UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA, ARQUITECTURA Y DISEÑO ENSENADA

SEGURIDAD DE DATOS PERSONALES EN E-SALUD BASADO EN CAOS

TESIS

que para cubrir los requisitos necesarios para obtener el grado de

INGENIERO EN ELECTRÓNICA

presenta:

OSCAR FRANCISCO ATONDO VALDEZ

Ensenada, Baja California, México. Abril de 2019.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA, ARQUITECTURA Y DISEÑO ENSENADA

SEGURIDAD DE DATOS PERSONALES EN E-SALUD BASADO EN CAOS

TESIS

Que para obtener el grado de Ingeniero en Electrónica presenta:

OSCAR FRANCISCO ATONDO VALDEZ

Aprobada por el siguiente comité:

Dr. Miguel Ángel Murillo Escobar
Director del comité

Dr. Fausto Abundiz Pérez
Miembro del comité

Dr. Adrián Arellano Delgado
Miembro del comité

Dra. Liliana Cardoza Avendaño
Miembro del comité

Dr. José Antonio Michel Macarty
Miembro del comité

Ensenada, Baja California, México. Abril de 2019

RESUMEN de la tesis de **Oscar Francisco Atondo Valdez**, presentada como requerimiento para obtener el grado de INGENIERO en ELECTRÓNICA, del programa de Licenciatura de la Universidad Autónoma de Baja California. Ensenada, Baja California, México. Abril de 2019.

SEGURIDAD DE DATOS PERSONALES EN E-SALUD BASADO EN CAOS

Resumen aprobado por:

Dr. Miguel Ángel Murillo Escobar
Director de tesis

En este trabajo de tesis, se diseña un algoritmo de encriptamiento basado en caos y se implementa a una interfaz gráfica para añadir seguridad al envío de información médica mediante correo electrónico.

Se analizan diferentes mapas caóticos con la finalidad de determinar cual de ellos cumple en mejor medida con las características necesarias para ser utilizado en un sistema de encriptamiento, obteniendo al Mapa Tent como el más apto para ser implementado con el algoritmo propuesto, dadas sus características de sensibilidad a las condiciones iniciales, no linealidad, secuencias obtenidas y velocidad de procesamiento.

Se le realizan distintos análisis estadísticos de seguridad para determinar la efectividad del algoritmo criptográfico propuesto, entre los cuales se encuentran la sensibilidad a la clave, sensibilidad a texto Claro, N-gramas y entropía de la información, así como el tiempo de encriptado que se demoraba el algoritmo en procesar los datos a encriptar. Adicionalmente, se desarrolla una interfaz gráfica en la que se simplifica la captura de datos de historiales clínicos de pacientes y con ayuda del algoritmo diseñado, se encripta la información capturada para su envío mediante correo electrónico de manera segura.

Se espera que el uso de este tipo de interfaces con sistemas de encriptado caótico mejore el proceso seguro de envío de información medica privada.

Palabras clave: telemedicina, e-Salud, caos, criptografía, análisis de seguridad, mapa Tent.

Abstract of the thesis presented by **Oscar Francisco Atondo Valdez**, as a requirement to obtain the ELECTRONICS ENGINEER degree, of the program of Barchelor's degree of the Autonomous University of Baja California. Ensenada, Baja California, Mexico. April 2019.

CHAOS BASED E-HEALTH PERSONAL DATA SECURITY

Abstract approved by:

Dr. Miguel Ángel Murillo Escobar
Thesis director

In this thesis work, an encryption algorithm based on chaos is designed and implemented in a graphical interface to add security to medical information sharing by e-mail.

Different chaotic maps are analyzed in order to determine which of them has the necessary characteristics to be used in an encryption system, obtaining the Tent Map as the most suitable to be implemented with the proposed algorithm, because of its characteristics of sensitivity to initial conditions, non-linearity, obtained sequences and processing speed.

Several security statistical analyzes are carried out to determine the effectiveness of the proposed cryptographic algorithm, among which are key sensitivity, clear text sensitivity, N-grams and entropy of the information, as well as the encryption time that the algorithm takes to process the data to be encrypted. Additionally, a graphical interface is developed in which the capture of data from clinical records of patients is simplified and, with the help of the designed algorithm, the captured information is encrypted for sending by e-mail in a secure way.

It is expected that the use of this type of interfaces with chaotic encryption systems will improve the secure process of sharing private medical information.

Keywords: telemediciene, e-Health, chaos, cryptography, security analysis, Tent map.

A mis padres y hermanas

Agradecimientos

A mis padres, Teresa y Francisco, por su amor incondicional, el esfuerzo que han realizado para brindarme un hogar en donde crecer, educación para poder superarme y la formación que me llevó a convertirme en la persona que soy ahora.

A mis hermanas, Nayeli y Karla, por darme un ejemplo a seguir para cumplir mis metas, su cariño y apoyo en todo momento.

Al Dr. Miguel Ángel Murillo Escobar, por toda su ayuda y consejos brindados. Gracias por el tiempo dedicado a apoyarme a desarrollar este trabajo de tesis.

A la Dra. Rosa Martha López Gutiérrez, por todo su atención y cariño durante el tiempo que estudié la carrera universitaria. Por su apoyo como maestra y coordinadora de la carrera de Ingeniería Electrónica.

Al Dr. César Cruz Hernández, por el conocimiento que me compartió como maestro, sus consejos y orientación fuera del horario de clases. Por inspirarme a realizar este trabajo de tesis y continuar con proyectos de investigación.

A los miembros de mi comité de tesis, Dra. Liliana Cardoza Avendaño, Dr. Fausto Abundiz Pérez, Dr. Adrián Arellano Delgado Gutiérrez y Dr. José Antonio Michel Macarty, por ser parte de mi equipo de formación profesional, por sus comentarios y tiempo dedicado para evaluar mi trabajo.

A mis amigos, Adilene, Estefanía, Alexis, Carlos, Daniel, David, Eduardo y José, por su ayuda en los momentos que lo necesité y compartir momentos memorables en mi vida.

A la Universidad Autónoma de Baja California (UABC), por brindarme un espacio integro donde realizarme profesionalmente y adquirir conocimiento invaluable para mi desarrollo personal. En especial a la Facultad de Ingeniería, Arquitectura y Diseño (FIAD), por convertirse en mi segunda casa durante 4 años.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo económico brindado a través del Proyecto de Grupos de Investigación en Ciencia Básica, Referencia 166654.

Ensenada, Baja California, México.
Abril de 2019

Oscar Francisco Atondo Valdez

Tabla de Contenido

Lista de Figuras

Lista de Tablas

Capítulo 1

Introducción

Es innegable la presencia que tienen las telecomunicaciones en cualquier aspecto de la sociedad actual, llegando a tal punto de no poder imaginar las relaciones humanas modernas sin este tipo de tecnologías. La apresurada evolución de las técnicas de telecomunicaciones que se ha presentado en los últimos años ha conseguido acortar distancias de una forma que parecía impensable hace algunas décadas, no solamente permitiendo mejorar la forma en la que los seres humanos se comunican en la Tierra, sino también optimizar el intercambio de información desde y hacia el espacio exterior. Estas grandes aportaciones han permitido que, gracias a que las comunicaciones a grandes distancias han sido simplificadas, otros campos de la ciencia y tecnología pudieran crecer exponencialmente.

Las telecomunicaciones son una forma de comunicación electrónica a distancia, que permite cumplir los requerimientos de enlace rápido que se necesitan en el mundo para la resolución de sus innumerables problemas y la entrega oportuna de la información, utilizando distintos canales de intercomunicación. Estos canales se logran a partir de una eficaz infraestructura como instalaciones, protocolos, servidores, telefonía convencional e IP, software, centros de carga, equipo de cómputo o redes sociales, que hacen posible vivir informado con las novedades tecnológicas.

Con el paso del tiempo se le han ido encontrando una gran cantidad de aplicaciones a los medios de telecomunicaciones, desde mensajería personal hasta el intercambio de información militar. Dentro de éstas aplicaciones también se encuentran las que tienen fines médicos, dando origen a la telemedicina. Esta, a su vez, forma parte del campo de la e-Salud, el cual es el término que se utiliza para referirse a las tecnologías de la información y comunicaciones que se enfocan en el sector sanitario [1].

La telemedicina permite lograr el cuidado de pacientes a distancia, como las aportaciones de Dickinson et al. en 2018, en donde se monitoreó de forma remota el ritmo cardíaco de personas con fallos en el corazón [2]; el almacenamiento y envío de información médica, como lo pueden ser historiales clínicos o resultados de laboratorio; y la medicina interactiva a distancia, para consultas médicas en línea [3, 4], por ejemplo. El amplio espectro de oportunidades que ofrece la telemedicina ha llamado la atención de

distintos centros de investigación en todo el mundo, con el fin optimizar sus procesos. Un tema de interés muy importante que necesita ser resuelto son los factores de inseguridad que pueden afectar al intercambio de información médica, para evitar que un intruso acceda a la información compartida.

Para resolver los problemas de seguridad que afectan a los sistemas de comunicaciones es habitual hacer uso de la criptografía, la cual, permite cambiar la estructura de un conjunto de datos para que solamente puedan ser descifrados por los usuarios deseados. Existen diversos tipos de sistemas criptográficos, la selección del más adecuado dependerá de: las características de la información que será compartida, del medio en el que se transmitirá, de los métodos empleados en el envío de la información y de los requerimientos específicos de los usuarios del sistema. La confiabilidad de un sistema criptográfico debe ser puesta a prueba por medio de distintos análisis matemáticos y estadísticos [5].

Dentro de la gran cantidad de sistemas criptográficos que se han elaborando a lo largo de los años se encuentran los que utilizan técnicas de criptografía no convencional, los cuales han ido ganando popularidad en los últimos años gracias a que dichas técnicas aportan características especiales a los sistemas, las cuales los hacen menos susceptibles a ataques de criptoanálisis, que es la ciencia que se encarga de estudiar los métodos para corromper un sistema criptográfico y descubrir el mensaje original a partir del mensaje cifrado.

Por otro lado, el estudio del caos también se ha ido expandiendo rápidamente en años recientes. Aunque su nombre sea popularmente relacionado con el desorden y la impredicibilidad, lo cierto es que la teoría del caos postula que más allá del desorden, hay sistemas cambiantes en el tiempo que a simple vista podrían parecer nada más que un conjunto de datos al azar, pero en realidad obedencen a reglas de comportamiento que pueden ser preestablecidas, es decir, elaborar conjuntos de datos pseudoaleatorios a partir de determinadas condiciones iniciales.

Con el paso del tiempo se han ido encontrando aplicaciones distintas para los fundamentos de la teoría del caos. Una de las primeras fue su uso en el modelado de sistemas meteorológicos [6], aunque también es muy común encontrarle usos en la medicina [7] y, como no podría ser de otra forma, en la criptografía no convencional.

1.1. Motivación

Como era de esperarse, el constante aumento del uso de las tecnologías de comunicación a distancia ha originado una cantidad enorme de datos que son transmitidos a cada segundo, aunque en un principio no se esperaba tal magnitud de información que circula actualmente por los medios de telecomunicaciones. Toda esta información viaja por distintos canales, los cuales no siempre son los más seguros, dejando expuestos los

datos que en gran parte de los casos son confidenciales.

En el año 2016, más de 16 millones de registros de pacientes fueron robados a organizaciones de atención médica en los Estados Unidos. Ese mismo año, el sector médico fue la quinta industria más afectada por ataques cibernéticos. Además, a principios de 2017 el Servicio Nacional de Salud británico fue paralizado por un ataque que bloqueó las computadoras que contenían muchos de sus registros y sistemas de control de citas médicas.

Los puntos vulnerables de los sistemas de información y comunicaciones con propósitos médicos no son debidos a que no se ha trabajo en el diseño de sistemas de seguridad para estos fines, pues existe una gran cantidad de trabajo realizado en este campo. El principal problema es que conforme se hacen avances en la creación de técnicas de información de datos, también se desarrollan nuevos procedimientos criptoanalíticos que logran vulnerabilizar los sistemas criptográficos de reciente creación. Por ello, es esencial que los métodos de encriptado nuevos cuenten con características que los hagan más resistentes a este tipo de ataques.

Debido a lo mencionado anteriormente, se demuestra la importancia del desarollo de técnicas de criptografía no convencional, pues gracias a las características que añaden a métodos de cifrado se crean sistemas más complejos y con un menor grado de vulnerabilidad.

Las crecientes investigaciones enfocadas en el caos originan cada vez más sistemas caóticos, favoreciendo los avances en el uso de este tipo de técnicas en la criptografía. No obstante, la valoración para determinar si un sistema caótico es capaz de ser aplicado en un algoritmo de cifrado debe ser realizada a través de una serie de pruebas estadísticas que ayuden a evidenciar su seguridad y uso eficiente de recursos.

1.2. Objetivos y alcances de la tesis

Debido al interés de proteger la información confidencial en telemedicina y el creciente uso de algoritmos criptográficos no convencionales, se desarrolló este trabajo de tesis planteando alcanzar el siguiente *objetivo general*:

Diseñar e implementar un algoritmo de cifrado caótico para el envío de historiales clínicos por correo electrónico.

Que para cumplir con el objetivo general, se plantea alcanzar los siguientes *objetivos particulares*:

1. Determinar el mapa caótico a utilizar en función de sus características de no linealidad, pseudoaleatoreidad, sensibilidad a condiciones iniciales y velocidad de procesamiento.

2. Diseñar un algoritmo criptográfico para cadenas de texto con base en caos digital.
3. Realizar pruebas de seguridad y eficiencia al algoritmo diseñado.
4. Diseñar una interfaz gráfica para simplificar la captura de datos médicos y emplear el algoritmo criptográfico diseñado.

1.3. Organización del manuscrito

Este trabajo de investigación esta compuesto por siete capítulos, los cuales son brevemente descritos a continuación:

- **Capítulo 1:** se presenta la introducción de este trabajo de tesis, la motivación y los objetivos a alcanzar.
- **Capítulo 2:** se describen algunos conceptos y momentos históricos de la telemedicina, así como las características de un sistema de telemedicina y su seguridad.
- **Capítulo 3:** se presenta una introducción al caos, se muestran las características de sistema caótico y algunas pruebas para elegir un mapa caótico para el algoritmo de cifrado caótico propuesto.
- **Capítulo 4:** se presentan las definiciones relacionadas a la criptografía, un marco histórico sobre su evolución a lo largo del tiempo, se detallan las características de los sistemas criptográficos y la importancia de la seguridad en ellos, además se mencionan algunos métodos de cifrado no convencional.
- **Capítulo 5:** se detalla el algoritmo de cifrado propuesto para información médica y las pruebas de seguridad que se le aplicaron.
- **Capítulo 6:** se muestra la interfaz gráfica creada para que los usuarios puedan utilizar de forma práctica el algoritmo de cifrado propuesto.
- **Capítulo 7:** se reportan las conclusiones de este trabajo de forma general y se da una perspectiva para trabajos futuros.

Capítulo 2

Telemedicina

En este capítulo se describen algunos conceptos y se presenta un marco histórico de la telemedicina. Se describen las características que debe tener un sistema de telemedicina y se menciona la importancia que tiene la seguridad informática en este tipo de sistemas.

2.1. Introducción

La definición más general de la palabra telemedicina es *medicina a distancia*, esto incluye la prestación de todo tipo de servicios sanitarios y el intercambio de información médica de forma remota [8]. En un sentido más amplio, la OMS define a la telemedicina como *el suministro de servicios de atención sanitaria, en los que la distancia constituye un factor crítico, utilizando las telecomunicaciones con objeto de intercambiar datos para diagnósticos, tratamientos, prevenir enfermedades, investigaciones y la formación de profesionales de la salud* [9].

El sector sanitario además de ser uno de los más activos en cuanto a la incorporación de nuevas tecnologías en el cuidado del paciente, existen otros condicionantes que han hecho aumentar el interés por la telemedicina [10], como son:

- Barreras de acceso entre la población y los servicios sanitarios.
- Necesidad creciente de manejo de la información por parte de los profesionales sanitarios.
- Una tendencia en aumento de la población a exigir una atención sanitaria de mayor calidad.
- Mayor disponibilidad de la infraestructura necesaria para desarrollar sistemas de telemedicina.

Relacionado al campo de la telemedicina, en el año 1999 surgió un nuevo término, e-Sauld. Este fue derivado directamente del campo del comercio electrónico y apareció como una expresión necesaria para describir el uso combinado de la comunicación

electrónica y las tecnologías de la información en el sector sanitario, tanto en aspectos relacionados con la gestión de los negocios en ese campo como para usos clínicos y educativos, tanto en el entorno local como a distancia [11]. Para muchos autores el término e-Salud esta enfocado al creciente uso de Internet en telemedicina, para prestar todo tipo de servicios entre proveedores (públicos y privados) y pacientes [12].

La Telemedicina aporta grandes comodidades y beneficios como la reducción de los tiempos de atención, diagnósticos y tratamientos oportunos, mejora en la calidad del servicio, reducción de costes de transporte, atención continua, tratamientos apropiados, disminución de riegos profesionales, posibilidad de interconsulta, mayor cobertura, campañas oportunas de prevención, entre otras.

La generalización en el uso de la telemedicina obliga a la población a plantearse si realmente ofrece respuestas que sean aceptables, tanto en calidad como en eficiencia, eficacia y efectividad, ofreciendo un margen de seguridad aceptable para sus usuarios [13].

2.2. Breve historia de la telemedicina

Es verdad que en la actualidad se han logrado avances espectaculares en el campo de telemedicina, pero llegar a este punto tomó muchos años de contribuciones. Si bien, el primer uso de se le dió a las tecnologías de telecomunicaciones para situaciones médicas es incierto, el concepto de medicina a distancia pudo haberse originado hace cientos de años. Un ejemplo de esto es el uso del heliógrafo (ver figura 2.1) en Europa en el siglo XIV, mediante el cual se transmitía información sobre la plaga de peste negra de aquella época [14].

Figura 2.1: Heliógrafo, instrumento para hacer señales telegráficas por medio de la reflexión de los rayos del Sol en un espejo movible.

Otros dispositivos que alentaron la evolución de la telemedicina fueron: El telégrafo, pues su uso en el ámbito militar facilitó los pedidos de suministros médicos; el teléfono en el inicio del siglo XX, el cual fue de gran utilidad para que la comunidad médica

pudiera mantenerse actualizada; y la radio, que partir de la Primera Guerra Mundial fue usada regularmente para informar el estado de salud de sus tropas y solicitar ayuda médica [15].

El primer caso documentado en el campo de la telemedicina que apareció en la literatura médica fue un proyecto iniciado en 1948, en el que se transmitían imágenes radiológicas entre dos puntos de Pensilvania, Estados Unidos, los cuales se encontraban a 38 kilómetros de distancia. Gracias a estas aportaciones un equipo de radiólogos en Canadá crearon un sistema de teleradiología en los años 50.

En 1964, se realizó el primer enlace de video interactivo entre el Instituto de Psiquiatría de Nebraska en Omaha y el hospital estatal Norfolk, los cuales quedaban a 180 kilómetros de distancia [16], pero sólo hasta 1967 se instaló el primer sistema completo de televisión interactiva entre paciente y médico en tiempo real, enlazando el aeropuerto de Boston con el hospital general de Massachusetts [17].

A principios de los años setenta se eligieron 26 lugares de Alaska para comprobar si las comunicaciones podrían mejorar la salud de los pueblos. Se utilizó el satélite I de la NASA que fue puesto en órbita en 1966, donde se realizó la transmisión de televisión a blanco y negro. Se determinó que el uso de vídeo a distancia aportaba beneficios en algunos casos que no eran de urgencias, debido a que los casos de urgencias no podían esperar a la agenda de consultas planificadas de acuerdo a la disponibilidad del satélite.

En 1986, la clínica Mayo instaló un sistema basado en satélite con la finalidad de unir las clínicas de Rochester, Jacksonville y Scottsdale. El sistema permitía una comunicación de vídeo con una tasa completa de imágenes (30fps), capaz de ser utilizado en varias ramas de la medicina.

Con el auge de Internet en la década de 1990, se inicia una fase de la telemedicina caracterizada por una disminución en los costos de producción de equipos electrónicos de telecomunicación, donde se destacan especialmente investigaciones financiadas por la Armada de los Estados Unidos para el monitoreo de sus ejércitos, telepresencia por cirugía robótica y nuevas tecnologías para análisis médicos complejos como los de daño cerebral.

En septiembre de 2009, el TATRC (Telemedicine and Advanced Technology Research Center) mostró sus innovadoras tecnologías en métodos de detección y tratamiento de lesión cerebral, asegurando que al facilitar la evaluación inicial de los pacientes con trauma craneoencefálico con el uso de la telemedicina se obtienen mejores desenlaces en términos de calidad de vida, mayor impacto en el pronóstico a corto y a largo plazo y reduciría el tiempo de reingreso de un soldado a la zona de combate [18].

2.3. Características de los sistemas de Telemedicina

Un sistema de telemedicina es un sistema complementario para las actividades médicas. Hoy en día se pueden encontrar sistemas capaces de transmitir audio, vídeo, imágenes y documentos por medio de diversos sistemas de telecomunicaciones, lo que fomenta el avance de la telemedicina. Para que un sistema de telemedicina pudiera funcionar correctamente deben de emplear como mínimo equipos capaces de comunicarse, medios de comunicación y estándares de interoperabilidad de información.

En distintos documentos se muestran las características de un sistema de telemedicina, a continuación se enumeran los más esenciales:

1. Separación geografía entre usuarios del sistema durante un encuentro clínico.
2. El empleo de las tecnologías de telecomunicaciones necesarias para la interacción.
3. Equipo de gestión del sistema.
4. Infraestructura organizacional.
5. Desarrollo de protocolos clínicos para orientación de los pacientes hacia diagnósticos y fuentes de tratamiento apropiados.
6. Creación de normas de comportamiento para el reemplazamiento de las normas del comportamiento cara-a-cara tradicionales.

Lógicamente, las soluciones de la telemedicina presentan elementos que dependerán de su entorno de aplicación médico, así como la cantidad de usuarios involucrados y la variedad de escenarios de uso. En la actualidad, los sistemas de telemedicina están evolucionando desde las técnicas clásicas de conexión punto a punto para aplicaciones dedicadas hacia sistemas interactivos de multimedia en red distribuido.

Se pueden distinguir dos modos de operación básicos en los sistemas de telemedicina [19], que son:

- **En tiempo real o síncrono,** el cual requiere la disponibilidad simultánea de los usuarios que estén involucrados en la sesión, como en una videoconferencia, por ejemplo.
- **En tiempo diferido o asíncrono,** en los que un diagnóstico o una consulta se puede realizar con dilación de minutos u horas. Utiliza principalmente el correo electrónico como medio de transferencia de información y constituye el mayor volumen de la actividad de telemedicina.

2.4. Seguridad en Telemedicina

Como ya semencionó en los párrafos previos, la telemedicina ofrece una gran cantidad de beneficios, pero también tiene sus desventajas [20] (ver tabla 2.1).

Ventajas	Desventajas
Optimización de recursos asistenciales	Intercambio de una información sensible
Mejora en la gestión de la demanda	Gran volumen de información almacenada
Reducción de las estancias hospitalarias	Compromiso de la confidencialidad
Disminución de los desplazamientos	Compromiso de la seguridad
Mejor comunicación entre profesionales	Amenaza en la continuidad en la asistencia
Mejor accesibilidad de los pacientes	Poca equidad en el acceso a la tecnología

Tabla 2.1: Ventajas y desventajas de la telemedicina

Con lo que se ha mencionado anteriormente, se puede observar que el común denominador en los sistemas de telemedicina es el intercambio de la información. No se puede negar que este intercambio tiene grandes ventajas, pero es difícil establecer sus límites. De este punto es de donde surge la principal amenaza hacia la telemedicina, el hecho que las tecnologías permitan este intercambio de información a gran escala y con diferentes formatos, puede comprometer la seguridad de la información y su confidencialidad.

En la figura 2.2 se observa que los puntos vulnerables de un sistema de telemedicina son sus aplicaciones, servicios e infraestructura, mientras que los ataques y amenazas se dan principalmente en los equipos de red y diagnóstico médico, seguido del recurso humano y el almacenamiento y acceso de la información.

Figura 2.2: Puntos donde recaen las vulnerabilidades y ataques potenciales de un sistema de telemedicina

En los últimos años se han realizado grandes avances en el campo de la seguridad en telemedicina [21, 22], pero debido a que el uso de sistemas de telemedicina se expande a gran velocidad, es importante seguir haciendo aportaciones a este ámbito.

2.5. Conclusiones

La telemedicina representa la unión de las tecnologías de la información, telecomunicación y los servicios en salud. La evolución de cada una de ellas influencía directamente el desarrollo de más y mejores sistemas de telemedicina que puedan brindar un mejor servicio, especialmente en zonas desprotegidas, acercar a profesionales, colaborar en la educación continua y mejorar la atención de los pacientes sin tener que salir de sus hogares.

El punto débil más relevante de la telemedicina son los problemas de seguridad. Dado que la mayor cantidad de volumen de datos en telemedicina se da por correo electrónico, y siendo este un medio que no cuenta en todos los casos con las medidas de seguridad adecuadas, convierte a la gran cantidad de información médica transmitida cada minuto en un blanco fácil para los delincuentes cibernéticos.

Capítulo 3

Caos

En este capítulo, se presenta una introducción al *caos*, pasando por su significado y primeros aportes a la teoría del caos. Se describen las características que debe tener un sistema caótico y la relevancia que tiene el cálculo de la exponente de Lyapunov para determinar si un sistema es caótico o no lo es. Adicionalmente se presentan distintos mapas caóticos de los cuales se eligió uno para ser utilizado en el algoritmo de cifrado caótico propuesto en este trabajo de tesis.

3.1. Introducción

En matemáticas y física, entre otras ciencias, se utiliza el término *caos* para referirse a un comportamiento aparentemente impredecible de los sistemas dinámicos no lineales determinísticos, es decir, sistemas en los que cada estado futuro está determinado por el previo. En términos generales, el caos determinista da lugar a trayectorias asociadas a la evolución temporal de forma muy irregular y aparentemente aleatoria, sin embargo, son totalmente dependientes de sus valores anteriores.

Uno de los problemas no lineales que trajo de cabeza a los físicos y matemáticos desde el siglo XVII, dentro de la modelización del Sistema Solar, fue el *problema de los n cuerpos*, que puede enunciarse de manera muy sencilla: *dados n cuerpos de distintas masas bajo atracción gravitacional mutua, se trata de determinar el movimiento de cada uno de ellos en el espacio.* Aunque el problema tiene un enunciado aparentemente de gran simplicidad, su solución no es en absoluto fácil. Newton resolvió geométricamente el problema de los dos cuerpos para dos esferas moviéndose bajo atracción gravitacional mutua en los Principia y posteriormente Daniel Bernoulli lo resolvió analíticamente en una memoria premiada por la Academia Francesa [23].

Múltiples físicos y matemáticos dedicaron sus esfuerzos a dar una respuesta al problema de los tres cuerpos y a la cuestión de la estabilidad del Sistema Solar, llegando a contabilizarse más de 800 trabajos al respecto hasta el año 1900 [24]. De entre ellos, el trabajo Henri Poincaré (ver figura 3.1) fue clave para la configuración de la Teoría de

los Sistemas Dinámicos y del Caos. Poincaré mostró que las dinámicas muy complicadas que generaba este sistema de los tres cuerpos era posible, en el sentido de que una pequeña perturbación en el estado inicial de la posición del un cuerpo, podría llevar eventualmente a trayectorias radicalmente distintas.

Figura 3.1: Henri Poincaré (1854-1912).

En 1963, Edward Lorenz, interesado en el problema de la convección en la atmósfera terrestre, simplificó de forma drástica las ecuaciones de Navier-Stokes de la mecánica de fluidos, conocidas por su complejidad. Se dice que Lorenz decidió reducir la cantidad de números decimales para los cálculos en la computadora de 6 a 3, con la finalidad de ahorrar tiempo. Lorenz esperaba encontrar resultados similares con 3 decimales que con 6, sin embargo, eran totalmente diferentes, repitió el experimento con diferentes grados de precisión y siempre eran diferentes [25].

A este descubrimiento, Lorenz lo llamó *dependencia sensitiva de las condiciones iniciales* y con ello creó la base de una nueva ciencia: el Caos. Dependencia sensitiva de las condiciones iniciales significa que, partiendo de dos puntos del espacio de fases, las dos trayectorias correspondientes acaban por diverger, aun siendo exageradamente cercanos los puntos en cuestión. Estos dos puntos representan conjuntos de condiciones iniciales y las trayectorias obtenidas representan la distinta evolución del sistema según sea el punto de partida.

Lorenz descubrió que su sistema contenía una dinámica muy anormal, las soluciones oscilaban irregularmente sin llegar a repetirse, pero en una región acotada del espacio de fases. Al ver el gráfico resultante, las trayectorias rondaban siempre alrededor de lo que ahora se define como atractor extraño (ver figura 3.2). Curiosamente, la forma del atractor se asemeja a la de las alas de una mariposa.

El sistema que desarrolló Lorenz está descrito por las siguientes ecuaciones diferen-

ciales no lineales:

$$\frac{dx}{dt} = \sigma(y - x), \tag{3.1a}$$

$$\frac{dy}{dt} = \rho x - y - xz, \tag{3.1b}$$

$$\frac{dz}{dt} = xy - \beta z \tag{3.1c}$$

donde x, y y z son los estados del sistema, x_0, y_0 y z_0 son las condiciones iniciales, σ, ρ y β son los parámetros de control y t es el tiempo.

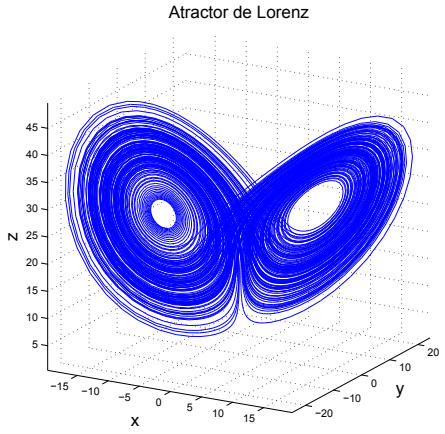

Figura 3.2: Atractor extraño del sistema de Lorenz en tres dimensiones.

3.2. Sistema caótico y sus propiedades

Un sistema dinámico es un sistema cuyo estado evoluciona con el tiempo. Durante muchos años, estos sistemas eran clasificados, dependiendo de su comportamiento, en estables e inestables, pero en la segunda mitad del siglo pasado se definió una nueva clasificación: el comportamiento caótico.

Un sistema estable genera fuerzas de atracción y lo mantiene confinado en un atractor de punto fijo u órbita periódica, mientras que en un sistema inestable se generan fuerzas de repulsión que expulsa a la trayectoria fuera del atractor. En este sentido, un sistema caótico presenta ambas características, es decir, existe un atractor que atrae la trayectoria del sistema, pero al mismo tiempo hay otras fuerzas que lo alejan de este y permanece confinado en una zona del atractor para trazar una trayectoria extraña [26].

Un sistema caótico se puede describir por un conjunto de ecuaciones diferenciales o en diferencias no lineales, que generan secuencias caóticas que son deterministas,

es decir, el valor futuro depende del valor actual. Los sistemas caóticos presentan las siguientes propiedades:

- *No linealidad.* No cumplen con el principio de superposición, es decir, un sistema complejo no puede ser descompuesto en un conjunto de sistemas simples.
- *Sensibilidad exponencial a condiciones iniciales y parámetros de control.* La dinámica o trayectoria del sistema caótico se verá altamente modificada si se varía ligeramente una condición inicial o parámetro de control.
- *Mezcla de datos.* Un pequeño rango de condiciones iniciales cubre la mayor parte del espectro caótico.
- *Ergodicidad.* Para cualquier condición inicial o parámetro de control, la trayectoria caótica se mantiene confinada en un espacio conocido como atractor extraño.
- *Exponente de Lyapunov positivo.* Un sistema de dimensión N posee N exponentes de Lyapunov; si uno de ellos es positivo, el sistema es caótico; si dos o más son positivos, el sistema es hipercaótico.
- *Atractor extraño con dimensión fractal.* La gráfica de fase del sistema genera lo que se conoce como atractor, que puede ser punto fijo (sistema estable), ciclo límite (sistema periódico) o atractor extraño (sistema caótico).

3.3. Exponente de Lyapunov

El exponente de Lyapunov es un análisis estadístico sobre la manera en que dos secuencias con condiciones iniciales extremadamente similares se comportan de forma iterativa, es una forma de poder cuantificar cuanto divergen. De esta forma, se puede utilizar el exponente de Lyapunov como un indicador del grado de sensibilidad a las condiciones iniciales y predictividad de la secuencia caótica [27].

El exponente de Lyapunov se determina con la siguiente expresión:

$$\lambda = \frac{1}{T}\ ln\ |\frac{f^n(x_n - \delta_0) - f^n(x_n)}{\delta_0}| \tag{3.2}$$

donde λ es el exponente de Lyapunov, T es el número de iteraciones, x_n es una condición inicial y δ_0 es la pequeña diferencia que se le añadirá a la condición inicial.

Los exponentes de Lyapunov son bastante útiles para determinar que tan caótico es un sistema. Este análisis (entre otras pruebas) se aplicó a cada uno de los mapas caóticos que se probaron, con el fin de determinar cual de ellos resultaría mas conveniente utilizar para el algoritmo de encriptado caótico propuesto.

3.4. Mapas Caóticos

En matemáticas, un mapa caótico es un sistema determinista que presenta comportamiento caótico. Los mapas pueden ser tanto de tiempo continuo como discreto. Es común encontrar el uso mapas caóticos para modelar una gran variedad de sistemas dinámicos.

Para evitar el uso de métodos de aproximaciones numéricas, en este trabajo se presentan mapas caóticos de tiempo discreto. Además se optó por limitar las dimensiones de los mapas a evaluar a 1, pues esto representa una menor cantidad de recursos computacionales a utilizar.

A continuación se presentan los mapas caóticos que se analizaron con ayuda del software *MATLAB* para determinar cual sería utilizado en el algoritmo propuesto.

3.4.1. Mapa Logístico

El mapa logístico de una dimensión es conocido como el sistema no lineal más sencillo que existe y que presenta claramente resultados caóticos. Está descrito por la siguiente ecuación [28]:

$$x_{n+1} = ax_n(x_n - 1) \tag{3.3}$$

donde $x_n \in (0, 1)$ es el estado del mapa discreto, x_0 es la condición inicial con valores entre $0 < x_0 < 1$ y a es el parámetro de control con $3.57 < a < 4$ para que el mapa genere secuencias caóticas.

En la figura 3.4, se observa la secuencia caótica generada por el mapa logístico, con $a = 3.999400522875507$, $x_0 = 0.500000000000000$ y $n = 250$.

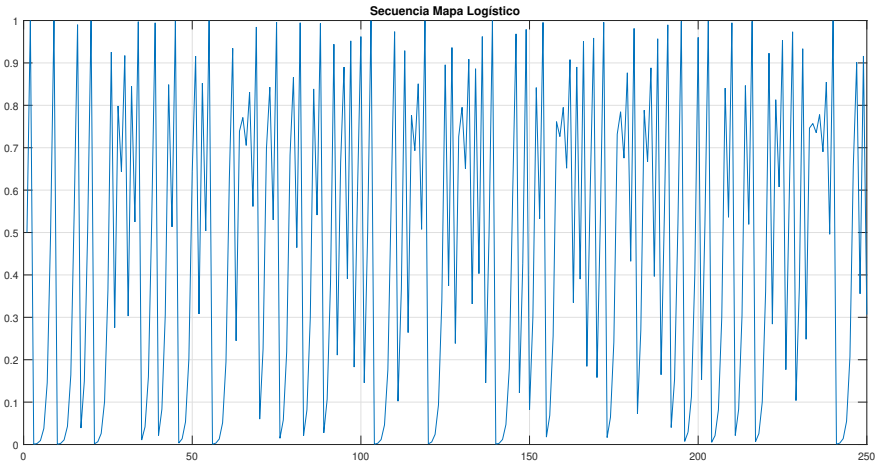

Figura 3.3: Estado x de la secuencia caótica del mapa logístico.

Para el cálculo del exponente de Lyapunov, se utiliza como condición inicial $x_0 = 0.500000000000000$, perturbación $\delta = 1 \times 10^{-6}$, número de iteraciones $T = 1000$ y parámetro de control $a = 3.999400522875507$. El valor obtenido fue $\lambda = 0.465728736335009$, y siendo este un valor positivo, se comprueba la presencia de caos en el mapa logístico.

3.4.2. Mapa Chebyshev

Los polinomios de Chebyshev son curvas de coseno con una perturbación en la escala horizontal, pero la vertical se mantiene constante. El mapa Chebyshev queda representado con la siguiente expresión:

$$x_{n+1} = cos(\alpha * cos^{-1}(x_n)) \tag{3.4}$$

donde $x_n \in (-1, 1)$ es el estado del mapa discreto, x_0 es la condición inicial, y α es el parámetro de control con $\alpha < 1$ para que el mapa genere secuencias caóticas.

En la figura 3.4 se observa la secuencia caótica generada por el mapa Chebyshev, con $\alpha = 3.999400522875507$, $x_0 = 0.000000000000000$ y $n = 250$.

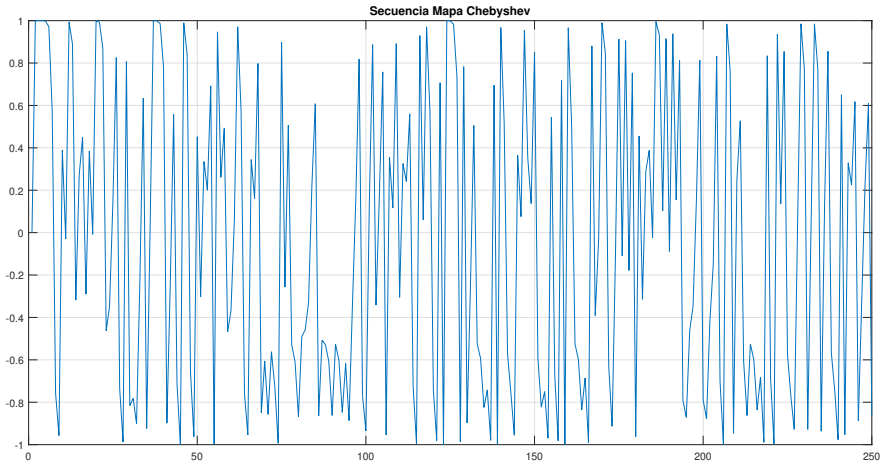

Figura 3.4: Estado x de la secuencia caótica del mapa Chebyshev.

Para el cálculo del exponente de Lyapunov, se utiliza como condición inicial $x_0 = 0.500000000000000$, perturbación $\delta = 1 \times 10^{-6}$, número de iteraciones $T = 1000$ y parámetro de control $\alpha = 3.999400522875507$. El valor obtenido fue $\lambda = 1.384187861062185$, comprobando el comportamiento caótico del mapa Chebyshev.

3.4.3. Mapa Tent

El *mapa Tent* esta compuesto por dos segmentos lineales. Este tipo de mapa tiene muchas variantes, pero para efectos de este trabajo se utiliza la función descrita por la siguiente expresión:

$$x_{n+1} = F(x_n, u) = \begin{cases} ux_n/2, & u < 0.5 \\ u(1 - x_n)/2, & u \geq 0.5 \end{cases} \tag{3.5}$$

donde $x_n \in (0, 1)$ es el estado del mapa discreto, x_0 es la condición inicial, y u es el parámetro de control con $u \in (2, 4]$ para que el mapa genere secuencias caóticas.

En la figura 3.5 se observa la secuencia caótica generada por el mapa Tent, con $u = 3.999400522875507$, $x_0 = 0.000000000000000$ y $n = 250$.

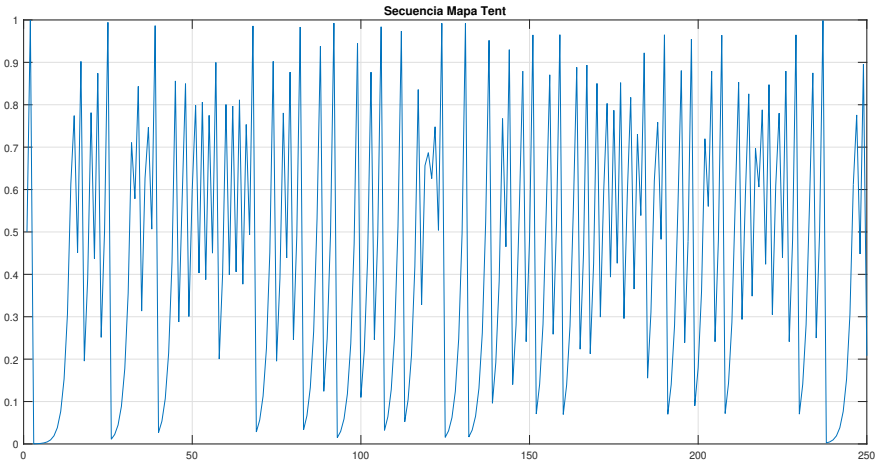

Figura 3.5: Estado x de la secuencia caótica del mapa Tent.

Para el cálculo del exponente de Lyapunov, se utiliza como condición inicial $x_0 = 0.500000000000000$, perturbación $\delta = 1 \times 10^{-6}$, número de iteraciones $T = 1000$ y parámetro de control $u = 3.999400522875507$. El valor obtenido fue $\lambda = 0.692997300054496$, confirmando que la secuencia obtenida con el mapa Tent presenta un comportamiento caótico.

3.4.4. Mapa caótico lineal por partes (PWLCM)

El *mapa caótico lineal por partes* (PWLCM por sus siglas en inglés) es uno de los mapas unidimensionales más famosos. Está compuesto por múltiples segmentos lineales y se define por la siguiente expresión:

$$x_{n+1} = F(x_n, p) = \begin{cases} x_n/p, & 0 < x_n < p \\ (x_n - p)/(0.5 - p), & p \leq x_n < 0.5 \\ F(1 - x_n, p), & 0.5 \leq x_n < 1 \end{cases} \tag{3.6}$$

En la figura 3.6 se presenta una secuencia caótica generada por el mapa PWLCM, con $p = 0.300000000000000$, $x_0 = 0.500000000000000$ y $n = 250$.

Para el cálculo del exponente de Lyapunov, se utiliza como condición inicial $x_0 = 0.500000000000000$, perturbación $\delta = 1 \times 10^{-6}$, número de iteraciones $T =$

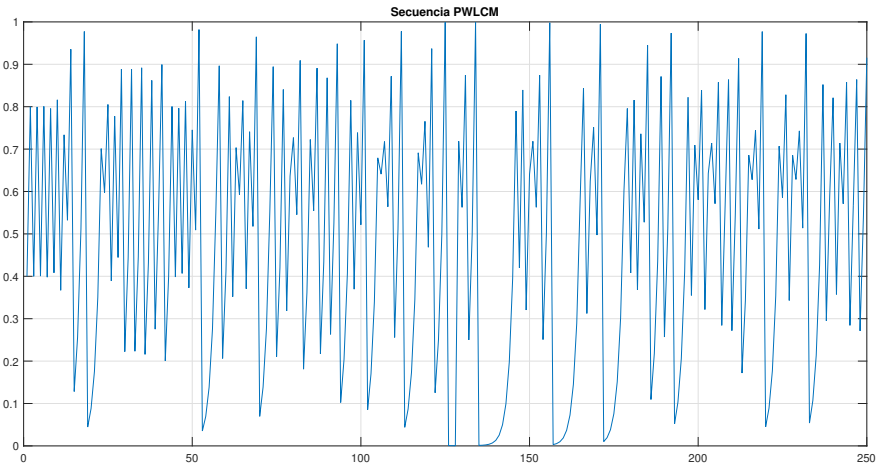

Figura 3.6: Estado x de la secuencia caótica del mapa PWLCM.

1000 y parámetro de control $p = 0.300000000000000$. El valor obtenido fue $\lambda = 1.256833150069977$, comprobando el comportamiento caótico del mapa PWLCM.

3.5. Comparación de mapas caóticos

Una vez confirmado el comportamiento caótico en los mapas presentados anteriormente, se compararon entre ellos para determinar cual se adecua mejor a las necesidades de un sistema de encriptado.

Además de generar secuencias pseudoaleatorias, para que un mapa caótico pueda ser utilizado en un algoritmo de cifrado, debe generar secuencias caóticas uniformes. Con secuencias de este tipo se reduce la probabilidad de que ciertos valores se repitan demasiado en la secuencia, a tal punto de producir un vulnerabilidad.

Para este trabajo de tesis se pretendía utilizar una mejora [29] a las secuencias caóticas obtenidas, la cual consiste en eliminar los primeros 3 dígitos después del punto de los valores caóticos generados por los mapas caóticos. Este proceso transforma todos aquellos valores muy cercanos a 0 (ejemplo 0.001321... a 0.321...) y muy cercanos a 1 (ejemplo 0.999231... a 0.231...), a valores totalmente diferentes. Esta mejora se aplicó a cada mapa caótico para comparar el efecto que tenía en cada uno de ellos.

La uniformidad de las secuencias es fácilmente observable haciendo uso de histogramas, estos ayudan a representar los datos obtenidos de forma que se pueda contabilizar el número de veces que se repite un valor en la secuencia obtenida.

Se obtuvieron los histogramas de cada mapa, utilizando los mismos parámetros y condiciones iniciales con las que se obtuvieron las exponentes de Lyapunov, pero en este caso se realizaron 1000 iteraciones en lugar de 250, para obtener más información.

En la figura 3.7, se observa que en la secuencia sin optimizar del mapa Logístico se tienen valores que se repiten hasta 24 veces, mientras que cuando la secuencia es optimizada este número baja hasta un máximo de 5 repeticiones el valor más frecuente.

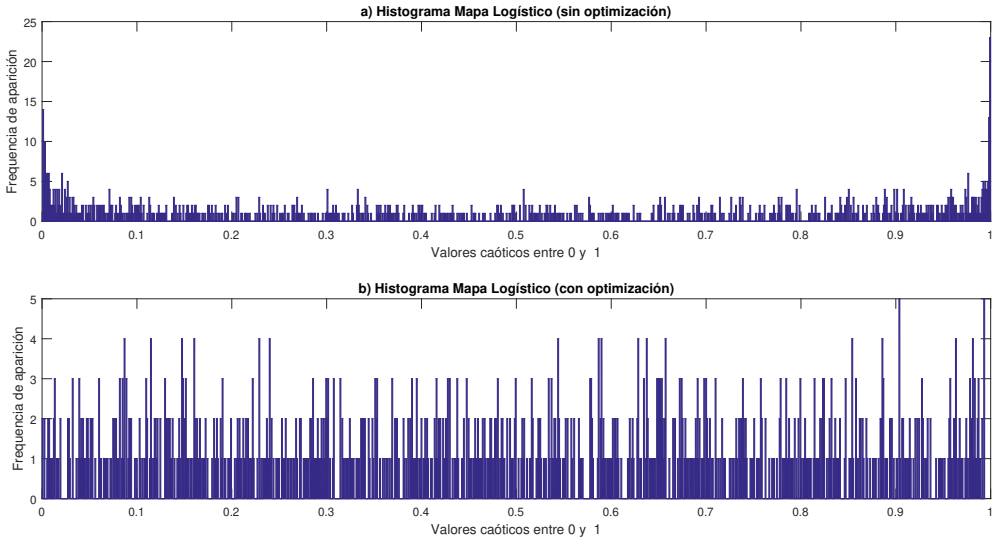

Figura 3.7: Histogramas de las secuencias caóticas obtenidas con el mapa Logístico: a) sin optimización, b) con optimización.

En la figura 3.8 se comparan los histogramas obtenidos con el mapa Chebyshev, en los cuales se tienen valores repetidos hasta 24 veces cuando no esta optimizado, y solamente 6 cuando si lo está.

En cuanto a los histogramas del mapa Tent que se observan en la figura 3.9, se obtuvo que con y sin optimización, el número de repeticiones de los valores mas frecuentes es de 5 para ambos casos. Esto quiere decir que, aún sin tener la optimización, el mapa Tent obtiene secuencias igual de uniformes que cuando si se encuentra optimizado.

Para los histogramas del mapa PWLCM de la figura 3.10, se obtuvieron resultados similares a los del mapa Tent, esto es, un máximo de 5 apariciones del valor más frecuente cuando está optimizado y cuando no lo está.

Además de los histogramas, se analizó el tiempo que le tomaba a al software MATLAB en obtener las secuencias de cada mapa, tanto sin optimización como con optimización. Los resultados de esta prueba se pueden observar en la tabla 3.1, en donde se obtuvo que el mapa que se procesa mas rápido es el logístico, seguido de los mapas Tent, Chebyshev y PWLCM, en ese orden.

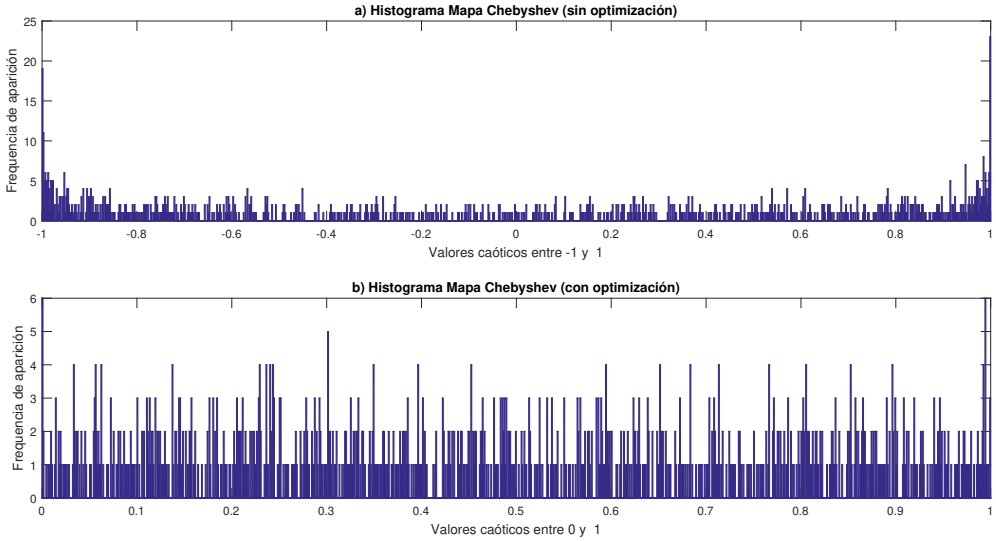

Figura 3.8: Histogramas de las secuencias caóticas obtenidas con el mapa Chebyshev: a) sin optimización, b) con optimización.

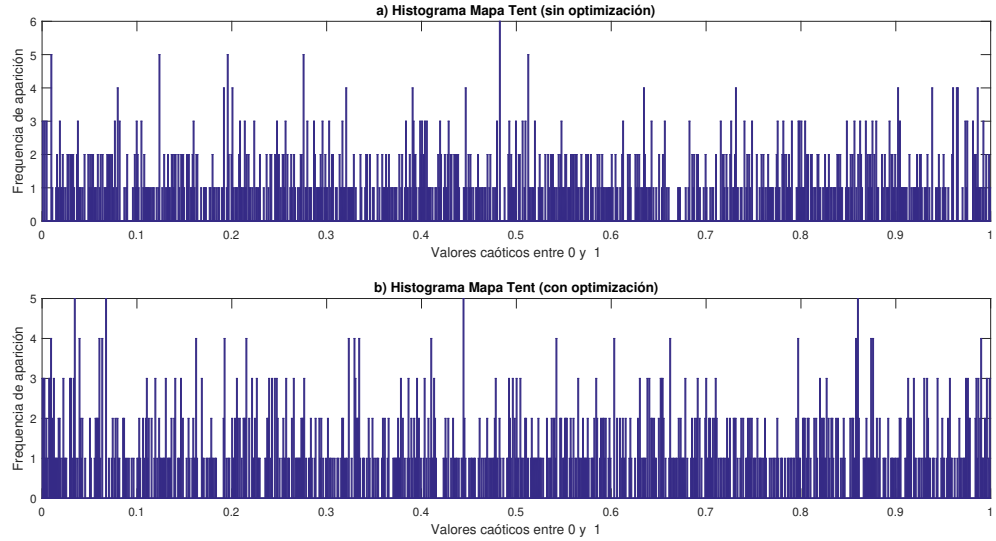

Figura 3.9: Histogramas de las secuencias caóticas obtenidas con el mapa Tent: a) sin optimización, b) con optimización.

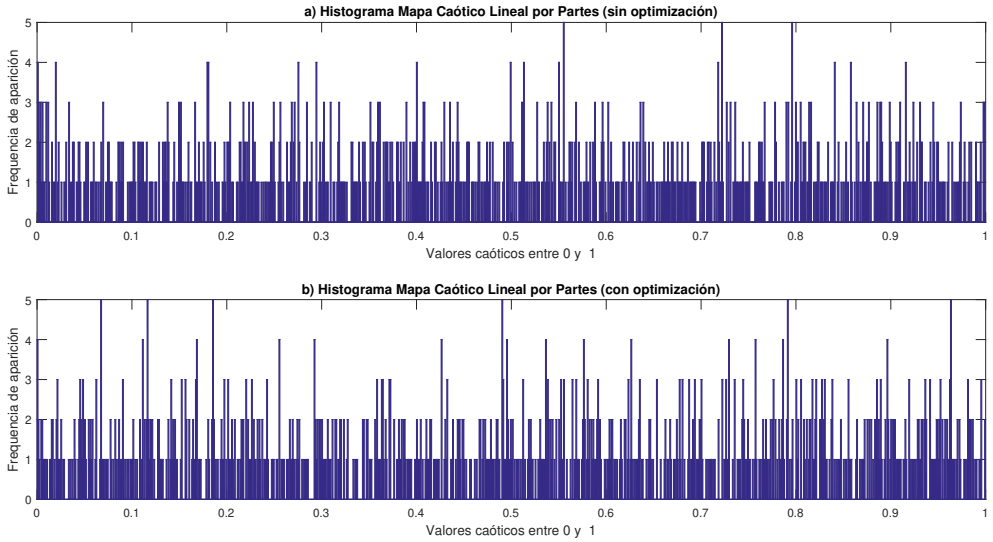

Figura 3.10: Histogramas de las secuencias caóticas obtenidas con el mapa PWLCM: a) sin optimización, b) con optimización.

Mapa	Sin Optimizar	Optimizado
Logístico	$0.323306mS$	$0.382486mS$
Chebyshev	$0.435351mS$	$0.500355mS$
Tent	$0.373177mS$	$0.438345mS$
PWLCM	$90.057476mS$	$90.154981mS$

Tabla 3.1: Tiempo de procesado de las secuencias de cada mapa caótico.

A partir de los datos obtenidos, tanto de los histogramas como de la prueba de tiempo, se obtuvo que el mapa logístico es el más rápido, sin embargo, requiere de la optimización para lograr secuencias uniformes. Por otro lado, los mapas Tent y PWLCM no necesitan ser optimizados para obtener uniformidad. La desventaja del mapa PWLCM es que tarda demasiado en ser procesado, en cambio el mapa Tent sin ser optimizado tarda menos en obtener secuencias con uniformidad similar a la de la secuencia del mapa logístico con optimización, pero en una menor cantidad de tiempo.

3.6. Conclusiones

Se presentaron las bases de la teoría del caos y se mencionaron las características de un sistema caótico. Se presentaron 4 mapas caóticos con el fin de seleccionar uno para ser implementado en el algoritmo propuesto en este trabajo de tesis.

Después de realizar las pruebas a cada mapa, se concluye que el mapa más adecuado a utilizar es el mapa Tent, pues con el se pueden obtener secuencias uniformes en una menor cantidad de tiempo, prescindiendo de la optimización para lograrlo.

Capítulo 4

Criptografía

En este capítulo se presentan las definiciones relacionadas a la criptografía, así como un marco histórico sobre la evolución y presencia de dicha ciencia a lo largo del tiempo. Se detallan las características de los sistemas criptográficos y los aspectos que se deben tomar en cuenta al evaluar la efectividad y eficiencia de los mismos. Así mismo, se mencionan algunos métodos de cifrado no convencional.

4.1. Introducción

La criptografía debe su nombre a los vocablos griegos kryptos (secreto) y graphein (escritura), dicho de otra forma, la criptografía es sinónimo de *escritura secreta*. En términos generales se podría decir que la criptografía es el estudio de la metodología y los principios de transformación de un mensaje legible a uno que no lo es, para posteriormente transformar el mensaje ilegible a su forma original [30].

De una forma más amplia, la criptografía es el estudio de las técnicas matemáticas relacionadas a los aspectos de seguridad de la información, como la confidencialidad, integridad de datos, autenticación de entidades, y autenticación del origen de la información. Esta ciencia hace uso extensivo de teoría de números, probabilidad, estadística, algebra, matemáticas discretas; apoyándose en la teoría de la información, complejidad computacional y teoría de cifrado. En particular, muchos algoritmos criptográficos modernos son diseñados y evaluados basándose en matemáticas discretas y combinatorias [31].

La contraparte de la criptografía es el criptoanálisis, pues este último se encarga de estudiar los sistemas criptográficos, pero no con el fin de mejorar su fiabilidad, sino con el de descubrir vulnerabilidades en ellos y romper su seguridad para tener acceso a la información secreta, en la mayoría de los casos utilizando los mismos fundamentos de la criptografía, tanto teóricos, como matemáticos y computacionales. Por otro lado, la criptología es la ciencia que engloba tanto a las técnicas de la criptografía como las del criptoanálisis [32, 33].

Aunque la criptografía parezca ser una rama de estudio compleja y con aplicaciones muy específicas, lo cierto es que es enormemente utilizada en asuntos que van desde los más básicos como la protección de mensajes de aplicaciones de chat, hasta cuestiones más delicadas como la seguridad de datos bancarios o incluso información militar. Con la gran abundancia de información relevante en el cambiante mundo digital actual, la batalla constante entre usuarios de dicha información y los intrusos que están al acecho de tales datos, ha provocado que la criptografía adquiera mayor importancia en los sistemas de comunicaciones y las técnicas criptográficas se encuentren en constante evolución [34, 35].

Visto lo anterior, el estudio de la criptografía no se limita solamente a un reducido grupo de personas trabajando en entornos secretos. La criptografía se ha convertido en una disciplina científica y cada vez más compañías cuentan con equipos de seguridad compuestos por personas que son capaces de diseñar, aplicar y evaluar algoritmos criptográficos. El conocimiento que es requerido para aplicar la criptografía en telecomunicaciones se ha vuelto más accesible con el paso del tiempo. Claramente esto ha provocado que el diseño de técnicas criptográficas para proveer seguridad a los sistemas de comunicaciones sea más sencillo.

Tradicionalmente el uso principal de la criptografía ha sido proveer confidencialidad a los sistemas de comunicaciones, encriptando los mensajes antes de ser enviados. Esto se extendió posteriormente a la provisión de autenticación, por ejemplo, utilizando un protocolo y algoritmo criptográfico para establecer si una entidad es realmente quién dice ser. Durante los últimos años se han mostrado nuevas aplicaciones para la criptografía. Algunos ejemplos son la protección de la integridad para datos confidenciales y el no rechazo de transacciones, que son importantes para los pagos electrónicos, pero también para otros fines. La mayor cantidad de aplicaciones descubiertas ha llevado a una creciente demanda de algoritmos criptográficos [36].

4.2. Historia de la Criptografía

Desde mucho antes de que aparecieran las primeras civilizaciones, los seres humanos han tenido la necesidad de comunicarse, es parte de su naturaleza y es un elemento clave para relacionarse los unos con los otros, pero, así como es necesario que sus mensajes sean conocidos, existen casos en donde un determinado mensaje es pensado para ser entendido solamente por un receptor específico y oculto para todos aquellos ajenos a este. Es de esperarse que, así como los sistemas de comunicación han cambiado con el paso de los años, las formas de mantener información confidencial hayan evolucionado también.

Aunque es difícil identificar el comienzo de los sistemas criptográficos, hay algunos rastros de evidencia que señalan el uso de la criptografía en los primeros sistemas de

escritura. Muchos modelos de *.esçritura secreta"* fueron inventados y reinventados durante miles de años, pero sin una mejora constante. Esta es una rama de estudio en la que la civilización china no dio grandes pasos, debido a la naturaleza especial de su idioma, que carece de un alfabeto de tamaño conveniente [37].

Con el paso de los años la criptografía ha ido desarrollándose conforme las necesidades humanas así lo han requerido. Los descubrimientos más antiguos de los que se ha tenido conocido sobre el uso de la criptografía se remontan hasta el antiguo Egipto, donde se han encontrado jeroglíficos no convencionales tallados en antiguos monumentos, cerca del año 1900 a.C., aunque, aun siendo mensajes que no eran entendidos por todos, no se han sido vistos como intentos serios de escritura secreta, sino más bien como métodos para generar misterio y curiosidad, sobre todo para el espectador más astuto [38].

El texto cifrado más antiguo que se ha descubierto pertenece a los escribas mesopotámicos, quienes a diferencia de los egipcios si tuvieron el objetivo fundamental de ocultar el significado de la escritura. La información cifrada está plasmada en tabletas de barro, las cuales datan del año 1500 a.C. aproximadamente, y en ellas se encuentra lo que parecen ser los detalles para la elaboración de un tipo barniz utilizado en la alfarería, lo cual debió tener un valor significativo en aquella época [39].

Aunque los métodos mencionados anteriormente fueron utilizados para transformar un mensaje en otro, el registro más antiguo de escritura secreta aplicado de una forma estrictamente enfocada a la transmisión de un mensaje con un nivel importante de confidencialidad fue en el 400 a.C., durante la guerra entre Atenas y Esparta, el ejército espartano empleaba la Scítala o Escítalo (ver figura 4.1), que puede ser considerado como el primer sistema de criptografía por transposición. El mensaje solo podía ser leído cuando éste, escrito en una tira de papel o cuero, se enrollaba sobre una vara del mismo largo y grosor, que poseía el destinatario deseado [40].

Figura 4.1: Escítalo, sistema criptográfico utilizado por el ejercito espartano.

Posteriormente, para el siglo I a.C., surgió el conocido *cifrado César*, llamado así dada la creencia de que fue utilizado por el emperador Julio César, aunque algunos historiadores indican que dicho emperador no lo utilizaba directamente, es un hecho que este método de cifrado fue utilizado en su misma época. El cifrado consiste en mover el carácter a representar cierto número de posiciones adelante (tres era la opción más

común) dentro del alfabeto a utilizar, el receptor del mensaje, solo debía sustituir cada letra del mensaje por su equivalente retrocediendo el lugar de las letras en el abecedario por el mismo número [41].

Durante la primera guerra mundial, el ministro de Asuntos Exteriores alemán, Arthur Zimmerman, envió un telegrama con texto cifrado al embajador de Alemania en Estados Unidos, con la intención de que lo reenviara al embajador alemán en México, donde proponía una Alianza entre Alemania y México, proponiendo a este último ayuda para que recuperara los territorios de Texas, Nuevo México y Arizona a cambio de servir como intermediario entre Japón y Alemania. El telegrama fue interceptado y descifrado por los servicios de inteligencia ingleses, que alertaron al ejército estadounidense, que se había mantenido al margen, quien se decidió a declararle la guerra al ejercito alemán, lo que contribuyó a dar fin al conflicto [42]. Esta fue la primera vez que el criptoanálisis había tenido una relevancia fundamental en un hecho histórico tan grande.

Para 1919 se registra la primera patente de una máquina criptográfica, la llamada máquina Enigma, creación del holandés Alexander Koch y el alemán Arthur Scherbius. Así la primera versión comercial fue puesta en venta en 1923 y se llamó Enigma-A. A esta primera versión le siguieron tres modelos Enigma B, C y D, siendo la última la más importante, ya que tuvo un gran éxito después de haber sido adquirida en 1926 por la marina alemana.

Figura 4.2: Máquina Enigma en el Museo Nacional de la Ciencia y la Tecnología Leonardo da Vinci, Milán.

En 1975 IBM presentó el sistema de cifrado de llave secreta DES (Data Encryption Standard). En 1977, el gobierno de Estados Unidos adoptó este método como estándar y también lo hicieron varios gobiernos del mundo. En 1981 DES fue estandarizado por la ANSI como ANSI X.3.92. y en 1998 fue descifrado en 56 horas por un ataque de fuerza bruta. Posteriormente, en el año de 1977, el MIT dio a conocer un poderoso algoritmo criptográfico llamado RSA que dada su robustez y efectividad es ampliamente utilizado

hoy en día [43].

4.3. Sistemas criptográficos y su clasificación

Un *sistema criptográfico* esta usualmente compuesto por un algoritmo conocido, el cual es dependiente de un parámetro llamado clave secreta para cifrar y descifrar información entre dos entidades (emisor y receptor).

Un sistema criptográfico puede definirse como una quintupla compuesta por los siguientes elementos [44]:

- m: representa el texto claro.
- c: representa el texto cifrado.
- K: representa la clave secreta.
- E: representa la función de cifrado.
- D: representa la función de descifrado.

En todo sistema criptográfico, se debe cumplir la siguiente condición

$$D_K(E_K(m)) = m \tag{4.1}$$

es decir, si un mensaje m es cifrado con una función E y una clave K y después se descifra con la misma clave K, se obtiene el mensaje original m.

En criptografía, los seis principios de Kerckhoffs relativos a las propiedades deseables de un sistema criptográfico son [45]:

1. Si el sistema no es teóricamente irrompible, al menos debe serlo en la práctica.
2. La efectividad del sistema no debe depender de que su diseño permanezca en secreto.
3. La clave debe ser fácilmente memorizable de manera que no haya que recurrir a notas escritas.
4. Los criptogramas deberán dar resultados alfanuméricos.
5. El sistema debe ser operable por una única persona.
6. El sistema debe ser fácil de utilizar.

Los sistemas criptográficos pueden ser clasificados en dos tipos, dependiendo del tipo de clave que utilicen [46]. Esta clasificación es la siguiente:

- *Simétricos o de clave privada:* son aquellos que utilizan una sola clave K tanto para cifrar como descifrar. La desventaja de estos sistemas de clave simétrica es que la transmisión de clave se debe gestionar de manera segura, esto quiere decir que el emisor y el receptor deben haber acordado una clave común por medio de un canal de comunicación seguro antes de poder intercambiar información confidencial por un canal de comunicación inseguro.

- *Asimétricos o de clave pública:* en este caso se utilizan dos claves, una pública K_p y otra privada K_q. La clave pública es utilizada para cifrar y la clave privada para descifrar. La ventaja es la seguridad en el manejo y tamaño de claves, ya que no es necesario que dos personas se pongan de acuerdo en que clave utilizar, lo único que tiene que hacer el emisor es conseguir la clave pública del receptor. Sin embargo, se requiere mayor tiempo de cálculo, mayor espacio de memoria y claves mayores de 1,024 bits para considerarse seguro, por lo que se traduce en mayor costo y tiempo de cálculo, comparándola con criptografía simétrica.

Los sistemas criptográficos también pueden ser clasificados dependiendo de la estructura del algoritmo de cifrado, esto es por:

- *Cifrado de flujo:* Cifran el mensaje bit a bit (o byte a byte) hasta terminarlo.

- *Cifrado de bloque:* Cifran el mensaje en bloques de k bits simultáneamente.

Dentro del cifrado simétrico existen dos conceptos relacionados con la teoría de la información y las comunicaciones seguras: la confusión y la difusión (el cifrado asimétrico se basa en operaciones matemáticas). Los algoritmos simétricos modernos como 3DES o AES implementan operaciones de confusión y difusión. A continuación se describen:

- *Confusión:* Consiste en permutar, es decir, **cambiar de posición**, cada elemento del texto claro (bit o bytes) de manera desordenada con respecto a la clave secreta y generar elementos cifrados.Este proceso tiene el objetivo de dificultar el descubrimiento de la clave con análisis estadístico.

- *Difusión:* Consiste en **cambiar el valor** a cada elemento del texto claro de manera desordenada con respecto a la clave secreta, para transformarlo en otro elemento del mismo alfabeto y generar elementos cifrados. Tiene la finalidad de ocultar cualquier relación estadística entre texto claro y texto cifrado.

Para los fines de este trabajo de tesis, se propone un algoritmo de cifrado simétrico, utilizando técnicas de difusión y confusión.

4.4. Criptografía no convencional

Las herramientas que se utiliza en la criptografía convencional para diseñar los algoritmos, son teoría de números, álgebra, curvas elípticas, entre otras, como los ya mencionados 3DES y AES. Por otro lado existe la criptografía no convencional en la que se utilizan herramientas matemáticas en estado de investigación como la criptografía cuántica [47, 48], criptografía con ADN [49, 50] y criptografía caótica [51, 52].

- **Criptografía cuántica**: Se basa en el principio de incertidumbre de Heisenberg, esto es, que al observar un sistema cuántico éste se perturba a si mismo, e impide que el observador conozca su estado exacto antes de la observación. Por tanto, si un sistema cuántico se utiliza para transferir información, alguien que quiera espiar la comunicación, o incluso el receptor previsto, podría verse impedido de obtener toda la información enviada por el emisor. La criptografía cuántica hace uso de dos canales de comunicación entre los dos participantes. Un canal cuántico, el cual, tiene una sola dirección y que generalmente consiste en una fibra óptica. El otro es un canal convencional, público y bidireccional. Para la transmisión de información, los datos binarios son codificados mediante fotones.

- **Criptografía ADN**: El ADN tiene propiedades que pueden ser útiles en un sistema de criptografía, tales como: capacidad de almacenar mucha información, paralelismo (fenómeno evolutivo, que produce un cambio equivalente en dos ramas de una agrupación contenida en un antepasado común) y poco consumo de potencia. En la actualidad operaciones de ADN basadas en suma, complemento, eliminación e inserción, son utilizadas para el cifrado de información.

- **Criptografía caótica**: Se basa en ecuaciones no lineales diferenciales o en diferencias, las cuales, generan secuencias desordenadas o caóticas, pero que son deterministas y que presentan sensibilidad a condiciones iniciales. No existe una fórmula simple que defina a un sistema caótico en cualquier punto dado, lo que se califica como un problema muy difícil de resolver, dotándolos de una ventaja para su aplicación en la criptografía, eliminando las desventajas fundamentales de la criptografía convencional. Las secuencias que produce un sistema caótico son utilizadas como referencia para transformar un texto claro a un texto cifrado mediante un algoritmo y una clave secreta que esta relacionada con las condiciones iniciales.

Las características de los sistemas caóticos los hacen aptos para ser utilizados en un sistema criptográfico, pues, además de añadir ventajas al sistema, su aplicación no es tan costosa como lo es en los sistemas de criptografía no convencional que utilizan tecnologías cuánticas o de ADN.

Un sistema de cifrado caótico digital de contener las siguientes propiedades [53]:

1. Se debe describir que sistema caótico utiliza el cifrado.

2. La degradación digital sebe ser evaluada, en caso de que se discretize un sistema continuo.

3. El sistema criptográfico debe ser fácil de implementar con base a costos aceptables y buena velocidad de cifrado.

4. La clave secreta debe ser claramente definida.

5. El espacio de claves debe ser especificada sólo para generar secuencias caóticas.

6. El efecto avalancha debe producirse para cualquier clave secreta: alta sensibilidad a la clave secreta.

7. Información parcial de la clave secreta, no debe revelar información parcial del texto claro, tampoco de la parte de la clave desconocida.

8. El proceso para generar secuencias caóticas a partir de la clave secreta debe estar claramente definido.

9. El cifrado debe tener alta sensibilidad al texto claro.

10. El cifrado debe generar un texto cifrado con distribución de probabilidad uniforme.

4.5. Seguridad de un sistema criptográfico

Los principios de Kerckhoffs indican que la seguridad de un sistema criptográfico debe recaer en la clave secreta y no sobre el algoritmo de cifrado. El algoritmo de cifrado se considera de dominio público. Un sistema criptográfico se considera vulnerado si un criptoanalista encuentra la forma de determinar la clave secreta y en consecuencia el texto claro.

Existen tres formas fundamentales de vulnerar un sistema criptográfico [53]:

1. **Ataques teóricos (lógicos):** Aplicar teoría de la información y criptoanálisis para quebrantar el algoritmo. La seguridad se evalúa mediante análisis basados en métodos matemáticos, donde se muestra que el mensaje cifrado y clave secreta no pueden ser revelados al implementar ataques conocidos.

2. **Ataques físicos:** Vulnerabilidades del sistema criptográfico, que se pueden aprovechar para quebrantar el algoritmo mediante ataques físicos, por ejemplo la información de tiempo, el consumo de energía, fugas electromagnéticas o incluso sonido, pueden proporcionar fuentes adicionales de información, que puede aprovecharse para romper el sistema criptográfico.

3. **Ataques humanos:** Persuadir o presionar a personas que poseen información privilegiada.

Por otra parte, el sistema criptográfico debe resistir los siguientes ataques criptoanalíticos (de tipo lógico), con base al principio de Kerckhoffs; es decir, se conoce todo sobre el sistema criptográfico, excepto la clave secreta:

1. *Ataques diferenciales.* Son ataques del tipo solo texto claro elegido y conocido, donde se debe mostrar alta sensibilidad del sistema criptográfico a la clave secreta y al texto claro, para que el sistema criptográfico pueda resistirlos.

2. *Ataques estadísticos.* Son ataques de histogramas y correlación, donde se debe mostrar mediante análisis de correlación e histogramas, la uniformidad del texto cifrado, para resistir estos ataques.

3. *Ataque exhaustivo.* Son ataques donde se tratan todas las posibles combinaciones de claves, por lo que, la clave debe contener más de 2^{100} opciones.

4.6. Conclusiones

La cripografía se encarga de salvaguardar la confidencialidad de la información que lo necesite. Ha formado parte de sucesos históricos relevantes y su evolución ha permitido conseguir sistemas criptográficos eficientes y seguros. Las técnicas de cifrado no convencionales han ganado popularidad años recientes, pero están aún en desarrollo. La teoría del caos puede ser aplicada en la criptografía para diseñar sistemas de cifrado no convencionales con altos niveles de seguridad.

Hay que tomar en cuenta las características que debe poseer un sistema de cifrado, sobre todo la importancia de la clave secreta. Las pruebas de seguridad en sistemas criptográficos son fundamentales para determinar sus fortalezas y/o vulnerabilidades.

Capítulo 5

Algoritmo de cifrado caótico propuesto

En este capítulo, se presentan los detalles del algoritmo de cifrado propuesto en este trabajo de tesis. Se propone utilizar una clave secreta simétrica de 128 bits representada por 32 caracteres hexadecimales para generar secuencias caóticas de dos mapas Tent y cifrar los datos de manera secuencial (flujo), basándose en una arquitectura de confusión y difusión. Finalmente, se presentan las pruebas de seguridad con las cuales se analizó el algoritmo de cifrado.

5.1. Introducción

Como se mencionó en la sección 2.2, el mayor volumen de la actividad de telemedicina proviene de sistemas de tiempo diferido, los cuales utilizan el correo electrónico como principal medio de transferencia de información. Uno de los principales elementos médicos enviados por correo electrónico es el historial clínico.

El *historial clínico* o *historia clínica* se puede definir como el *registro escrito de los datos sociales, preventivos y médicos de un paciente, obtenidos directa o indirectamente, y constantemente puestos al día*, es decir, se trata del documento que debe recoger la información del paciente en sus áreas: social (datos personales, datos laborales, datos familiares, datos educacionales, entre otros), preventiva (vacunaciones del paciente, por ejemplo) y asistencial (patologías atendidas en ocasiones anteriores y el seguimiento de las mismas), que permitirá darle seguimiento sanitario de cada individuo [54].

El registro del historial clínico construye un documento principal en un sistema de información sanitario, imprescindible en su vertiente asistencial, administrativa, y además constituye el registro completo de la atención prestada al paciente durante su enfermedad, de lo que se deriva su trascendencia como documento legal.

Un historial clínico electrónico puede ser fácilmente representado por una cadena

de texto. El cifrado de texto es el principal punto de enfoque de la criptografía. Esto representa el punto de partida para establecer los requisitos necesarios en el algoritmo de cifrado propuesto.

En este trabajo de tesis se propone un algoritmo para crifrado de historiales clínicos en formato de texto, compuestos de caracteres ASCII imprimibles (ver figura 5.1). Para la estructura de los historiales se tomó como referencia la representación de datos médicos en las fichas propuestas por Padney et al. en 2018 [55] (ver figura 5.2).

Caracteres imprimibles ASCII

Binario	Dec	Hex	Represen-tación	Binario	Dec	Hex	Represen-tación	Binario	Dec	Hex	Represen-tación
0010 0000	32	20	espacio ( )	0010 0001	33	21	!	0010 0010	34	22	"
0010 0011	35	23	#	0010 0100	36	24	$	0010 0101	37	25	%
0010 0110	38	26	&	0010 0111	39	27	'	0010 1000	40	28	(
0010 1001	41	29	)	0010 1010	42	2A	*	0010 1011	43	2B	+
0010 1100	44	2C	,	0010 1101	45	2D	-	0010 1110	46	2E	.
0010 1111	47	2F	/	0011 0000	48	30	0	0011 0001	49	31	1
0011 0010	50	32	2	0011 0011	51	33	3	0011 0100	52	34	4
0011 0101	53	35	5	0011 0110	54	36	6	0011 0111	55	37	7
0011 1000	56	38	8	0011 1001	57	39	9	0011 1010	58	3A	:
0011 1011	59	3B	;	0011 1100	60	3C	<	0011 1101	61	3D	=
0011 1110	62	3E	>	0011 1111	63	3F	?	0100 0000	64	40	@
0100 0001	65	41	A	0100 0010	66	42	B	0100 0011	67	43	C
0100 0100	68	44	D	0100 0101	69	45	E	0100 0110	70	46	F
0100 0111	71	47	G	0100 1000	72	48	H	0100 1001	73	49	I
0100 1010	74	4A	J	0100 1011	75	4B	K	0100 1100	76	4C	L
0100 1101	77	4D	M	0100 1110	78	4E	N	0100 1111	79	4F	O
0101 0000	80	50	P	0101 0001	81	51	Q	0101 0010	82	52	R
0101 0011	83	53	S	0101 0100	84	54	T	0101 0101	85	55	U
0101 0110	86	56	V	0101 0111	87	57	W	0101 1000	88	58	X
0101 1001	89	59	Y	0101 1010	90	5A	Z	0101 1011	91	5B	[
0101 1100	92	5C	\	0101 1101	93	5D	]	0101 1110	94	5E	^
0101 1111	95	5F	_	0110 0000	96	60	`	0110 0001	97	61	a
0110 0010	98	62	b	0110 0011	99	63	c	0110 0100	100	64	d

Figura 5.1: Caracteres ASCII imprimibles.

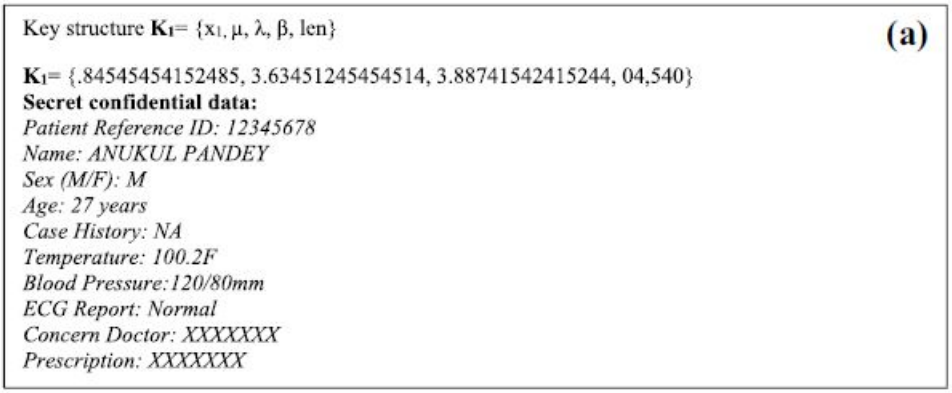

Figura 5.2: Ficha de historial clínico propuesta por Padney et al., 2018.

A partir de los datos de un historial clínico se puede generar una cadena de texto, permitiendo que la información fuera fácilmente manipulable por el algoritmo criptográfico presentado en esta tesis. El algoritmo fue pensado para utilizarse en conjunto de una interfaz gráfica que será presentada en el siguiente capítulo.

El algoritmo de cifrado propuesto en este trabajo, está basado en el trabajo de Murillo, 2015 [29]. Para efectos de este trabajo, se sustituyó el uso del mapa logístico por el mapa Tent para generar las secuencias caóticas. Como se vió en la sección 3.5, el mapa Tent genera secuencias caóticas más uniformes que el mapa logístico sin necesitar de algún tipo de optimización, permitiendo ahorrar tiempo en el cifrado.

El algoritmo presentado se basa en las siguientes características criptográficas [44]:

- *Cifrado simétrico*. Se utiliza la misma clave secreta para cifrar y descifrar.
- *Arquitectura de confusión y difusión*. Se emplean métodos para cambiar la posición y el valor a cada elemento claro en una sola operación.
- *Cifrado a flujo*. El algoritmo cifra cada elemento del texto claro, uno a la vez hasta terminarlo. En este caso, se utilizan elementos que son representados por 8 bits (1 byte).
- *Cifrado no convencional*. El algoritmo utiliza **secuencias caóticas del mapa Tent** (Sec. 3.4.3), que son determinadas por la clave secreta para generar secuencias pseudoaleatorias, permitiendo realizar el proceso de confusión y difusión.

El uso de mapas caóticos unidimensionales como el mapa Tent, tienen, habitualmente, ciertas desventajas si se utilizan para fines criptográficos. Algunas de las *desventajas* son [56]:

- Rangos caóticos discontinuos.
- Distribución de datos no uniforme.
- Espacio de claves pequeño.
- Periodicidad en rangos caóticos.

Sin embargo, los sistemas caóticos unidimensionales tienen *ventajas* como [57]:

- Estructura simple.
- Fácil de implementar en sistemas digitales.
- Poco consumo de memoria y recursos físicos.
- Generación de datos a alta velocidad.

Aún con las desventajas mencionadas, en la sección 3.5 se confirmó que en el caso de las secuencias caóticas generadas con el mapa Tent sí se presenta uniformidad, por lo que la desventaja de encontrar secuencias no uniformes que se da normalmente en los mapas caóticos unidimensionales no se presenta. Esto deja al mapa Tent con más ventajas que desventajas para su implementación en sistemas de cifrado.

En la figura 5.3 se muestra el **diagrama de bloques del proceso de cifrado**, el cual, procede de la siguiente manera: primero, se itera el mapa Tent 2 con base en la clave secreta, después se calcula el valor de Z que tiene relación con el texto claro y secuencias caóticas del mapa Tent 2, se continua con la iteración del mapa Tent 1 con base en la clave secreta y el valor de Z para realizar los procesos de confusión y difusión sobre el texto claro, y finalmente se agrega el valor de Z al criptograma para que el usuario autorizado pueda descifrar correctamente la información.

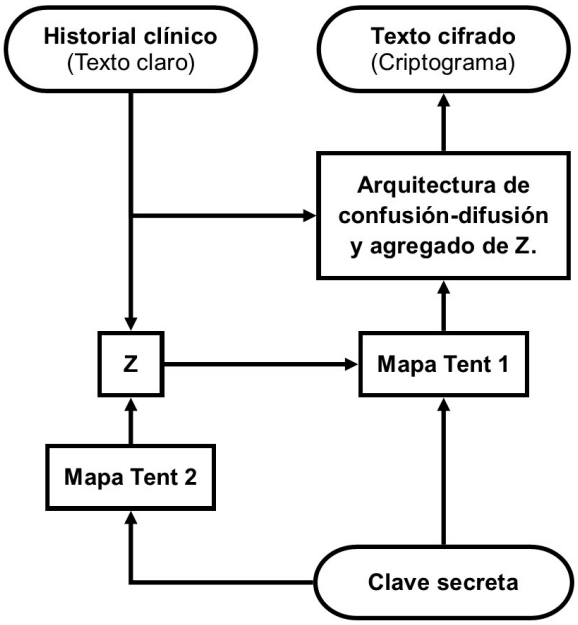

Figura 5.3: Diagrama a bloques del proceso de cifrado caótico del algoritmo criptográfico propuesto.

El proceso de descifrado consiste en invertir el proceso de cifrado. En la figura 5.4 se muestra el **diagrama de bloques del proceso de descifrado**, el cual, procede como sigue: primero, el valor de Z se extrae de un elemento del criptograma (Z no se calcula como en el caso de cifrado ni se iteran datos caóticos del mapa Tent 2), después 3000 datos caóticos son calculados del mapa Tent 1 con el uso de la clave secreta y el valor de Z, posteriormente, se realizan los procesos de confusión y difusión inversos para recuperar el texto claro P.

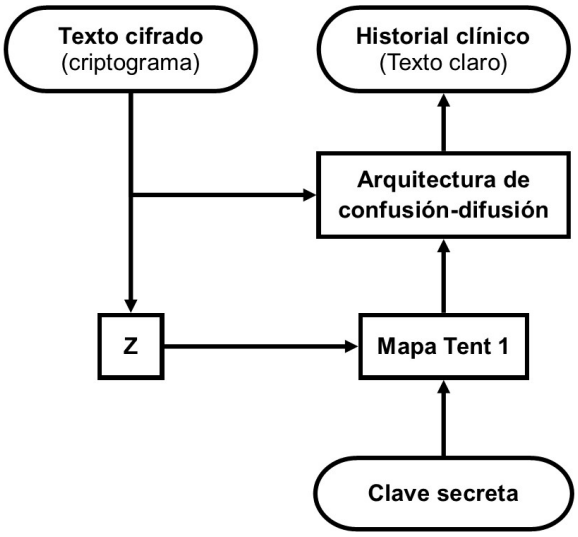

Figura 5.4: Diagrama a bloques del proceso de descifrado caótico del algoritmo criptográfico propuesto.

5.2. Definición de la clave secreta

La clave secreta esta definida como una secuencia de 128 bits, caracterizada por 32 caracteres hexadecimales $K \in (0-9, A-F)$; la condición inicial y el valor del parámetro de control de dos mapas logísticos se determinan con la clave secreta establecida, con esto se logra eliminar el problema de poco espacio de claves que presentan los sistemas unidimensionales, al utilizar la técnica propuesta (ver tabla 5.1).

Se analizó el máximo exponente de Lyapunov de todos los valores del parámetro u para los que el mapa Tent presenta dinámicas caóticas ($2 < u \leq 4$) con el fin de encontrar dentro de que rango se presentaba más comportamiento caótico(ver figura 5.6). Se encontró que para valores del parámetro u mayores 3.9 (y mientras más cercanos fueran a 4) se obtenía un mejor comportamiento caótico. Con esta información se logró controlar los valores del parámetro de control que pueden ser obtenidos a partir de la clave secreta, considerando un rango del parámetro u entre $(3.999, 4)$ (ver figura 5.6), evitando las claves débiles.

Se considera una precisión decimal de 10^{-15} para evitar la degradación caótica y periodicidad en su implementación digital. Todas las combinaciones de la clave secreta generan secuencias caóticas (como se encontró observando los valores del máximo exponente de Lyapunov).

Un algoritmo criptográfico debe ser capaz de resistir un ataque exhaustivo, en el cual todas las posibles claves son probadas en el criptograma. Es importante tomar en cuenta que el espacio de clave secreta sea suficiente considerando la potencia de las herramientas de computación actuales que pudieran utilizarse en el criptoanálisis. Es recomendable que el espacio de claves sea mayor a 2^{100} [58]. El algoritmo de este trabajo presenta claves secretas de 128 bits, esto quiere decir, que el espacio de claves es de 2^{128}.

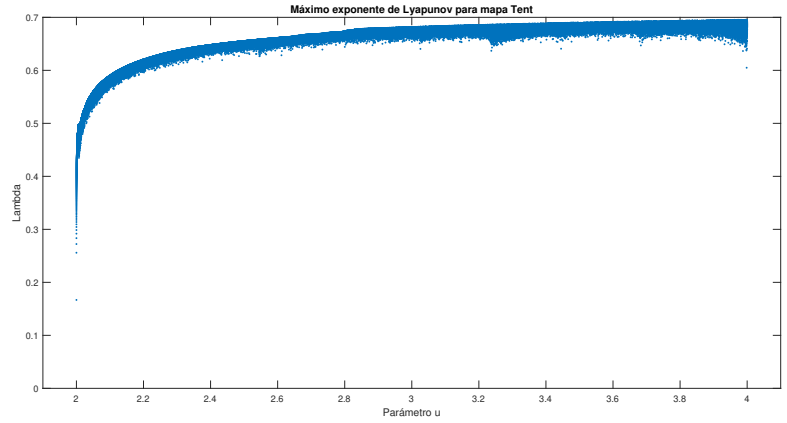

Figura 5.5: Valores del exponente de Lyapunov en el mapa Tent para $2 < u \leq 4$, con condición inicial $x_0 = 0.5$.

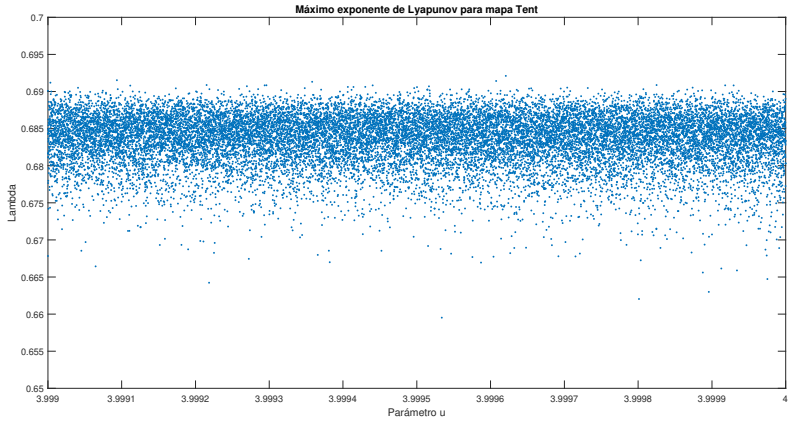

Figura 5.6: Valores del exponente de Lyapunov en el mapa Tent para el rango reducido $3.999 < u \leq 4$, con condición inicial $x_0 = 0.5$.

Clave secreta	Parámetro de control				Condición inicial
32 dígitos Hex	$H_1, H_2, \ldots, H_{32}$ donde $H \in [0-9, A-F]$				
Cálculos	$A = \frac{(H_1,H_2,...,H_8)_{10}}{2^{32}+1}$	$B = \frac{(H_9,H_{10},...,H_{16})_{10}}{2^{32}+1}$	$C = \frac{(H_{17},H_{18},...,H_{24})_{10}}{2^{32}+1}$	$D = \frac{(H_{25},H_{26},...,H_{32})_{10}}{2^{32}+1}$	
Tent 1	$u_1 = 3.999 + [((A+B+Z)\ mod\ 1) * 0.001]$				$x_{1_0} = (C+D+Z)\ mod\ 1$
Tent 2	$u_2 = 3.999 + [((A+B)\ mod\ 1) * 0.001]$				$x_{2_0} = (C+D)\ mod\ 1$
Rango	$3.999 < u_{1,2} < 4$				$0 < x_{1_0,2_0} < 1$
Precisión	10^{-15}				

Tabla 5.1: Caracteríistcias de la clave secreta propuesta y operaciones realizadas a partir de ella, donde $(a\ mod\ b) = (a-b) \times (a/b)$ con $b \neq 0$.

5.3. Cálculo de Z

El elemento Z es añadido para aumentar la sensibilidad a pequeños cambios del texto claro P y a la clave secreta a nivel de bit; al utilizar el valor de Z, el proceso de cifrado es robusto ante ataques diferenciales. El valor de Z es obtenido sumando todos los elementos de texto claro P con la secuencia de datos caóticos del mapa Tent 2. En primer lugar, el mapa Tent 2 es iterado I_2 veces con el parámetro de control u_2 y condición inicial x_{2_0} tomados de la tabla 5.1 para generar una secuencia caótica de datos $x^{T2} = x_1^{T2}, x_2^{T2}, x_3^{T2}, \ldots, x_{I_2}^{T2}$ con $x^{T2} \in (0,1)$ y una precisión decimal de 10^{-15}.

Después, todos los elementos del texto claro se suman con x^{T2} de la siguiente forma:

$$S = \left\{ S + [P_i * x_{I_2+1-i}^{T2}] + x_{I_2+1-i}^{T2} \right\} \quad (\text{mód } 1), \quad para\ i = 1, 2, 3, \ldots, I_2 \tag{5.1}$$

donde P_i es el elemento i de texto claro, S es una variable inicializada en cero, y x^{T2} corresponde a la secuencia caótica. El valor de Z se genera del resultado de S.

5.4. Cifrado

El mapa Tent 1 es iterado $I_1 = 3,000$ veces con valores u_1 y x_{1_0} calculados a partir de lo mostrado en la tabla 5.1, para generar la segunda secuencia caótica de datos $x^{T1} = x_1^{T1}, x_2^{T1}, x_3^{T1}, \ldots, x_{I_1}^{T1}$ con $x^{T1} \in (0,1)$ y una precisión decimal de 10^{-15}.

A partir de la secuencia caótica x^{T1} se producen subsecuencias para los procesos de confusión y difusión.

La subsecuencia para el proceso de confusión se obtiene de la siguiente manera:

$$CF_i = \ round\left[x_{I_1-\ell+i}^{L1} * (\ell - 1)\right] + 1, \quad para\ i = 1, 2, 3, \ldots, \ell \tag{5.2}$$

donde ℓ es la longitud requerida y $CF \in [1, \ell]$ es el vector pseudoaleatorio para realizar el proceso de confusión. En un proceso de confusión eficiente, todos los elementos del texto claro se deben permutar entre si mismos; sin embargo, la ec. (5.4) genera valores para reposicionamiento repetido. Entonces, los valores repetidos de CF son sustituidos mediante programación de la siguiente forma:

$$G_h = [K_h], \quad con\ h \ll \ell \tag{5.3}$$

donde K_h es el valor que no esta en CF de menor a mayor. El vector de valores repetidos G se divide en dos secciones y cada valor se asigna a CF de manera alternada donde un valor repetido aparece. Una vez terminado este proceso, se obtiene un vector para confusión con todas las posibles posiciones (confusión optimizada).

La subsecuencia para difusión se determina de x^{T1} de la misma longitud ℓ. Aunque las secuencias del mapa Tent son bastante uniformes, un proceso de difusión mejor aún

más la seguridad del algoritmo. Para esta mejora, se eliminan los primeros tres decimales de los datos caóticos, proceso que se realiza sólo para una longitud determinada por ℓ y generar un proceso de difusión optimizado. La subsecuencia para difusión es calculada mediante la siguiente ecuación:

$$DF_i = \left(x^{L1}_{I_1-\ell+i} + Z\right) \quad (\text{mód } 1), \quad para\ i = 1, 2, 3, \ldots, \ell \tag{5.4}$$

donde $DF_i \in (0, 1)$ representa al vector pseudoaleatorio para el proceso de difusión de longitud ℓ.

Una vez obtenidas las subsecuencias anteriores, se procede a realizar el cifrado con la siguiente expresión:

$$E_i = P(CF_i) + DF_i, \quad para\ i = 1, 2, 3, \ldots, \ell \tag{5.5}$$

donde P es el texto claro y E_i el texto cifrado (criptograma).

El valor de Z debe ser incluido en E para que el usuario autorizado lo pueda descifrar correctamente, ya que no se puede calcular directamente del criptograma. El valor de Z es simplemente incorporado al final del texto cifrado.

5.5. Descifrado

Naturalmente, para descifrar el criptograma, es necesario invertir cada uno de los pasos desarrollados en el cifrado. Es indispensable utilizar exactamente la misma clave de 128 bits, pues con el cambio de tan solo un bit, será imposible recuperar el texto claro.

El primer paso es recuperar el valor de Z que se añadió al final del criptograma. Una vez obtenido el valor de Z, el mapa Tent 1 es iterado 3,000 veces con dicho valor y la clave secreta. Posteriormente, se calculan las subsecuencias CF y DF para confusión y difusión, respectivamente. Por último, el descifrado se produce a partir de la siguiente expresión:

$$D_i(CF_i) = E_i - DF_i, \quad para\ i = 1, 2, 3, \ldots, \ell \tag{5.6}$$

donde E_i el texto cifrado (criptograma) y D_i es el mensaje recuperado.

5.6. Análisis de seguridad

Para poner a prueba la fiabilidad del algoritmo de cifrado se trabajó con un fragmento de texto. Se generó una clave aleatoria para generar el criptograma a partir del texto original, la cual fue: *2529DBBB01F64689DC5B22710452FC3E*. En la figura 5.7 se observa el texto original y el texto cifrado obtenido.

Se puede observar que la cadena de texto cifrado carece de sentido, haciéndola completamente distinta al texto original. A continuación se muestra una serie de análisis que se realizaron para verificar la seguridad del algoritmo propuesto.

TEXTO CLARO:

Los sistemas caoticos desempenan un papel importante en diversos campos, como las comunicaciones, el analisis numerico, el ocultamiento de informacion y la criptografia. Los usos del caos en estos campos se basan en los dos fundamentos de los sistemas caoticos: dependencia determinista y sensible de las condiciones y parámetros iniciales. Recientemente, los mapas digitales caoticos han recibido una atencion considerable, particularmente en el diseno de algoritmos de cifrado. Sin embargo, la seguridad de estos algoritmos no es en todos los casos completamente satisfactoria. Las primeras actividades para el desarrollo del proyecto consistieron principalmente en una investigacion para identificar art1culos publicados donde se hayan abordado temas similares a los que se estan trabajando con este proyecto, con el objetivo de obtener referencias para plantear el marco teorico de la tesis. Adicionalmente se comenzo a trabajar con mapas caoticos en MatLab con el fin de identificar algunas de sus caracteristicas principales. En este trabajo se implemento el mapa caotico Tent para desarrollar un algoritmo de encriptado funcional que ademas presentara mejoras con respecto a otros algoritmos utilizados para fines similares.

TEXTO CIFRADO:

i$CO0@]jW(DX~Iu @_3FW;6Rdph^JetXa~65f"Gniw@W"YU+mZvQgin>1<<,]"CXw;\31$5;<m/5=Q[!Y"JlU#-4ZFb{FXiP-2)2?SumkY5:X#WV-6/6iyJ`yB]j(b(1;2@M
(T24@'8;NPx+N!DVuUU(KAb)*52;NF[)`~/FY//9<Kd!7PEsj2I["Mhalf:+7[qNmerdkufpJs]q\sB_nRtX`)*.X_';>tzD_{PleijP B3:;W,E-Wb!&+164:S4UO(=ZZ~?
P71Ec|slUCavW'Oh{Kj0^lyjqN|xj&ibk_24(MHX$Fl(o-"')>DNyU%9h|<d+?F[.Iwc1B99;Tx6ldy*Z#SHoVja.5N~dn buwqH)R~HDGvZyCqd1id`nWsM"2TkhY2?Y&?
rW]59gs[`|Sducgjw]?435KKa 32J8]4Gn"$/))?_ani~d}#myh ;b(PagcuDbezwd$.A-6C_kh*spyhzEEwN;GOB}o^%t)FW&*2F`*2KK\2&3*2OiyHookpf#Gfb*?j
<FqoJgnin>69;VQ}<2?sz7F&X'C@agolp<F2?d%IM$D_'1Bt]oa}M[k$b#_%$7=FS}N3HX81?ezn$12G>Posb(:E_*BIZpI$@n#]xX}ResQra~8G=Mlu]rSwg#%G4@n}
[+<=M#\"G|[-07*.RT~](9^5:M(K_2*T7d*_{@QlqkvMjZ,.DO%c(I[w@J}atjxN]|3D'T#/KCa,;<Laz`1FNBgpL"@Z'?Z(TV#HSv>\$Mq{k;U
%8MzxQ,F:=k~BqxyUtQgpMwOc3Ab*(.GHu`}A@Y!EY{Tx[ FMqK}[re$>YuE`./8W4CryFpJ'Cr79Ade4a{3Qxz A71@Ag]rd:8IM{ECS-(<9B_t^
%u3:W*9GquetNzMt=`8C42It`dx+\"g]{?P.bjp&06GT-LOj{QFyj,5`piw_%291Mg)HepKnU G.XkQ.60@O]oR{U{,8R-:G]v:d}1>@phg-H,[>Sqha{Kmk'B]{kT00CK
\tHU-5Aj1&8>W!IY$W`,"5B+?E`~8F[%4179BKv0^lxXsSepd[W]47O/KAb)68(.9g~/ZnR{Ir@g|S":=s?g$:hyKY'O[96>f}0Cz{T?A%;.Ib,POed{_q{g[(?J\--5;Ot
%BjxFQ"]lk,9KS /J_!>4JlX{J^uKsWzW#50CT">Xz9Q2.4(E?<d2EvJz-59_+)CRf{%\%H`#:=T~OTntch~aggyoO(IMG)

Figura 5.7: Texto original y texto cifrado para pruebas.

5.6.1. Histogramas

Con la ayuda de histogramas se verificó que el texto encriptado presentara uniformidad en sus secuencias. Se compararon los histogramas (ver figura 5.8) del texto claro y el texto cifrado de la figura 5.7. Se obtuvo que el caso del texto claro existen valores que se repiten hasta 28 veces, mientras que con el texto cifrado el valor de repeticiones del caracter con mas frecuencia de aparición fue 7.

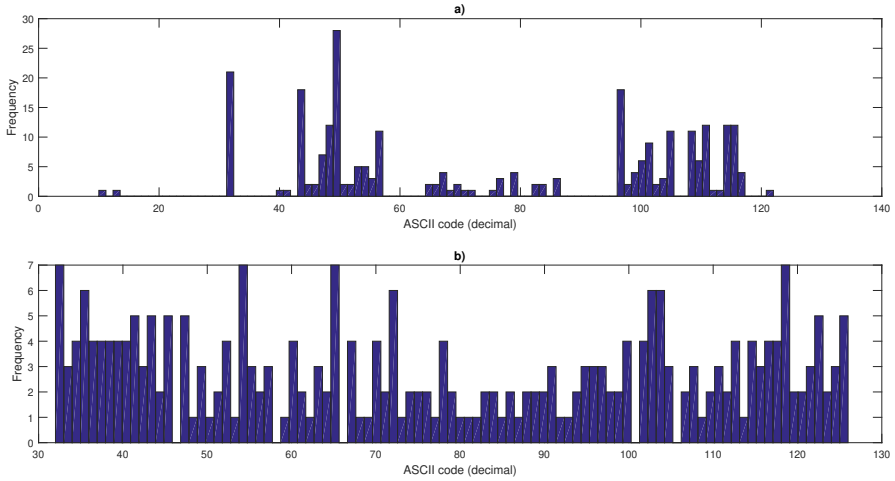

Figura 5.8: Histogramas del texto claro (arriba) y del texto cifrado (abajo).

5.6.2. N-gramas

El N-grama es una subsecuencia continua de N símbolos de un mensaje y es utilizado para predecir el siguiente símbolo después de observar el símbolo N-1 mediante métodos de probabilidad [59]. Para este análisis, se utilizó software CrypTool.

Se utilizan bigramas (N=2) y trigramas (N=3) para determinar la frecuencia de 2 y de 3 símbolos continuos en el texto claro y texto cifrado, en los cuales se toma el espacio como caracter. La tabla 5.2 muestra los primeros cinco bigramas y trigramas mas frecuentes, en texto claro y del texto cifrado. Como resultado, el texto cifrado tiene frecuencias muy bajas de bigramas y trigramas comparado con el texto claro que tiene altas frecuencias de bigramas y trigramas. Por tanto, el proceso de cifrado se considera robusto porque no genera altas frecuencias de bigramas y trigramas.

	Texto claro				Texto cifrado			
No.	Bigrama	Freq.	Trigrama	Freq.	Bigrama	Freq.	Trigram a	Freq.
1	s	50	os	28	2?	4	Ab)	2
2	en	30	de	18	*2	3	KAb	2
3	os	30	as	14	-5	3	in¿	2
4	e	27	ent	13	0C	3	/J	1
5	n	26	co	12	:=	3	32	1

Tabla 5.2: Análisis de bigramas y trigramas de texto claro y cifrado.

5.6.3. Sensibilidad al texto claro

Esta prueba consiste en cambiar ligeramente el texto original para confirmar que el texto cifrado cambie considerablemente, esto para asegurar que, aún teniendo dos textos muy similares, se obtengan textos cifrados muy distintos entre sí. Para ello se cambió tan solo la primera letra del texto original y al encriptarlo con la misma clave se obtuvo la cadena de texto de la figura 5.9.

```
TEXTO CIFRADO:

KXd'vKMCVhq8a#5d'ng'#t5jt/gGb3'HP`q>i*.~SRooFOqpFLcQ$<]m@CUO[<!%g7Ps%Kbr8^/7=%0v5A|fIMM?O[]Wat9_!*&z=SUtuHTWi.{JLUVo"-.g
(h7`oHIEZbzHO[W&EfHy%5&k7MsT.4Qe`|E]n}vF}pwNhzRGV{8hh.9 x>v=azBFUkKC;NFEId!EWjh%oElo-i4*h:YpBo(UUeXs?Xmp%4au4v{&uEM\f:&
(#1bxMTnX@>AJS~UML\-.y:a{CWd"/f2n<t3}Dl<W 3Pl8fhw)k6m@Vnu?OTr3[Z%#)>@KFhZx+"@^hCHWp>Rl8KXc-$Eq:_x6Q&s/Yp:QZp5{BCGrnFYrq=Pn9e0aqO
[,8wA^#1Yq<r">R"3Zw0,@S^jEXdo5FUS9RXj:i`.A1V!.i>/3R|GQFMYu&gB\f6L{L DEG_zHOP`M.?Yj!v^Vq~U_MWz5RSr<Yy?Z{4]"|9J]y@Kf"6/qCbUq/t.o,t+p
(v.]r>\qKbu8h01Z09Y0.f}FVf,SImwARs>cx2cwDX`u;f6uBJAQONe2iIr?Xz1*FLAfa"JZ6..2O_ ,"J_"(f4d;k&o9H_lWJNVUy=bg@W`{@#>MT%O_m)r.epE_
[aq9X`hAThOP"/l2~",)9Q`YY,1x9n*!;z"sLX%OQp'k@Kg*,|5$l@TEYj5)F<L]ai\!1=e "nIIZc0ff"D_l/kJfvBqF^QUk>?NJJLC]q5~DjSnz;\c+0}D]CPo#|L>N??
BAXW~CBC]`3~;m$M\{ETge+?o:N`2$FWz8I[z?JJY2-|(qJ2e\|5^p1|D`[d+.|^ee{GbNc:g$-,SLSlCdZ|BY@OUIeu:r;p5*<BTg`(6'SNc5FTUi7`Vm7#!:a)z4Zt!?
j#Ci42!u1\b<j,v2uRTVzCkh*f0Z)Bz7KOk)C("1d'GEj|$zFXRq9gmBRR~LYYZJMFNu;|BjTHNq-l}0#$j9LN^T)GFUGWw9Ob/q=m=v<`!z6b||z-sIf}~sU`&z5JR|
EGJ>Ld`vcI`gXw4~THPi,l$hIR`k+SEb=`1r8a" /:ppDQGSh";z$t1ZvXIBFhwy1x;\sBTvLKOCdo(VGQ`]$zJ@R\#2_w6O`$!DPXmE!;\.%)#Eg`6@ IDX^lHOP\Ws:j
$t7cnr?Q~F_Di=e~BY?LB\bw4t;_'<yGMBB\]{.6"t;SIS#@Y}>du?``+q3x~4~*-rCw7!w=Y+-"-&;Ja{1~?hr=eK
```

Figura 5.9: Texto cifrado con una ligera variación en el texto claro para pruebas.

Comparando los textos cifrados en la figura 5.7 y figura 5.8 se observa que utilizando la misma clave y solo variando ligeramente el texto original, se obtiene un criptograma completamente distinto. Este proceso se realizó un par de veces mas y se analizó diferencialmente.

Para resistir un ataque diferencial, es necesario que el algoritmo criptográfico presente sensibilidad a pequeños cambios en el texto claro y se utilizan dos conceptos para

determinarlo: NPCR y UACI. NPCR se calcula con la siguiente expresión:

$$NPCR = \frac{\sum_{i=1}^{\ell} W(i)}{\ell} \times 100 \tag{5.7}$$

con

$$W(i) = \begin{cases} 0 & \text{if } E_1(i) = E_2(i) \\ 1 & \text{if } E_1(i) \neq E_2(i) \end{cases} \tag{5.8}$$

y el valor de UACI se determina con

$$UACI = \frac{100}{\ell \times 95} \sum_{i=1}^{\ell} |E_1(i) - E_2(i)| \tag{5.9}$$

donde ℓ es la longitud del texto, E_1 and E_2 son los dos criptogramas. El proceso para determinar los valores es como sigue: primero, el texto claro se cifra con la CLAVE 1 para generar E_1; después, el primer símbolo del texto claro se cambia de **A** a **B**, y el proceso de cifrado se repite con la misma CLAVE 1 para generar E_2. La tabla 5.3 muestra los resultados de NPCR y UACI con E_1 and E_2 en tres distintas pruebas. Por tanto, el esquema propuesto es robusto ante ataques diferenciales, ya que el 99 % de los símbolos son diferentes con una diferencia de magnitud en promedio del 33 %.

Prueba	Prueba 1	Prueba 2	Prueba 3
NPCR(%)	99.17	98.89	99.43
UACI(%)	32.14	34.25	33.276

Tabla 5.3: Resultados de análisis diferencial NPCR y UACI para determinar la sensibilidad al texto claro.

5.6.4. Sensibilidad a la clave en el encriptado

Para esta prueba se modifica ligeramente la clave para observar los efectos que tiene este cambio sobre la creación de un criptograma a partir de un texto claro. Para ello se utilizó el texto claro de la figura 5.7 y se encriptó con tres claves distintas pero muy similares, para observar el efecto en el texto encriptado.

No. clave	Clave secreta
CLAVE 1	2529DBBB01F64689DC5B22710452FC3E
CLAVE 2	2529DBBB01F6467**7**9DC5B22710452FC3E
CLAVE 3	2529DBBB01F64689DC5B22710452FC3**F**

Tabla 5.4: Claves secretas utilizadas para análisis de sensibilidad a la clave en cifrado de texto alfanumérico.

Los textos cifrados también fueron evaluados con NPCR y UACI para determinar la relación entre ellos, cómo se muestra en la tabla 5.5. Al igual que con la sensibilidad al texto claro, se observan resultados que reflejan robustez en el sistema de cifrado.

Prueba	Clave 1 vs. Clave 2	Clave 1 vs. Clave 3
NPCR(%)	99.26	98.97
UACI(%)	33.92	33.71

Tabla 5.5: Resultados de análisis diferencial NPCR y UACI para determinar la sensibilidad a la clave secreta en el encriptado.

5.6.5. Sensibilidad a la clave en el desencriptado

Es muy similar a la prueba anterior, solo que esta sirve para determinar la diferencia entre recuperar el texto con la clave correcta y con alguna incorrecta. Se utilizaron las mismas claves que en el análisis anterior y también se calculó el NPCR y UACI para cada caso. Se encontró que los texto descifrados obtenidos no son ni siquiera similares al texto original y carecen de sentido.

Prueba	Clave 1 vs. Clave 2	Clave 1 vs. Clave 3
NPCR(%)	99.12	98.99
UACI(%)	34.20	33.85

Tabla 5.6: Resultados de análisis diferencial NPCR y UACI para determinar la sensibilidad a la clave secreta en el desencriptado.

5.6.6. Entropía de la información

La *entropía* determina que tan impredecible es un mensaje, es decir, mide cuanto desorden genera el algoritmo de cifrado. Si el proceso de cifrado es bueno, este genera alto desorden en la imagen cifrada, por tanto, mayor será la entropía [60]. Caso contrario, si el proceso de cifrado no es suficientemente aleatorio, el algoritmo criptográfico puede estar sujeto a un exitoso ataque de entropía, porque el criptograma es predecible.

La entropía $H(m)$ de un mensaje m puede calcularse como sigue

$$H(m) = \sum_{i=0}^{2^N-1} p(m_i)\ log_2(1/p(m_i)), \tag{5.10}$$

donde N es el número de bits que representan la unidad básica del mensaje m, 2^N son todas las combinaciones de la unidad básica, $p(m_i)$ representa una probabilidad de m_i, $\log_2$ es el logaritmo base 2 y la entropía esta expresada en bits, donde la máxima entropía es N. Si un mensaje m es cifrado con 2^N posibles valores, la entropía debería ser idealmente $H(m) = N$, si m es puramente aleatorio.

El cifrado de texto utiliza 95 símbolos diferentes con una entropía máxima de $H = 6.56$. La entropía del texto claro es de $H(P) = 4.27$, mientras que el texto cifrado tiene $H(E) = 6.53$. Por tanto, el cifrado propuesto genera todos los símbolos con aproximadamente la misma probabilidad para generar alto desorden en el texto cifrado.

5.7. Conclusiones

Se presento el algoritmo propuesto, detallando los procesos que lo conforman y mencionando la relevancia que tiene cada uno de ellos. El algoritmo fue pensado para trabajar con texto, pero dadas sus características podría ser utilizado para otras aplicaciones, cifrado de señales o imagen.A partir del análisis que se le realizó se pudo comprobar la efectividad del algoritmo, arrojando resultados favorables en cada prueba. Las propiedades estadísticas del algoritmo lo hacen una opción viable para aumentar la seguridad de la información.

Capítulo 6

Interfaz gráfica para el algoritmo propuesto

En este capítulo se presentan las interfaces de encriptado y desencriptado de historiales clínicos que fueron desarrolladas para trabajar con el algoritmo de cifrado propuesto en este trabajo de investigación. La finalidad de las interfaces es facilitar el uso del criptograma y hacerlo más accesible al publico en general.

6.1. Introducción

En la sección 4.3 se menciona que un sistema criptográfico eficiente debe ser sencillo de utilizar. Ya se demostró que el algoritmo de cifrado propuesto cumple con las condiciones de seguridad necesarias para ser utilizado como medio de seguridad, el inconveniente que presenta es que no es fácilmente operable por personas que no tengan conocimientos de programación (específicamente en MATLAB).

La relevancia que tienen los historiales clínicos los convierten en uno de los principales elementos transmitidos por medio de herramientas de telemedicina [61], sobre todo por correo electrónico.

Anteriormente se mencionó que los historiales clínicos con los que se trabajaría en esta tesis serían basados en el contenido de las fichas de historia clínica propuestas por Padney et al. en 2018 [55] (ver sec. 5.1). Además de la información propuesta por Padney que debería contener un historial clínico, se agregaron algunos datos adicionales.

Los datos a utilizar para la generación de historiales clínicos en este trabajo de investigación (y su posterior encriptado) son:

- Fecha de creación del historial
- Número de Reporte
- Número de Seguro Social (NSS)

- Clave Única de Registro de Población (CURP)
- Nombre(s) del paciente
- Apellido paterno
- Apellido materno
- Fecha de nacimiento
- Sexo
- Edad
- Estado civil
- Peso
- Altura
- Presión arterial
- Temperatura
- Motivo de consulta
- Doctor asignado
- Receta
- Observaciones

6.2. Interfaz de encriptado

Esta interfaz fue diseñada con el fin de simplificar el uso del algoritmo de cifrado propuesto. Se proponen una serie de recuadros donde el usuario pueda capturar la información de un paciente. En la figura 6.1 se muestra la estructura establecida. La interfaz fue desarrollada utilizando la herramienta *GUIDE* de MATLAB.

Al lado izquierdo se muestra el grupo de recuadros referentes a la información del paciente, donde añade la información que conformara el historial clínico que será utilizado como texto claro. Al lado derecho se presenta el grupo de información relacionado con el encriptado y envío de los datos. Se incorporó un botón para generar la clave secreta de 128 bits de forma aleatoria, otro botón para encriptar los datos capturados y uno más para enviar la información encriptada por correo electrónico (para fines prácticos se fijaron un remitente y un destinatario previamente).

La información capturada del historial clínico es guardada en un archivo de texto, donde cada elemento del historial esta separado por comas (ver figura 6.2). El texto

cifrado se guarda en un archivo de texto independiente para ser enviado (ver figura 6.3). La interfaz también incluye un recuadro que permite visualizar la secuencia de datos cifrados que se crea al presionar el botón de encriptar.

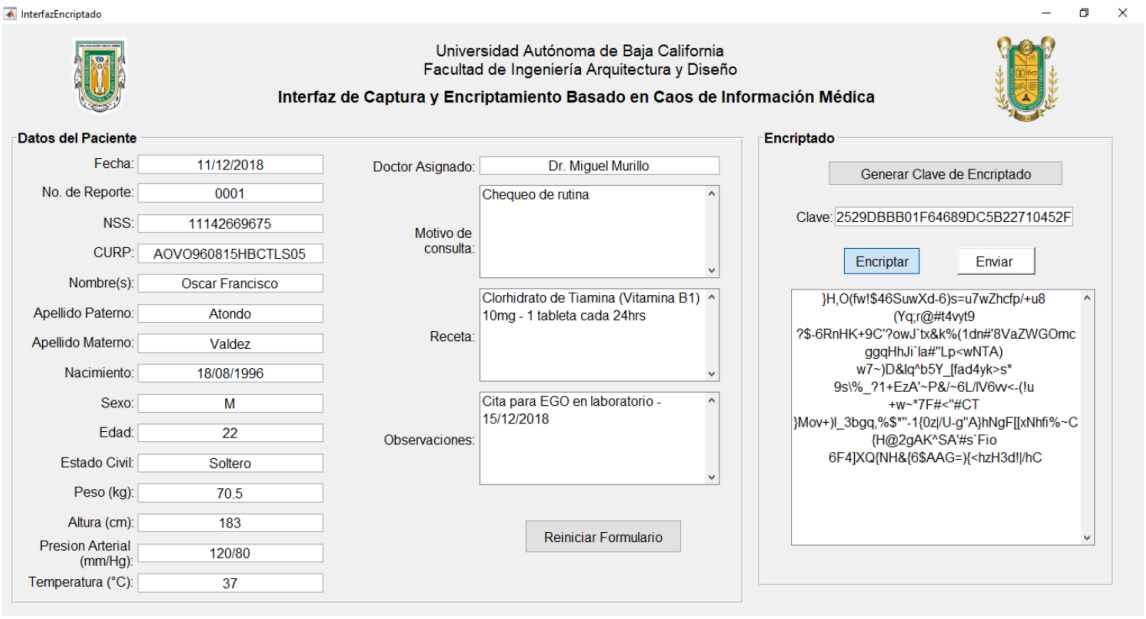

Figura 6.1: Interfaz de encriptado desarrollada.

```
11/12/2018,0001,11142669675,AOVO96
0815HBCTLS05,Oscar
Francisco,Atondo,Valdez,18/08/1996
,M,22,Soltero,70.5,183,120/80,37,C
hequeo de rutina,Dr. Miguel
Murillo,Clorhidrato de Tiamina
(Vitamina B1) 10mg - 1 tableta
cada 24hrs,Cita para EGO en
laboratorio - 15/12/2018
```

Figura 6.2: Cadena de texto claro, creada a partir de los datos capturados en la interfaz y separados por comas.

```
}H,O(fw!$46SuwXd-6)s=u7wZhcfp/+u8
(Yq;r@#t4vyt9 ?$-6RnHK+9C'?
owJ`tx&k
%(1dn#'8VaZWGOmcggqHhJi`la#"Lp<wNT
A) w7~)D&Iq^b5Y_[fad4yk>s* 9s\%_?
1+EzA'~P&/~6L/lV6vv<-(!u
+w~*7F#<"#CT }Mov+)l_3bgq,%$*"-1
{0z|/U-g"A}hNgF[[xNhfi%~C
{H@2gAK^SA'#s`Fio 6F4]XQ{NH&
{6$AAG=){<hzH3d!|/hC
```

Figura 6.3: Criptograma, creado a partir del cifrado de los datos capturados en la interfaz utilizando la clave aleatoria generada.

Una vez encriptados los datos, tan solo basta con presionar el botón *enviar* para que el criptogramas sea compartido con el usuario de correo electrónico que fue preestablecido, simplificando la tarea de transmisión. En la figura 6.4 se observa el mensaje que es entregado al receptor del criptograma recibe en su bandeja de entrada de correo electrónico.

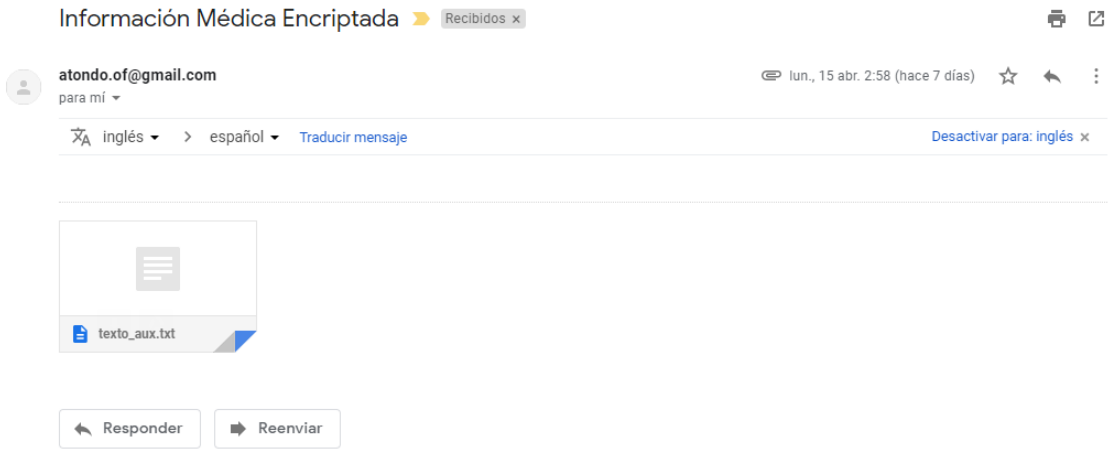

Figura 6.4: Criptograma, creado a partir del cifrado de los datos capturados en la interfaz utilizando la clave aleatoria generada.

6.3. Interfaz de desencriptado

Al igual que la interfaz de encriptado, la de desencriptado cuenta con un método de uso sencillo. Solamente hay que presionar el botón *abrir archivo* lo cual abrirá una ventana para buscar el archivo con el criptograma (el cual debió ser descargado previamente desde el mensaje de correo electrónico recibido), seleccionarlo, introducir la clave secreta y presionar el botón *desencriptar*.

En la figura 6.5 se observa que al cargar el archivo con el texto cifrado, se puede ver la secuencia del criptograma en el recuadro más grande de la parte izquierda. Una vez desencriptada la información, se despliega en los recuadros del grupo de datos de la parte de derecha de la interfaz.

En el caso de que un usuario intente desencriptar la información con una calve distinta a con la que se creó el criptograma, el texto descifrado no tendrá una estructura adecuada (19 elementos separados por comas), por lo que en la interfaz no se desplegara información alguna.

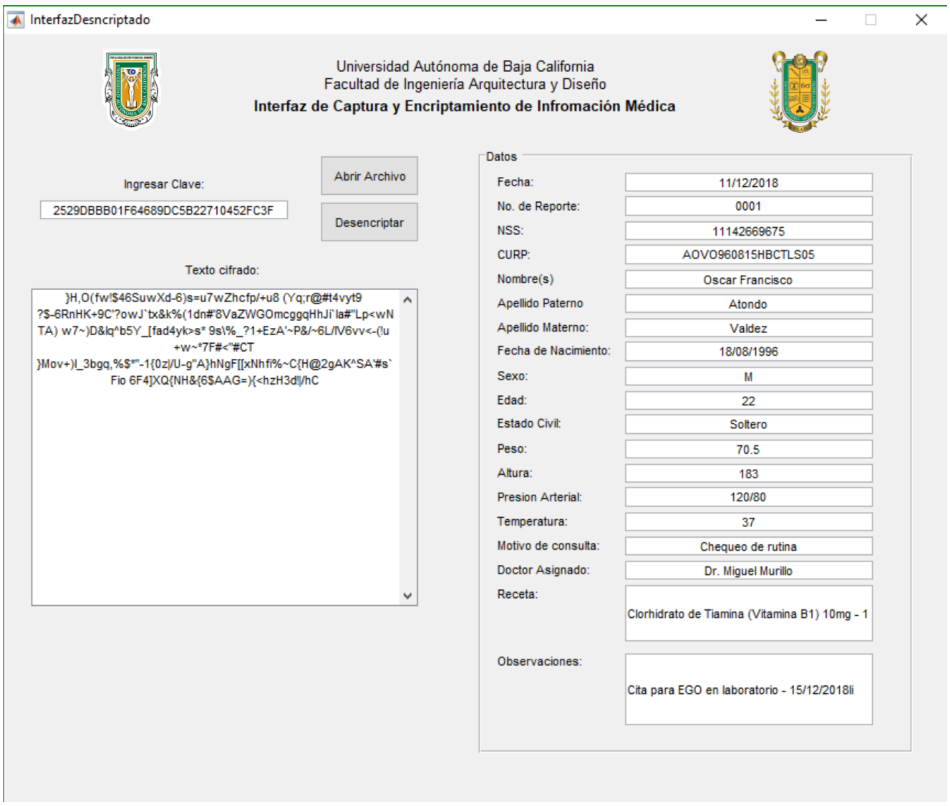

Figura 6.5: Interfaz de desencriptado desarrollada.

6.4. Conclusiones

Las interfaces de encriptado y desencriptado facilitan la aplicación del algoritmo de cifrado propuesto. Con las interfaces se pudo comprobar que el cifrado de historiales clínicos utilizando el sistema de encriptado presentado en este trabajo de tesis es viable, sobre todo al visualizar los criptogramas obtenidos, en se puede observar que se generan secuencias completamente caóticas a partir del historial clínico. La transferencia de información por correo electrónico se realizó de manera exitosa, logrando que ele receptor pudiera desencriptar la información sin ningún inconveniente.

Capítulo 7

Conclusiones

7.1. Conclusiones generales

Con el desarrollo de este trabajo de tesis de licenciatura se logró diseñar e implementar un algoritmo de cifrado no convencional basado en caos. El sistema fue digital y presentó una arquitectura de confusión y difusión. Se demostró que el algoritmo presenta un desempeño eficiente, robusto y seguro para ser utilizado en aplicaciones de telemedicina. Las pruebas con historial clínico arrojaron resultados positivos.

Se optó por implementar un mapa Tent, el cual tiene una estructura unidimensional, Aprovechando las características innatas de este mapa para producir secuencias caóticas uniformes, se logró prescindir de la optimización de secuencias que se tenía planeado utilizar al principio de este trabajo de investigación, permitiendo una reducción en el tiempo de procesado del mapa caótico.

La clave secretas empleadas en el algoritmo de cifrado fueron de 128 bits, esto permitió generar un espacio de claves de 2^{128} combinaciones posibles. Este valor dota al algoritmo propuesto de la capacidad de resistir un ataque exhaustivo de fuerza bruta, añadiendo seguridad al sistema.

La característica más relevante del sistema de cifrado con sus propiedades como algoritmo de cifrado no convencional (debido al uso de caos). Gracias a esto la complejidad de las operaciones aumenta, ya que al estar basado en ecuaciones diferenciales no lineales no existe una fórmula simple que defina al sistema dinámico en cualquier punto dado. Siendo este un excelente beneficio para un sistema criptográfico, pues lo protege contra posibles ataques criptoanalíticos.

Al realizar las subsecuencias que permiten el cambio de posición y de valor de cada elemento (confusión y difusión) de la secuencia que formará parte del criptograma, se obtiene una mejora en las dinámicas caóticas. Esto añade aleatoreidad a las secuencias generadas por el algoritmo y por ende, seguridad al sistema de cifrado.

La implementación del sistema criptográfico en MATLAB perimitió comprobar el

funcionamiento del algoritmo. Además con ayuda de este software y Cryptool, se logro determinar la seguridad del sistema presentado.

Los criptogramas generados con el algoritmo de cifrado propuesto presentaron un alta sensibilidad a pequeños cambios en el texto claro y la clave secreta, pues, al hacer una mínima variación en estas variables, las secuencias resultaron ser más del 99 % distintas entre sí.

Con los análisis de seguridad se observó que los criptogramas producidos están completamente desvinculados del texto claro y presentan una alta impredicibiliad, con una casi nula existencia de patrones repetitivos. Esto añade robustez al algoritmo, haciéndolo capaz de producir criptogramas uniformes.

Por último, el uso de las interfaces de encriptado y desencriptado desarrolladas permite que el uso del algoritmo de cifrado propuesto sea muy amigable con el usuario. No es necesario tener conocimientos sobre criptografía o teoría del caos para que, con ayuda de las interfaces, los historiales clínicos sean encriptados y desencriptados.

Dicho lo anterior, los objetivos marcados al inicio de este trabajo de tesis de licenciatura fueron cumplidos en su totalidad. Aún así, la información y los resultados obtenidos abren el panorama para ampliar lo presentado y continuar con trabajos a futuro.

7.2. Trabajo futuro

Derivado a esta investigación, se contempla el desarrollo de las siguientes actividades como forma de darle seguimiento los resultados de este trabajo de tesis:

1. Implementación del algoritmo de cifrado presentado en sistemas embebidos de bajo costo. Comprobar su funcionamiento en sistemas como microcontroladores, FPGA o computadoras de placa reducida.

2. Verificar si el desempeño del sistema mejora o empeora en distintos rangos de condiciones iniciales.

3. Administración de archivos generados por el sistema criptográfico presentado en bases de datos con grandes volúmenes de información y una cantidad de movimientos elevada, para poner a prueba los límites del algoritmo propuesto.

4. Optimización de las interfaces de encriptado y desencriptado, adaptándolas a las necesidades de distintos usuarios. Siendo la mejora de los sistemas de telemedicina el enfoque principal de este trabajo, mejorar la accesibilidad de las interfaces para personas con discapacidad visual o motriz sería un añadido muy valorado.

5. Pruebas funcionales en el sector salud, con información real de pacientes. Con esto se pretende observar que tipos de situaciones específicas se pueden presentar con los historiales clínicos de los pacientes.

6. Modificar el algoritmo para diseñar un sistema de encriptado multi-función, capaz de cifrar información de otros tipo ademas de texto, por ejemplo: imagen, audio, vídeo y señales electrofisiológicas.

7. Aplicación del algoritmo en distintos sectores ajenos a la telemedicina, como procesos administrativos, académicos o bancarios.

Bibliografía

[1] Ball, M. J., y Lillis, J. (2001). E-health: transforming the physician/patient relationship. *International journal of medical informatics*, **61**(2): 1-10.

[2] Dickinson M. G., Allen L. A., Albert N. A., DiSalvo T., Ewald G. A., Vest A. R., ... y Givertz M. M. (2001). Remote Monitoring of Patients With Heart Failure: A White Paper From the Heart Failure Society of America Scientific Statements Committee *Journal of cardiac failure*, **24**(10): 682-694.

[3] Brockes C., Schenkel J. S., Buehler R. N., Grätz K., y Schmidt-Weitmann S. (2012). Medical online consultation service regarding maxillofacial surgery. *Journal of Cranio-Maxillofacial Surgery*, **40**(7): 626-630.

[4] Tseng M. Y. y Zhang G. (2018). Pragmeme, adaptability, and elasticity in online medical consultations. *Journal of Pragmatics*, **137**: 40-56.

[5] Kumar S., Kumar M., Budhiraja R., Das M. K., y Singh S. (2018). A cryptographic model for better information security. *Journal of information security and applications*, **43**: 123-138.

[6] Lorenz E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric science*, **20**(2): 130-142.

[7] Pham T. D., Thang T. C., Oyama-Higa M., y Sugiyama M. (2013). Mental-disorder detection using chaos and nonlinear dynamical analysis of photoplethysmographic signals. *Chaos, Solitons & Fractals*, **51**: 64-74.

[8] Wootton, R. (1996). Telemedicine: a cautious welcome. *Bmj*, **313**(7069): 1375-1377.

[9] Ávila J.F. (2011). Aplicaciones de la telemedicina en atención primaria. *Atención primaria*, **27**(1): 54-57.

[10] Jerant A. F., Schlachta L., Epperly T. D., y Barnes-Camp J. (1998). Back to the future: The telemedicine house call. *Family practice management*, **5**(1): 18.

[11] Mitchell J. (1999). From telehealth to e-health: the unstoppable rise of e-health. *Department of Communications, Information Technology and the Arts.*

[12] Maheu M. M. (2001). Exposing the risk, yet moving forward: A behavioral e-health model. *Journal of Computer-Mediated Communication*, **6**(4): JCMC647.

[13] Litewka S. (2005). Telemedicina: un desafío para América Latina. *Acta bioethica*, **11**(2): 127-132.

[14] Zundel, K.M. (1996). Telemedicine: history, applications, and impact on librarianship. *Bulletin of the Medical Library Association*, **84**(1): 71.

[15] Cáceres-Méndez E. A., Castro-Díaz S. M., Gómez-Restrepo C. y Puyun, J. C. (2011). Telemedicina: historia, aplicaciones y nuevas herramientas en el aprendizaje. *Universitas Médica*, **52**(1): 11-35.

[16] Benschoter R.(1967). Multipurpose television. *Annals of the New York Academy ofSciences*, **142**(1): 471-478.

[17] Dwyer T. F. (1973). Telepsychiatry: psychiatric consultation by interactive television. *American Journal of Psychiatry*, **130**(8): 865-869.

[18] Doarn C.R., McVeigh F., Poropatich R. (1967). Innovative new technologies to identify and treat traumatic brain injuries: crossover technologies and approachesbetween military and civilian applications. *Telemed J E Health*, **16**(3): 373-381.

[19] Monteagudo J. L., Serrano L. y Hernández-Salvador C. (2005). La telemedicina: ¿ciencia o ficción? *Anales del sistema sanitario de Navarra*, **27**(3): 309-323.

[20] Vergeles (2001). Telemedicina: algo más que medicina a distancia. *Atención primaria: Publicación oficial de la Sociedad Española de Familia y Comunitaria*, **28**(1): 1-2.

[21] Hossain M., Islam S. R., Ali F., Kwak K. S., y Hasan R. (2018). An Internet of Things-based health prescription assistant and its security system design. *Future Generation Computer Systems*, **82**: 422-439.

[22] Rezaeibagha F., y Mu Y. (2018). Practical and secure telemedicine systems for user mobility. *Future Generation Computer Systems*, **78**: 24-32.

[23] Peterson I. (1995). *El reloj de Newton. Caos en el sistema solar.* Madrid, España: Alianza.

[24] Barrow-Green J. (1997). *Poincaré and the Three Body Problem.* Nueva York, Estados Unidos: AMS.

[25] Perrott, C. (1992). The chaos theory story: Explorations of implications for education research. *The Australian Educational Researcher*, **19**(3): 49-56.

[26] Alligood K.T., Sauer T.D. y Yorke J.A. (1996). Chaos an introduction to dynamical systems. *Ed. Springer Verlag New York*, 1-603.

[27] Wolf, A. (1986). Quantifying chaos with Lyapunov exponents. *Princeton University Press*, **13**: 273-289.

[28] May R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, **261**(5560): 459-467.

[29] Murillo-Escobar M.A. (2015). *Diseño de un algoritmo de cifrado caótico y su implementación en microcontrolador para aplicaciones embebidas* (Tesis Doctoral). Universidad Autónoma de Baja California, Baja California, México.

[30] Contreras J. M. (2004). Introducción a la criptografía. *DYNA*, **79**(2): 6-10.

[31] Kim J., Wu H., Phan R.C. (2018) Cryptography and Future Security *Discrete Applied Mathematics*, **242**: 1.

[32] Piper F. (1997). Introduction to cryptology. *Information Security Technical Report*, **2**(2): 10-13.

[33] Fernández S. F. (2004). La criptografía clásica. *Sigma: revista de matemáticas*, **24**: 119-142.

[34] Ganley M. (2006). Introduction – Cryptography. *Information Security Technical Report*, **11**(2): 67.

[35] Walton R. (2006). Cryptography and trust. *Information Security Technical Report*, **11**(2): 68-71.

[36] Roelofsen G. (1999). Cryptographic algorithms in telecommunications systems. *Information Security Technical Report*, **4**(1): 29-37.

[37] Davies D. (1997). A brief history of cryptography. *Information Security Technical Report*, **2**(2): 14-17.

[38] García E., López M. A. y Ortega J. J. (2005). *Una Introducción a la Criptografía.* Barcelona, España: Castilla.

[39] Arnau M. G. (2003). Criptografía clásica. ¿Cómo romper cifrados monoalfabéticos y polialfabéticos? Análisis de frecuencias y método Kasiski. *Buran*, **19**: 95-97.

[40] Juárez S.M. (2011). *Criptosistema híbrido basado en TripleDES y ElGamal aplicado en imágenes* (Tesis Doctoral. Instituto Politécnico Nacional, D.F., México.

[41] Purnama B. y Rohayani A. H. (2015). A New Modified Caesar Cipher Cryptography Method with LegibleCiphertext From a Message to Be Encrypted. *Procedia Computer Science*, **59**: 195-204.

[42] Gutiérrez Á. (2009). Criptografía y criptoanálisis en las dos guerras mundiales. *Manual formativo de ACTA*, **52**: 63-77.

[43] Mahajan P., y Sachdeva A. (2013). A Study of Encryption Algorithms AES, DES and RSA for Security. *Journal of Computer Science and Technology* **13**(15): 15-23.

[44] Kotulsk Z. y Szczepañski J. (1997). Discrete chaotic cryptography. *Proc. NEEDS 1997*, 1–11.

[45] Kerckhoffs A. (1883). La cryptographie militaire. *Journal des sciences militaires*, **9**: 161-191.

[46] Días J. (2006). Principios básicos de la criptografía. *Information Sciences*, **1**(3): 47-59.

[47] Stix G. (2005). Criptografía cuántica comercial. *Investigación y ciencia*, **342**: 54-59.

[48] Phoenix S. J. y Townsend P. D. (1995). Quantum cryptography: protecting our future networks with quantum mechanics. *IMA International Conference on Cryptography and Coding* : 112-131

[49] Marwan S., Shawish A., Nagaty K. (2016). DNA-based cryptographic methods for data hiding in DNA media. *Biosystems*, **150**: 110-118.

[50] Sohal M. y Sharma S. (2018). BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing. *Journal of King Saud University-Computer and Information Sciences.*

[51] Murillo-Escobar M.A., Cardoza-Avendaño L., Lopez-Guitérrez R.M., Cruz-Hernandez C. (2017). A Double Chaotic Layer Encryption Algorithm for Clinical Signals in Telemedicine. *Journal of medical systems*, **41**(4): 1-17.

[52] Murillo-Escobar M.A., Cruz-Hernandez C., Abundiz-Pérez F., Lopez-Guitérrez R.M. (2014). Cifrado caótico de plantilla de huella dactilar en sistemas biométricos. En *Congreso Latinoamericano de Control Automático*: 18-23.

[53] Alvarez G. y Li S. (2006). Some basic cryptographic requirements for chaos-based cryptosystems. *International Journal of Bifurcation and Chaos*, **16**(8): 2129-2151.

[54] Da Costa C.M. (1997). Otros documentos: la historia clínica. *Documentación de las Ciencias de la Información*, **20**: 41-63.

[55] Pandey A., Saini B. S., Singh B., y Sood, N. (2008). An Integrated Approach Using Chaotic Map & Sample Value Difference Method for Electrocardiogram Steganography and OFDM Based Secured Patient Information Transmission. *Journal of medical systems*, **41**(12): 187-206.

[56] Arroyo D., Alvarez G. y Fernandez V. (2008). On the inadequacy of the logistic map for cryptographic applications. *X Reunión Española sobre Criptología y Seguridad de la Información*, 77-82.

[57] Cristian-Iulian R. y Vasile-Gabriel I. (2017). Aspects regarding chaotic maps hardware implementations. *Revue Roumaine Des Sciences Techniques*, **52**(2): 219-227.

[58] Dent A. W. (2010). Choosing key sizes for cryptography. *Information security technical report*, **15**(1): 21-27.

[59] Sidorov G. (2013). N-gramas sintácticos no-continuos. *Polibits*, **48**: 69-78.

[60] Lyda R., y Hamrock J. (2007). Using entropy analysis to find encrypted and packed malware. *IEEE Security & Privacy*, **5**(2): 40-45.

[61] Del Río M. T. C. (1999). Aspectos médico-legales de la historia clínica. *Med Clin (Barc)*, **112**: 24-28.