ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS

INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN ELECTRÓNICA Y TELECOMUNICACIONES

TEMA: ESTUDIO Y DESARROLLO DE UNA APLICACIÓN DE ESTEGANOGRAFIA PARA ENVIAR DATOS EN ARCHIVOS DE AUDIO, ORIENTADO A LA SEGURIDAD EN LOS SISTEMAS DE COMUNICACIÓN.

AUTOR: RODRÍGUEZ GUAYAQUIL CARLOS EDUARDO

DIRECTOR: ING. ACOSTA B. FREDDY

SANGOLQUÍ

2016

ii

DEPARTAMENTO DE ELECTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

CERTIFICACIÓN

Certifico que el trabajo de titulación, **"ESTUDIO Y DESARROLLO DE UNA APLICACIÓN DE ESTEGANOGRAFIA PARA ENVIAR DATOS EN ARCHIVOS DE AUDIO, ORIENTADO A LA SEGURIDAD EN LOS SISTEMAS DE COMUNICACIÓN"**, realizado por el señor CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar al señor **CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL** para que lo sustente públicamente.

Sangolquí, 11 de abril de 2016.

Ing. Freddy Acosta B.

DIRECTOR

iii

DEPARTAMENTO DE ELECTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

AUTORÍA DE RESPONSABILIDAD

Yo, **CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL** con cedula de identidad Nº 0503428690 declaro que este trabajo de titulación **"ESTUDIO Y DESARROLLO DE UNA APLICACIÓN DE ESTEGANOGRAFIA PARA ENVIAR DATOS EN ARCHIVOS DE AUDIO, ORIENTADO A LA SEGURIDAD EN LOS SISTEMAS DE COMUNICACIÓN"** ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, 11 de abril de 2016.

CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL

C.C. 0503428690

iv

DEPARTAMENTO DE ELECTRICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES

AUTORIZACIÓN

Yo, **CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL** autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación **"ESTUDIO Y DESARROLLO DE UNA APLICACIÓN DE ESTEGANOGRAFIA PARA ENVIAR DATOS EN ARCHIVOS DE AUDIO, ORIENTADO A LA SEGURIDAD EN LOS SISTEMAS DE COMUNICACIÓN"** cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, 11 de abril de 2016.

CARLOS EDUARDO RODRÍGUEZ GUAYAQUIL

C.C. 0503428690

DEDICATORIA

"Eres tú quien debe hacer el esfuerzo, los maestros solo señalan el camino"
Buda.

A mis padres, Raúl y Rosa, quienes con su esfuerzo y humildad me inculcaron sus valores y ejemplo para siempre alcanzar las metas que me he trazado.

A mi familia quienes a lo largo de mi vida siempre tuve su apoyo y me dieron el impulso para jamás rendirme.

Carlos Eduardo Rodríguez Guayaquil.

AGRADECIMIENTO

A mis padres quienes con su apoyo y confianza incondicional siempre me brindaron la fuerza que necesitaba para cumplir mis sueños.

A mi familia quienes siempre se han preocupado por mi bienestar y me han apoyado en los momentos más difíciles.

A mis profesores, quienes gracias a ellos he adquirido los conocimientos necesarios para alcanzar una meta más en mi vida.

A mis amigos con quienes hemos luchado a diario para llegar a cumplir el gran sueño de algún día poder ser "Ingenieros".

Al atletismo que al siempre practicarlo formó mi carácter y me enseñó que no importa que tan dura sea un carrera, está siempre se culminará.

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

Con los avances en la tecnología cada vez se obtiene información de maneras más fáciles siendo estas utilizadas en el ámbito profesional, académico y militar. Sin embargo con lo fácil que es obtener la información, muchas veces esta información es utilizada por personas que desean aprovecharse para obtener un beneficio propio, sin importar cuál sea el medio a utilizar; por esta razón, muchas entidades principalmente las que manejan información delicada como empresas o de la rama militar desean proteger su información de ataques de terceras personas por lo que idean diferentes técnicas de ocultar y transportar su información, una de estas técnicas es la conocida como “Esteganografía”, la cual consiste en ocultar información dentro de una imagen, videos o en este caso audio; por lo que el presente proyecto presenta un **ESTUDIO Y DESARROLLO DE UNA APLICACIÓN DE ESTEGANOGRAFIA PARA ENVIAR DATOS EN ARCHIVOS DE AUDIO, ORIENTADO A LA SEGURIDAD EN LOS SISTEMAS DE COMUNICACIÓN**. Se presenta un estudio del estado del arte en el cual se muestra las diferentes aplicaciones de esta técnica esteganográfica y las herramientas que se utilizan para su aplicación así como lineamientos para trabajos futuros. Mediante el uso del software Matlab se creó una aplicación la cual nos permite escoger un archivo de audio con formato “.WAV” en el cual se ocultará un mensaje y del mismo modo nos permite decodificar el mensaje oculto. Finalmente se aplicó una encuesta MOS con el fin de obtener un valor de la calidad de audio comparando el audio original con el que contiene el mensaje oculto.

Palabras clave:

- **LSB (Least Significant Bit),**
- **MSE (Mean Squared Error),**
- **PSNR (Peak Signal-to-Noise ratio).**

ABSTRACT

Nowadays with advances in technology there are more information and the ways to get it are easier, these information can be used in many different application like professional, academic, and military, but is too easy to get information. This information is used to people who want to take advantage for their own benefit, many entities mainly important companies or military branch want to protect their information from other persons attacks, for that reason their thinks other forms to hide and transport the information, one of this is the steganography, this consists in hiding information inside of image, videos or in this case in audio, this project present a STUDY AND DEVELOPMENT OF AN APPLICATION OF STEGANOGRAPHY TO SEND DATA IN AUDIO FILES, ORIENTED TO THE SECURITY IN THE COMMUNICATION SYSTEMS, the project shows a study of state of art where shows different application and tools of steganography, also has a guidelines for future works. Using software MatLab, an application was created, this allows to choose an audio file with format ".wav" where the data will be hidden, and decode the hidden message created in the reception. Finally was applicate a survey MOS type to obtain an answer to know if the result was satisfactory.

KEYWORDS:

- **LSB (Least Significant Bit),**
- **MSE (Mean Squared Error),**
- **PSNR (Peak Signal-to-Noise ratio).**

CAPÍTULO 1

INTRODUCCIÓN

1.1. ANTECEDENTES

La esteganografía no es una técnica actual ya que esta ha sido utilizada hace mucho tiempo específicamente desde el año 470 A.C., hoy en día es una técnica muy utilizada en la seguridad para ocultar información delicada que solo se desea que una persona o un grupo selecto de personas vea. (Checa, 2014)

La esteganografía se la puede confundir con la criptografía, sin embargo la principal diferencia es que la criptografía codifica la información dejándolos ininteligibles, sin embargo la esteganografía oculta la información utilizando una portadora sin la necesidad de transmitirlo cifrado.

En la actualidad la esteganografía es un método muy utilizado para ocultar información, sin embargo en el Ecuador esta técnica apenas se utiliza debido a que no se toma en cuenta lo susceptible que es la información a ataques de personas maliciosas.

La esteganografía más utilizada es la que se aplica en imágenes debido a que en esta se utiliza un cuadro de imagen RGB para ocultar la información, sin embargo se pueden aplicar en otros archivos como lo es en audio y últimamente en voz IP siendo esta la más efectiva pero a la vez la más difícil

de conseguir debido a que no se puede prever durante cuánto tiempo se establecerá la conversación.

1.2. JUSTIFICACIÓN E IMPORTANCIA

Debido a los grandes avances en la tecnología que se tiene hoy en día los métodos de ataque son cada vez más frecuentes, y más efectivos, logrando de este modo obtener información confidencial, que puede ser usado de diferentes maneras, ya sea para comercializarse o chantajear a una persona en específico.

La esteganografía trata del estudio y aplicación de varias técnicas que permiten ocultar información de diferentes maneras, ya sea por palabras, imágenes, o en el caso de esta tesis por medio de audio.

La esteganografía en imágenes se basa en ocultar la información en cuadros de imagen RGB donde a cada cuadro se le asigna un dato que este llevara oculto por lo que la cantidad de información que esta puede contener depende del tamaño de la imagen, del mismo modo la esteganografía en audio se basa en ocultar información en portadoras para que de este modo el archivo no sufra ningún cambio, más sin embargo un cambio notorio que este nos puede dar es en el peso del archivo debido a que este al contener más información llegaría a tener un mayor peso.

La esteganografía en audio es una técnica no muy conocida y por ende no es aplicada a varios aspectos de seguridad sin embargo su utilización es en el ámbito estudiantil para adquirir más practica con este tipo de técnica, empresarial para enviar información a áreas específicas de la empresa y que solo un grupo selecto de personas puedan obtener la información; o una de las aplicaciones más importantes que se le puede dar a este tipo de técnica es el ámbito de seguridad militar ya que se necesita de una forma de enviar y

recibir información de manera secreta, la esteganografía es una excelente técnica, debido a que estos tienen en su poder códigos de alto riesgo, ya sea para confirmación de una orden, o el permiso de un ataque sea este a pequeña o gran escala.

1.3. ALCANCE DEL PROYECTO

El propósito general de este proyecto es el de dar seguridad a la información que se la considera importante ya sea en el ámbito privado o público, por lo que la esteganografía es un método muy eficaz para ocultar información.

En primer lugar se realizó una investigación sobre si en el país se ha utilizado este método de ocultar información en archivos de audio, por lo que se investigó diferentes tesis y paper`s publicados en diferentes instituciones encontrando solo sobre la esteganografía en imágenes por lo que se investigó sobre trabajos realizados en el exterior.

Luego de tener un mayor conocimiento sobre los trabajos realizados en el exterior se procedió a realizar el estudio de la esteganografía y sus diferentes métodos de implementación en varios tipos de archivos.

Una vez definido el método más adecuado se procedió a realizar un algoritmo para ocultar información en archivos de audio, para lo cual se necesitó de un entorno de desarrollo como lo es Matlab.

Una vez implementado el algoritmo se lo puso a prueba en diferentes entornos como:

- Se ocultó información en archivos de audio y se envió dicho archivo por correo electrónico y se descifro la información que esta lleva oculta.

- Se colocó el archivo de audio en un reproductor de música y luego se descifro para obtener la información oculta.

Al momento de finalizar con la investigación y desarrollo del algoritmo se concluyó cual fue el porcentaje de éxito de ocultar información y descifrarla, y si esta aplicación pudo ser ocupada por alguna entidad privada o pública.

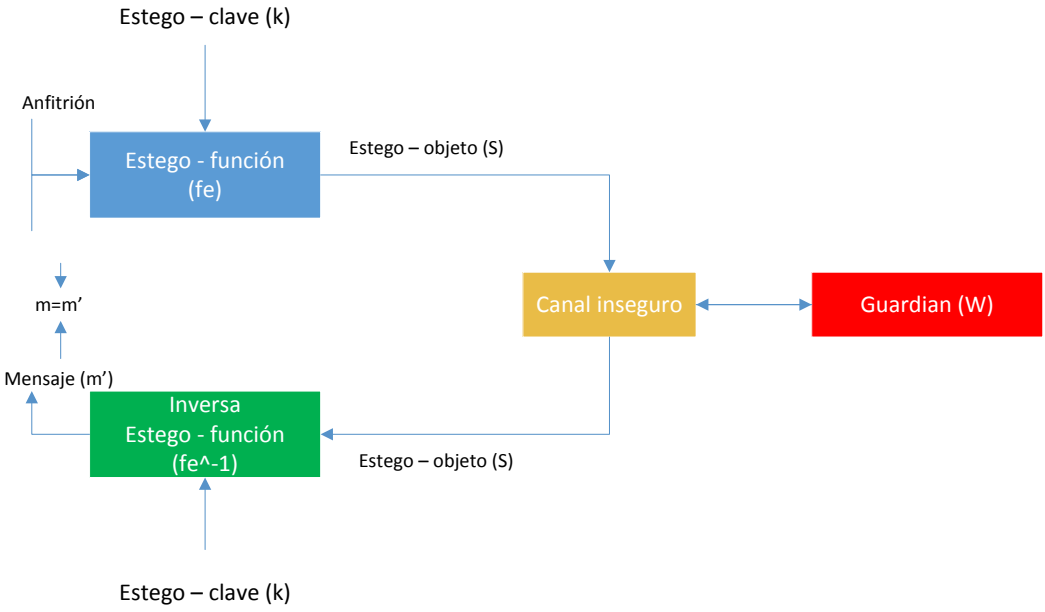


Figura 1 Esquema tipico de la esteganografía

1.4. OBJETIVOS

1.4.1. General

Estudiar y desarrollar una aplicación que permita enviar datos de manera oculta en archivos de audio, con la utilización de la técnica esteganográfica para seguridad en sistemas de comunicación.

1.4.2. Específicos

- Investigar y conocer los diferentes métodos esteganográficos actuales.
- Desarrollar una aplicación para él envió de información oculta con la utilización de Matlab como plataforma de desarrollo.
- Realizar pruebas en diferentes escenarios para verificar la calidad del audio y analizar en el dominio de la frecuencia.
- Documentar el trabajo realizado a lo largo de la investigación.

1.5. ESTRUCTURA DEL PROYECTO

El presente documento se divide en los siguientes capítulos:

Capítulo 1. *Introducción*. En este capítulo se puede tener un referente a los antecedentes, justificación e importancia que conlleva a plantear el trabajo de investigación.

Capítulo 2. *Seguridad de la información con la esteganografía*. En este capítulo se describe sobre la información histórica sobre la esteganografía de

que trata y como esta ha ido evolucionando al pasar de los años, del mismo modo se analizara el modelo esteganográfico utilizado para ocultar información e diferentes archivos.

Capítulo 3. *Esteganografía técnicas y aplicaciones*. En este capítulo se dará un estudio a las diferentes técnicas esteganográficas existentes y utilizadas en este trabajo de investigación, del mismo modo se dará a conocer las diferentes aplicaciones en las que se utiliza la técnica de la esteganografía.

Capítulo 4. *Implementación del programa.* En este capítulo se podrá tener una breve explicación del software utilizado “Matlab” para implementar la aplicación esteganográfica. Se tendrá una explicación del método utilizado para insertar datos en un archivo de audio y del mismo modo una explicación para recuperar la información, para finalizar se podrá observar el esquema grafico del programa tanto para ocultar como recuperar la información.

Capítulo 5. *Pruebas y análisis de resultados.* En este capítulo se realizaron las pruebas en diferentes escenarios con el audio con mensaje oculto y el audio original mediante pruebas MOS realizado a diferentes personas y se analizaron los resultados obtenidos en cada escenario propuesto.

Capítulo 6. *Conclusiones y Líneas de trabajos futuros*. En este capítulo se concluyó sobre los resultados obtenidos en el capítulo 5 y en el desarrollo de la aplicación creada, del mismo modo se dio una explicación de trabajos futuros que se puede realizar a partir del trabajo de investigación realizado.

Bibliografía. En este capítulo se tiene las fuentes de consulta que se han utilizado en el desarrollo del proyecto de investigación.

CAPÍTULO 2

SEGURIDAD DE LA INFORMACIÓN CON LA ESTEGANOGRAFIA.

2.1. INTRODUCCIÓN A LA INFORMACIÓN OCULTA

Para poder comprender de mejor manera de que se trata la esteganografía es necesario tener un breve conocimiento de algunos conceptos básicos:

- **Archivo encubridor:** Es el archivo el cual sirve como base en donde se ocultara la información o mensaje.

- **Mensaje:** Son los datos los cuales van hacer ocultos en un archivo.

- **Archivo con esteganografía:** Es el archivo el cual contiene la información oculta.

- **Atacante:** Entidad la cual desea obtener información de manera ilegal.

- **Seguridad:** Permite asegurar que los recursos del sistema se utilizan de manera en la que se espera. (Casierra, 2009)

La esteganografía viene de las palabras griegas “stegos” que significa cubierta y “graphos” que significa escritura, se la puede definir como el arte de ocultar un mensaje “huésped” en otro mensaje “anfitrión”. El principal

objetivo de la esteganografía es el de ocultar información en un archivo sin levantar sospecha de la información oculta, debido a las implicaciones que tiene la esteganografía en la seguridad de las comunicaciones y el crecimiento por la seguridad en la red la esteganografía es un tema muy estudiado. (Orbegozo, 2011)

Debido a que dependiendo de la cantidad de información que se desee ocultar en el archivo anfitrión, este podría verse afectado de manera considerable debido a que lo recomendado es que el mensaje anfitrión debe ser mayor al mensaje huésped, de esta manera el ocultar el mensaje sin alterar significativamente al mensaje anfitrión.

2.2. HISTORIA DE LA ESTEGANOGRAFÍA

La esteganografía no es una técnica reciente debido a que esta data del año 470 A.C. principalmente fue usado por los griegos. (Checa, 2014)

En la antigua Grecia el método esteganográfico más usado era el de tatuar mensajes en las cabezas rapadas de los soldados, sin embargo el mayor problema que tiene esto es que se debe esperar el cabello para poder ocultar el mensaje en sus cabezas.

En 1462 nació Johannes von Heidelberg uno de los primeros en escribir sobre los mensajes ocultos, llamando a su primera obra “La esteganografía” (forma parte de una trilogía) ver Figura 2 Portada del libro “La esteganografía”en donde describe como ocultar mensajes sin provocar ninguna sospecha, bajo esta escritura se declaró que el uso de esta técnica era mágica y su obra fue prohibida debido a que argumentaban que utilizan a los espíritus para enviar mensajes secretos. Años después de su muerte se argumentaba que la técnica de Johannes von Heidelberg tenía una debilidad y era que tanto el emisor como el receptor deben conocer del método

esteganográfico utilizado para poder descifrar el mensaje, sin embargo esta debilidad también es considerada su mayor fortaleza ya que solo las personas que conocen la técnica pueden conocer la información oculta. (Casierra, 2009)

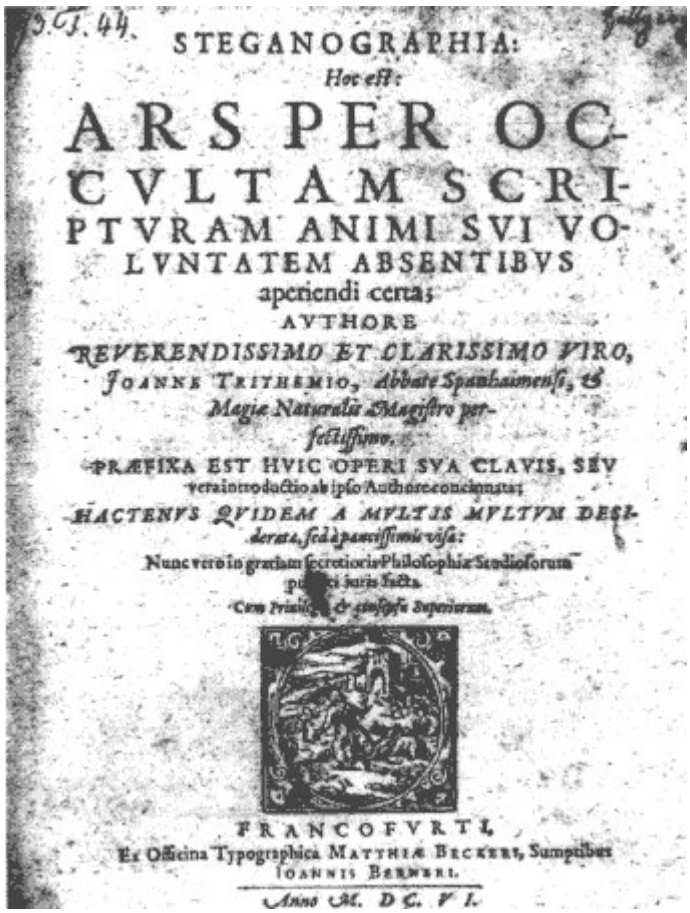

STEGANOGRAPHIA:
Hoc eſt:
ARS PER OC-
CVLTAM SCRI-
PTVRAM ANIMI SVI VO-
LVNTATEM ABSENTIBVS
aperiendi certa;
AVTHORE
REVERENDISSIMO ET CLARISSIMO VIRO,
IOANNE TRITHEMIO, Abbate Spanhaimenſi, &
Magiæ Naturalis Magiſtro per-
fectiſſimo.
PRÆFIXA EST HVIC OPERI SVA CLAVIS, SEV
vera introductio ab ipſo Authore concinnata;
HACTENVS QVIDEM A MVLTIS MVLTVM DESI-
derata, ſed à pauciſſimis viſa:
Nunc vero in gratiam ſecretioris Philoſophiæ Studioſorum
publici juris facta.
Cum Privilegio & conſenſu Superiorum.
FRANCOFVRTI,
Ex Officina Typographica MATTHIÆ BECKERI, Sumptibus
IOANNIS BERNERI.
Anno M. DC. VI.

Figura 2 Portada del libro “La esteganografía”

Fuente: (Garnacho, Juan M, & Castro)

Gaspar Schott científico alemán, escribió el libro Schola Steganographica (1665) Figura 3 Portada del Schola Steganographica, siendo este uno de los más importantes en lo que respecta a la esteganografía, debido a que se aleja de lo esotérico y lo mágico para enfocar la esteganografía desde el punto de vista de la técnica y la ciencia.

Figura 3 Portada del Schola Steganographica

Fuente: (Gámez, 2008)

En el campo de las telecomunicaciones cualquier persona que tenga acceso a información puede robar y manipular el mensaje cifrado con relativa facilidad, pero aun si la persona conoce de la existencia del mensaje oculto las probabilidades de que lo obtenga son muy bajos, debido a que no conoce el método que se utilizó para ocultar información y al momento de analizar el archivo no tiene certeza de que este contenga algún mensaje oculto.

2.3. MODELO DE LA ESTEGANOGRAFÍA

Para poder dar un modelo de comunicación entre el emisor y el receptor, las aplicaciones esteganográficas siguen el modelo descrito en la Figura 4 Ejemplo de modelo esteganográfico en donde el emisor escoge un objeto el cual servirá como anfitrión, el cual puede ser transmitido por diferentes sistemas de comunicación sin levantar sospecha. (Checa, 2014)

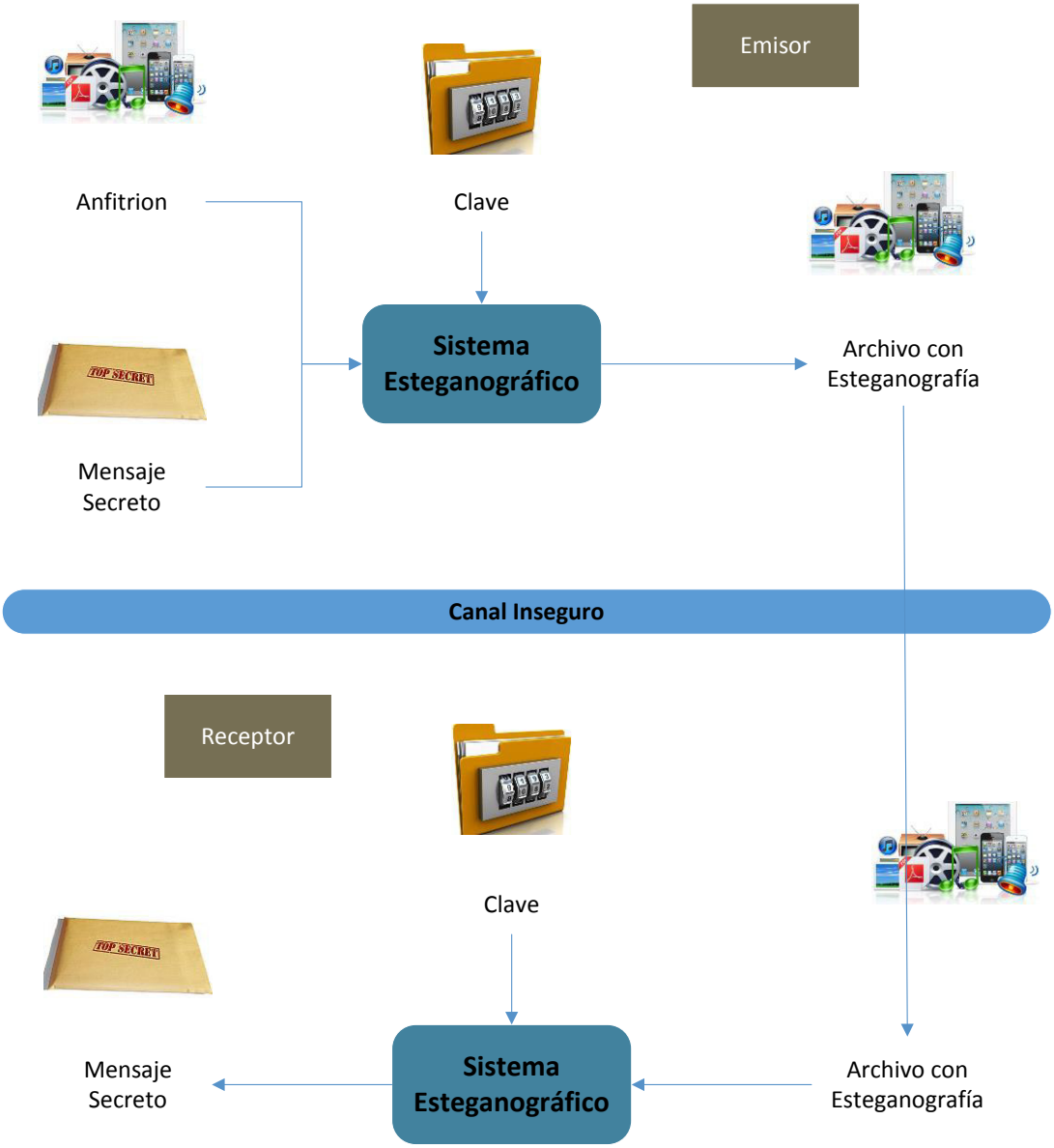


Figura 4 Ejemplo de modelo esteganográfico

Al momento de seleccionar un anfitrión, se crea un mensaje secreto el cual será colocado dentro del anfitrión mediante el uso de una técnica esteganográfica, del mismo modo se coloca una clave para dar más seguridad al mensaje secreto y de este modo crear el archivo esteganográfico el cual contiene tanto al anfitrión como al mensaje secreto, el cual se enviara por algún sistema de comunicación.

Luego del que el archivo esteganográfico ha sido enviado y recibido por el receptor, este puede recuperar siempre y cuando conozca el método esteganográfico y la clave utilizada para ocultar información.

Una de las características más importantes sobre la esteganografía para dar mayor seguridad, es jamás utilizar el mismo anfitrión, ni el mismo sistema de comunicación por segunda vez, debido a que esto levanta sospechas y el archivo esteganográfico podría sufrir algún tipo de ataque.

La clave que se coloca no siempre es necesaria, debido que la técnica esteganográfica utilizada de por sí ya es considerada segura, sin embrago para dar más seguridad al mensaje secreto se lo recomienda hacerlo.

2.3.1. Características de un sistema esteganográfico.

1. **Capacidad:** Es el número de bits que se puede esconder dentro del mensaje anfitrión.

2. **Robustez:** Es la capacidad del sistema para poder someterse a diferentes escenarios ya sean adición de ruido, escalado, transformaciones, etc. sin perder el mensaje secreto escondido en el anfitrión.

3. **Invisibilidad:** Es la capacidad de que el mensaje secreto pase sin ser detectado en diferentes tipos de análisis.

4. **Seguridad:** Es la capacidad de que el mensaje secreto se enfrente a diferentes ataques y que este se encuentre libre de peligros y de riesgos. (Orbegozo, 2011)

2.4. ESTUDIO DEL ESTADO DEL ARTE.

Con el gran avance de la tecnología, la continua evolución y expansión de la información, está es más propensa a ataques, por lo que cada día se crean diferentes herramientas y métodos de protección de la misma.

Hoy en día una herramienta muy útil es la esteganografía debido a que esta al tener diferentes técnicas se la puede utilizar de diferentes maneras y ocultar la información dentro de un anfitrión, comúnmente la esteganografía se la confunde con la criptografía, pero estas tienen un gran diferencia debido a que la criptografía cifra los datos y los oculta dentro de un anfitrión, la esteganografía oculta los datos dentro del mismo anfitrión de este modo el mensaje secreto se vuelve parte del anfitrión.

En el siglo V A.C. el historiador griego Heródoto con el fin de comenzar una revolución en contra de los persas, rapo a uno de sus esclavos y le tatuó un mensaje esperó a que le volviera a crecer el cabello y de este modo oculto el mensaje. Cuando el esclavo llegó a donde se encontraba Aristágoras de Mileto, este le rapo la cabeza y pudo ver el mensaje que le habían mandado. (Vico, 2010)

El ejemplo esteganográfico más usado es el de los prisioneros que fue una idea propuesta por Gustavus J. Simmons. En este artículo, Simmons analiza cómo se pueden comunicar dos prisioneros que se encuentran en dos celdas diferentes pero se les permite comunicar mediante mensajes supervisados y estos idean una forman de enviar mensajes ocultos sin levantar sospechas. Los dos prisioneros llamados Alice y Bob, se envían mensajes entre sí, pero a estos les añaden un código de redundancia para verificar que el guardia Willie, no ha modificado los mensajes con los cuales se comunican. Un día Willie se da cuenta de que Alice y Bob se comunican entre ellos mediante los mensajes, sin embargo el no conoce el método por lo que al tratar de alterar los mensajes Alice y Bob se dan cuentan y deciden cambiar los bits con los

cuales elaboraban un plan de fuga. De este modo Simmons dio a conocer a la esteganografía como un método de ocultar mensajes dentro de un anfitrión. (Miriam, 2012)

En el siglo XV el científico italiano Giovanni Battista Della Porta descubrió como poner un mensaje dentro de un huevo cocinado. El método era sencillo solo se preparaba una tinta que contenía vinagre y escribía en la superficie del huevo al haber sido cocido el huevo este absorbía la tinta y el mensaje se ocultaba dentro de la superficie cocida del huevo y la única forma de leerlo era rompiendo la cascara.

En 1857, Brewster sugirió la posibilidad de ocultar mensajes secretos mediante reducción fotográfica en un espacio no mayor que un punto de tinta. Una de las historias documentadas de la utilización de este método es en la guerra Franco_Prusiana de 1870-1871. (Miriam, 2012)

En la segunda guerra mundial se utilizó una técnica muy ingeniosa para la época, que consistía en usar microfilmes en los puntos de las “ies” o en signos de puntuación para de este modo poder enviar mensajes. Los prisioneros usaban las “i, j, t y f” para ocultar mensajes en “código morse”, el ultimo sistema para enviar mensaje era considerado el más ingenioso este fue llamado “Null Cipher” el cual consistía en enviar un mensaje oculto dentro de un mensaje común, el secreto yacía en que en un mensaje común se leía solo cierta posición de la palabra y de este modo se lograba descifrar el mensaje oculto.

En la tesis “Esteganografia em audio e imagen utilizando a técnica LSB” hace referencia a cómo utilizar la esteganografía en audio e imagen con la técnica LSB, sin embargo solo la implementa en imagen debido a que es la más común y en audio solo realiza una breve explicación mas no la implementa debido a que no es muy utilizada. (Cantanhede, 2009)

Sin embargo en la tesis “Implementación de un sistema esteganográfico para inserción de textos en señales de audio” realiza la implementación del programa para archivos de audio y nos da información sobre el oído humano y como este no puede detectar si un archivo sufrió alguna alteración. (Casierra, 2009)

En resumen, la esteganografía históricamente ha tenido una utilización muy amplia debido a que sido utilizada en el área de conflictos militares, espionaje, la política y hoy en día por personas particulares para poder proteger información que se considera delicada, esto cada vez se está volviendo un habito común debido a que con el avance de la tecnología los ataques a la información es cada vez más frecuente.

2.5. FORMATOS DE AUDIO DIGITAL

Hoy en día existen una gran variedad de formatos digitales, esto surgió principalmente debido a las diferentes plataformas de reproducción existentes, a continuación se va a detallar algunos de los formatos más utilizados.

Tabla 1

Características de los formatos de audio más utilizados

Formato	Desarrolladores	Extensión	Aplicación
WAV	Microsoft e IBM	.wav	Windows lo utiliza para su propio sistema.
WMA	Microsoft	.wma	Compresión de audio.
MP3	MPGE	.mp3	Intercambio de archivos musicales por Internet.
AAC	AT&T, Sony	.aac	Utilizado en iTunes.

2.5.1. Formato WMA

El formato WMA (Windows Media Audio) que fue desarrollado por Microsoft. Una de las características más importantes de este tipo de formato es que los archivos que se encuentran bajo la extensión .wma tienden a ser mucho más pequeños en espacio de disco que incluso los mp3.

La desventaja que tiene este tipo de formato es la pérdida de calidad de audio, debido a que al comprimir el archivo de audio la calidad de sonido se ve afectada.

Tabla 2

Caracterisiticas del formato WMA

Desarrollador	Microsoft
Extensión	.wma
Canales	Estéreo, mono
Frecuencia de muestreo	48000 Hz, 44100 Hz

2.5.2. Formato MP3

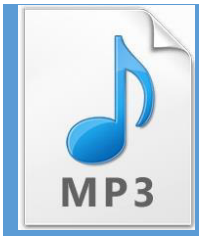
El formato MP3 es uno de los más populares debido a que los archivos bajo la extensión .mp3 ocupan un espacio en disco duro muy pequeño, se toma por ejemplo que si en un CD se puede almacenar 10 archivos con extensión .wav, los archivos con el formato .mp3 pueden almacenar 100 archivos en el mismo cd.

El formato MP3 no nació en un laboratorio o empresa grande, el formato MP3 nació en Internet debido a que varios usuario se unieron y formaron el grupo “Moving Picture Expert Group”, con el fin de poder intercambiar archivos de música por Internet y de este modo nació el formato MP3.

La razón por la cual este tipo de formato es tan utilizado es debido a que elimina el rango de frecuencia que el oído humano no escucha, específicamente los que se encuentran mayores a 20 KHz y menores a 20 Hz, por lo cual el archivo final creado con extensión .mp3 tiende a ser de menor tamaño en disco duro.

Tabla 3

Caracteristicas del formato MP3

	**Desarrollador**	**Moving Picture Expert Grouo (MPEG)**
	Extensión	.mp3
	Canales	Estéreo, mono
	Frecuencia de muestreo	48000 Hz, 44100 Hz

2.5.3. Formato AAC

Este tipo de formato es utilizado por Apple mediante su software iTunes, el cual tiene casi las mismas características que ".mp3", es decir, se basa en la eliminación de rango de frecuencias que el oído humano no escucha, sin embargo con este tipo formato se obtiene una mejor calidad de audio y un menor espacio de disco duro ocupado.

AAC utiliza variables de frecuencias de bits, los cuales adaptan los bits utilizados para codificar los datos del audio, y además utiliza sonidos polifónicos los cuales crean un sonido envolvente mejor conocido como "Surround".

Tabla 4

Caracteristicas del formato AAC

	**Desarrollador**	**AT&T, Sony**
	Extensión	.aac, .mp4
	Canales	Estéreo, mono
	Frecuencia de muestreo	24000 Hz, 22500 Hz

2.6. Archivo o formato WAV

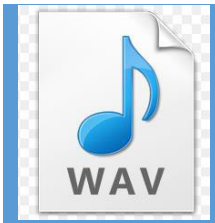
El formato WAV (Waveform Audio File) fue desarrollado por Microsoft e IBM, es uno de los formatos de audio digital actuales con excelente calidad debido a que este no posee ningún tipo de compresión de datos, este tipo de archivo es utilizado en los sistemas operativos como Windows para los sonidos de su propio sistema, estos se escuchan al momento de prender y apagar las computadoras.

El formato WAV es uno de los formatos más utilizados profesionalmente, debido a que este al poseer la particularidad de soportar diferentes tipos de codec's actuales específicamente el formato PCM el cual no tiene ningún tipo de compresión, con esto nos genera una calidad de audio excelente debido a que el sonido se captura a 44100 Hz y a 16 Bits.

La mayor desventaja que posee el formato WAV, es el espacio en disco que ocupa, generalmente es de 10 Mb por cada minuto grabado, lo que con lleva que este tipo de formato no sea utilizado en Internet, normalmente se utiliza el formato MP3.

Tabla 5

Caracteristicas del formato WAV



Desarrollador	Microsoft, IBM
Extensión	.wav
Canales	Estéreo, mono
Frecuencia de muestreo	44100 Hz, 22050 Hz

Este tipo de formato es el más utilizado por archivos de audio, debido a que no contiene ningún tipo de compresión de datos su calidad de audio es

muy alta, los desarrolladores fueron Microsoft e IBM, su frecuencia de muestro va desde los 11000 Hz a 48000 Hz, otra de la características además de su excelente calidad de audio, es su compatibilidad con todos los reproductores de audio multimedia.

El archivo WAV en si consta de 2 partes, al ser formatos RIFF la primera parte representa el formato de los datos además de la muestra y en su segunda parte se encuentran los datos del sonido.

Tabla 6

Cabecera WAV

Nombre del Campo	Tamaño del Campo (Bytes)	Descripción
ChunkID	4	Formato wav compuesto de 2 partes: Formato y datos
ChunkSize	4	
Format	4	
Subchunk1 ID	4	Contiene la descripción del sonido en los datos.
Subchunk1 Size	4	
AudioFormat	2	
NumChannels	2	
SampleRate	4	
ByteRate	4	
BlockAlign	2	
BitsPerSample	2	
Subchunk2 ID	4	Contiene el tamaño y los datos del sonido.
Subchunk2 Size	4	
Data	Subchunk2Size	

En la Tabla 7 Detalle de la cabecera WAV se da a más detalle una descripción de la cabecera WAV.

Tabla 7

Detalle de la cabecera WAV

Nombre	Descripción
ChunkID	Contiene RIFF en ACII (52 49 46 46)
ChunkSize	36 + SubChunk2Size
Format	Contiene WAVE en ASCII (57 41 56 45)
Formato WAVE está formado por "formato" y "Datos"	
Subchunk1 ID	Contiene fmt en ASCII (66 6d 74 20)
Subchunk1 Size	16 para PCM (modula la señal)
AudioFormat	PCM = 1 (identificador de PCM)
NumChannels	Mono =1, Stereo = 2
SampleRate	8000, 44100 , etc.
ByteRate	SampleRate * NUmChannel * BitsPerSample / 8
BlockAlign	NumChannel * BitsPerSample / 8
BitsPerSample	8 bits = 8, 16 Bits = 16, etc.
ExtraParamSize	Si está en PCM, este campo no existe
ExtraParams	Para parámetros extras
"Datos" contiene el tamaño de los datos y el sonido	
Subchunk2 ID	Contiene DATA en ASCII (64 61 74 61)
Subchunk2 Size	NumSample * NumChannels * BitsPerSample / 8
Data	Datos del sonido

A continuación se dará un ejemplo con una secuencia de bits, para entender de mejor manera cómo funciona el formato.

52 49 46 46 24 08 00 00 57 41 56 45 66 6d 74 20 10 00 00 00 01 00 02 00 22 56 00 00 88 58 01 00 04 00 10 00 64 61 74 61 00 08 00 00 00 00 00 00 24 17 1e f3 3c 13 3c 14 16 f9 18 f9 34 e7 23 a6 3c f2 24 f2 11 ce 1a 0d

Tabla 8

Informacion de bits

Secuencia	Descripción
Chunk Descriptor	
52 49 46 46	ChunkID = RIFF
24 08 00 00	ChunkSize = 2084
57 41 56 45	WAVE
FMT SubChunk	
66 6d 74 20	FMTF
10 00 00 00	SunChunkSize = 16
01 00	AudioFormat = 1
02 00	NumChannels = 2
22 56 00 00	SampleRate = 22050
88 58 01 00	ByteRate = 88200
04 00	BlockAlign = 4
10 00	BitsPerSample = 16
Data Subchunk	
64 61 74 61	DATA
00 08 00 00	SubChunk2Size = 2048
00 00	Muestra 1, muestra del canal izquierdo
00 00	Muestra 1, muestra del canal derecho
24 17	Muestra 2, muestra del canal izquierdo
1e f3	Muestra 2, muestra del canal derecho
3c 13	Muestra 3, muestra del canal izquierdo
3c 14	Muestra 3, muestra del canal derecho

CAPÍTULO 3

ESTEGANOGRAFÍA TÉCNICAS Y APLICACIONES

3.1. TERMINOLOGÍA

Para tener un mejor entendimiento y evitar confusiones, es necesarios tener un breve conocimiento de la terminología básica utilizada en las técnicas de ocultar información.

- El *archivo portador* es aquel el cual va a servir como base para ocultar la información, el archivo portador puede ser de un archivo de audio, imagen video o texto, para el caso específico de esta tesis será un archivo de audio con formato ".wav".

- Se conoce como información oculta a la información que se enviará de forma secreta en el archivo portador.

- Al momento de introducir la información oculta en el archivo portador se tiene como resultado un *archivo esteganográfico* el cual comúnmente es conocido como *estego-audio*, para esta tesis en concreto al archivo creado se lo nombrará como el usuario desee para evitar levantar sospecha sobre el archivo y la información que este tiene. (Checa, 2014)

3.2. TÉCNICAS ESTEGANOGRÁFICAS

Debido a las diferentes técnicas en las que se puede utilizar la esteganografía, la utilización de estas depende del tipo de archivo en la que se trabaje; ya sea en imágenes, audio, video o texto. Las más utilizadas son las de LSB, algoritmos y transformadas de separación del espectro. (Cantanhede, 2009)

3.2.1. Técnica de inserción del bit menos significativo

La técnica LSB consiste en sustituir el bit menos significativo de la codificación de un pixel de una imagen o muestra de un audio por la información que se desea ocultar. De este modo se puede modificar el bit menos significativo de un byte del archivo de audio o imagen, sin alterar al anfitrión en ningún aspecto notable.

Para tener más claro cómo funciona la técnica de LSB se dará un ejemplo de cómo se sustituiría el bit del mensaje a ocultar por el que tiene el archivo de audio o imagen.

El anfitrión se encuentra de color amarillo y los bit`s que cambiarían se encuentran señalados de color morado.

10110101 – 11010010 – 11001010 – 11001010 – 11011101

El texto anterior nos señala con negrillas y de un color especifico los bit´s que van a ser cambiados por los bits 11010, y de este modo se oculta el mensaje dentro de un anfitrión. (Cantanhede, 2009)

Mensaje oculto:

10110101 – 11010011 – 11001010 – 11001011 – 11011100

Como se pudo apreciar anteriormente al cambiar los bits del anfitrión por los del mensaje a ocultar de los bits menos significativos, cambiaron 3 de los 5 en total por lo que el cambio no es total si no parcial, de este modo se puede ocultar el mensaje sin alterar completamente al anfitrión.

LSB en audio

El LSB en audio es muy parecido al aplicado en imágenes, con la gran diferencia que en audio se debe cambiar los bits menos significativos del audio directamente más no los canales RGB como en imágenes.

El sistema auditivo humano puede escuchar diversos tipos de frecuencias entre 20 Hz a los 20 kHz Figura 5 Rango de frecuencias, por lo que los seres humanos tiene la capacidad de distinguir una gran cantidad de intervalos de sonido, sin embargo no los escucha todos ni puede percibir cambios mínimos realizados.

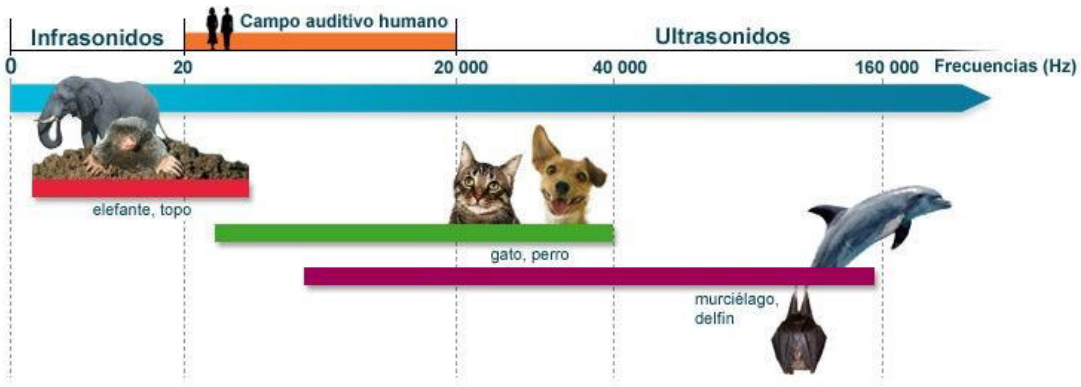


Figura 5 Rango de frecuencias

Fuente: (Chacon, s.f.)

Para poder ocultar el mensaje secreto en el anfitrión es necesario que el mensaje y el archivo de audio sean convertidos en binario para de este modo

se pueda intercambiar los bits menos significativos del audio por los bits del mensaje secreto, sin embargo se debe tomar en cuenta que los archivos .wav tienen especificaciones en su cabecera por lo que es necesario identificarla y no alterarla en ningún momento, debido a que este es la información más importante de los archivos de audio.

3.2.2. Técnicas basadas en algoritmos y transformadas

Las técnicas esteganográficas basadas en algoritmos y transformadas se basan en la transformada de Fourier, transformada directa de cosenos (estás técnicas son utilizadas principalmente en imágenes), la transformada de wavelet (utilizada principalmente en archivos de audio específicamente en los archivos .wav), dependiendo de qué o en donde se aplique las técnicas pueden ayudar a mejor la esteganografía en los archivos portadores. (Cantanhede, 2009)

Lo que se debe tomar en cuenta respecto a estas técnicas es que se debe utilizar de manera adecuada, debido a que cada una de ellas puede ayudar a ocultar de mejor manera el mensaje secreto sin ser detectado, si se lo utiliza de manera incorrecta los archivos específicamente de audio pueden verse comprometido en su calidad de sonido.

3.2.3. Técnicas para esteganografía en videos

Para utilizar la esteganografía en audio una de las técnicas más utilizadas es la del método de transformada discreta del coseno (DCT). Esta técnica es muy utilizada en la compresión de imágenes y videos. (Cantanhede, 2009)

Las técnicas para videos es muy parecida a la de las imágenes debido a que los video en si son imágenes continuas y sonido al mismo tiempo, por lo que la cantidad de información que se puede ocultar es mayor, sin embargo no se debe exagerar con la cantidad de información a ocultar por lo que en mayor cantidad oculta, más fácil será detectar algún cambio en el video.

3.3. APLICACIONES DE LA ESTEGANOGRAFÍA

A continuación se va explicar las aplicaciones que tiene la esteganografía en diferentes ámbitos que actualmente se utilizan.

3.3.1. Aplicaciones militares

En el ámbito militar es en donde se da mayor prioridad a la seguridad en la información debido a que el contenido que puede tener se considera en mayor parte confidencial, por lo que su seguridad es muy valiosa. En donde se ha visto mayormente utilizado es en las guerras debido a que al momento de intercambiar información existe una gran posibilidad de ataques y que el contenido sea filtrado.

3.3.2. Derechos de autor

Esta aplicación se usa específicamente para garantizar los derechos de autor del trabajo que un autor realizó. Para esta aplicación el autor debe dar su autorización para poder ocultar su información dentro del trabajo que realizo, y de este modo poder garantizar que su trabajo no se va haber envuelto en algún tipo de plagio, y que reciba el reconocimiento que merece por el trabajo que realizo.

3.3.3. Aplicaciones médicas

Este tipo de aplicación esteganográfica se podría decir que no es muy utilizada o casi nula, debido a que normalmente cuando un médico da su diagnóstico sobre lo que tiene su paciente este normalmente tiene una imagen como una radiografía y en un papel aparte la información de su paciente, debido a que el tema de la salud es muy importante se debería colocar la información del paciente dentro de los exámenes médicos que este se realizó, de este modo solo el médico y el paciente sabrían que problema tiene y a quien pertenece el examen echo, de este modo se aseguraría la información de ataques de personas externas que desean dañar al paciente.

3.3.4. Control de acceso

Este tipo de aplicación se la utiliza específicamente para poder evitar la copia o falsificación de algún documento que permita el acceso a alguna entidad pública o privada, normalmente para dar acceso a algún lugar específico hoy en día las tarjetas magnéticas son muy utilizadas y se creía que eran muy seguras, sin embargo con el pasar del tiempo estas se volvieron fácilmente falsificadas debido a que no tiene ninguna protección, al momento de utilizar la esteganografía en las tarjetas magnéticas se puede asegurar que aunque alguien logra copiar o falsificar una tarjeta, no podrá tener la información que se encuentra dentro de ella y de este modo no podrá tener acceso a algún lugar específico. (Cantanhede, 2009)

3.4. TIPOS DE ATAQUES

Debido a que la esteganografía es un método de ocultar información siempre está expuesta a ataques, por lo que existen diferentes métodos para

obtener el contenido que lleva un archivo, sin embargo la mayoría de los ataques que se realizan sirven como pauta para saber cuál es el fallo de la técnica utilizada y poder corregirla; los ataques más comunes utilizados hoy en día son los ataques estructurales y los ataques estadísticos los cuales van hacer explicados a continuación.

3.4.1. Ataques estructurales

Al momento que se ingresa información en algún archivo tiende a sufrir algún tipo de cambio, por más mínimo que este sea si alguien llega a notarlo el mensaje vendría a estar en un inminente ataque.

El método más común utilizado para este tipo de ataque se trata el de compactar el archivo varias veces, para de este modo eliminar todos los bits semejantes o redundantes que el archivo contenga; al momento de realizar esto el mensaje oculto llegaría a sufrir algún tipo de cambio o incluso llegar a ser inentendible por el ataque que sufrió.

3.4.2. Ataques estadísticos

Este tipo de ataque se basa en utilizar las estadísticas en varios archivos y de este modo poder estimar en cual archivo se encuentra algún tipo de mensaje oculto, normalmente se lo realiza tomando varios archivos que en si son iguales pero uno de ellos tiene una información oculta, al momento de utilizar las estadísticas analizan el contenido del archivo y de este modo pueden llegar a localizar el archivo con mensaje oculto.

CAPÍTULO 4

IMPLEMENTACIÓN DEL PROGRAMA

4.1. IMPLEMENTACIÓN DEL PROGRAMA

Para la implementación de este proyecto se utilizó el software “MATLAB R2010a” donde se elaboraron dos pantallas: la pantalla “GUIpresen” es la presentación del proyecto (ver Figura 6 Presentación de la portada), donde se detalla de que se trata el tema, muestra una imagen de fondo de la institución, contiene el nombre del autor, del director de tesis y del año realizado; tiene un botón “Continuar” el cual nos permite ir a la siguiente presentación de pantalla y poder observar la aplicación realizada, también tiene el botón “Salir” el cual permite cerrar la ventana principal; la segunda pantalla “GUI” (ver Figura 7 Presentación de la aplicación) se tiene la aplicación realizada en donde se puede seleccionar un audio, ocultar o extraer algún mensaje; y observar la señal del audio y si esta ha sufrido algún cambio.

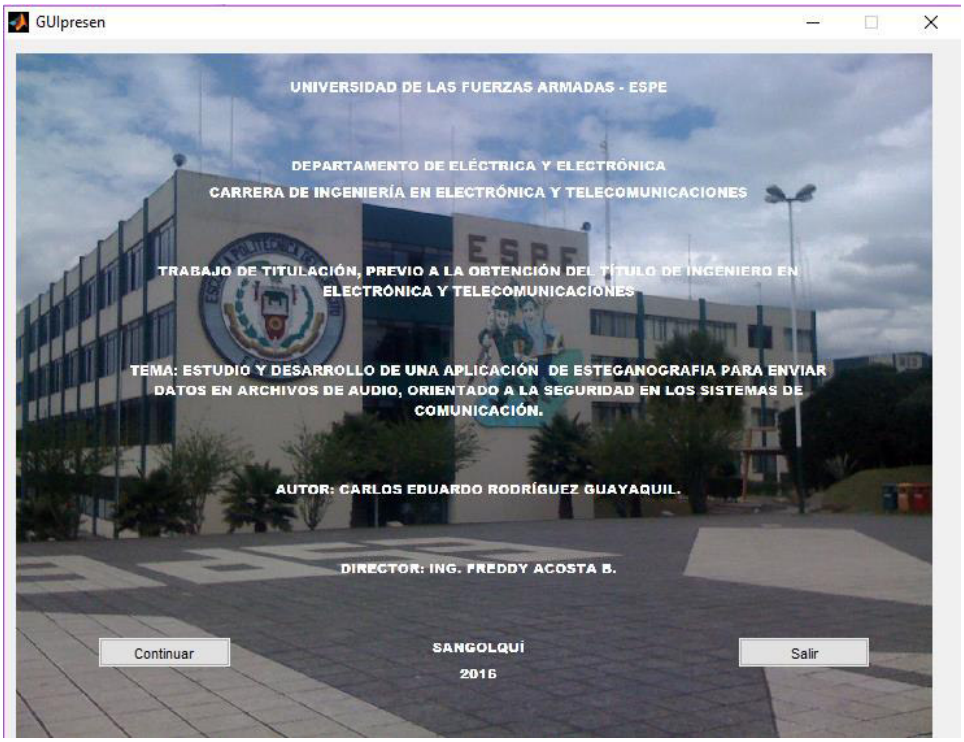


Figura 6 Presentación de la portada

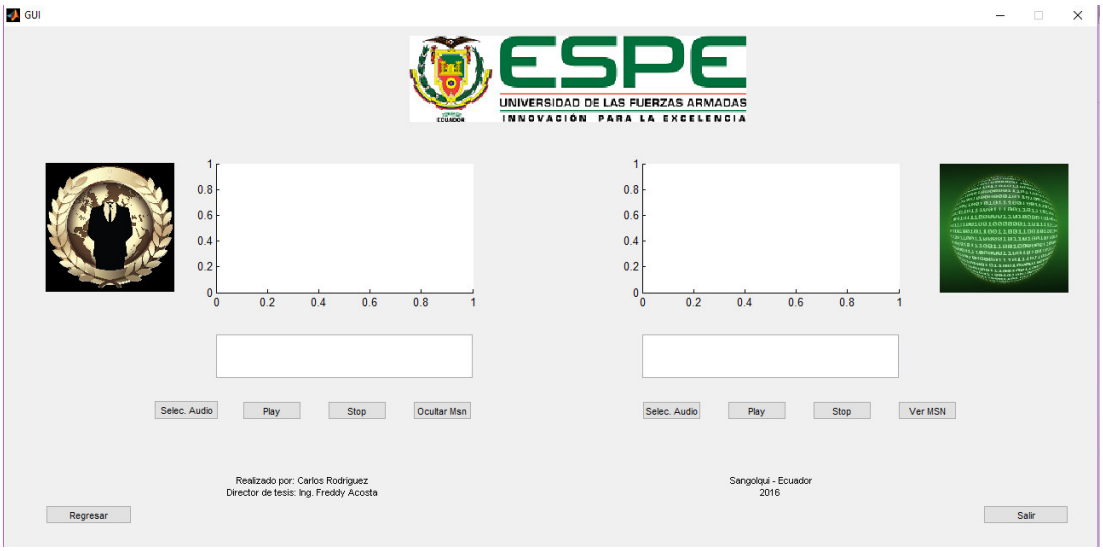


Figura 7 Presentación de la aplicación

4.1.1. PROGRAMA DE INSERCIÓN DE DATOS

Para la inserción de datos en un archivo de audio se establece el siguiente diagrama de flujo (ver Figura 8 Diagrama de flujo de inserción de datos), el cual permite comprender de mejor manera su funcionamiento:

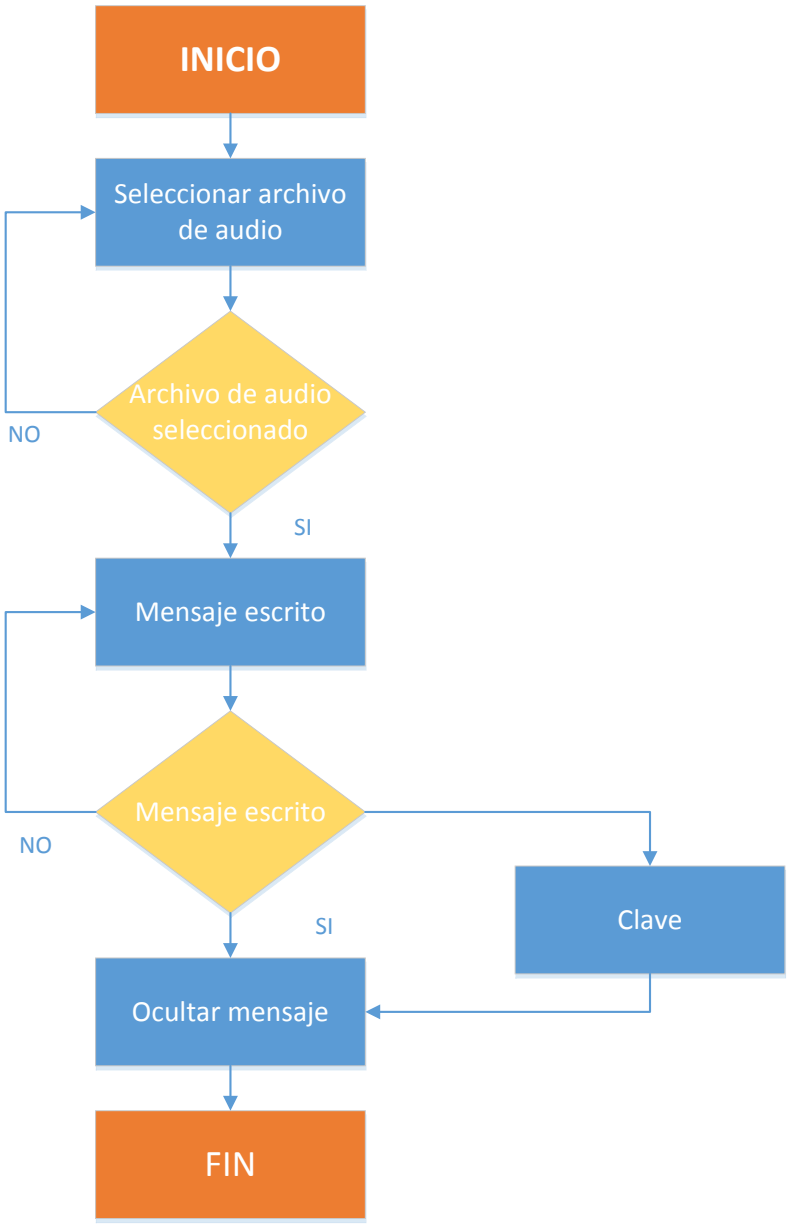


Figura 8 Diagrama de flujo de inserción de datos

Una vez entendido el diagrama de flujo anterior se procede a implantar el programa y de este modo se tiene la siguiente interfaz gráfica en la Figura 9 Interfaz gráfica de ocultar mensaje en audio:

Figura 9 Interfaz gráfica de ocultar mensaje en audio

Para poder ocultar un mensaje en un archivo de audio en la aplicación realizada, es necesario seguir los siguientes pasos:

1. Seleccionar el audio donde se desea ocultar el mensaje (la extensión del audio debe ser .wav). Figura 10 Selección del audio en el programa.

Figura 10 Selección del audio en el programa

2. Una vez seleccionado el audio, se procede a dar “play”, y se podrá escuchar el audio y observar la señal que este posee. Figura 11 Reproducción del audio seleccionado

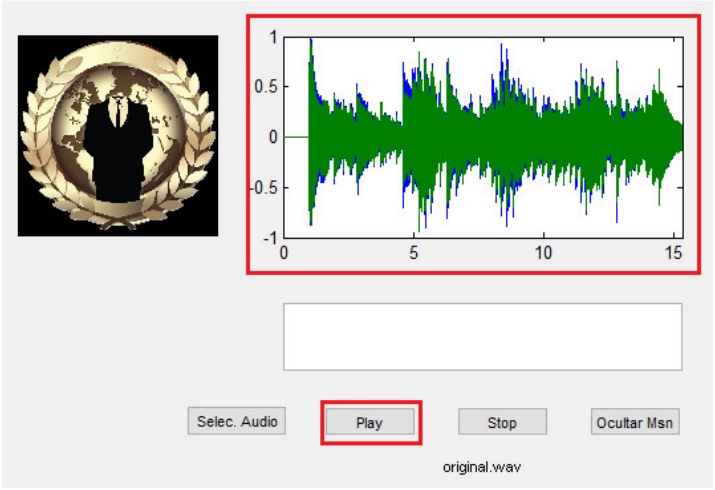


Figura 11 Reproducción del audio seleccionado

3. Una vez seleccionado el audio, se procede a escribir el mensaje y a ocultarlo en el archivo de audio. Figura 12 Mensaje a ser oculto

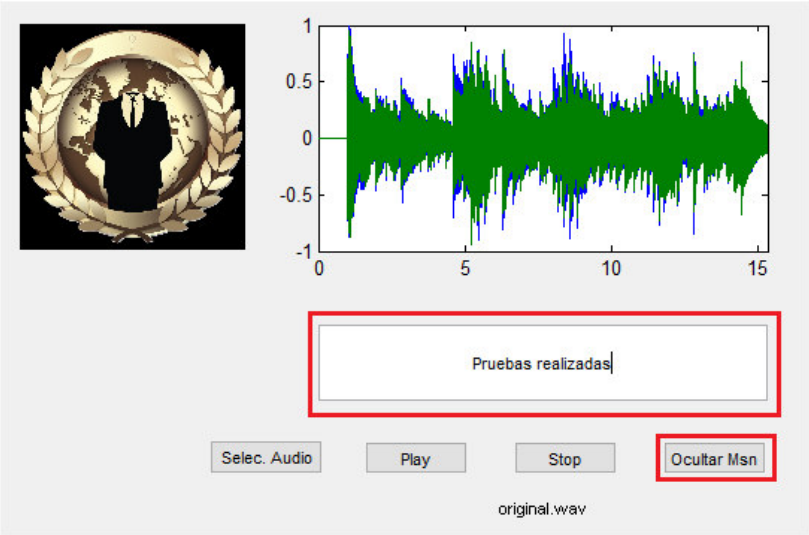


Figura 12 Mensaje a ser oculto

Para poder ocultar el mensaje dentro del archivo de audio se realiza los siguientes pasos dentro del programa creado:

Para colocar el mensaje “Pruebas realizadas” comenzamos tomando cada letra y espacio y tomamos su valor decimal:

Figura 13 Caracteres en decimal `msg_double=double(msg);`

```
msg_double =

  Columns 1 through 12

    80   114   117   101    98    97   115    32   114   101    97   108

  Columns 13 through 18

   105   122    97   100    97   115
```

Figura 13 Caracteres en decimal

Los valores numéricos tomados anteriormente los convertimos a binario y los colocamos en matriz:

Figura 14 Matriz de caracteres `msg_bin=de2bi(msg_double,8);`

```
msg_bin =

     0     0     0     0     1     0     1     0
     0     1     0     0     1     1     1     0
     1     0     1     0     1     1     1     0
     1     0     1     0     0     1     1     0
     0     1     0     0     0     1     1     0
     1     0     0     0     0     1     1     0
     1     1     0     0     1     1     1     0
     0     0     0     0     0     1     0     0
     0     1     0     0     1     1     1     0
     1     0     1     0     0     1     1     0
     1     0     0     0     0     1     1     0
     0     0     1     1     0     1     1     0
     1     0     0     1     0     1     1     0
     0     1     0     1     1     1     1     0
     1     0     0     0     0     1     1     0
     0     0     1     0     0     1     1     0
     1     0     0     0     0     1     1     0
     1     1     0     0     1     1     1     0
```

Figura 14 Matriz de caracteres

Al momento de crear la matriz el siguiente paso es saber cuál es el tamaño, debido a que esto nos ayudara a saber cuanta información va a ser cambiada.

Figura 15 Tamaño de la matriz `[m,n]=size(msg_bin);`

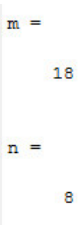


Figura 15 Tamaño de la matriz

Finalmente colocamos la matriz en una columna continua, la cual va hacer colocado en los bits menos significativos del archivo de audio original.

Figura 16 Bits colocados en columna `msg_bin_re=reshape(msg_bin,m*n,1);`

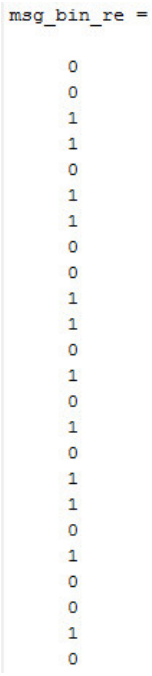


Figura 16 Bits colocados en columna

4. El paso final es nombrar al nuevo archivo de audio con el mensaje oculto. Figura 17 Nombre al nuevo archivo de audio

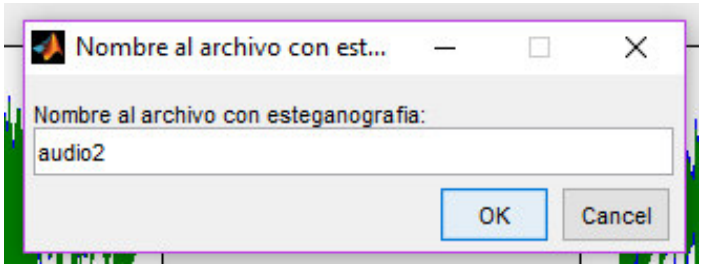


Figura 17 Nombre al nuevo archivo de audio

4.1.2. PROGRAMA DE RECUPERACIÓN DE DATOS

Para extraer los datos en un archivo de audio con esteganografía se establece el siguiente diagrama de flujo (ver Figura 18 Diagrama de flujo de extracción de datos), el cual permite comprender de mejor manera su funcionamiento:

Figura 18 Diagrama de flujo de extracción de datos

Una vez entendido el diagrama de flujo anterior se procede a implantar el programa y de este modo se tiene la siguiente interfaz gráfica Figura 19 Interfaz gráfica de extracción del mensaje del audio:

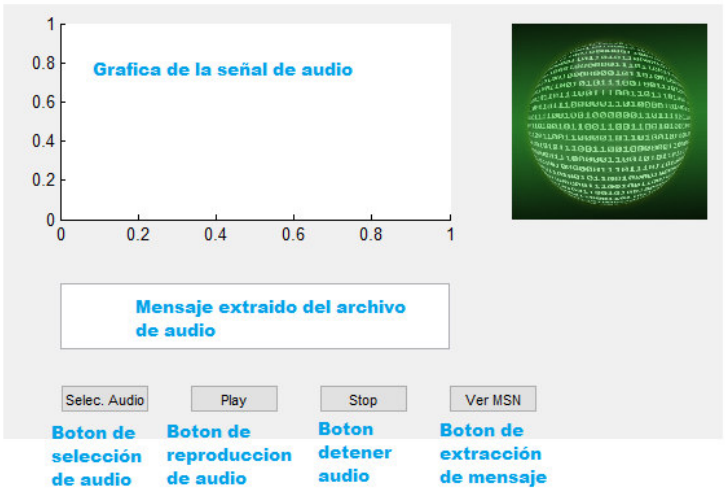


Figura 19 Interfaz gráfica de extracción del mensaje del audio

Para poder extraer el mensaje de un archivo de audio en la aplicación realizada, es necesario seguir los siguientes pasos:

1. Seleccionar el audio con esteganografía de donde se extraerá el mensaje oculto Figura 20 Selección del audio.

Figura 20 Selección del audio

2. Una vez seleccionado el audio se procede a dar “play” y se podrá escuchar el audio y observar la señal que este posee Figura 21 Reproducción del audio con esteganografía.

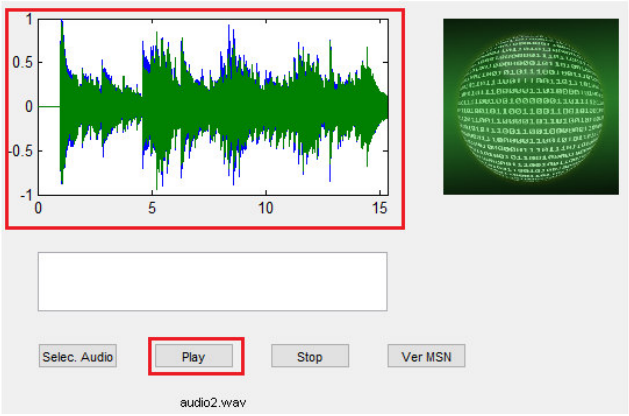


Figura 21 Reproducción del audio con esteganografía

3. Procedemos a dar click en “Ver MSN”, si la información de la clave que posee el archivo con esteganografía es correcta se podrá ver el mensaje oculto Figura 22 Extracción del mensaje oculto, caso contrario no se podrá extraer nada del archivo.

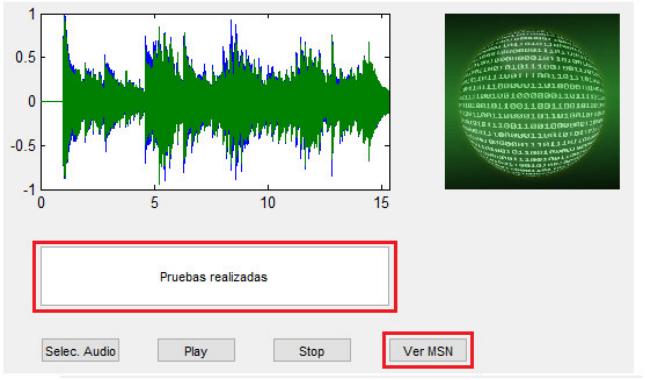


Figura 22 Extracción del mensaje oculto

La clave que contiene el archivo de audio con esteganografía es la seguridad que posee para que la información no sea extraída con facilidad, la clave que se colocó en el archivo es un vector de 8 bit´s el cual fue se lo puede poner en cualquier parte del archivo de audio en el caso específico de esta tesis fue colocado después de la cabecera WAV, para que de este modo si alguien desea obtener la información del archivo, primero la clave deberá ser la

correcta caso contrario no se obtendrá nada del archivo de audio con esteganografía.

Para poder extraer el mensaje contenido en el archivo de audio con esteganografía, se deben seguir los siguientes pasos en el programa creado:

Extraemos la clave del archivo de audio con esteganografía y comparamos si la clave es la correcta:

Figura 23 Clave del mensaje `clave=bitget(datos(1:8),lsb)';`

```
clave =

     1     1     1     0     1     1     1     0
```

Figura 23 Clave del mensaje

Si la clave extraída y la que contiene el programa son las mismas, se procede a extraer el mensaje oculto:

```
if clave==[1 1 1 0 1 1 1 0]
```

Extraemos la longitud del texto oculto para saber la cantidad de información que se encuentra oculta:

Figura 24 Cantidad de bits ocultos se extra el texto oculto.

```
% Extraer la longitud del texto de la primera 9 al 28 muestras de
        m_bin=zeros(10,1);
        n_bin=zeros(10,1);

        m_bin(1:10)=bitget(datos(9:18),lsb);
        n_bin(1:10)=bitget(datos(19:28),lsb);

% Convertir la longitud a decimal
        len=bi2de(m_bin')*bi2de(n_bin');

        msg_bin=zeros(len,1);
```

```
len =

   144
```

Figura 24 Cantidad de bits ocultos

Al momento de tener la cantidad de información que se encuentra oculta, se procede a extraerla, en este caso se extrae 144 bit`s que se colocaron en el archivo de audio y se los coloca en una matriz:

Figura 25 Matriz de bits, se extrae la matriz oculta.

```
%extraer el mensaje del archivo wav
        msg_bin(1:len)=bitget(datos(29:28+len),lsb); %extraer en una
columna
        msg_bin_re=reshape(msg_bin,len/8,8);    %lo convertimos en
una matriz
```

```
msg_bin_re =

     0     0     0     0     1     0     1     0
     0     1     0     0     1     1     1     0
     1     0     1     0     1     1     1     0
     1     0     1     0     0     1     1     0
     0     1     0     0     0     1     1     0
     1     0     0     0     0     1     1     0
     1     1     0     0     1     1     1     0
     0     0     0     0     0     1     0     0
     0     1     0     0     1     1     1     0
     1     0     1     0     0     1     1     0
     1     0     0     0     0     1     1     0
     0     0     1     1     0     1     1     0
     1     0     0     1     0     1     1     0
     0     1     0     1     1     1     1     0
     1     0     0     0     0     1     1     0
     0     0     1     0     0     1     1     0
     1     0     0     0     0     1     1     0
     1     1     0     0     1     1     1     0
```

Figura 25 Matriz de bits

Por último, de los bits tomamos su valor decimal y por último en letras las cuales corresponden al mensaje oculto Figura 26 Mensaje oculto:

```
        msg_double=bi2de(msg_bin_re); %convertimos a numeros
decimales

        msg=char(msg_double)'  %los numeros se convierten en letras
    end
```

```
msg =

Pruebas realizadas
```

Figura 26 Mensaje oculto

La cantidad máxima de información que se puede ocultar en un archivo de audio depende de la capacidad de bits que este contenga, todo depende de la duración en minutos del archivo de audio.

Para saber cuál es la cantidad de información que se puede ocultar en aun archivo de audio se debe tomar en cuanto el tiempo de duración y la velocidad de bits; para el archivo de audio con el que se trabajó el tiempo de duración es de 15 segundos y una velocidad de 1411 kbps; por lo que la cantidad de información que se podría ocultar es la siguiente:

$$cantidad = tiempo\ de\ duración * velocidad\ de\ transmision$$

$$cantidad = 15\ s * 1411\ \frac{bits}{s}$$

$$cantidad = 21165\ bits$$

Con el cálculo anterior se puede determinar que se podría cambiar un total de 21165 bits y para conocer la cantidad de letras solo se debe dividir para 8, y de este se modo se conocería su cantidad máxima:

$$letras = \frac{cantidad}{8}$$

$$letras = \frac{21165}{8}$$

$$letras = 2645$$

CAPÍTULO 5

PRUEBAS Y ANÁLISIS DE RESULTADOS

Para las pruebas realizadas con el audio original y con el audio con mensaje oculto, se estableció varios escenarios en donde mediante una encuesta MOS se establecerá si existe algún tipo de variación de un audio con respecto al otro. Figura 27 Escenario para la realización de la encuesta MOS

Figura 27 Escenario para la realización de la encuesta MOS

Para la realización del proyecto de investigación fue necesario tener a disposición un archivo de audio con extensión .wav para poder ocultar la

información, del mismo modo fue necesario tener el software Matlab en donde se creó el programa para ocultar la información en archivos de audio.

Una vez creado el programa en Matlab y haber ocultado el mensaje en el archivo de audio con extensión .wav, se procedió a realizar diferentes escenarios con los audios esteganográficos, se propuso 3 escenarios debido a que los archivos de audio se debían acoplar a situaciones cotidianas de cualquier persona, para de este modo poder comparar bajo qué condiciones es más factible compartir el archivo esteganográfico sin que este sufra un cambio significativo o levante alguna sospecha, posteriormente se realizó una encuesta MOS a un grupo de personas para ver si ellos detectan algún cambio con respecto al audio original.

1. **Escenario 1:**
    - En este escenario después de haber creado los archivos con esteganografía, se los dejo en la carpeta que previamente estaba direccionada, sin moverles a ningún otro lugar hasta el día en que se extrajera el mensaje oculto y poder determinar si habían sufrido algún cambio, para luego ponerlos a prueba en la encuesta MOS.

2. **Escenario 2:**
    - El escenario propuesto consiste en que los archivos de audio con esteganografía después de haber sido creados, se los pasara a un dispositivo portátil en este caso fue un Smartphone, donde los archivos de audio pasaron durante una semana reproduciéndose periódicamente, y al momento de extraer la información oculta poder determinar si el contenido sufrió algún cambio por el continuo uso, para luego ser sometidos a la encuesta MOS.

3. **Escenario 3:**
    - En este escenario los archivos de audio con esteganografía se los comprimió en un archivo .rar, que posteriormente fue enviado por correo electrónico y se los extrajo, a continuación se procedió a reproducir el archivo de audio con esteganografía para comprobar si había sufrido algún cambio, inmediatamente se procedió a extraer el mensaje oculto y de este modo poder realizar la encuesta MOS.

En los escenarios propuestos anteriormente se obtuvo como resultado que todos los audios con esteganografía al ser extraídos el mensaje que tenían oculto, los mensajes se encontraban completos y sin ningún signo de alteración; el escenario donde se notó un poco el deterioro la calidad de audio fue en el tercero ya que en este los audios fueron comprimidos y enviados por internet y debido a este pudieron ser afectados un poco.

Para determinar si existe algún cambio o se nota alguna perturbación de los audios con esteganografía se procedió a realizar una encuesta MOS a un grupo de 50 personas, donde la pregunta era la siguiente:

¿Respecto al audio original, califique los audios con mensaje oculto?
En caso de notar algún cambio comente donde y que cambio noto

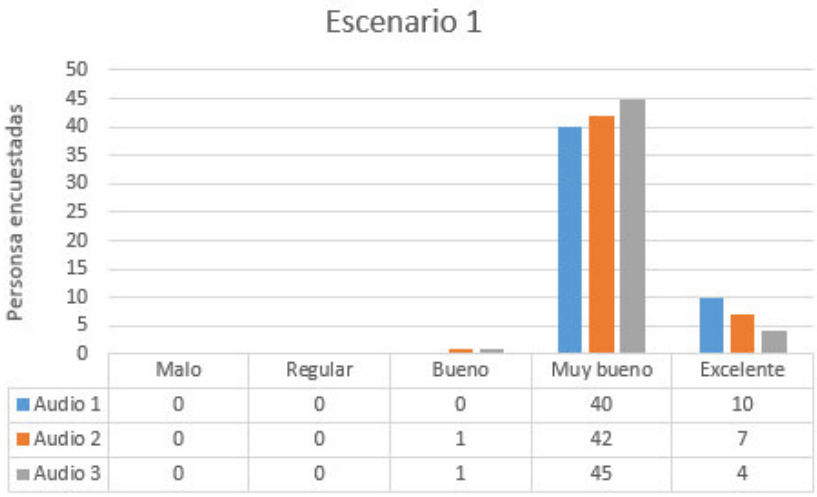
(1) Malo	(2) Regular	(3) Bueno	(4) Muy Bueno	(5) Excelente

Tabla 9 Tabla de calificación de audios

	Calificación	Comente donde noto algún cambio
Audio 1		
Audio 2		
Audio 3		

Dando como premisa la pregunta anterior se procedió a encuestar al grupo de personas y se obtuvo los siguientes resultados:

- En el primer escenario se obtuvo resultados muy satisfactorios, debido a que la mayoría de los encuestados notaron que la calidad de audio no vario en absoluto respecto al audio original, por lo que en la Figura 28 Resultados Escenario 1 se puede ver graficados la respuesta de los encuestados.



	Malo	Regular	Bueno	Muy bueno	Excelente
Audio 1	0	0	0	40	10
Audio 2	0	0	1	42	7
Audio 3	0	0	1	45	4

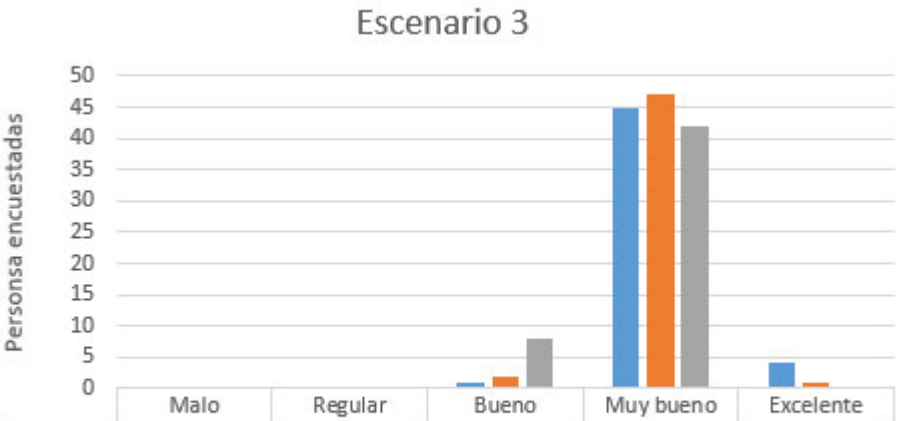
Figura 28 Resultados Escenario 1

- En el segundo escenario se obtuvo resultados muy similares al del primer escenario, sin embargo un pequeño grupo de personas tres exactamente, pudieron notar que el audio había sufrido un pequeño descenso en su calidad de sonido, por lo que en la Figura 29 Resultados Escenario 2 se puede notar la respuesta de los encuestados.

	Malo	Regular	Bueno	Muy bueno	Excelente
Audio 1	0	0	0	42	8
Audio 2	0	0	1	45	4
Audio 3	0	0	2	47	1

Figura 29 Resultados Escenario 2

- En el tercer escenario debido a que los audios fueron comprimidos, las personas encuestadas notaron un cambio en la calidad de audio respecto al audio original, por lo que en la Figura 30 Resultados Escenario 3 se puedo notar que los resultados varían respecto a los escenarios 1 y 2.



	Malo	Regular	Bueno	Muy bueno	Excelente
Audio 1	0	0	1	45	4
Audio 2	0	0	2	47	1
Audio 3	0	0	8	42	0

Figura 30 Resultados Escenario 3

La encuesta MOS nos permite determinar si un grupo de personas es capaz de detectar algún cambio del archivo con esteganografía con respecto al archivo original, sin embargo esta prueba es muy subjetiva y todo depende de las condiciones del ambiente de donde se realice las pruebas y de las personas en sí mismo; por lo que hay la necesidad de tener valores estadísticos y poder determinar con valores reales que diferencia existe entre los 2 archivos.

Para poder tener un valor estadístico real se tomó en cuenta dos parámetros los cuales son MSE y PSNR, los cuales nos van a permitir determinar la diferencia entre el archivo de audio original y el archivo con esteganografía.

En primer lugar se empezó a extraer el MSE de los archivos de audio con esteganografía respecto al audio original, por consiguiente se obtuvo que el MSE en los 3 escenarios propuestos tuvo un valor insignificante Figura 31 MSE de los archivos de audio con esteganografía, debido a que la información oculta en los archivos con audio con esteganografía es mínima comparada con la cantidad que se puede ocultar.

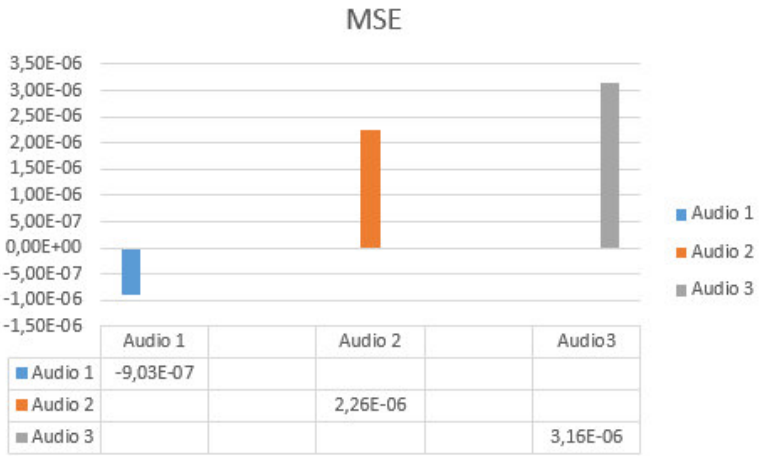


Figura 31 MSE de los archivos de audio con esteganografía

Una vez determinado el MSE de los archivos de audio con esteganografía se puede determinar el PSNR de dichos archivos por lo que al momento de comparar con el audio original Figura 32 PSNR de los archivos de audio con esteganografía, se pudo determinar que el ruido que existe en estos archivos de audio con esteganografía es mínimo ya que los valores obtenidos son muy altos como para considerar que existe algún daño.

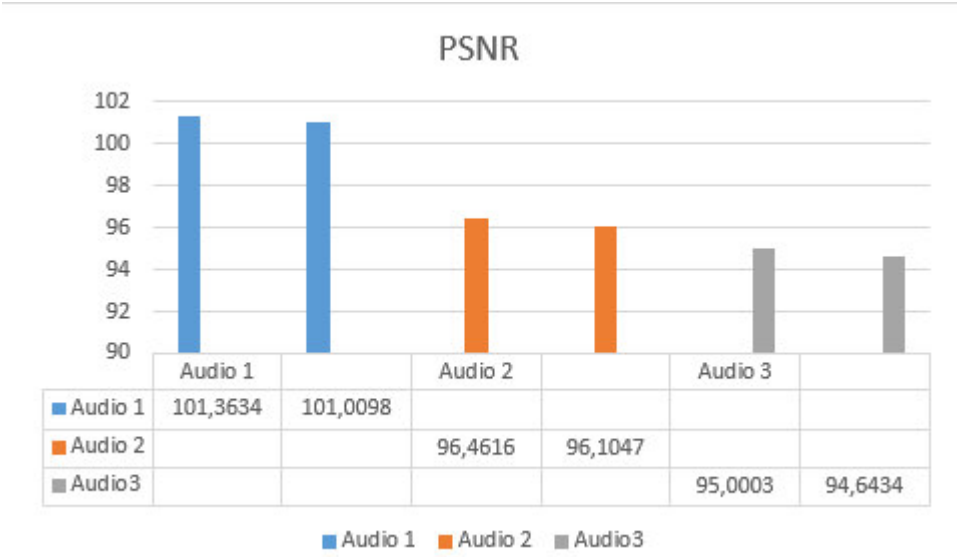


Figura 32 PSNR de los archivos de audio con esteganografía

CAPÍTULO 6

CONCLUSIONES Y LÍNEAS DE TRABAJOS FUTUROS

6.1. CONCLUSIONES

- Se logra ocultar información en archivos de audio con formato .wav, desde una palabra simple como “hola”, hasta un texto completo con numeración y símbolos.

- La información que se oculta dentro del archivo de audio puede comprometer la calidad de sonido, debido a que cuanta más información se oculta dentro del archivo, más cambios va a sufrir debido al reemplazo de los últimos bits por los del mensaje a ser enviado.

- En los diferentes escenarios de prueba, se puede concluir que los archivos que se encontraban en la computadora y se colocan en un reproductor de música, son los que menos sufrieron un cambio y no pueden ser detectados por las personas encuestadas.

- Los archivos comprimidos y enviados por internet sufren un ligero cambio en su calidad de audio, esto se debe ya que al ocultar información cambiando los bits menos significativos y ser comprimidos para ser enviados por internet se ven afectados dos veces y la calidad de audio puede sufrir alguna variación.

- Una característica fundamental que se determinó para realizar la implementación de la esteganografía en archivos de audio, es que los archivos deben tener una extensión .wav, debido a que su calidad de audio es superior a otros formatos y es el estándar para trabajar sobre la plataforma de Matlab.

- Al momento de comenzar a implementar el programa de esteganografía para ocultar información, se debe tener en cuenta que la cabecera del archivo de audio no puede sufrir ninguna alteración, ya que la cabecera contiene información específica de archivos wav.

- Al momento de realizar un análisis estadístico se tomó en cuenta el valor de MSE de los archivos de audio con esteganografía con respecto al original, se puede determinar que los valores obtenidos eran ínfimos, esto se debe a que la técnica utilizada fue debidamente aplicada y la cantidad de información oculta es pequeña comparada con la capacidad que posee el archivo de audio original.

- Respecto al PSNR obtenido de los archivos de audio, se determinó que el ruido o la afectación realizada al momento de ocultar información fue mínima, ya que los valores obtenidos están sobre el 90%, y se puede determinar que la calidad de audio no fue afectada en su totalidad.

- En la encuesta MOS realizada a diferentes personas, la mayoría no noto ningún cambio en la calidad de audio respecto al audio original, sin embargo hubo un grupo reducido de personas que pudo detectar ligeros cambios, mas no pudieron detectar si los cambios era ruido o la calidad de audio había disminuido.

6.2. LÍNEAS DE TRABAJOS FUTUROS

- Con la culminación del presente trabajo de investigación se da paso a diferentes trabajos futuros debido a la versatilidad que el tema posee en diferentes aspectos relacionados a la seguridad.

- El próximo desafío a realizarse en este trabajo de investigación, es de integrar el programa a diferentes formatos de audio, ya sea para archivos con formato .wav, .mp3, y otros.

- Al momento de finalizar este trabajo de investigación una línea de trabajo futuro que se presentó, es el de implantar la esteganografía en archivos de video, debido a que los videos en si son una serie de imágenes consecutivas junto a una pista de audio; de este modo se puede aplicar la esteganografía con la técnica de LSB al audio de archivos de videos.

Bibliografía

B, A. (s.f.). *computo.* Obtenido de http://computo.fismat.umich.mx/~karina/tesisLicenciatura/apend_b.html

Bustamante, C. P. (24 de 5 de 2016). *El lado del mal.* Obtenido de http://www.elladodelmal.com/2016/05/esteganografia-con-ficheros-de-audio.html

Cantanhede, H. S. (2009). *Esteganografia em Audio e Imagem utilizando a técnica LSB.* Catalao: Universidade Federl De Goiás.

Casierra, J. P. (2009). *Implementacao de um sistema esteganografico para insercao de textos em sinais de áudio.* Recife: Universidad Federal de Pernambuco.

Chacon, A. (s.f.). *Emaze*. Obtenido de https://www.emaze.com/@AIIIIZCI/sonido

Checa, E. A. (2014). *Implementación del Algoritmo Esteganográfico F5 para imágenes JPEG a Color.* Quito: Escuela Politécnica Nacional.

clases06. (s.f.). *eumus.* Obtenido de http://www.eumus.edu.uy/eme/ensenanza/electivas/dsp/presentaciones/clase06.pdf

criptored. (s.f.). *criptored*. Obtenido de http://www.criptored.upm.es/crypt4you/temas/privacidad-proteccion/leccion7/images/image018.gif

Daniel, & Santiago. (11 de 7 de 2015). *myslide.* Obtenido de http://myslide.es/documents/senales-de-audio-matlab.html

G, C. J. (17 de 11 de 2006). *Slideshare.* Obtenido de http://es.slideshare.net/cjumbo/introduccin-a-la-esteganografa

Gabriel, P. (2010). *numerical.* Obtenido de http://www.numerical-tours.com/matlab/audio_3_gabor/

Gámez, I. B. (2008). *Técnica de inserción en video aprovechando el mismo ancho de banda.* Mexico D.F.: Instituto Politecnico Nacional.

Garnacho, A. R., Juan M, E.-T., & Castro, J. C. (s.f.). *Portal.uc3m.* Obtenido de http://portal.uc3m.es/portal/page/portal/inst_juan_velazquez_velasco/cursos_seminarios/seminario_descubriendo_reverso_internet_web_mining/Hernandez-Esteganografia.pdf

Gimenez, M. (21 de 9 de 2015). Obtenido de http://esteganografia.cursodesearchmarketing.com/esteganografia-digital-gamopedica/

Hamdaqa, M., & Tahvildari, L. (2011). *ReLACK: A Reliable VoIO Steganography.* Waterloo: University of Waterloo.

ihiu01. (s.f.). *nereida.* Obtenido de http://nereida.deioc.ull.es/~pcgull/ihiu01/cdrom/matlab/contenido/node2.html

Intech. (s.f.). *Intech open.* Obtenido de http://cdn.intechopen.com/pdfs-wm/21375.pdf

Karina, G. M. (s.f.). *Revistas bolivianas.* Obtenido de http://www.revistasbolivianas.org.bo/pdf/rits/n8/n8a28.pdf

lightlink. (s.f.). *lightlink*. Obtenido de http://www.lightlink.com/tjweber/StripWav/WAVE.html

MathWorks. (2016). *Mathworks.* Obtenido de http://www.mathworks.com/products/matlab/features.html

Matlab. (s.f.). *Matlab*. Obtenido de http://1.bp.blogspot.com/-1pAyd5lpZ-c/T1EGjP5le4I/AAAAAAAAKlY/PwMypHbmRlU/s1600/matlab.png

Mazurczyk, W., Szaga, P., & Szczypiorski, K. (2015). *Using Transcoding for Hidden Communication in IP Telephony.* Poland: Warsaw University of Technology.

Miriam, H. P. (2012). *Mensajes subliminales.* Mexico.D.F.: Escuela Superior de Ingenieria Mesanica y Electrica Unidad Zacatenco.

Muñoz, D. A. (2 de 1 de 2014). *Criptored.* Obtenido de http://www.criptored.upm.es/crypt4you/temas/privacidad-proteccion/leccion7/leccion7.html

Orbegozo, I. S. (2011). *Técnicas de auto escalado de clous computing aplicadas al estegoanálisis.* Madrid: Universidad Complutesne De Madrid.

Pérez Días, N. A., & Rodríguez Clemente, C. A. (2010). *Mensajes subliminales en formato wav.* Mexico, D.F.: Escuela Superior de Ingenieria Mecanica y Electrica Unidad Zacatenco.

Peris, J. (s.f.). *IT campus Academy.* Obtenido de http://itcampusacademy.com/ingenieria-software/tienes-un-oido-fino-cifrado-y-esteganografia-de-datos-en-ficheros-de-audio-y-un-test-de-estegonanalisis-con-tu-oreja/

Sapp. (s.f.). *Sapp*. Obtenido de http://soundfile.sapp.org/doc/WaveFormat/

Vico, J. D. (2010). *Esteganografia y estegoanalisis: Ocultacion de datos en streams de audio vorbis.* Madrid: Universisdad Politecnica de Madrid.