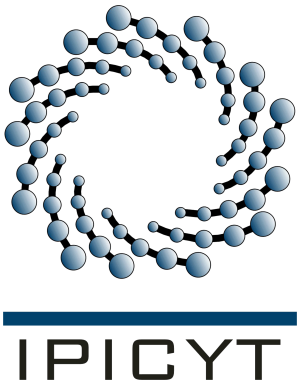


INSTITUTO POTOSINO DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA A.C.

POSGRADO EN CONTROL Y SISTEMAS DINÁMICOS

Generación de cajas de sustitución útiles en cifrado por bloques con base en dinámica caótica.

Tesis que presenta

M.C.S.D. Bahia Betzavet Cassal Quiroga

Para obtener el grado de

Doctor en Control y Sistemas Dinámicos

Director de la Tesis:

Dr. Eric Campos Cantón

San Luis Potosí, S.L.P., 08 de Octubre de 2021

Constancia de aprobación de la tesis

La tesis **Generación de cajas de sustitución útiles en cifrado por bloques con base en dinámica caótica** presentada para obtener el Grado de Doctor en Control y Sistemas Dinámicos fue elaborada por **Bahia Betzavet Cassal Quiroga** y aprobada el **08 de 10 de 2021** por los suscritos, designados por el Colegio de Profesores de la División de Matemáticas Aplicadas del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

_____

Dr. Haret-Codratian Rosu Barbus
(Sinodal)

_____

Dr. Juan Gonzalo Barajas Ramírez
(Sinodal)

_____

Dr. Guillermo Huerta Cuéllar
(Sinodal)

_____

Dr. José Tuxpan Vargas
(Sinodal)

_____

Dr. Eric Campos Cantón
(Director de tesis)

Créditos Institucionales

Esta tesis fue elaborada en la División de Matemáticas Aplicadas del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. Eric Campos Cantón.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología CONACYT-424203 y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.

IPICYT

Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 020 del Libro Primero de Actas de Exámenes de Grado del Programa de Doctorado en Control y Sistemas Dinámicos está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 8 días del mes de octubre del año 2021, se reunió a las 16:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

Dr. Juan Gonzalo Barajas Ramírez	**Presidente**	**IPICYT**
Dr. Eric Campos Cantón	**Secretario**	**IPICYT**
Dr. Haret-Codratian Rosu Barbus	**Sinodal**	**IPICYT**
Dr. José Tuxpan Vargas	**Sinodal**	**IPICYT**

a fin de efectuar el examen, que para obtener el Grado de:

DOCTORA EN CONTROL Y SISTEMAS DINÁMICOS

sustentó la C.

Bahia Betzavet Cassal Quiroga

sobre la Tesis intitulada:

Generación de cajas de sustitución útiles en cifrado por bloques con base en dinámica caótica

que se desarrolló bajo la dirección de

Dr. Eric Campos Cantón

El Jurado, después de deliberar, determinó

APROBARLA

Dándose por terminado el acto a las 18:30 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición de la interesada y para los fines que a la misma convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 8 días del mes de octubre de 2021.

Dr. Marcial Bonilla Marín
Secretario Académico

Mtra. Ivonne Lizette Cuevas Vélez
Jefa del Departamento del Posgrado

IPICYT
SECRETARÍA ACADÉMICA
INSTITUTO POTOSINO DE
INVESTIGACIÓN CIENTÍFICA
Y TECNOLÓGICA, A.C.

Dedicatoria

A mi familia.

Agradecimientos

Primero quiero agradecer a Dios por acompañarme todos los días, por darme la fortaleza, capacidad para elegir, seguir, terminar mis estudios y este trabajo de tesis. De manera muy especial quiero agradecer a mi esposo Héctor, gracias por permitirme formar parte de tu vida, gracias por tu amor, por tu tiempo, por tu comprensión, por ser como eres. Gracias por impulsarme hacer una mejor persona, apoyarme, aguantarme y sobre todo por motivarme hacer las cosas de la mejor manera posible.

A mi madre Verónica, mi hermana Andrea, sobrinos y a la familia Cassal Quiroga, gracias por brindarme su apoyo, confianza y amor a pesar de la distancia, por estar en los momentos mas felices y difíciles de mi vida. Gracias por ayudarme a cumplir esta meta, por las porras, besos y buenos deseos. Los amo.

A mi director de tesis, el Dr. Eric Campos no me queda mas que agradecerle infinitamente por tomarme como su alumna de doctorado aunque se que tenía grandes deficiencias. Por haber depositado su confianza en mí para realizar este trabajo. Gracias por toda su paciencia, por sus comentarios, por hacer esas preguntas en las reuniones que me hicieron dar más de mi para desenvolverme y llegar a este punto culminante en esta etapa de mi vida. Siempre le estaré muy agradecida por toda su orientación y apoyo durante todo este tiempo. Muchas gracias por todo.

A mis sinodales, el Dr. Haret-Codratian Rosu, el Dr. Juan Gonzalo Barajas, el Dr. Guillermo Huerta y el Dr. José Tuxpan gracias por tomarse su tiempo para leer y revisar este trabajo de tesis.

A todos los investigadores de la División de Matemáticas Aplicadas, por contribuir en mi formación, gracias por sus consejos y enseñanzas, impulsándome para seguir adelante.

A CONACYT por la beca de doctorado, a la fundación Sofía Kovaleskaia, a la Sociedad Matemática Mexicana e IPICyT, por el apoyo otorgado y facilitarme los medios para desarrollar mi proyecto doctoral.

A mis amigos y compañeros del IPICyT, muchas gracias por brindarme recuerdos

felices dentro del instituto, por su paciencia, consejos y convivencia.

A la familia Gilardi Velazquez muchísimas gracias por todo su tiempo, comprensión, apoyo, por escucharme cuando lo necesitaba, por siempre tener un abrazo y un beso para mí. Gracias por permitirme formar una familia con ustedes, por abrigarme como una hija, los amo y los amare siempre.

A las personas que en este tiempo se han convertido en una parte muy importante de mi vida, Adriana, Ismael, Aide, Omar, Aletse, Rodolfo, Pablo, José Luis, Roberto y muchas personas mas que no acabaría de nombrar, esas personas que me han demostrado que no se necesitan lazos de sangre para quererlas. Muchísimas gracias por su infinita paciencia, por su compañía, por su apoyo incondicional, por todos los momentos que pasamos juntos,y por la convivencia todos estos años. Mis respetos y admiración para cada uno de ustedes. Los quiero mucho.

A la familia Campos Méndez, mi familia de corazón, gracias por su calidez, por brindarme su hogar y por hacerme sentir una mas de ustedes.

A todos los que por descuido no han sido incluidos en estas líneas, pero han sido fundamentales en mi formación como profesional y personal, mis más sinceros agradecimientos.

Finalmente, a la vida que me dio la oportunidad de encontrármelos en el camino y contar con cada uno de ustedes.

Resumen

En este trabajo de tesis, se presentan algoritmos para el diseño de cajas de sustitución de $n \times n$-bits (conocidas como S-boxes por su nombre en inglés), basadas en series de tiempo de un sistema dinámico discreto con comportamiento caótico. Los elementos de una caja de sustitución de $n \times n$-bits son obtenidos de secuencias generadas por series de tiempo con distribución uniforme. En particular, las secuencias con distribución uniforme, usadas en esta tesis, son generadas a través de dos series de tiempo con retardo del mapa logístico. El objetivo de utilizar estas dos series con retardo, es ocultar el tipo de mapeo que se está utilizando, por lo que se evita la distribución en forma de "U"del mapa logístico y se obtienen elementos de la caja de sustitución no correlacionados. Los algoritmos que se proponen son simples y garantizan la generación de cajas de sustitución, componente principal en el cifrado de bloque, que cumplen los siguientes criterios: biyectividad; no linealidad; estricto criterio de avalancha; criterio de independencia de los bits de salida; criterio de distribución equiprobable XOR entrada/salida; y probabilidad lineal máxima esperada. Las cajas de sustitución que cumplen estos criterios son conocidas comúnmente como "buenas cajas de sustitución". Finalmente, se desarrolla una aplicación basada en el principio de cifrado polialfabético, a través de las cajas de sustitución generadas por los algoritmos propuestos.

Abstract

In this thesis work, algorithms are presented for the design of substitution boxes of $n\ timesn$ -bits (known as S-boxes by their English name), based on time series of a discrete dynamic system with chaotic behavior. The elements of a substitution box of $n\ timesn$ -bits are obtained from sequences generated by time series with uniform distribution. In particular, the sequences with uniform distribution, used in this thesis, are generated through two time series with delay of the logistic map. The objective of using these two series with delay is to hide the type of mapping being used, thus avoiding the Üshaped distribution of the logistic map and obtaining uncorrelated substitution box elements. The proposed algorithms are simple and the generation of substitution boxes, the main component in block encryption, which meet the following criteria: bijectivity; non-linearity; strict avalanche criteria; criterion of independence of the output bits; equiprobable distribution criterion XOR input/output; and expected maximum linear probability. Replacement boxes that meet these criteria are commonly known as "good replacement boxes". Finally, an application is developed based on the polyalphabetic encryption principle, through the substitution boxes generated by the proposed algorithms.

Índice general

Índice de figuras

Índice de tablas

Capítulo 1

Introducción

1.1. Orígenes de la criptografía

La Real Academia Española define la criptografía como "el arte de escribir con clave secreta o de un modo enigmático", hoy en día se le considera como la ciencia dedicada a estudiar métodos para codificar (cifrar o encriptar) mensajes, para evitar que su contenido pueda ser leído por un tercero no autorizado.

La criptografía es probablemente una de las disciplinas más antiguas relacionadas con las matemáticas. El objetivo principal de esta ciencia es garantizar el secreto de la comunicación entre dos entidades (personas, organizaciones, etc.), y en segundo lugar, cerciorarse que la información enviada es auténtica en un doble sentido, es decir: la persona que envía el mensaje (información) sea efectivamente quien dice ser y que el contenido del mensaje enviado, habitualmente denominado criptograma, no haya sido modificado en tránsito.

Los orígenes de la criptografía tienen lugar en tiempos antiguos con la invención de la escritura, se originó el deseo de transmitir información de manera rápida y segura. Fue así como surgieron los primeros mecanismos para esconder o disfrazar la información. Por miles de años, reyes, reinas y generales se han apoyado en una comunicación secreta para gobernar sus países y dirigir sus ejércitos. Pero al mismo tiempo, han tenido que

estar al tanto de las consecuencias de que sus mensajes caigan en manos equivocadas, revelando secretos valiosos a naciones rivales o fuerzas opositoras. La amenaza de una posible intercepción enemiga, ha motivado el desarrollo de técnicas para disfrazar o esconder los mensajes con la intención de que sólo el receptor pueda entenderlos. El uso de la criptografía comenzó con métodos de envío de mensajes confidenciales por medio de escritura secreta en la antigüedad.

La utilización de métodos criptográficos se remonta a 2000 años antes de Cristo, en Esparta, donde los espartanos usaban un sistema de escritura secreta llamado Scytale (Figura. 1.1a [1]) durante los enfrentamientos con Atenas. Otro ejemplo es el cifrado del general romano Julio César, el cual consiste en reemplazar cada letra del mensaje por otra que se encuentra un número fijo de posiciones más adelante en el alfabeto. Por ejemplo, en la Figura 1.1b [2] se muestra un desplazamiento hacia la izquierda de 3 posiciones del alfabeto en español, es decir una **B** en el texto original se convierte en una **E** en el texto codificado [1].

(a) (b)

Figura 1.1: (a) Escítala de Esparta , (b) Cifrado César.

Posteriormente, con la llegada de la Revolución Industrial, se crearon máquinas de rotores para cifrar mensajes, la primera máquina de este tipo que se tiene registro se le conoce como la máquina de Hebern, la cual poseía un rotor. Para los primeros años de 1900's se inventó la máquina Enigma, esta máquina electromecánica es tal vez la más famosa utilizada para cifrar y descifrar los códigos durante la Segunda Guerra Mundial (ver Figura. 1.2 [3]), fue usada por los Nazis en la Segunda Guerra Mundial.

[1] https://n9.cl/qkz3p
[2] https://n9.cl/gql7z
[3] https://n9.cl/6ikb

Figura 1.2: Máquina Enigma.

Se basaba en los cifrados de sustitución y en la máquina de Hebern, de tal forma que cambiaba una letra por otra de forma mecánica por medio de tres rotores, la llave de esta máquina estaba dada por la posición inicial de los rotores y era capaz de producir 17 576 combinaciones. En 1935 la máquina contaba con 4 rotores y para esa fecha los alemanes la tomaron para uso oficial y exclusivo militar, con las modificaciones que realizaron la máquina era capaz de producir 10 967 424 combinaciones, por esta razón se decía que era indescifrable. Con un funcionamiento avanzado para su época, la máquina Enigma Alemana fue la que dio inicio a las primeras computadoras que se utilizaron para cifrar y descifrar códigos, además de servir como base para cifrados modernos [2, 3].

En general, los cifrados históricos basan su funcionamiento en letras, incluso la máquina Enigma. Sin embargo, en 1917 Gilbert Vernam ingeniero de la compañía AT&T propuso el cifrado conocido como Vernam [4], en el cual se proponía un algoritmo de cifrado completamente diferente, que basa su funcionamiento en el sistema binario. Esto significó el comienzo de la era digital.

Posteriormente, Joseph Mauborgne propuso que la llave del cifrado fuera una secuencia totalmente aleatoria, a este cifrado se le conoce como **one-time-pad**. Este cifrado sigue teniendo una influencia en los cifrados modernos, ya que fue el primero y único cifrado para el que existe una demostración de seguridad perfecta, la cual fue propuesta por Claude Shannon [1].

Con el fin de que la comunicación sea segura, se requiere de un sistema que tanto el

que envía (emisor) como el que recibe (receptor) el mensaje lo conozcan, pero que permanezca desconocido ante cualquier posible enemigo. El sistema, denominado esquema de cifrado (criptosistema), consiste generalmente de un método para esconder (cifrar) la información y una pieza de información adicional, llamada llave, que proporciona el acceso a un cifrado en particular. Además, el alfabeto (alfabeto plano) usado para escribir el mensaje original, no necesariamente es el mismo utilizado para escribir el mensaje cifrado (alfabeto del cifrado).

Una clasificación para la criptografía es la división en criptografía de substitución y criptografía de transposición. En la criptografía de transposición, las letras de un mensaje se re-acomodan generando un anagrama que aumenta el número de posibles combinaciones a medida que se incrementa el número de letras en el mensaje. La criptografía de substitución cambia cada letra en el mensaje original (o mensaje que se desea enviar, también llamado texto plano) por una letra diferente. Esto significa que cada letra es reemplazada por otra pero se mantiene su posición. Un ejemplo de criptografía de substitución se da con el esquema de cifrado del César antes mencionado.

En general cuando se quiere enviar un mensaje, el texto plano mediante transformaciones es convertido en un texto sin sentido; al que se le denomina texto cifrado. Es importante mencionar que para que esta transformación tenga sentido es necesario que este proceso requiera de una llave, de tal forma que solo cuando se aplica la llave correcta el proceso es reversible.

Hoy en día, nos encontramos en la era de la información, una época donde el intercambio de información por medios electrónicos como Internet, correo electrónico, telefonía celular, comercio electrónico, tarjetas inteligentes, entre otros, resulta algo cotidiano. Si esta información no es protegida adecuadamente es posible que se haga mal uso de la misma, por esta razón es necesario y fundamental contar con sistemas criptográficos seguros.

Cuando se almacena o transmite información valiosa o secreta, a menudo resulta insuficiente, inaplicable y costoso protegerla sólo de manera física. Por lo que se deben emplear otras técnicas más apropiadas y eficientes; la criptografía se ha desarrollado como un conjunto de técnicas para solucionar este tipo de situaciones.

La criptografía aplicada a la seguridad de las transacciones electrónicas de datos, ha adquirido fundamental importancia en los últimos tiempos. Cada día millones de usuarios generan e intercambian grandes volúmenes de información en diversos campos, tales como archivos financieros y de índole legal, informes médicos, servicios bancarios a través de Internet, conversaciones telefónicas y transacciones de comercio electrónico. Estos y otros ejemplos de aplicaciones merecen un tratamiento especial desde el punto de vista de seguridad, no sólo en el transporte de dicha información, sino también en lo que respecta al almacenamiento de la misma. El uso de técnicas de criptografía es especialmente aplicable en este sentido.

1.2. Criptografía

La criptografía es una rama de la criptología (Figura 1.3) que se enfoca en solucionar problemas de confidencialidad, integridad y autenticación [5]. La propiedad de confidencialidad garantiza que la información es accesible sólo para aquellos autorizados a tener acceso, esta propiedad es principalmente provista por los cifrados simétricos; la integridad es el proceso que permite saber si un mensaje llega a su destino completo y sin alteraciones, provista principalmente en los protocolos; la autenticación es el proceso por el cual se determina la identidad de un usuario, esta propiedad es proporcionada por los cifrados asimétricos.

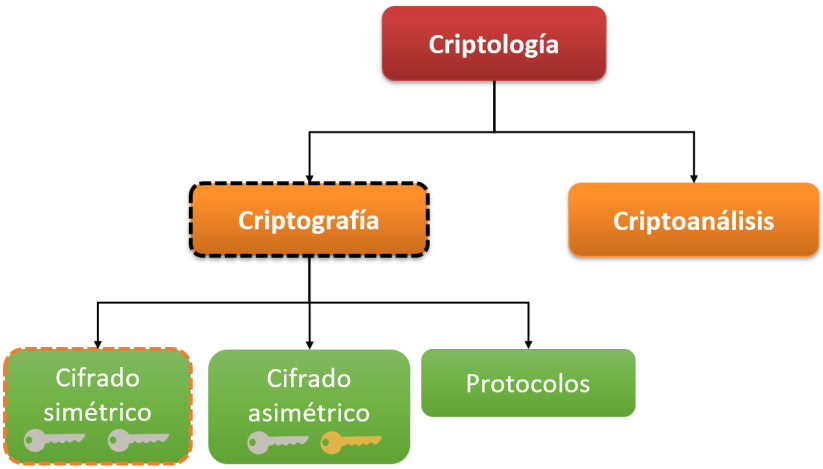


Figura 1.3: Visión general del campo de la Criptología.

Los protocolos usualmente se utilizan en conjunto con los cifrados simétricos y asimétricos, entre ellos se encuentran las funciones Hash y los códigos de autentificación del mensaje (MACs por su sigla en inglés).

Los cifrados asimétricos son aquellos en los que se utilizan dos llaves diferentes (una llave privada y una pública) en cada uno de los extremos de la comunicación para cifrar los mensajes y descifrarlos. Cada usuario tendrá una llave pública y otra privada. Las llaves públicas y privadas se generan simultáneamente y están ligadas la una de la otra. Esta relación debe ser muy compleja para que resulte muy difícil que se obtenga una a partir de la otra.

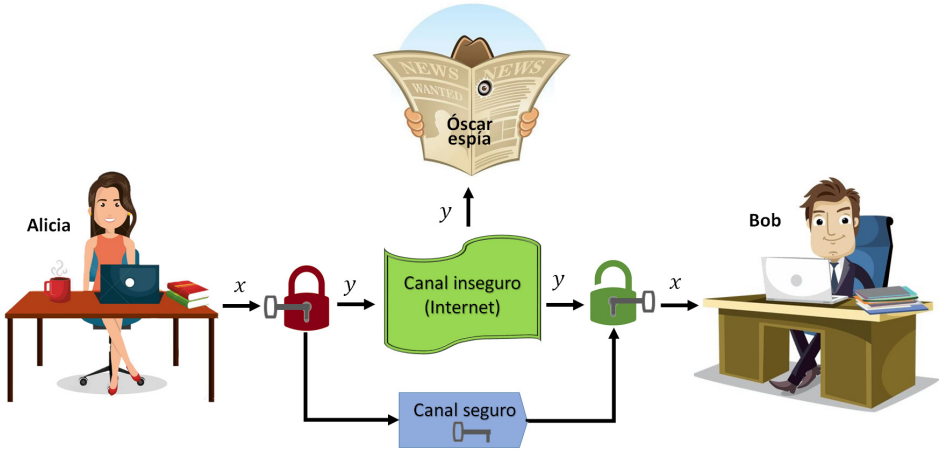


Figura 1.4: Criptosistema de llave simétrica, permite el intercambio de información entre dos usuarios por medio de un canal inseguro haciéndolo ilegible para un tercero.

La criptografía de llave simétrica (cifrado simétrico) permite a dos usuarios (Alicia y Bob nombres fueron usados por primera vez en [6]) compartir información cifrada por medio de un canal inseguro, haciéndolo ilegible para un tercero (Oscar), mientras que la llave utilizada para cifrar y descifrar los mensajes es enviada mediante un canal seguro, Figura 1.4. Existen dos clases dentro de esta clasificación, los cifrados en flujo y bloques, Figura. 1.5.

Los cifrados en flujo, se basan en cifrar un elemento del texto plano (mensaje a cifrar) y un elemento de la llave (pseudo-aleatoria generada por una semilla), para generar un texto cifrado. En el caso de los cifrados en bloques, dividen el mensaje a cifrar en diversos fragmentos de longitud fija (típicamente de 64, 128 o 256 bits) y aplican el

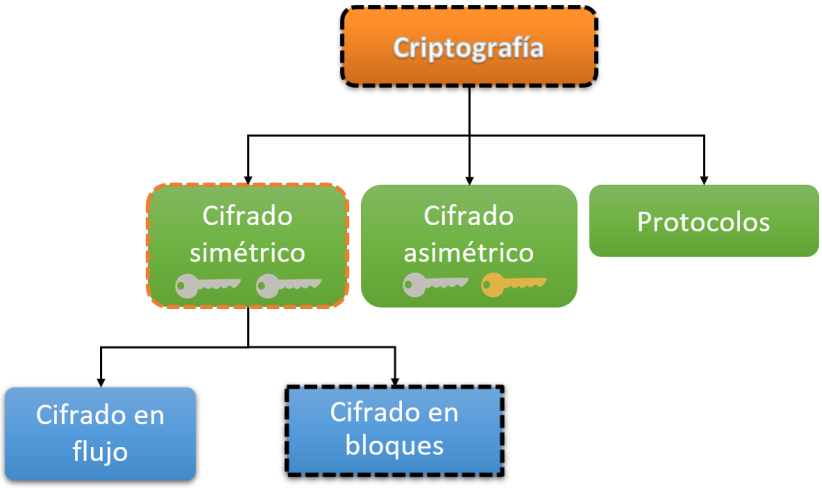


Figura 1.5: Principales áreas dentro de la Criptografía.

algoritmo de cifrado en bloque repetidas veces al fragmento o bloque de información, usando la misma llave para obtener bloques de texto cifrado de tamaño fijo.

Nos enfocaremos en los criptosistemas basados en cifrados en bloque los cuales dividen el mensaje a cifrar en diversos fragmentos, por lo general de longitud fija, y aplican a cada uno de ellos una transformación criptográfica. La mayoría de estos algoritmos se basan en los conceptos de confusión y difusión propuestos por Shannon.

1.3. Cifrados en bloque

En 1973 el Buro Nacional de Estándares (NBS por su sigla en inglés), ahora conocido como Instituto Nacional de Estándares y Tecnología (NIST por su sigla en inglés), lanzó una convocatoria para crear un cifrado y establecerlo como un estándar en Estados Unidos. La idea era encontrar un cifrado que fuera seguro y pudiera ser usado en una variedad de aplicaciones. En 1974 recibieron una propuesta realizada por un grupo de criptógrafos que trabajan para la compañía IBM, el algoritmo fue llamado cifrado Lucifer. Lucifer es una familia de cifrados desarrollados por Horst Feistel a finales de la década de 1960 y fue uno de los primeros cifrados en bloque que operó sobre información digital. En particular Lucifer trabajaba con bloques de 64 bits usando una llave de 128 bits. Después de ser analizado por agencias gubernamentales, las que a su vez propusieron algunos cambios, este cifrado fue bautizado bajo el nombre de

Data Encryption Standard (DES). Uno de los cambios más significativos que tuvo este cifrado, fue que DES se diseñó específicamente para resistir ataques de criptoanálisis diferencial [7].

Finalmente, en 1977 la versión final del cifrado DES fue dada a conocer públicamente [8], en la cual se describía el funcionamiento completo del algoritmo, sin embargo, algunos criterios de diseño como las cajas de sustitución, nunca fueron descritos.

Originalmente el cifrado DES se concibió para ser el estándar durante 10 años hasta 1987, sin embargo, fue hasta el año de 1999 cuando fue remplazado.

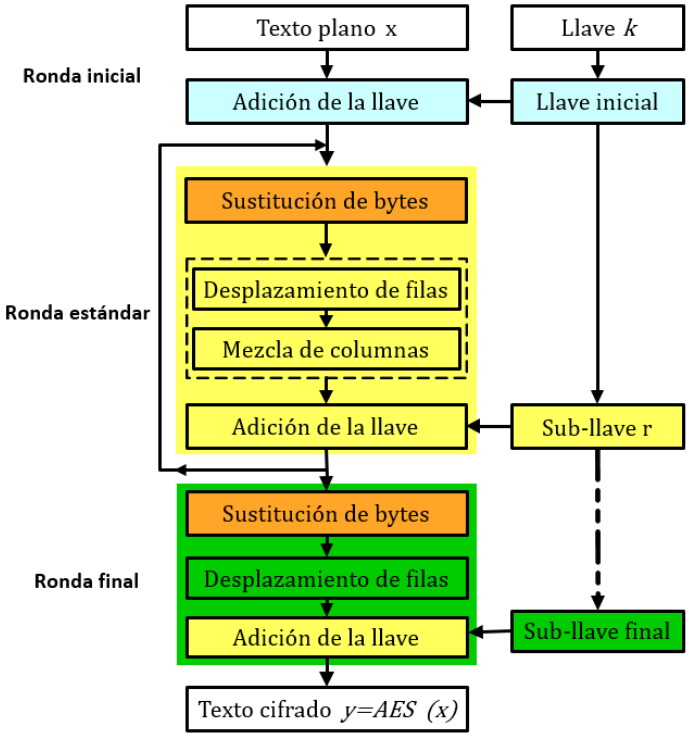


Figura 1.6: Diagrama general del algoritmo AES.

Por otra parte, en 1997 el NIST lanzó una nueva convocatoria para establecer un nuevo cifrado el cual llevaría el nombre de Advanced Encryption Standard (AES), a diferencia del cifrado DES sería seleccionado de un concurso abierto administrado por el NIST. Los requerimientos que se marcaban en la convocatoria eran que en este nuevo cifrado debería de manejar bloques de 128 bits y soportar tres diferentes longitudes de llave: 128, 192 y 256 bits. Fue en 2001 que el cifrado Rijndael propuesto por dos criptógrafos Belgas Joan Daemen y Vincent Rijmen, fue anunciado oficialmente

como el nuevo estándar para Estados Unidos [9]. En la Figura 1.6 se muestra un diagrama general del algoritmo AES el cual consta de tres etapas, la primera es la adición de la llave, la segunda es una ronda estándar compuesta de sustituciones de bytes, desplazamiento de filas, mezcla de columnas y una ronda final [1].

Muchos de los cifrados actuales se han inspirado en el DES cuyo principal componente son las cajas de sustitución. Estas cajas proveen la propiedad de confusión a los criptosistemas y son muy utilizadas en los cifrados por bloques convencionales. Una propuesta muy prometedora que ha crecido en los últimos años es la de criptosistemas basados en sistemas caóticos, debido a que poseen propiedades que son análogas en el campo de la criptografía, a este campo se le ha denominado como criptografía caótica.

1.4. Criptografía basada en caos

En los últimos años, se han publicado muchos artículos que se centran en el estudio de sistemas criptográficos basados en caos [10–19], esto es, debido a la relación que existe entre las propiedades de los sistemas caóticos y las propiedades criptográficas. En [20] se da la relación entre estas propiedades, por ejemplo, la confusión está relacionada con la ergodicidad, la propiedad de difusión con sensibilidad a las condiciones iniciales y la dinámica determinista con la pseudoaleatoriedad determinista.

Existen varios trabajos en donde se expone la relación que existe entre estas dos áreas [20–22]. Como se mencionó anteriormente los conceptos de confusión y difusión son esenciales en la criptografía. La ergodicidad es una fuente de confusión, esto es debido a que los puntos que forman las trayectorias, viajan por todo el espacio fase de forma no ordenada sin importar la entrada, de tal forma que para cualquier entrada se tiene la misma distribución de salida. Por otro lado, la sensibilidad a las condiciones iniciales y a los parámetros de bifurcación generan difusión, ya que un pequeño cambio en la entrada se transforma en un gran cambio en la salida, al realizar una modificación muy pequeña en la condición inicial de un sistema caótico se tendrían órbitas muy diferentes mientras que en los criptosistemas un cambio muy pequeño en la llave o en el texto plano llevarían a salidas diferentes.

Propiedades caóticas	Propiedades criptográficas	Descripción
Ergodicidad	Confusión	La salida tiene la misma distribución de probabilidad para cualquier entrada.
Sensibilidad a las condiciones iniciales	Difusión	Una pequeña variación en la entrada puede causar un gran cambio en la salida.
Dinámica determinista	Pseudoaleatoriedad determinista	Un proceso determinista puede causar un comportamiento pseudoaleatorio.

Tabla 1.1: Relaciones entre propiedades Caóticas y Criptográficas.

Los sistemas dinámicos poseen la propiedad de ser deterministas, esto es, la evolución de un sistema a partir de una condición inicial se puede reproducir siempre y cuando se tengan las mismas condiciones iniciales y parámetros, esto se traduce en comportamiento pseudo-aleatorio ya que es posible reproducir exactamente el mismo comportamiento bajo ciertas condiciones. Podemos decir que la similitud de propiedades entre estos sistemas es directa ya que sus propiedades son análogas (Tabla 1.1 [20]), como se muestra a continuación:

- Sensibilidad respecto a la llave: al invertir un bit de la llave el texto cifrado es completamente diferente.
- Sensibilidad respecto al texto plano: al invertir un bit del texto plano se genera un texto cifrado completamente diferente.
- No debe existir ningún patrón en el texto cifrado el cual pueda relacionar al texto plano.
- Sensibilidad al parámetro de bifurcación: una pequeña variación en el parámetro es suficiente para generar trayectorias diferentes incluso si comienzan en la misma condición inicial.
- Sensibilidad a las condiciones iniciales: dos trayectorias que comienzan en puntos iniciales muy cercanos, se separan exponencialmente una de otra.

- Ergodicidad: los puntos que forman las trayectorias en el espacio fase están uniformemente distribuidos.

Un resumen las principales similitudes y diferencias entre mapeos caóticos y algoritmos criptográficos se muestran en la Tabla 1.2 [23]. Sin embargo aún existe un largo camino por recorrer ya que hay algunas diferencias en el funcionamiento de los criptosistemas clásicos y los criptosistemas basados en caos.

Algoritmos criptográficos	**Sistemas caóticos**
Espacio fase: conjunto finito de enteros	Espacio fase: subconjunto números reales
Métodos algebraicos	Métodos analíticos
Vueltas	Iteraciones
Llave (Booleana)	Parámetros (reales)
Espacio de llaves discreto	Espacio de llaves continuo
Realizaciones digitales por medio de aritmética con enteros	Realizaciones digitales por medio de aritmética con punto flotante

Tabla 1.2: Semejanzas y diferencias entre sistemas caóticos y criptográficos.

En los criptosistemas clásicos tales como el DES y el AES, la llave está definida por conjunto de bits en un espacio de llaves discreto, mientras que en los sistemas basados en caos la llave está definida por números reales en un espacio de llaves continuo. Además en los sistemas clásicos la confusión y la difusión se logra por medio de varias vueltas de un algoritmo mientras que en los sistemas basados en caos se logra por medio de iteraciones.

Una diferencia importante radica en que el proceso de cifrado en los criptosistemas clásicos está definido sobre conjuntos finitos y en tiempo discreto, mientras que el comportamiento caótico evoluciona sobre números reales y puede desarrollarse en tiempo continuo o discreto, esto ha llevado a que la criptografía caótica se divida en dos ramas las cuales estudiaremos a continuación.

En los cifrados basados en sistemas continuos existe una gran variedad de formas y algoritmos para cifrar la información, de forma general podemos tener la siguiente clasificación:

- Generadores de bits pseudo-aleatorios: se basan en las órbitas que pueden generar los mapeos, por medio de alguna transformación se convierten los valores reales

a valores binarios y así generar una secuencia que se le conoce como key-stream. Revisiones de este tema se pueden encontrar en [24, 25]

- Cifrados en flujo: están basados en los generadores de bits, toman como entradas el texto plano y la secuencia de bits (keystream) de tal forma que para cifrar la información se usa la operación XOR y a la salida se tiene el texto cifrado [26–29].

- Cifrados en bloque: este tipo de cifrados basan su funcionamiento en cajas de sustitución, por un lado se pueden definir estas cajas de sustitución por medio de mapeos [30–33] y por otro lado definir algoritmos que usen dichas cajas y cifren la información por medio de sustitución y permutación [34–37].

- Cifrados de sustitución y difusión: recientemente se han propuesto varios algoritmos, este tipo se basan en el principio de los cifrados en bloque sin embargo logran el proceso de cifrado en solo una vuelta [38–40], en las siguientes referencias se realiza un análisis a fondo de este tipo de cifrados [41, 42].

En estos algoritmos la condición inicial y los parámetros de bifurcación son usados como llave del sistema, además la principal ventaja de este tipo de cifrados basados en sistemas discretos radica en que el proceso de cifrado se puede aplicar a cualquier archivo multimedia como son texto, imágenes, audio, vídeo. Existen en la literatura diversas referencias en donde tratan a fondo el tema de cifrados basados en sistemas discretos entre las que se pueden mencionar [21, 43–45].

La construcción de cajas de sustitución criptográficamente seguras, es un componente muy importante para el diseño de criptosistemas seguros. En este sentido se han desarrollado algunos algoritmos utilizando sistemas dinámicos discretos. Por ejemplo, en [10–14], la generación de cajas de sustitución se introdujo a través de una única serie de tiempo de un mapa o combinando dos series de tiempo de mapas diferentes. Sin embargo, estos algoritmos no garantizan que las series utilizadas tengan una distribución uniforme, como en nuestro enfoque basado en dos series caóticas de tiempo de retardo derivadas del mapa logístico. De la misma manera, hay algoritmos basados

en sistemas caóticos continuos [15–17]. También hay algoritmos basados en la mezcla de series de tiempo de sistemas dinámicos continuos y discretos [18, 19] y en [46] el algoritmo se construye a través de series de retardo de tiempo.

La ventaja de utilizar sistemas dinámicos caóticos discretos es que de una iteracción a otra, los elementos de las series de tiempo no están correlacionados. Sin embargo, esto no sucede si se utiliza un sistema dinámico caótico continuo, los elementos de las series de tiempo están fuertemente correlacionados. Por lo tanto, se necesitan muchas iteraciones y el cálculo de la información mutua entre los elementos de las series para poder decir cuando están descorrelacionados, lo que implica un mayor costo computacional.

En los esquemas de encriptación basados en el caos, las secuencias pseudoaleatorias basadas en mapas caóticos se usan generalmente como libreta de un solo uso (one time pad) para el cifrado de mensajes. Dado que los esquemas de encriptación, basados en un mapa caótico de baja dimensión, tienen una complejidad computacional baja, pueden analizarse con un bajo costo computacional utilizando las funciones de iteración y correlación [38]. Las series caóticas con retardo tienen un comportamiento complejo y borran la traza del mapeo que las genera. Usando estas series de tiempo, las cajas de sustitución pueden diseñarse y proporcionar un aumento de las no linealidades asociadas a las cajas de sustitución, de tal forma que se garanticen buenas propiedades estadísticas en los generadores.

Las redes de comunicaciones modernas se han extendido enormemente los límites para hacer posible la comunicación y transmisión de información. Asociado a este acelerado crecimiento existe un incremento en la demanda de las técnicas criptográficas, el cual ha originado un crecimiento en el estudio de la criptografía [5].

1.5. Motivación y Objetivo

Las cajas de sustitución son la principal componente de los cifrados por bloques [1]. Estas cajas de sustitución le dan a los sistemas criptográficos la propiedad de confusión descrita por Shannon y se usan en cifrados en bloque convencionales, como el estándar de cifrado de datos (DES) y el estándar de encriptación avanzada (AES). En estos

criptosistemas la seguridad depende principalmente de las propiedades de las cajas de sustitución que se utilizan, siendo éstas fijas y de carácter público.

Dada la relación que existe entre las propiedades de los sistemas caóticos y las propiedades de los criptosistemas, en este trabajo de investigación se presenta la generación de cajas de sustitución mediante la teoría de sistemas caóticos, de forma tal que los criterios de una buena caja de sustitución sean satisfechos.

1.5.1. Objetivo general

El objetivo general de esta tesis es diseñar un algoritmo para generar cajas de sustitución útiles en cifrado por bloques por medio de sistemas dinámicos con comportamiento caótico que satisfagan los siguientes criterios: biyectividad, no linealidad, estricto criterio de avalancha, Criterio de independencia de los bits de salida, Distribución equiprobable XOR entrada/salida, y Probabilidad lineal máxima esperada.

La estructura de esta tesis esta organizada de la forma siguiente:

- En el capítulo 2 se introducen conceptos básicos necesarios para caracterizar los sistemas dinámicos y se dan a conocer los criterios utilizados en la literatura para generar cajas de sustitución criptográficamente seguras.

- En los capítulos 3 y 4 se proponen dos algoritmos para generar cajas de sustitución basados en generadores pseudoaleatorios caóticos. Estos algoritmos son validados utilizando técnicas de análisis de seguridad y rendimiento que proceden de la rama de la criptografía.

- Es en el capítulo 5 se presenta una aplicación de codificado para imágenes basado en los algoritmos para generar cajas de sustitución diseñados en los capítulos 3 y 4.

- Para finalizar, en el capítulo 6 se muestran las conclusiones que se han obtenido en este trabajo de tesis, así como algunas recomendaciones para futuros trabajos que den continuidad al trabajo iniciado en esta tesis.

Capítulo 2

Sistemas dinámicos y criptografía.

En este capítulo se introducirán algunos conceptos básicos sobre sistemas dinámicos y los criterios que se utilizan para validar que las cajas de sustitución son buenas para criptografía.

2.1. Sistemas dinámicos

Un sistema dinámico puede ser descrito como un conjunto de reglas que determinan la evolución de las variables de estado a partir de un cierto estado inicial, con respecto del tiempo.

Definición 1 *[47] Un sistema dinámico determinista es una terna* $\{T, X, F^t\}$ *tal que:*

1. T *es llamado conjunto tiempo;*
2. X *es llamado espacio de estados* ($X \subset \mathbb{R}^n$)*;*
3. $F^t : X \to X$ *es una familia de operadores de evolución parametrizados por* $t \in T$ *que satisfacen las propiedades*

   a) $F^0 = id$ *donde* id *es la identidad del mapa sobre* X, $idx = x$ *para todo* $x \in X$. *Esta propiedad implica que el sistema no cambia su estado espontáneamente.*

b) *$F^{t+s} = F^t \circ F^s$, significa que $F^{t+s} = F^t(F^s x)$, $\forall x \in X$ así como $t, s \in T$ tal que ambos lados de la ecuación están definidos. Esencialmente esta propiedad establece que el campo vectorial del sistema no varía en el tiempo, con esto se dice que el sistema es autónomo.*

Se acostumbra llamar mapeo a F^t. La aplicación F^t determina la evolución del sistema con respecto al tiempo, propiciando todos los cambios que desarrolle el sistema en un futuro.

Los sistemas dinámicos se clasifican según la naturaleza del conjunto de tiempo T. Diremos que un sistema dinámico es **continuo** cuando su espacio de tiempo sea $\mathbb{R}_{\geq 0}$ y se considera **discreto** cuando su espacio de tiempo T sea $\mathbb{Z}_{\geq 0}$.

2.1.1. Sistemas dinámicos discretos

Los sistemas dinámicos de tiempo discreto se describen por medio de ecuaciones en diferencias

$$x_{n+1} = f(x_n), \tag{2.1}$$

a la expresión anterior se le conoce como relación de recurrencia o función iterativa, donde x es la variable de estado, $n \in \mathbb{Z}_{\geq 0}$ es el tiempo y f es una función $f : \mathbb{R} \to \mathbb{R}$. Si se comienza con un valor inicial x_0 entonces sus iteraciones nos describen la órbita:

$$\{x_i : \quad i = 0 \longrightarrow \infty\} = \{x_0, x_1, x_2, \ldots x_n, x_{n+1}, \ldots\}. \tag{2.2}$$

La órbita más sencilla es la que se mantiene constante, si especificamos esta definición para sistemas discretos tenemos que un punto fijo se define de la siguiente forma:

Definición 2 *[48] Un punto fijo, o punto de periodo uno, es un punto en el cual $x_{n+1} = f(x_n) = x_n$ para todo n.*

De forma gráfica se dibujan x_{n+1} contra x_n y la intersección de la diagonal con f es el punto fijo.

Definición 3 *[48] Un punto periódico de periodo N es un punto en el cual* $x_{n+N} = f^N(x_n) = x_n$ *para todo n.*

Es importante notar que la presencia de algún punto fijo estable implica la obtención de órbitas periódicas de periodo 1, mientras que al tener puntos fijos inestables se tiene la posibilidad de tener órbitas de periodo N. Para encontrar puntos periódicos de periodo dos en un mapeo, es necesario encontrar los puntos donde se intersectan $f^2(x)$ con la diagonal, de forma similar para encontrar puntos fijos de periodo tres, se deben de encontrar los puntos de intersección de $f^3(x)$ con la diagonal. El trabajo de Sarkovskii [49] muestra que si un mapeo contiene puntos periódicos de periodo tres, entonces es posible encontrar puntos periódicos de todos los periodos.

Teorema 1 *[48] Sea* x^* *un punto fijo del mapeo* $f(x)$ (2.1)*. El punto fijo es estable si:*

$$\left| \frac{d}{dx} f\left(x^*\right) \right| < 1, \tag{2.3}$$

el punto fijo es inestable si:

$$\left| \frac{d}{dx} f\left(x^*\right) \right| > 1. \tag{2.4}$$

2.1.2. Sistemas caóticos.

El caos es un fenómeno que no es fácil de clasificar o identificar, por esta razón varios autores han hecho un esfuerzo para dar una definición de caos en las trayectorias de un sistema dinámico siendo la de Robert Devaney [50] una de las más acertadas.

Definición 4 *[51] Sea* X *un conjunto. El mapa* $f : X \to X$ *una función continua, se dice que el sistema es caótico en* X *si satisface las siguientes condiciones:*

1. f *es topológicamente transitiva: para cualquier par de conjuntos abiertos no vacíos* $U, W \subset X$*, existe cierto* $m > 0$ *tal que* $f^m(U) \cap W = \emptyset$*.*

2. *los puntos periódicos de* f *son densos en* X*.*

3. *f tiene dependencia sensible a las condiciones iniciales. Esto es, existe un $\varepsilon > 0$ tal que, para cualquier $x \in X$ en cualquier vecindad Z de x, existe un $y \in Z$ & $m \geq 0$ tal que $|f^m(x) - f^m(y)| > \varepsilon$.*

Banks y colaboradores [52], demostraron que la transitividad topológica y la existencia de un conjunto denso de puntos periódicos implican la dependencia sensible de las condiciones iniciales por lo cual esta última propiedad es redundante. Por otro lado Vellekoop y Berglund [51] mostraron que para mapeos continuos, la transitividad implica que el conjunto de órbitas periódicas es denso, con lo cual transitividad implica caos. Existe otra herramienta útil que indica si un sistema dinámico es caótico, la cual es conocida como el exponente de Lyapunov y se basa en la condición de sensibilidad a las condiciones iniciales [53]. Básicamente mide la tasa de divergencia exponencial entre trayectorias cercanas, lo cual indica la dependencia sensitiva del sistema a las condiciones iniciales. Donde d_n es la distancia entre dos órbitas al tiempo n, separadas inicialmente por una distancia d_0.

$$\frac{d_n}{d_0} \equiv e^{\lambda n}, \lambda = \frac{1}{n} ln \left| \frac{d_n}{d_0} \right|. \tag{2.5}$$

Para que el resultado sea significativo, la divergencia exponencial se debe medir en la vecindad de la trayectoria de referencia; por lo tanto, los exponentes de Lyapunov se calculan a cada iteración y se toma la media al cabo de un número N de iteraciones suficientemente grande,

$$\lambda = \lim_{N \longrightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \ln \left| \left(\frac{d_z}{d_0} \right)_n \right|, \tag{2.6}$$

donde d_0 y d_z son las distancias inicial y final en la iteración n, λ es el exponente de Lyapunov.

Teorema 2 *[48] Si el exponente de Lyapunov es positivo $\lambda > 0$, entonces el sistema es caótico; si el exponente de Lyapunov es negativo $\lambda < 0$, entonces la órbita es periódica y cuando el exponente de Lyapunov es cero $\lambda = 0$, ocurre una bifurcación.*

Otra herramienta que se usa para el estudio de sistemas dinámicos son los diagra-

mas de bifurcación, los cuales son una representación gráfica del comportamiento de las órbitas en función de un parámetro. En estos diagramas, dependiendo de lo que se grafique se puede ver el cambio de estabilidad de los puntos de equilibrio, el cambio de periodo en las órbitas y la propiedad que se conoce como cascadas de periodo dos. En la cual puntos de periodo uno se convierten en puntos de periodo dos, posteriormente en puntos de periodo cuatro y así sucesivamente hasta llegar a caos, que es el comportamiento útil para generar secuencias pseudo-aleatorias.

2.1.3. Entropía

La entropía es una medida de desorden que se puede usar para caracterizar una serie de tiempo, definida por Shannon [54], y se expresa de la siguiente manera:

$$H(s) = -\sum_{i=0}^{2^n-1} P(s_i) log_2 P(s_i). \tag{2.7}$$

La información puede verse como una secuencia de símbolos, que forman parte de un alfabeto finito.

Cada elemento del alfabeto utilizado es representado por 8 bits, esto es, con 8 bits puede ser representado cualquier elemento del alfabeto. Así tenemos 2^8 símbolos que componen el alfabeto $S = \{s_0, s_1, \ldots, s_{255}\}$ y también se puede asociar una probabilidad a cada elemento $P(s_0)$, $P(s_1), \ldots, P(s_{255})$.

La entropía de la información está asociada con la frecuencia de aparición de cada símbolo. El símbolo que proporciona más información es el que tiene menos probablilidad de aparecer. Por lo tanto, para maximizar el valor de entropía, todos los elementos deben tener la misma distribucion de probabilidad, el valor máximo de entropía que se puede obtener es 8.

2.1.4. Análisis de fluctuaciones sin tendencia

Es posible determinar la correlación de una serie temporal a través del análisis de fluctuaciones sin tendencia (DFA) que fue desarrollado por Peng y colaboradores [55]. Por lo tanto, la técnica del DFA nos indica si el sistema propuesto genera series de

tiempo no correlacionadas. El DFA es una herramienta importante para la detección de autocorrelaciones de largo alcance en series de tiempo con no estacionariedades. El DFA está basado en un análisis de escalamiento a partir de la teoría de la caminata aleatoria. Las dos ventajas principales del DFA sobre muchos otros métodos son la detección de correlaciones de largo alcance de una señal incrustada en series de tiempo aparentemente no estacionarias, y esta técnica evita la detección espuria de aparentes correlaciones de largo alcance que son un artefacto de no estacionariedad.

El procedimiento de DFA consiste en los siguientes cinco pasos:

* Se calcula la media de la serie de tiempo $\bar{x}$.

* Se integran las series intermedias por intervalos de la serie de tiempo (de tamaño N).

$$y(k) = \sum_{i=1}^{k} [x(i) - \bar{x}]. \tag{2.8}$$

* La serie integrada es dividida en bloques de igual tamaño n. La tendencia local es obtenida por mínimos cuadrados y es removida de cada bloque.

* La raíz de la fluctuación cuadrada promedio de esta serie integrada y sin tendencia es calculada por:

$$f(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [y(k) - y_n(k)]^2}. \tag{2.9}$$

* Las fluctuaciones pueden ser caracterizadas por un exponente de escalamiento η, la pendiente de la recta que relaciona $logF(n)$ con $log\ n$,

$$f_m(n) \sim n^{\eta}. \tag{2.10}$$

El exponente de escalamiento η para distintos regímenes se define de la siguiente manera:

1 Si $\eta < 0.5$, DFA define correlaciones anti-persistentes de ley de potencia de largo alcance.

2 Si $\eta \approx 0.5$, DFA define el ruido blanco (datos no correlacionados).

3 Si $0.5 < \eta < 1$, DFA define el ruido de las correlaciones de la ley de potencia.

4 Si $\eta \approx 1$, DFA define ruido rosa $(1/f)$.

5 Si $\eta > 1$, DFA define un comportamiento no estacionario o ilimitado.

6 Si $\eta \approx 1.5$, DFA define el movimiento Browniano o el ruido Gaussiano.

En el presente trabajo se utilizará el DFA para determinar que las series de tiempo utilizadas son no correlacionadas, esto quiere decir que el valor de exponente de escalamiento requerido es $\eta \approx 0.5$.

2.2. Criterios para cajas de sustitución criptográficamente seguras.

Se ha realizado una recopilación de seis criterios reportados en la literatura para evaluar que las cajas de sustitución sean criptográficamente seguras. Estos criterios son: biyectividad, no linealidad, criterio estricto de avalancha, criterio de independencia de los bits de salida, distribución equiprobable XOR entrada/salida y máxima probabilidad lineal esperada.

2.2.1. Biyectividad

Sea $S(x)$ una caja de sustitución y $\mathbb{F} = \{0, 1\}$, la cual es biyectiva si y sólo si sus funciones Booleanas f_i satisfacen la siguiente condición:

$$wt(a_1 \cdot f_1 \oplus a_2 \cdot f_2 \oplus \cdots \oplus a_n \cdot f_n) = 2^{n-1}, \tag{2.11}$$

donde $a_i \in \mathbb{F}$, $(a_1, a_2, \cdots, a_n) \neq (0, 0, \cdots, 0)$ y $wt(\cdot)$ es el peso de Hamming [56, 57], se garantiza que la caja de sustitución es biyectiva.

2.2.2. No linealidad

Una caja de sustitución cumple con el criterio de no linealidad si la función Booleana f es altamente no lineal, es decir si encontramos la distancia minima entre el conjunto de la funciones de f y las funciones afin siendo el valor máximo que se puede obtener 112.

Definición 5 *[58] La no linealidad de una función Booleana $f : \mathbb{F}^n \to \mathbb{F}$ se denota por*

$$N_f = \min_{l \in A_{w,c}(x)} d_H(f, l), \tag{2.12}$$

donde $A_{w,c}(x)$ es un conjunto de funciones afines, $d_H(f, l)$ es la distancia de Hamming entre f y l.

La distancia mínima entre dos funciones Booleanas se puede describir mediante el espectro de Walsh [59]:

$$\min_{l \in A_{w,c}(x)} d_H(f, l) = 2^{n-1}(1 - 2^{-n} \max_{\omega \in \mathbb{F}^n} |\hat{S}_{(f)}(\omega)|), \tag{2.13}$$

donde el espectro de Walsh de $f(x)$ se define de la siguiente manera:

$$\hat{S}_{(f)}(\omega) = |\hat{F}_f(\omega)|_{\omega \in \mathbb{B}^n} = \sum_{x \in \mathbb{F}^n} (-1)^{f(x) \oplus x \bullet \omega}, \tag{2.14}$$

con $\omega \in \mathbb{F}^n$ y $x \bullet \omega$ es el producto punto entre x y ω

$$x \bullet \omega = x_1 \cdot \omega_1 \oplus \cdots \oplus x_n \cdot \omega_n. \tag{2.15}$$

2.2.3. Estricto criterio de avalancha

Este criterio fue introducido por primera vez por Webster y Tavares [60]. Una función Booleana f satisface el estricto criterio de avalancha (SAC por su sigla en inglés), si cada uno de los bits de salida cambia con una probabilidad de un medio cuando un solo bit de la entrada es cambiado. Entonces, formalmente una función Booleana f satisface SAC, si y sólo si

$$\sum_{x \in \mathbb{F}^n} f(x) \oplus f(x \oplus e_i) = 2^{n-1}, \quad \forall i : 1 \leq i \leq n, \tag{2.16}$$

donde $e_i \in \mathbb{F}^n$ tal que $wt(e_i) = 1$.

2.2.4. Independencia de los bits de salida

El criterio de independencia de bits de salida (BIC por su sigla en inglés) es otro criterio que se debe satisfacer para una caja de sustitución, introducido por Webster y Tavares [60]. Esto significa que todas las variables de avalancha deben ser independientes por pares, para un conjunto dado de vectores de avalancha generados por el cambio de un solo bit del texto plano.

Adam y Tavares introdujeron otro método para medir el BIC para funciones Booleanas, f_i y f_j $(i \neq j)$ son dos bits de salida de la caja de sustitución, este es si $f_i \oplus f_j$ es altamente no lineal y satisface el SAC [57]. Además, $f_i \oplus f_j$ se puede probar calculando la Distancia Dinámica (DD). La DD de una función f se puede definir como:

$$DD(f) = \max_{\substack{d \in \mathbb{F}^n \\ wt(d)=1}} \frac{1}{2} \left| 2^{n-1} - \sum_{x=0}^{2^n-1} f(x) \oplus f(x \oplus d) \right|. \tag{2.17}$$

Si el valor de DD es un entero pequeño y cercano a cero, la función f satisface el SAC de $f_i \oplus f_j$ y por lo tanto satisface el BIC.

2.2.5. Distribución equiprobable XOR entrada/salida

Biham y Shamir [7] introdujeron el criptoánalisis diferencial, el cual ataca a las cajas de sustitución más rapido que el ataque de fuerza bruta. Es deseable que una caja de sustitución tenga uniformidad diferencial. Esto se puede medir con la probabilidad diferencial máxima esperada (MEDP por su sigla en inglés).

La MEDP para un mapa dado S puede calcularse midiendo la resistencia diferencial, la cual se define de la siguiente manera:

$$DP_f = \max_{\Delta x \neq 0, \Delta y} \left(\frac{\#\{x \in \mathbb{F}^n | S(x) \oplus S(x \oplus \Delta x) = \Delta y\}}{2^m} \right), \tag{2.18}$$

donde 2^n es la cardinalidad de todos las posibles valores de entrada (x), Δx y Δy son denominados diferencias de entrada y salida, respectivamente, para S. Por lo tanto, el valor mas pequeño de DP_f proporciona una mejor propiedad criptográfica, esto es, su resistencia al criptoánalisis diferencial.

2.2.6. Probabilidad lineal máxima esperada

La probabilidad lineal máxima esperada (MELP por su sigla en inglés) es el valor máximo de desequilibrio de un evento [13]. Dadas dos máscaras seleccionadas al azar a & b, a se utiliza para calcular la máscara de todos los valores posibles de una entrada x, y se usa b para calcular la máscara de los valores de salida de la caja de sustitución correspondientes. La paridad de la máscara de bits de entrada a es igual a la paridad de los bits de salida de la máscara b. El MELP de una caja de sustitución dada se puede calcular mediante la siguiente ecuación:

$$LP_f = \max_{a,b \in \mathbb{F}^n \setminus \{0\}} \left(2^{-n} \sum_{x \in \mathbb{F}^n} (-1)^{a \cdot x + b \cdot f(x)} \right)^2 . \tag{2.19}$$

Cuanto más cerca esté MELP de cero, mayor será la resistencia al ataque de criptoanálisis lineal.

Capítulo 3

Algoritmo para generar cajas de sustitución.

En este capítulo, se presenta el algoritmo propuesto para generar cajas de sustitución dinámicas, el cual está basado en un generador de números pseudoaleatorios criptográficamente seguro (CSPRNG).

3.1. Análisis del mapeo logístico.

El mapeo logístico es uno de los sistemas unimodales más famosos y estudiado, se puede encontrar una amplia variedad de referencias en la literatura en donde se exponen sus propiedades [48, 61, 62].

El mapa logístico es un modelo demográfico de tiempo discreto análogo a la ecuación logística introducida por Pierre François Verhulst en el siglo XIX, que se describe en la siguiente ecuación diferencial:

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{K}\right), \tag{3.1}$$

donde x es la variable de estado del sistema, r es un parámetro relacionado con la tasa de máximo crecimiento de la población, y K es la llamada capacidad de carga (es decir, la población máxima sostenible). Entonces $x \leq K$, cuando $x = K$ la población deja de crecer. Robert May [63] popularizó esta ecuación diferencial a uno de los siste-

mas dinámicos discretos más famosos, el mapeo logístico, que se define de la siguiente manera:

$$f_\alpha(x_i) = \alpha x_i(1 - x_i), \tag{3.2}$$

donde x_i es la variable de estado del mapeo logístico y α es el parámetro del sistema, $f_\alpha : [0, 1] \to [0, 1]$, para el parámetro de bifurcación $\alpha \in [0, 4]$ y $x_0 \to [0, 1]$. Sin embargo, en el contexto de las matemáticas, los valores del parámetro α no están restringidos al intervalo $[0, 4]$, por lo que matemáticamente es posible considerar valores negativos [64], por está razón, el mapa logístico ahora se estudia en el intervalo $[-2, 0)$ para fines criptográficos. Ahora, se estudia el comportamiento del mapeo en los dos intervalos y se asegura que con $\alpha \in [-2, 4]$ las órbitas no se escapan al infinito para algunas condiciones iniciales. En [64] los autores observaron la dinámica del sistema en estos dos intervalos y encontraron información útil para aplicaciones de acciones, por otro lado, estos intervalos fueron útiles para construir un generador de números pseudo-aleatorios criptográficamente seguro [65].

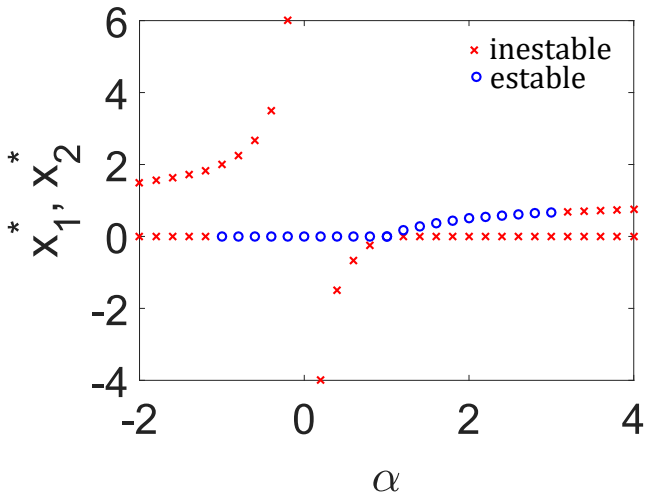


Figura 3.1: Estabilidad de los puntos fijos donde una cruz y un círculo denotan puntos fijos repulsivos y atractivos, respectivamente.

El sistema dinámico (3.2) presenta uno o dos puntos fijos ubicados en $x_1^* = 0$ y en $x_2^* = \frac{\alpha-1}{\alpha}$, para $\alpha \neq 0$. La Figura 3.1 representa la estabilidad de los puntos fijos donde un cruz y un círculo denotan puntos fijos repulsivos y atractivos, respectivamente. Los puntos fijos cambian su estabilidad según el parámetro α, esto es, cuando $|f'_\alpha(x_1^*)| < 1$

y $|f'_\alpha(x_2^*)| < 1$ entonces los puntos fijos x_1^* y x_2^* son estables, respectivamente, y son inestables cuando $|f'_\alpha(x_1^*)| > 1$ y $|f'_\alpha(x_2^*)| > 1$. El interés fue el último caso porque el sistema presenta comportamiento complejo, esto es, que ambos puntos fijos son repulsivos, $|f'_\alpha(x_1^*)| = |\alpha| > 1$ y $|f'_\alpha(x_2^*)| = |-\alpha + 2| > 1$. El punto fijo x_1^* es repulsivo para $\alpha < -1$ y $\alpha > 1$. Por otro lado, el punto fijo x_2^* es repulsivo para $\alpha < 1$ pero $\alpha \neq 0$, y $\alpha > 3$. Por consiguiente los valores de interés son $\alpha \in [-2, -1] \cup [3, 4]$, ésta es la condición para tener ambos puntos fijos repulsivos.

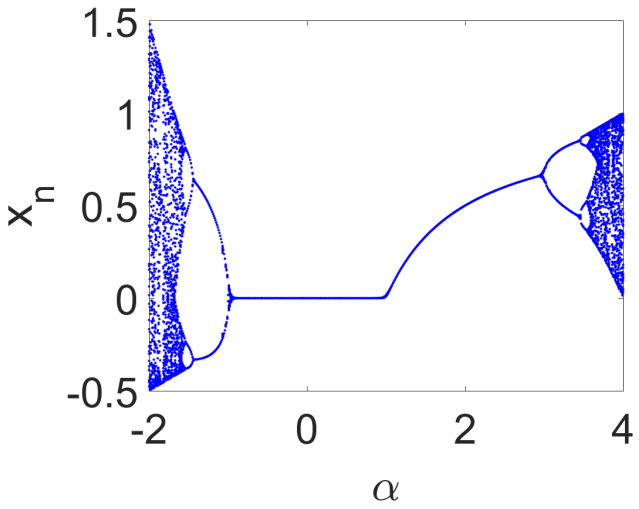


Figura 3.2: Diagrama de bifurcación para el mapeo logístico dado por la Ec. (3.2).

El sistema dinámico (3.2) bifurca cuando $|f'_\alpha(x_1^*)| = 1$ o $|f'_\alpha(x_2^*)| = 1$, esto sucede para x_1^* cuando $\alpha = -1$ ó 1, y para x_2^* los valores de las bifurcaciones se dan con $\alpha = 1$ y 3. Es posible analizar el comportamiento del sistema mediante un diagrama de bifurcación, mostrado en la Figura 3.2. El diagrama muestra los valores que toman las órbitas en función del parámetro α, la ruta hacia el caos son bifurcaciones de duplicación de período en las cuales se observan en $\alpha = 3$ y en $\alpha = -1$. Existen intervalos para el parámetro α cerca de -2 y 4 donde el mapeo logístico $f_\alpha(x)$ se comporta de manera caótica.

Existen varios enfoques para demostrar que un sistema es caótico, uno de ellos demuestra que los sistemas dinámicos cumplen con la definición dada por Devaney [50], otro enfoque se basa en el exponente de Lyapunov [66], [67]. El exponente de Lyapunov de la Ec. (2.6) se muestra en la Figura. 3.3. La gráfica de los exponentes

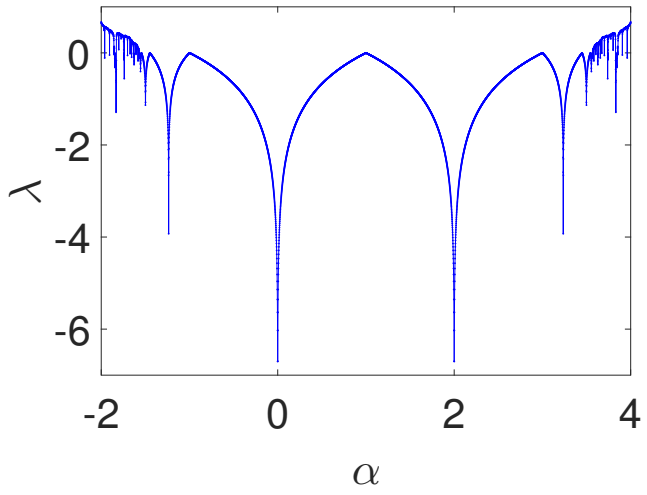


Figura 3.3: Exponente de Lyapunov en función del parámetro α.

de Lyapunov es simétrica con respecto a $\alpha = 1$, el comportamiento caótico del mapeo logístico aparece para valores del parámetro α cerca de -2 y 4. La estabilidad de los puntos fijos está relacionada con los valores del exponente de Lyapunov, por ejemplo, cuando $\alpha \in (-1, 3)$ las órbitas del sistema convergen a un punto fijo, y cuando los puntos fijos son inestables, se producen órbitas periódicas estables que empiezan a bifurcar en periodo dos hasta que aparece el caos.

Se propone utilizar el mapeo logístico para generar series de tiempo con una distribución uniforme, sin evidenciar el mapeo utilizado. Para lograr esto, se propone un enfoque basado en dos series de tiempo caóticas del mapeo logístico. Según el análisis de los exponentes de Lyapunov, los valores de α se seleccionan arbitrariamente dentro de la región del caos, por lo que se consideran -2 y 4.

En la Figura 3.4 se presenta la gráfica del mapeo logístico para ambos valores del parámetro α, -2 y 4, se muestran las curvas en triángulos azules y en cruces negras, respectivamente. El mapeo logístico para estos valores de parámetros es invariante en diferentes intervalos:

$$\begin{aligned} f_{-2} &: [-0.5, 1.5] \to [-0.5, 1.5]; \\ f_4 &: [0, 1] \to [0, 1]. \end{aligned} \tag{3.3}$$

Vale la pena decir que las series de tiempo generadas con ambos valores de paráme-

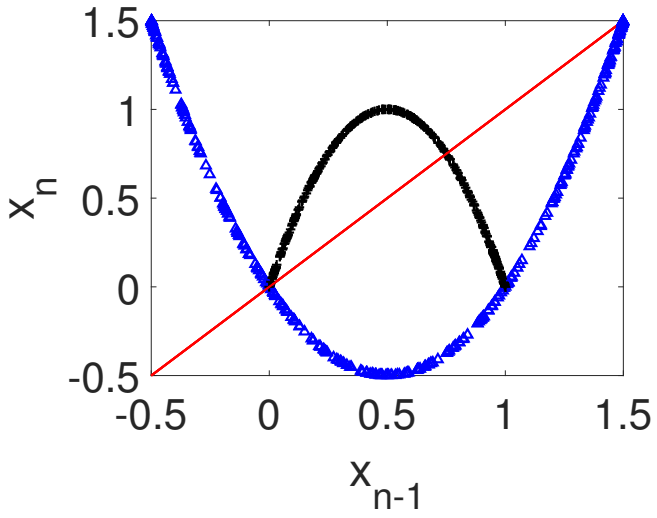


Figura 3.4: Mapeo logístico para valores de $\alpha = -2$ y $\alpha = 4$, en triángulos azules y cruces negras respectivamente.

tros tienen una distribución en forma de U [68, 69].

3.2. Dinámica simbólica

El uso de la dinámica simbólica para analizar series temporales ha sido una herramienta útil para estudiar el comportamiento caótico. En general, un análisis regular por medio de dinámica simbólica consiste en transformar una serie de tiempo en una secuencia de símbolos, conocidos como patrones ordinarios [70], al considerar la relación de orden entre los datos en la serie de tiempo.

Al contar el número de veces que aparece cada patrón ordinario en la secuencia, se pueden estimar las probabilidades de los diferentes patrones, a partir de una secuencia de n elementos, $x_1, \cdots, x_i, \cdots, x_n$, la transformación simbólica cambia las secuencias de n elementos en m símbolos o patrones ordinarios.

Nuestro interés es generar los patrones ordinarios, usando la relación de orden estricta "$<$", en un conjunto dado por tres elementos consecutivos x_i, x_{i+1}, x_{1+2} de la secuencia el cual es llamado análisis de pasos (SWA por su sigla en inglés). Nuestros patrones ordinarios se identifican por los siguientes seis símbolos:

$$
\begin{aligned}
c_1 &: \quad x_i \quad < \quad x_{i+1} \quad < \quad x_{i+2}, \\
c_2 &: \quad x_i \quad < \quad x_{i+2} \quad < \quad x_{i+1}, \\
c_3 &: \quad x_{i+2} \quad < \quad x_i \quad < \quad x_{i+1}, \\
c_4 &: \quad x_{i+2} \quad < \quad x_{i+1} \quad < \quad x_i, \\
c_5 &: \quad x_{i+1} \quad < \quad x_i \quad < \quad x_{i+2}, \\
c_6 &: \quad x_{i+1} \quad < \quad x_{i+2} \quad < \quad x_i.
\end{aligned}
\tag{3.4}
$$

Los símbolos c_i, con $i = 1, \ldots, 6$, se muestran en la Figura 3.5.

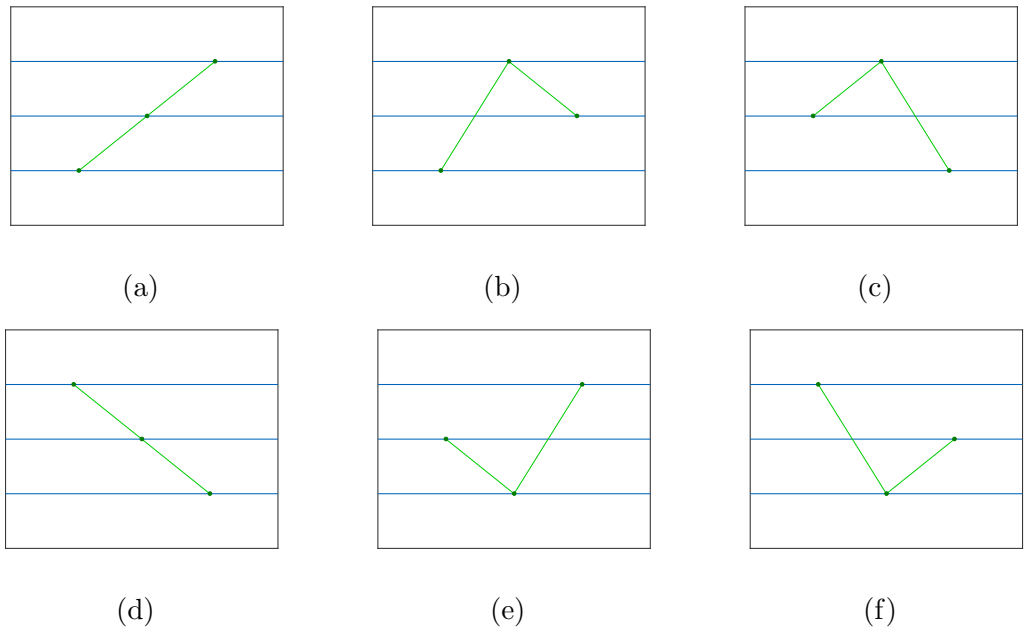
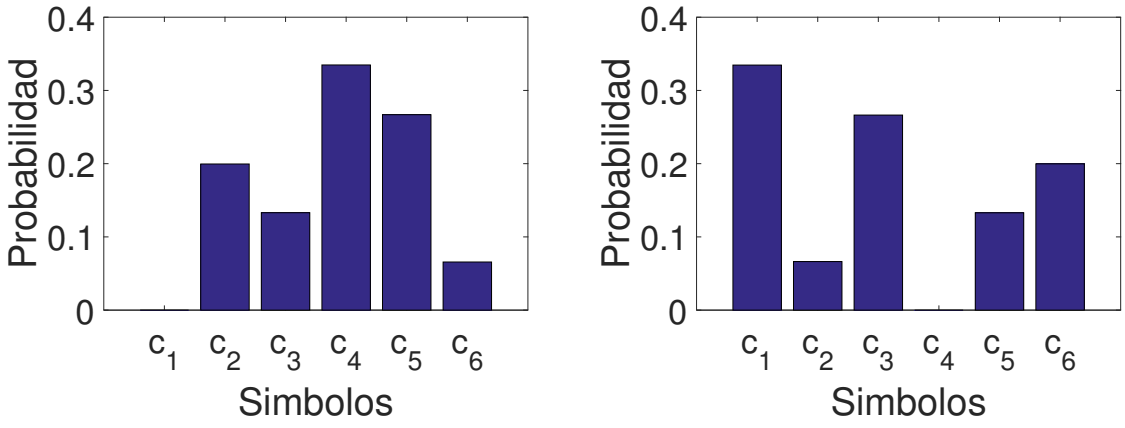


Figura 3.5: (a) $c_1 : x_i < x_{i+1} < x_{i+2}$; (b) $c_2 : x_i < x_{i+2} < x_{i+1}$; (c) $c_3 : x_{i+2} < x_i < x_{i+1}$, (d) $c_4 : x_{i+2} < x_{i+1} < x_i$; (e) $c_5 : x_{i+1} < x_i < x_{i+2}$; (f) $c_6 : x_{i+1} < x_{i+2} < x_i$.

Se analizan los seis patrones ordinarios para dos series temporales caóticas generadas por (3.2) con 177700 elementos y los valores de parámetro $\alpha = -2$ y $\alpha = 4$. El análisis muestra que para cada parámetro hay un símbolo que no ocurre, específicamente, los símbolos son c_1 y c_4 para $\alpha = -2$ y $\alpha = 4$, respectivamente Figura 3.6.

Además, la distribución de probabilidad de estos seis patrones ordinarios no es uniforme, para el caso de $\alpha = -2$ se tiene que la probabilidad de cada símbolo es $c_1 = 0$, $c_2 = 0.19$, $c_3 = 0.13$, $c_4 = 0.33$, $c_5 = 0.26$ y $c_6 = 0.06$. Observe que el símbolo c_1 nunca ocurre, esto implica que la secuencia no contiene tres puntos consecutivos tales

que $x_i < x_{i+1} < x_{i+2}$, Figura 3.6 (a) . Para $\alpha = 4$ las probabilidades obtenidas son $c_1 = 0.33$, $c_2 = 0.06$, $c_3 = 0.26$, $c_4 = 0$, $c_5 = 0.13$ y $c_6 = 0.19$. En este caso el símbolo c_4 nunca ocurre, por lo que la secuencia no contiene tres puntos consecutivos tales que $x_{i+2} < x_{i+1} < x_i$, Figura 3.6 (b). Adicionalmente, se calculó la entropía (sección 2.1.3) de una serie temporal generada con el mapa logístico cuyo valor obtenido es de 7.72. Estas características no son útiles para propósitos criptográficos ya que revelan un patrón interno sobre el sistema que se utiliza.



(a) (b)

Figura 3.6: Probabilidades de ocurrencia de c_i para series de tiempo con (a) $\alpha = -2$ y (b) $\alpha = 4$, respectivamente.

El exponente de escalamiento para las órbitas obtenidas con el mapeo logístico es $\eta \approx 0.5$, por lo que el DFA indica que la serie temporal del mapa logístico presenta propiedades de ruido blanco.

3.3. Generador de números pseudoaleatorios criptográficamente seguro (CSPRNG)

Considérese una órbita $x_0, x_1, x_2, \ldots$ del mapeo logístico (3.2) con condición inicial $x_0 \in I$, donde I está determinado por el parámetro $\alpha \in \{-2, 4\}$, $f_\alpha : I \to I$. Sean $M1$ y $M2$ dos series de tiempo generadas con el mapeo logístico mediante las siguientes consideraciones: *i)* dadas dos condiciones iniciales arbitrarias x_{01}, x_{02}, de manera tal

que, $x_{01} \neq x_{02}$; *ii)* dos valores del parámetro de bifurcación diferentes α_1 y α_2; y *iii)* l -unidades de memoria para cada serie de tiempo $x_{(i-k_{l-1})1}$, ..., $x_{(i-k_2)1}, x_{(i-k_1)1}, x_{i1}$ y $x_{(i-k_{l-1})1}$, ..., $x_{(i-k'_2)2}, x_{(i-k'_1)2}, x_{i2}$, de tal forma que las órbitas tienen una distribución uniforme independientemente de la distribución en forma de U del mapeo logístico. Con el fin de ilustrar el algoritmo, se han elegido valores de parámetros de bifurcación como $\alpha_1 = -2$ y $\alpha_2 = 4$ para las series de tiempo $M1$ y $M2$, respectivamente. Estos valores de parámetros aseguran que el sistema (3.2) tenga un comportamiento caótico en ambos casos.

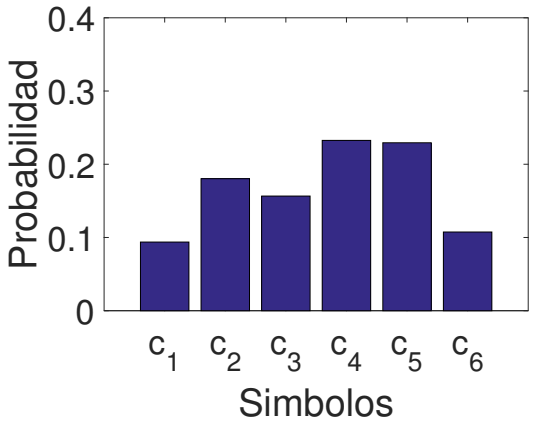


Figura 3.7: Probabilidad de los símbolos de $M_1 + M_2$.

El SWA aplicado a $M_1 + M_2$ se muestra en la Figura 3.7, y señala que ahora todos los símbolos c_i son posibles. Sin embargo, la ocurrencia de los símbolos no tiene una distribución uniforme. El valor de entropía calculado para esta serie mixta es $H = 4.36$ eso muestra una gran disminución en comparación con la serie de tiempo única del mapa logístico. El exponente de escalamiento del DFA $\eta = 0.419$ indica la aparición de un tipo de correlación, cabe recordar que se busca en criptografía tener series de tiempo des-corre . Estos resultados muestran que las series mixtas que utilizan dos series temporales del mapa logístico no cumplen los requisitos y es necesario considerar características adicionales.

Para garantizar que el generador presente buenas propiedades estadísticas, es necesario generar series de tiempo con distribución uniforme y también es deseable eliminar la forma del mapeo logístico en estas nuevas series de tiempo. Esto se logra mediante

el número de retardos involucrados. Existen muchas combinaciones de retardos que pueden des-correlacionar la forma del mapeo logístico y las series de tiempo, pero cada unidad de retardo necesita memoria y tiempo de procesamiento. Ahora se analiza la serie de tiempo $M2 = m_{02}, m_{12}, m_{22}, \ldots$ con dos unidades de memoria, para $\alpha = 4$, los elementos m_{i2} de la serie de tiempo $M2(x_{(i-k1)2}, x_{i2})$ se obtienen de la siguiente manera:

$$m_{i2} = M2(x_{(i-k1)2}, x_{i2}) = x_{(i-k1)2} + x_{i2}, \, mod \; 1, \tag{3.5}$$

donde $k_1 = 5$. Al graficar $m_{(n-1)2}$ contra m_{n2}, Figura 3.8, es posible distinguir la forma del mapeo logístico. Ya que la longitud del retardo no es importante, la forma del mapa logístico siempre permanece, por lo que es necesario considerar más unidades de memoria.

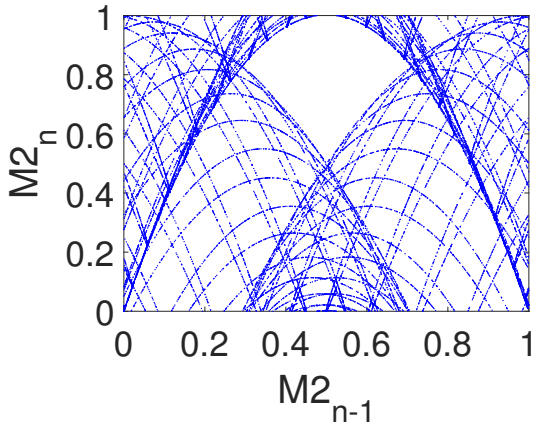


Figura 3.8: Mapa del primer retorno de la serie temporal $M2(x_{(i-k1)2}, x_{i2})$ considerando dos unidades de memoria.

Ahora consideramos tres unidades de memoria para obtener los elementos de la serie de tiempo $M2(x_{(i-k2)2}, x_{(i-k1)2}, x_{i2})$ dado lo siguiente:

$$m_{i2} = M2(x_{(i-k2)2}, x_{(i-k1)2}, x_{i2}) = x_{(i-k2)2} + x_{(i-k1)2} + x_{i2}, \, mod \; 1, \tag{3.6}$$

donde $k_1 = 10$ y $k_2 = 5$. Ahora, para este caso de tres unidades de memoria, que son la cantidad mínima para obtener una nube de puntos en $(m_{(n-1)2}, m_{n2})$, Figura 3.9.

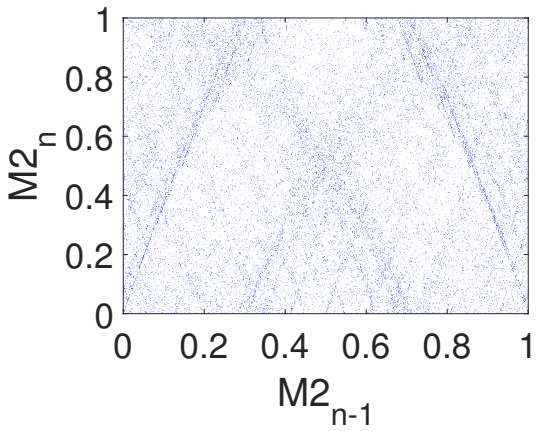


Figura 3.9: Mapa del primer retorno de la serie temporal $M2(x_{(i-k2)2}, x_{(i-k1)2}, x_{i2})$ considerando tres unidades de memoria.

La forma del mapeo logístico casi desaparece, por lo que tres unidades de memoria son suficientes. El problema al considerar más unidades de memoria, tiene un precio computacional de almacenamiento de información. Por este motivo, se utilizan dos retardos k_1, k_2 y el estado actual de la serie temporal del mapeo logístico. Además, los retardos no deben ser contiguos para evitar patrones regulares que afecten directamente los resultados de la prueba.

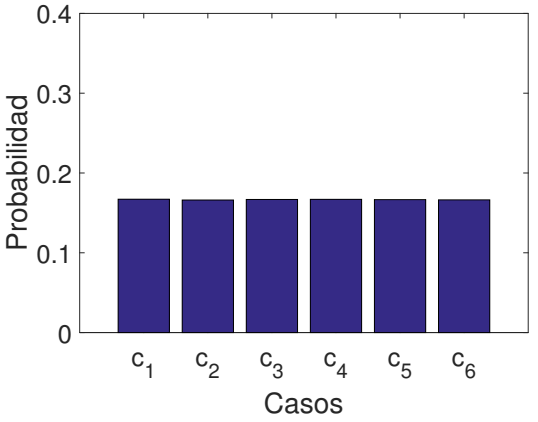


Figura 3.10: Probabilidad de ocurrencia de los símbolos definidos por el SWA para una serie de tiempo con retardos Ec.(3.6).

Ahora se aplica el SWA a la serie m_{i2} y el resultado se muestra en la Figura 3.10, lo que señala que todos los símbolos son posibles. Adicionalmente, los símbolos tienen una buena aproximación a una distribución uniforme. El valor de entropía obtenido es

$H = 7.995$, el cual muestra un incremento en comparación con las series mixtas sin considerar los retrasos de tiempo.

El exponente de escalamiento obtenido por el DFA es $\eta = 0.4965$ que es un valor cercano al valor deseado 0.5. Estos resultados muestran, que al considerar los retrasos están más cerca de los requisitos necesarios y se pueden realizar para obtener una distribución uniforme en SWA.

Se consideran diferentes retardos, esto es, $k_2 = k_2' = 10$, $k_1 = 6$ y $k_1' = 5$ para ambas series de tiempo $M1$ y $M2$. por lo tanto, estas series están formadas por la suma de dos estados de retardo $x_{(i-10)1}$ & $x_{(i-5)1}$ y el estado actual x_{i1} de la órbita $x_{01}, x_{11}, x_{21}, \ldots$, en $M1$. De la misma manera para $M2$, $x_{(i-10)2}$, $x_{(i-6)2}$ y x_{i2} de la órbita $x_{02}, x_{12}, x_{22}, \ldots$. Los valores de la serie de tiempo están limitados por la operación *mod 1*, lo que garantiza que $M1, M2 \in [0, 1) \subset \mathbb{R}$. Explícitamente $M1(x_{(i-10)1}, x_{(i-5)1}, x_{i1})$ y $M2(x_{(i-10)2}, x_{(i-6)2}, x_{i2})$ se expresan de la siguiente manera:

$$m_{i1} = M1(x_{(i-10)1}, x_{(i-5)1}, x_{i1}) = x_{(i-10)1} + x_{(i-5)1} + x_{i1}, \ \textit{mod 1}, \tag{3.7}$$

$$m_{i2} = M2(x_{(i-10)2}, x_{(i-6)2}, x_{i2}) = x_{(i-10)2} + x_{(i-6)2} + x_{i2}, \ \textit{mod 1}. \tag{3.8}$$

Finalmente, estas series de tiempo $M1 = m_{01}, m_{11}, m_{21}, \ldots$ y $M2 = m_{02}, m_{12}, m_{22}, \ldots$ dadas por (3.7) y (3.8), respectivamente, se mezclan y la operación *mod 1* es aplicada nuevamente, este proceso genera una nueva serie de tiempo Z_i dada como sigue:

$$Z_i = m_{i1} + m_{i2}, \ \textit{mod 1}. \tag{3.9}$$

A partir de ahora, Ec. (3.9) se conoce como mapeo con retardo para el cual $Z_i \in [0, 1) \subset \mathbb{R}$.

El objetivo de utilizar este enfoque es que con la combinación de dos series de tiempo con retardo representados por Z_i, es posible descartar la forma del mapeo caótico utilizado. Por ejemplo, la serie de tiempo x_n puede revelar el mapeo si x_n contra x_{n-1} se representa gráficamente, como se muestra en la Figura 3.11 (a), el mapeo logístico aparece. A diferencia de la serie de tiempo z_n, no puede revelar el mapa si z_n

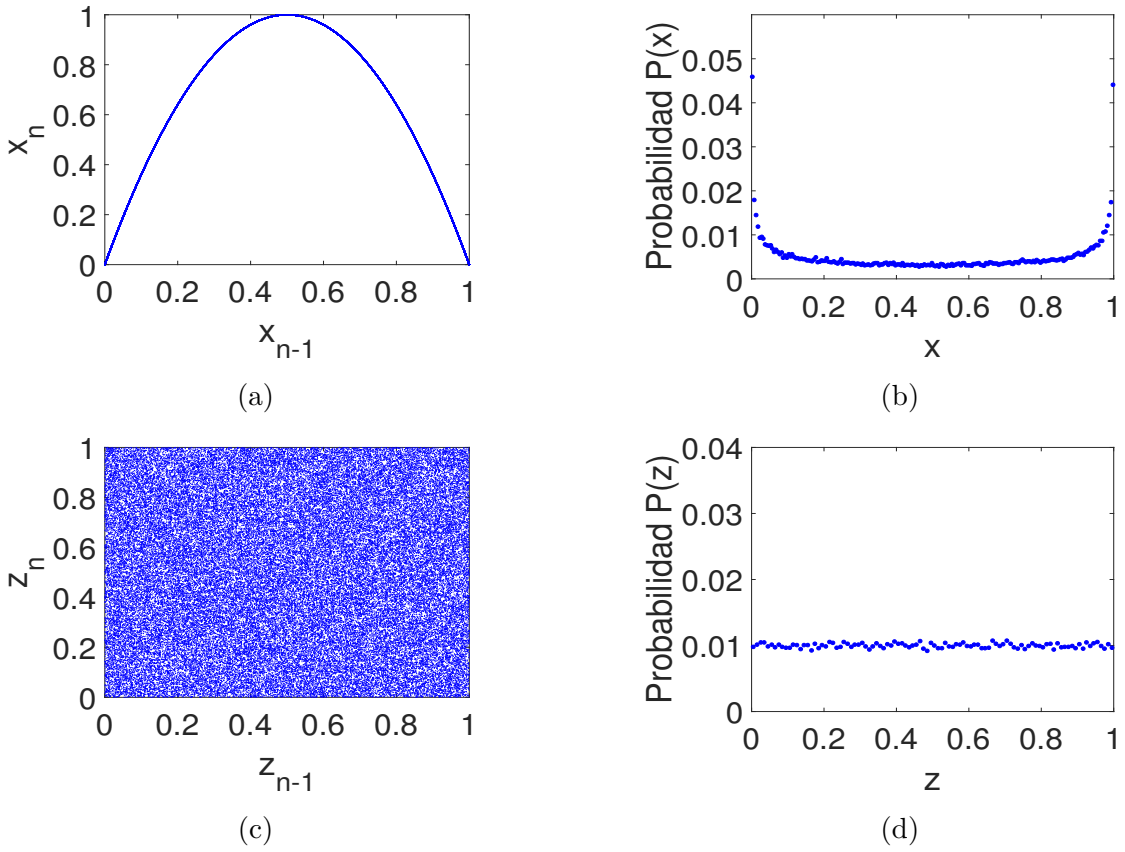


Figura 3.11: (a) Mapeo Logístico dado por x_n contra x_{n-1}; (b) Distribución de probabilidad en "forma de U" del mapeo logístico; (c) Mapeo con retardo dado por z_n contra z_{n-1}; (d) Distribución de probabilidad uniforme del mapeo con retardo.

contra z_{n-1} se representa gráficamente (Figura 3.11 (c)), los retardos utilizados no son revelados, y tampoco aparece la forma del mapeo logístico. Además, esto nos permite cambiar la característica de la distribución de probabilidad en "forma de U" [68] por una distribución de probabilidad uniforme en las series de tiempo obtenidas x_n y z_n, ver la Figura 3.11 (b) y (d), respectivamente. Ésta es una característica importante, en comparación con los esquemas basados en caos, ya que facilita la construcción de las cajas de sustitución porque todos los valores tienen la misma probabilidad de ocurrencia a diferencia con los esquemas basados en sistemas caóticos simples.

El SWA aplicado a Z se muestra en la Figura 3.12, y señala que ahora todos los símbolos son posibles y la aparición de los símbolos muestra una distribución uniforme. El valor de entropía calculado para este caso fue $H = 7.997$. La ley de escala revelada por el DFA $\eta = 0.4985$ indica un valor más cercano a 0.5, que es el valor deseado.

Estos resultados muestran que al considerar series de retardos mixtos, los requisitos se cumplen y podrían utilizarse para generar cajas de sustitución.

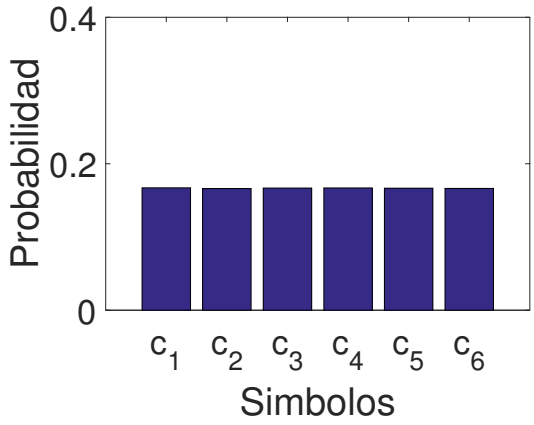


Figura 3.12: Probabilidad de ocurrencia de los símbolos definidos por el SWA para una serie de tiempo con retardos Ec.(3.6).

Para obtener una serie de tiempo binaria (s) útil para sistemas criptográficos, se construye la dinámica simbólica de Z_i. Así los elementos de s son números binarios, esto es, $s_i(Z_i) \in \{0, 1\}$. Un requisito necesario para la dinámica simbólica es obtener ceros y unos con la misma probabilidad, por lo tanto, el proceso para obtener las series binarias es como sigue:

$$s_i = \begin{cases} 0, & \text{for } \ 0 < Z_i \leq 0.5, \\ 1, & \text{for } \ 0.5 < Z_i < 1. \end{cases} \tag{3.10}$$

En la Figura 3.13 se muestra el diagrama del generador propuesto.

3.4. Algoritmo para generar cajas de sustitución vía CSPRNG.

El algoritmo propuesto para la creación de cajas de sustitución de $n \times n$ [71] es descrito en los siguientes pasos:

Paso 1 Seleccione las condiciones iniciales x_{01} y x_{02} para CSPRNG para generar un flujo de bits $s_0, s_1, s_2, \ldots$

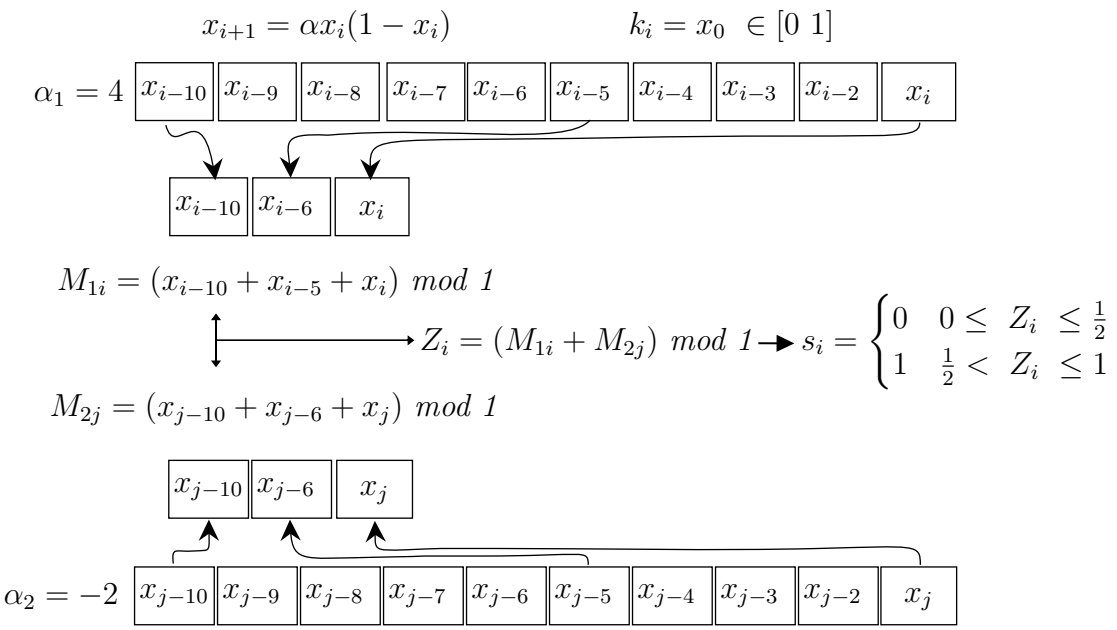


Figura 3.13: Diagrama del generador basado en retardos.

Paso 2 Genere la secuencia de bloques de n-bits cada una, $C_0 = (s_0, s_1, \ldots, s_{n-1})$, $C_1 = (s_n, s_{n+1}, \ldots, s_{2n-1})$, $C_2 = (s_{2n}, s_{2n+1}, \ldots, s_{3n-1})$, ...

Paso 3 Convierta los bloques $C_0, C_1, C_2, \ldots$ de n-bits a números enteros D_0, D_1, D_2,...

Paso 4 Deseche los elementos repetidos D's para seleccionar 2^n valores diferentes. La regla para descartar un elemento es la siguiente: si $D_i = D_j$ con $i < j$ entonces descartar D_j.

Paso 5 Crear la caja de sustitución con 2^n elementos diferentes de D's.

Una vez que el procedimiento ha finalizado, el algoritmo propuesto devuelve una $n \times n$ caja de sustitución con 2^n valores distintos. Nótese que D_0 es el primer elemento de la caja de sustitución, pero el segundo elemento podría no ser D_1 si $D_0 = D_1$, sin embargo, se han generado los suficientes 2^n elementos para construir la caja de sustitución. Cada bloque C's está compuesto por n bits, $s_j, s_{j+1}, \ldots, s_{j+n-1}$, que están relacionadas con las funciones f_i, con $i = 1, \ldots, n$.

Para ilustrar se muestra el siguiente **ejemplo**, para $n = 8$, $x_{01} = 0.8147$, $x_{02} = 0.9058$, $\alpha_1 = 4$ y $\alpha_2 = -2$ se obtiene la caja de sustitución de 8×8 presentada en

la siguiente Tabla 3.1. Esta caja de sustitución propuesta tiene las propiedades de confusión y difusión, que son de vital importancia para los cifrados en bloque.

64	46	150	174	220	26	233	224	148	170	143	247	225	212	90	124
44	204	59	61	43	121	129	2	109	164	103	249	16	237	27	35
216	184	81	213	161	169	89	199	140	38	239	48	163	193	21	147
222	217	70	196	195	192	234	41	47	15	14	42	98	190	186	36
242	51	60	87	24	104	189	55	118	111	231	120	8	226	7	141
85	9	73	101	3	197	12	66	82	110	65	25	165	176	80	181
125	31	218	74	68	52	149	95	182	19	112	5	136	79	214	34
158	50	188	137	28	191	155	84	105	126	92	179	162	152	200	0
171	142	240	203	88	160	32	202	99	18	100	97	145	53	194	93
245	119	185	20	235	123	134	139	128	116	173	76	17	132	209	135
83	168	57	56	223	30	91	4	22	122	102	221	208	131	71	86
39	114	252	10	172	201	177	77	94	246	54	175	183	108	156	45
219	210	40	130	113	153	13	166	58	23	253	215	238	33	198	248
229	227	96	206	107	144	67	254	115	167	244	106	180	157	255	241
207	243	228	187	49	78	251	37	62	1	205	117	29	178	75	236
11	250	146	6	151	69	138	133	72	232	211	127	159	63	154	230

Tabla 3.1: Caja de sustitución generada por el algoritmo propuesto.

3.5. Desempeño de las cajas de sustitución

En esta sección, se examina el desempeño de las cajas de sustitución obtenidas para confirmar su inmunidad, especialmente contra el criptoanálisis diferencial y lineal, para ello seis criterios criptográficos importantes y conocidos de las cajas de sustitución de 8×8 son calculados.

3.5.1. Biyectividad

El valor calculado de la caja de sustitución propuesta es $2^{n-1} = 128$ que es el valor deseado, con $n = 8$, de acuerdo con la fórmula (2.11). Por lo tanto, el criterio de biyectividad se cumple, es decir, la caja de sustitución propuesta es uno a uno, sobreyectiva y balanceada; que es uno de los criterios primarios en criptografía.

3.5.2. No linealidad

La no linealidad es uno de los principales criterios para el diseño de cualquier caja de sustitución, que garantiza que la caja de sustitución propuesta no sea una función lineal entre los vectores de entrada y los vectores de salida. La no linealidad simboliza el grado de disimilitud entre la función Booleana f y las funciones afines l. Si la distancia de Hamming es mínima entre las funciones Booleanas, se dice que tiene una no linealidad alta, esto es, mediante la reducción del espectro Walsh en (2.13). Una caja de sustitución contiene n funciones Booleanas y la no linealidad de cada función Booleana debe calcularse. Las no linealidades de la caja de sustitución propuesta son las siguientes: 104, 104, 102, 104, 96, 102, 100 y 102, respectivamente. La alta no linealidad garantiza la capacidad de resistir ataques poderosos como el criptoanálisis diferencial.

3.5.3. Estricto criterio de avalancha

El estricto criterio de avalancha (SAC) o también llamado efecto de avalancha se utiliza para indicar la aleatoriedad de una caja de sustitución cuando una entrada tiene un cambio. Los resultados obtenidos de este criterio son mostrados en la Tabla 3.2. En la caja de sustitución propuesta, se obtiene un SAC máximo de 0.5781, el mínimo es 0.3906, y el valor promedio de 0.5012 que es cercano al valor deseado de 0.5. Basados en estos resultados, se puede concluir que la caja de sustitución generada por el método propuesto cumple con la propiedad de SAC.

0.5781	0.4844	0.5000	0.4219	0.4844	0.5156	0.4063	0.5469
0.5156	0.5000	0.4688	0.5156	0.5469	0.3906	0.5469	0.4375
0.5469	0.5000	0.5000	0.5469	0.4063	0.5156	0.4531	0.5313
0.4531	0.5156	0.5000	0.4531	0.5313	0.5313	0.4844	0.4688
0.5156	0.5469	0.4844	0.5313	0.5313	0.5625	0.5625	0.5469
0.4063	0.4844	0.5000	0.4063	0.5625	0.5625	0.4844	0.5313
0.4219	0.4063	0.5313	0.5313	0.4219	0.5625	0.4844	0.4844
0.5469	0.5156	0.5469	0.5625	0.4531	0.5625	0.5781	0.4531

Tabla 3.2: Resultado del criterio SAC de la caja de sustitución propuesta.

3.5.4. Independencia de bits de salida

El criterio BIC garantiza que no exista un patrón estadístico o dependencia entre los vectores de salida. El BIC de la caja de sustitución es calculado mediante el método que se describe en la Subsección 2.2.4, los resultados obtenidos se muestran en las Tablas 3.3, 3.4 y 3.5. El valor promedio de BIC- No linealidad es 103.8571, el valor promedio de BIC-SAC es 0.5066 y el máximo valor de DD es 8, lo cual indica que la caja de sustitución satisface el BIC.

0	104	104	106	104	106	106	102
104	0	106	98	102	104	102	104
104	106	0	104	102	96	104	104
106	98	104	0	106	100	106	104
104	102	102	106	0	102	100	102
106	104	96	100	102	0	104	108
106	102	104	106	100	104	0	106
102	104	104	104	102	108	106	0

Tabla 3.3: Resultado del criterio de BIC-No linealidad de la caja de sustitución generada.

0	0.5020	0.5176	0.5137	0.5293	0.5098	0.4727	0.5059
0.5020	0	0.4980	0.4844	0.5039	0.5313	0.5156	0.5000
0.5176	0.4980	0	0.5039	0.4941	0.5313	0.5000	0.5020
0.5137	0.4844	0.5039	0	0.5117	0.4980	0.5020	0.5020
0.5293	0.5039	0.4941	0.5117	0	0.5234	0.5000	0.5137
0.5098	0.5313	0.5313	0.4980	0.5234	0	0.5039	0.5000
0.4727	0.5156	0.5000	0.5020	0.5000	0.5039	0	0.5156
0.5059	0.5000	0.5020	0.5020	0.5137	0.5000	0.5156	0

Tabla 3.4: Resultado del criterio de BIC-SAC de la caja de sustitución generada.

0	2	6	2	4	6	4	6
2	0	2	2	2	2	4	2
6	2	0	6	8	2	6	4
2	2	6	0	4	2	8	4
4	2	8	4	0	2	0	2
6	2	2	2	2	0	2	8
4	4	6	8	0	2	0	0
6	2	4	4	2	8	0	0

Tabla 3.5: La DD de la caja de sustitución generada (Criterio BIC–SAC).

4	3	3	4	3	3	3	3	4	3	3	3	3	3	4	4
3	3	4	3	3	4	3	3	4	3	3	4	4	4	4	3
4	3	4	3	4	4	3	3	3	3	3	3	3	4	3	3
4	3	3	3	4	4	4	4	3	4	5	4	3	2	3	3
5	4	4	3	3	3	4	4	4	3	5	3	3	3	3	3
3	3	3	4	4	3	5	4	3	3	3	5	5	3	3	3
3	3	3	3	3	4	4	3	3	3	4	3	3	2	3	3
3	2	3	3	3	4	3	3	3	3	3	4	3	3	3	3
3	3	3	5	5	3	3	4	3	4	3	2	5	3	3	3
3	3	3	4	3	4	3	3	3	4	3	3	4	3	4	3
4	3	4	3	2	3	3	4	3	3	3	3	3	4	3	3
3	4	3	3	3	3	3	3	3	4	3	3	3	3	4	4
3	3	3	3	3	4	3	3	2	4	3	3	4	4	3	3
4	3	4	3	4	4	3	4	4	3	4	4	3	3	3	3
3	4	3	3	3	3	3	3	3	4	4	3	3	3	3	3
3	3	3	5	4	5	4	3	3	5	3	3	4	3	5	/

Tabla 3.6: Criterio de distribución equiprobable XOR entrada/salida para la caja de sustitución generada.

3.5.5. Distribución equiprobable XOR entrada/salida

El criterio de distribución equiprobable XOR entrada/salida analiza las diferencias entre los pares de entrada con las diferencias entre los pares de salida para descubrir los bits de la llave. La idea es que la diferencia de entrada tenga la misma probabilidad diferencial de salida para resistir el criptoánalisis diferencial. El criterio de la caja de sustitución es calculado por la Ec.(2.18) y los resultados son presentados en la Tabla 3.6. El valor máximo de la caja de sustitución generada por el método propuesto es 5, lo cual indica que la caja de sustitución satisface el criterio de distribución equiprobable XOR entrada/salida.

3.5.6. Probabilidad lineal máxima esperada

El valor del criterio de MELP es calculado a través de aproximaciones lineales de la función no lineal. El objetivo final es recuperar los bits de la llave. MELP estudia la correlación entre los bits de entrada y salida. Este criterio es calculado por la Ec.(2.18) y el valor promedio es 0.0176. Por último, presentamos nuestros resultados que contrastan con algunos resultados presentados en diferentes artículos publicados que utilizan otros enfoques.

	Biyectividad	No linealidad			SAC			BIC			I/O XOR	MELP
		mín.	máx.	prom.	mín.	máx.	prom.	SAC	No linealidad	DD		
Skipjack [79]	123	100	108	105.12	0.3906	0.5938	0.5027	0.5003	104.03	109	0.0469	0.0137
APA [72]	128	112	112	112	0.4375	0.5625	0.5007	0.4997	112	112	0.0156	0.0039
Gray [73]	128	112	112	112	0.4375	0.5625	0.4998	0.5026	112	112	0.0156	0.0039
AES [74]	128	112	112	112	0.4531	0.5625	0.5049	0.5046	112	112	0.0156	0.0039
Ref. [10]	128	98	107	103.25	0.3828	0.5938	0.5059	0.5033	104.21	108	0.0469	0.0166
Ref. [30]	129	103	109	104.87	0.3984	0.5703	0.4966	0.5044	102.96	109	0.0391	0.0176
Ref. [75]	128	96	106	103	0.3906	0.6250	0.5039	0.5010	100.35	106	0.5000	0.0220
Ref. [76]	128	112	112	112	0.4219	0.5469	0.5115	0.4982	108.71	112	0.0313	0.0120
Ref. [14]	128	102	108	105.25	0.4375	0.5781	0.5056	0.5019	103.78	108	0.0391	0.0244
Ref. [17]	128	104	110	106.25	0.4219	0.5938	0.5039	0.5059	103.35	108	0.0391	0.0198
Ref. [77]	128	102	108	106	0.4219	0.5938	0.5002	0.5016	104.42	108	0.0391	0.0220
Ref. [78]-1	128	106	108	106.75	0.3906	0.6094	0.4941	0.5013	104.28	108	0.0391	0.0156
Ref. [78]-2	128	106	108	106.75	0.4063	0.5938	0.4971	0.5008	102.92	106	0.0391	0.0198
Propuesta 1	128	96	104	101.75	0.3906	0.5781	0.5012	0.5066	103.42	108	0.0391	0.0176

Tabla 3.7: Comparación de cajas de sustitución basadas en caos y cajas de sustitución usadas en cifrados en bloque tradicionales.

3.6. Resultados comparativos

En esta sección se presenta una comparativa del desempeño de la caja de sustitución creada con el algoritmo descrito en la Subsección 3.4. En la Tabla 3.7 se muestran los valores de los criterios para nuestra caja de sustitución y el conjunto de cajas de sustitución reportadas en la literatura (cajas de sustitución estándar y basadas en caos). En esta Tabla 3.7, se puede observar que la caja de sustitución generada cumple con el criterio más importante que es la biyectividad y logra una buena aproximación con el resto de los criterios [10, 14, 17, 30, 72–78]. Principalmente muestra mejor desempeño en las pruebas relacionadas con ataques (MELP y distribución equiprobable XOR entrada/salida). Además es una metodología con base en un sistema con operaciones simples que genera secuencias con comportamiento complejo.

Capítulo 4

Algoritmo para la generación de cajas de sustitución basado en el mapeo logístico extendido

En este capítulo se presenta una generalización al mapeo logístico, para ampliar su rango y dominio, así como su análisis dinámico. Esto con la finalidad de modificar el algoritmo para la generación de cajas de sustitución propuesto en el capitulo 3 y mejorar su desempeño.

4.1. Análisis del mapeo logístico extendido

El mapeo logístico extendido es definido por:

$$f_{\alpha e}(x_i) = \alpha x_i(2^n - x_i), \tag{4.1}$$

donde x es la variable de estado, α y n son los parámetros del sistema. Esta estructura permite la modificación del tamaño del dominio y rango de la función, esto es, $f_{\alpha e} : [0, 2^n] \rightarrow [0, 2^n]$, para el parámetro de bifurcación $\alpha \in [0, \frac{4}{2^n}]$ y $x_0 \rightarrow [0, 255]$. Sin

embargo, como en el caso del mapeo logístico normalizado, los valores del parámetro α no están restringidos para valores positivos [64]. Ahora, se estudia el comportamiento del mapeo en el intervalo $\alpha \in [\frac{-2}{2^n}, \frac{4}{2^n}]$ en el que se asegura que las órbitas no se escapan al infinito para algunas condiciones iniciales.

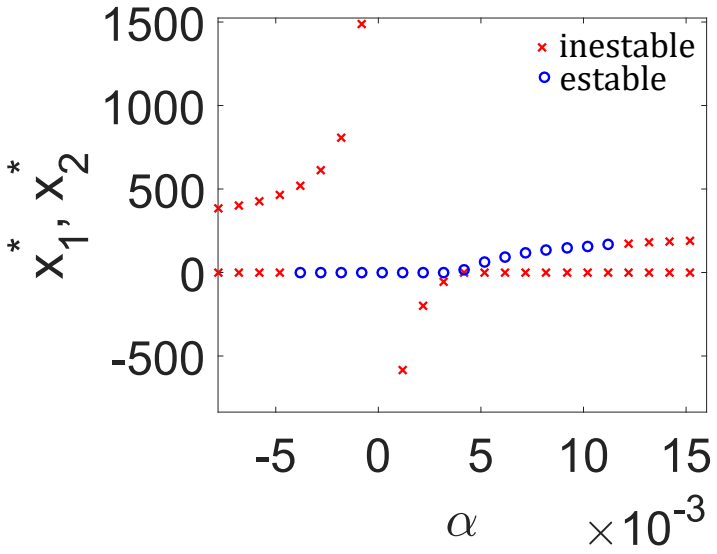


Figura 4.1: Estabilidad de los puntos fijos donde la cruz y el círculo denotan puntos fijos repulsivos y atractivos, respectivamente.

Al ser solo una modificación en el tamaño del mapeo las propiedades dinámicas del sistema se preservan. Esto puede ser observado en el siguiente ejemplo para el caso de $n = 8$, en la Figura 4.1 se muestra la estabilidad de los puntos fijos; donde una cruz y un círculo denotan puntos fijos repulsivos y atractivos, respectivamente.

En la Figura 4.2 el diagrama muestra los valores que toman las órbitas en función del parámetro α y la ruta hacia el caos son bifurcaciones de duplicación de período, al igual que en el caso del mapeo logístico normalizado. Y finalmente en la Figura 4.3, se presenta el cálculo del exponente de Lyapunov, el comportamiento caótico del mapeo logístico extendido aparece para valores del parámetro α cerca de $\frac{-2}{2^8}$ y $\frac{4}{2^8}$.

El objetivo es utilizar el mapeo logístico extendido para generar series de tiempo con distribución de probabilidad uniforme, sin evidenciar el mapeo utilizado. Basados en el análisis realizado en la sección 3.3, se propone un enfoque basado en dos series de tiempo caóticas del mapeo logístico extendido. Según el análisis de los exponentes de Lyapunov, los valores de α se seleccionan arbitrariamente dentro de la región del caos,

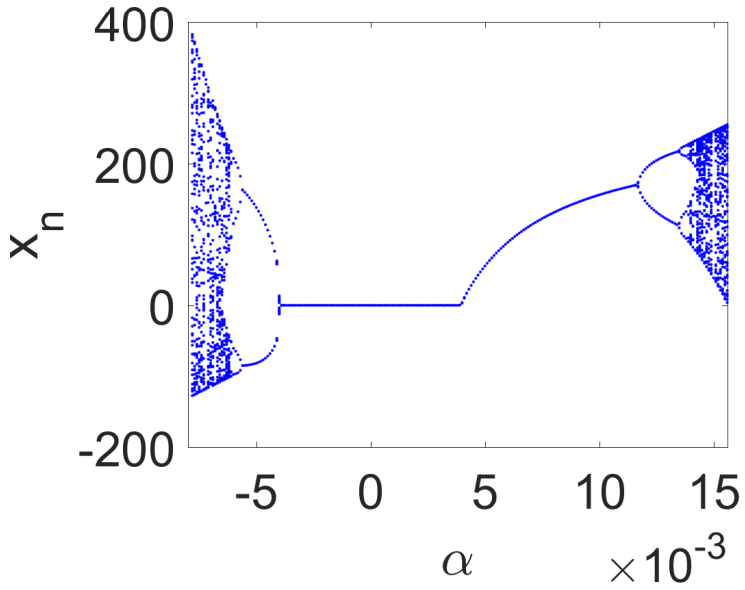


Figura 4.2: Diagrama de bifurcación del mapeo logístico extendido dado por la Ec. (4.1).

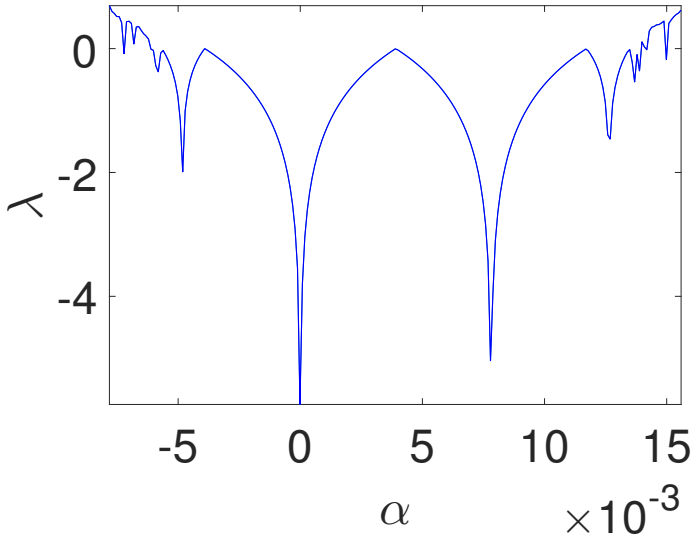


Figura 4.3: Exponente de Lyapunov en función del parámetro α.

por lo que se consideran $\alpha = \frac{-2}{2^8}$ y $\frac{4}{2^8}$.

En la Figura 4.4 se presenta la forma del mapeo logístico para ambos valores de parámetros $\alpha = \frac{-2}{2^8}$ y $\alpha = \frac{4}{2^8}$ en triángulos azules y en cruces negras, respectivamente. El mapeo logístico para estos valores de parámetros es invariante en diferentes intervalos

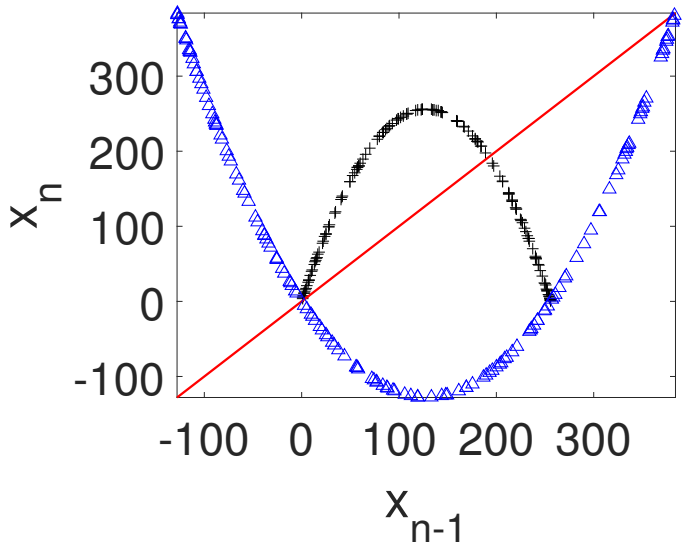


Figura 4.4: Mapeo logístico para $\alpha = \frac{-2}{2^8}$ en triángulos azules y para $\alpha = \frac{4}{2^8}$ en cruces negras.

de la siguiente manera

$$\begin{aligned} f_{\frac{-2}{2^8}} &: [-128, 384] \to [-128, 384]; \\ f_{\frac{4}{2^8}} &: [0, 255] \to [0, 255]. \end{aligned} \tag{4.2}$$

4.2. Generador de números pseudo-aleatorios con mapeo extendido

Siguiendo la metodología presentada en la sección 3.3, consideramos la órbita x_0, x_1, $x_2, \ldots$ del mapeo logístico extendido (4.1) dada la condición inicial $x_0 \in I$. El intervalo I está determinado por el parámetro $\alpha \in \{-\frac{-2}{2^8}, \frac{4}{2^8}\}$, $f_\alpha : I \to I$. Sean $M1$ y $M2$ dos series de tiempo generadas con el mapeo logístico extendido mediante las siguientes consideraciones: *i)* dadas dos condiciones iniciales arbitrarias x_{01}, x_{02}, de manera tal que, $x_{01} \neq x_{02}$; *ii)* dos valores del parámetro de bifurcación diferentes α_1 y α_2; y *iii)* l -unidades de memoria para cada serie de tiempo $x_{(i-k_{l-1})1}$, ..., $x_{(i-k_2)1}, x_{(i-k_1)1}, x_{i1}$ y $x_{(i-k_{l-1})1}$, ..., $x_{(i-k'_2)2}, x_{(i-k'_1)2}, x_{i2}$. Así las órbitas tienen una distribución uniforme independientemente de la distribución en forma de U del mapeo logístico.

Con el fin de ilustrar el algoritmo, se han elegido valores de parámetros de bifurcación como $\alpha_1 = \frac{-2}{2^8}$ y $\alpha_2 = \frac{4}{2^8}$ para las series de tiempo $M1$ y $M2$, respectivamente. Estos valores de parámetros aseguran que el sistema (4.1) tiene comportamiento caótico en ambos casos.

De la misma forma que en la sección 3.3, para garantizar que el generador presente buenas propiedades estadísticas y eliminar la forma del mapeo, se consideran series de tiempo con retardos diferentes, esto es, $k_2 = k_2' = 10$, $k_1 = 6$ y $k_1' = 5$ para ambas series de tiempo $M1$ y $M2$. Por lo tanto, estas series están formadas por la suma de dos estados de retardo $x_{(i-10)1}$ & $x_{(i-5)1}$ y el estado actual x_{i1} de la órbita $x_{01}, x_{11}, x_{21}, \ldots$, en $M1$. De la misma manera para $M2$, $x_{(i-10)2}$, $x_{(i-6)2}$ y x_{i2} de la órbita $x_{02}, x_{12}, x_{22}, \ldots$. Los valores de la serie de tiempo están limitados por la operación $mod\ 2^8$, lo que garantiza que $M1, M2 \in [0, 2^8) \subset \mathbb{R}$. Explícitamente $M1(x_{(i-10)1}, x_{(i-5)1}, x_{i1})$ y $M2(x_{(i-10)2}, x_{(i-6)2}, x_{i2})$ se expresan de la siguiente manera:

$$m_{i1} = M1(x_{(i-10)1}, x_{(i-5)1}, x_{i1}) = x_{(i-10)1} + x_{(i-5)1} + x_{i1},\ mod\ 2^8, \tag{4.3}$$

$$m_{i2} = M2(x_{(i-10)2}, x_{(i-6)2}, x_{i2}) = x_{(i-10)2} + x_{(i-6)2} + x_{i2},\ mod\ 2^8. \tag{4.4}$$

Finalmente, estas series de tiempo (4.3) y (4.4), se suman y la operación $mod\ 2^8$ es aplicada, este proceso genera una nueva serie de tiempo Z_i dada por:

$$Z_i = m_{i1} + m_{i2},\ mod\ 2^8. \tag{4.5}$$

Con estas series de tiempo con retardo (4.5), nos permite cambiar la característica de la distribución de probabilidad en "forma de U"[68] con una distribución de probabilidad uniforme en las series de tiempo obtenidas x_n y z_n, como fue mostrado en la Sección 3.3. Ésta es una característica importante, en comparación con los esquemas basados en caos, ya que facilita la construcción de las cajas de sustitución porque todos los valores tienen la misma probabilidad de ocurrencia a diferencia con los esquemas basados en sistemas caóticos simples.

Para obtener una serie de tiempo de números enteros s se construye la dinámica simbólica de Z_i. Así los elementos de s son números enteros, esto es, $s_i(Z_i) \in \{0, 2^8\}$.

Un requisito necesario para la dinámica simbólica es obtener cada uno de los enteros con la misma probabilidad, por lo tanto, el proceso para obtener las series es el siguiente:

$$s_i = \lfloor Z_i \rfloor. \tag{4.6}$$

Este proceso puede ser visto bajo el esquema mostrado en la Figura 4.5.

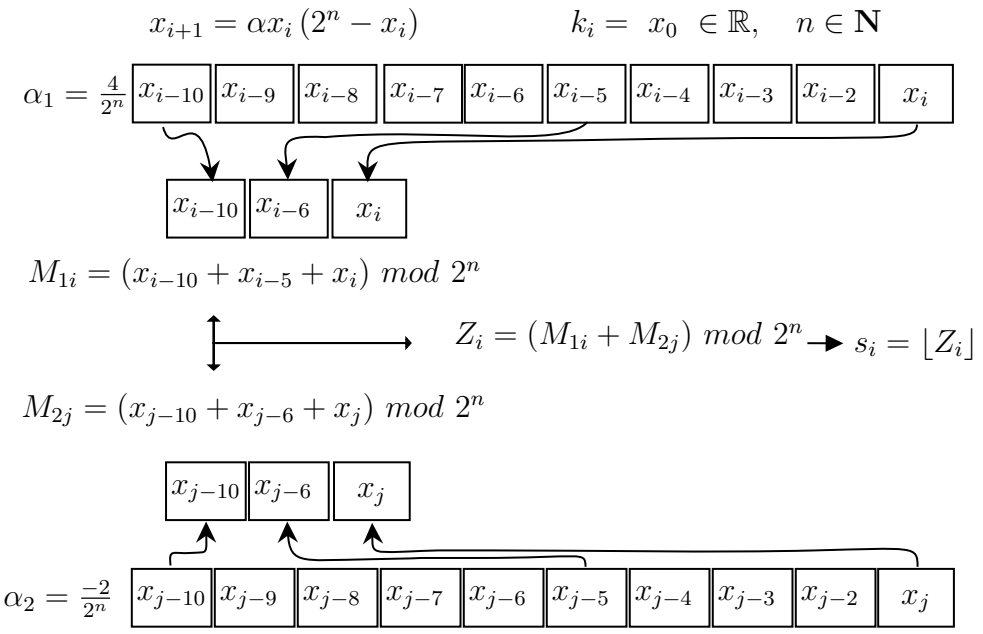


Figura 4.5: Esquema generador de números eneros pseudoaleatorios.

4.3. Algoritmo para generar cajas de sustitución usando el mapeo logístico extendido

En esta sección se propone un algoritmo para generar cajas de sustitución el cual está basado en PRNG que utiliza el mapeo logístico extendido [69]. Los pasos para generar el algoritmo son de manera similar a la estructura anterior Sección 3.4, a continuación se muestran:

Paso 1 Seleccione las condiciones iniciales x_{01} y x_{02} para PRNG para generar un flujo de bytes $s_0, s_1, s_2, \ldots$

Paso 2 Deseche los elementos repetidos D's para seleccionar 2^n valores diferentes. La regla para descartar un elemento es la siguiente: si $s_i = s_j$ con $i < j$ entonces descartar s_j.

Paso 3 Crear la caja de sustitución con 2^n elementos diferentes de s's.

Cuando finaliza el algoritmo propuesto da como resultado una caja de sustitución de $n \times n$ con 2^n valores distintos.

A modo de **ejemplo**, para $n = 8$, $x_{01} = 191$, $x_{02} = 209$, $\alpha_1 = \frac{1}{64}$ y $\alpha_2 = \frac{-1}{158}$ se deriva la caja de sustitución de 8×8 mostrada en la siguiente Tabla 4.1. Las propiedades de confusión y difusión están implícitas en la caja de sustitución propuesta.

8	195	2	130	142	128	75	60	40	248	178	117	225	34	169	212
85	3	244	222	122	246	110	206	181	95	131	89	18	81	104	37
16	151	118	239	228	199	154	149	5	236	42	14	220	45	237	47
1	240	254	9	41	243	64	135	229	53	21	103	73	173	6	214
78	97	98	31	230	59	231	241	189	120	235	234	87	226	249	217
51	137	233	204	207	105	24	213	114	48	183	187	17	201	132	245
106	170	172	140	58	148	200	57	164	202	92	30	180	113	68	152
36	156	134	29	232	141	115	82	205	223	193	102	251	174	46	76
192	32	123	136	147	33	247	70	49	129	72	211	88	7	255	126
168	66	138	83	71	61	112	171	127	167	165	23	12	186	26	56
108	175	65	133	198	27	54	111	124	4	84	39	13	96	44	52
143	162	216	93	91	100	190	194	15	210	43	69	107	203	50	221
161	185	79	209	11	94	101	0	22	159	224	166	182	63	35	238
150	67	252	25	10	90	208	77	121	176	116	119	163	144	177	99
86	38	191	158	62	139	74	218	157	160	197	80	19	55	125	28
179	184	153	188	215	145	155	20	196	109	219	146	242	250	253	227

Tabla 4.1: Caja de sustitución obtenida con el algoritmo propuesto.

4.4. Desempeño de las cajas de sustitución propuestas

En esta sección, presentamos el desempeño basados en los seis criterios criptográficos importantes de las cajas de sustitución.

4.4.1. Biyectividad

Una caja de sustitución cumple con el criterio de biyectividad si todos los posibles vectores de entrada se asignan a distintos vectores de salida, es decir, el valor ideal es 128 el cual es obtenido de la fórmula (2.11). El valor calculado de la caja de sustitución propuesta es $2^{n-1} = 128$, que es el valor deseado. Por esta razón el criterio de biyectividad se cumple, esto es, la caja de sustitución propuesta es uno a uno, sobreyectiva y balanceada.

4.4.2. No linealidad

Cada caja de sustitución contiene n funciones Booleanas, en este caso $n = 8$. Por consiguiente, ocho no linealidades 96, 104, 106, 102, 104, 102, 108 y 96 se obtienen de la caja de sustitución propuesta por medio la Ec. (2.13). Este resultado muestra que la caja de sustitución es altamente no lineal y difícil de abordar de forma lineal por un criptoanalista.

4.4.3. Estricto criterio de avalancha

El objetivo de este criterio es estimular la caja de sustitución en la entrada y observar el efecto de avalancha creado en los bits de salida. Idealmente, si un solo bit en la entrada se cambia, la mitad de los bits de salida deben cambiar y por lo tanto cambiar su estado. Siguiendo el procedimiento de la Ec. (2.16) se calcula una matriz de dependencia proporcionada en la Tabla 4.2, se obtiene un SAC máximo de 0.6094, el mínimo es 0.4219, y el valor promedio de 0.5059 que es cercano al valor deseado de 0.5. Basados en estos resultados, se puede concluir que la caja de sustitución generada por el método propuesto cumple con la propiedad de SAC.

4.4.4. Independencia de los bits de salida

El BIC de la caja de sustitución es calculado por el método detallado en la Subsección 2.2.4. Los resultados son mostrados en las Tablas 4.3, 4.4 y 4.5. El valor promedio de

0.4688	0.4688	0.4531	0.5313	0.4844	0.5156	0.4531	0.5469
0.4844	0.4688	0.5781	0.5938	0.5469	0.4688	0.5313	0.5469
0.5156	0.5156	0.5156	0.5000	0.5938	0.5469	0.5469	0.4688
0.4844	0.5625	0.5000	0.5000	0.4375	0.5000	0.5625	0.4688
0.5625	0.4844	0.5313	0.5313	0.4844	0.5156	0.4844	0.4375
0.5000	0.4844	0.5781	0.5313	0.4688	0.4688	0.5000	0.4531
0.5156	0.4688	0.5000	0.4531	0.5781	0.4688	0.5156	0.4219
0.4844	0.4531	0.6094	0.5625	0.5313	0.4531	0.4375	0.5469

Tabla 4.2: Criterio SAC de la caja de sustitución propuesta.

BIC- No linealidad es 103.50, el valor promedio de BIC-SAC es 0.5050 y el máximo valor de DD es 12, lo cual indica que la caja de sustitución satisface el BIC. El criterio BIC garantiza que no existe un patrón estadístico o dependencia entre los vectores de salida.

0	106	106	108	106	104	102	102
106	0	102	104	100	106	106	106
106	102	0	100	106	102	104	102
108	104	100	0	104	104	104	100
106	100	106	104	0	104	104	96
104	106	102	104	104	0	104	102
102	106	104	104	104	104	0	104
102	106	102	100	96	102	104	0

Tabla 4.3: Resultado del criterio de BIC-No linealidad de la caja de sustitución generada.

0	0.4785	0.5176	0.5098	0.5039	0.5195	0.5195	0.4707
0.4785	0	0.5000	0.4922	0.5195	0.5137	0.4941	0.5000
0.5176	0.5000	0	0.5469	0.5137	0.5020	0.4863	0.4980
0.5098	0.4922	0.5469	0	0.5098	0.5176	0.5137	0.4922
0.5039	0.5195	0.5137	0.5098	0	0.5176	0.4863	0.5156
0.5195	0.5137	0.5020	0.5176	0.5176	0	0.4785	0.4863
0.5195	0.4941	0.4863	0.5137	0.4863	0.4785	0	0.5371
0.4707	0.5000	0.4980	0.4922	0.5156	0.4863	0.5371	0

Tabla 4.4: Criterio de BIC-SAC de la caja de sustitución generada.

4.4.5. Distribución equiprobable XOR entrada/salida

El criterio de distribución equiprobable XOR entrada/salida de la caja de sustitución generada es calculado mediante la Ec. (2.18) y los resultados son mostrados en la Tabla 4.6. El valor máximo es 12, lo cual indica que la caja de sustitución satisface el criterio

0	0	4	4	8	12	2	2
0	0	4	4	4	2	6	2
4	4	0	6	2	4	2	0
4	4	6	0	4	10	4	8
8	4	2	4	0	4	4	12
12	2	4	10	4	0	8	6
2	6	2	4	4	8	0	2
2	2	0	8	12	6	2	0

Tabla 4.5: DD de la caja de sustitución generada.

y es resistente ante criptoanálisis diferencial. Esto significa que los bits de entrada y los bits de salida en una caja de sustitución tienen la misma distribución de probabilidad.

6	8	6	8	6	6	6	10	8	10	6	6	6	8	6	6
8	6	6	6	6	8	8	8	8	6	6	4	12	6	8	6
8	4	8	10	8	6	8	6	6	8	6	10	6	6	8	8
6	6	6	6	8	6	6	4	8	6	6	6	6	8	8	6
6	6	6	6	6	6	8	6	6	8	8	6	6	6	8	6
6	6	6	6	6	6	6	6	8	8	6	6	6	4	6	6
6	8	8	6	6	6	8	6	8	8	6	6	6	6	8	6
8	6	6	6	6	8	8	6	6	6	6	8	8	6	6	6
6	6	6	6	8	8	6	8	8	6	6	6	6	8	6	6
8	8	6	8	6	8	6	6	6	6	6	8	6	6	8	6
6	6	6	6	8	6	6	6	6	10	6	6	6	6	6	6
6	6	6	6	8	6	8	6	6	8	6	6	6	6	6	6
8	4	8	8	6	6	6	6	8	6	8	6	6	6	8	6
8	6	6	6	8	8	6	8	8	6	8	6	8	6	6	6
6	8	6	6	6	8	8	8	8	8	6	6	6	6	6	6
6	8	8	6	6	6	6	8	6	6	8	6	8	6	6	0

Tabla 4.6: Distribución equiprobable XOR entrada/salida de la caja de sustitución generada.

4.4.6. Probabilidad lineal máxima esperada

El valor del criterio de MELP es calculado a través de aproximaciones lineales de la función no lineal. El objetivo final es recuperar los bits de la llave. MELP estudia la correlación entre los bits de entrada y salida. Este criterio es calculado por la Ec. (2.18) y el valor promedio es 0.0156.

4.5. Resultados comparativos

En esta sección se hace una comparación del desempeño de la caja de sustitución creada con el algoritmo descrito en la Subsección 4.3. En la Tabla 4.7 se muestran los valores de los criterios para las cajas de sustitución propuestas en este trabajo (secciones 3.6 y 4.5) y un conjunto de cajas de sustitución ampliamente conocido (estándar y basadas en caos). En esta Tabla 4.7, se puede observar que la caja de sustitución generada con el mapeo logístico extendido cumple con los criterios [10, 14, 17, 30, 72–78]. Principalmente muestra mejor desempeño en las pruebas relacionadas con ataques (MELP y distribución equiprobable XOR entrada/salida). Además es una metodología con base en un sistema con operaciones simples que genera secuencias con comportamiento complejo.

	Biyectividad	**No linealidad**			**SAC**			**BIC**			**I/O XOR**	**MELP**
		mín.	**máx.**	**prom.**	**mín.**	**máx.**	**prom.**	**SAC**	**No linealidad**	**DD**		
Skipjack [79]	123	100	108	105.12	0.3906	0.5938	0.5027	0.5003	104.03	109	0.0469	0.0137
APA [72]	128	112	112	112	0.4375	0.5625	0.5007	0.4997	112	112	0.0156	0.0039
Gray [73]	128	112	112	112	0.4375	0.5625	0.4998	0.5026	112	112	0.0156	0.0039
AES [74]	128	112	112	112	0.4531	0.5625	0.5049	0.5046	112	112	0.0156	0.0039
Ref. [10]	128	98	107	103.25	0.3828	0.5938	0.5059	0.5033	104.21	108	0.0469	0.0166
Ref. [30]	129	103	109	104.87	0.3984	0.5703	0.4966	0.5044	102.96	109	0.0391	0.0176
Ref. [75]	128	96	106	103	0.3906	0.6250	0.5039	0.5010	100.35	106	0.5000	0.0220
Ref. [76]	128	112	112	112	0.4219	0.5469	0.5115	0.4982	108.71	112	0.0313	0.0120
Ref. [14]	128	102	108	105.25	0.4375	0.5781	0.5056	0.5019	103.78	108	0.0391	0.0244
Ref. [17]	128	104	110	106.25	0.4219	0.5938	0.5039	0.5059	103.35	108	0.0391	0.0198
Ref. [77]	128	102	108	106	0.4219	0.5938	0.5002	0.5016	104.42	108	0.0391	0.0220
Ref. [78]-1	128	106	108	106.75	0.3906	0.6094	0.4941	0.5013	104.28	108	0.0391	0.0156
Ref. [78]-2	128	106	108	106.75	0.4063	0.5938	0.4971	0.5008	102.92	106	0.0391	0.0198
Propuesta 1	128	96	104	101.75	0.3906	0.5781	0.5012	0.5066	103.42	108	0.0391	0.0176
Propuesta 2	128	96	108	102.25	0.4219	0.6094	0.5059	0.5050	103.50	108	0.0469	0.0156

Tabla 4.7: Comparación de cajas de sustitución basadas en caos y cajas de sustitución usadas en cifrados en bloque tradicionales.

Capítulo 5

Aplicación de las cajas de sustitución para el codificado de imágenes

El cifrado de Alberti [80] fue uno de los primeros cifrados polialfabéticos en los que su principio es la sustitución, esto empleando múltiples alfabetos de modo que la salida tenga una distribución uniforme. Hoy en día, el cifrado de Alberti está considerando como una codificación ya que solo sustituye un elemento por otro. Tomando esta idea de çifrados polialfabéticos", se presenta una aplicación de las cajas de sustitución en la codificación de imágenes, donde una intensidad particular de un píxel es sustituida por diferentes intensidades en la misma ronda. Para medir la efectividad de la función de codificado se realiza un análisis estadístico sobre la información codificada para determinar la aleatoriedad que produce la función.

5.1. Pruebas estadísticas para cifrados de imágenes.

En esta sección presentamos las pruebas de seguridad que se utilizan para evaluar las funciones de cifrado al utilizar imágenes. Se dice que cualquier criptosistema de imagen es bueno, si el algoritmo de cifrado oculta todos los atributos de una imagen

plana, y la imagen cifrada es totalmente aleatoria [81].

5.1.1. Correlación de pixeles

La correlación determina la relación entre dos variables. En otras palabras, la correlación es una medida que calcula el grado de similitud entre dos variables. El coeficiente de correlación es una medida útil para juzgar la calidad de cifrado de cualquier sistema criptográfico [82]. En general uno de los principales problemas al cifrar imágenes se debe a que los valores de los pixeles pueden estar altamente correlacionados en cualquier dirección (vertical, horizontal y diagonal). Sin embargo, la función de cifrado debe ser capaz de producir cambios en los valores de todos los pixeles y de esta forma obtener baja correlación en pixeles adyacentes. Para lograr este objetivo se debe obtener una distribución uniforme de pixeles en la imagen cifrada sin importar la distribución de la imagen en la entrada, esto es precisamente lo que se define como concepto de confusión. Para poder cuantificar y comparar la correlación de pixeles adyacentes en diferentes direcciones y en diferentes imágenes, se define el coeficiente de correlación r_{xy} el cual está dado por la covarianza de dos pixeles entre el producto de la desviación estándar de cada uno de los pixeles, como se muestra a continuación:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}}, \tag{5.1}$$

con

$$cov(x,y) = \frac{1}{\eta}\sum_{i=1}^{\eta}(x_i - E(x))(y_i - E(y)), \tag{5.2}$$

$$E(x) = \frac{1}{\eta}\sum_{i=1}^{\eta} x_i \quad , \quad E(y) = \frac{1}{\eta}\sum_{i=1}^{\eta} y_i, \tag{5.3}$$

$$D(x) = \frac{1}{\eta}\sum_{i=1}^{\eta}(x_i - E(x))^2 \quad , \quad D(y) = \frac{1}{\eta}\sum_{i=1}^{\eta}(y_i - E(y))^2, \tag{5.4}$$

donde x, y representan dos pixeles adyacentes, y η es el número total de pares (x, y) en este trabajo tomamos $\eta = 2000$. Se espera que la imagen original tenga un coeficiente de correlación cercano a 1, el cual indica que los pixeles adyacentes tienen valores

muy similares, por otro lado, se espera que la imagen cifrada tenga un coeficiente de correlación cercano a 0, indicando que los valores de pixeles adyacentes son diferentes.

5.1.2. Entropía

Previamente definida en la sección 2.1.3. Se calcula la entropía a la imagen original y a la imagen codificada en escala de grises, se espera que el valor de la entropía en la imagen codificada incremente con respecto a la entropía de la imagen original a un valor cercano a 8.

5.1.3. Calidad de cifrado.

Un problema importante en los algoritmos de cifrado es la evaluación de la calidad del cifrado. La desviación en los valores de píxeles entre la imagen original y la imagen cifrada es un buen parámetro para expresar la calidad del cifrado [82]. La aleatoriedad introducida en la imagen cifrada ayuda a ocultar las características de la imagen de texto plano. La calidad de cifrado es buena, si la desviación (cambios) de píxeles es máxima e irregular entre la imagen de texto plano y la imagen cifrada.

Un buen proceso de cifrado crea grandes cambios en el valor de los pixeles, donde estos nuevos valores de los pixeles deben de ser completamente diferentes a los de la imagen original. También se requiere que estos cambios sean irregulares así entre más cambios se encuentren en los valores de los pixeles más efectivo será el algoritmo de cifrado y por lo tanto tendrá una mejor calidad. La calidad de cifrado fue definida en [83] y se representa como el promedio de los cambios de cada nivel en la escala de grises:

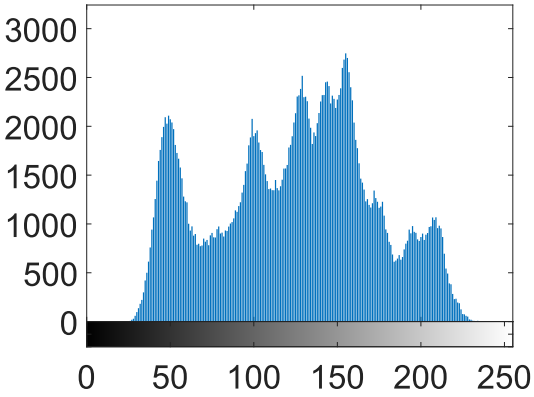
$$Q = \frac{\sum_{L=0}^{255} |H_L(C) - H_L(P)|}{256}, \tag{5.5}$$

donde L es el nivel de la escala de grises, $H_L(C)$ y $H_L(P)$ son el número de repeticiones de cada valor de gris en la imagen cifrada y en la imagen original respectivamente.

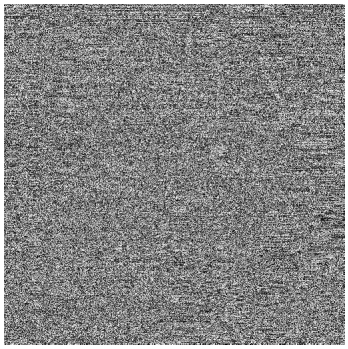
5.2. Función de codificado basada en las cajas de sustitución

Pensando en el principio del cifrado polialfabético, se generan cajas de sustitución con los algoritmos reportados en las secciones 3.4 y 4.3, para codificar una imagen en escala de grises bajo un esquema de cajas de sustitución. La propuesta utilizada, es mediante la aplicación de las cajas de sustitución a la imagen original, para ello se genera una caja de sustitución distinta por cada renglón de la imagen; es decir cada caja de sustitución modifica los pixeles de la imagen original por renglones obteniéndose la imagen codificada.

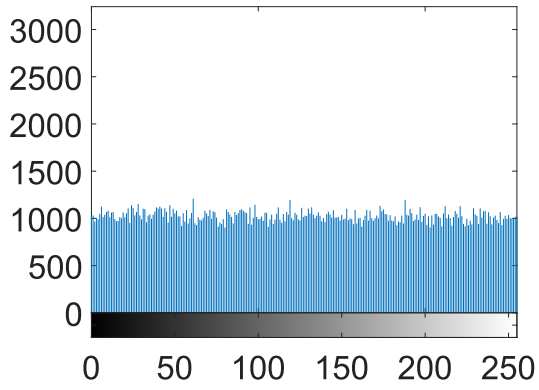
(a) Imagen original de Lenna (P).



(b) Distribución de pixeles de P.



(c) Imagen codificada de Lenna (C).



(d) Distribución de pixeles de C.

Figura 5.1: Codificado de imagen en escala de grises bajo un esquema de cajas de sustitución dinámicas.

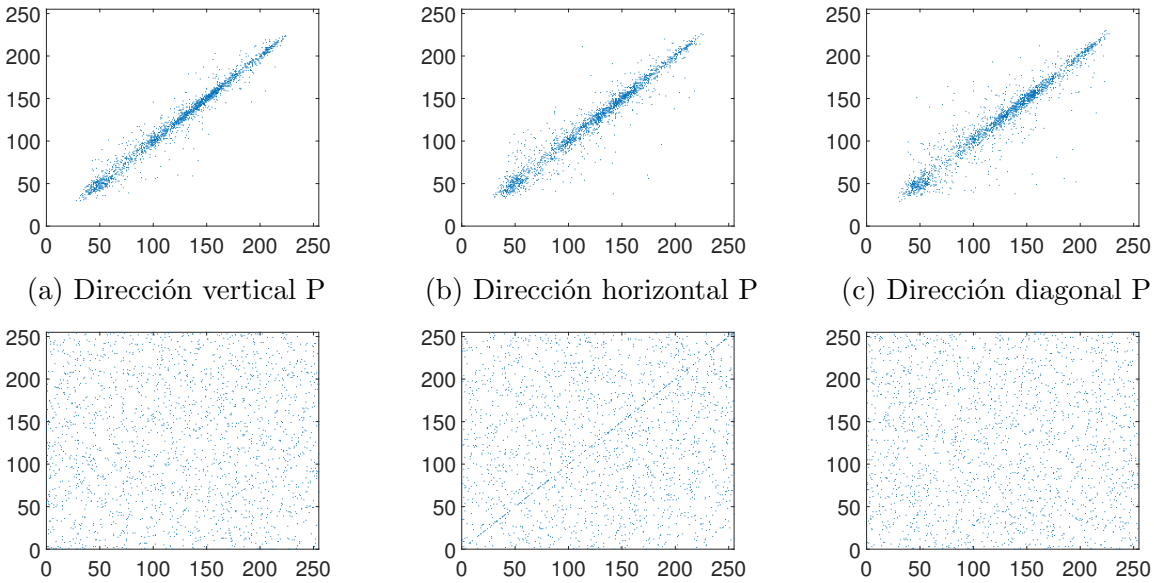
La Figura 5.1 muestra el codificado de una imagen en escala de grises bajo un es-

quema de cajas de sustitución donde 5.1a-5.1c representan la imagen de Lenna original y la imagen codificada junto con la distribución de pixeles en escala de grises de P y C dadas en 5.1b-5.1d. En la criptografía, siempre se desea una distribución uniforme en los elementos del mensaje cifrado, ya que esta propiedad se logró mediante la simple aplicación de distintas cajas de sustitución por renglón a la imagen original.

Dirección	Imagen original	Imagen codificada
Horizontal	0.9696	0.0665
Vertical	0.9862	0.0305
Diagonal	0.9570	0.0298

Tabla 5.1: Correlación de pixeles adyacentes en distintas direcciones.

En la tabla 5.1 se muestran los resultados obtenidos de la prueba de correlación de pixeles con valores cercanos a 1 para la imagen original y 0 para la imagen codificada.



(a) Dirección vertical P (b) Dirección horizontal P (c) Dirección diagonal P

(d) Dirección vertical C (e) Dirección horizontal C (f) Dirección diagonal C

Figura 5.2: Diagramas de dispersión de pixeles adyacentes de la imagen de Lenna (P) e imagen cifrada (C) en diferentes direcciones.

Adicionalmente en la Figura 5.2 se exhiben los diagramas de dispersión de pixeles adyacentes en diferentes direcciones, de 5.2a-5.2c la correlación para la imagen de Lenna original y de 5.2d-5.2f la correlación para la imagen de Lenna codificada. En estas se puede observar que la imagen original presenta una alta correlación mientras que en la

imagen codificada esta desaparece.

El cálculo de la entropía de la imagen de Lenna original y codificada se muestran en la tabla 5.2, en la cual se puede observar que para la imagen codificada el valor de la entropía aumenta con respecto a la imagen original tomando un valor cercano a 8.

	Imagen original	Imagen codificada
Lenna	7.4452	7.9974

Tabla 5.2: Entropía de las imágenes.

Finalmente la tabla 5.3 presenta el valor de la calidad de cifrado para la imagen de Lenna bajo el esquema de cajas de sustitución en comparación con la aplicación de una caja de sustitución estática, de los resultados obtenidos se puede observar que la calidad del cifrado es mayor en el esquema de cajas de sustitución.

Lenna (caja por renglón)	Lenna (una sola caja)
662.0391	70.4219

Tabla 5.3: Calidad del cifrado para imágenes codificadas.

Capítulo 6

Conclusiones

En este trabajo de tesis se abordó el problema de diseño de algoritmos de cajas de sustitución útiles para criptografía basadas en mapeos caóticos. Con base en el análisis dinámico y de series de tiempo mediante dinámica simbólica, entropía de la información y el DFA del mapeo logístico, se determino que la mezcla de dos series de tiempo con retardo permite tener una distribución de probabilidad uniforme en los elementos de la serie, además de ocultar la traza del mapeo caótico utilizado. Utilizando estas series de tiempo del mapeo logístico, se desarrollaron dos algoritmos para la generación de cajas de sustitución útiles para criptografía. El primero, mediante series de tiempo con elementos reales en el intervalo $[0, 1]$ y el segundo, basado en un generador de números enteros en el intervalo $[0, 2^n]$.

Se evaluó el desempeño de las cajas generadas con los algoritmos propuestos llevando a cabo pruebas estadísticas determinadas (criterios de una buena caja de sustitución). Los resultados obtenidos muestran que todos los criterios se cumplen, los cuales fueron comparados con otras cajas de sustitución reportadas en la literatura, mostrando una alta inmunidad para resistir ataques de criptoanálisis diferencial y lineal.

Por último, se presentó una aplicación de los algoritmos de cajas de sustitución propuestos para el codificado de una imagen en blanco y negro basado en el principio del cifrado polialfabético, se evaluó la imagen codificada utilizando algunas pruebas de

cifrado obteniendo resultados satisfactorios.

Trabajo a futuro

Como trabajo a futuro se propone trabajar en la implementación de los algoritmos de generación de cajas de sustitución en algún algoritmo de cifrado por bloques ya existente. Adicionalmente se podrá trabajar en el diseño de un algoritmo de cifrado basado en sustituciones dinámicas.

Apéndice A

Productividad

Publicaciones en revistas

- B. B. Cassal-Quiroga, E. Campos-Cantón, **Generation of Dynamical S-Boxes for Block Ciphers via Extended Logistic Map**, *Mathematical Problems in Engineering*, Vol. 2020, (2020).

Congresos internacionales

- 2015 Workshop on chaotic and nonlinear dynamics in circuits and systems, **Póster of Dynamical S-Boxes for Block Ciphers via Extended Logistic Map**.

Congresos nacionales

- 1er encuentro para la divulgación e investigación en el estudio de sistemas complejos y sus aplicaciones, **Plática: Aplicación del mapeo logístico en la criptografía, 2019.**
- 2do Congreso Nacional de circuitos y sistemas, **Póster: Aplicación de mapeos caóticos a sistemas criptográficos, 2019.**

- 51 Congreso Nacional de la Sociedad Matemática Mexicana, **Plática: Sistemas criptográficos basados en sistemas caóticos, 2018.**
- Segundo Encuentro de Mujeres Matemáticas Mexicanas, **Póster: Algoritmos criptográficos basados en caos, 2018.**

Bibliografía

[1] C. Paar, y J. Pelzl, *Understanding cryptography: a textbook for students and practitioners*. Springer Science & Business Media, 2009.

[2] D. Kahn, *The Codebreakers: The comprehensive history of secret communication from ancient times to the internet*. Simon and Schuster, 1996.

[3] S. Singh, *The code book: the science of secrecy from ancient Egypt to quantum cryptography*. Anchor, 2000.

[4] G. S. Vernam, "Cipher printing telegraph systems: For secret wire and radio telegraphic communications," *Journal of the AIEE*, vol. 45, no. 2, pp. 109–115, 1926.

[5] A. J. Menezes, P. C. Van Oorschot, y S. A. Vanstone, *Handbook of Applied Cryptography*. CRC Press, 1997.

[6] R. L. Rivest, A. Shamir, y L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[7] E. Biham, y A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[8] D. E. Standard, "Federal information processing standards publication 46," *National Bureau of Standards, US Department of Commerce*, vol. 4, 1977.

[9] J. Daemen, y V. Rijmen, *The design of Rijndael: AES-the advanced encryption standard*. Springer Science & Business Media, 2013.

[10] G. Jakimoski, y L. Kocarev, "Chaos and cryptography: block encryption ciphers based on chaotic maps," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 2, pp. 163–169, 2001.

[11] G. Chen, "A novel heuristic method for obtaining S-boxes," *Chaos, Solitons & Fractals*, vol. 36, no. 4, pp. 1028 – 1036, 2008.

[12] Y. Wang, K. W. Wong, X. Liao, y T. Xiang, "A block cipher with dynamic S-boxes based on tent map," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 7, pp. 3089 – 3099, 2009.

[13] D. Lambić, "A novel method of s-box design based on chaotic map and composition method," *Chaos, Solitons & Fractals*, vol. 58, pp. 16 – 21, 2014.

[14] A. Belazi, M. Khan, A. A. A. El-Latif, y S. Belghith, "Efficient cryptosystem approaches: S-boxes and permutation–substitution-based encryption," *Nonlinear Dynamics*, vol. 87, no. 1, pp. 337–361, 2017.

[15] F. Özkaynak, y A. B. Özer, "A method for designing strong S-boxes based on chaotic Lorenz system," *Physics Letters A*, vol. 374, no. 36, pp. 3733 – 3738, 2010.

[16] G. Liu, W. Yang, W. Liu, y Y. Dai, "Designing S-boxes based on 3-D four-wing autonomous chaotic system," *Nonlinear Dynamics*, vol. 82, no. 4, pp. 1867–1877, 2015.

[17] Ü. Çavuşoğlu, A. Zengin, I. Pehlivan, y S. Kaçar, "A novel approach for strong s-box generation algorithm design based on chaotic scaled Zhongtang system," *Nonlinear Dynamics*, vol. 87, no. 2, pp. 1081–1094, 2017.

[18] R. Guesmi, M. A. B. Farah, A. Kachouri, y M. Samet, "A novel design of chaos based S-boxes using genetic algorithm techniques," in *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, pp. 678–684, IEEE, 2014.

[19] Y. Tian, y Z. Lu, "S-box: Six-dimensional compound hyperchaotic map and artificial bee colony algorithm," *Journal of Systems Engineering and Electronics*, vol. 27, no. 1, pp. 232–241, 2016.

[20] G. Alvarez, y S. Li, "Some basic cryptographic requirements for chaos-based cryptosystems," *International Journal of Bifurcation and Chaos*, vol. 16, no. 08, pp. 2129–2151, 2006.

[21] L. Kocarev, "Chaos-based cryptography: a brief overview," *IEEE Circuits and Systems Magazine*, vol. 1, no. 3, pp. 6–21, 2001.

[22] C. I. Rîncu, y A. Serbanescu, "Chaos-based cryptography a possible solution for information security," *Series III: Mathematics, Informatics, Physics, Bulletin of the Transilvania University of Brasov*, vol. 2, no. 51, pp. 113–126, 2009.

[23] I. Mishkovski, L. Kocarev, y S. Lian, *Chaos-Based Public-Key Cryptography*, pp. 27–65. Springer Berlin Heidelberg, 2011.

[24] S. R. T. Addabbo, A. Fort, y V. Vignoli, *Chaos based generation of true random bits*, pp. 355–377. Springer Berlin Heidelberg, 2009.

[25] T. Addabbo, A. Fort, S. Rocchi, y V. Vignoli, *Digitized Chaos for Pseudo-random Number Generation in Cryptography*, pp. 67–97. Springer Berlin Heidelberg, 2011.

[26] H. Kwok, y W. K. Tang, "A fast image encryption system based on chaotic maps with finite precision representation," *Chaos, solitons & fractals*, vol. 32, no. 4, pp. 1518–1529, 2007.

[27] X. Wang, X. Wang, J. Zhao, y Z. Zhang, "Chaotic encryption algorithm based on alternant of stream cipher and block cipher," *Nonlinear Dynamics*, vol. 63, no. 4, pp. 587–597, 2011.

[28] S. Mazloom, y A. M. Eftekhari-Moghadam, "Color image encryption based on coupled nonlinear chaotic map," *Chaos, Solitons & Fractals*, vol. 42, no. 3, pp. 1745–1754, 2009.

[29] H. Liu, y X. Wang, "Color image encryption based on one-time keys and robust chaotic maps," *Computers & Mathematics with Applications*, vol. 59, no. 10, pp. 3320–3327, 2010.

[30] G. Tang, y X. Liao, "A method for designing dynamical S-boxes based on discretized chaotic map," *Chaos, Solitons & Fractals*, vol. 23, no. 5, pp. 1901–1909, 2005.

[31] I. Hussain, T. Shah, M. A. Gondal, y H. Mahmood, "An efficient approach for the construction of lft S-boxes using chaotic logistic map," *Nonlinear Dynamics*, vol. 71, no. 1-2, pp. 133–140, 2013.

[32] I. Hussain, T. Shah, H. Mahmood, y M. A. Gondal, "Construction of s8 liu j S-boxes and their applications," *Computers & Mathematics with Applications*, vol. 64, no. 8, pp. 2450–2458, 2012.

[33] A. P. Nguyen, y T. D. Nguyen, "Determining quality of S-boxes using pseudo random sequences generated from stream ciphers," in *International Conference on Algorithms and Architectures for Parallel Processing*, pp. 72–79, Springer Berlin Heidelberg, 2012.

[34] Y. Liu, S. Tian, W. Hu, y C. Xing, "Design and statistical analysis of a new chaotic block cipher for wireless sensor networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 8, pp. 3267–3278, 2012.

[35] L. Kocarev, y G. Jakimoski, "Logistic map as a block encryption algorithm," *Physics Letters A*, vol. 289, no. 4-5, pp. 199–206, 2001.

[36] D. Yang, X. Liao, Y. Wang, H. Yang, y P. Wei, "A novel chaotic block cryptosystem based on iterating map with output-feedback," *Chaos, Solitons & Fractals*, vol. 41, no. 1, pp. 505–510, 2009.

[37] X. Y. Wang, y X. M. Bao, "A novel block cryptosystem based on the coupled chaotic map lattice," *Nonlinear Dynamics*, vol. 72, no. 4, pp. 707–715, 2013.

[38] Y. Zhou, L. Bao, y C. P. Chen, "Image encryption using a new parametric switching chaotic system," *Signal Processing*, vol. 93, no. 11, pp. 3039 – 3052, 2013.

[39] Z. H. Guan, F. J. Huang, y W. J. Guan, "Chaos-based image encryption algorithm," *Physics Letters A*, vol. 346, no. 1-3, pp. 153–157, 2005.

[40] C. Fu, B. Lin, Y. Miao, X. Liu, y J. Chen, "A novel chaos-based bit-level permutation scheme for digital image encryption," *Optics communications*, vol. 284, no. 23, pp. 5415–5423, 2011.

[41] K. Wong, *Image Encryption Using Chaotic Maps*, pp. 333–354. Springer Berlin Heidelberg, 2009.

[42] S. J. Li, C. Q. Li, G. R. Chen, N. G. Bourbakis, y K. T. Lo, "A general quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks," *Signal Processing: Image Communication*, vol. 23, no. 3, pp. 212–223, 2008.

[43] L. Kocarev, y S. Lian, *Chaos-based cryptography: Theory, algorithms and applications*. Springer Science & Business Media, 2011.

[44] G. Alvarez, J. M. Amigó, D. Arroyo, y S. Li, *Lessons Learnt from the Cryptanalysis of Chaos-Based Ciphers*, pp. 257–295. Springer Berlin Heidelberg, 2011.

[45] J. M. Amigó, *Chaos-Based Cryptography*, pp. 291–313. Springer Berlin Heidelberg, 2009.

[46] F. Özkaynak, y S. Yavuz, "Designing chaotic S-boxes based on time-delay chaotic system," *Nonlinear Dynamics*, vol. 74, no. 3, pp. 551–557, 2013.

[47] Y. A. Kuznetsov, *Elements of applied bifurcation theory*. Springer Science & Business Media, 2013.

[48] S. Lynch, *Dynamical Systems with Applications using Maple*. Birkhäuser Boston, 2010.

[49] O. M. Sharkovsky, "Coexistence of the cycles of a continuous mapping of the line into itself," *Ukrainskij matematicheskij zhurnal*, vol. 16, no. 01, pp. 61–71, 1964.

[50] R. L. Devaney, *An introduction to chaotic dynamical systems*. Westview Press, 2003.

[51] M. Vellekoop, y R. Berglund, "On intervals, transitivity = chaos," *The American Mathematical Monthly*, vol. 101, no. 4, pp. 353–355, 1994.

[52] J. Banks, J. Brooks, G. Cairns, G. Davis, y P. Stacey, "On Devaney's definition of chaos," *The American mathematical monthly*, vol. 99, no. 4, pp. 332–334, 1992.

[53] G. Tancredi, A. Sánchez, y F. Roig, "A comparison between methods to compute Lyapunov exponents," *The Astronomical Journal*, vol. 121, no. 2, p. 1171, 2001.

[54] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.

[55] C. K. Peng, S. Havlin, H. E. Stanley, y A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 82–87, 1995.

[56] C. Adams, y S. Tavares, "Good S-boxes are easy to find," in *Advances in Cryptology—CRYPTO'89 Proceedings* (G. Brassard, ed.), pp. 612–615, Springer New York, 1990.

[57] C. Adams, y S. Tavares, "The structured design of cryptographically good S-boxes," *Journal of Cryptology*, vol. 3, no. 1, pp. 27–41, 1990.

[58] Y. Tian, y Z. Lu, "Chaotic S-box: Intertwining logistic map and bacterial foraging optimization," *Mathematical Problems in Engineering*, vol. 2017, pp. 1–11, 2017.

[59] W. Millan, "How to improve the nonlinearity of bijective S-boxes," in *Information Security and Privacy* (C. Boyd and E. Dawson, eds.), pp. 181–192, Springer Berlin Heidelberg, 1998.

[60] A. F. Webster, y S. E. Tavares, "On the design of S-boxes," in *Advances in Cryptology — CRYPTO '85 Proceedings* (H. C. Williams, ed.), pp. 523–534, Springer Berlin Heidelberg, 1986.

[61] R. A. Holmgren, *A first course in discrete dynamical systems*. Springer Science & Business Media, 1996.

[62] S. N. Elaydi, *Discrete Chaos*. Chapman & Hall/CRC, 2000.

[63] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, pp. 459–467, 1976.

[64] D. S. Dendrinos, y M. Sonis, "Socio-spatial stocks and antistocks; the logistic map in real space," *The Annals of Regional Science*, vol. 27, no. 4, pp. 297–313, 1993.

[65] M. García-Martínez, y E. Campos-Cantón, "Pseudo-random bit generator based on lag time series," *International Journal of Modern Physics C*, vol. 25, no. 04, p. 1350105, 2014.

[66] C. Li, y G. Chen, "Estimating the Lyapunov exponents of discrete systems," *Chaos*, vol. 14, no. 2, pp. 343–346, 2004.

[67] C. Yang, C. Q. Wu, y P. Zhang, "Estimation of Lyapunov exponents from a time series for n-dimensional state space using nonlinear mapping," *Nonlinear Dynamics*, vol. 69, no. 4, pp. 1493–1507, 2012.

[68] J. Urías, E. Campos, y N. F. Rulkov, *Random Finite Approximations of Chaotic Maps*, pp. 231–242. Springer New York, 2006.

[69] B. B. Cassal-Quiroga y E. Campos-Cantón, "Generation of dynamical S-boxes for block ciphers via extended logistic map," *Mathematical Problems in Engineering*, vol. 2020, 2020.

[70] C. Bandt, y B. Pompe, "Permutation entropy: A natural complexity measure for time series," *Phys. Rev. Lett.*, vol. 88, p. 174102, 2002.

[71] B. B. Cassal-Quiroga, y E. Campos-Cantón, "Generation of dynamical S-boxes via lag time chaotic series for cryptosystems," *arXiv preprint arXiv:1902.06412*, 2019.

[72] L. Cui, y Y. Cao, "A new s-box structure named affine-power-affine," *International Journal of Innovative Computing, Information and Control*, vol. 3, no. 3, pp. 751–759, 2007.

[73] M. T. Tran, D. K. Bui, y A. D. Duong, "Gray s-box for advanced encryption standard," in *2008 International Conference on Computational Intelligence and Security*, pp. 253–258, 2008.

[74] J. Daemen, y V. Rijmen, "AES proposal: Rijndael." Available: `http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf`, 1999.

[75] M. Khan, T. Shah, H. Mahmood, M. A. Gondal, y I. Hussain, "A novel technique for the construction of strong S-boxes based on chaotic Lorenz systems," *Nonlinear Dynamics*, vol. 70, no. 3, pp. 2303–2311, 2012.

[76] A. Belazi, A. A. A. El-Latif, A. V. Diaconu, R. Rhouma, y S. Belghith, "Chaos-based partial image encryption scheme based on linear fractional and lifting wavelet transforms," *Optics and Lasers in Engineering*, vol. 88, pp. 37–50, 2017.

[77] F. U. Islam, y G. Liu, "Designing s-box based on 4D-4 wing hyperchaotic system," *3D Research*, vol. 8, no. 1, pp. 1–9, 2017.

[78] F. Özkaynak, "Construction of robust substitution boxes based on chaotic systems," *Neural Computing and Applications*, vol. 31, no. 8, pp. 1–10, 2017.

[79] I. Hussain, T. Shah, M. A. Gondal, y Y. Wang, "Analyses of SKIPJACK s-box," *World Appl. Sci. J.*, vol. 13, no. 11, pp. 2385–2388, 2011.

[80] L. B. Alberti, y A. Zaccagnini, "A treatise on ciphers, trans," *A. Zaccagnini (Turin: Galimberti, 1997)*, 1997.

[81] M. García-Martínez, y E. Campos-Cantón, "Pseudo-random bit generator based on multi-modal maps," *Nonlinear Dynamics*, vol. 82, no. 4, pp. 2119–2131, 2015.

[82] J. Ahmad, y F. Ahmed, "Efficiency analysis and security evaluation of image encryption schemes," *computing*, vol. 23, p. 25, 2010.

[83] H. E. D. H. Ahmed, H. M. Kalash, y O. S. F. Allah, "Encryption quality analysis of the RC5 block cipher algorithm for digital images," *Optical Engineering*, vol. 45, no. 10, p. 107003, 2006.