Análisis del efecto Avalancha sobre RipeMD320

Ignacio Ballesteros González

21 de mayo de 2017

Índice

Resumen

Para la realización de la práctica primera de la asignatura de *Seguridad de las Tecnologías de la Información* se ha realizado el análisis del efecto *avalancha* sobre el algoritmo de *hash RipeMD320*.

1. Introducción

El algoritmo de *hash RipeMD* se desarrolló en la universidad *Katholieke Universiteit Leuven* y se publicó su primera versión en 1996. El algoritmo se mejoró en siguientes versiones (*160, 256*) para evitar colisiones hasta llegar a la versión *RipeMD320* que será la que aquí se analice.

1.1. Herramientas utilizadas

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

Java 1.8 Plataforma de desarrollo.

Eclipse Entorno de desarrollo para programación en *Java. Neon 3 Release (4.6.3)*

Bouncy Castle Librería con funciones criptográficas, entre ellas incluida la necesaria para *RipeMD320*.

Apache Commons Math Librería *Java* 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. Comenzamos el análisis con una cadena (**A**) *pseudo-aleatoria uniformemente distribuida* de 320 *bits*.
2. Calculamos el *hash RipeMD320* sobre la cadena **A**.
3. Variamo un *bit* aleatorio de la cadena **A**. Esta nueva cadena será la cadena **B**.
4. Calculamos el *hash RipeMD320* sobre la cadena **B**.
5. Calculamos la *distancia de Hamming* entre el *hash* de la cadena **A** y **B**.
6. Generamos una nueva cadena origen y repetimos estos pasos n veces, dependiendo del tamaño de la muestra que queramos analizar.
7. Aplicamos un análisis sobre las *distancias de Hamming* obtenidas.

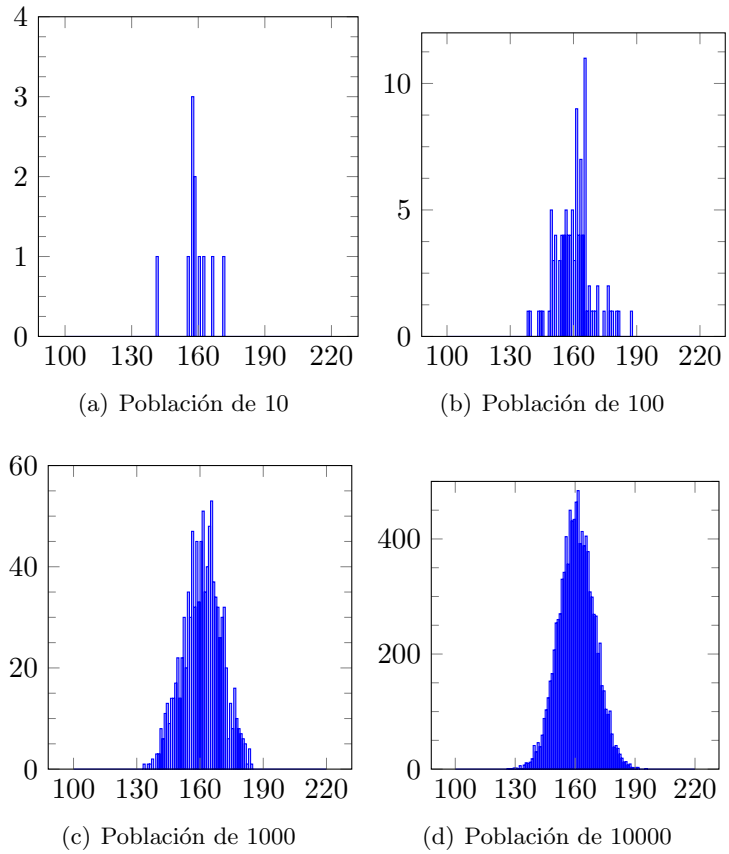
3. Datos recogidos

Los análisis se ha realizado según el tamaño de la muestra. Los datos estadísticos asociados se presentan en la siguiente tabla:

Población	Media	Moda	Mediana	Desviación	Asimetría	Curtosis
10	158.36	157	158	7.43	-0.82	3.05
100	159.93	165	161	8.68	0.29	0.85
1000	160.64	165	161	9.10	-0.11	-0.23
10000	160.29	161	160	8.99	0.02	0.04
100000	160.23	161	160	8.97	0.00	0.00
1000000	160.24	161	160	8.97	0.00	-0.01
100000000	160.03	160	160	8.97	0.00	0.00

Cuadro 1: Tabla de medidas estadísticas

En orden creciente según la población, visualización de los datos en un histograma. En el eje y se representan el número de ocurrencias de la *distancia de Hamming* dada (eje x).



(a) Población de 10 (b) Población de 100

(c) Población de 1000 (d) Población de 10000

Figura 1: Histogramas de población pequeña

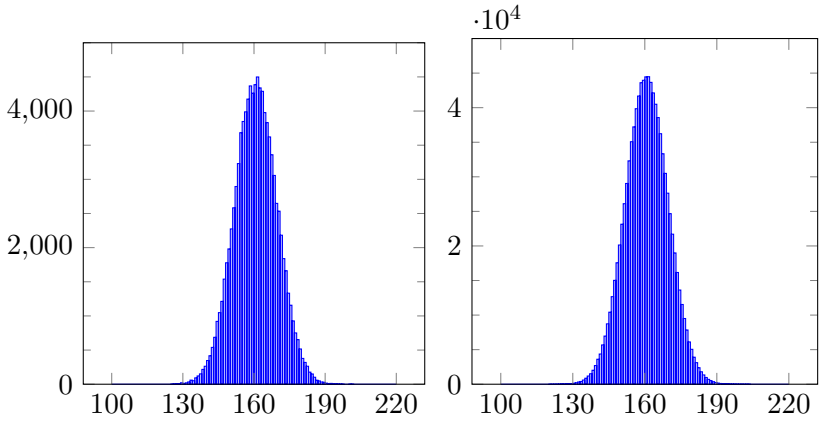
4. Valoración de los resultados

Cuando se trata de comprobar el efecto *avalancha* sobre una función *hash* 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 *RipeMD320* las cadenas son de *320 bits*, por lo que la *distancia de Hamming* asociada en un cambio de *bit* se concentraría cerca de 160.

Como podemos comprobar según los datos estadísticos del *Cuadro 1*, la media, mediana y moda tienden a 160. Por otro lado, cuanto más simétrica sea la distribución, más *uniformemente distribuida* será y más se aproximará al comportamiento de un *Oráculo Aleatorio*.

En base a los datos obtenidos de simetría cercana a 0 (a falta de cifras significativas), una curtosis que indica una fuerte proximidad a la mediana, la cercanía entre *mediana*, *media* y *moda*; **se determina que la función de *hash RipeMD320* exhibe un fuerte *efecto avalancha***.



(a) Población de 100000 (b) Población de 1000000

Figura 2: Muestras de población mediana

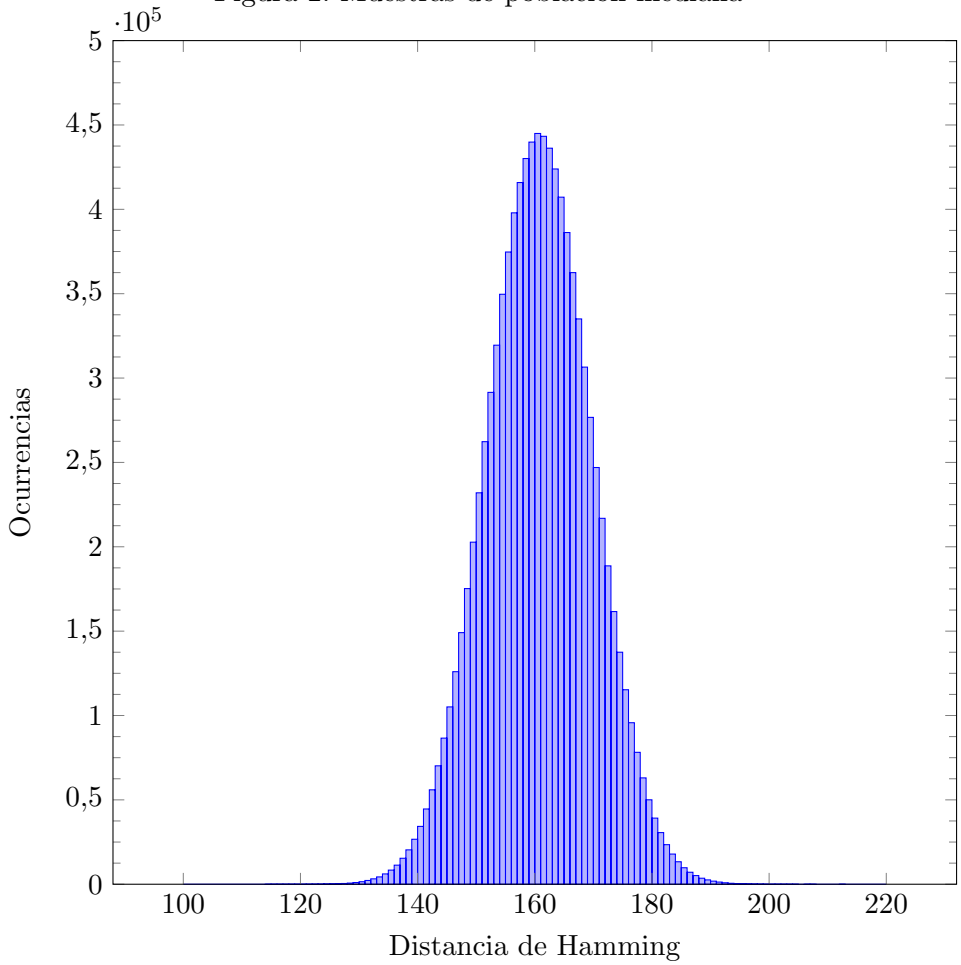


Figura 3: Muestra de gran tamaño