MA-53G Intr. a la Criptografía **Semestre Primavera, 1996**

Cátedra No. 7 : 27 de Agosto

Transcriptor: Gregorio González

1 Criptoanálisis Diferenciable

1.1 El M'etodo de Criptoan'alisis Diferenciable

El método de Criptoanálisis Diferenciable fue presentado por Eli Biham y Adi Shamir en 1990.[1] Este método permite realizar un ataque de texto plano elegido que resulta más eficiente que un ataque de fuerza bruta contra el sistema DES.

El criptoanálisis diferenciable se basa en el análisis de la propagación de diferencias entre dos textos planos a través de las funciones del criptosistema cuando se usa la misma clave. Básicamente, se toma un par de textos planos, posiblemente al azar, pero con una cierta "distancia"elegida. Según la clave usada, la diferencia en la entrada llevará con mayor o menor probabilidad a ciertas diferencias en la salida (texto cifrado). El criptoanálisis diferenciable aprovecha el hecho que estas distribuciones en la diferencia de salida no son uniformes. Al obtener la diferencia entre los textos cifrados, se puede descartar aquellas claves con las cuales esta diferencia tiene probabilidad cero. A medida que se analizan más pares de textos, se irán descartando claves hasta que una clave aparecerá como la más probable. Esta es la clave correcta.

El método sirve para atacar un criptosistema o una parte de él que tenga la estructura que se muestra en la Figura 1.1.

Se define la diferencia según el criptosistema a atacar. Para el criptosistema DES la diferencia adecuada es el XOR. La condición que debe cumplir esta operación (y que se usa para elegirla) es la siguiente:

$$\triangle B = B - B' = f(X,k) - f(X',k) = \triangle X.$$

[1] E. Biham y A.Shamir, 1.- *"Differential Cryptanalysis of DES-like Cryptosystems"* , Advances in Cryptology–CRYPTO '90 Proceedings, Springer Verlag, 1991, pp. 2-21; 2.- *"Differential Cryptanalysis of DES-like Cryptosystems"* , Journal of Cryptology, v.4, n. 1, 1991, pp 3-72; 3.- *"Differential Cryptanalysis of FEAL and N-Hash"* , Advances in Cryptology–EUROCRYPT '91 Proceedings, Springer-Verlag, 1991, pp. 1-16.

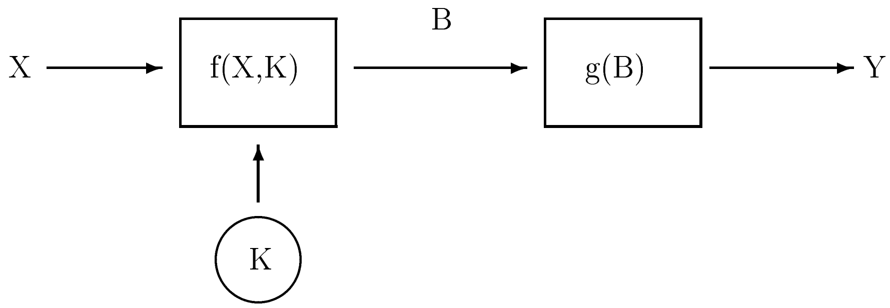

Figura 1: Estructura de un criptosistema atacable mediante criptoan'alisis diferenciable.

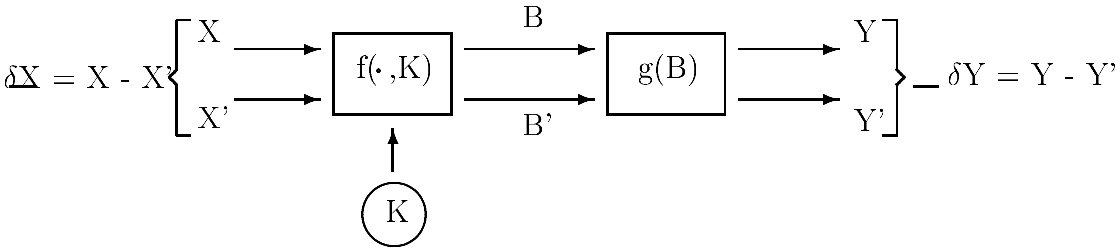

Figura 2: Propagación de las diferencias a través del criptosistema.

Es decir, que la diferencia no se altere con f, independientemente de la clave (ver Figura 1.1). Se define entonces:

$$IN(\delta,\tau) = \{B : g(B) - g(B+\delta) = \tau\}.$$

En el caso de DES:

$$IN(\delta,\tau) = \{B : g(B) \oplus g(B \oplus \delta) = \tau\},$$

Y constru'imos:

$$K(X,\delta,\tau) = \{k : f(X,k) \in IN(\delta,\tau)\}.$$

Notemos que si encriptamos $X,\ X' = X - \delta$ y obtenemos $Y,\ Y'$ tal que $\triangle Y = \tau$ entonces necesariamente $k \in K(X,\delta,\tau)$.

Lo que se desea es que el cardinal de $K(X,\delta,\tau)$ sea mucho menor que la cantidad de claves posibles. Este conjunto básicamente representa aquellas claves para las cuales las diferencias obtenidas tienen probabilidad mayor que cero.

El ataque procede de la siguiente forma:

1. Crear contadores para todas las claves posibles
2. Elegir X, δ, determinar X', Y, Y', y $\tau = \triangle Y$
3. Calcular $IN(\delta, \tau)$
4. Incrementar los contadores para las claves en $K(X, \delta, \tau)$

El contador de la clave correcta se incrementa en cada paso. Si para una clave dada, alguna de las diferencias obtenidas con un par de textos planos tiene probabilidad cero de ocurrir, su contador no se incrementa en ese paso. El proceso desde el paso 2 al 4 se repite hasta que sólo queda una clave con el contador igual al número de pares de texto analizados. Dicha clave es la correcta.

1.2 Problemas del Criptoan'alisis Diferenciable

A pesar de su aparente simplicidad, este ataque puede resultar impr'actico en varios casos. Primero que nada, hay una probabilidad bastante baja de tener 'exito hasta que se pasa de cierto umbral, es decir, mientras no se verifiquen suficientes pares de textos es imposible distinguir la clave correcta de las dem'as que no se han descartado. Adem'as, este ataque es costoso desde el punto de vista del almacenamiento de datos, pues requiere contadores para las 2^{48} claves posibles (en el caso de DES). Otro problema es el n'umero de textos planos necesario para descartar todas las claves menos una. Se puede probar que un criptosistema es inmune al criptoan'alisis diferenciable demostrando que requiere analizar m'as textos planos que los que es posible generar con el tamaño de bloque utilizado.

1.3 Un Ejemplo: Criptoan'alisis de 3-DES

Ignoraremos las permutaciones de entrada y salida pues son f'aciles de deshacer (ya que son de dominio p'ublico, i.e. conocidas). Vamos a tratar de descubrir la tercera subclave, ya que la clave completa se puede obtener a partir de ella ciclando sobre las 2^8 posibilidades (ya que esto no es demasiado costoso). Recordemos que:

$$
\begin{aligned}
L_i &= R_{i-1} \\
R_i &= L_{i-1} \oplus f(R_{i-1}, K_i) \quad i = 1, 2, 3.
\end{aligned}
$$

Donde la entrada es L_0R_0 y la salida es L_3R_3. Luego:

$$\begin{aligned} R_3 &= L_2 \oplus f(R_2, K_3) \\ &= R_1 \oplus f(R_2, K_3) \\ &= L_0 \oplus f(R_0, K_1) \oplus f(R_2, K_3). \end{aligned}$$

Sea entonces $L_3'R_3'$ otra salida, correspondiente a la entrada $L_0'R_0'$. Sigue que:

$$\begin{aligned} \triangle R_3 &= R_3 \oplus R_3' \\ &= \triangle L_0 \oplus f(R_0, K_1) \oplus f(R_0', K_1) \oplus f(R_2, K_3) \oplus f(R_2', K_3) \end{aligned}$$

El ataque es de texto plano elegido, as'i que en particular podemos tomar las entradas de forma que $R_0 = R_0'$. Luego:

$$\begin{aligned} \triangle R_3 &= \triangle L_0 \oplus f(R_2, K_3) \oplus f(R_2', K_3) \\ \triangle R_3 \oplus \triangle L_0 &= f(R_2, K_3) \oplus f(R_2', K_3). \end{aligned}$$

Recordemos que en DES, la estructura de f es la que se muestra en la Figura 1.3: Sigue que:

$$\begin{aligned} \triangle R_3 \oplus \triangle L_0 &= P(C) \oplus P(C') \\ &= P(C \oplus C'), \end{aligned}$$

de donde se ve que:

$$\triangle C = P^{-1}(\triangle R_3 \oplus \triangle L_0).$$

Luego $\triangle C_i,\ i = 1, \ldots, 8$ son conocidos. Veamos que $\triangle B_i,\ i = 1, \ldots, 8$ tambi'en son conocidos.

$$E(R_2) \oplus K_3 = B \quad y \quad E(R_2') \oplus K_3 = B'$$

luego

$$\triangle B = E(R_2) \oplus E(R_2').$$

Es decir, $\triangle B$ no depende de la clave.

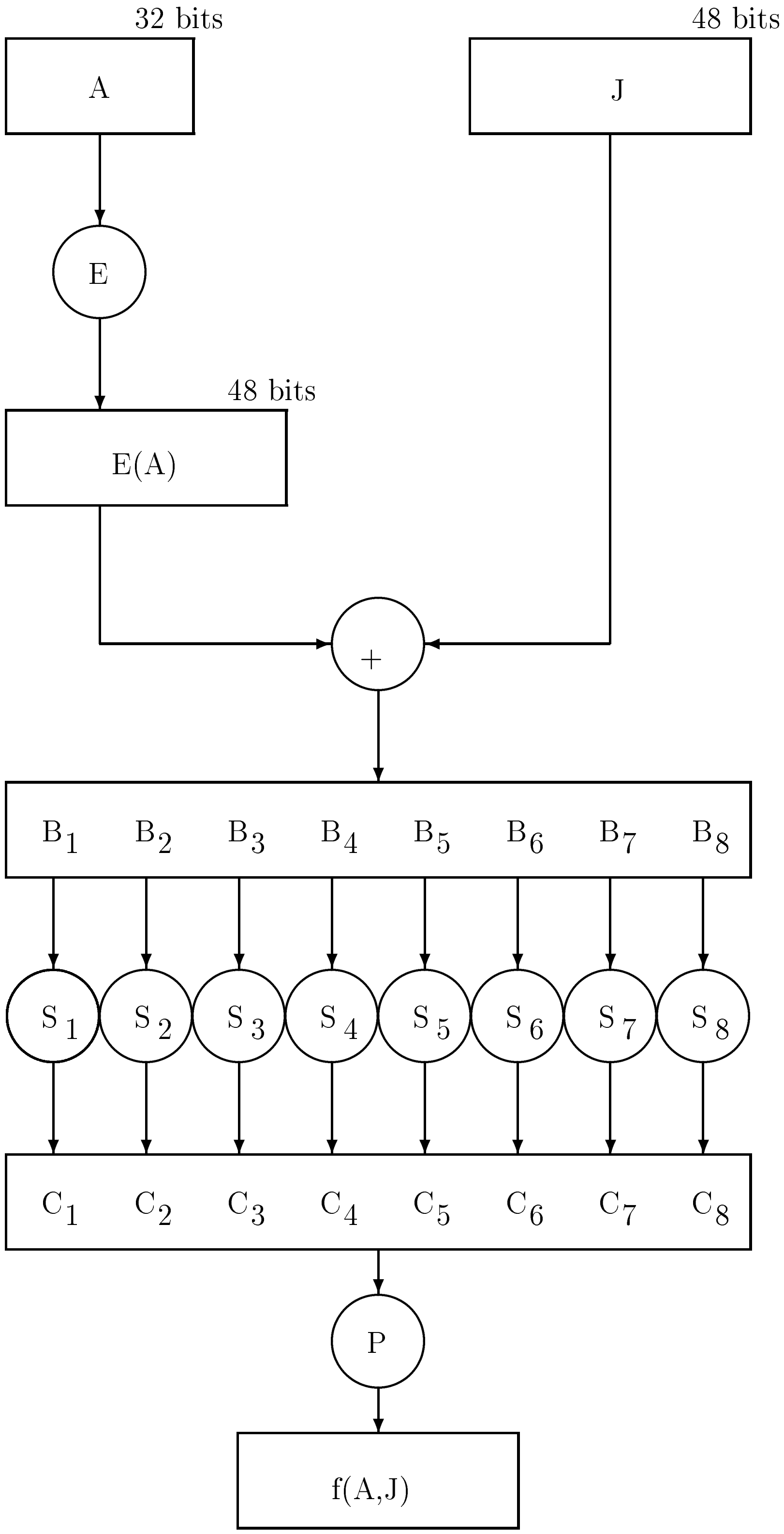

Figura 3: Estructura de f en el criptosistema DES.

Como $L_3 = R_2,\ L'_3 = R'_2$, se tiene que:

$$\triangle B = E(L_3) \oplus E(L'_3)$$

Luego $\triangle B_i,\ i = 1, \ldots, 8$ se pueden obtener a partir de la informaci'on que poseemos. Con esto podemos montar un ataque de criptoan'alisis diferenciable contra cada una de las 8 cajas-S. Con ello obtenemos 6 bits de la clave K_3 por cada caja, en total 48 bits. Los restantes 8 se obtienen por fuerza bruta sin mayor problema (ya que son $2^8 = 256$ posibles combinaciones).

1.4 Criptoan'alisis Diferenciable y DES

El criptoan'alisis diferenciable funciona contra DES en cualquiera de sus modos de operaci'on, ECB, CBC, CFB y OFB con la misma complejidad computacional, es decir, ninguno de estos modos presenta ventajas respecto a los otros en cuanto a resistencia contra este ataque.

En la siguiente tabla presentamos un resumen de los mejores ataques diferenciales contra DES con distintos números de pasos [2] . La primera columna es el número de pasos de la implementaci'on. Las siguientes dos son el número de textos elegidos o textos conocidos que deben ser examinados para efectuar el ataque (el criptoanálisis diferenciable puede adaptarse a una forma con texto plano conocido en vez de elegido de manera relativamente simple). La cuarta columna es el número de estos textos que fueron analizados realmente en estos ataques más exitosos.

Número de Pasos	Textos Elegidos	Textos Conocidos	Textos Analizados
8	2^{14}	2^{38}	4
9	2^{24}	2^{44}	2
10	2^{24}	2^{43}	2^{14}
11	2^{31}	2^{47}	2
12	2^{31}	2^{47}	2^{21}
13	2^{39}	2^{52}	2
14	2^{39}	2^{51}	2^{29}
15	2^{47}	2^{56}	2^{7}
16	2^{47}	2^{55}	2^{36}

La resistencia de DES se puede mejorar aumentando el número de pasos, como se puede apreciar en la tabla. Un ataque de criptoanálisis diferenciable de texto elegido

[2] E. Biham y A. Shamir, *"Differential Cryptanalysis of the Data Encryption Standard"*, Springer-Verlag, 1993.

contra un DES con 17 o 18 pasos toma aproximadamente el mismo tiempo que un ataque de fuerza bruta. [3] Con 19 pasos o más, el criptoanálisis diferenciable se hace imposible porque requiere m'as textos planos elegidos que los que se pueden generar con el tamaño de bloque que utiliza DES (más de 2^{64} textos, donde los textos son de 64 bits).

Este ataque depende mucho de la estructura de las cajas-S. Estas est'an optimizadas contra el criptoanálisis diferenciable, es decir, están constru'idas de manera de hacer que el ataque sea lo más difícil posible.

¿Por qu'e DES es tan resistente al criptoan'alisis diferenciable? ¿Por qu'e est'an optimizadas las cajas-S para hacer que este ataque sea lo m'as dif'icil que sea posible? ¿Por qu'e tiene la cantidad justa de pasos para que sea ineficiente, pero ni uno m'as? La respuesta es porque los diseñadores conoc'ian este ataque. Don Coppersmith, de la IBM, escribi'o: [4]

> El diseño aprovechaba ciertas t'ecnicas de criptoan'alisis, principalmente la t'ecnica del "criptoan'alisis diferenciable", que no eran conocidas en la literatura publicada. Tras discusiones con NSA, se decidi'o que entregar las consideraciones del diseño revelar'ia la t'ecnica del criptoan'alisis diferenciable, una t'ecnica poderosa que puede ser utilizada contra varios criptosistemas. Esto a su vez debilitar'ia la ventaja competitiva que los Estados Unidos disfrutaba sobre el resto de los pa'ises en materia de criptograf'ia.

Adi Shamir desafi'o a Coppersmith a decir que no hab'ia encontrado ning'un ataque m'as fuerte contra DES desde entonces, pero Coppersmith no ha respondido. Sabiendo que la NSA conoc'ia el criptoan'alisis diferenciable m'as de 15 años antes que el mundo acad'emico, existe cierta preocupaci'on en cuanto a qu'e m'etodos habr'an desarrollado en esos 15 años de ventaja, durante los cuales ciertamente sus investigadores no deben haber estado descansando precisamente.

---

[3] B. Schneier, *"Applied Cryptography", Wiley, 1996, p. 289.*

[4] D. Coppersmith, 1.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, Technical Report RC 18613, IBM T.J. Watson Center, Dic. 1992; 2.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, IBM Journal of Research and Development, v. 38, n. 3, Mayo 1994, pp. 243-250

1.5 Ideas Alternativas

Existen otros m'etodos con ciertas similitudes al criptoan'alisis diferenciable que se han utilizado tambi'en contra DES, con mayor o menor 'exito. Entre ellas tenemos el Criptoan'alisis de Clave Relacionada (Related Key Cryptanalysis) y el Criptoan'alisis Lineal.

Criptoan'alisis de Clave Relacionada

La clave de DES se rota cierta cantidad de bits en cada paso: 2 tras cada paso salvo los pasos 1,2,9 y 16, en los cuales rota uno s'olo. ¿Por qu'e?

El criptoan'alisis de clave relacionada es similar al criptoan'alisis diferencial, pero examina las diferencias entre claves. El ataque en s'i es muy distinto: El criptoanalista toma deliberadamente cierta relaci'on entre dos claves, pero no conoce las claves mismas. Se encriptan datos con ambas claves. En la versi'on de texto plano conocido, se dispone de el texto plano y el texto cifrado con ambas claves. En la versi'on con texto plano elegido, el criptoanalista elige el texto plano a encriptar con ambas claves (pero no las claves).

Un DES modificado, en el cual la clave se rota exactamente dos bits en cada paso, es menos seguro. Utilizando criptoan'alisis de clave relacionada se puede quebrar esta variante usando 2^{17} textos elegidos de clave elegida o 2^{33} textos conocidos de clave elegida. [5].

Este ataque es bastante impr'actico, pero es interesante por tres motivos:

- Uno, es el primer ataque criptoanal'itico contra el sistema de generaci'on de subclaves de DES.
- Dos, este ataque es independiente del n'umero de pasos del criptosistema, es igualmente efectivo contra DES con 16, 32 o 1000 pasos.
- Y tres, DES es inmune a este ataque. La variaci'on en el n'umero de bits que se rotan frustra el ataque de clave relacionada.

[5]E. Biham, 1.- *"New Types of Cryptanalitic Attacks Using Related Keys"*, Technical Report #753, Computer Science Department, Technion–Israel Institute of Technology, Sep. 1992.; 2.- *"New Types of Cryptanalitic Attacks Using Related Keys"*, Journal of Cryptology, v. 7, n. 4, 1994, pp. 229-246.

Criptoan'alisis Lineal

El Criptoan'alisis Lineal es otro tipo de ataque, inventado por Mitsuru Matsui. [6] Este ataque usa aproximaciones lineales para describir la acci'on de un criptosistema de bloque (en este caso DES). Esto significa que si tomamos el "o exclusivo" (XOR) de algunos de los bits del texto plano, el XOR de algunos del texto cifrado, y tomamos el XOR de ambas cosas, obtendremos un bit que ser'ia el XOR de algunos bits de la clave. Esto es una aproximaci'on lineal, y coincidir'a con la realidad con cierta probabilidad p. Si $p \neq \frac{1}{2}$, este sesgo puede ser aprovechado. Utilizando textos planos y sus textos cifrados se puede ir adivinando los valores de los bits de la clave. Mientras m'as datos se tiene, m'as confiable es la estimaci'on. Mientras mayor es el sesgo (i.e. mientras mayor es $\mid p - \frac{1}{2} \mid$) mayor es la tasa de 'exitos con la misma cantidad de datos conocidos.

Si las aproximaciones lineales fuesen insesgadas, entonces resultar'ian en 32 de 64 entradas posibles (a cada caja-S. Este ataque est'a dirigido a dicha parte del algoritmo DES). Si hay una caja con sesgo suficientemente grande, el ataque puede funcionar. La caja-S m'as sesgada es S_5. De hecho, el segundo bit de entrada es igual al XOR de los cuatro de la salida s'olo en 12 casos. Esto se traduce en una probabilidad de $\frac{3}{16}$, i.e. un sesgo de $\frac{5}{16}$, y es el sesgo m'as extremo de todas las cajas-S. (Seg'un el libro de Schneier, Shamir not'o este sesgo, pero no pudo encontrar una manera de explotarlo).

Contra un DES de 16 pasos, este ataque puede encontrar la clave con un promedio de 2^{43} textos conocidos. Una implementaci'on en software de este ataque quebr'o una clave DES en 50 d'ias usando 12 estaciones HP9735 [7] . Es el ataque m'as efectivo contra DES que se conoce a la fecha.

El criptoan'alisis lineal depende fuertemente de la estructura de las cajas-S, y las cajas-S del DES no est'an optimizadas contra este ataque. De hecho, el ordenamiento de las cajas-S elegido para DES est'a entre el 9% y el 16% que ofrecen menor protecci'on contra este ataque. [8] De acuerdo a Don Coppersmith [9], la resistencia contra el criptoan'alisis lineal "no fu'e parte del criterio de diseño del DES". O bien

[6]M. Matsui, *"Linear Cryptanalysis Method for DES Cypher"*, Advances in Cryprology–EUROCRYPT '93 Proceedings, Springer-Verlag, 1994, pp. 386-397.

[7]M. Matsui, *"The First Experimental Cryptanalysis of the Data Encryption Standard"*, Advances in Cryptology–CRYPTO '94 Proceedings, Springer-Verlag, 1994, pp. 1-11.

[8]B. Schneier, *"Applied Cryptography"*, Segunda Edici'on, Wiley, 1996, p. 293.

[9]D. Coppersmith, 1.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, Technical Report RC 18613, IBM T.J. Watson Center, Dic. 1992; 2.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, IBM Journal of Research and Development, v. 38, n. 3, Mayo 1994, pp. 243-250

no conoc'ian el criptoan'alisis lineal, o conoc'ian otro ataque mucho m'as poderoso, haciendo que la resistencia contra 'este tuviera prioridad sobre la resistencia contra el criptoan'alisis lineal.

EL criptoan'alisis lineal es m'as reciente que el diferencial, y pueden haber mayores mejoras a su rendimiento en los años venideros. Algunas ideas se han propuesto, [10] pero no es claro que puedan ser utilizadas efectivamente contra un DES completo (con sus 16 pasos). Sin embargo, funcionan bastante bien contra variantes de menos pasos.

1.6 Direcciones a Explorar

Se han hecho algunos trabajos intentando extender el concepto de criptoan'alisis diferenciable a derivadas de orden superior [11] Lars Knudsen usa "derivadas parciales" para atacar DES de 6 pasos; requiere 32 textos elegidos y 20.000 encriptaciones. [12] Todav'ia es muy pronto para saber si estas extensiones facilitar'an quebrar un DES de 16 pasos. Otra v'ia de ataque es el Criptoan'alisis Lineal-Diferenciable, que combina ambas t'ecnicas antes presentadas. Susan Langford y Hellman tienen un ataque contra 8-DES que descubre 10 bits de la clave con una probabilidad de 'exito de un 80% con 512 textos planos elegidos y un 95% de probabilidad de 'exito con 768 textos planos elegidos, [13]. Tras el ataque se requiere una b'usqueda de fuerza bruta de los bits restantes ($2^4$6 claves posibles). Mientras que este ataque consume un tiempo similar a ataques previos, requiere mucho menos texto plano. Lamentablemente parece que no se extiende con facilidad a DES con m'as pasos. Sin embargo, este ataque es nuevo y el trabajo sigue. Es posible que haya un descubrimiento grande en alg'un momento durante los pr'oximos años. Tal vez mezclando ambos enfoques ... ¿Qui'en sabe?

[10] Ver referencias [1270],[811] del libro de Schneier.

[11] Ver referencias [702,161,927,858,860] del libro de Schneier.

[12] L.R. Knudsen, *"Applications of Higher Order Differentials and Partial Differentials"*, K.U. Leuven Workshop on Cryptographic Algorithms, Springer-Verlag, 1995.

[13] S.K. Langford y M.E. Hellman, *"Cryptanalysis of DES"* presentado en la conferencia 1994 RSA Data Security, Redwood Shores, California, 12-14 de Enero 1994

1.7 El Criterio de Diseño de las Cajas-S

Tras la salida a la luz p'ublica del criptoan'alisis diferenciable, IBM public'o el verdadero criterio de diseño de las cajas-S. [14] El criterio para las cajas S es:

- Cada caja-S tiene 6 bits de entrada y 4 de salida (pues era el mayor tamaño que se pod'ia implementar en un s'olo chip con la tecnolog'ia de 1974).
- Ning'un bit de salida de una caja-S podr'ia estar demasiado cerca de ser una funci'on lineal de los bits de entrada.
- Si se fijan el bit de m'as a la izquierda y el de m'as a la derecha de la entrada de una caja-S, y se var'ia el resto, cada salida posible de 4 bits se obtiene exactamente una vez.
- Si dos entradas a una caja-S difieren exactamente en un bit, las salidas deben ser distintas en al menos dos bits.
- Si dos entradas a una caja-S difieren en sus primeros dos bits y son id'enticas en los 'ultimos dos, las dos salidas no pueden ser iguales.
- Para cualquier diferencia no nula de 6 bits entre las entradas, no m'as de 8 de los 32 pares de entradas que tienen esa misma diferencia pueden dar la misma diferencia en la salida.
- Un criterio similar al anterior, pero para el caso de tres cajas-S activas (i.e. para cualquier par de diferencias no nulas, no m'as de cierta cantidad de los tripletes de entradas puede dar el mismo par de diferencias de salida).

Claramente el sexto criterio apunta a defender las cajas-S contra el criptoan'alisis diferenciable, agrandando los conjuntos de claves posibles para cada par de diferencias de entrada y salida.

El art'iculo explica estos criterios y sus razones de ser. Generar cajas-S es sumamente f'acil hoy en d'ia, pero era una tarea complicada en aquella 'epoca. Se dice que tuvieron computadores corriendo durante meses para fabricar las famosas cajas-S.

[14] D. Coppersmith, 1.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, Technical Report RC 18613, IBM T.J. Watson Center, Dic. 1992; 2.- *"The Data Encryption Standard (DES) and Its Strength against Attacks"*, IBM Journal of Research and Development, v. 38, n. 3, Mayo 1994, pp. 243-250