} z_j x^j$. Mediante el producto entre esta función y el polinomio de realimentación *C(x)* de un RDRL que permite generar la secuencia se puede construir el polinomio de grado menor que L, *P(x)=C(x)Z(x)*. Al polinomio *P(x)* se le conoce como polinomio de estado inicial por corresponderse sus coeficientes con los contenidos del estado inicial del RDRL de polinomio de realimentación *C(x)*.

La expresión de la función generatriz como el cociente entre ambos polinomios $\frac{P(x)}{C(x)}$ se puede utilizar para calcular la complejidad lineal de la secuencia $\{z_n\}$ teniendo en cuenta que si $mcd\{P(x), C(x)\}= 1$, entonces el polinomio *C(x)* es el polinomio de realimentación del equivalente lineal de $\{z_n\}$, y por tanto su grado coincide exactamente con la complejidad lineal de $\{z_n\}$.

Una de las caracterizaciones más útiles de las secuencias filtradas viene dada por la Forma Algebráica Normal (FAN) de la función no lineal *f*, que la expresa como suma de productos:

$$f(x_1, x_2, ..., x_L) = a_{1,1}x_1 + a_{2,2}x_2 + \cdots + a_{L,L}x_L + a_{1,2}x_1x_2 + \cdots + a_{L-1,L}x_{L-1}x_L + \cdots + a_{1,2,...,L}x_1x_2\cdots x_L. \quad (1)$$

En particular, dada una secuencia filtrada $\{z_n\}$, si se denota como $\widehat{s_{n-t}} \in \{0,1\}^{2^L-1}$ al vector formado por el primer período de la secuencia $\{s_{n-t}\}$ $\forall t \in \{1, 2, ..., L\}$ generada a partir del RDRL de partida, y como $\widehat{z_n} \in \{0,1\}^{2^L-1}$ al vector formado por el primer período de la secuencia filtrada $\{z_n\}$, se tiene la siguiente expresión de coeficientes binarios:

$$\widehat{z_n} = a_{1,1}\widehat{s_{n-1}} + a_{2,2}\widehat{s_{n-2}} + \cdots + a_{L,L}\widehat{s_{n-L}} + a_{1,2}\widehat{s_{n-1}}\widehat{s_{n-2}} +$$

$$\cdots + a_{1,2,\ldots,L}\widehat{s_{n-1}}\widehat{s_{n-2}}\cdots\widehat{s_{n-L}}. \qquad (2)$$

El orden del filtrado no lineal que produce la secuencia $\{z_n\}$ viene dado por el máximo de los órdenes de los términos que aparecen en la expresión anterior con coeficientes unitarios.

2 Resultados principales

En esta sección analizamos el caso de los filtrados no lineales de segundo orden, es decir, aquellos tales que su FAN está formada por términos lineales y productos de dos etapas dentro de un mismo estado del RDRL:

$$z_j = f(s_{j-1}, s_{j-2}, \ldots, s_{j-L}) = a_{1,1}s_{j-1} + a_{2,2}s_{j-2} + \cdots + a_{L,L}s_{j-L} +$$

$$a_{1,2}s_{j-1}s_{j-2} + a_{1,3}s_{j-1}s_{j-3} + \cdots + a_{L-1,L}s_{j-(L-1)}s_{j-L}. \qquad (3)$$

Este tipo de filtrados ha sido estudiado en [4] y [1], mostrándose también algunos resultados parciales en [9]. En este trabajo abordamos el caso general.

Suponemos sin pérdida de generalidad que el estado inicial del RDRL es $(s_{-1}, s_{-2}, \ldots, s_{-(L-1)}, s_{-L})$= (0,0,...,0,1).

Si en la expresión (3) de z_j se extraen como factores comunes de cada sumando las etapas mayores que forman cada producto se obtiene la siguiente:

$$z_j = s_{j-1}(a_{1,1} + a_{1,2}s_{j-2} + a_{1,3}s_{j-3} + \cdots + a_{1,L}s_{j-L}) +$$
$$s_{j-2}(a_{2,2} + a_{2,3}s_{j-3} + a_{2,4}s_{j-4} + \cdots + a_{2,L}s_{j-L}) + \cdots +$$
$$s_{j-(L-1)}(a_{L-1,L-1} + a_{L-1,L}s_{j-L}) +$$

$$s_{j-L}a_{L,L} = \sum_{i=1}^{L} s_{j-i} \sum_{l=0}^{L-i} a_{i,i+l}s_{j-(i+l)}, \qquad (4)$$

dado que $s_j^2 = s_j$, $\forall j$ por tratarse de datos binarios.

De esta manera, la función generatriz de $\{z_n\}$ admite la siguiente expresión:

$$Z(x) = \sum_{j=0}^{\infty} z_j x^j = \sum_{j=0}^{\infty} x^j \sum_{i=1}^{L} s_{j-i} \sum_{l=0}^{L-i} a_{i,i+l}s_{j-(i+l)}. \qquad (5)$$

Teniendo en cuenta la condición impuesta sobre el estado inicial se tiene que (sustituyendo j-i por j en (5)):

$$Z(x) = \sum_{j=0}^{\infty}\sum_{i=1}^{L} x^{j+i}s_j \sum_{l=0}^{L-i} a_{i,i+l}s_{j-l} = \sum_{i=1}^{L}\sum_{l=0}^{L-i} a_{i,i+l}x^i \sum_{j=0}^{\infty} x^j s_j s_{j-l}. \quad (6)$$

Si denotamos como $S_l(\mathrm{x})$ a la serie infinita $\sum_{j=0}^{\infty} x^j s_j s_{j-l}$, gracias a la relación de recurrencia lineal se obtiene que:

$$S_l(x) = \sum_{j=0}^{\infty} x^j s_{j-l} \sum_{k=1}^{L} c_k s_{j-k} \quad (7)$$

De nuevo, usando la restricción impuesta sobre el estado inicial se consigue la siguiente expresión (sustituyendo j-k por j en (7)):

$$S_l(x) = \sum_{j=0}^{\infty}\sum_{k=1}^{L} x^{j+k}s_{j+k-l}c_k s_j = \sum_{k=1}^{L} c_k x^k \sum_{j=0}^{\infty} x^j s_j s_{j-(l-k)} =$$

$$\sum_{k=1}^{L} c_k x^k S_{l-k}(x). \quad (8)$$

De donde se obtiene el siguiente sistema de L ecuaciones:

$$\begin{cases} S_1(x) = c_1 x S_0(x) + c_2 x^2 S_{-1}(x) + \cdots + c_L x^L S_{1-L}(x) \\ S_2(x) = c_1 x S_1(x) + c_2 x^2 S_0(x) + \cdots + c_L x^L S_{2-L}(x) \\ \vdots \\ S_L(x) = c_1 x S_{L-1}(x) + c_2 x^2 S_{L-2}(x) + \cdots + c_L x^L S_0(x) \end{cases} \quad (9)$$

Las incógnitas de subíndice negativo se pueden calcular, utilizando de nuevo la restricción sobre el estado inicial, de la siguiente forma:

$$S_{-i}(x) = \sum_{j=0}^{\infty} x^j s_j s_{j+i} = x^{-i}(s_{-i}s_0 + s_{1-i}s_1 x + s_{2-i}s_2 x^2 + \cdots + s_{-1}s_{-1+i}x^{-1+i} + s_0 s_i x^i + s_1 s_{1+i}x^{1+i} + \cdots) =$$

$$x^{-i}\sum_{j=0}^{\infty} x^j s_{j-i}s_j = x^{-i}S_i(x). \quad (10)$$

Por tanto, sustituyendo estas incógnitas en el sistema de ecuaciones (9) se tiene el siguiente:

$$\begin{cases} S_1(x) = c_1 x S_0(x) + c_2 x S_1(x) + \cdots + c_L x S_{L-1}(x) \\ S_2(x) = c_1 x S_1(x) + c_2 x^2 S_0(x) + \cdots + c_L x^2 S_{L-2}(x) \\ \vdots \\ S_L(x) = c_1 x S_{L-1}(x) + c_2 x^2 S_{L-2}(x) + \cdots + c_L x^L S_0(x) \end{cases} \tag{11}$$

Si se despeja la incógnita S_0(x) de cada ecuación se obtiene la forma definitiva del sistema de ecuaciones:

$$\begin{cases} c_1 x S_0(x) = (1 + c_2 x) S_1(x) + c_3 x S_2(x) + \cdots + c_L x S_{L-1}(x) \\ c_2 x^2 S_0(x) = (c_1 x + c_3 x^2) S_1(x) + \cdots + c_L x^2 S_{L-2}(x) \\ \vdots \\ c_L x^L S_0(x) = c_{L-1} x^{L-1} S_1(x) + c_{L-2} x^{L-2} S_2(x) + \cdots + S_L(x) \end{cases} \tag{12}$$

El sistema (12) en forma matricial queda expresado como sigue:

$$\begin{pmatrix} 1 + c_2 x & c_3 x & \cdots & c_L x & 0 \\ c_1 x + c_3 x^2 & 1 + c_4 x^2 & \cdots & 0 & 0 \\ c_2 x^2 + c_4 x^3 & c_1 x + c_5 x^3 & \cdots & 0 & 0 \\ & & \vdots & & \\ c_{L-4} x^{L-4} + c_{L-2} x^{L-3} & c_{L-5} x^{L-5} + c_{L-1} x^{L-3} & \cdots & 0 & 0 \\ c_{L-3} x^{L-3} + c_{L-1} x^{L-2} & c_{L-4} x^{L-4} + c_L x^{L-2} & \cdots & 0 & 0 \\ c_{L-2} x^{L-2} + c_L x^{L-1} & c_{L-3} x^{L-3} & \cdots & 1 & 0 \\ c_{L-1} x^{L-1} & c_{L-2} x^{L-2} & \cdots & c_1 x & 1 \end{pmatrix}.$$

$$\begin{pmatrix} S_1(x) \\ S_2(x) \\ \vdots \\ S_{L-1}(x) \\ S_L(x) \end{pmatrix} = \begin{pmatrix} c_1 x \\ c_2 x^2 \\ \vdots \\ c_{L-1} x^{L-1} \\ c_L x^L \end{pmatrix} \cdot S_0(x). \tag{13}$$

El valor de S_0(x) se puede calcular mediante la relación de recurrencia lineal:

$$S_0(x) = \sum_{j=0}^{\infty} x^j s_j = \sum_{j=0}^{\infty} x^j \sum_{k=1}^{L} c_k s_{j-k}. \tag{14}$$

Utilizando la restricción impuesta sobre el estado inicial se obtiene lo siguiente:

$$S_0(x) = c_L + \sum_{k=1}^{L} c_k x^k \sum_{j=0}^{\infty} x^j s_j = c_L + \sum_{k=1}^{L} c_k x^k S_0(x). \qquad (15)$$

De donde, por ser c_L=1, se deduce finalmente la expresión de S_0(x):

$$S_0(x) = \frac{1}{1 + \sum_{k=1}^{L} c_k x^k}. \qquad (16)$$

De todo el desarrollo teórico anterior se desprende el siguiente resultado que proporciona una expresión de la función generatriz de las secuencias filtradas mediante funciones de orden dos.

Teorema 2.1. *Dadas una secuencia binaria* $\{s_n\}$ *producida mediante un RDRL de polinomio de realimentación* $C(x) = 1 + c_1 x + \cdots + c_L x^L$ *con coeficientes binarios y* $c_L = 1$, *y una función no lineal f de orden dos definida mediante* $f(x_1, x_2, ..., x_L) = a_{1,1}x_1 + \cdots + a_{L,L}x_L + a_{1,2}x_1x_2 + \cdots + a_{L-1,L}x_{L-1}x_L$ *también de coeficientes binarios, se tiene que la función generatriz de la secuencia filtrada* $\{z_n\}$=$\{f(s_{n-1}, s_{n-2}, ..., s_{n-L})\}$ *se puede expresar de la forma:*

$$Z(x) = \sum_{i=1}^{L} \sum_{l=0}^{L-i} a_{i,i+l} x^i S_l(x), \qquad (17)$$

donde $S_l(x)$ *queda definida mediante el sistema de ecuaciones (13).*

Desarrollando el determinante de la matriz de coeficientes del sistema de ecuaciones (13) por la última columna y utilizando que c_L=1, se observa que en la diagonal secundaria del menor de orden *L-1* resultante están presentes las *L-1* potencias sucesivas de x, mientras que en todos los elementos de la diagonal principal aparece un 1. Por ambas razones se obtiene, aplicando la regla de Cramer, que el determinante de la matriz de los coeficientes es necesariamente un polinomio de grado $\sum_{i=1}^{L-1} i = \frac{L(L-1)}{2}$. Concretamente este polinomio resulta ser:

* si L es impar, el producto de $\frac{L-1}{2}$ polinomios irreducibles de grado L,

* si L es par, el producto de $\frac{L}{2} - 1$ polinomios irreducibles de grado L y un polinomio irreducible de grado $\frac{L}{2}$.

Por otra parte, los numeradores de los valores asociados a cada incógnita $S_l(x)$, $l = 1, 2, ..., L$ son polinomios de grados menores o iguales que $\frac{L(L-1)}{2}$ multiplicados por S_0(x). Si se denota como $l_0 = \min\{l/ c_l \neq 0\}$ y se definen los valores $p_{l_0} = l_0$ y $p_l \geq l_0$ $\forall l \neq l_0$, de cada uno de esos polinomios siempre se puede sacar como factor común la potencia x^{p_l} debido a la forma de la matriz columna del término independiente.

En conclusión, las incógnitas $S_l(x)$ se pueden expresar de la forma siguiente:

$$S_l(x) = \frac{x^{p_l} \sum_{j=0}^{\frac{L(L-1)}{2}-p_l} b_j x^j}{\sum_{j=0}^{\frac{L(L+1)}{2}} d_j x^j}, \tag{18}$$

siendo $b_j, d_j \in GF(2)$.

Según el teorema 2.1 se tiene que

$$Z(x) = \sum_{i=1}^{L} \sum_{l=0}^{L-i} a_{i,i+l} x^i S_l(x). \tag{19}$$

Si se denota como $i_0 = \min\{i/ \exists a_{i,i+l} \neq 0\}$, siempre se puede sacar $x^{i_0+l_0}$ como factor común en el numerador de $Z(x)$. Por tanto, según el cálculo de la complejidad lineal a partir de la expresión como cociente entre polinomios primos entre sí que define la función generatriz, se puede deducir que la complejidad lineal de la secuencia filtrada $\{z_n\}$ es mayor o igual que $i_0 + l_0$, tal y como indica el siguiente resultado:

Corolario 2.1. *En las condiciones del teorema 2.1, la función generatriz de la secuencia filtrada* $\{z_n\} = \{f(s_{n-1}, s_{n-2}, ..., s_{n-L})\}$ *se puede expresar de la forma:*

$$Z(x) = x^{i_0+l_0} \sum_{i=1}^{L} \sum_{l=0}^{L-i} a_{i,i+l} x^{i-i_0} \frac{x^{p_l-l_0} \sum_{j=0}^{\frac{L(L-1)}{2}-p_l} b_j x^j}{\sum_{j=0}^{\frac{L(L+1)}{2}} d_j x^j}, \tag{20}$$

siendo $l_0 = \min\{l/\ c_l \neq 0\}$, $p_{l_0} = l_0$, $p_l \geq l_0\ \forall l \neq l_0$, $i_0 = \min\{i/\ \exists a_{i,i+l} \neq 0\}$, *y* b_j, d_j *coeficientes binarios.*

Por tanto, la complejidad lineal de la secuencia filtrada $\{z_n\}$ *es mayor que* $i_0 + l_0$.

El mayor valor que puede tomar la expresión $i_0 + l_0$ es *2L-1*, que resulta en el caso concreto en que el polinomio de realimentación del RDRL de partida es $C(x) = 1 + x^L$ y la función no lineal es $f(s_{j-1}, s_{j-2}, ..., s_{j-L}) = a_{L-1,L-1}s_{j-(L-1)} + a_{L,L}s_{j-L} + s_{j-(L-1)}s_{j-L}$.

Observación 2.1. *Para acercarnos al valor máximo de la cota de la complejidad lineal* $i_0 + l_0$, *basta considerar aquellos RDRLs cuyos polinomios de realimentación tengan únicamente potencias altas* ($i_0 \to L$) *y aquellas funciones no lineales cuyos productos estén formados exclusivamente por etapas altas de la secuencia* ($l_0 \to L - 1$).

Considérese ahora la función no lineal consistente en una suma de productos de orden dos de etapas equidistantes:

$$f(s_{j-1}, s_{j-2}, ..., s_{j-L}) = a_{1,1+\delta}s_{j-1}s_{j-(1+\delta)} + a_{2,2+\delta}s_{j-2}s_{j-(2+\delta)} + \cdots + a_{L-\delta,L}s_{j-(L-\delta)}s_{j-L}. \quad (21)$$

En este caso, según el teorema 2.1 la función generatriz queda de la forma:

$$Z(x) = S_\delta(x)\left[a_{1,1+\delta}x + a_{2,2+\delta}x^2 + \cdots + a_{L-\delta,L}x^{L-\delta}\right]. \quad (22)$$

Si se denota como $i_0 = \min\{i/\ a_{i,i+\delta} = 1\}$, se tiene que en el numerador de *Z(x)* siempre se pueden sacar como factores comunes los polinomios x^{i_0} y $1 + a_{i_0+1,i_0+1+\delta}x + a_{i_0+2,i_0+2+\delta}x^2 + \cdots + a_{L-\delta,L}x^{L-\delta-i_0}$. Por tanto, si este último polinomio es primo con el denominador de *Z(x)*, entonces siempre se puede garantizar una complejidad lineal de valor $L - \delta + p_\delta$. Esto se muestra en el siguiente resultado:

Teorema 2.2. *Dadas una secuencia binaria* $\{s_n\}$ *producida mediante un RDRL de polinomio de realimentación* $C(x) = 1 + c_1x + \cdots + c_Lx^L$ *con coeficientes binarios y* $c_L = 1$, *y una función no lineal f de orden dos definida mediante*

$$f(x_1, x_2, ..., x_L) = a_{1,1+\delta}x_1x_{1+\delta} + a_{2,2+\delta}x_2x_{2+\delta} +$$

$$\cdots + a_{L-\delta,L}x_{L-\delta}x_L, \tag{23}$$

también de coeficientes binarios, se tiene que la función generatriz de la secuencia filtrada $\{z_n\}=\{f(s_{n-1}, s_{n-2}, \ldots, s_{n-L})\}$ *se puede expresar de la forma:*

$$Z(x) = x^{i_0+p_\delta}(1 + a_{i_0+1,i_0+1+\delta}x + a_{i_0+2,i_0+2+\delta}x^2 + \cdots + a_{L-\delta,L}x^{L-\delta-i_0}) \frac{\sum_{j=0}^{\frac{L(L-1)}{2}-p_\delta} b_j x^j}{\sum_{j=0}^{\frac{L(L+1)}{2}} d_j x^j}, \tag{24}$$

siendo $p_\delta \geq \min\{l/\ c_l \neq 0\}$, $i_0 = \min\{i/\ \exists a_{i,i+\delta} = 1\}$, *y* b_j, d_j *coeficientes binarios.*

Corolario 2.2. *En las condiciones del teorema 2.2, si el polinomio* $1 + a_{i_0+1,i_0+1+\delta}x + a_{i_0+2,i_0+2+\delta}x^2 + \cdots + a_{L-\delta,L}x^{L-\delta-i_0}$ *es primo con el polinomio* $\sum_{j=0}^{\frac{L(L+1)}{2}} d_j x^j$, *entonces la complejidad lineal de la secuencia filtrada* $\{z_n\}$ *es mayor o igual que* $L - \delta + p_\delta$.

Observación 2.2. *Si* L *es impar, la condición del corolario 2.2 siempre se verifica ya que el denominador está formado por un producto de polinomios irreducibles de grado* L. *Por el contrario, si* L *es par, la condición anterior se verifica sólo cuando el polinomio* $1 + a_{i_0+1,i_0+1+\delta}x + a_{i_0+2,i_0+2+\delta}x^2 + \cdots + a_{L-\delta,L}x^{L-\delta-i_0}$ *es primo con el polinomio de grado* $\frac{L}{2}$ *recíproco del polinomio* $\prod_{i=0}^{\frac{L}{2}-1}(x + \alpha^{2^i(2^{\frac{L}{2}}+1)})$, *siendo* α *una raíz del polinomio* $x^L + c_1x^{L-1} + \cdots + c_{L-1}x + c_L$.

3 Ejemplo ilustrativo

En esta sección desarrollamos un ejemplo completo donde se ponen de manifiesto los resultados teóricos anteriores a la vez que se demuestran de forma práctica las conclusiones obtenidas.

Considérese el RDRL de polinomio de realimentación $C(x) = 1 + x^5 + x^6$ mostrado en la figura 3.

La relación de recurrencia lineal es $s_j = s_{j-5} + s_{j-6}$ y los coeficientes son $c_i = 0\ \forall i \neq 5, 6,\ c_5 = c_6 = 1$.

La secuencia producida por este RDRL a partir del estado inicial $(s_{-1},\ s_{-2},\ s_{-3},\ s_{-4},\ s_{-5},\ s_{-6})$= (0,0,0,0,0,1) es la secuencia de período 2^6-1=63 cuyo primer período es

$$1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 1,\ 1,\ 1,\ 1,\ 0,\ 1,\ 0,$$
$$0,\ 0,\ 1,\ 1,\ 1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 0,\ 1,\ 1,\ 0,\ 1,$$
$$1,1,0,1,1,0,0,1,1,0,1,0,1,0,1,1,1,1,1. \qquad (25)$$

La secuencia (25) asimismo puede obtenerse mediante el RDRL cíclico puro de polinomio de realimentación $C(x) = 1 + x^{63}$ y polinomio de estado inicial dado por ese primer período de manera que su función generatriz se puede expresar de la forma siguiente:

$$Z(x) = \frac{\frac{1+x^5+x^6+x^{10}+x^{12}+x^{15}+x^{16}+x^{17}+x^{18}+x^{20}+x^{24}+x^{25}+x^{26}+x^{29}+x^{32}}{1+x^{63}} + \frac{x^{34}+x^{35}+x^{37}+x^{38}+x^{39}+x^{41}+x^{42}+x^{45}+x^{46}+x^{48}+x^{50}+x^{52}+x^{53}+x^{54}+x^{55}+x^{56}+x^{57}}{1+x^{63}}}{} =$$
$$\frac{1}{1+x^5+x^6}. \qquad (26)$$

De ahí se deduce que efectivamente la complejidad lineal de la secuencia (3.1) es 6 y su equivalente lineal es el RDRL de polinomio de realimentación $1 + x^5 + x^6$ con estado inicial (0, 0, 0, 0, 0, 1).

Consideremos ahora una función no lineal de orden dos aplicada sobre las etapas de este RDRL y analicemos la secuencia filtrada cuyos términos vienen dados por

$$z_j = f(s_{j-1},\ s_{j-2},\ s_{j-3},\ s_{j-4},\ s_{j-5},\ s_{j-6}) = \sum_{i=1}^{6}\sum_{l=0}^{6-i} a_{i,i+l} s_{j-i} s_{j-(i+l)}. \qquad (27)$$

Del análisis teórico realizado en la sección anterior se obtiene el siguiente sistema de ecuaciones independiente de la función no lineal específica:

$$\begin{pmatrix} 1 & 0 & 0 & x & x & 0 \\ 0 & 1 & x^2 & x^2 & 0 & 0 \\ 0 & x^3 & 1+x^3 & 0 & 0 & 0 \\ x^4 & x^4 & 0 & 1 & 0 & 0 \\ x^5 & 0 & 0 & 0 & 1 & 0 \\ x^5 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} S_1(x) \\ S_2(x) \\ S_3(x) \\ S_4(x) \\ S_5(x) \\ S_6(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ x^5 \\ x^6 \end{pmatrix} \frac{1}{1+x^5+x^6}. \qquad (28)$$

El determinante de la matriz de los coeficientes del sistema (28), tal y como ya se ha dicho, es el producto de varios polinomios irreducibles de grados 6 y 3:

$$\begin{vmatrix} 1 & 0 & 0 & x & x & 0 \\ 0 & 1 & x^2 & x^2 & 0 & 0 \\ 0 & x^3 & 1+x^3 & 0 & 0 & 0 \\ x^4 & x^4 & 0 & 1 & 0 & 0 \\ x^5 & 0 & 0 & 0 & 1 & 0 \\ x^5 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} =$$

$$1 + x^3 + x^8 + x^{10} + x^{11} + x^{12} + x^{15} =$$

$$(1 + x + x^3)\,(1 + x^2 + x^4 + x^5 + x^6)\,(1 + x + x^4 + x^5 + x^6). \qquad (29)$$

Por otra parte, las soluciones del sistema de ecuaciones (28), aplicando la regla de Cramer resultan de la forma siguiente:

$$\begin{aligned}
S_1(x) &= \frac{x^6(1+x^3+x^5+x^6+x^9)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}, \\
S_2(x) &= \frac{x^{12}(1+x+x^2)(1+x)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}, \\
S_3(x) &= \frac{x^{15}}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}, \\
S_4(x) &= \frac{x^{10}(1+x^3+x^5)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}, \\
S_5(x) &= \frac{x^5(1+x^3+x^6+x^8+x^9+x^{10})}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}, \\
S_6(x) &= \frac{x^6(1+x^3+x^5+x^{12}+x^{14}+x^{15})}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}.
\end{aligned}$$

(30)

Al permanecer todos los denominadores de los $S_i(x)$ en (30) de grado 15 se deduce que para el RDRL escogido cualquier función que sea un solo producto de orden dos tiene complejidad lineal máxima de valor 15.

Consideramos ahora las cuatro funciones no lineales siguientes correspondientes a las señaladas respectivamente como a), b), c) y d) en la figura 4:

$$\begin{aligned}
z_j^1 &= f_1(s_{j-1},\ s_{j-2},\ s_{j-3},\ s_{j-4},\ s_{j-5},\ s_{j-6}) = s_{j-3}\,s_{j-6} + s_{j-4}\,s_{j-5}, \\
z_j^2 &= f_2(s_{j-1},\ s_{j-2},\ s_{j-3},\ s_{j-4},\ s_{j-5},\ s_{j-6}) = s_{j-3}\,s_{j-5} + s_{j-5}\,s_{j-6}, \\
z_j^3 &= f_3(s_{j-1},\ s_{j-2},\ s_{j-3},\ s_{j-4},\ s_{j-5},\ s_{j-6}) = s_{j-4}\,s_{j-6} + s_{j-5}\,s_{j-6},
\end{aligned}$$

$$z_j^4 = s_{j-1}\, s_{j-4} + s_{j-2}\, s_{j-5} + s_{j-3} s_{j-6}. \tag{31}$$

Para cada una de ellas, las funciones generatrices correspondientes quedan de la siguiente forma:

$$Z^1(x) = x^3 S_3(x) + x^4 S_1(x) = \frac{x^{10}(1+x^3)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)},$$

$$Z^2(x) = x^3 S_2(x) + x^5 S_1(x) = \frac{x^{11}(1+x^3+x^4+x^5+x^6+x^7+x^9)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)},$$

$$Z^3(x) = x^4 S_2(x) + x^5 S_1(x) = \frac{x^{11}(1+x^3+x^6+x^8+x^9)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)},$$

$$Z^4(x) = x S_3(x) + x^2 S_3(x) + x^3 S_3(x) = \frac{x^{16}(1+x+x^2)}{(1+x+x^3)\,(1+x^2+x^4+x^5+x^6)\,(1+x+x^4+x^5+x^6)(1+x^5+x^6)}.$$

(32)

De donde se concluye que salvo la primera secuencia filtrada que tiene una complejidad lineal de valor 15, las otras tres secuencias tienen una complejidad lineal máxima de valor 21.

Obsérvese que las dos últimas funciones se corresponden con los casos recomendados en el desarrollo teórico: por un lado f_3 es una función cuyos productos están formados exclusivamente por términos altos, mientras que por otro, f_4 es una suma de productos de fases equidistantes.

4 Conclusiones

En este trabajo hemos abordado el estudio de la complejidad lineal de los filtrados no lineales de segundo orden.

Hemos obtenido expresiones para las funciones generatrices de las secuencias producidas en el caso general y en el caso particular de la suma de productos de etapas equidistantes. A raíz de dichas expresiones hemos deducido cotas inferiores de la complejidad lineal y algunas recomendaciones para el diseño de los generadores en cuestión.

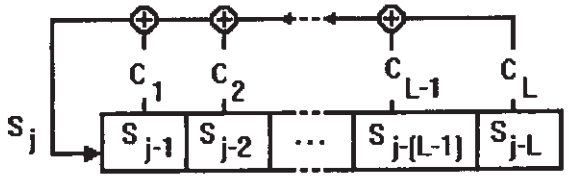

Figura 1: RDRL

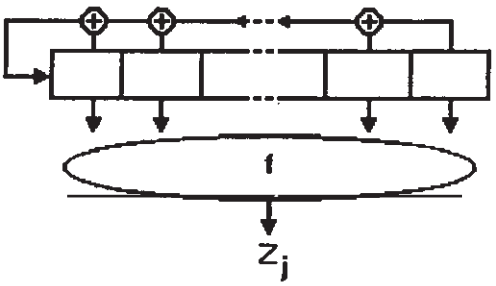

Figura 2: Filtrado no lineal

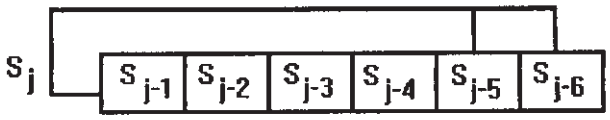

Figura 3: Ejemplo $L = 6$

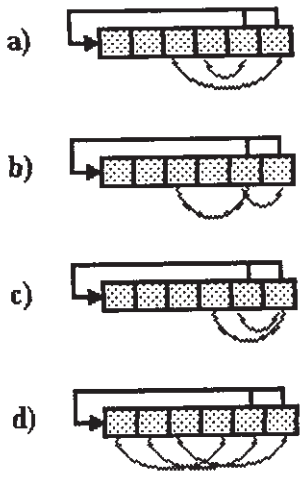

Figura 4: Ejemplos de filtrados no lineales

References

[1] P. Caballero Gil, A. Fúster Sabater, "Equivalente Lineal Descompuesto del Filtrado no Lineal y Resultados sobre Complejidad Lineal", Revista de la Academia Canaria de Ciencias, Vol. VI, No. 1, 1994.

[2] A. Fúster Sabater, P. Caballero Gil, "On the Linear Complexity of Nonlinearly Filtered PN-Sequences", Advances in Cryptology-ASIACRYPT'94, Lecture Notes in Computer Science Vol. 917, Springer-Verlag, 1995.

[3] S.W. Golomb, "Shift Register Sequences", San Francisco: Holden-Day, 1967. Edición revisada, Aegean Park Press, Laguna Hills, 1982.

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

[5] E.L. Key, "An Analysis of the Structure and Complexity of Nonlinear Binary Sequence Generators", IEEE Transactions on Information Theory, Vol. IT-22, Nov. 1976.

[6] P.V. Kumar, R.A. Scholtz, "Bounds on the Linear Span of Bent Sequences", IEEE Transactions on Information Theory, Vol. IT-29, Nov. 1983.

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

[8] J.L. Massey, "Shift-Register Synthesis and BCH Decoding", IEEE Transactions on Information Theory, Vol. IT-15, Jan. 1969.

[9] J.L. Massey, 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, Springer-Verlag, 1994.

[10] K.G. Paterson, "New Lower Bounds on the Linear Complexity of Nonlinearly Filtered m-Sequences", enviado a IEEE Transactions on Information Theory, 1996.

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

Dpto. de Estadística, I. O. y C.
Facultad de Matemáticas.
Universidad de La Laguna.
38271 La Laguna. Tenerife.

Laboratorio de Criptografía. C.S.I.C.
Serrano 144.
28006 Madrid.

Recibido: 14 de Enero de 1997
Revisado: 15 de Marzo de 1997