Ciencias Matemáticas, Vol. 35, No. 1, 2021, Pag. 37-40
Recibido 11-2020, Aceptado 04-2021

Criptoanálisis algebraico a cifrados en bloques ligeros
Algebraic cryptanalysis of lightweight block ciphers

Roberto Labrada Claro[1], Miguel Angel Borges Trenard[2]*, Mijail Borges Quintana[2]

Resumen El campo de la Criptografía Ligera es relativamente nuevo, su esencia consiste en la necesidad de encontrar compromisos entre ligereza y seguridad. En este trabajo realizamos un estudio de SIMON y SIMECK, los cuales son cifrados en bloques ligeros, y aplicamos un ataque algebraico sobre estos cifrados. Comparamos nuestros resultados con los obtenidos por otros autores.

Abstract The field of Lightweight Cryptography is relatively new, its essence consists in the need to find commitments between lightness and security. In this work we present a study of SIMON and SIMECK, which are lightweight block ciphers, and we applied an algebraic attack on these ciphers. We compare our results with those obtained by other authors.

Palabras Clave
cifrados en bloques ligeros — criptoanálisis algebraico — base de Gröbner —

Keywords
lightweight block ciphers — algebraic cryptanalyis — Gröbner basis

[1] *Dirección de Criptografía, Ministerio del Interior, Santiago de Cuba, Cuba, evilrobertolc@gmail.com.*
[2] *Departamento de Matemática, Universidad de Oriente, Santiago de Cuba, Cuba, borgestrenard2014@gmail.com, mijail@uo.edu.cu.*
***Autor para Correspondencia**

1. Introducción

La Criptografía Ligera es relativamente nueva, la misma se centra en buscar un compromiso entre ligereza y seguridad. Una interrogante que guía esta tendencia pudiese ser: ¿cómo se puede llegar a altos niveles de seguridad utilizando pequeñas potencias de cálculo? En años recientes han recibido considerable atención los denominados Cifrados en Bloques Ligeros (**CBL**). En comparación con los cifrados tradicionales, los **CBL** tienen dos propiedades principales:

1) Las aplicaciones sobre dispositivos restringidos no requieren como regla el cifrado de grandes masas de datos, por ello los **CBL** no necesitan tener gran capacidad para el procesamiento.

2) Los **CBL** son usualmente implementados a nivel de hardware, en particular en plataformas tales como microcontroladores de 8 bits.

Los diseñadores de cifrados ligeros tienen que atender la relación costo-seguridad-desempeño. Para los cifrados en bloques ligeros la longitud de la clave del cifrado es la que genera la relación costo-seguridad, la cantidad de rondas del cifrado provee de la relación seguridad-desempeño y la arquitectura de hardware provee la relación costo-desempeño. Usualmente, dos de esas tres relaciones pudiesen ser optimizadas, siendo muy complejo optimizar las tres al mismo tiempo.

Siguiendo la anterior idea, la presentación pública en 2013 del trabajo de especialistas de la Agencia de Seguridad Nacional de los Estados Unidos ([3]), sobre las Familias de Cifrados en Bloques Ligeros SIMON y SPECK, tuvo impacto en la comunidad científica criptográfica, por las características muy buenas de los cifrados y también por el lugar y forma de procedencia de los mismos.

Aunque los cifrados en bloques ligeros se implementan sobre dispositivos restringidos, no sucede lo mismo con los criptoanálisis que a ellos se les realizan. Diversos estudios se han realizado sobre criptoanálisis a cifrados en bloques ligeros, ver por ejemplo: [6], en que se realiza un ataque algebraico a cinco rondas del cifrador en bloques SIMON, [1], donde se muestra que la versión más pequeña de SIMON exhibe un marcado efecto diferencial. En la mayoría de los casos, las observaciones presentadas no llevan abiertamente a un ataque, sino que proveen de bases para futuros análisis a las variantes de cifrados especificadas.

Se hace entonces necesario la programación, comprensión y conocimientos sobre vulnerabilidades de este tipo de cifrados ante los ataques más conocidos, tanto de forma computacional, como en sus fundamentos matemáticos. En esa dirección se encuentra el propósito del presente trabajo, el cual resume una parte de la tesis de maestría [4].

2. Desarrollo

2.1 Familia de Cifrados en Bloques Ligeros

Comprender el funcionamiento interno de los cifrados es el primer paso para encontrar debilidades o fortalezas en ellos, requisito indispensable al momento de ofrecer un criterio de seguridad y decidir qué cifrado usar y bajo cuáles condiciones. Por lo anterior, a continuación se describen las funciones de cifrado y descifrado de los algoritmos escogidos y algunas de sus propiedades, así como se discute sobre paquetes de cálculos para la ejecución de estos cifrados.

2.1.1 Familia SIMON

SIMON [3] es un cifrado liviano, diseñado por la Agencia de Seguridad Nacional de los Estados Unidos, publicado en el verano del 2013, puede ser implementado tanto en software como en hardware y posee un alto desempeño en una variedad de dispositivos. Para ser tan flexible como posible, sus diseñadores presentaron a la familia SIMON de 10 algoritmos, la cual consiste en una combinación de diferentes tamaños de bloques y claves.

Función de Ronda

SIMON pertenece a la familia de cifrados en bloques del tipo Feistel. El cifrado y descifrado se basan en las operaciones XOR, AND y la rotación circular $\lll j$ (rotar j bits a la izquierda). Para cada ronda (r_i), el esquema de Feistel opera con dos listas de n-bits, una izquierda (L_{i-1}) y otra derecha (R_{i-1}), creando el estado de ronda de $2n$-bits. La mitad izquierda pasa a través de una función en cada ronda:

$$f(L_{i-1}) = (L_{i-1} \lll 1)\&(L_{i-1} \lll 8) \oplus (L_{i-1} \lll 2),$$

donde & denota la operación AND y $\oplus$ denota a XOR. Luego, $f(L_{i-1})$ se suma a la otra mitad R_{i-1} y a la clave de ronda K_i

$$L_i = f(L_{i-1}) \oplus R_{i-1} \oplus K_i,$$

para crear la nueva parte izquierda L_i y definir la nueva mitad derecha como $R_i = L_{i-1}$.

Generación de las claves de ronda

La función de generación de las claves de ronda es similar a la función de ronda; utiliza las operaciones XOR y la rotación circular, solo que ahora es hacia la derecha, $\ggg j$ (rotar j veces a la derecha). Para evitar propiedades lineales y simetrías en las rotaciones circulares, se agrega a la generación de las claves una sucesión de constantes de rondas z_j. Una constante c también se suma junto con z_j, donde $c = (2n-4)$. Dependiendo del número $m = 2,3,4$, escogido para la entrada, la generación de las claves será: $k_{i+m} =$

$$\begin{aligned} c \oplus (z_j)_i \oplus k_i (I \oplus (\ggg 1))(\ggg 3) k_{i+1}, \text{ si } m = 2, \\ c \oplus (z_j)_i \oplus k_i (I \oplus (\ggg 1))(\ggg 3) k_{i+2}, \text{ si } m = 3, \\ c \oplus (z_j)_i \oplus k_i (I \oplus (\ggg 1))(\ggg 3) k_{i+3} \oplus k_{i+1}, \text{ si } m = 4. \end{aligned} \quad (1)$$

para $0 \leq i < T - m$, donde T es el número de rondas.

2.2 Modelación algebraica de 5 rondas del SIMON

En [2] se obtiene una representación de las transformaciones del cifrado, como un sistema de ecuaciones polinómicas en varias variables, reducido a 5 rondas del SIMON32. Interpretando el algoritmo y reordenando las ecuaciones de forma tal que las variables desconocidas se encuentren en el miembro izquierdo y sustituyendo la variable desconocida x_3 por el valor conocido C_R, el sistema de ecuaciones quedaría expresado como sigue:

$$x_0 \oplus k_1 = f(P_L) \oplus P_R, \quad (2)$$
$$x_1 \oplus f(x_0) \oplus k_2 = P_L, \quad (3)$$
$$x_2 \oplus f(x_1) \oplus x_0 \oplus k_3 = 0, \quad (4)$$
$$f(x_2) \oplus x_1 \oplus k_4 = C_R, \quad (5)$$
$$x_2 \oplus k_5 = f(C_R) \oplus C_L, \quad (6)$$
$$k_5 \oplus k_1 \oplus k_2 \oplus (k_4 \ggg 3) \oplus (k_2 \ggg 1) \oplus (k_4 \ggg 4) = D. \quad (7)$$

donde (P_L, P_R) y (C_L, C_R) son los lados izquierdo y derecho de los textos claros y cifrado respectivamente, k_i es la i-ésima clave de ronda, x_{i-1} es la parte izquierda de la salida después de la ronda $i-1$, y D es una constante conocida, que es la que permite generar cada clave k_{i+m} a partir de las siguientes, según las ecuaciones (1).

Para un par o dos pares de texto claro y cifrado, la cantidad de incógnitas supera a la cantidad de ecuaciones, a partir de 3 pares, se observa una tendencia a que el número de ecuaciones es superior al de incógnitas, que es justo lo que se recomienda, tanto para la unicidad de la solución, como para la posibilidad de resolverlo. El siguiente paso es transformar las 5 ecuaciones vectoriales a sus respectivas representaciones por componentes.

2.2.1 Ataque

Para calcular la solución de los sistemas, se utilizó el paquete de bases de Gröbner que posee el sistema de cálculo simbólico MAPLE (Versión 18). Todos los cálculos realizados se llevaron a cabo en un computador con un procesador Intel Core i7-4790, a 3.6 GHz, con 16 GB de RAM. Para realizar los experimentos se tomó en cada intento una clave y los respectivos textos claros de forma aleatoria, realizándose 100 intentos por cada experimento, el tiempo resultante mostrado en las Tablas denota el promedio de tiempo de los ataques. Las opciones en la columna de resultados son S (cuando se obtuvo éxito, es decir, se obtuvo la clave) y N cuando no se pudo obtener, que los recursos del computador no fueron suficientes para realizar los cálculos.

Para cada par en cuestión, se obtuvieron las ecuaciones de ronda correspondientes. Luego de alcanzado el sistema polinómico, se creó el álgebra de polinomios junto con las ecuaciones del campo, para garantizar que las soluciones se mantuviesen en el campo de trabajo y no en alguna extensión del mismo. Los resultados se listan a continuación (ver página siguiente).

2.3 Modelación algebraica de 5 rondas del SIMECK

La forma de proceder del cifrado SIMECK [7] es muy parecida a la de SIMON y se puede apreciar similitud entre

Tabla 1. Resultados del criptoanálisis algebraico al SIMON32/64.

Rondas, Pares, Tiempo(s)	Resultado
5, 1, 32752.425	N
5, 2, 1059.797	S
5, 3, 2.687	S
6, 3, 13.000	S
7, 3, 27290.922	N
7, 4, 24350.594	N
7, 5, 31623.891	N
7, 6, $1{,}94481172 * 10^5$	N

las primitivas de los cifrados SIMON y SIMECK (por limitaciones de espacio no se detalla en este documento). Surge entonces de forma natural la interrogante siguiente: dada las pocas diferencias de los algoritmos y teniendo en cuenta el ataque algebraico ya realizado sobre el SIMON, ¿el cifrado SIMECK ofrece mayor o menor resistencia que el cifrado SIMON?

Con el objetivo de responder tal interrogante, para hacerle un ataque al cifrado SIMECK y poder realizar comparaciones, se decidió proceder de la misma forma para realizar el criptoanálisis algebraico. Utilizando el mismo método, se logró obtener una representación de las transformaciones del cifrado como un sistema de ecuaciones polinómicas en varias variables.

2.3.1 Ataque

Para cada par en cuestión, se obtuvieron las ecuaciones de ronda correspondientes, luego de alcanzado el sistema polinómico, se creó el álgebra de polinomios junto con las ecuaciones del campo, para garantizar que las soluciones se mantuviesen en el campo de trabajo y no en alguna extensión del mismo. Los resultados se listan en la Tabla 2.

Tabla 2. Resultados del criptoanálisis algebraico al SIMECK32/64.

Rondas, Pares, Tiempo(s)	Resultado
5, 1, 23949.844	N
5, 2, 1606.359	S
5, 3, 2.903	S
6, 3, 17651.625	S
6, 4, 408.266	S
7, 3, 11697.141	N
7, 4, 15716.469	N
7, 5, 19288.454	N
7, 6, 33762.937	N

2.4 Resumen y contraste de los ataques

2.4.1 SIMON

Desde la vista a luz pública del cifrado SIMON, varios artículos han intentado atacarlo, véase por ejemplo [2], donde se realiza un ataque algebraico a texto claro conocido de una versión reducida del SIMON, utilizando un software del tipo "SAT-solver" ([5]). El ataque al SIMON en [2] se realizó en un computador con procesador Intel Core i7 2.70 GHz, con 16 GB de RAM, y fue implementado con el SAT-solver CryptoMiniSat2. Ellos atacaron 5 y 6 rondas del cifrado logrando:

- 5 rondas del SIMON32/64, con tres pares de texto claro y sus respectivos textos cifrados, el tiempo de ejecución fue de un promedio de 3.75s, logrando obtener la clave.
- 6 rondas del SIMON32/64, con tres pares de texto claro y sus respectivos textos cifrados, el tiempo de ejecución fue de un promedio de 290.7s, logrando obtener la clave.
- Para dos pares, según afirman los autores, se obtenían sistemas que no tenían solución única.

Otro objetivo que nos llevó a experimentar con el SIMON vino directamente de lo comentado en el párrafo anterior, ya que nos surgió la siguiente interrogante: entre el empleo de SAT-solver o el cálculo mediante Bases de Gröbner, ¿con cual método de solución de sistemas de ecuaciones polinómicas se obtiene mejor resultado? En tal sentido, se pudo apreciar que se mejoró el tiempo de búsqueda de la clave logrado en [2], lográndose incluso calcular la clave con sólo 2 pares, lo cual no se obtuvo en [2]; en general, el resultado con las bases de Gröbner en MAPLE fue mejor que con los sistemas SAT-solvers utilizados en [2].

2.4.2 SIMECK

El cifrado SIMECK es un derivado de los cifrados SIMON y SPECK ([3]), el cual según los autores, "ha tomado de cada uno lo mejor para constituir un cifrado más seguro y más eficiente". Resultó entonces interesante ver cómo se comportaban ambos cifrados bajo el mismo ataque. La conclusión que se obtuvo fue que, para cada experimento, el tiempo que demoró calcular la clave del cifrado SIMECK se mantuvo siempre por encima del tiempo que demoró hallar la clave del SIMON, que era lo esperado.

3. Conclusiones

Los resultados descritos en este trabajo muestran un estudio de los cifrados pertenecientes a la Familia de Cifrados en Bloques Ligeros y una valoración sobre su resistencia al criptoanálisis algebraico. Se estudió con detalle el modo de operar de las funciones de cifrado y descifrado de los algoritmos escogidos, así como algunas propiedades que poseen. También se programaron paquetes de cálculo para el cifrado y descifrado de los principales miembros de estas familias.

Se halló la modelación algebraica de importantes operaciones de cifrado y fueron programadas algunas técnicas para el criptoanálisis algebraico. Estas técnicas se aplicaron a los cifrados SIMON y SIMECK.

Como resultado directo de las experimentaciones, se llegó a la conclusión que en general, el resultado con las bases de

Gröbner en MAPLE fue mejor que con los sistemas "SAT-solvers", utilizados en [2], lo que constituye una alternativa al ataque antes mencionado.

Referencias

[1] AlKhzaimi, Hoda y Martin M Lauridsen: *Cryptanalysis of the SIMON Family of Block Ciphers.* IACR Cryptology ePrint Archive, 2013:543, 2013.

[2] Astrid, Berghult: *A practical comparison between algebraic and statistical attacks on the lightweight cipher SIMON*, 2016.

[3] Beaulieu, Ray, Stefan Treatman-Clark, Douglas Shors, Bryan Weeks, Jason Smith y Louis Wingers: *The SIMON and SPECK lightweight block ciphers*. En *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, páginas 1–6. IEEE, 2015.

[4] Claro, R. Labrada: *Criptoanálisis Algebraico a Cifrados en Bloques Ligeros*. Tesis en opción al título de Máster en Matemática. Universidad de La Habana, Cuba, 2018.

[5] Gomes, Carla P, Henry Kautz, Ashish Sabharwal y Bart Selman: *Satisfiability solvers*. Foundations of Artificial Intelligence, 3:89–134, 2008.

[6] M.A. Borges-Trenard, R. Labrada Claro: *Ataque algebraico a cinco rondas del cifrado en bloques SIMON*. Congreso Internacional Compumat. La Habana, Cuba, 2017.

[7] Yang, Gangqiang, Bo Zhu, Valentin Suder, Mark D Aagaard y Guang Gong: *The simeck family of lightweight block ciphers*. En *International Workshop on Cryptographic Hardware and Embedded Systems*, páginas 307–329. Springer, 2015.