Análisis del efecto Avalancha sobre XTEA

Ignacio Ballesteros González

11 de diciembre de 2017

Índice

Resumen

Para la realización de la segunda práctica de la asignatura de *Seguridad de las Tecnologías de la Información* se ha realizado el análisis del efecto *avalancha* sobre el algoritmo de cifrado *XTEA*.

1. Introducción

El algoritmo `XTEA` fue desarrollado...

1.1. Herramientas utilizadas

Para la posible replicación de este análisis, se detallan las herramientas utilizadas en el desarrollo del mismo:

Python 2 y 3 Plataforma de desarrollo.

Emacs Entorno de desarrollo para programación en *Python*. *Emacs 25.3.2*

xtea Implementación del algoritmo de cifrado *xtea*. *xtea (0.4.0)*

scipy Librería *Python* para cálculos estadísticos.

LaTeX Realización de figuras y de este mismo documento.

2. Estructura del análisis

Para la realización del análisis se ha seguido el siguiente procedimiento:

1. Generamos una muestra de estudio (*pseudo-aleatoria*) compuesta de una clave **(K)**, un vector de inicialización **(VI)** y un texto a cifrar **(T)**.
2. Realizamos un cifrado con *XTEA* para obtener la cadena cifrada **C**.
3. Para comprobar los distintos efectos avalancha modificamos un bit en la clave, el vector inicial y el texto **(K', VI', T')**.
4. Para cada modificación, se aplica de nuevo el cifrado con las combinaciones de clave, vector inicial y texto:
   - (K, VI, T')
   - (K, VI', T)
   - (K', VI, T)
5. Por cada resultado del cifrado, se calcula la distancia de Hamming con respecto a la cadena cifrada original **C**.
6. Aplicamos un análisis estadístico sobre los histogramas generados.

3. Datos recogidos

El cifrado *XTEA* trabaja con bloques de 128 *bits*. Por este motivo, hemos elegido una generación de cadenas de entrada de 128 *bits* (16 *bytes*) para estudiar el efecto avalancha.

Durante sucesivas iteraciones con distintos tamaños de muestra, cuando nos acercamos a las 500 cadenas de muestra, el error de la varianza desciende por debajo del 5 %.

Dado que el cifrado tiene 3 parámetros de entrada (clave, vector inicial, texto a cifrar), el estudio del efecto avalancha lo tendremos que hacer sobre variaciones en estos 3 parámetros.

3.1. Variación del texto de entrada

Los datos recogidos variando en un *bit* el texto a cifrar (y manteniendo clave y vector inicial) son los siguientes:

Muestra	Media	Mediana	Desviación	Variance	Asimetría	Curtosis
500.00	1.00	1.00	0	0	0	-3.00

Cuadro 1: Tabla de medidas estadísticas

Es muy llamativo cómo nos encontramos con unos datos muy distintos a los posibles esperados cuando se produce un efecto avalancha. Que no haya desviación, ni varianza y la distribución sea simétrica, implica que todos los valores se encuentran en 1. Es decir, cuando se modifica un bit en la entrada, solo se modifica un bit en la salida. En la figura 1 se puede observar cómo la distribución del histograma dista mucho de ser binomial. De hecho, en todos los casos comprobados, la variación es un único *bit*.

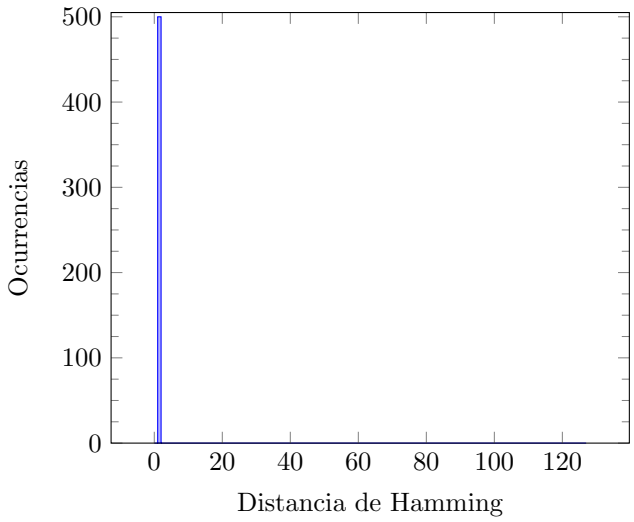

Figura 1: Análisis sobre una población de 500

3.2. Variación en la clave

Con una variación de un **bit** en la clave de cifrado y manteniendo el texto de entrada original obtenemos los siguientes datos:

Muestra	Media	Mediana	Desviación	Variance	Asimetría	Curtosis
500.00	64.27	64.00	5.55	30.83	0.08	0.15

Cuadro 2: Tabla de medidas estadísticas

Y el con el histograma en la figura 2.

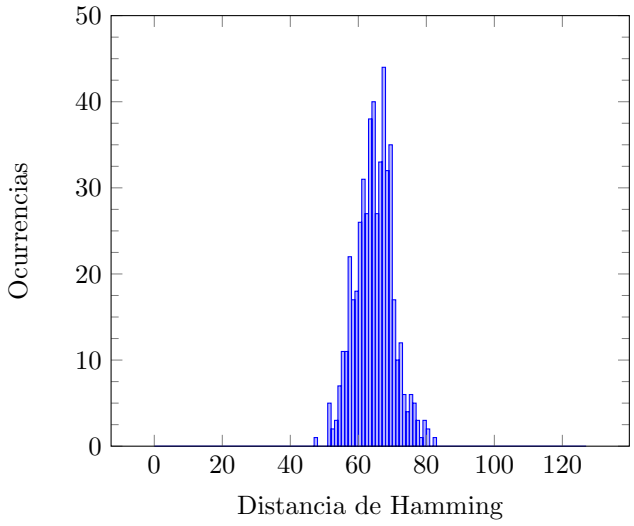

Figura 2: Análisis sobre una población de 500

Si nos fijamos en los errores obtenidos respecto a una distribución binomial ($p = 128, n = 0{,}5$), el error en la varianza es del 0,036 ($3'6\,\%$)

3.3. Variación en el vector inicial

Con una variación de un **bit** en el vector de inicialización y manteniendo el texto de entrada original obtenemos los siguientes datos:

+

Muestra	Media	Mediana	Desviación	Variance	Asimetría	Curtosis
500.00	64.31	64.00	5.66	32.04	-0.09	0.34

Cuadro 3: Tabla de medidas estadísticas

Y el con el histograma en la figura 2.

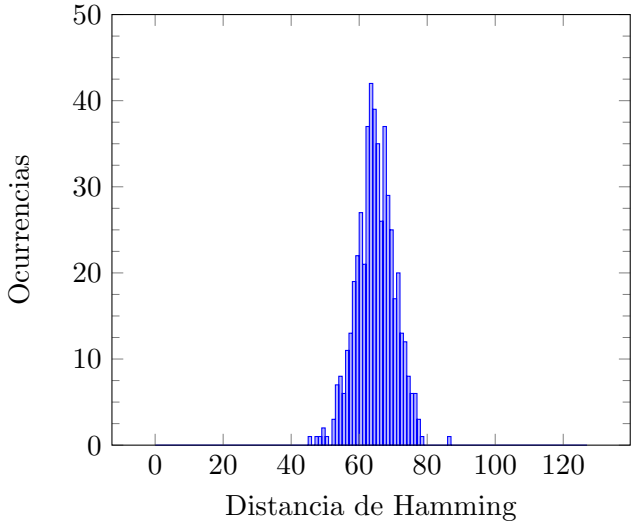

Figura 3: Análisis sobre una población de 500

En el caso del vector inicial, el error en la varianza obtenido es de 0,1 %. Una aproximación muy buena para la binomial.

4. Valoración de los resultados

Cuando se trata de comprobar el efecto *avalancha* sobre una función de cifrado se espera que cuando se cambie una pequeña parte de la cadena de entrada (por ejemplo, un *bit*), aproximadamente la mitad de los *bits* de la cadena de salida cambien también.

Que solo cambie aproximadamente la mitad de los bits quiere decir que la *distancia de Hamming* asociada a la modificación de un solo *bit* sea de aproximadamente la **mitad** de la cadena. En en el caso de la función *XTEA* las cadenas son de *128 bits*, por lo que la *distancia de Hamming* asociada en un cambio de *bit* se concentraría cerca de 64. Este análisis se ha hecho con esta longitud de cadena porque el cifrado toma bloques de *128 bits*.

El resultado arrojado por la variación de un *bit* en la cadena de texto de entrada provoca que solo se cambie un *bit* en la cadena de salida. Esto indica que no hay un buen efecto avalancha en el cifrado.

Sin embargo, si nos centramos en el cambio de un *bit* en la clave o en el vector inicial, observamos que el efecto avalancha se produce satisfactoriamente.

Con estos resultados, podemos decir que el cifrado *XTEA* posee parcialmente la propiedad del efecto avalancha.