Calculando Equivalentes Débiles de Filtrados No Lineales

A. Fúster-Sabater
Inst. Tecnologías Físicas y de la Información
Consejo Superior Investigaciones Científicas
Email: amparo@iec.csic.es

P. Caballero-Gil
Departamento de Informática
Universidad de La Laguna
Email: pcaballe@ull.edu.es

***Resumen*—Dada una secuencia binaria generada con una función filtro no lineal aplicada sobre un registro de desplazamiento lineal o LFSR (Linear Feedback Shift Register), siempre es posible generar la misma secuencia a partir de cualquier otro LFSR de la misma longitud mediante el uso de otra función filtro. El problema aún sin resolver es el cálculo de la función filtro equivalente para cada LFSR. En este trabajo se analiza el caso en el que se utiliza un LFSR recíproco para generar un equivalente de un filtrado no lineal de partida mediante el cálculo de la relación específica entre ambas funciones filtro. Además, desde un punto de vista criptográfico, el método aquí desarrollado permite determinar filtrados equivalentes inseguros de otros que son aparentemente seguros. Este resultado puede considerarse como una demostración de que, para que un generador de secuencia cifrante pueda llegar a ser considerado totalmente seguro, debe cumplir diferentes propiedades, algunas de las cuales son aún desconocidas.**

***Palabras clave*—Criptografía de clave secreta (*Secret key cryptography*), Generador pseudoaleatorio (*Pseudorandom generator*), Cifrado en flujo (*Stream cipher*), Filtrado no lineal, LFSR.**

I. Introducción

Un cifrado en flujo está compuesto por un generador de secuencia cifrante cuya secuencia pseudoaleatoria de salida se suma módulo 2 con los bits del texto en claro. Dado que la operación de cifrado es muy rápida, se considera que los cifrados en flujo son en general más eficientes que cualquier otro tipo de cifrado. Esta es la razón principal por la que resultan especialmente adecuados para las comunicaciones inalámbricas como las de telefonía móvil, Wi-Fi o Bluetooth.

Una de las formas más habituales de construir generadores de secuencia cifrante es mediante el uso de un generador pseudoaleatorio conocido como registro de desplazamiento lineal o LFSR (Linear Feedback Shift Register) [9], cuya secuencia de salida es la imagen de una función lineal aplicada sobre sus estados sucesivos. Si se cumplen determinadas condiciones, esta estructura produce secuencias con características muy deseables para uso criptográfico. En particular, si su polinomio característico es primitivo, la secuencia generada, llamada *m*-secuencia, tiene algunas propiedades muy útiles, tales como un período grande y una buena distribución estadística de ceros y unos. Sin embargo, la secuencia producida por un LFSR nunca debe utilizarse como secuencia cifrante en un cifrado en flujo porque la linealidad inherente de la estructura podría ser fácilmente utilizada para romper el cifrado.

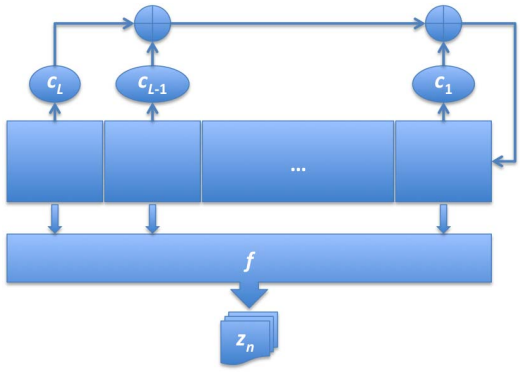

Figura 1. Filtrado No Lineal

Un interesante generador de secuencia cifrante basado en un LFSR es el filtrado no lineal, que produce una secuencia cifrante a partir de la salida de una función booleana no lineal aplicada sobre los estados de un LFSR. En particular, un filtrado no lineal consta de dos partes (ver Fig. 1):

1. Un LFSR de longitud L, con polinomio característico $P(x) = x^L + c_1 \cdot x^{L-1} + \cdots + c_{L-1} \cdot x + c_L$ de coeficientes binarios que, a partir de un estado inicial IS (Initial State), genera una secuencia de salida $\{a_n\}$.
2. Una función booleana no lineal $F : GF(2)^L \rightarrow GF(2)$, llamada función filtro, cuyas variables de entrada son los L bits de los sucesivos estados del LFSR, y cuya imagen es la secuencia binaria $\{Z_n\}$.

Aunque las secuencias producidas por LFSR están bien estudiadas, no puede decirse lo mismo de las secuencias obtenidas con filtrados no lineales.

Este trabajo trata sobre la relación entre los diferentes filtrados no lineales que producen exactamente la misma secuencia. El objetivo principal es mostrar que, aunque el estudio de las propiedades de un generador puede llevar a la conclusión de que las propiedades de la secuencia generada son buenas, a veces esa deducción puede ser errónea. En particular, este trabajo muestra que dos estructuras con niveles de seguridad aparentemente diferentes pueden producir la misma secuencia cifrante. De hecho, este resultado puede verse como una demostración de que el nivel de seguridad real de un generador es siempre el nivel de seguridad del elemento más débil de la clase de generadores equivalentes.

ISBN: 978-84-9717-323-0

La organización de este trabajo es de la siguiente forma. La sección II incluye una breve revisión de algunos trabajos relacionados. En la sección III, tras los preliminares necesarios, se aborda el problema del recuento general de filtrados no lineales equivalentes, así como un estudio de la relación entre ellos. Después, la sección IV presenta una breve explicación de la propuesta, que está basada en el nuevo concepto de filtrados recíprocos, e introduce un nuevo método para el cálculo de equivalentes débiles de filtrados no lineales mediante un ejemplo didáctico. Finalmente, la sección V describe algunas conclusiones y posibles líneas futuras de investigación.

II. Trabajos Relacionados

Una herramienta bastante útil para estudiar secuencias binarias es el algoritmo de Berlekamp-Massey [12], que determina el LFSR más corto que permite generar cualquier secuencia binaria finita de entrada. La longitud de dicho LFSR se conoce como complejidad lineal de la secuencia. Algunas cotas generales, tanto inferiores como superiores, de la complejidad lineal de las secuencias filtradas han sido publicadas en [10] y [5], mientras que cotas más ajustadas pero para casos específicos pueden encontrarse en [16] y [17].

El proyecto eSTREAM [3] representa el esfuerzo más importante en cuanto al diseño de cifrados en flujo. Fue un proyecto de varios años que tenía como objetivo promover el diseño de sistemas de cifrado en flujo eficientes adecuados para su adopción generalizada. Como resultado fueron escogidos siete generadores de secuencia cifrante con dos perfiles diferentes, software y hardware. Uno de ellos, el llamado SOSEMANUK, es un generador basado en un LFSR donde la longitud del LFSR utilizado es 10 y el contenido de cada etapa es un elemento de $GF(2^{32})$. Dicho generador responde a principios de diseño similares a los del generador SNOW 2.0, predecesor del generador SNOW 3G, que constituye el núcleo de los algoritmos de protección de la confidencialidad y de la integridad en la cuarta generación de comunicaciones de telefonía móvil, LTE y LTE-Advanced, [19].

Existen varias referencias bibliográficas interesantes de ataques criptográficos contra filtrados no lineales.

El primer ataque de correlación básico contra un filtrado no lineal fue publicado en [18]. En él, las correlaciones entre la *m*-secuencia $\{a_n\}$ producida por el LFSR y la secuencia filtrada $\{Z_n\}$ son utilizadas para construir un generador equivalente consistente en una combinación no lineal de varios LFSR. Los principales inconvenientes de ese ataque son la enorme cantidad de tiempo requerido para calcular las correlaciones necesarias, y el requisito de que la función filtro F debe tener una alta correlación con una función afín. Tras definir la no linealidad de una función booleana como la distancia de Hamming mínima entre esa función y una función afín, una consecuencia práctica del ataque de correlación básico es que todo diseñador debe elegir siempre funciones filtro altamente no lineales para los filtrados no lineales. Luego, el concepto subyacente a los ataques de correlación básicos fue mejorado, proponiéndose el ataque de correlación rápido descrito en [13]. Dos desventajas comunes de las diferentes versiones de esos ataques son el gran número de bits de secuencia cifrante interceptados que son necesarios para llevarlos a cabo, y la suposición de que la función filtro no es altamente no lineal.

Un ataque general de inversión fue propuesto en [8] contra cualquier función filtro. Como consecuencia, se obtuvo una caracterización sencilla de los filtrados que son resistentes contra dicho ataque de inversión. Por otra parte, los trabajos [6] y [4] propusieron el llamado ataque por decimación contra cualquier generador de secuencia cifrante basado en un LFSR. La idea es considerar una secuencia decimada de la secuencia cifrante interceptada de manera que la secuencia decimada pueda ser generada a partir de una secuencia generada por el LFSR decimada. Sin embargo, según [16] si la longitud L del LFSR es un número primo, entonces el ataque por decimación no proporciona ninguna ventaja.

En los últimos años se han publicado varios ataques algebraicos contra cifrados en flujo. En ellos, el atacante utiliza los bits de la secuencia interceptada para establecer un sistema no lineal de ecuaciones polinómicas en función de los bits generados por el LFSR. El principal problema con respecto a este tipo de ataques es que, como se muestra en [7], el problema para obtener la solución de un sistema no lineal de ecuaciones multivariantes es NP-duro, incluso si todas las ecuaciones son de segundo grado y el cuerpo subyacente es $GF(2)$. Para hacer frente a este problema, se propuso el conocido como método algebraico XL [2] que permite resolver un sistema no lineal de ecuaciones cuadráticas para algunos filtrados no lineales. Con el fin de incrementar la resistencia frente a este ataque, la función filtro debe ser no sólo altamente no lineal, sino también tener una gran distancia con respecto a aproximaciones de pequeño grado algebraico.

Los ataques basados en el equilibrio tiempo-memoria-datos (time-memory-data tradeoff) [1] pueden evitarse fácilmente en los filtrados no lineales mediante el uso de LFSR de gran longitud. Hay otro ataque interesante, llamado de suposición y determinación (guess and determine) [14], que explota la relación entre los valores internos (tales como la recurrencia lineal en el LFSR), y la relación utilizada para construir la secuencia cifrante a partir de esos valores internos. Como su nombre indica, este ataque trata de adivinar algunos valores internos para luego usar las relaciones mencionadas y determinar otros valores internos. Después de un ataque de ese tipo, el cifrado se considera roto cuando un estado interno completo ha podido ser determinado a partir de valores adivinados. Este tipo de ataque puede evitarse mediante la elección adecuada del polinomio del LFSR.

Uno de los trabajos más estrechamente relacionados con el presente es [11], donde se propuso un ataque por transformación lineal contra un filtrado no lineal. La idea detrás de ese ataque es transformar el generador dado en un filtrado no lineal equivalente con el mismo LFSR pero con una función filtro más adecuada para algunos de los ataques mencionados.

Otro trabajo reciente es [15], donde se define una clase de equivalencia de los filtrados no lineales, demostrando que un número importante de propiedades criptográficas no son invariantes entre los elementos de la misma clase de equiva-

lencia. Los propios autores reconocen que la determinación del cifrado equivalente más débil de dicha clase es una tarea muy difícil debido a que el tamaño de la clase de equivalencia es muy grande. El presente trabajo no estudia dicha clase de equivalencia completa, sino sólo uno de sus elementos, que hemos identificado que en muchos casos conduce a un generador equivalente más débil que el filtrado de partida.

En conclusión, cada ataque contra el filtrado no lineal suele conducir a nuevas conclusiones sobre propiedades deseables del LFSR y/o de la función filtro. En consecuencia, uno de los principales temas de investigación con respecto a los filtrados no lineales es sobre cómo construir una buena función booleana para aumentar la resistencia contra los ataques mencionados. Este trabajo se ocupa de esta cuestión ya que demuestra que las propiedades del generador no siempre garantizan la seguridad de las secuencias producidas.

III. Estudio General de Filtrados Equivalentes

En esta sección se obtiene el número de filtrados equivalentes y luego se analiza la relación entre ellos.

Dado un filtrado no lineal consistente en un LFSR con polinomio característico $P_1(x)$ y función filtro $F_1(x)$, siempre es posible generar la misma secuencia con cualquier otro LFSR de la misma longitud y otra función filtro.

Si α es una raíz del polinomio característico $P_1(x)$ a la vez que un elemento primitivo de $GF(2^L)$ y $(k, 2^L-1) = 1$, entonces se tiene que α^k es también un elemento primitivo de $GF(2^L)$. Por tanto, se concluye que hay $\phi(2^L-1)$ elementos primitivos de $GF(2^L)$. En particular, los L conjugados de cualquier elemento (que son las sucesivas potencias cuadradas), por ejemplo, $\alpha, \alpha^2, \alpha^4, \alpha^8, ..., \alpha^{2^{L-1}}$, son elementos primitivos de $GF(2^L)$ además de raíces del mismo polinomio, que puede calcularse mediante la expresión $\prod_{i=0}^{L-1}(x-\alpha^{2^i})$ en $GF(2^L)$.

Por tanto, hay $\phi(2^L-1)/L$ polinomios primitivos de $GF(2^L)$, cada uno con L raíces que son todos los conjugados de un elemento primitivo. Puesto que cada uno de estos polinomios define un LFSR de longitud L, hay $\Phi(2^L-1)/L$ LFSR diferentes de longitud L, cada uno correspondiente a un conjunto de conjugados de un elemento primitivo de $GF(2^L)$.

En conclusión, dado que cualquier secuencia obtenida con un filtrado no lineal puede ser generada mediante una función filtro sobre cada LFSR, entonces hay $\Phi(2^L-1)/L$ filtrados no lineales diferentes que pueden ser utilizados para generarla.

La relación entre dos elementos primitivos, α y β, raíces de dos polinomios característicos de dos LFSR diferentes de longitud L viene dada por la expresión $\beta = \alpha^k$ siendo $mcd(k, 2^L-1) = 1$ y $k \neq 2^i \cdot$ j (mod 2^L-1) con $i, j > 0$.

Esta información sobre la relación entre los polinomios característicos $P_1(x)$ y $P_2(x)$ de dos LFSR podría ayudar a definir la relación entre las dos funciones filtro $F_1(x)$ y $F_2(x)$ que forman parte de los dos generadores equivalentes que producen la misma secuencia filtrada.

Como se puede ver en la tabla I, los casos $k = 1$ y $k = 2^{L-1}-1$ siempre determinan diferentes conjuntos de raíces conjugadas que definen diferentes LFSR. De hecho, los polinomios correspondientes a las raíces α y $\beta = \alpha^{2^{L-1}-1}$ son siempre recíprocos.

Tabla I
Ejemplos de recuento de filtrados equivalentes

L	3	4	5	6
2^L-1	7	15	31	63
N. filtrados	2	2	6	6
k por filtro	1,2,4	1,2,4,8	1,2,4,8,16	1,2,4,8,16,32
	3,5,6	7,11,13,14	3,6,12,24,17	5,10,20,40,17,34
			5,10,20,9,18	11,22,44,25,50,37
			7,14,28,25,19	13,26,52,41,19,38
			11,22,13,26,21	23,46,29,58,53,43
			15,30,29,27,23	31,62,61,59,55,47

Cualquier *m*-secuencia $\{a_n\}$ puede escribirse en función de las raíces del polinomio característico del LFSR mediante la función traza, de modo que $a_n = Tr(\alpha^n) = \sum_{i=0}^{L-1}\alpha^{n2^i}$. En consecuencia, dada una secuencia $\{a_n\}$ generada por un LFSR con polinomio $P_1(x)$ y raíz α, y otra secuencia $\{b_n\}$ generada por otro LFSR con polinomio $P_2(x)$ y raíz β tal que $\beta = \alpha^k$, se tiene que $a_n = \sum_{i=0}^{L-1}\alpha^{n2^i}$ y $b_n = \sum_{i=0}^{L-1}\alpha^{kn2^i}$. Esto se muestra con un ejemplo en la tabla II.

Si dos filtrados definidos por los correspondientes polinomios y funciones filtro $(P_1(x), F_1(x))$ y $(P_2(x), F_2(x))$ generan la misma secuencia, entonces se tiene que:

$$F_1(a_n, a_{n+1}, ..., a_{n+L-1}) =$$
$$= F_1(\sum_{i=0}^{L-1}\alpha^{n2^i}, \sum_{i=0}^{L-1}\alpha^{(n+1)2^i}, ..., \sum_{i=0}^{L-1}\alpha^{(n+L-1)2^i}) =$$
$$F_2(b_n, b_{n+1}, ..., b_{n+L-1}) =$$
$$= F_2(\sum_{i=0}^{L-1}\alpha^{kn2^i}, \sum_{i=0}^{L-1}\alpha^{k(n+1)2^i}, ..., \sum_{i=0}^{L-1}\alpha^{k(n+L-1)2^i}).$$

La forma algebraica normal de una función booleana permite escribir la secuencia generada por un filtrado $(P_1(x), F_1(x))$ en función de una raíz α del polinomio $P_1(x)$ y coeficientes binarios, de la siguiente manera:

$$F_1(a_n, a_{n+1}, ..., a_{n+L-1}) =$$
$$= c_0 a_n + \cdots + c_{L-1}a_{n+L-1} + c_{0,1}a_n a_{n+1} +$$
$$+ \cdots + c_{L-2,L-1}a_{n+L-2}a_{n+L-1} + \cdots +$$
$$+ c_{0,1,...L-1}a_n a_{n+1}\cdots a_{n+L-1} =$$
$$= c_0\sum_{i=0}^{L-1}\alpha^{n2^i} + \cdots + c_{L-1}\sum_{i=0}^{L-1}\alpha^{(n+L-1)2^i} +$$
$$+ c_{0,1}\sum_{i=0}^{L-1}\alpha^{n2^i}\sum_{i=0}^{L-1}\alpha^{(n+1)2^i} + \cdots +$$
$$+ c_{L-2,L-1}\sum_{i=0}^{L-1}\alpha^{(n+L-2)2^i}\sum_{i=0}^{L-1}\alpha^{(n+L-1)2^i} +$$

Tabla II
EJEMPLOS DE RELACIONES ENTRE RAÍCES, POLINOMIOS Y *m*-SECUENCIAS

Raíces	Polinomio	*m*-secuencia
$\alpha, \alpha^2, \alpha^4, \alpha^8, \alpha^{16}$	$x^5+x^4+x^3+x^2+1$	$\{a_n\}$
$\alpha^{15}, \alpha^{30}, \alpha^{29}, \alpha^{27}, \alpha^{23}$	recíproco= $x^5+x^3+x^2+x+1$	$\{b_n\}$ inverso de $\{a_n\}$
$\alpha^3, \alpha^6, \alpha^{12}, \alpha^{24}, \alpha^{17}$	$\prod_{i=0}^{4}(x-\alpha^{3\cdot 2^i}) = x^5+x^4+x^2+x+1$	$\{c_n\}$
$(\alpha^3)^{15}=\alpha^{14}, \alpha^{28}, \alpha^{25}, \alpha^{19}, \alpha^7$	recíproco= $x^5+x^4+x^3+x+1$	$\{d_n\}$inverso de $\{c_n\}$
$\alpha^5, \alpha^{10}, \alpha^{20}, \alpha^9, \alpha^8$	$\prod_{i=0}^{4}(x-\alpha^{5\cdot 2^i}) = x^5+x^3+1$	$\{e_n\}$
$(\alpha^5)^{15}=\alpha^{13}, \alpha^{26}, \alpha^{21}, \alpha^{11}, \alpha^{22}$	recíproco= x^5+x^2+1	$\{f_n\}$ inverso de $\{e_n\}$

$$+\cdots+c_{0,1,\ldots L-1}\sum_{i=0}^{L-1}\alpha^{n2^i}\sum_{i=0}^{L-1}\alpha^{(n+1)2^i}\cdots\sum_{i=0}^{L-1}\alpha^{(n+L-1)2^i}.$$

Por tanto, si la expresión se divide en cosets (conjuntos de enteros $E\cdot 2^i \ mod(2^L-1)$ con $0\le i\le L-1$), entonces la función se puede expresar como:

$$F_1(a_n, a_{n+1}, ..., a_{n+L-1}) =$$
$$=\sum_{i=0}^{L-1} C_{coset1}\alpha^{n\cdot coset1\cdot 2^i} + C_{coset2}\alpha^{n\cdot coset2\cdot 2^i}+\cdots$$

con $C_{cosetj}\in GF(2^L)$.

Los pesos de los cosets cuyos coeficientes son distintos de cero en la expresión anterior proporcionan información sobre el orden de la función. En particular, si se conoce la relación $\beta=\alpha^k$ entre dos filtrados $(P_1(x), F_1(x))$ y $(P_2(x), F_2(x))$ que generan la misma secuencia, entonces:

$$F_2(b_n, b_{n+1}, ..., b_{n+L-1}) =$$
$$=\sum_{i=0}^{L-1} D_{coset1}\alpha^{n\cdot coset1\cdot 2^i} + D_{coset2}\alpha^{n\cdot coset2\cdot 2^i}+\cdots$$

con $D_{cosetj}\in GF(2^L)$.

Entonces, los cosets que aparecen en ambas expresiones están vinculados de manera que para cada coset $cosetv$ en la primera expresión, existe otro coset $cosetw$ en la segunda:

$$\sum_{i=0}^{L-1} C_{cosetv}\alpha^{n\cdot cosetv\cdot 2^i} = \sum_{i=0}^{L-1} D_{cosetw}\alpha^{n\cdot cosetw\cdot 2^i}.$$

IV. FILTRADOS RECÍPROCOS

A partir de los resultados mostrados en la sección anterior, si se consideran dos LFSR con polinomios recíprocos $P_1(x)$ y $P_2(x)$, se pueden obtener dos conclusiones.

1. Si se aplica la misma función filtro $F(x)$ a ambos LFSR, se generan secuencias diferentes. Se puede utilizar el algoritmo de Berlekamp-Massey sobre las secuencias filtradas resultantes y a partir de las factorizaciones de los polinomios obtenidos, se concluye que siempre se corresponden exactamente con los mismos cosets.
2. Para generar la misma secuencia con esos LFSR, se deben utilizar dos funciones filtro diferentes $F_1(x)$ y $F_2(x)$. Dado que la factorización del polinomio obtenido con el algoritmo de Berlekamp-Massey corresponde a los cosets complementarios especulares en los grupos definidos por cada uno de los LFSR, sobre el orden de las funciones filtro influyen los pesos de esos cosets. En particular, se puede concluir que $orden(F_i) = max(L$-(peso de cada clase lateral vinculado a la factorización del polinomio de la secuencia)).

Por tanto, si existe un filtrado que produce una secuencia en la que la factorización del polinomio sólo corresponde a cosets de peso $> L/2$, entonces existe un filtrado equivalente que es menos fuerte ya que tiene orden $< L/2$. Con respecto a este filtrado equivalente, es bien sabido que el LFSR es recíproco del original. Sin embargo, descubrir cómo es la función filtro es más complejo.

Por otra parte, si una función filtro tiene orden $\sim L/2$, dado que el orden viene dado por el máximo de los pesos de los cosets asociados a la factorización, entonces existe un filtrado equivalente de orden $\ge L/2$, pues ese grado viene dado por el máximo de los pesos de los cosets. En consecuencia, si se usa un LFSR recíproco, se sabe que su peso es al menos $L-L/2$. Esto puede verse como una demostración de la conocida recomendación de uso de funciones filtro de orden $\sim L/2$.

De todo lo anterior se puede concluir que para cualquier filtrado, siempre puede obtenerse un filtrado equivalente para generar la misma secuencia, llamado filtrado recíproco. Con el fin de determinar el filtrado recíproco para cualquier filtrado conocido, el procedimiento propuesto incluye los siguientes cuatro pasos básicos:

1. Determinar las relaciones entre las raíces de los polinomios característicos del LFSR inicial y su recíproco.
2. Expresar ambas *m*-secuencias mediante la función traza.
3. Calcular los coeficientes de los cosets en la expresión de la función filtro.
4. Elegir los coeficientes adecuados para construir la función filtro recíproca.

Este procedimiento se ilustra con un ejemplo didáctico.

Ejemplo:

Dado un LFSR de longitud $L=5$, polinomio característico $P_1(x)=x^5+x^3+1$, y estado inicial $IS_1=(1,0,0,0,0)$, se aplica la función filtro de orden 4:

$$F_1(a_0, a_1, a_2, a_3, a_4) =$$
$$=a_0a_1a_3a_4+a_0a_2a_3a_4+a_0a_1a_4+a_0a_1a_3+a_1a_3a_4+$$
$$+a_0a_3a_4+a_1a_2+a_1a_3+a_2a_4+a_0a_2+a_0a_3+a_1+a_2+a_3$$

para producir la secuencia filtrada de periodo $2^5 - 1$:

$$0010110110101101110000100101011.$$

El LFSR recíproco tiene polinomio característico $P_2(x) = x^5 + x^2 + 1$, cuya raíz β se relaciona con la raíz α de $P_1(x)$ según la expresión $\beta = \alpha^{2^{5-1}-1} = \alpha^{15}$. Además, gracias al inverso modular de 15 (mod 31), puede obtenerse la relación inversa $\alpha = \beta^{29}$.

Al mismo tiempo, las *m*-secuencias $\{a_n\}$ y $\{b_n\}$ obtenidas de $P_1(x)$ y $P_2(x)$, respectivamente, pueden expresarse mediante sus expresiones traza:

$$a_n = \alpha^n + \alpha^{2n} + \alpha^{4n} + \alpha^{8n} + \alpha^{16n}$$
$$b_n = \beta^n + \beta^{2n} + \beta^{4n} + \beta^{8n} + \beta^{16n}.$$

En consecuencia, las funciones filtro F_1 y F_2 pueden expresarse en función de los $\phi(2^5 - 1)/5 = 6$ cosets $\{15, 11, 7, 5, 3, 1\}$. En particular, el coeficiente C_{15} correspondiente al coset de orden máximo 4 puede conseguirse mediante el test de presencia de raíces [16], mientras que los coeficientes C_7 y C_{11} correspondientes a los cosets de orden 3 pueden calcularse mediante agrupación de términos:

$$C_{15} = \alpha^6, C_7 = \alpha^{24}, C_{11} = \alpha^4.$$

A partir de estos valores se puede concluir que no existen más cosets de menor peso en la expresión de la función filtro F_1. Por tanto,

$$F_1 = C_{15}\alpha^{15n} + C_{15}^2\alpha^{30n} + C_{15}^4\alpha^{29n} + C_{15}^8\alpha^{27n} + C_{15}^{16}\alpha^{23n} +$$
$$+ C_7\alpha^{7n} + C_7^2\alpha^{14n} + C_7^4\alpha^{28n} + C_7^8\alpha^{25n} + C_7^{16}\alpha^{19n} +$$
$$+ C_{11}\alpha^{11n} + C_{11}^2\alpha^{22n} + C_{11}^4\alpha^{13n} + C_{11}^8\alpha^{26n} + C_{11}^{16}\alpha^{21n}.$$

Si se sustituye $\alpha = \beta^{29}$ en esa expresión, la función filtro F_2 que genera la misma secuencia se puede expresar como:

$$F_2 = C_{15}\beta^{29\cdot 15n} + C_{15}^2\beta^{29\cdot 30n} + C_{15}^4\beta^{29\cdot 29n} +$$
$$+ C_{15}^8\beta^{29\cdot 27n} + C_{15}^{16}\beta^{29\cdot 23n} + C_7\beta^{29\cdot 7n} + C_7^2\beta^{29\cdot 14n} +$$
$$+ C_7^4\beta^{29\cdot 28n} + C_7^8\beta^{29\cdot 25n} + C_7^{16}\beta^{29\cdot 19n} + C_{11}\beta^{29\cdot 11n} +$$
$$+ C_{11}^2\beta^{29\cdot 22n} + C_{11}^4\beta^{29\cdot 13n} + C_{11}^8\beta^{29\cdot 26n} + C_{11}^{16}\beta^{29\cdot 21n} =$$
$$= C_{15}\beta^n + C_{15}^2\beta^{2n} + C_{15}^4\beta^{4n} + C_{15}^8\beta^{8n} + C_{15}^{16}\beta^{16n} +$$
$$+ C_7\beta^{17n} + C_7^2\beta^{3n} + C_7^4\beta^{6n} + C_7^8\beta^{12n} + C_7^{16}\beta^{24n} +$$
$$+ C_{11}\beta^{9n} + C_{11}^2\beta^{18n} + C_{11}^4\beta^{5n} + C_{11}^8\beta^{10n} + C_{11}^{16}\beta^{20n}.$$

En consecuencia, se puede concluir que en esta expresión sólo aparecen los cosets 1, 3 y 5. Por otra parte, sus coeficientes D_1, D_3 y D_5 vienen dados por:

$$D_1 = C_{15} = \alpha^6 = \beta^{29\cdot 6} = \beta^{19}$$
$$D_3 = C_7^2 = \alpha^{24\cdot 2} = \alpha^{17} = \beta^{29\cdot 17} = \beta^{28}$$
$$D_5 = C_{11}^4 = \alpha^{4\cdot 4} = \alpha^{16} = \beta^{29\cdot 16} = \beta^{30}.$$

Dado que el peso máximo de los cosets en esa expresión es 2, se analizan los términos no lineales de orden 2 en la expresión de F_2. Como antes, para cada término no lineal de orden 2, se pueden obtener los coeficientes D_3 y D_5 mediante el test de presencia de raíces, mientras que el coeficiente D_1 correspondiente al coset de peso 1 puede calcularse agrupando términos, como se muestra en la tabla III.

Tabla III
COEFICIENTES DE LOS COSETS 3, 5 Y 1 PARA TODOS LOS POSIBLES TÉRMINOS DE ORDEN 2

	D_3	D_5	D_1
b_0b_1	β^{19}	β^{30}	β^{16}
b_0b_2	β^{7}	β^{29}	β
b_0b_3	β	β^{19}	β^{17}
b_0b_4	β^{14}	β^{27}	β^{2}
b_1b_2	β^{22}	β^{4}	β^{17}
b_1b_3	β^{10}	β^{3}	β^{2}
b_1b_4	β^{4}	β^{24}	β^{18}
b_2b_3	β^{25}	β^{9}	β^{18}
b_2b_4	β^{13}	β^{8}	β^{3}
b_3b_4	β^{28}	β^{14}	β^{19}

Una versión interesante del problema de la mochila discreta definido mediante los coeficientes de la tabla III es entonces resuelto de manera que para cada una de las dos primeras columnas correspondientes a los cosets de peso máximo, se calculan los elementos cuya suma coincide con el correspondiente coeficiente conocido. En particular, la solución muestra que los coeficientes correspondientes a los productos b_0b_2, b_1b_2, b_1b_3, b_1b_4 y b_3b_4 dan dos valores:

$$D_3 = \beta^7 + \beta^{22} + \beta^{10} + \beta^4 + \beta^{28} = \beta^{28}$$
$$D_5 = \beta^{29} + \beta^4 + \beta^3 + \beta^{24} + \beta^{14} = \beta^{30}.$$

Este resultado aplicado sobre la última columna implica que, para obtener la suma final $D_1 = \beta^{19}$, tienen que incluirse los elementos lineales $b_1 + b_2 + b_4$ en la función filtro F_2:

$$D_1 = \beta + \beta^{17} + \beta^2 + \beta^{18} + \beta^{19} + \beta + \beta^2 + \beta^4 = \beta^{19}.$$

Por tanto, se obtiene la expresión final de la función filtro equivalente:

$$F_2(b_0, b_1, b_2, b_3, b_4) =$$
$$= b_0b_2 + b_1b_2 + b_1b_3 + b_1b_4 + b_3b_4 + b_1 + b_2 + b_4.$$

Esta función aplicada sobre el LFSR recíproco con polinomio característico $P_2(x) = x^5 + x^2 + 1$ y estado inicial $IS_2 = (1, 0, 0, 1, 0)$, produce la misma secuencia que el filtrado de entrada

$$0010110110101101110000100101011.$$

F_2 es una función de orden 2 con el mismo número de términos de orden 2 y de orden 1 que F_1, pero sin términos de orden 3 ni de orden 4. Por tanto, desde un punto de vista criptográfico, el atacante podría lanzar un ataque más eficaz contra F_2 que contra F_1, aunque ambos filtrados generan exactamente la misma secuencia.

Por tanto, este ejemplo muestra que el método propuesto se puede aplicar sobre cualquier filtrado conocido para producir

un filtrado equivalente, que en el caso del LFSR recíproco es de un orden inferior. Esta es una demostración de que para algunos generadores aparentemente seguros pueden existir equivalentes más débiles, y lo que es más importante, que estos equivalentes pueden ser calculados.

V. Conclusiones y Trabajos Futuros

Este trabajo ha abordado el problema del cálculo de filtrados no lineales equivalentes que producen la misma secuencia que un filtrado conocido. En particular, se presenta el análisis de un caso en el que un LFSR recíproco se utiliza para definir un filtrado equivalente. De hecho, en esas condiciones existen relaciones específicas entre ambas funciones filtro que permiten la definición de un método específico para calcular la función filtro equivalente de una de partida. El estudio concluye que el generador equivalente puede tener un nivel de seguridad inferior al del filtrado original. Por tanto, el método propuesto permite la construcción de equivalentes más débiles que los filtrados de partida. En conclusión, este trabajo muestra que dos estructuras con niveles de seguridad aparentemente diferentes en función de sus propiedades, pueden de hecho producir exactamente la misma secuencia cifrante, por lo que en realidad ambos generadores debe ser considerados tan inseguros como el más débil de los dos.

Dada la dificultad del tema, todavía quedan muchas cuestiones abiertas. En particular, una de ellos es el desarrollo de métodos óptimos para la elección de los coeficientes correspondientes que aparecen en la última fase del método propuesto. Además, un estudio similar al que se muestra en este trabajo, pero sobre otros equivalentes que no se basen en el LFSR recíproco, podría ser útil para llevar a cabo potenciales ataques contra los filtrados no lineales.

Agradecimientos

Investigación financiada por el MINECO y la fundación Europea FEDER mediante los proyectos TIN2011-25452 e IPT-2012-0585-370000.

Referencias

[1] Biryukov, A., Shamir, A. Cryptanalytic time/memory/data tradeoffs for stream ciphers. Advances in Cryptology, ASIACRYPT00, Lecture Notes in Computer Science 1976, pp. 1-13. Springer-Verlag, 2000.

[2] Courtois, N., Klimov, A., Patarin, J., Shamir, A. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. Advances in Cryptology, EUROCRYPT00, Lecture Notes in Computer Science 1807, pp. 392-407. Springer-Verlag, 2000.

[3] eSTREAM: the ECRYPT Stream Cipher Project. Available from http://www.ecrypt.eu.org/stream/

[4] Filiol, E. Decimation attack on stream ciphers. Advances in Cryptology, INDOCRYPT 2000, Lecture Notes in Computer Science 1977, pp. 31-42. Springer-Verlag, 2000.

[5] Fuster-Sabater A., Caballero-Gil, P. On the linear complexity of nonlinearly filtered pn-sequences. Advances in Cryptology, ASIACRYPT94, Lecture Notes in Computer Science 917, pp. 80-90. Springer-Verlag, 1995.

[6] Games, R.A., Rushanan, J.J. Blind synchronization of m- sequences with even span. Advances in Cryptology, EUROCRYPT93, Lecture Notes in Computer Science 765, pp. 168-180. Springer-Verlag, 1994.

[7] Garey, M.R., Johnson, D.S. Computers and Interactability. Freeman and Company, 1979.

[8] Golic, J.D., Clark, A., Dawson, E. Generalized inversion attack on nonlinear filter generators. IEEE Transactions on Computers, 49(10), pp. 1100-1109, 2000.

[9] Golomb, S.W., Shift Register-Sequences, Aegean Park Press, Laguna Hill, 1982.

[10] Key, E.L. An analysis of the structure and complexity of nonlinear binary sequence generators. IEEE Transactions on Information Theory, 22(6), pp. 732-736, 1976.

[11] Lohlein, B. Design and analysis of cryptographic secure keystream generators for stream cipher encryption. PhD thesis, Faculty of Electrical and Information Engineering, University of Hagen, Germany, 2001.

[12] Massey, J. L. Shift-register synthesis and BCH decoding, IEEE Transactions on Information Theory, IT-15 (1), pp. 122-127, 1969.

[13] Meier W., Staffelbach, O.J. Fast correlation attacks on stream ciphers. Journal of Cryptology, 1(3), pp. 159-176, 1989.

[14] Pasalic, E. On guess and determine cryptanalysis of LFSR-based stream ciphers. IEEE Transactions on Information Theory, 55(7), pp. 3398–3406, 2009.

[15] Ronjom, S., Cid, C. Nonlinear equivalence of stream ciphers. Fast Software Encryption, pp. 40–54. Springer-Verlag, 2010.

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

[17] Schneider, M. Methods of generating binary pseudo-random sequences for stream cipher encryption. PhD thesis, Faculty of Electrical Engineering, University of Hagen, Germany, 1999.

[18] Siegenthaler, T. Decrypting a class of stream ciphers using ciphertext only. IEEE Transactions on Computers, 100(1), pp. 81–85, 1985.

[19] SNOW 3G specification. Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 and UIA2. Available from http://www.3gpp.org/DynaReport/35216.htm