Criptoanálise Linear do Sistema AES Simplificado Modificado por Sequências Caóticas

José A. P. Artiles, Daniel P. B. Chaves, Cecilio Pimentel

***Resumo*— Este artigo descreve a criptoanálise linear para o esquema *Advanced Encryption Standard* (AES) simplificado com a introdução de uma técnica de randomização desta cifra empregando mapas caóticos. Mostra-se que com a utilização de mapas caóticos o adversário precisa de uma quantidade substancialmente maior de pares de texto claro e texto cifrado para descobrir bits de chave com certa confiabilidade. Com este resultado é possível entender o impacto dos bits caóticos no algoritmo AES original.**

***Palavras-Chave*— Cifra de bloco, criptoanálise linear, mapas caóticos, padrão avançado de criptografia simplificado.**

***Abstract*— This paper describes the linear cryptanalysis of the simplified Advanced Encryption Standard (AES) with the introduction of a randomization technique employing chaotic maps. It is shown that with the use of chaotic maps the adversary needs a larger number of pairs of plaintext and ciphertexts to discover key bits with a certain reliability. Given these results, it is possible to understand the impact of the chaotic bits in the original AES algorithm.**

***Keywords*— Block chipher, linear cryptanalysis, chaotic maps, simplified advanced encryption standard.**

I. Introdução

O *Advanced Encryption Standard* (AES) é o algoritmo padrão adotado pelo Instituto Nacional de Padrões e Tecnologia (NIST, *National Institute of Standards and Technology*) para substituir o *Data Encryption Standard* (DES) para a proteção de informações sigilosa. Trata-se do Rijndael com comprimento do bloco de dados fixo em 128 bits [1], podendo, em sua versão original, operar também com blocos de 192 e 256 bits. O NIST detalha este padrão em seu FIPS PUB 197 (*Federal Information Processing Standards Publications 197*), um documento aberto para consulta geral [2]. Em [3] foi apresentado um algoritmo AES modificado (a saída de cada S-Box é somada com bits gerados por um mapa caótico) com a finalidade da redução da complexidade e número de operações realizadas no algoritmo original. O algoritmo modificado é denominado de AES1 e uma versão simplificada (denominada de AES2) também foi proposta em [3]. Estes algoritmos foram analisados por diferentes quantificadores comumente usados para avaliar tanto a aleatoriedade do texto cifrado como a capacidade de resistência à ataques estatísticos [3]. Outras análises também devem ser realizadas em algoritmos criptográficos, como por exemplo a sua robustez contra a criptoanálise linear (CL). Um trabalho precursor da CL foi introduzido por Matsui [4] em 1992. Em 1993, esta técnica foi usada como um ataque ao DES [5].

Departamento de Eletrônica e Sistemas, Universidade Federal de Pernambuco, Recife-PE, Brasil, E-mails: joseantonio.artiles@ufpe.br, daniel.chaves@ufpe.br, cecilio@ufpe.br. Este trabalho foi parcialmente financiado pelo CNPq e FACEPE.

O esforço computacional para avaliar a eficácia do CL no Rijndael pode ser proibitivo, como solução, emprega-se uma versão mais simples deste algoritmo. Em [6] foi proposto o algoritmo AES simplificado (SAES) que apresenta o comprimento do bloco de dados menor que o AES, sem contudo perder a essência do algoritmo original. Isso significa que, ao entender o algoritmo SAES e expandir seus conceitos, pode-se entender o comportamento da criptoanálise do AES de forma mais simples. O objetivo deste trabalho é propor algoritmos SAES modificados por sequências caóticas que seguem os mesmos princípios do AES1 e AES2 (denominados de SAES1 e SAES2) e estudar a CL para estes sistemas. Verifica-se que estes são consideravelmente mais robustos contra a CL que o SAES. Uma nova versão simplificada (SAES3), que estabelece um compromisso entre complexidade computacional e segurança entre o SAES1 e SAES2 também é proposta neste trabalho.

O restante deste artigo está organizado em quatro seções. A Seção II descreve o algoritmo SAES. A CL para o sistema SAES é analisada na Seção III e esta análise é estendida para os sistemas SAES1, SAES2, SAES3 na Seção IV. Uma comparação da robustez destes sistemas contra CL é realizada nesta seção. As conclusões deste trabalho são resumidas na Seção V.

II. Preliminares

A. Algoritmo AES simplificado

No SAES [6], cada bloco de entrada (texto claro) tem 16 bits $\{x_0, \cdots, x_{15}\}$ e a chave original também tem 16 bits $\{k_0, \cdots, k_{15}\}$. Considera-se duas iterações, sendo necessário expandir os 16 bits da chave criando a primeira sub-chave $\{k_{16}, \cdots, k_{31}\}$ e a segunda sub-chave $\{k_{32}, \cdots, k_{47}\}$, totalizando 48 bits.

Na primeira iteração, a chave original é somada bit a bit (módulo 2) com o texto claro e realiza-se as mesmas etapas do algoritmo AES original (substituição realizada por S-Boxes, deslocamento de linhas e mistura de colunas). Na segunda iteração só realiza-se a etapa de substituição e somas de sub-chaves, como ilustra a Fig. 1. A saída do SAES é o texto cifrado $\{y_0, \cdots, y_{15}\}$. O conjunto de operações realizadas em cada iteração é descrito a seguir.

1) **Substituição**: Os bits de entrada da etapa de substituição são dados por $a_i = x_i \oplus k_i$, para $\{i = 0, 1, \cdots, 15\}$. Esta etapa compreende 4 S-Boxes idênticas que operam de forma paralela, tendo cada S-box 4 bits (*nybble*) de entrada e 4 bits de saída. Os bits de saída são obtidos por um mapeamento não linear e reversível através de operações definidas em $GF(2^4)$, geradas pelo polinômio primitivo $P(x) = x^4 + x + 1$.

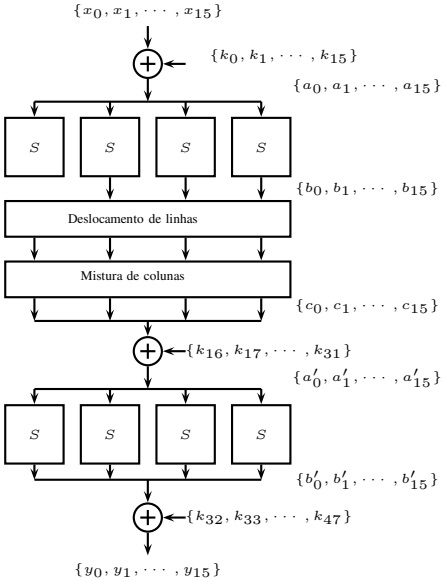

Fig. 1. Diagrama de blocos do algoritmo SAES com duas iterações.

Seja a_0, a_1, a_2, a_3 a entrada de uma S-Box. Inicialmente, esta sequência é invertida em $GF(2^4)$ (0000 não é invertível, então, nesta etapa, sua saída é 0000). A sequência de entrada invertida $a_0^-, a_1^-, a_2^-, a_3^-$ é utilizada para obter a saída da S-box da seguinte forma:

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_0^- \\ a_1^- \\ a_2^- \\ a_3^- \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{1}$$

2) **Deslocamento das linhas**: Neste bloco a sequência $\{b_0, \cdots, b_3, b_4, \cdots, b_7, b_8, \cdots, b_{11}, b_{12}, \cdots, b_{15}\}$ é mapeada em $\{b_0, \cdots, b_3, b_{12}, \cdots, b_{15}, b_8, \cdots, b_{11}, b_4, \cdots, b_7\}$, sendo b_i os bits da saída da etapa de substituição.

3) **Mistura de colunas**: A etapa de mistura de colunas realiza uma mistura nos bits de saída da etapa de deslocamento de linhas, com a finalidade de misturar saídas de S-box distintas. Este mapeamento é detalhado em [6].

4) **Adição de sub-chave**: Nesta etapa soma-se bit a bit (módulo 2) os 16 bits de saída do bloco mistura de colunas com os bits da sub-chave $\{k_{16}, \cdots, k_{31}\}$, finalizando a primeira iteração. Os bits resultantes desta soma são a entrada do bloco de substituição da segunda iteração. Os bits de saída deste bloco $\{b'_1, b'_2, \cdots, b'_{15}\}$ são somados com os bits da sub-chave $\{k_{32}, \cdots, k_{47}\}$ para gerar o texto cifrado:

$$y_i = b'_i + k_{i+32} \tag{2}$$

De forma geral podemos definir os bits da chave como:

$$\begin{aligned} k_{16} &= k_0 \oplus l_0 \oplus 1 \\ k_i &= k_{i-16} \oplus l_{i-16}, \{i = 17, 18, \cdots, 23\} \\ k_i &= k_{i-16} \oplus l_{i-24}, \{i = 32, 33, 36, \cdots, 39\} \\ k_{34} &= k_{18} \oplus l_{10} \oplus 1 \\ k_{35} &= k_{19} \oplus l_{11} \oplus 1 \\ k_i &= k_{i-8} \oplus k_{i-16}, \{i = 24, 25, \cdots, 31, 40, 41, \cdots, 47\} \end{aligned} \tag{3}$$

sendo $(l_0, \cdots l_{15})$ as saídas das S-Boxes. Observa-se em (3) a existência 16 bits de chave $\{i = 24, 25, \cdots, 31, 40, 41, \cdots, 47\}$ que são combinação linear de outros bits de chaves, sendo preciso encontrar os outros 32 bits.

III. Criptoanálise Linear

A CL explora relações lineares dos bits de entrada e de saída das S-boxes que apresentam alta probabilidade, uma vez que as S-boxes, geralmente, são os únicos elementos não lineares em uma cifra de bloco. A CL é um ataque de texto claro conhecido, isto é, o adversário conhece em um conjunto de pares de texto claro e texto cifrado obtidos com a mesma chave. A ideia da CL é encontrar equações da forma:

$$\sum_{k \in S_1} x_k \oplus \sum_{l \in S_2} y_l = \left(\sum_{m \in S_3} k_m \right) \oplus t \tag{4}$$

que apresente probabilidade maior que $0,5$, sendo t um bit com valor 0 ou 1, x_k denota o k-ésimo bit de texto claro, y_l denota o l-ésimo bit de texto cifrado, k_m representa o m-ésimo bit da chave e cada S_i é um subconjunto de $\{0, 1, \cdots, 15\}$. Para cada equação, o adversário avalia o lado esquerdo de (4) para cada par de texto claro e texto cifrado e estima a probabilidade do lado direito está correto. Define-se a probabilidade que a ℓ-ésima equação esteja correta por p_ℓ, em que o bit t é escolhido para o qual $0,5 \leq p_\ell \leq 1$. Se um algoritmo de cifragem mostra uma tendência que (4) seja satisfeita com probabilidade $1/2$, isto é uma evidência de um algoritmo criptográfico robusto para esta criptoanálise. Quanto mais afastada a probabilidade p_ℓ é de $1/2$, mais eficaz é a CL.

A. Criptoanálise Linear do SAES

Esta seção analisa a CL do esquema SAES introduzido na Seção II-A. A ideia central consiste em encontrar equações lineares correspondentes aos bits de entrada e de saída das S-boxes que possuem probabilidade maior que $0,5$. Existem para cada S-box, 256 equações para todas as possíveis combinações de entrada e de saída, sendo possível extrair 12 equações com probabilidade $0,75$ entre as entradas e saídas de cada S-box, como por exemplo, $a_0 \oplus b_2 = 0$, $a_2 \oplus a_3 \oplus b_3 = 1$. Dada a existência de 4 S-boxes por iteração, na saída da primeira etapa de substituição obtém-se 48 equações que dependem dos bits de texto claro, dos bits da chave e os bits de saída da primeira etapa de substituição $\{b_0, b_1, \cdots, b_{15}\}$, como por exemplo, temos $b_5 \oplus x_5 = k_5$ e $b_8 \oplus b_{11} \oplus x_9 = k_9$, cada equação com probabilidade $0,75$.

Como o adversário tem conhecimento de pares de texto claro $\{x_0, \cdots, x_{15}\}$ e de texto cifrado $\{y_0, \cdots, y_{15}\}$ é preciso realizar o mesmo procedimento na segunda etapa de substituição. Dado que entre a saída da primeira etapa de substituição (primeira iteração) e a entrada da segunda etapa de substituição todas as operações são lineares, pode-se escrever $\{a'_1, \cdots, a'_{15}\}$ em função dos bits $\{b_0, \cdots, b_{15}\}$, bem como $\{b'_0, \cdots, b'_{15}\}$ em função de $\{y_0, \cdots, y_{15}\}$, obtendo assim equações da forma $b_5 \oplus b_8 \oplus b_{11} \oplus y_{12} \oplus y_{15} = k_{29} \oplus k_{44} \oplus k_{47} \oplus 1$ na segunda etapa de substituição. Para obter equações da forma

mostrada em (4) deve ser realizada uma combinação (soma módulo 2) de equações de cada etapa, como por exemplo:

$$\begin{array}{c} b_5 \oplus x_5 = k_5 \\ b_8 \oplus b_{11} \oplus x_9 = k_9 \\ \hline b_5 \oplus b_8 \oplus b_{11} \oplus y_{12} \oplus y_{15} = k_{29} \oplus k_{44} \oplus k_{47} \oplus 1 \\ x_5 \oplus x_9 \oplus y_{12} \oplus y_{15} = k_5 \oplus k_9 \oplus k_{29} \oplus k_{44} \oplus k_{47} \oplus 1 \end{array} \quad (5)$$

Como cada equação tem uma probabilidade de ser satisfeita, a probabilidade da equação resultante será discutida na seguinte subseção.

1) Combinação de equações: Como é apresentado em [7], expressões obtidas na primeira e segunda iteração da etapa de substituição são consideradas uma variável aleatória binária. Seja X uma variável aleatória referente a uma expressão da primeira iteração e similarmente Y uma variável aleatória referente a segunda iteração, de tal forma que $P(X = 1) = p_1$ e $P(Y = 0) = p_2$, sendo $0,5 < p_1, p_2 < 1$. Considera-se que X e Y são variáveis aleatórias independentes. Então:

$$\begin{aligned} P(X \oplus Y = 1) &= P(X = 1, Y = 0) + P(X = 0, Y = 1) \\ &= P(X = 1)P(Y = 0) + P(X = 0)P(Y = 1) \\ &= p_1 p_2 + (1 - p_1)(1 - p_2). \end{aligned} \quad (6)$$

Quando $p_1 = p_2 = p$, obtém-se $q \triangleq P(X \oplus Y = 1) = 2p^2 - 2p + 1$. É importante frisar que $0,5 \leq q \leq p$ no intervalo $0,5 \leq p \leq 1$, uma vez que

$$q - p = 2p^2 - 3p + 1 = (2p - 1)(p - 1). \quad (7)$$

Para o intervalo válido para p, o termo $(p - 1)$ tem valores negativos e o termo $(2p - 1)$ valores positivos, obtendo $q - p \leq 0$. Seguindo um raciocínio análogo, pode ser mostrado que a combinação de equações com probabilidades distintas é limitado superiormente por:

$$q = p_1 p_2 + (1 - p_1)(1 - p_2) \leq \min(p_1, p_2). \quad (8)$$

No intervalo válido de p_1 e p_2, existem duas regiões de mínimos para q, em $p_1 = 0,5$ ou $p_2 = 0,5$, e um valor de máximo quando $p_1 = p_2 = 1$. Portanto, a probabilidade de combinação de equações com probabilidades distintas é limitada pela menor delas atingindo o menor valor igual a $0,5$ quando umas das probabilidades é 0,5.

Realizando uma combinação entre as equações obtidas na primeira e segunda iteração da etapa de substituição obtemos equações da forma (4). A soma de $b_5 \oplus x_5 = k_5$ e $b_8 \oplus b_{11} \oplus x_9 = k_9$, resulta em uma equação com probabilidade:

$$\begin{aligned} q &= 2p^2 - 2p + 1 \\ &= 2(0,75)^2 - 2(0,75) + 1 = 0,625 \end{aligned} \quad (9)$$

que somada com $b_5 \oplus b_8 \oplus b_{11} \oplus y_{12} \oplus y_{15} = 1 \oplus k_{29} \oplus k_{44} \oplus k_{47}$ obtém-se uma equação com probabilidade:

$$\begin{aligned} q_1 &= p_1 p_2 + (1 - p_1)(1 - p_2) \\ &= (0,75)(0,625) + (1 - 0,75)(1 - 0,625) \\ &= 0,5625. \end{aligned} \quad (10)$$

Realizando este processo em todas as equações obtidas nas duas etapas de substituição, obtemos 32 equações linearmente independentes (quantidade de bits no algoritmo de chaves que tem influencia direta de S-box). A questão principal é, quantos pares de texto claro e texto cifrado n são necessários para o adversário quebrar a cifra (com uma certa confiabilidade) usando as 32 equações. Define-se que o adversário quer uma confiabilidade de 95% que os bits da chave obtidos sejam corretos.

Seja W uma variável aleatória definida como o número de pares de texto claro e texto cifrado para o qual o lado direito de cada equação obtida da forma (5) é o valor correto (0 ou 1) do lado esquerdo dividido por n, para uma dada chave. Cada par de texto claro e texto cifrado é uma realização do experimento e a probabilidade do lado direito da equação seja correto é q_1. Cada realização é independente e podemos descrever este processo como uma distribuição binomial normalizada pela quantidade de pares texto claro texto cifrado n. Então, a média de W é q e sua variância é $\sigma^2 = q(1 - q)/n$.

Para a CL é desejado que $P(W \geq 0,5)$. Usando a confiabilidade estabelecida, então $P(W \geq 0,5) = \sqrt[32]{0,95} = 0,9984$. Para um valor suficientemente grande de n, W tende a uma variável aleatória normal. Definindo uma variável aleatória normal $Z = (W - q)/\sigma$ de média zero e variância unitária, temos que:

$$\begin{aligned} P(W \geq 0,5) &= P\left(Z \geq \frac{0,5 - q}{\sigma}\right) \\ &= 1 - Q\left(\frac{\sqrt{n}(q - 0,5)}{\sqrt{q(1 - q)}}\right) = 0,9984 \end{aligned} \quad (11)$$

em que a função $Q(x)$ é a área da cauda de uma normal de média zero e variância unitária. Para o caso em que $q = 0,5625$, obtém-se:

$$Q(0,126\sqrt{n}) = 0,0016. \quad (12)$$

O argumento da função $Q(x)$ que satisfaz (12) é 2,94, então a partir de $0,126\sqrt{n} = 2,94$, obtém-se $n = 544,55$. Desta forma, necessita-se 545 pares de texto claro e texto cifrado para achar os bits de chave com uma confiabilidade de 95 %. Na próxima seção, uma análise similar será feita para o sistema SAES modificado por uma sequência caótica.

IV. Criptoanálise Linear do SAES Modificado com Sequência Caótica

Nesta seção analisa-se a CL para três algoritmos baseados no SAES com as S-Boxes modificadas por uma sequência caótica. Estes são denominados de SAES1, SAES2, SAES3.

A. SAES1

Neste algoritmo, os 4 bits de saída de cada S-Box são somados (XOR) bit a bit com uma sequência **h** gerada a partir de um mapa caótico. Para cada S-box no caminho de dados, utiliza-se dois bits caóticos (z_0, z_1), com representação polinomial dada por $c(x) = z_0 x + z_1$. Como as operações de cada S-box são definidas em GF(2^4), multiplica-se $c(x)$ por um polinômio primitivo $p(x) = x^3 + x + 1$ em GF(2^3), obtendo assim o polinômio $h(x) = c(x)p(x) \mod P(x)$.

Os coeficientes deste polinômio formam uma sequência $\mathbf{h} = (h_0, h_1, h_2, h_3)$. Este mapeamento é dado por:

$$\begin{array}{rcl} (z_0, z_1) & \blacktriangleright & (h_0, h_1, h_2, h_3) \\ (0,0) & \blacktriangleright & (0,0,0,0) \\ (0,1) & \blacktriangleright & (1,0,1,1) \\ (1,0) & \blacktriangleright & (0,1,0,1) \\ (1,1) & \blacktriangleright & (1,1,1,0) \end{array} \quad (13)$$

Dois bits caóticos (z_2, z_3) são usados na obtenção das sub-chaves de forma similar, totalizando quatro bits caóticos para cifrar um bloco de 16 bits de texto claro. Para simplificar a análise, na segunda iteração utiliza-se a mesma sequência $\mathbf{h}$ usada na iteração anterior. Observa-se em (13) que h_0 e h_2 são iguais a z_1, h_1 é igual a z_0, e h_3 é igual a $z_0 \oplus z_1$. No caminho de dados, os bits de saída das S-boxes na primeira iteração do SAES1 se relacionam com os bits de saída do SAES da seguinte forma:

$$[\hat{b}_i, \hat{b}_{i+1}, \hat{b}_{i+2}, \hat{b}_{i+3}] = [b_i, b_{i+1}, b_{i+2}, b_{i+3}] \oplus [z_1, z_0, z_1, z_0 \oplus z_1], \quad (14)$$

sendo $i = \{0, 4, 8, 12\}$. De forma similar são obtidos os bits na saída da segunda iteração $(\hat{b}'_0, \cdots, \hat{b}'_{15})$.

A utilização dos bits caóticos tem um impacto direito nos bits de saída da etapa de substituição. Na saída da primeira etapa de substituição são modificados pela presença dos bits caóticos, obtendo $[\hat{b}_i, \hat{b}_{i+1}, \hat{b}_{i+2}, \hat{b}_{i+3}] = [b_i, b_{i+1}, b_{i+2}, b_{i+3}] \oplus [z_1, z_0, z_1, z_0 \oplus z_1]$, para $i = \{0, 4, 8, 12\}$. De forma similar no algoritmo de obtenção de sub chaves, os bits serão modificados $[\hat{k}_i, \hat{k}_{i+1}, \hat{k}_{i+2}, \hat{k}_{i+3}] = [k_i, k_{i+1}, k_{i+2}, k_{i+3}] \oplus [z_3, z_2, z_3, z_2 \oplus z_3]$, para $i = \{16, 20, 24, 28, 32, 36, 40, 44\}$. Na saída da segunda iteração os bits caóticos tem um impacto direto nos bits de texto cifrado, sendo este $[\hat{y}_i, \hat{y}_{i+1}, \hat{y}_{i+2}, \hat{y}_{i+3}] = [y_i, y_{i+1}, y_{i+2}, y_{i+3}] \oplus [z_1, z_0, z_1, z_0 \oplus z_1]$, com $i = \{0, 4, 8, 12\}$. O algoritmo SAES1 tem a mesma estrutura que o SAES, portanto as equações obtidas no algoritmo SAES1 são similares, bastando substituir b_i, y_i e k_i por $\hat{b}_i$, $\hat{y}_i$ e $\hat{k}_i$, respectivamente. Por exemplo, a equação $b_5 \oplus b_8 \oplus b_{11} \oplus y_{12} \oplus y_{15} = 1 \oplus k_{29} \oplus k_{44} \oplus k_{47}$ do algoritmo SAES, torna-se:

$$\hat{b}_5 \oplus \hat{b}_8 \oplus \hat{b}_{11} \oplus \hat{y}_{12} \oplus \hat{y}_{15} = \hat{k}_{29} \oplus \hat{k}_{44} \oplus \hat{k}_{47} \oplus 1$$

ou

$$b_5 \oplus z_0 \oplus b_8 \oplus z_1 \oplus b_{11} \oplus z_0 \oplus z_1 \oplus y_{12} \oplus z_0 \oplus y_{15} \oplus z_0 \oplus z_1 = k_{28} \oplus z_3 \oplus k_{44} \oplus z_2 \oplus k_{47} \oplus z_2 \oplus z_3 \oplus 1$$

ou

$$b_5 \oplus b_8 \oplus b_{11} \oplus y_{12} \oplus y_{15} = k_{29} \oplus k_{44} \oplus k_{47} \oplus z_1 \oplus 1.$$

Realizando os mesmos procedimentos que no algoritmo SAES, obtém-se 40 equações (com probabilidade $q = 0,5625$) que são divididas em 7 grupos, dependendo da combinação de bits caóticos em cada uma delas. O número de equações em cada grupo é mostrado na Tabela I. Cada grupo individualmente não apresenta 32 equações linearmente independentes, no entanto existem combinações de equações de grupos distintos que permitem obter este número de equações. Por exemplo, a soma módulo 2 de equações dos grupos z_0 com $z_1 \oplus z_2$, obtém-se 48 equações com probabilidade $q_1 = 2q^2 - 2q + 1 = 0,5078$, e somando estas equações com as do grupo $z_0 \oplus z_1 \oplus z_2$, obtém-se equações linearmente independentes suficientes que não dependem dos bits caóticos, cada uma com probabilidade:

$$\begin{aligned} q_2 &= (0,5625)(0,5078) + (1 - 0,5625)(1 - 0,5078) \\ &= 0,5009. \end{aligned} \quad (15)$$

Para o caso em que a probabilidade é $0,5009$ o adversário precisa de $n = 2.667.777$ pares de texto claro e texto cifrado (para achar este valor de n deve-se substituir $q = 0,5009$ em (11) e proceder de forma análoga ao parágrafo após (12)). Outras combinações de equações também podem ser realizadas, mas a probabilidade das mesmas é mais próxima de $0,5$, o que aumenta o valor de n.

B. SAES2

O algoritmo SAES2 é uma versão simplificada do SAES1 em que o bloco de mistura de colunas é eliminado. Devido à eliminação deste blobo, as equações obtidas na etapa de substituição da segunda iteração são modificadas devido à não existência de mistura entre S-box diferentes. Na segunda etapa de iteração obtemos 48 equações, como por exemplo $\hat{b}_5 \oplus \hat{y}_{12} \oplus \hat{y}_{15} = \hat{k}_{29} \oplus \hat{k}_{44} \oplus \hat{k}_{47}$, e conhecendo que $\hat{b}_5 \oplus x_5 = k_5$, podemos combinar ambas equações para obter uma equação da forma (4). Um procedimento similar é realizado para obter 48 equações da forma (4) com probabilidade $0,625$. A presença dos bits caóticos vai dividir as 48 equações em 4 grupos, como mostra a Tabela II.

Uma combinação (soma módulo 2) dos grupos $\{z_0 \oplus z_2\}$ e $\{z_1 \oplus z_3\}$ conduz a 96 equações do grupo $z_0 \oplus z_1 \oplus z_2 \oplus z_3$ com probabilidade $0,53125$, que ao serem combinadas com equações do grupo $z_0 \oplus z_1 \oplus z_2 \oplus z_3$, obtém-se equações que não dependem dos bits caóticos, com probabilidade $0,5078$, precisando o adversário de $n = 35.518$ pares de texto claro e texto cifrado.

C. SAES3

A retirada do bloco mistura de colunas no SAES2 acarreta em uma perda da difusão de bits de S-boxes distintas. Um novo algoritmo, o SAES3, visa compensar esse efeito. Neste os blocos deslocamento de linhas e mistura de colunas são substituídos por um bloco mistura aleatória. Um deslocamento cíclico à direita aleatório é realizado nos bits de saída das 4 S-boxes $(b_0, \cdots b_{15})$ do caminho de dados dependendo

TABELA I
GRUPOS DE EQUAÇÕES DEPENDENDO DOS BITS CAÓTICOS PARA O SAES1

Bits caóticos	Quantidade de equações
z_0	12
z_3	4
$z_1 \oplus z_2$	4
$z_1 \oplus z_3$	8
$z_2 \oplus z_3$	4
$z_0 \oplus z_1 \oplus z_2$	4
$z_0 \oplus z_2 \oplus z_3$	4

TABELA II
GRUPOS DE EQUAÇÕES DEPENDENDO DOS BITS CAÓTICOS PARA O SAES2

Bits caóticos	**Quantidade de equações**
Não	20
$z_0 \oplus z_2$	12
$z_1 \oplus z_3$	8
$z_0 \oplus z_1 \oplus z_2 \oplus z_3$	8

TABELA III
EQUAÇÕES DEPENDENDO DOS BITS CAÓTICOS SAES-3

Deslocamento	Bits caóticos	No. Equações	Probabilidade
01	z_0	4	0,5625
	z_1	4	0,5625
	z_2	4	0,5625
	$z_0 \oplus z_1$	8	0,5625
	$z_1 \oplus z_2$	4	0,5625
	$z_1 \oplus z_3$	4	0,625
	$z_2 \oplus z_3$	4	0,5625
	$z_1 \oplus z_2 \oplus z_3$	4	0,625
	$z_0 \oplus z_2 \oplus z_3$	4	0,625
	$z_0 \oplus z_1 \oplus z_3$	8	0,5625
10	z_1	8	0,625
	z_3	8	0,625
	$z_0 \oplus z_2$	4	0,5625
	$z_0 \oplus z_3$	8	0,5625
	$z_1 \oplus z_2$	8	0,5625
	$z_0 \oplus z_1 \oplus z_2$	8	0,625
	$z_0 \oplus z_1 \oplus z_2 \oplus z_3$	4	0,5625
11	z_0	8	0,5625
	z_2	8	0,5625
	$z_0 \oplus z_1$	4	0,5625
	$z_2 \oplus z_3$	4	0,5625
	$z_0 \oplus z_3$	4	0,5625
	$z_0 \oplus z_1 \oplus z_3$	8	0,5625
	$z_0 \oplus z_2 \oplus z_3$	4	0,5625
	$z_1 \oplus z_2 \oplus z_3$	8	0,625

dos bits caóticos $z_0 z_1$. Pode-se realizar quatro deslocamentos possíveis, $z_0 z_1 = 00$ sem deslocamento, $z_0 z_1 = 01$ deslocamento de um bit, $z_0 z_1 = 10$ deslocamento de dois bits e $z_0 z_1 = 11$ deslocamento de três bits. Cada deslocamento ocorre com a mesma probabilidade, $1/4$, e para cada um deles existem 48 equações possíveis. A Tabela III mostra a quantidade de equações que dependem dos bits caóticos para cada deslocamento, em que equações têm probabilidades $0,625$ ou $0,5625$. Para $z_0 z_1 = 00$ tem-se o algoritmo SAES2.

Como foi realizado anteriormente para o SAES1 e SAES2, pela combinação de equações pode-se obter 32 equações linearmente independentes que não dependem dos bits caóticos. Por exemplo, a combinação dos grupos de z_0 com z_1 na Tabela III, com deslocamento 01 resulta em 16 equações que dependem de $z_1 \oplus z_0$ com probabilidade $0,5078$. Combinando estas 16 equações com as 8 equações que dependem de $z_0 \oplus z_1$ (para o mesmo deslocamento) obtém-se 128 equações, sendo possível extrair a partir destas 32 equações linearmente independentes com probabilidade $0,5009$.

Este procedimento também pode ser realizado para os outros deslocamentos com a determinação de 32 equações linearmente independentes com probabilidade $0,5009$ que não dependem dos bits caóticos. Os procedimentos realizados na Tabela II são similares aos realizados no algoritmo SAES2, sendo que as 32 equações apresentam uma probabilidade $0,5078$. A probabilidade de obter 32 equações linearmente independentes é:

$$q = \frac{1}{4}(0,5078) + \frac{3}{4}(0,5009) = 0,5026. \quad (16)$$

Desta forma o adversário precisa de $n = 319.660$ pares de texto claro e texto cifrado. Outras combinações podem ser realizadas, mas as probabilidades das equações encontradas apresentam um valor menor que $0,5026$.

TABELA IV
RESULTADOS DA CL PARA OS ALGORITMOS PROPOSTOS

Algoritmo	Probabilidade	Pares
SAES	0,5625	545
SAES1	0,5009	2.667.777
SEAS2	0,5078	35.518
SAES3	0,5026	319.660

A Tabela IV apresenta uma comparação dos valores da probabilidade de equações linearmente independentes que são satisfeitas corretamente e do número de pares texto claro e texto cifrado que é necessário para que a CL seja bem sucedida com uma confiabilidade de 95 %. A introdução de bits caóticos leva a um aumento considerável na quantidade de pares de texto claro e texto cifrado necessários em comparação com o algoritmo SAES. O algoritmo SAES1 apresenta o melhor desempenho para este análise, mas também é o algoritmo mais complexo, realizando um maior número de operações. O algoritmo SAES3 apresenta robustez contra a CL significantemente melhor que o SAES2 com uma complexidade similar (a única diferença entre estes é que o deslocamento para o SAES3 depende dos bits caóticos).

V. CONCLUSÕES

Os algoritmos apresentados em [3] foram modificados com estrutura similar para um análise mais simples da CL. Estes algoritmos de cifra de bloco modificados (com a introdução de bits caóticos na etapa de substituição de bytes e na geração de sub-chaves) aumentam o valor de pares de texto claro e texto cifrado necessários para achar de bits de chaves com certa confiabilidade. Como trabalhos futuros, pode-se realizar uma análise similar para outras técnicas de criptoanálise, como a cripotanálise diferencial, amplamente utilizada na literatura.

REFERÊNCIAS

[1] J. Daemen and V. Rijmen, *The design of Rijndael: AES-The Advanced Encryption Standard*. Springer-Verlog, 2002.

[2] Federal Information Processing Standards (FIPS), "Advanced Encryption Standard (AES)," NIST, Tech. Rep. FIPS 197, November 2001. [Online]. Available: http://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf

[3] J. A. P. Artiles, D. P. B. Chaves, , and C. Pimentel, "Cifragem de imagens usando cifras de bloco e sequências caóticas," *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, São Pedro*, pp. 1–5, Setembro 2017.

[4] M. Matsui and A. Yamagishi, "A new method for known plaintext attack of FEAL cipher," *Workshop on the Theory and Application of of Cryptographic Techniques*, vol. 658, pp. 81–91, 1992.

[5] M. Matsui, "Linear cryptanalysis method for DES cipher," *Advances in Cryptology EUROCRYPT93*, vol. 765, pp. 386–397, 1994.

[6] M. Mohammad, E. Schaefer, and S. Wedig, "A simplified Rijndael algorithm and its linear and differential cryptanalysis," *Cryptologia*, vol. 27, no. 2, pp. 148–177, 2003.

[7] D. Mansoori and H. Khaleghi, "Linear cryptanalysis on second round simplified AES," *The 8th International Conference on Advanced Communication Technology*, pp. 1210–1214, 2006.