Tipo de artículo: Artículo original

Análisis estadístico de las sucesiones de salida del Generador de Secuencias Pseudoaleatorias: Fortuna

Statistical Assessment of the output succession of the Pseudo Random Number Generation: Fortuna

Laura Castro Argudín[1*], **https://orcid.org/0000-0003-2522-349X**
Yaniel Medina Pérez[2], **https://orcid.org/0000-0002-6290-9442**

[1] Unidad de Proyectos Avanzados, Dirección de Criptografía. Correo electrónico: **lastro9711@gmail.com**
[2] Departamento de Criptología, Facultad 3, Universidad del Ministerio del Interior. Correo electrónico: **yanielmedinaperez@gmail.com**

*** Autor para correspondencia: lastro9711@gmail.com**

Resumen

La Dirección de Criptografía (DC) es el órgano responsabilizado con trazar las políticas en materia criptográfica, por lo que entre sus principales tareas tiene la asimilación de sistemas criptográficos públicos, para su posible empleo, debido a lo cual, la evaluación de los mismos es un aspecto imprescindible a tener en cuenta. El generador de secuencias pseudoaleatorias mixto denominado Fortuna (Kohno, (2010)) es un ejemplo de los generadores públicos de interés para la DC. El trabajo presentado aporta una valoración estadística de las sucesiones de salidas de este Generador de Números Pseudoaleatorios Criptográficamente Seguro basado en el modelo original descrito por los autores. Las pruebas son analizadas mediante el paquete de pruebas estadísticas propuestas en el año 2008, por el Instituto Nacional de Estándares y Tecnologías de los Estados Unidos (NIST por sus siglas en inglés) así como por el conjunto de Test del DieHard. El estudio realizado arroja que el generador Fortuna posee buenas propiedades estadísticas analizando algunas de las sucesiones de salidas, por lo que se recomienda su uso en aplicaciones criptográficas.

Palabras clave: DieHard, Fortuna, generadores de secuencias pseudoaleatorias, NIST, PRNG.

Abstract

The Cryptography Directorate (DC) as an organization. It is responsible for tracking cryptographic policies, so among its main tasks has the assimilation of public cryptographic systems, for possible use, due to which, the evaluation of them is an essential aspect to take into account. The mixed generator of pseudorandom sequences called Fortuna (Kohno, (2010)) is an example of the public generators of interest for DC. The presented work provides a statistical evaluation of the successions of PRNG Fortuna departures based on the original model described by the authors. The tests are analyzed through the statistical test package proposed by the National Institute of Standards and Technologies of the United States (NIST), as well as by the DieHard test suite. The study shows that the generator of Fortuna has good statistical properties that analyze some of the successions of products. For this reason, its use is recommended in cryptographic applications

Keywords: *DieHard, Fortuna, Pseudorandom Number Generation, NIST, PRNG*.

Recibido: 04/01/2022
Aceptado: 28/03/2022

Introducción

En la actualidad con el desarrollo de las tecnologías y las comunicaciones se ha hecho cada vez más necesario la protección de la información. Los números aleatorios han jugado un papel importante en el desarrollo de la Criptografía. Muchos protocolos y algoritmos necesitan de ellos para su funcionamiento.

Los generadores de secuencias pseudoaleatorias normalmente producen secuencias de cualidad insuficiente para ser usado en escenarios criptográficos. Por tal motivo es importante tener una valoración estadística de las sucesiones salidas de los mismos para poder garantizar entre otras primitivas: confidencialidad, integridad, así como autentificación y no repudio. (Kelsey, 1999). Teniendo en cuenta las irregularidades expuestas anteriormente en la situación problemática se plantea como problema científico: ¿Cómo son las características estadísticas del generador de secuencias pseudoaleatorias mixto Fortuna? Teniendo en cuenta como objeto de estudio: Los generadores de secuencias pseudoaleatorias y como campo de acción: el generador de secuencias pseudoaleatorias mixto, Fortuna; con el objetivo general de: Analizar el generador de secuencias pseudoaleatorias mixto Fortuna mediante sus características estadísticas para su utilización en aplicaciones criptográficas en el país.

Materiales y métodos

Respondiendo al objetivo que se planteó anteriormente y al problema científico se aplicaron los siguientes métodos utilizados fueron los siguientes:

Histórico-lógico: Estudia toda la trayectoria histórica del objeto, su tendencia, las etapas más significativas de su desarrollo y sus conexiones históricas fundamentales de forma cronológica y lógica. Permite profundizar en los referentes teóricos metodológicos de la investigación. Se utilizó para el estudio de trabajos anteriores y utilizarlos como punto de referencia y comparación con los resultados obtenidos.

Hipotético-Deductivo: Permite adelantar y verificar las nuevas hipótesis sobre la realidad, establece nuevas predicciones a partir del sistema de conocimientos que se tiene. Consiste en deducir y explicar leyes e hipótesis de menor nivel de generalidad y abstracción a partir de propuestas de mayor nivel de generalidad, abstracción y lógica.

Métodos empíricos:

El Análisis de contenido y comunicación interpersonal: Recopila información bibliográfica acerca del fenómeno a investigar y desarrollar. Se utiliza para analizar los documentos relacionados con la investigación.

Materiales

Implementar fuentes naturales puede ser de mucho trabajo. Las fuentes típicamente tienen que ser integradas en los diversos controladores de hardware del sistema operativo. Esto es complicado de hacer a nivel de usuario.

Para la implementación del Fortuna en este trabajo se utilizó una semilla aleatoria tomada del Visual Studio 2012 para así poder comprobar las sucesiones de salidas del propio generador que es básicamente el AES en modo contador con los requerimientos del propio PRNG Fortuna descrito por los autores. Fue implementado para generar tantos datos aleatorios como se deseen cambiando la clave cada 1 MB como se exige en su publicación.

Pruebas de aleatoriedad

Se pueden aplicar varias pruebas estadísticas a una secuencia para intentar comparar y evaluar su aleatoriedad. (Diffie, (1979)). Hay un número infinito de posibles pruebas estadísticas, cada una de las cuales evalúa la presencia o ausencia de un "patrón" que, si se detecta, indicaría que la secuencia no es aleatoria. Debido a que hay tantas pruebas para juzgar si una secuencia es aleatoria o no, ningún conjunto finito específico de pruebas se considera completo. Además, los resultados estadísticos deben interpretarse con cuidado y precaución para evitar conclusiones incorrectas sobre una secuencia de un generador específico.

Conjunto de Test estadísticos propuestos por el NIST

Test estadísticos del NIST: (AndrewRukhin, abril del 2010). Es el resultado de la colaboración entre la División de Seguridad Computacional y la División de Ingeniería Estadística del NIST. Esta batería consiste en quince test estadísticos desarrollados para medir la aleatoriedad de secuencias binarias producidas por generadores de números aleatorios (RNG por sus siglas en inglés) y generadores de números pseudo-aleatorios (PRNG por sus siglas en inglés).

Estas son las 15 pruebas: (Federal Information Processing Standards, noviembre 26 del 2001)

1- Test de Frecuencia

2-Test de Frecuencia en bloque

3-Test de Rachas

4-Test de Rachas Largas en Bloques

5-Test Rango de Matrices Binarias
6-Test Transformada Discreta de Furier
7-Test Alfabetos Coincidentes Solapados
8-Test Alfabetos no Coincidentes Solapados
9-Test Estadística Universal de Maurer
10-Test Complejidad Lineal
11-Test Serie
12-Test Entropía Aproximada
13-Test Suma Acumulativa
14-Test Excursión Aleatoria
15-Test Variante Excursión Aleatoria

El NIST establece dos enfoques para interpretar los resultados empíricos obtenidos:
1-El análisis de la proporción de secuencia que pasan los Test estadísticos.
2-El análisis de la distribución de los P-valores en busca de uniformidad.

En caso de que uno de ellos falle, la hipótesis nula H_0 definida como "la secuencia analizada es aleatoria"; debe ser rechazada.

Para el análisis de proporción se establece un intervalo de confianza mediante la ecuación $p \pm 3 \sqrt{\frac{p(1-p)}{m}}$ donde: $p = 1 - \alpha$ y $\alpha = 0{,}01$.

Para cada Test se calcula un valor de proporción dividiendo la cantidad de P-valores $\geqslant 0.01$ entre el total de muestras analizadas.

Conjunto de Test estadísticos del DIEHARD

Es un conjunto de 14 Test estadísticos que mide la calidad de una sucesión de números para evaluar su aleatoriedad. Está determinado como uno de los conjuntos de Test más rigurosos de los existentes. (Foundation, 1985)

1-Test Paradoja del Cumpleaños

2-Test Permutaciones Solapadas

3-Test Rango de matrices

- Rango de matrices $31x31$ y $32x32$

- Rango de matrices $6x8$

4-Test del Mono

5-Test de flujo sobre palabras de 20 bits

6-Test Contador de unos en una secuencia de bytes

7-Test Contador de unos en un byte específico

8-Test Aparcamiento

9-Test Distancia mínima

10-Test Círculo aleatorio

11-Test Reducción

12-Test Sumas solapadas

13-Test Rachas

14-Test de los Dados

Resultados y discusión

Cada muestra consiste en 10^3 secuencias de 10^6 bits cada una, generadas de forma aleatoria como exige el Fortuna. En este caso el rango de proporción para concluir si las muestras son aleatorias es (0,980561, 0,99943928) para todas las pruebas excepto el Test de Excursiones Aleatorias y la Variante de Excursiones Aleatorias, en cuyo caso las muestras analizadas son de tamaño 599 y por tanto la proporción debe ser superior a 0,977804.

Análisis de los resultados de los test del NIST

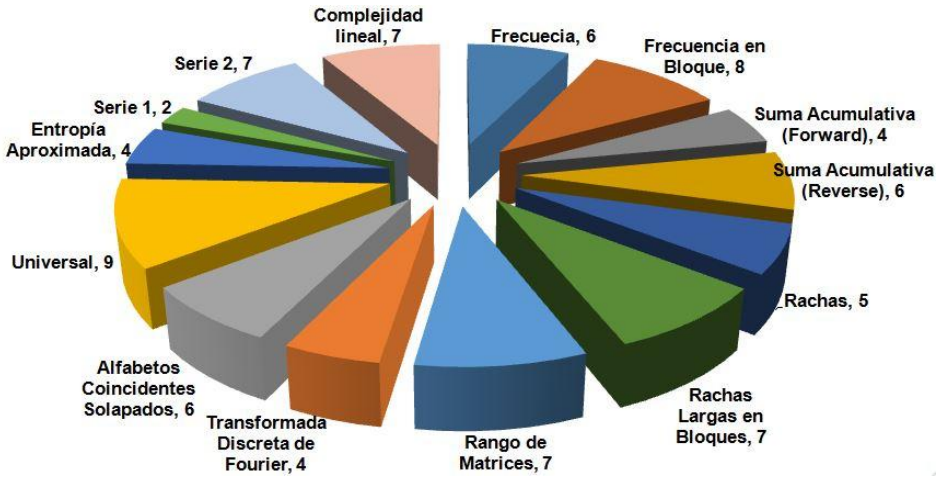

Figura 1: Cantidad de muestras que fallaron por Test

Análisis de Proporción

En las siguientes gráficas se puede notar que todos los valores de proporción se encontraban dentro del intervalo adecuado por lo que se acepta la hipótesis nula (H_0). Por ejemplo, en el test de Frecuencia se observó una proporción de 0.9940, lo cual significa que 994 de las 1000 secuencias analizadas resultaron aleatorias, fallando solo 6 de estas como era de esperar.

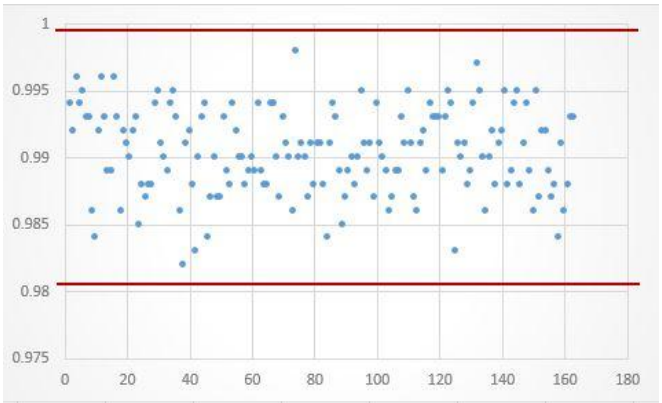

Figura 2: Análisis de Proporción

Análisis de Proporción de los Test de Excursión Aleatoria y Variante de Excursión Aleatoria

En el caso de los Test de Excursión Aleatoria y Variantes de Excursión Aleatoria, que solo se le pudieron realizar a 599 secuencias como se dijo anteriormente, se observa que todos están dentro del intervalo de confianza, por lo que se acepta la hipótesis nula (H_0).

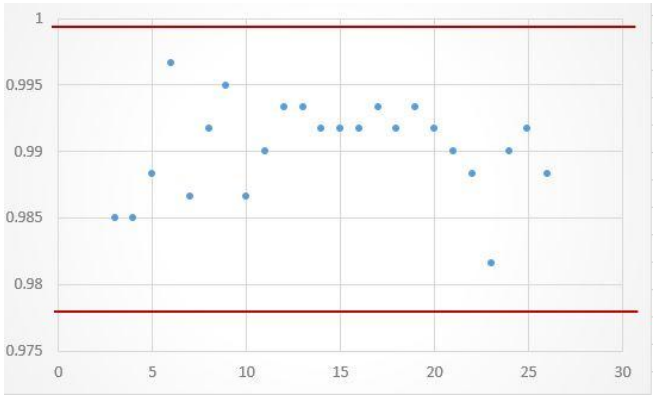

Figura 3: Análisis de proporción de los Test de Excursión Aleatoria y Variante de Excursión Aleatoria

Análisis de los resultados de los test del DieHard

Para analizar los resultados estadísticos del conjunto de pruebas del DieHard, se tomó cada muestra como una secuencia de 13 MB generada de forma aleatoria. En este caso el análisis consiste en observar que los P-valores se encuentran en el intervalo (0.0001, 0.9999) dejando un margen de falla lo más amplio posible, ya que en realidad los P-valores obtenidos en muestras no aleatorias están cercanos a 10^{-6} o a $1 - 10^{-6}$.

Las siguientes gráficas corresponden a los P-valores obtenidos en una sola muestra generada por el PRNG Fortuna. A pesar de que los mismos muestran valores muy cercanos a los extremos del intervalo, aún se encuentran en el margen establecido.

De esta forma, al igual que en el caso del conjunto de pruebas del NIST, se puede suponer que cualquier gama generada por uno de estos algoritmos se comporte de forma aleatoria.

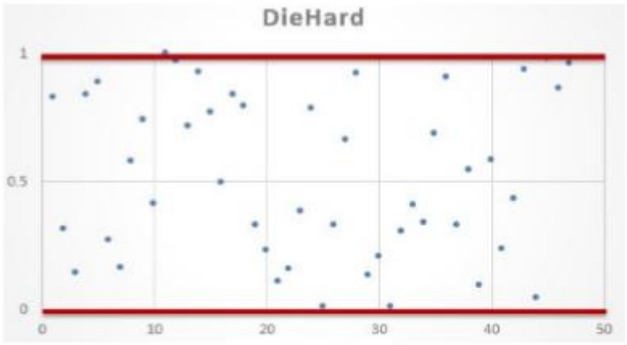

Figura 4: Resultados de los Test del DieHard

Conclusiones

El objetivo de este trabajo fue cumplido de forma satisfactoria. Para ello se implementó el generador Fortuna y se generaron sucesiones. El principal aporte del trabajo fue analizar la aleatoriedad de las sucesiones de salida del PRNG Fortuna, concluyendo que las muestras analizadas resultaron pseudo-aleatorias. Como resultado de este trabajo, la Dirección de Criptografía cuenta con una valoración estadística del PRNG Fortuna, que le permitirá tener un criterio de decisión en cuanto a su posible uso en aplicaciones que busquen garantizar la seguridad de la información con técnicas criptográficas.

Conflictos de intereses

Declaramos que los autores de esta investigación no tenemos conflicto de intereses

Contribución de los autores

1. Conceptualización: Laura Castro Argudín, Yaniel Medina Pérez
2. Curación de datos: Laura Castro Argudín, Yaniel Medina Pérez
3. Análisis formal: Laura Castro Argudín, Yaniel Medina Pérez
4. Adquisición de fondos: Laura Castro Argudín, Yaniel Medina Pérez
5. Investigación: Laura Castro Argudín, Yaniel Medina Pérez
6. Metodología: Laura Castro Argudín, Yaniel Medina Pérez
7. Administración del proyecto: Laura Castro Argudín, Yaniel Medina Pérez

8. Recursos: Laura Castro Argudín, Yaniel Medina Pérez
9. Software: Laura Castro Argudín, Yaniel Medina Pérez
10. Supervisión: Laura Castro Argudín, Yaniel Medina Pérez
11. Validación: Laura Castro Argudín, Yaniel Medina Pérez
12. Visualización: Laura Castro Argudín, Yaniel Medina Pérez
13. Redacción – borrador original: Laura Castro Argudín, Yaniel Medina Pérez
14. Redacción – revisión y edición: Laura Castro Argudín, Yaniel Medina Pérez

Financiamiento

La investigación fue financiada por los autores.

Referencias

AndrewRukhin, J. S. (abril del 2010). Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. *NIST Special Publication 800-2210*, ii-G1.

Diffie, W. a. ((1979)). Privacy and Authentication: An Introduction to Cryptography. *Proceedings of the IEEE*, pp. 397–427.

Dworkin, M. J. (Agosto 04, 2015). Federal Information Processing Standards Publication. SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions.. *FIPS PUB 202*, 56 paginas.

Federal Information Processing Standards. (noviembre 26 del 2001). Announcing the Advanced Encryption Standard (AES. *Federal Information Processing Standards Publication 197*, 47 páginas.

Foundation, P. a. (1985). *The Diehard Battery of Tests of Randomness.*

Kelsey, J. S. (1999). *"Notes on the Design and Analysis of the Yarrow Cryptographic Pseudorandom Number Generator"*. Minneapolis: Counterpane Systems; 101 E Minnehaha Parkway, Minneapolis, MN 55419, USA.

Kohno, N. F. (2010). *Cryptography Engineering. Design Principles and Practical Applications.* United States of America: Wiley Publishing, Inc. .

Technology, National Institute of Standards and Tecnology. ((2001)). Recommendation for Block Cipher Modes of Operation. *NIST Special Publication 800-38A.*, 59 paginas.

V., D. J. (2002). *The Design of Rijndael: AES - The Advanced Encryption Standard.* Springer-Verlag.