Cifrador de bloque con clave de 128 bits

Andrés Francisco Farías[1] – Andrés Alejandro Farías[2]

[1] DACEFyN – UNLaR (afarias665@yahoo.com.ar)
[2] DACEFyN – UNLaR (andres_af86@hotmail.com)

Resumen. Desarrollo de un cifrador de bloque, con clave de 128 bits, sobre la estructura de una red de Feistel, con un algoritmo de cifrado basado en funciones booleanas de cuatro variables y registros de desplazamiento de retroalimentación lineal con función de filtrado no lineal. La particularidad de las funciones es que fueron elegidas por sus propiedades criptográficas como ser balanceadas y tener alta no linealidad. Finalmente la salida del texto cifrado fue sometida a pruebas estadísticas de aleatoriedad

Palabras Clave: Cifrador, clave, función booleana, no linealidad.

Key Words: Cipher, key, boolean function, non-linearity.

1 Introducción

El trabajo presentado se trata de un cifrador de bloque, basado en un una red de Feistel, la que permite el cifrado y descifrado utilizando la misma estructura, donde para el caso del descifrado se utilizan las subclaves cambiando el orden [1], [2].

La clave adoptada es de 16 carácteres, es decir 128 bits y se utilizan para el algoritmo de cifrado, funciones booleanas de cuatro variables balanceadas, que cumplen con el Criterio de Avalancha Estricta (SAC, sigla en inglés) y de alta no linealidad junto con registros de desplazamiento de retroalimentación lineal con función de filtrado no lineal (NLFSR, sigla en inglés).

2 Esquema del cifrador

El cifrado de bloque, se denomina así por realizar el proceso de cifrado por bloques de textos de igual longitud, para este trabajo se utilizó bloques de 256 bits. Luego esos bloques son ensamblados siguiendo distintas disposiciones.

Básicamente la estructura del cifrador está conformada por:

- Red de Feistel para cifrado
- Red de Feistel para descifrado
- Clave y subclaves
- Algoritmo de cifrado
- Modo de operación

2.1 Red de Feistel para Cifrado

El proceso de cifrado consiste en dividir el texto plano en bloques de 256 bits, y a cada uno de ellos, se los somete a los procedimientos de la Red de Feistel, indicados en Fig. 1.

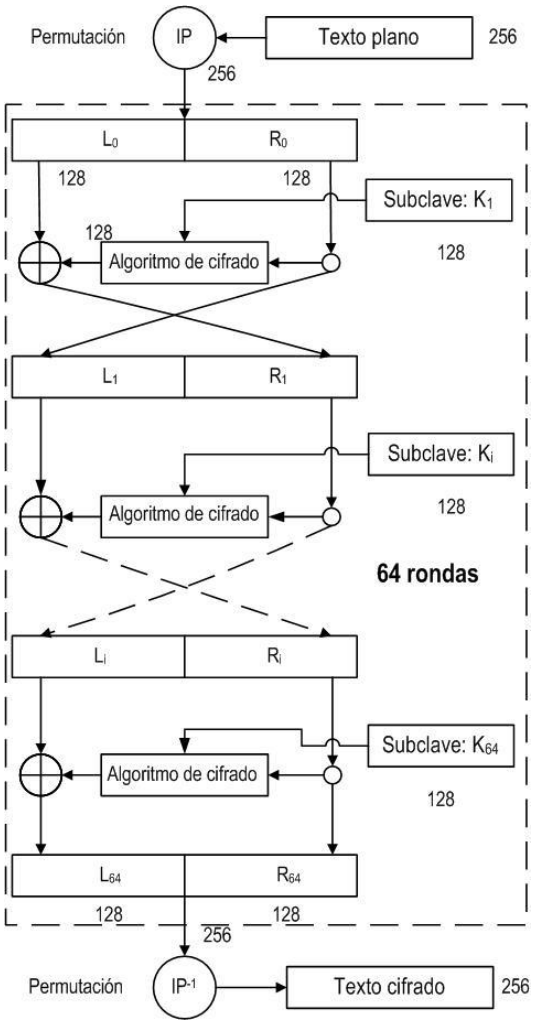

Fig. 1. Red de Feistel de cifrado

Cada bloque antes de ingresar a la red sufre una permutación dada por una matriz PI, luego de ello se divide al bloque en dos bloques, uno izquierdo y otro derecho, de 128 bits cada uno, a partir de ese momento esos bloques entran en las 64 rondas, en las que aparece el algoritmo de cifrado y las 64 subclaves. Finalmente los bloques resultantes del final de las rondas se concatenan para un formar un bloque de 256 bits,

que es sometido a una nueva permutación IPI, que da como resultado el texto cifrado.

2.2 Red de Feistel para descifrado

La Red de Feistel para descrifrado es similar a la anterior, pero en este caso se toma el texto cifrado y se lo divide en bloques de 256 bits, que es sometido a una permutación PI, y dividido en dos bloques de 128 bits, uno izquierdo y otro derecho, Fig. 2.

Se realizan las 64 rondas con las 64 subclaves del cifrado, ingresadas en orden inverso. Las matrices PI e IPI son las mismas que se utilizaron para el cifrado:

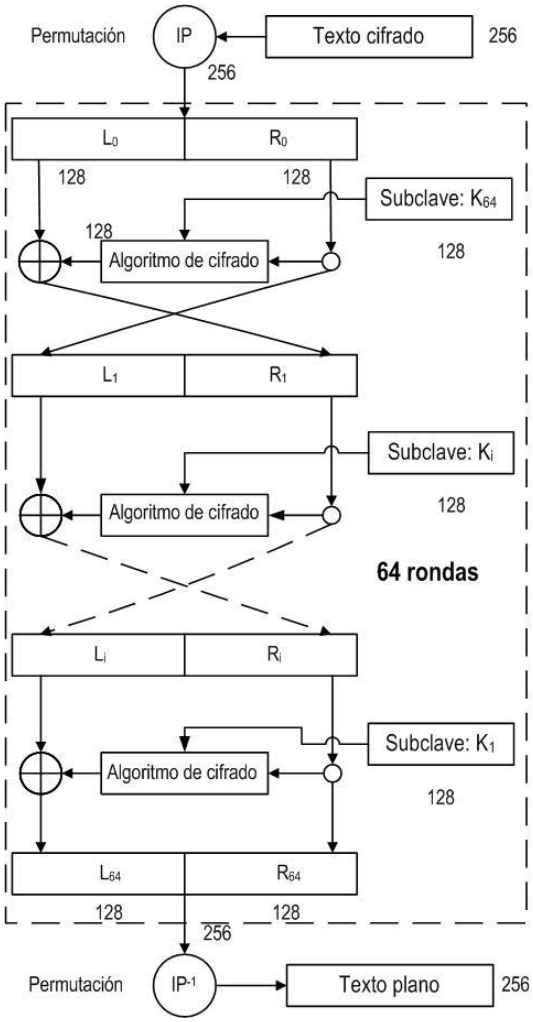

Fig. 2. Red de Feistel de descifrado

2.3 Clave y subclaves

Como se dijo anteriormente la clave está conformada con 16 carácteres lo que equivale a 128 bits, de las que se generan 16 subclaves de 128 bits, siguiendo los pasos que se muestran en Fig. 3 :

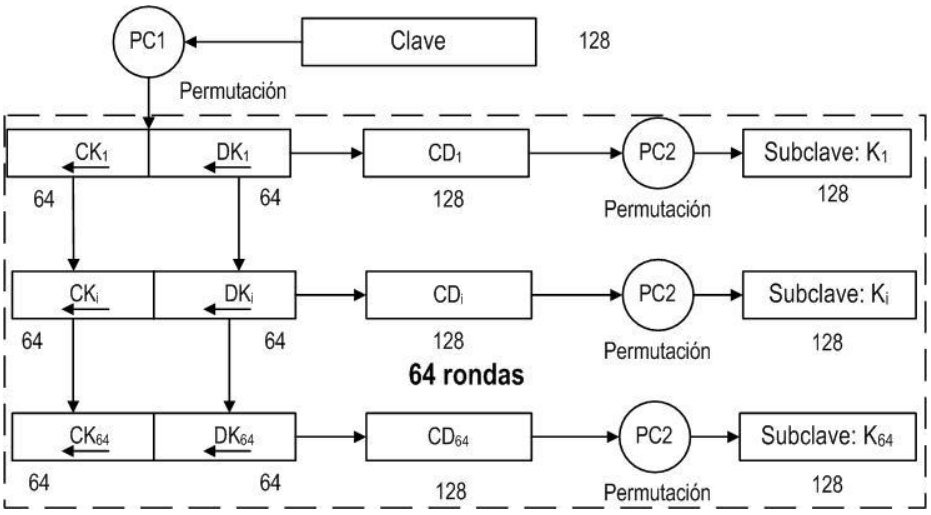

Fig. 3. Tratamiento de las Subclaves

La clave es sometida a una permutación según la matriz de permutación PC1; luego se divide el bloque de 128 bits resultante en dos bloques de 64 bits, los que sufren desplazamiento de las posiciones de los bits de manera de tener 64 pares de bloques de 64 bits que corresponderán a las 64 subclaves.

Esos pares son ensamblados y luego sometidos a la permutación PC2, para obtener las 64 subclaves finales. Los pares cuyos números de orden son: 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 y 60 se desplazan un bit a la izquierda, el resto se desplazan dos bits en el mismo sentido.

2.4 Algoritmo de cifrado

El algoritmo de cifrado tiene la configuración que se indica en la Fig. 4, donde la mitad del bloque de texto, la parte derecha de 128 bits, es sometida a una operación XOR con la subclave de 128 bits.

La secuencia de 128 bits resultante, se divide en bloques de cuatro bits que alimentarán posteriormente a las funciones boolenas de cuatro variables. De esa operación resultarán bloques de 4 bits modificados, los que serán ensamblados para obtener el resultado final de de la función de Feistel, un bloque de 128 bits.

Este bloque conformará los estados iniciales que alimentarán al NLFSR de 128 bits, que correrá una vez cargado, 128 ciclos para obtener los 128 bits del salida del

algoritmo de cifrado, Fig. 4.

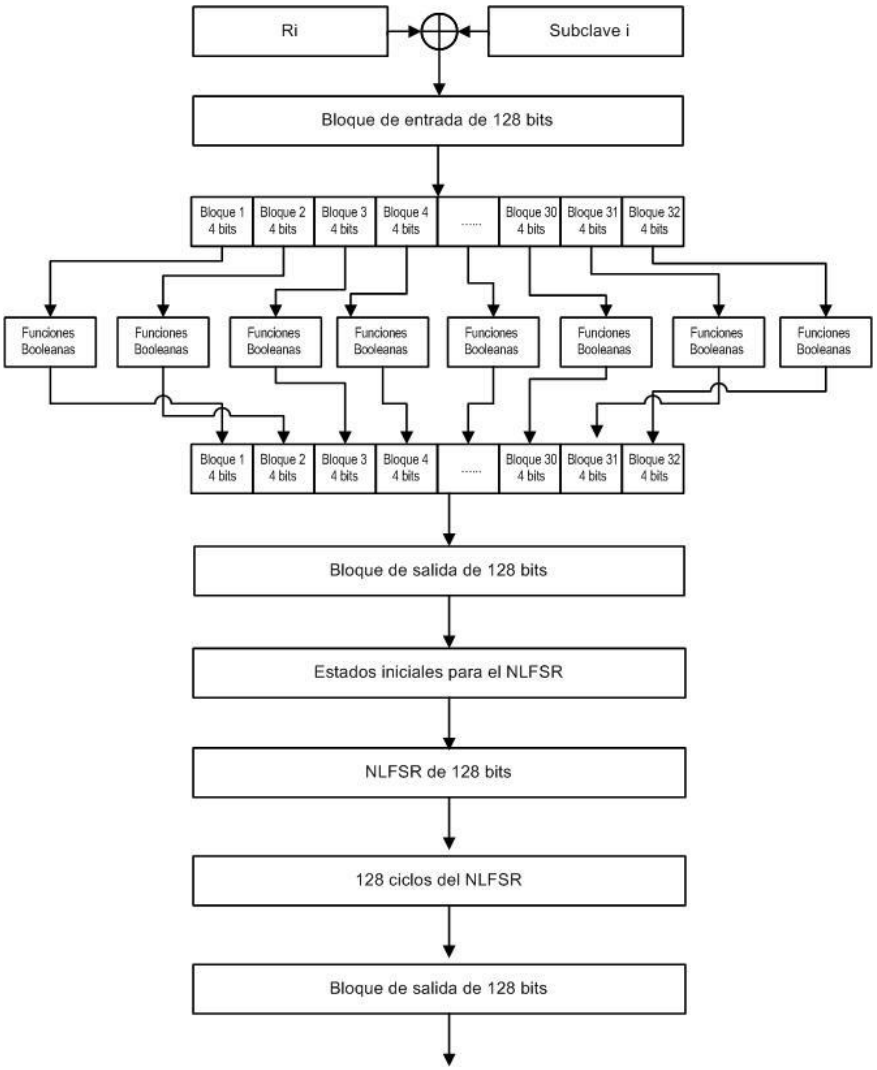

Fig. 4. Algoritmo de cifrado

Bloques de 4 bits: A continuación en Fig. 5, se indican las rutas de alimentación de las cuatro funciones booleanas a partir de las 4 salidas de un bloque, dando como resultado un bloque con bits modificados.

Esta operación se repite para los otros 32 bloques de 4 bits, donde aparecen otras funciones booleanas, dsipuestas de la siguiente forma, para los primeros 64 bloques se utilizan 64 funciones booleanas y para los siguientes 64 bloques se disponen de las mismas 64 funciones booleanas.

Como se dijo estas funciones tienen buenas propiedades criptográficas.

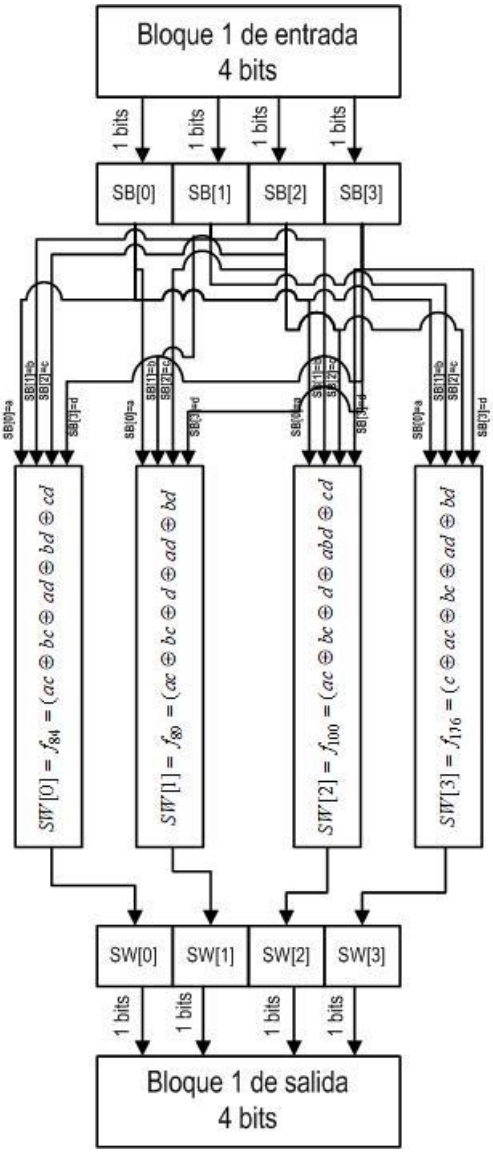

Fig. 5. Bloque de 4 bits

NLFSR: Está compuesto por un registro de desplazamietno de retroalimentación lineal de de 128 bits (LFSR, sigla en inglés), del tipo Fibonacci con un polinomio primitivo de conexión, y una función booleana de filtrado no lineal, de cuatro variables de las mismas características que el resto de las funciones utilizadas, Fig. 6.

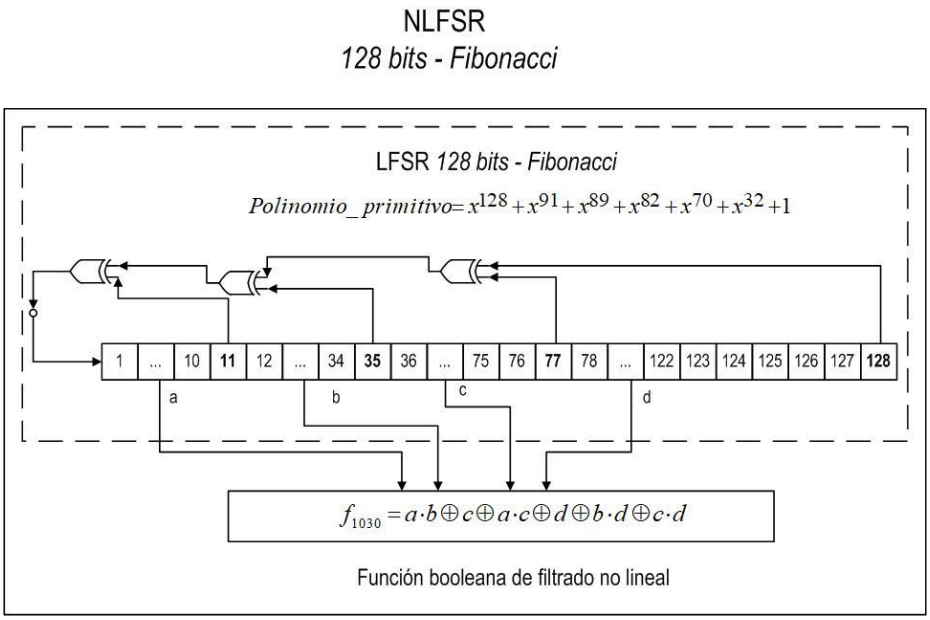

Fig. 6. Bloque de 4 bits

2.5 Modo de operación

Los bloques son ensamblados siguiendo el modo: Electronic Code Book (ECB) Mode [3], tanto para el cifrado como para el descrifrado, como muestra la Fig. 7:

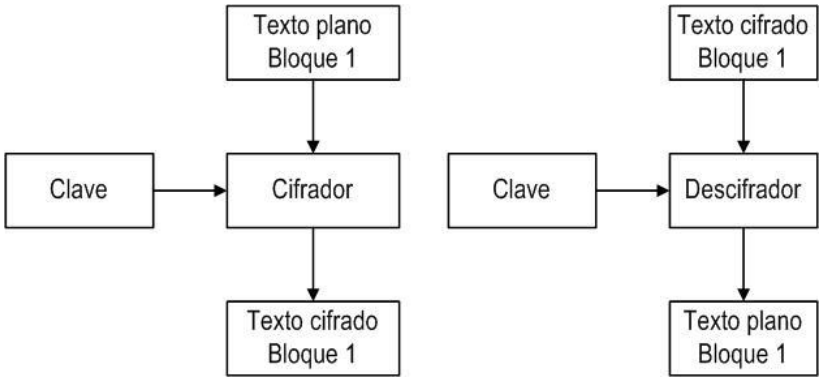

Fig. 7. Electronic Code Book (ECB)

3 Permutación

Se recurre a una matriz con una distribución aleatoria de las posiciones, para obtenerla se utiliza un generador de números aleatorios, en esta ocasión se adopta un generador congruencial multiplicativo [4].

3.1 Generador congruencial multiplicativo

El generador tiene la siguiente expresión:

$$x_{i+1} = (a_x \cdot x_i)mod\ m_x$$

Donde: $a_x = multiplicador$, $m_x = módulo$ y $x_0 = semilla$

Tabla 1. Matriz IP

Matriz	módulo	multiplicador	semilla
IP	1048576	1279	1153
PC1	1048576	1597	1531
PC2	1048576	1933	1759

4 Propiedades criptográficas de las funciones booleanas

A continuación se indican algunas de las propiedades criptográficamente más significativas, adoptadas para este trabajo [5], [6], [7].

- **Función Balanceada:** Una función booleana de n-variables f es balanceada si $\mathrm{w(f)} = \mathrm{2n} - 1$. Esta propiedad es deseable para evitar ataques criptodiferenciales. La función es balanceada cuando el primer coeficiente del espectro de Walsh-Hadamard, es igual a cero: $\mathrm{F}(0) = 0$.
- **No Linealidad:** Valores altos de esta propiedad reducen el efecto de los ataques por criptoanálisis lineal. La No Linealidad de una función booleana puede ser calculada directamente de la transformada de Walsh-Hadamard: $NL_f = \frac{1}{2} \cdot (2^n - |WH_{max}(f)|)$
- **Grado Algebraico:** El grado algebraico de una función, es el número de entradas más grande que aparece en cualquier producto de la Forma Normal Algebraica. Es deseable que sean valores altos.
- **SAC:** El Criterio de Avalancha Estricto requiere los efectos avalancha de todos los bits de entrada. Una función booleana se dice que satisface SAC sí y solo sí, $f(x \oplus u)$, es balanceada para toda u con w(u)=1:.

4.1 Tabla de resultados

Para funciones booleanas de cuatro variables, hay un total de 12.870 balanceadas, de las cuales aplicando los criterios anteriores se eligen 64, indicados en la Tabla 2.

Tabla 2. Funciones booleanas adoptadas

Bloques	Bits	Función
1 y 17	1	$f_{84} = (ac \oplus bc \oplus ad \oplus bd \oplus cd)$
1 y 17	2	$f_{89} = (ac \oplus bc \oplus d \oplus ad \oplus bd)$
1 y 17	3	$f_{100} = (ac \oplus bc \oplus d \oplus abd \oplus cd)$
1 y 17	4	$f_{176} = (c \oplus ac \oplus bc \oplus ad \oplus bd)$
2 y 18	1	$f_{199} = (c \oplus ac \oplus bc \oplus d \oplus cd)$
2 y 18	2	$f_{381} = (c \oplus abc \oplus ad \oplus bd \oplus cd)$
2 y 18	3	$f_{468} = (c \oplus d \oplus ad \oplus bd \oplus cd)$
2 y 18	4	$f_{536} = (ab \oplus bc \oplus ad \oplus bd \oplus cd)$
3 y 19	1	$f_{541} = (541ab \oplus bc \oplus d \oplus ad \oplus cd)$
3 y 19	2	$f_{547} = (ab \oplus bc \oplus d \oplus bd \oplus acd)$
3 y 19	3	$f_{621} = (ab \oplus ac \oplus ad \oplus bd \oplus cd)$
3 y 19	4	$f_{627} = (ab \oplus ac \oplus d \oplus ad \oplus bcd)$
4 y 20	1	$f_{630} = (ab \oplus ac \oplus d \oplus bd \oplus cd)$
4 y 20	2	$f_{762} = (ab \oplus ac \oplus bc \oplus bd \oplus cd)$
4 y 20	3	$f_{773} = (ab \oplus ac \oplus bc \oplus ad \oplus cd)$
4 y 20	4	$f_{787} = (ab \oplus ac \oplus bc \oplus ad \oplus bd)$
5 y 21	1	$f_{971} = (ab \oplus c \oplus ac \oplus bd \oplus cd)$
5 y 21	2	$f_{983} = (ab \oplus c \oplus ac \oplus ad \oplus bcd)$
5 y 21	3	$f_{1017} = (ab \oplus c \oplus ac \oplus d \oplus ad)$
5 y 21	4	$f_{1098} = (ab \oplus c \oplus bc \oplus bd \oplus acd)$
6 y 22	1	$f_{1107} = (ab \oplus c \oplus bc \oplus ad \oplus cd)$
6 y 22	2	$f_{1153} = (ab \oplus c \oplus bc \oplus d \oplus bd)$
6 y 22	3	$f_{1300} = (b \oplus ab \oplus bc \oplus ad \oplus cd)$
6 y 22	4	$f_{1316} = (b \oplus ab \oplus bc \oplus d \oplus bd)$
7 y 23	1	$f_{1550} = (b \oplus ab \oplus ac \oplus bd \oplus cd)$
7 y 23	2	$f_{1579} = (b \oplus ab \oplus ac \oplus ad \oplus bcd)$
7 y 23	3	$f_{1585} = (b \oplus ab \oplus ac \oplus d \oplus ad)$
7 y 23	4	$f_{1623} = (b \oplus ab \oplus c \oplus ac \oplus ad)$
8 y 24	1	$f_{1957} = (b \oplus ab \oplus c \oplus bc \oplus bd)$
8 y 24	2	$f_{2172} = (b \oplus abc \oplus ad \oplus bd \oplus cd)$
8 y 24	3	$f_{2266} = (b \oplus d \oplus ad \oplus bd \oplus cd)$
8 y 24	4	$f_{2352} = (b \oplus ac \oplus bc \oplus abd \oplus cd)$
9 y 25	1	$f_{88} = (ac \oplus bc \oplus ad \oplus bd \oplus abd \oplus cd)$
9 y 25	2	$f_{94} = (ac \oplus bc \oplus dad \oplus abd \oplus bcd)$
9 y 25	3	$f_{99} = (ac \oplus bc \oplus d \oplus bd \oplus abd \oplus acd)$
9 y 25	4	$f_{124} = (ac \oplus bc \oplus abc \oplus ad \oplus bd \oplus cd)$
10 y 26	1	$f_{158} = (ac \oplus bc \oplus abc \oplus d \oplus abd \oplus cd)$
10 y 26	2	$f_{244} = (c \oplus ac \oplus abc \oplus ad \oplus bd \oplus bcd)$
10 y 26	3	$f_{328} = (c \oplus bc \oplus abc \oplus ad \oplus bd \oplus acd)$
10 y 26	4	$f_{387} = (c \oplus abc \oplus ad \oplus bd \oplus abd \oplus cd)$
11 y 27	1	$f_{418} = (c \oplus abc \oplus d \oplus abd \oplus acd \oplus bcd)$
11 y 27	2	$f_{538} = (ab \oplus bc \oplus ad \oplus bd \oplus cd \oplus acd)$
11 y 27	3	$f_{544} = (ab \oplus bc \oplus d \oplus ad \oplus acd \oplus bcd)$
11 y 27	4	$f_{556} = (ab \oplus bc \oplus d \oplus abd \oplus cd \oplus acd)$
12 y 28	1	$f_{579} = (ab \oplus bc \oplus abc \oplus ad \oplus bd \oplus cd)$

12 y 28	2	$f_{604} = (ab \oplus bc \oplus abc \oplus d \oplus bd \oplus acd)$
12 y 28	3	$f_{622} = (ab \oplus ac \oplus ad \oplus bd \oplus cd \oplus bcd)$
12 y 28	4	$f_{633} = (ab \oplus ac \oplus d \oplus bd \oplus acd \oplus bcd)$
13 y 29	1	$f_{640} = (ab \oplus ac \oplus d \oplus abd \oplus cd \oplus bcd)$
13 y 29	2	$f_{664} = (ab \oplus ac \oplus abc \oplus ad \oplus bd \oplus cd)$
13 y 29	3	$f_{679} = (ab \oplus ac \oplus abc \oplus d \oplus ad \oplus bcd)$
13 y 29	4	$f_{764} = (ab \oplus ac \oplus bc \oplus bd \oplus cd \oplus acd)$
14 y 30	1	$f_{769} = (ab \oplus ac \oplus bc \oplus bd \oplus abd \oplus c)$
14 y 30	2	$f_{774} = (ab \oplus ac \oplus bc \oplus ad \oplus cd \oplus bcd)$
14 y 30	3	$f_{780} = (ab \oplus ac \oplus bc \oplus ad \oplus abd \oplus cd)$
14 y 30	4	$f_{788} = (ab \oplus ac \oplus bc \oplus ad \oplus bd \oplus bcd)$
15 y 31	1	$f_{789} = (ab \oplus ac \oplus bc \oplus ad \oplus bd \oplus acd)$
15 y 31	2	$f_{794} = (ab \oplus ac \oplus bc \oplus d \oplus ad \oplus bd)$
15 y 31	3	$f_{808} = (ab \oplus ac \oplus bc \oplus d \oplus ad \oplus cd)$
15 y 31	4	$f_{819} = (ab \oplus ac \oplus bc \oplus d \oplus bd \oplus cd)$
16 y 32	1	$f_{832} = (ab \oplus c \oplus ac \oplus bc \oplus ad \oplus bd)$
16 y 32	2	$f_{974} = (ab \oplus c \oplus ac \oplus bd \oplus acd \oplus bcd)$
16 y 32	3	$f_{990} = (ab \oplus c \oplus ac \oplus ad \oplus abd \oplus bcd)$
16 y 32	4	$f_{998} = (ab \oplus c \oplus ac \oplus ad \oplus bd \oplus cd)$

5 Pruebas Estadísticas de Aleatoriedad

Las secuencias binarias entregadas por el cifrador, son sometidas a algunas de las pruebas de aleatoriedad [8], de la Norma NIST Special Publication 800-22, del trabajo de Rukhin (et al.) [9].

5.1 Prueba de Frecuencia:

El propósito de esta prueba es determinar si el número de unos y ceros en una secuencia es aproximadamente el mismo que se espera de una secuencia verdaderamente aleatoria. La prueba evalúa la cercanía de la fracción de unos a ½, que es decir, el número de unos y ceros en una secuencia debe ser aproximadamente el mismo. Todas las pruebas posteriores dependen de la aprobación de esta prueba.

$$P_{value} = \left(\frac{S(obs)}{\sqrt{2}}\right).$$

5.2 Prueba de Frecuencia en un Bloque:

El objetivo de esta prueba es determinar si la frecuencia de unos en un bloque de M bits es aproximadamente $M / 2$, como se esperaría bajo un supuesto de aleatoriedad.

$$P_{value} = igamc\left(\frac{N}{2}, \frac{\chi^2(obs)}{2}\right)$$

5.3 Prueba de Rachas:

Una racha de longitud k consta de exactamente k bits idénticos y está acotada antes y después con un poco del valor opuesto. El propósito de la prueba de rachas es determinar si el número de rachas unos y ceros de varias longitudes es lo esperado

para una secuencia aleatoria.

$$P_{value} = erfc\left(\frac{(|V_n(obs) - 2n\pi(1-\pi)|)}{2\sqrt{2n\pi(1-\pi)}}\right)$$

5.4 Prueba de Rachas de Unos en un Bloque.

El propósito de esta prueba es determinar si la longitud de la ejecución más larga de las dentro de la secuencia probada es consistente con la longitud de la serie más larga de las que cabría esperar en una secuencia aleatoria. Tenga en cuenta que una irregularidad en la longitud esperada de la serie más larga implica que también hay una irregularidad en la longitud de la serie más larga de ceros.

$$P_{value} = igamc\left(\frac{K}{2}, \frac{\chi^2(obs)}{2}\right)$$

5.5 Prueba de Entropía Aproximada:

El enfoque de esta prueba es la frecuencia de todas las posibles superposiciones patrones de m bits en toda la secuencia. El propósito de la prueba es comparar la frecuencia de bloques superpuestos de dos longitudes consecutivas / adyacentes (m, m + 1) contra el resultado esperado para un secuencia aleatoria.

$$P_{value} = igamc\left(2^{m-1}, \frac{\chi^2}{2}\right)$$

6 Interpretación de los resultados

Se verificaron cien muestras de 25.600 bits cada una, para el cifrador, con un nivel de significancia de $\alpha = 0{,}01$. La hipótesis nula es:

$$H_0 \rightarrow \text{p\_valor} > 0{,}01$$

A partir de los resultados se realizan dos procedimientos para la interpretación de los mismos:

- Proporción de muestras que pasan las pruebas.
- Prueba de Uniformidad de los p-valor
  - Tabla de frecuencia e histograma
  - Prueba de Bondad de Ajuste

Se aplica la prueba de Bondad de Ajuste χ^2aplicando la siguiente expresión:

$$\chi^2 = \sum_{i=1}^{10} \frac{\left(F_i - \frac{s}{10}\right)^2}{\frac{s}{10}}$$

Donde: F_i = Frecuencia de la clase i s = Cantidad de muestras

El primero se puede realizar considerando los resultados de todas las pruebas o en forma individual, el segundo se realiza en forma individual. En todos los casos se deben superar todas las pruebas para aceptar los resultados.

6.1 Proporción de muestras que pasan las pruebas

Para el análisis de los resultados, se determina la proporción de muestras que superan las pruebas, y con esos datos se construye un gráfico de puntos, luego se verifica si los mismos caen dentro de los límites superior e inferior, donde k es el número de muestras.

Tenemos k = 100 y el nivel de significancia de: α = 0,01.

$$LS, LI = (1 - \alpha) \pm 3 \cdot \sqrt{\alpha \cdot (1 - \alpha)/k}$$

Tabla 3. Pruebas

	Pruebas		Proporción	Superior	Inferior
1	Frecuencias	99	0,99	1,0198	0,96
2	Frecuencias en un Bloque	100	1	1,0198	0,96
3	Rachas	100	1	1,0198	0,96
4	Rachas de Unos en un Bloque	99	0,99	1,0198	0,96
5	Entropía Aproximada	99	0,99	1,0198	0,96

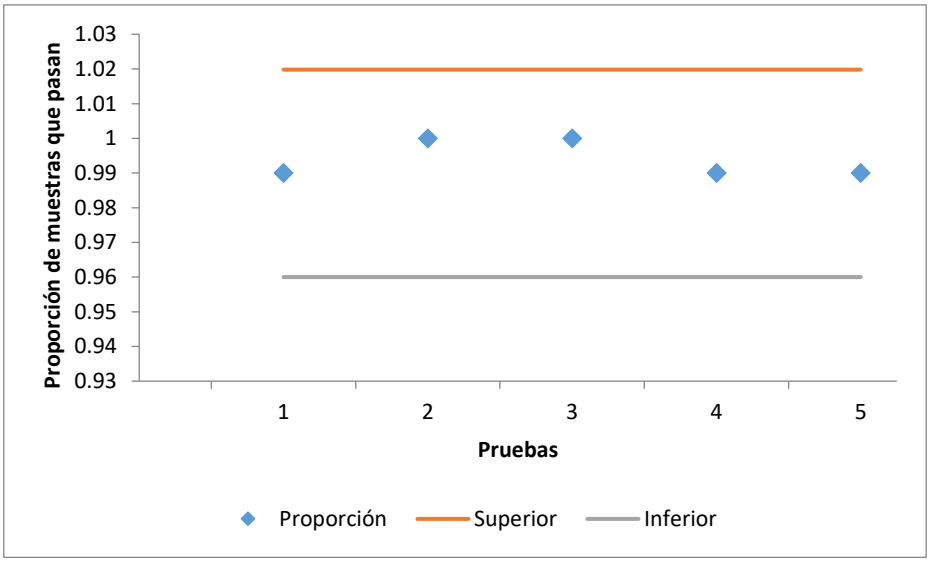

Fig. 8. Gráfico de puntos

6.2 Prueba de bondad de ajuste

En la Tabla 4, se indican los resultados del cálculo de los p valores, correspondientes a todas las pruebas utilizadas y todos cumplen la condición de:

$$\text{p valor} \geq 0{,}0001$$

Tabla 4. Pruebas de bondad de ajuste

Prueba	1	2	3	4	5
	0,00	0,10	0,10	0,00	0,10
	0,90	0,90	2,50	0,40	0,10
	2,50	0,10	0,40	0,00	3,60
	0,40	0,40	0,10	0,10	0,90
	0,40	0,90	0,00	0,40	0,40
	0,40	0,10	0,10	0,10	0,40
	0,00	0,40	0,00	0,40	0,10
	0,00	3,60	0,00	0,40	0,10
	0,90	0,10	0,90	1,60	0,40
	0,10	0,40	0,90	0,40	2,50
Chi cuadrada	5,60	7,00	5,00	3,80	8,60
P valor	**0,78**	**0,64**	**0,83**	**0,92**	**0,47**

6.3 Análisis de las pruebas

Dadas las pruebas sobre las cien secuencias correspondientes al uso de cien claves distintas aplicadas al mismo texto plano, y considerando las interpretaciones vistas en los puntos 6.1 y 6.2, el cifrador cumple con las condiciones de aleatoriedad.

7. Comparación de frecuencias de carácteres

Gráficos de frecuencias de carácteres de:texto plano y texto cifrado, Fig. 9.

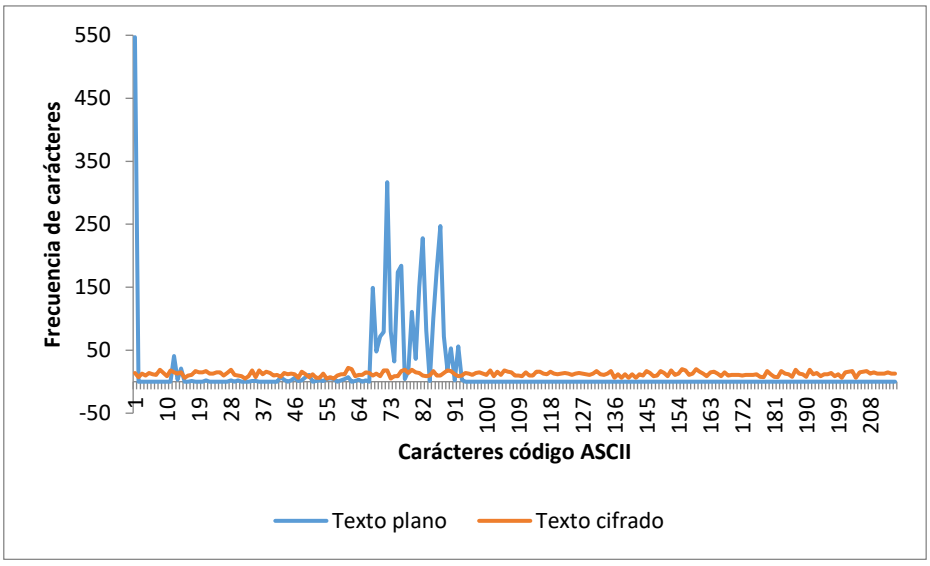

Fig. 9. Frecuencias de caracteres del texto plano y cifrado

Conclusiones

Se ha presentado un cifrador de bloque con algunas características interesantes tales como clave de mayor longitud y la utilización de algoritmo de cifrado basado en funciones booleanas de cuatro variables de buenas propiedades criptográficas y NLFSR de 128 bits.

Sobre este diseño se pueden implementar otras variantes para lograr futuras versiones que contemplen entre otras cosas: bloques mayores, claves más largas, otros modos de operación y algoritmos de cifrado más complejos.

La respuesta de esta versión fue buena y entregó un texto cifrado con una frecuencia de carácteres con cierta uniformidad, lo que hace difícil un criptoanálisis basado en la estadística de aparición de carácteres. También las secuencias de cifrado superaron las pruebas estadísticas de aleatoriedad.

Los cifrados son herramientas útiles cuando se necesita dar seguridad a información de tipo confidencial, esta propuesta es una contribución para este objetivo.

Referencias

1. Karakoç, F., Demirci, H., Harmanc, A.: AKF: A Key Alternating Feistel Scheme for Lightweight Cipher Designs, Information Processing Letters. 115, 359--367 (2015)
2. Bogdanov, A.: Analysis and Design of Block Cipher Constructions. Fakultat fur Elektrotechnik und Informationstechnik an der Ruhr-Universitat Bochum (2009)
3. García Méndez, P.: Descripción Polinomial de los Sistemas de Cifrado DES y AES. Universidad Autónoma Mexicana, México (2011)
4. Fishman, G.: Multiplicative Congruential Random Number Generators with Modulus 2ß : An Exhaustive Analysis for ß = 32 and a Partial Analysis for ß = 48. Mathematics of Computation. 54. (189), 33--344 (1990)
5. Braeken, A.: Cryptographic Properties of Boolean Functions and S-Boxes. Faculteit Ingenieurswetenschappen. Katholieke Universiteit Leuven (2003)
6. Elhosary, A., Hamdy, N., Farag, I., Rohiem, I.: State of the Art in Boolean Functions Cryptographic Assessment. International Journal of Computer Networks and Communications Security.1. (3), 88--94 (2013)
7. Clark, J., Jacob, J., Maitra, S., Stanica, P.: Almost Boolean Functions: The Design of Boolean Functions by Spectral Inversion. Computational intelligence. 20. (3), 450—462 (2004)
8. Duta, C., Mocanu, B., Vladescu, F., Gheorghe, L.: Randomness Evaluation Framework Of Cryptographic Algorithms. International Journal on Cryptography and Information Security (IJCIS), Vol. 4, No. 1, (2014)
9. Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., and Vo, S., “A Statistical Prueba Suite for Random and Pseudorandom Number Generators for Cryptographic Applications”, National Institute of Standards and Technology, (2000).