12-2018

Steganography A Data Hiding Technique

Naga Ranijth Kumar Kesa
nkkesa@stcloudstate.edu

Steganography–A Data Hiding Technique

by

Naga Ranjith Kumar Kesa

A Starred Paper

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree

Master of Science in

Information Assurance

December, 2018

Starred Paper Committee:
Mark B. Schmidt, Chairperson
Lynn A. Collen
Balasubramanian Kasi

Abstract

Steganography implements an encryption technique in which communication takes place by hiding information. A hidden message is the combination of a secret message with the carrier message. This technique can be used to hide the message in an image, a video file, an audio file or in a file system. There are large variety of steganography techniques that will be used for hiding secret information in images. The final output image is called as a stego-image which consists of a secret message or information. Imperceptibility, payload, and robustness are three most important parameters for audio steganography. For a more secure approach, encryption can be used, which will encrypt the secret message using a secret key and then sent to the receiver. The receiver after receiving the message then decrypts the secret message to obtain the original one. In this paper, compared steganography with cryptography, which is an encrypting technique and explained how steganography provides better security in terms of hiding the secret message. In this paper, the various techniques are illustrated, which are used in steganography and studying the implementation of those techniques. Also, demonstrated the implementation process of one of the steganography techniques. A comparative analysis is performed between various steganographic tools by using the sample test images and test data. The quality metrics such as PSNR and SSIM are calculated for the final output images which are used for rating the tools. This paper also discusses about the Steganalysis which is known as the process of identifying the use of steganography.

Keywords: Steganography, Cryptography, Steganalysis, Stego-image, PSNR, SSIM.

Table of Contents

List of Tables

List of Figures

Figure Page

Figure Page

Chapter I: Introduction

Introduction

As the advancement of the internet increased, it has become an important factor in information technology and plays a vital role in communication. The security of information is becoming a bigger concern. Cryptography is the technique which secures the communication. There are various methods developed for encrypting and decrypting the information, which secures the message. Due to the increase of the technology, sometimes cryptography is not enough for keeping the information as secret. It also important to retain the existence of the information secret. Steganography is the technique which is used to implement it. It is achieved by hiding the information inside other information, thus the existence of communicated message is hidden. This chapter provides the information about how steganography is different from cryptography and also how the steganography process is performed.

During the Second World War, Germans developed the Microdot technique. Using that technique, they have decreased the size of the information such as photographs to the typed period size. It is very difficult to detect, as the cover message is sent over a channel which contains the hidden message on one period of the paper. In today's world steganography is most commonly used on computers with networks as the delivery channels and digital data as the carriers (Provos & Honeyman, 2003).

Steganography is different from the cryptography because cryptography focuses on keeping information secret whereas steganography focuses on making the existence of the information secret. Though both ways are used to protect the data/information from outsiders, the technology is not perfect and can compromised. Once it is suspected or revealed that the hidden information exists, the steganography purpose is defeated partly. Steganography can be

strengthened by combining it with the cryptography. It is known that watermarking is a method used for hiding the trademark information in software, images and music. It is not considered as original form of steganography (Patel, & Tahilraman, 2016).

In steganography the message is hidden in the image, but watermarking will add something on top of the image for example a word “Confidential”, which will become part of the picture. There is a misconception that steganography is related or similar to encryption, but in real they are different. Encryption is a technology which converts the message from a readable to an unreadable format for protecting the sensitive data. Whereas, in steganography the information is hidden from the plain view and it is not mandatory to be encrypted.

The main drawback or disadvantage of the encryption is that the information is encrypted and sent over a channel and if someone captures an email or the data stream, then it raises suspicion that the data is encrypted when they see it. The person who monitors the network traffic will investigate why this is encrypted and will be using various tools for figuring out the encrypted message. In short, it can be said that encryption will provide confidentiality not secrecy.

Problem Statement

The internet is considered the most powerful tool of information and communication technology. The underlying issue has always been security that is provided to secure the information. Unfortunately, sometimes it is not enough to keep the contents of a message secret but also to send the secret information securely. How a secret and confidential information is hidden and communicated securely, and which will be the best way for communicating. These are the things to know, for achieving the safe communication.

Nature and Significance of the Problem

The purpose was to analyze the information hiding techniques that may help the users sharing the information so that, such information will reach the intended person(s) without being detected by other computer users (intruders or attackers) when carrying out day to day tasks and organizational activities.

Objective of the Study

The biggest challenge faced by information users is when they try to hide information from those who may not be authorized to see the information. The purpose of this study is to know how steganography, an information hiding technique, helps to overcome the problems faced by them and to test and evaluate the validity, utility and usability of various techniques and testing the quality of available steganography tools by hiding and encrypting information using images and keys respectively.

Study Questions/Hypotheses

1. How the exchange of information can be secured by hiding the existence of secret information using Steganography?
2. What are various tools and techniques that are used in steganography and the applications where the steganography is being used?
3. What are Pros and Cons of using Steganography over the other data security mechanisms?

Summary

This chapter has covered the introduction to the steganography and its process. The purpose of steganography in securing the exchange of information over the internet. The

objective and the driving force towards this research. Also, the research questions which will be addressed as part of the research.

Chapter II: Background and Review of Literature

Introduction

Steganography is the art of communication with the invisible information, i.e., it plays an important role in information security. The term steganography literally means “covered writing” which is derived from Greek. There are three elements to hide the information using Steganography: the cover image which hides the secret message, the secret message and the stego-image (which is cover object with message embedded inside it). The image (stego-image) being used for steganography purposes must be same as the original image, as to avoid drawing suspicion to the stego image. Data invisibility and image embedding capacity are two primary requirements that have been extensively researched in different steganography techniques. This chapter will provide the details about the process of steganography (Provos & Honeyman, 2003).

Steganography is the technique which hides the information in such a way that no third person other than the receiver knows that there is a secret message hidden inside the information that is transferred. And the main advantage of this technique is no other person except the receiver who is intended to receive the information can be suspected that there is hidden information present in the message that is being passed over a channel. Cryptography is the similar type of technique in which the main aim is to protect the data from knowing to the unauthorized persons expect the receiver who is intended to receive the message. But in cryptography, the original message is converted form human readable format to unreadable format. If a third person or an attacker sees the message and he finds it is encrypted and will give a sign of suspicion and then he tries to decode the information which leads to the leak of the secret message. So, in case of the steganography the chance of suspicion is less compared to the

cryptographic technique because the main objective of the steganography is to make the existence of the secret information or message invisible to the third person apart from the sender and the receiver.

In the present world of Information Technology (IT), there is huge advancement in the technology, which is directly proportional to the increase of the users of the internet. There are multiple purposes for which the technology is used and due to the increase in the technology the security has become the major concern and it became a critical step for the organizations for providing it. Just like each coin has two sides, even the rise of technology is impacted in both a positive and a negative way. It became a tough task to overcome the attacks on the information on the internet. Nowadays, most of the data is being stored in the cloud storage. The access levels play a major role in accessing the information. The authorization and authentication are the major parts and providing them accurately and reviewing them on a regular basis is the initial step. And protecting the passwords for that information is the major task. Many organizations will be having the confidential data like the banking sector and importantly in the military organizations the data should be more and more secure and safe. Leaking of information from them will lead to a severe problem. Apart from the storage of the data, many online transactions will happen like purchasing, transferring of funds in the banking sector and storing the personal information on the social media are to be secure. There are many attacks for the data like the phishing, third party attacks and social engineering which make people to share their personal information. The main problem and where the security is necessary, is when the sensitive data like the personal details, passwords and credit card details are captured by the third persons or attackers. Hence, the cryptography technique was invented initially for sending the messages

secretly by producing a cipher text which is known as encoded or encrypted information. The cipher text or the encrypted message will contain the original message in the form which the humans cannot read or decrypt and even by the computer unless proper key is known to decrypt it. Since the cipher text makes the attacker to suspect on the message because of the reason it is not in a readable format. So, steganography will eliminate the factor of suspecting the data by the attacker since the message is hidden inside another information and makes the existence of the message that is to be hidden invisible.

Background Related to the Problem

Hiding information into a medium requires following elements.

a) The cover medium(C) that holds the secret message.

b) The secret message (M), it can be a plain text, an image file or any other type of data.

c) The steganography techniques which are going to be used to hide the information.

d) A stego-key (K) which will be used for hiding and un-hiding the message.

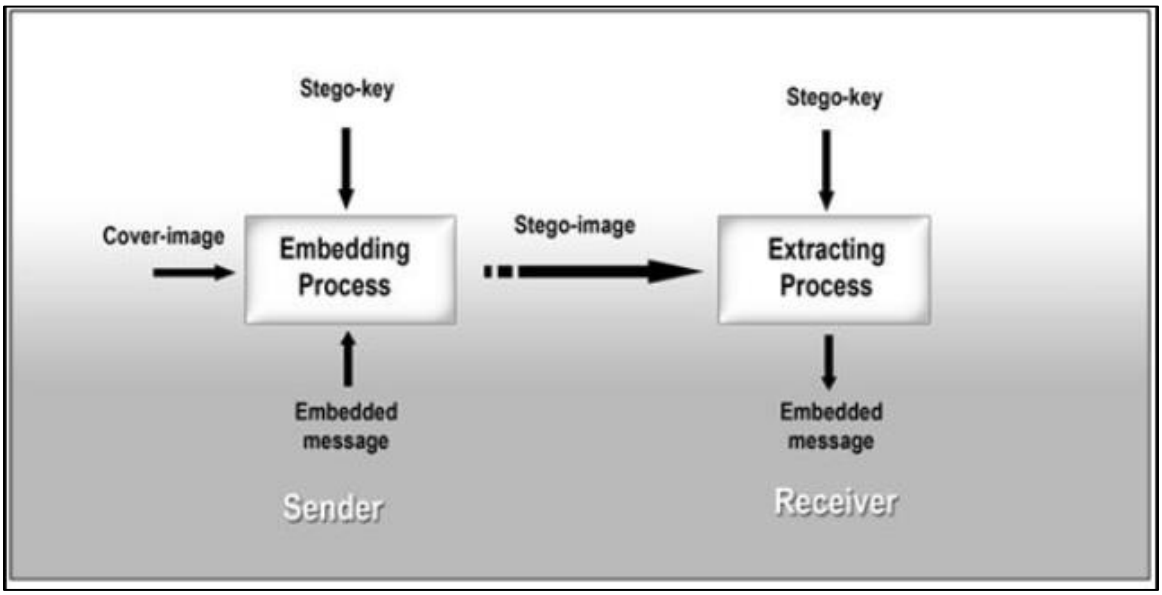

Figure 1: Steganography process (student web, n.d.).

In the modern style, considering the cover medium, steganography can be divided into five types:

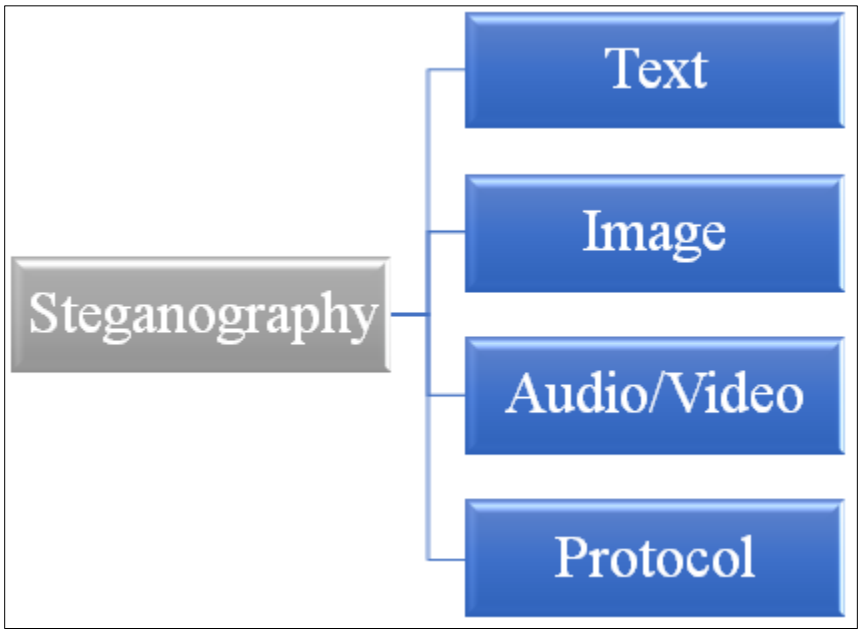

Figure 2: Types of steganography (Learning & Exercise, 2017).

Text steganography. It is the very common technique of steganography in which the information which is to be hidden is placed in a text file. After the invention of the internet and different type of digital file formats, it has decreased its importance. Text stenography using digital files is not used very often because the text files have a very small amount of excess data. It is a difficult task to know the presence of hidden text if we use text steganography. Text steganography has less noise compared to an image. So, hiding the information using the text will result in low embedding capacity. There are several methods in the text steganography (Banerjee & Indradip, 2011)

***Modifying spaces*.** In the cover text the data is hidden by modifying the spaces. A word processor can perform (1) the spaces between the words in a sentence, (2) the spaces which are at the ending of each line, (3) the spaces immediate to the punctuation marks. Generally, the word processor automatically adjusts the spaces for justifying the right margin, they cannot be controlled manually by the user. So that type of word processor will be rewritten to (1) allow a user to control for the spaces. (2) contain a list of the blank spaces precise sizes in a document, so the hidden bits can retrieve (Banerjee & Indradip, 2011).

In the primitive word processor, the spaces will be having a fixed size, a bit is hidden at the end of each sentence by adding 2 or more spaces to each sentence, where the number of spaces indicates the value of the bits. One space will represent hidden 0 and two spaces represent hidden 1 (Banerjee & Indradip, 2011).

***Syntactic methods*.** In this method, it will be achieved by altering the text and keeping the meaning of the text as it is. The method explained previously is vulnerable because of the reason that it uses the spaces and punctuations which are noticeable if they are inconsistently used they

make the observer to be suspicious of the message. This method is so difficult to implement because it is harder to computer to make understand (Banerjee & Indradip, 2011).

Example. If the below message is shared by the sender to the receiver, when an attacker captures this information it looks like a normal message to him or her. But there is secret message hidden in the sentence.

Since everyone can read, encoding text in

neutral sentences is doubtfully effective

If we consider the first letter of every word as highlighted in the below sentence, then we can able to get "Secret Inside" which is a secret message which will not be known to the attacker.

Since everyone can read, encoding text in

neutral sentences is doubtfully effective

"Secret Inside"

So, this conversation of considering the first letters of the word is made between the sender and receiver offline and hence the existence of the secret message is made invisible to the attacker.

Image steganography. This is the popular method in which images are used as the cover medium for steganography. A message is inserted in a digital image by using an algorithm and the secret key. There are various ways for embedding the secret message into an image. Also, the secret key is an optional thing unless it is required. Though, it adds additional security for the information that is being transferred. The result which is a stego-image is sent to the receiver. On the receiver side, the stego-image is processed by the extraction of algorithm

using the same secret key. In the process of communicating the stego-image, other than the authenticated persons nobody can be able to notice the existence of the secret message which is hidden in the image though they identify the transmission of a stego-image. So, it overcomes the problem of suspicion by the attacker or unauthorized persons who capture the information during the communication (Banerjee & Indradip, 2011).

The overall process of the image steganography is to hide the sensitive data or information inside a cover image without the degradation of the original image, hence providing the security by which no unauthorized person can access the information which is hidden. There are different methods by which an image steganography is achieved. Below is the classification of the techniques of the image steganography.

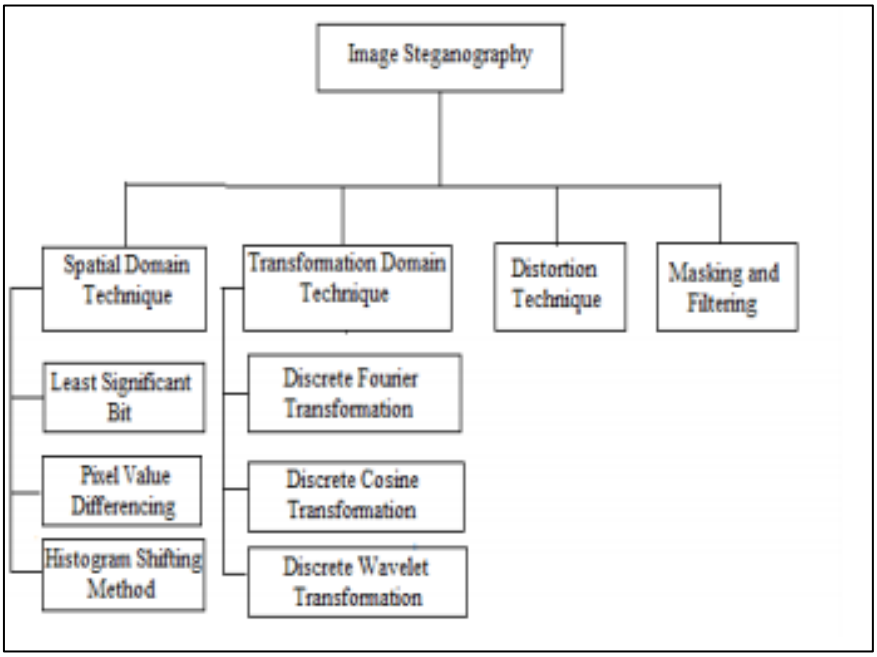

Figure 3: Image steganography techniques (Banerjee & Indradip, 2011)

Spatial domain methods. In the spatial domain method of steganography, for hiding the information it directly changes the pixel values of the image. It means the bits of the pixel values

of an image are replaced with the bits of the secret message. There is furthermore classification in the spatial domain methods. Out of those Least Significant Bit (LSB) is the most commonly used method. Below are the various methods used in the spatial domain technique (Emam & Marwa, 2015).

Least Significant Bit (LSB). In this method the least significant bits of an image are replaced with the bits of the secret message. For example, if a 'c' is to be hidden inside the image, the ASCII value of that letter is converted into a binary number. The obtained binary number is stored in the least significant bit of image binary format. The data is stored in an image. The principle for this method is if we change the least significant bits then there will be a minor change in the image that will be not visible to the human eye (Emam & Marwa, 2015). So, if we pass the stego-image (the image obtained after inserting the secret message) is exchanged over a carrier, the secret message inside will not be known to any unauthorized person except the receiver who is the intended person to receive the information. The stego-image and the original image looks the same, visually we cannot detect the change. The images which are used as the cover image are of two types one is 24-bit images and the other one is 8-bit images. It varies in the limit of the storing the information. In 24-bit images we can able to store three bits of secret information inside each pixel of the image whereas in 8-bit images we can only store one bit of secret information inside each pixel of an image. The main benefit of the least significant bit method is it easy for implementing, has a high message payload and has a smaller chance of the original image quality degradation (Emam & Marwa, 2015).

Pixel Value Differencing (PVD). In this method, the cover image is the gray scale image, which is having the secret message as a long bit-stream. This method was proposed for

hiding the confidential information into the images which are called 256 gray valued. Pixel value differencing method is proposed by considering the fact that the human eyes are able to observe the minor changes in the smooth areas in the image, but they are not able to observe larger relative changes at the edges in the image (Raja, Vanugopal, & Lalit, 2004).

This method will be using the difference between the neighboring pixels for determining number of bits that can embedded in an image. The larger the amount of difference is, the more bits of the secret message can be concealed inside the cover image. So, if there is a larger difference between the neighboring pixels, then we are able to insert more number of bits of the secret message. The image is scanned in a zigzag manner starting from upper left corner of an image. After scanning the image, it divides the image which is used to hide the information into the number of blocks in which each block composed of non-overlapping two consecutive pixels (Emam & Marwa, 2015).

In the block, the difference between two pixels will be used for categorizing the properties of the smoothness of the cover image. In this way, we can know where the pixels are located at. If the value of the difference is smaller than it shows that they are in the smooth area. If the value of the difference is larger, then they are around the edge area. So, the secret data bits are stored in the edge areas because if we store them at the smoother areas then they can be easily observed by the human eye.

Histogram shifting method. For representing an image graphically histograms are used. The histogram will represent the density and value of the particular pixel. The pixel is plotted for each part of an image. The histograms are useful for identifying the tonal distribution, pixel distribution and density of the colors. Apart from identifying these details a histogram will also

provide the lowest and highest values of the pixel in the graph. Histogram shifting is a method which is used for extracting or modifying the certain group of pixels from an image. The highest value in the histogram will be called maxima and the lowest value in the histogram will be called minima. The highest and lowest value of a histogram are called maxima and minima respectively (Raja et al., 2004).

The purpose of the maxima and minima values is to set a limit. When the value of the pixel is modified during the embedding process, the value should not exceed the maxima and minima limit. For manipulating an image, there are various algorithms which supports the functionality of the histogram. The number of pixels which constitutes for forming the peak in histogram of an image which is used as cover will equals to the capacity of hiding. The reason it determines the capacity of hiding is only a single peak is used in a cover image.

***Transformation domain technique*.** This method is used to hide the secret message in particular areas of image that is used as cover (Raja et al., 2004). By performing this process, it makes them stronger against different operations of image processing such as compression, enhancement and cropping. There are various transformation domain methods. For hiding the information, the basic approach is transforming the image that is used as the cover, pull the coefficients and lastly inserting the transformation. There are various Transformation domain techniques which are classified as following:

Discrete Fourier Transformation (DFT) technique. In this technique the embedding of the secret message is performed in frequency domain. This method is a complex way of hiding the secret message inside the frequency domain of an image. When this technique is applied for hiding the information, it converts the image which is used as the cover for hiding the data from

spatial domain to the frequency domain and then each pixel which is in the spatial domain are transformed into 2 parts, one is real and the other one is imaginary. The secret message bits are embedded in the real part of the frequency domain by excluding the first pixel. Inverse Discrete Fourier Transformation is applied after embed process which converts into spatial domain from the frequency domain. So, in order to extract the hidden message from the image, the image is converted from spatial domain to frequency domain. The we need to apply the DFT then followed by extraction algorithm to retrieve the original source image (Raja et al., 2004).

*Discrete Cosine Transformation (DCT)***.** This method transforms the image from the spatial domain to the frequency domain and then the image is separated into the spectral sub-bands based on the image visual quality. The image visual quality will be categorized as high, middle and low frequency components. In the figure, it shows the frequency components in which F_L indicates the low frequency component, F_H indicates higher frequency component. F_M will be using as the embedding region for providing the additional resistance to the techniques of lossy compression, avoiding the remarkable alteration of the cover image (Raja et al., 2004).

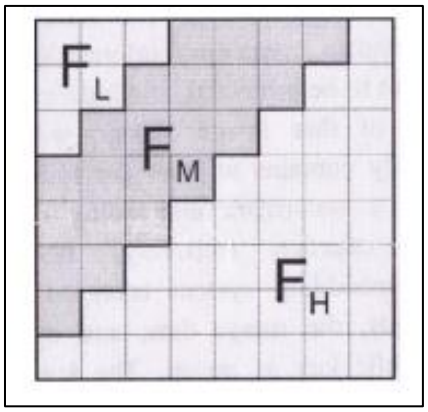

Figure 4: DCT regions (Raja et al., 2004).

*Discrete Wavelet Transformation (DWT) technique***.** This method is a new approach of applications of the wavelets. It is similar to the technique of least significant bit storing of the pixel but instead of replacing the bits of the original pixels in an image the data is stored into wavelet coefficients. The advantage of Discrete Wavelet Transformation over the Fourier Transformation is, DWT performs multi resolution and local analysis (Raja et al., 2004).

Distortion technique**.** In this technique, the data is stored by the concept of signal distortion. In this technique it is necessary to know about the actual cover image while performing the decoding process. During the process of encoding it applies the series of alterations for the cover image and then during the decoding process it will check the various differences comparing original cover image with the cover image which is distorted for recovering the secret information that is hidden in the cover image. By using the Distortion technique, it creates a stego object by the sender who sends the secret message by changing the changes which are sequential to cover image. So, the sequence of the changes made to the cover image indicates the particular secret message which is to be transmitted. The secret message will be encoded at pixels which are chosen pseudo-randomly in the image. The process of identifying the message by the receiver is when there is difference in the stego-image comparing with the cover image at the pixel of given message then the bit of the message is “1” or else it is “0”. The sender able to do the modifications to the “1” pixel values without affecting the statistical properties of original image. The receiver should be having the original cover image in order to retrieve the secret message which is the limitation for this technique. The cover image which is used for hiding the secret message is used only once that is by the sender in all other steganography techniques except this Distortion technique. So, if an authorized person have the

access for the cover image or captures the cover image then the intruder by doing the operations like rotating, scaling or cropping, he can able to detect the secret message from the stego-image easily (Raja et al., 2004).

Audio steganography. Audio steganography has a concern about putting an information in safe cover speech in a secure and robust manner. Communication, robustness, security and transmission are necessary for broadcasting important information to required sources while declining the access to the unauthorized persons. We can make an audible sound to be inaudible in the existence of other louder sound (Bhattacharyya, Bamerjee, & Sanyal, 2011). By the use of this property we can able to select a channel through which a message to be sent or hidden. Existing audio steganography software can embed messages in WAV and MP3 sound files. The below are the methods which are mostly used in the audio steganography LSB coding

***Parity coding*.** It is one strong audio technique of steganography. Instead of separating the signal into independent samples, it creates separate samples by breaking a signal and insert each bit of the message that should be secret from parity bit. For suppose, if parity bit of the selected portion is not matched then the secret message will be encoded, so this method will invert the Least Significant Bit of any one of samples among the selected region (Bhattacharyya et al., 2011).

***Phase coding*.** In this technique, the initial segment of the audio phase is replaced with reference phase which represents the information that is secret. The segments phase which are remaining is adjusted for preserving the proportion of the phase between the segments (Bhattacharyya et al., 2011).

Spread spectrum (SS). This method will spread the secret information across the spectrum of the frequency of audio signal. In this method the secret information is spread across the spectrum of the frequency of sound signal using the code that is independent of the original signal. So, finally the signal will use the bandwidth which is larger than what it originally requires for the transmission (Banerjee & Indradip, 2011).

The disadvantages of using the existing methods which are like parity coding and spread spectrum is, the human audibility is sensitive towards the noise and hence can detect the slightest noise that is introduced inside the sound file and one more problem associated with them is the robustness. Phase coding is having the disadvantage which is the very low transmission rate of the data because of the reason that the message which is to be hidden is encoded in first segment of the signal only. So, this method is used when it requires the small amount of data that needs to be transferred.

There are various data hiding techniques which are used to conceal the secret message/information inside an audio file, out of those Least significant bit (LSB) method is the easy and simple way for hiding the secret information within the digital audio signal/file in which the least significant bit of an audio file is replaced with the binary message. So, by using the lease significant bit we can store the larger amount of secret information that can be concealed within the audio file.

In this method, least significant bit of binary which is equivalent to each sample of a digitized audio signal/file will be replaced with the binary which is equivalent to the secret message. A program should be developed which reads the audio file (which the data needs to be embed) bit by bit and saves them into another file.

Video steganography. It is a technique used to hide any kind of files of any type of extension embed into a carrying Video file.

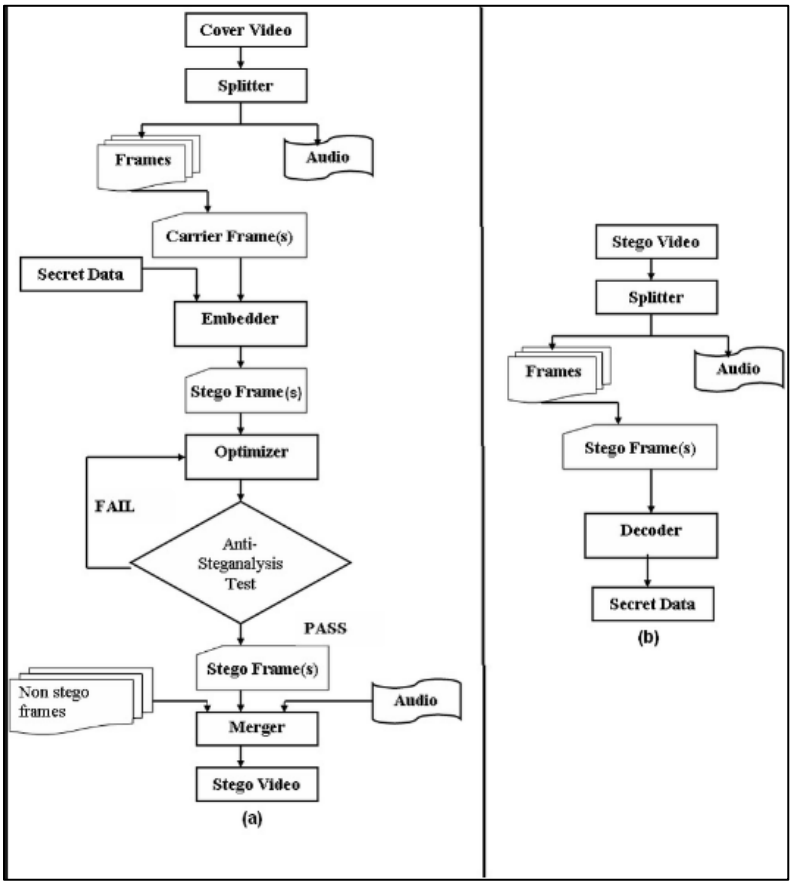

Figure 5: Video steganography process (Research gate, 2013).

Protocol steganography. It is used for inserting the information inside the network protocols like TCP/IP. The information will be hidden in some fields of header part of the TCP/IP packet which are either optional or never used (Bhattacharyya et al., 211).

Literature Related to the Problem

A digital image is demonstrated using a 2-D matrix at each grid point (i.e., pixel) of the color intestines. Typically, colored images utilize 24 bits, whereas, gray images use 8 bits to describe the color model, such as RGB model. To conceal information inside cover-image there

are several techniques in Steganography system. The Least Significant Bit (LSB) is one of the main techniques in spatial domain image Steganography. The spatial domain techniques are simple, easy to implement and manipulate the pixel bit values of the cover-image to embed the information which is to be hidden. The secret bits should not be changed and are directly written to the pixel bytes of the cover image. The LSB based image steganography which is the lowest significant bit in the byte value of the image pixel, embeds the secret in the least significant bits of pixel values of the cover image (Patel & Tahilraman, 2016).

The approach is Transform technique also known as Transform Domain Embedding, embed the message by modulating coefficients that are in transform domain, like the Discrete Cosine Transform (DCT) used in JPEG compression. Filtering and Masking techniques, are used to hide the information by the marking of an image, normally restricted to 24 bits and gray scale images, which is like the paper watermarks. These techniques will perform an analysis of image, thereby the information is placed in particular areas so that the secret message is better integral to cover image rather than just hiding the message in noise level.

In 2015, Abhilasha Ramdas Bhagat, A. Prof. Ashish B Dhembhare. had worked on a paper “An Efficient and Secure Data Hiding Technique–Steganography”. A framework was proposed by them for detecting the LSB steganography by making use of the media files which are digital as the cover objects. They have stated that Steganography can calculate a strong estimate of length of message which is hidden in the sample of LSBs for a larger group of the digital media file contents like image and audio, consists of signal which has correlated samples. “In the traditional steganography techniques principle was either to replace a certain part of the

frequency components of the carrier image, or to replace all the least significant bits of a multi-valued image with the secret data. (Bhagat & Dhembhare, 2015).

Literature Related to the Methodology

Watermarking and Fingerprinting are the other two techniques that seem to be same as Steganography. Watermarking can be used for providing the hidden copyright notices or any other certification licenses. But in the case of Fingerprinting, it uses the content of each copy and makes a unique information to the receiver. There are different types of carriers for steganography techniques a text message, an image file, an executable program file or an audio file. And, there are some steganography conditions for a steganography algorithm to be successful. They are invisibility, robustness against image manipulation, payload capacity, Robustness against statistical attacks, unsuspicious file and independent of file format (Patel & Tahilraman, 2016).

In 2016, Mr. Shruhad Kumar J. Patel1, Nikunj V. Tahilraman have been performed a work “Information Hiding Techniques: Watermarking, Steganography.” They have given an overview of steganography and watermarking techniques. Classification of various techniques for steganography is presented in this paper (Patel & Tahilraman, 2016).

In 2016, Palwinder Singh. have been proposed “A Comparative Study of Audio Steganography Techniques” for providing effective protection to the information over network. The popularity and availability of audio digital signals made researchers to choose them as a preferred choice to convey secret message. So, this paper has a comparative study of various audio steganography approaches and their techniques (Sing, 2016).

Summary

This chapter has given the idea of how the steganography is done and various techniques that are used to perform. This chapter also included the literature review which is related to the problem statement that how the data hiding is achieved and also review on the methodology used by steganography.

Chapter III: Methodology

Introduction

There are various techniques of steganography and can be used based on the purpose and need. It is so important to use correct technique based on the requirement otherwise the purpose is not fulfilled and result in failure of the process. The main purpose of the steganography is hiding the confidential information and it should be handled in an efficient manner. This chapter will discuss the various techniques of the steganography and its application.

Steganography is classified into three categories.

1) Pure Steganography is based on the assumption that no other party is aware of the communication i.e. where there is no stego key.
2) Secret key Steganography is most susceptible to interception i.e. where the stego key is exchanged prior to the communication.
3) Public key Steganography where a private key and public key is used for secure communication.

The techniques which are used for hiding information are receiving much attention today. This is mainly because of the fear in using the encryption services are becoming illegal. Also, the copyright owners who are willing for tracking the confidential property copyright that are against the unauthorized access and will be used in the digital materials such as book, film, music and software that are using the digital watermarks.

There are various ways for hiding the information inside the digital images. Below are some of the approaches:

- Least significant bit insertion
- Masking and filtering
- Algorithms and transformations

Each of these techniques have the different degrees of success

Least significant bit insertion. Bit of the secret message is obtained by the Least Significant Bit of some or all of the bytes inside an image. Digital images are mainly two types (i) 8-bit images and (ii) 24-bit images. So, based on the bits of the images which are used as the cover image that is used for hiding the secret information, the amount of secret data we can store is determined. The higher the bit size of the images the greater number of bits in the pixel of the image are stored. Basically, the bit numbers of the image will indicate the number of bits used for representing the color. Generally, the 8 bits will represent the 256 distinct colors whereas 24 bits will represent the 16.7 million distinct colors. We can insert three bits of secret information in each pixel in 24-bit images, one in each LSB position of the three eight-bit values. "Decreasing or increasing the value by changing the LSB does not change the appearance of the image; much so the resultant stego image looks almost same as the cover image. In 8-bit images, one bit of information can be hidden" (Chan, 2002).

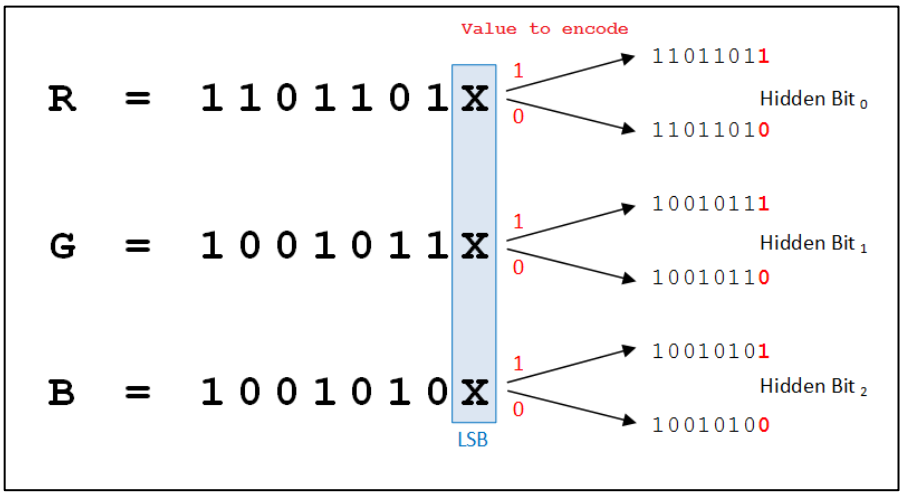

Figure 6: LSB process (KITPLOIT, 2017).

Masking and filtering. Masking and filtering are the two techniques that work with the analysis of the image and usually applied on 24-bits or gray scale images. This technique opposes to hide information inside of the data but actually extends an image by masking the secret message over the original data. Masking, Filtering and paper watermarks are similar in hiding information by marking an image. Digital watermarks include information such as ownership, license, or copyright. While in case of traditional steganography which conceals information, the use of watermarks will extend he information as it becomes as an attribute for the cover image (Provos & Honeyman, 2003).

Masking technique is more suitable than LSB with lossy JPEG images and adds redundancy to the hidden information. It might be helpful in protecting against some of the image processing like rotating and cropping. Masking techniques will be hiding the information in a way where the hidden message will be more integral for the cover image than to simply hiding the data within the "noise" level.

Algorithms and transformations. This steganography technique is used for hiding the data by using the mathematical functions which are in the compression algorithms. The basic idea is hiding the bits of the data in least significant coefficients.

The main advantage of the JPEG images comparative to other formats is the lossy compression methods of it. By using them we can store the images of high quality inside the small files or data. The compressed data will be stored as an integer(s), but the quantization process calculations will require the floating-point calculations which will be rounded. Errors are introduced by the rounding will define the lossy characteristic of a JPEG compression method. The discrete cosine transform (DCT) technique is used by the JPEG images for the achievement of the compression of an image.

The DCT is "a technique for expressing a waveform as a weighted sum of cosines". In a JPEG file, the image is made up of DCT coefficient. When a file is steganographically embedded into a JPEG image, the relation of these coefficients is altered. Instead of actual bits in the image being changed as in LSB steganography, it is the relation of the coefficients to one another that is altered (Raja & Vanugopal, 2004).

Design of the Study

This study involves analyses of various techniques of steganography and its scope of usage. A qualitative and quantitative approach is best suitable for my research because it will be helpful which technique is robust and which method is helpful for the specific requirement. I will also be implementing one of the techniques of steganography by using a tool. Initially, we will take an image that is to be used to hide the confidential data and then by using a tool we will be

inserting the confidential data into the image then compare the two images. Also, the confidential information is extracted from the stego-image by using the same tool.

Data Collection

The information is collected from the steganography related journals and articles, detailed information provided on the web regarding the various techniques of steganography and includes the study of previous research papers to know the opinions which will be used for my comparative analysis. For presenting the steganography method involves a laptop or PC, steganography tool, confidential data that is to be hidden and an image or any file in which the confidential data is to be hidden.

Data Analysis

Cryptography. Data used in cryptography refers to the stored digital information. The word security refers to protecting the assets. Data security means preventing the unauthorized access into any websites, computers and personal or organizational databases by applying some privacy measures. Cryptography is one of the data security technique which is used for the data security. Cryptography will be mainly used for the data protection. It is very helpful to the users to send the information in a safe and secure way. It helps in the process of authentication before accessing to the files or the data. Cryptography technique is having a multiple way to make the exchange of the important information in a confidential and secret way. There are so many various cryptographic techniques which are available currently, out of all those AES (Advanced Encryption Standard) is considered as powerful technique. In the present world the expectations out of the Information Security is to have confidentiality, nonrepudiation, authenticity and integrity. The confidentiality of the information which is communicated over the internet is the

most crucial worry and issue to the users and organizations respectively. In an organization, there are many internal documents which are so confidential, and leakage of that information may put the entire organization at risk (Kumari, 2017).

Goals of Cryptography. There are many goals that can achieved by the use of cryptography. The below are the goals of cryptography

1. Confidentiality
2. Authentication
3. Data Integrity
4. Non – repudiation
5. Access Control

Confidentiality. It is the main important goal which is to ensure that no third person should understand other than sender and receiver of the information. Since cryptography uses the key to encrypt and decrypt the message, confidentiality is achieved when no other person can access/read the information other that who have the decipher key (Kumari, 2017).

Authentication: It means the process of the verification of receiver's identity to make sure he/her or system is the intended receiver of the information. In short, it is the process of verification of each other (sender and receiver) before the exchange of the information (Kumari, 2017).

Data integrity. It means to make sure the data or information which is exchanged is not modified during the travel of information from sender to receiver via communication channel. Because the information may get changed by the intruders intentionally or by accidentally which results in the delivery of incorrect information to the sender. Data Integrity will confirm that the

data is unchanged since the date it was created or during the transmission. Hashing is the technique which is used to serve this purpose. This hash function will be used by both the sender and receiver. The sender will generate a message digest typically a hash code of the information that is being sent and send along with the message and the receiver will generate the hash code and compare with the message digest that is sent by the sender. If both matches, then the receiver can confirm that the data is not modified during the transmission of the data (Kumari, 2017).

Non-repudiation. It ensures that a sender had actually sent the message and the message was received by the intended party, so that the receiver cannot say that the information was not sent by the sender. For example, if non-repudiation was enabled in a transaction, an order which is electronically placed once then the purchaser cannot disagree the order (Kumari, 2017).

Access control. It is a process which prevents the unauthorized access to the resources which will protect the confidential data from being misused. The purpose of this goal is to make the resources available to the authorized users. It defines the permission levels of the information or the resources. For example, a manager in an organization can see the information related to the employees under him but, the employees cannot see their manager information (Kumari, 2017).

These above goals discussed can be achieved all at a time in single application or may be only one of the goals can be implemented based on the requirement and need of the application.

Cryptography vs. steganography. Even though Steganography and Cryptography are related closely, they are completely different in the approach. They can be referred as “cousins”. Cryptography is a process of encrypting the data into a form that if an intruder or an unauthorized person access it or hacks it, the data will make no sense for them. By using the

cryptography, a third person cannot able to understand or read the information since the data will be encrypted. However, it will create a suspect on the data easily as the data is in encoded form. It will create a curiosity for the intruders to know what was encoded and tries to reveal it. But in case of Steganography the message will be concealed, and it will be undetectable to a third person and hence there will no suspicion that a confidential or a secret data is being transmitted. In Steganography, the process is to hide a confidential or secret message in an "open" information/message. The open message can be any of the files like data streams, text, IP packets, audio, images etc. The message that needs to hidden will be embedded in the certain parts of host or it may cause generation of a new file. For example, if we want to send a text file then by using steganography we will hide that text file in an image and will send the image to the receiver. If an unauthorized person see the content of transmission, he will see as a normal image and will not know a secret file is inside an image (Patel, 2003).

Steganography and network security. As it is hard to defend steganography and very complicated to detect it, how you can defend your network against steganography? It is always better and a good idea to incorporate a process to safeguard our network resistant to the misuse of the steganography. This will include a scenario where the employees of an organization sending the important or confidential information to outside of organization or get some malicious information from an attacker who is unknown using the same internal network. An important thing and a first step for this kind of approach is "know your network." Because if we know about the network completely then we can implement the safeguard measures to defend or identify the use of steganography. One of the methods for detecting the steganography usage is to "look for obvious and repetitive patterns which may point to the identification or signature of

a steganography tool or hidden message" (Patel, 2003). It is easy to detect if it can be seen to a normal human eye.

Steganography is receiving a greater attention increasingly and hence the need of tools for steganalysis also widely spread. There are many tools of steganalysis and there are some basic common guidelines in all tools which involves in the detection of the usage of steganography tools. The basic thing that involves in the detection is to concentrate on the large files, check whether a bitmap image contains a greater number of colors that are duplicate. This would or will indicate that there is some data is embedded in an image. Also, having a look at the size of the files and its properties will reveal lot of information towards identifying the use of steganography. Anything that is found to be unusual will raise an immediate doubt. If you can able to find what was the tool used to hide the message, you can get the same tool and can compare the two files by making a known file with similar properties of the suspected file and hide a sample message in it. In a rare case, if you are having the original file, you can able to do the comparison analysis.

Steganography is mainly used in a place where the governments and/or organizations who will not allow the usage of encrypted communications. To fight against the usage of the steganography tools or software in a workplace or within a network which you are using, it is mandatory for making as part of written policy of information security that any of the employees of an organization are not allowed or should not use any kind of steganographic programs over the organization's network. If you intend to put any restrictions or any kind of exceptions that are also needing to be mentioned in your written policy of information security. The security policy also contain the details addressing emailing or/and receiving of sound files, text files, images

mainly on the systems or desktops which has the sensitive information. You should also demand or specify that the sensitive data should contain the digital watermarks which are the trusted ones. Watermarks are used to improve the security of the files. Also, the use of the chat rooms, group lists as well as forums on your workstations could be handled should be written on the security policy you created (Patel & Tahilraman, 2016). You may need to address the firewall issues. It needs the filters to restrict the access to the pornography in the organization as it is the most popular medium for holding the hidden messages. You also should set up a monitoring procedures and policies for the employee's activities on the web, email and attachments they send in the emails. Sometimes the employee should communicate to the external teams like the application teams and with the vendor. The policies should be good enough to monitor the communication through external links also.

Conclusion for cryptography and steganography. Cryptography and Steganography looks similar in the context that both of the techniques are used for a purpose of the network defense and they are fundamentally different in their goals. Cryptography technique is used for protecting the contents of the data or information. By doing so it achieves confidentiality, but it will not provide the secrecy. At this point, steganography comes into play because Steganography will hide the fact of the existence of the message. So, it will provide the confidentiality along with the secrecy of the message or information. However, Steganography can be used with cryptography to achieve the strong defense of the information. Because of the rapid changes which are happening in world, steganography is being used excessively and parallelly many improvements are seen towards the increase of its efficiency. There is lot of increase in the attention towards the uses of steganography and due to which it can be a security

threat as well. Malicious intent of using the steganography can be harmful to the network security. We should consider proper measures in advance and be prepared with the proper security plan. The internet now-a-days is being a source for the hiding of messages or data that passed over the internet. In this area of secret communications and steganography, the development is predicted to grow continuously in the coming years. Even we can see the governments will be addressing the steganography utilization in its defense against the terrorism also for the growth in forensic science region. Federal intelligence agencies will be enhancing the measures they are using to use the steganography tools. With it, “the ease in use and availability of steganography tools has law enforcement concerned in trafficking of illicit material via web page images, audio and other files transmitted through the internet.” (Patel, 2003). As of now regulations on the use of steganography is not implemented fully by the governments which left the technique open for the use which gives the room for harmful intent.

Steganalysis. There are several forms for the analysis and attacks on the hidden/secret information like disabling, extracting, detecting and destroying the hidden information. An approach of the attack is dependent on the information that is available for the Steganalyst. Steganalyst is a person who is working or attempting towards the detection of information streams that are based on Steganography.

Steganography-only attack	Only the steganography medium is available for analysis.
Known-carrier attack	The carrier that is, the original cover and steganography media are both available for analysis.
Known-message attack	The hidden message is known.
Chosen-steganography attack	The steganography medium and tool (or algorithm) are both known.
Chosen-message attack	A known message and steganography tool (or algorithm) are used to create steganography media for future analysis and comparison. The goal in this attack is to determine corresponding patterns in the steganography medium that may point to the use of specific steganography tools or algorithms.
Known-steganography attack	The carrier and steganography medium, as well as the steganography tool or algorithm, are known.

Figure 7: Different types of steganalysis approaches (Bhattacharyya & Souvik, n.d.).

Image based steganalysis. Steganalysis is science of discovering the hidden message or information. The main purpose of the Steganalysis is breaking the steganography and the goal of it is discovering the stego image which has the secret information hidden in it. All the algorithms related to steganalysis is will be dependent on the algorithms of the steganography putting the statistical differentiation between the stego and cover image. Steganalysis mainly deals with the three key categories (Bhattacharyya et al., 2011).

- Visual attacks
- Statistical attacks
- Structural attacks

Visual attacks. In this type attacks, with the help of a computer or through a keen inspection by naked eye, we can reveal the existence of the hidden message or information which will be helpful for separating the image (stego image) into different bit planes and can be used for in depth analysis (Banerjee & Indradip, 2011).

Statistical attacks. These are most powerful type of attacks as well as successful, as they identify very small changes in the statistical behavior of the images (Banerjee & Indradip, 2011). These attacks are further classified into (a) Passive attack and (b) Active attack.

Passive attacks will involve identifying the existence or non-existence of an embedded algorithm or covert message used, etc.

Active attacks are used for investigating the embedded data length or hidden data location or a secret key that was used in embedding.

Structural attacks. Generally, when a data that is to be hidden in an image or some carrier data files, the format of the carrier data will be changed. So, identifying these types of structural changes will help us in finding the existence of an image (Banerjee & Indradip, 2011).

Types of image based steganalysis. Steganalysis can be considered as two class pattern classification problem, aims for determining that whether the medium used for testing is a stego medium or a cover medium.

*Targeted steganalysis***.** It is the technique which works on the specific type of a stego-system and at times limited to only image format. After studying and performing some analysis on the embedding algorithm, we can find the statistics of an image that will change after embedding. The results which obtained from this targeted steganalysis will be very accurate, but these techniques tend to be inflexible because of the reason that in most of the cases there will be no path for extending them to the other embedding algorithms. Also, when the targeted steganalysis is found to be successful, it means it is having the higher probability than a random guessing, it also helps in the steganographic techniques for expanding and becoming more secure (Banerjee & Indradip, 2011).

*Blind steganalysis***.** It is the technique that is designed for working on all the types of image formats and embedding techniques. In short, this algorithm learns about the difference between the statistical properties of the stego and pure images and tells about the differences between them. The process of learning is done by the training of machine or system on large image database. These are usually not that accurate as the targeted, but it is a lot more expandable (Banerjee & Indradip, 2011).

Semi-blind steganalysis. It works on the specific range of the different stego systems. The range of them (stego-systems) can be dependent on domain they hide on, that is transform or spatial (Banerjee & Indradip, 2011).

***Specific approaches of image based steganalysis*.** Specific steganalytic method will take advantage of insecure part of the steganographic algorithm (Bhattacharyya et al., 2011).

Attacking LSB steganography. It is one of the most important among the spatial steganographic techniques. Accordingly, during the initial stages of development of the steganalysis, most of the work has done on the process of steganalyzing LSB steganography. Many of the steganalysis methods related to the LSB steganography was proved to be most successful, like RS analysis, weighted stego analysis, chi-square attack etc (Banerjee & Indradip, 2011).

Attacking LSB matching steganography. It should be noted that equal trend of frequency of the occurrence of Pairs of Values (PoVs) no longer present in the LSB matching steganography. So, many steganalysis methods related to the LSB steganography became invalid. LSB matching also generally known as ±k steganography was may be designed in a

context of the additive noise which is independent of cover image. The process is described below.

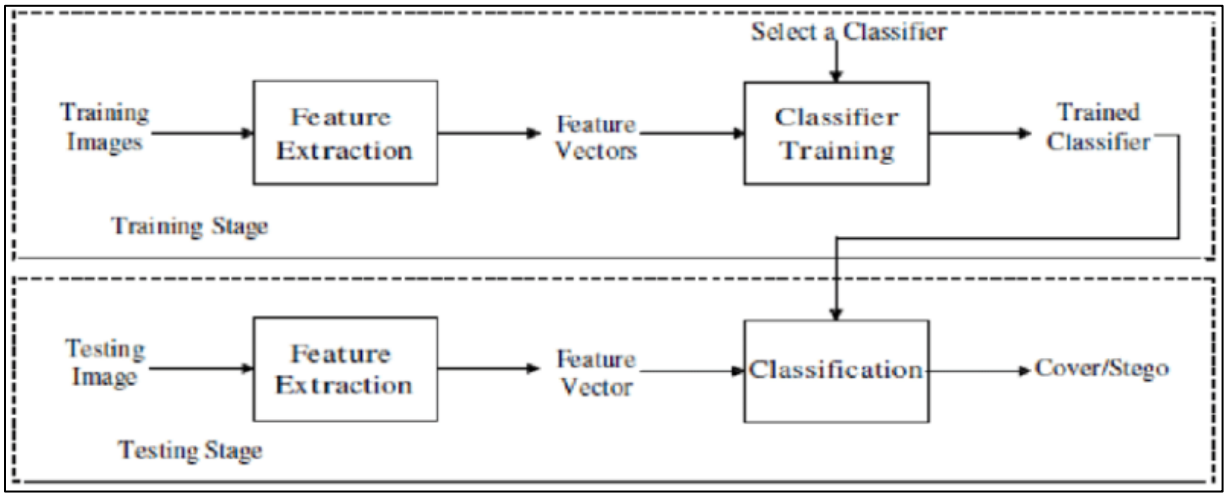

Figure 8: The process of a universal steganalytic method (Bhattacharyya & Souvik, n.d.).

In the above process, the step of feature extraction is used in the training as well testing stage. The main purpose of this is mapping an input mage from high dimensional image space to the low dimensional feature space. The main aim of the training stage is obtaining a trained classifier. There are many classifiers that are effective can be selected, such as neural network (NN), Fisher linear discriminant, Support vector machine. Classifier forms the decision boundaries to separate feature space into negative and positive regions by using the feature vectors that are extracted from training images. In testing stage, by using the trained classifier which has decision boundaries, in the feature space an image that is in question is divided accordingly to its feature vector's domination. If the location of the feature vector is in the region where the label of the classifier is positive, then the testing image is considered as positive class which is a stego image. If not, the testing image is considered as negative class which is cover image (Banerjee & Indradip, 2011).

Below are some of the universal steganalysis features

***Image quality feature*.** By using the steganographic techniques to hide an image somehow may be more or even less, it will cause some degradation to an image. For gauging the amount of distortion, image quality measures (IQMs) serves as the quantitative metrics which are based on the image features. “The statistical evidence left by steganography may be captured by a group of IQMs and then exploited for detection” (Banerjee & Indradip, 2011)**.**

Calibration based feature. The feature-based classification is applied together with context of calibration to implement a blind detector which is specific to JPEG images. The word calibration here means, using stego image as the side information some of the cover image parameters may be recovered approximately. By doing so, the process of calibration will increase the features sensitivity towards the embedding changes while the image to image variations are suppressed (Banerjee & Indradip, 2011).

Moment based feature. The impact on the cover image due to the steganography is considered as stego-noise. As the noise has be added, there will be some statistical changes to the image. It is effective and important to observe this kind of changes in the wavelet domain. Lyu and Farid has used an assumption that a PDF of the sub band coefficients of wavelet to that of prediction error of sub band coefficients will be changed post embedding the data. (Emam & Marwa, 2015) said “a 3-level wavelet decomposition, the first four PDF moments, i.e., mean, variance, skewness, and kurtosis, of the subband coefficients at each high-pass orientation (horizontal, vertical and diagonal direction) of each level are taken into consideration as one set of features” (Banerjee & Indradip, 2011).

Correlation based feature. By hiding the data in an image, there may be a distortion of local correlation of an image. Here, for a spatial image correlation means the inter-pixel dependency and for a JPEG image correlation means inter-block or intra-block DCT coefficient dependency (Chan, 2002) modeled the inter-pixel dependency by Markov chain and depicted it by a gray-level co-occurrence matrix (GLCM) in practice.

Text based steganalysis. Text media usage as cover channel for the secret communication became brought to more attention. So, by the increase of the attention created the increase of the concerns on the text steganalysis. Compared to the other type of cover media like audio, image and video, currently it is very harder to identify the hidden messages in the text. Generally, the text steganalysis make us of the fact that any embedding information will changes the statistical properties of the stego texts to some extent. So, it is very important to notice the modifications or changes in the stego texts (Banerjee & Indradip, 2011).

Based on the earlier work, text steganalysis can be classified in three categories: linguistics, format-based and invisible character based. Linguistics is different from the other two categories in which it will attempt to identify the secret messages inside the natural language texts. In the linguistic steganography, for concealing the information the semantic, lexical and syntactic properties of the texts are modified by making sure their meaning is not changed. Because of the polysemia of semantics and diversity of the syntax, it is tough to identify the modifications done in stego texts. Till now, there are many linguistic methods are introduced. For all those, designed the special features for extending the syntactical or semantic alterations of stego texts. If the size of a text or the content in the text file is large enough, then the differences in between Stego texts and Natural texts serves as the evidence and hence the performance of the

detections is normal. But, if he size of texts are small, the rate of detection will be decreased dramatically.

Audio steganalysis algorithms. Audio steganalysis is difficult because of the existence of the advanced schemes in audio steganography and the nature of the audio signals are high capacity streams of data demand the need for challenging the statistical analysis scientifically (Banerjee & Indradip, 2011).

***Phase and echo steganalysis*.** Zeng has introduced the steganalysis algorithms for detecting the echo steganography based on peak frequency statistical moments. “The phase steganalysis algorithm explores the fact that phase coding corrupts the extrinsic continuities of unwrapped phase in each audio segment, causing changes in the phase difference” **.** In every audio segment there will be a phase difference and the statistical analysis on it will be used for monitoring the alterations and train classifiers for differentiating the hided audio signal from clean audio signal. In the algorithm of echo steganalysis, by using the short window extracting it examines the peak frequency. After that it calculates the peak frequency’s eight high order center moments as a feature vectors which are fed to support vector machine and that is used as the classifier for differentiating the audio signals without and with data (Banerjee & Indradip, 2011).

***Universal steganalysis based on recorded speech*.** Johnson introduced a comprehensive universal steganalysis algorithm which serve as the base for the study on statistical regularities of the recorded speech. The statistical model of them will decay an audio signal which is the recorded speech by the use of basic functions that are localized in both the frequency and time domains in STFT.

Video steganalysis methodology. It focuses on the data which was hidden in the frames of a video.

***Video steganalysis exploring the temporal correlation between frames*.** One of the techniques for video steganalysis uses the repeated information that is present in temporal domain as deterrent against the secret messaged that are embedded by using the spread spectrum steganography technique. The earlier study based on the approaches of linear collusion, it is successful in finding the watermarks that are hidden having the low energy and with good precision. The results of the simulation will prove that superiority of methods of temporal based over the pure spatial methods in the detection of secret message (Banerjee & Indradip, 2011).

***Video steganalysis based on Asymptotic Relative Efficiency (ARE)*.** There was an algorithm of video steganalysis which incorporates the asymptotic relative efficiency-based detection. This algorithm will be suited for the applications in which a subset of video frames are alone watermarked with the confidential message and should not be all of them. The stego video is believed to have a sequence of the correlated image frames. The phases of signal processing underline the fact of the existence of embedded information in sequence of frames by use of motion estimation scheme. The detection phase will be based on the ARE, where both the watermarked confidential message and cover video are examining to be the random variables (Banerjee & Indradip, 2011).

Summary

This chapter includes introduction of the methodology which is used in Steganography and give the idea how this study is carried out in design of study. It also says about the sources

used for the research, how the data is collected and minimum requirements for the implementation.

Chapter IV: Analysis of Results

Introduction

In this paper, we use three images and a text file. Out of three images, two of them are used as test images. The third image and text file are used as concealed information. These are used to test the different steganographic tools. The images are taken based on their type and properties to test the software in superior manner. A passphrase is used in the process of hiding the secret data wherever it is necessary and applicable. The passphrase will be used for encryption, where the confidential data is encrypted prior to the hiding of the message or data inside the carrier images.

Further, the output images are analyzed in terms of 2 image quality metrics which are SSIM and PSNR. PSNR is the ratio between maximum amounts of power of signal to distorting noise power. The value of PSNR ranging between 0 and 100. The more the PSNR value, the better the quality of the image since the amount of error is low. SSIM is “Structural Similarity Index” and its value will be ranging between 0 and 1. 1 means that both images are compared, and they are exactly same. SSIM is the better metric compared to PSNR as it measures similarity between the images in the same way as human eye do and hence will be more accurate. In this paper, I used Imatest on Matlab for finding the SSIM and PSNR values between the two images.

PSNR is abbreviated as “Peak Signal-to-Noise Ration”. It is defined as the ratio between the maximum power of signal and power of the corrupting noise which affects the representation. It is measured in decibels and the range is 0 to 100. This ratio is used as the quality measurement between the final output image and the original image (MathWorks documentation, n.d.).

SSIM is abbreviated as “Structural Similarity”. It is used to measure the similarity between the original image and the final output image. It requires the original and compressed or final output image which should be same image. It tells about whether the two images are similar or not and it cannot predict which image is better among two (Imatest Documentation, n.d.).

Data Presentation

Below are the sample test images that are used in the experiment described in this paper. A text file is also used as part of the experiment.

Figure 9: Test image 1 and image 2.

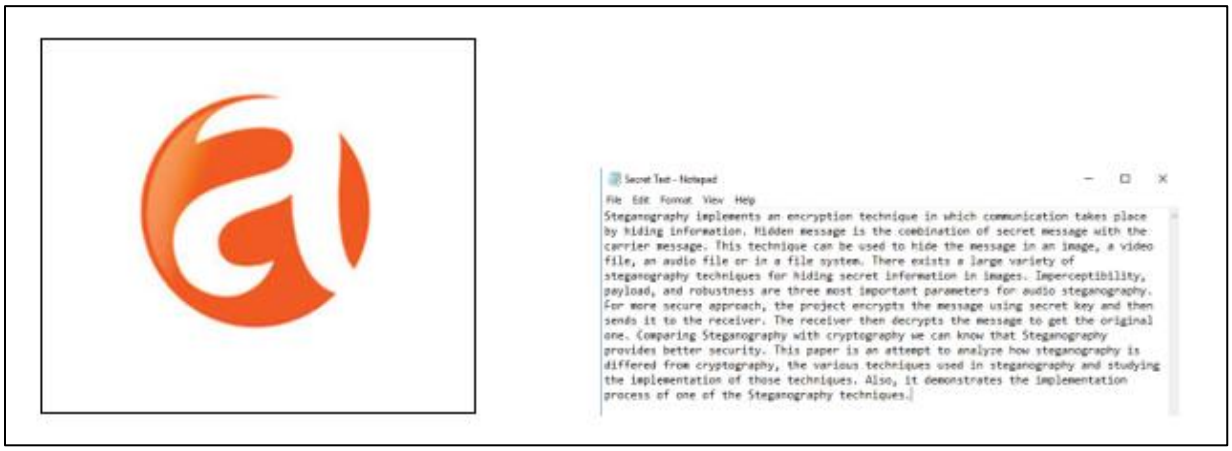

Figure 10: Secret image and text.

Table 1

Image Analysis

Image/File Name	**Used For**	**Image Size**	**Image Dimensions**	**Image Format**
Image1	Carrier	385KB	768x512	BMP
Image2	Carrier	128KB	1200x822	JPG
Secret Image	Hiding Data	8KB	251x201	JPG
Secret Text	Hiding Data	4KB	-	TXT

Steganographic software tools. The following are the different tools which are used in the experiment described in this paper. All the test images are used on all the tools and hided the data in the images and further analyzed on the quality of an image post completion of the hiding process by all the tools.

Hide N send. It is a portable software application which is used for hiding the files inside the JPG images. The secret message which is hided can be of any type such as xlxs, txt or docx etc. There are options for choosing the settings for concealment, encryption and hash algorithms. A passphrase needs to be applied when we click on ‘HIDE’ button. The hidden file from the JPG image can be extracted by using the ‘EXTRACT’ button and should provide the destination folder path where the image to be extracted and saved.

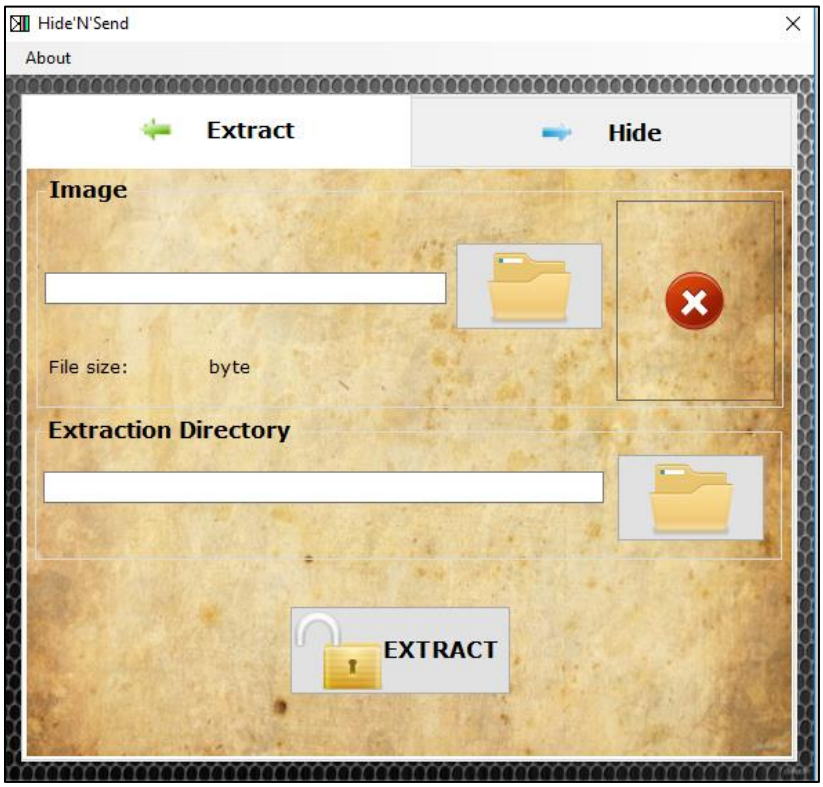

Figure 11: Hide N Send software.

CryptaPix. It CryptaPix will be used for steganography, encryption and management of data and images. It supports large number of image file types which include BMP, PNG, GIF and JPG. It also uses AES algorithm of encryption for encrypting the images before hiding them.

Apart from the above mentioned this software also provides many functions which are related to image editing like cropping, resizing, removing red eye and rotating from images.

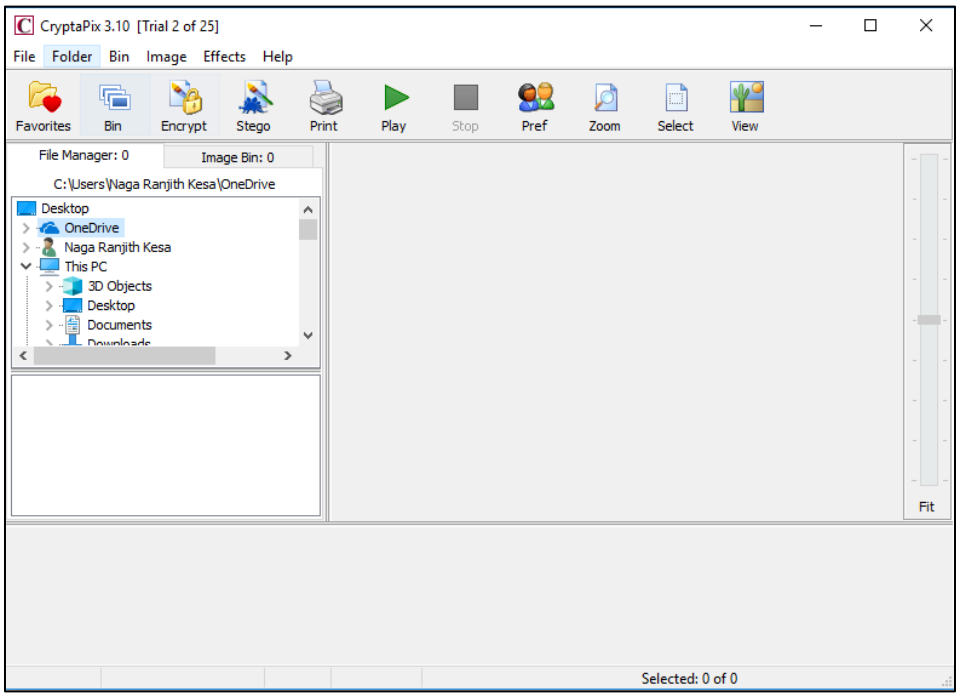

Figure 12: CryptaPix software.

QuickStego. QuickStego will be used to hide the text in images and hidden text messages can be read by only QuickStego users. It supports BMP, PNG, GIF and JPG image file types. It is compatible with windows. The text which is to be hidden can be added by typing or by loading it from a TXT file. After hiding the text in the image, the final output image will be saved in BMP format. The user interface of the software is easy to understand.

Figure 13: QuickStego software.

VSL. VSL will be used for hiding the images in any format. This software will make use of F5, Karhunen-Loeve Transform technique and LSB algorithms for hiding the data. It has many distortion filter and options for analyzing and decoding the images.

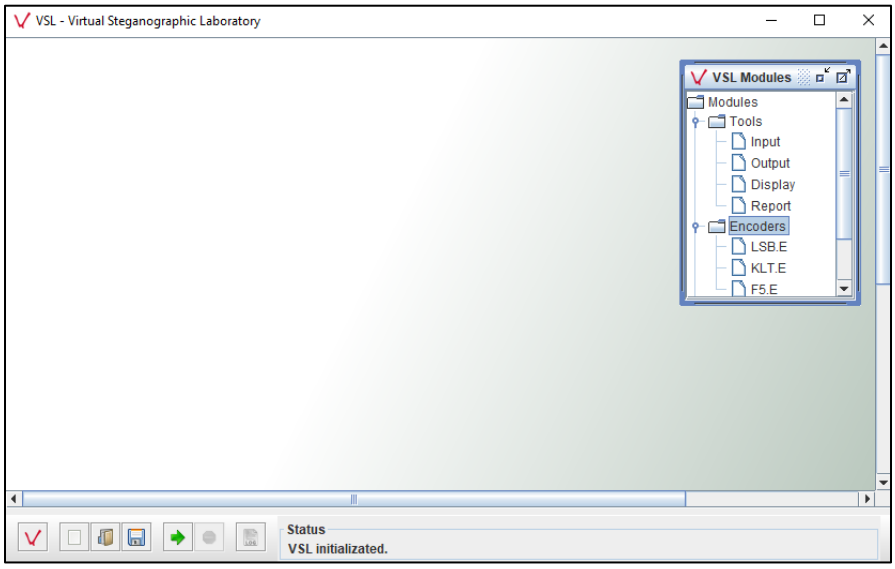

Figure 14: VSL software,

Steg tool. Steg will be hiding the important data inside BMP, TIF, JPG or PNG images. It allows you to hide a text message as well inside the specified image. After hiding the data, the final image consists of hidden data can be saved in PNG or TIF format. It is cross platform and portable program. The graphical user interface of Steg is easy.

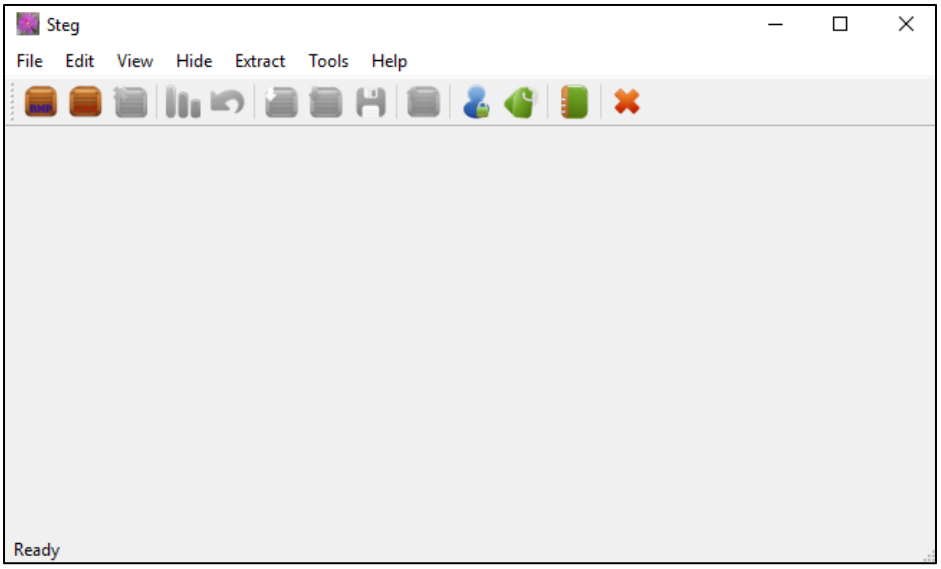

Figure 15: Steg software.

Table 2

Comparison of Software Tools of Steganography

Steganographic Software	**Software Size**	**Software Description**	**Software Creator**	**Software Resources**
Steg 1.0.0.2	7.1MB	Steg will be hiding the important data inside BMP, TIF, JPG or PNG images. It allows you to hide a text message as well inside the specified image. After hiding the data, the final image consists of hidden data can be saved in PNG or TIF format. It is cross platform and portable program. The graphical user interface of Steg is easy.	Fabio	https://www.softpedia.com/get/Security/Encrypting/Steg.shtml
CryptaPix 3.10	2.3MB	CryptaPix will be used for steganography, encryption and management of data and images. It supports large number of image file types which include BMP, PNG, GIF and JPG. It also uses AES algorithm of encryption for encrypting the images before hiding them. Apart from the above mentioned this software also provides many functions which are related to image editing like cropping, resizing, removing red eye and rotating from images.	Briggs Softworks	https://www.softpedia.com/get/Security/Encrypting/CryptaPix.shtml
VSL	1.48MB	VSL will be used for hiding the images in any format. This software will make use of F5, Karhunen-Loeve Transform technique and LSB algorithms for hiding the data. It has many distortion filter and options for analyzing and decoding the images.	Michal Wegrzyn	https://sourceforge.net/projects/vsl/
Quick Stego	1.7MB	QuickStego will be used to hide the text in images and hidden text messages can be read by only QuickStego users. It supports BMP, PNG, GIF and JPG image file types.	Cybernescence	https://www.softpedia.com/get/Security/Encrypting/QuickStego.shtml
Hide N Send	536KB	Hide N Send is the portable application will be used for hiding the secret files behind the JPG images. The hiding file can be of DOCX, TXT, XLSX types.	MRP Lab	https://www.softpedia.com/get/Security/Encrypting/Hide-N-Send.shtml

Table 3

Features of Software Tools of Steganography

Steganographic Software	**Carrier Image Formats**	**Encryption Support**	**Steganographic Algorithm**
Steg 1.0.0.2	JPEG (JPG), TIFF, PNG, BMP	NA	LSB
CryptaPix 3.10	BAY, BMP, CRW, CR2, CUR, DCR, DCX, DIB, EMF, FAX, GIF, G3F, G3N, ICB, ICO, JIF, JPC, JPE, JPG, JP2, J2C, J2K, MRW, NEF, ORF, PBM, PCX, PEF, PGM, PIX, PNG, PPM, PSD, PXM, RAF, RAW, RLE, SRF, TGA, TIF, VDA, VST, WBMP, WMF, XIF, X3F.	256-bit AES algorithm	LSB
VSL	Any	NA	LSB, Karhunen-Loeve Transform technique, F5Algorithm.
Quick Stego	BMP, JPG, GIF	NA	LSB
Hide N Send	JPG	AES, RC2, RC4	M-F5, M-LSB, F5, LSB

Implementation. The below are the steps performed for obtaining the results by using the sample images and the steganographic software tools. Below are the scenarios tested as part of the implementation by using each software

Secret image concealed in Test image1(BMP)

Secret image concealed in Test image2(JPG)

Secret text concealed in Test image1(BMP)

Secret text concealed in Test image2(JPG)

Hiding of data using Steg tool. Click on file select “Open generic image” for the images other than JPEG type and select “Open JPEG image” option for the JPEG images. The image that is selected will display in both Original media and Modified media columns. Then click on “Hide” option in the menu bar and select the “Hide Data” option. After selecting a dialog box will be open and then select the image or file that needs to be embedded and then click ‘Open’.

The data will be inserted, and a dialog will be pop up stating “Data successfully hidden” as shown below. Click on ‘OK’.

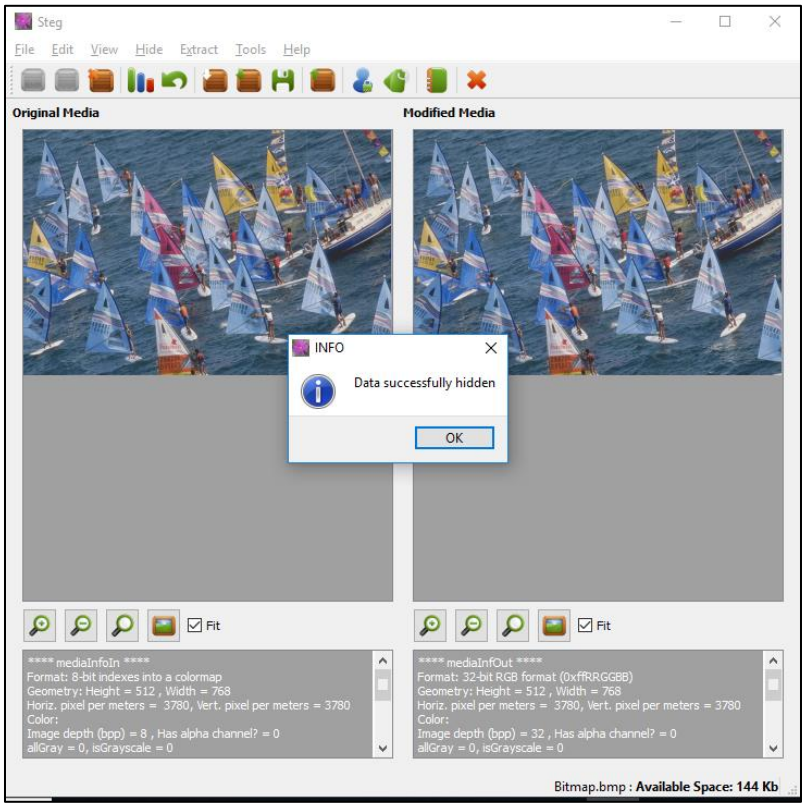

Figure 16: Hiding data using Steg

After successful hiding of data then click on ‘Save’ button as highlighted in below screenshot. A dialog box will be open and choose a destination folder to save the final output image with hidden data.

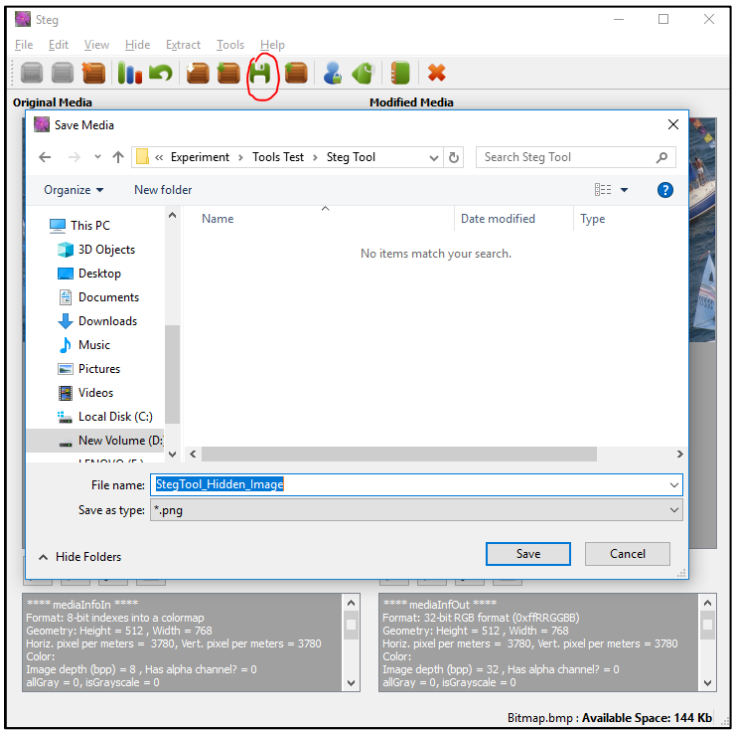

Figure 17: Saving the final output image with hidden data in Steg software.

Extraction of data using Steg tool. After obtaining the final output image, it will be sent to the intended receiver of the secret message. At receiver end the recipient need to use the same Steg tool for extracting the secret message hidden in the image sent. Click on file select “Open generic image” for the images other than JPEG type and select “Open JPEG image” option for the JPEG images. A dialog box is opened as shown below

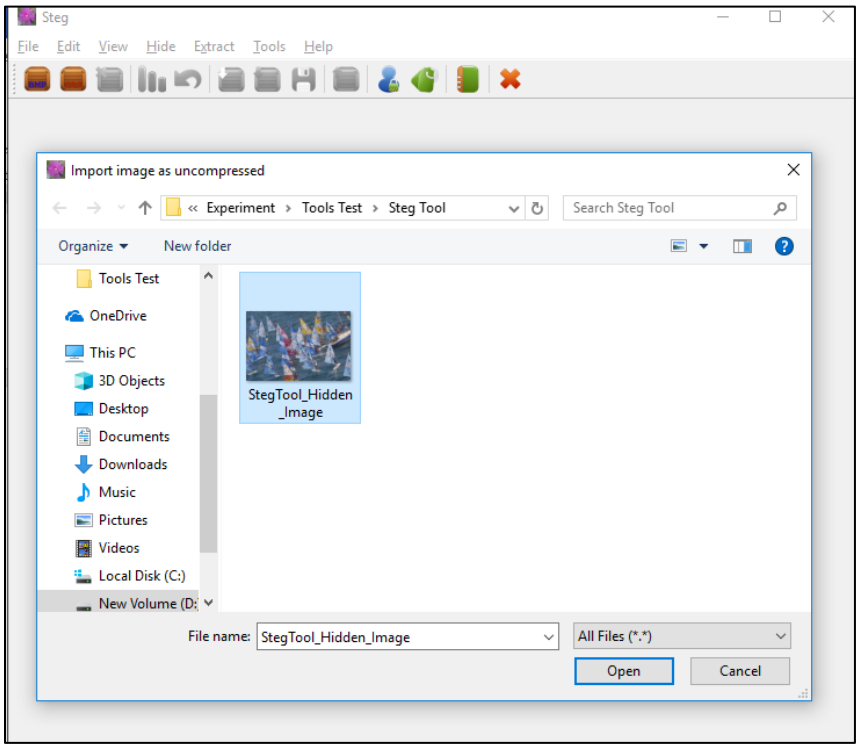

Figure 18: Opening the image with hidden data to extract in Steg software.

The image that is selected will display in both Original media and Modified media columns. Click on “Extract” in the menu bar and then select ‘Extract data’. A dialog box will be displayed as shown in the below screenshot to select the directory in which the secret message to be saved. We can select only the directories in which it needs to be saved. After choosing the directory click on ‘Choose’ then the data will be extracted to the selected destination and a popup will be displayed with a message as shown in Figure 20.

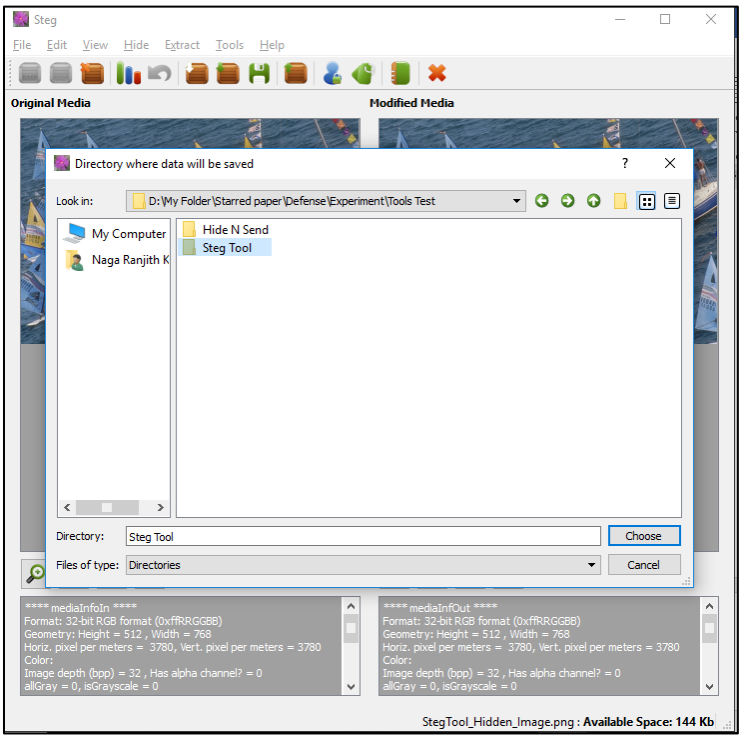

Figure 19: Saving the extracted secret message from stego image in Steg software.

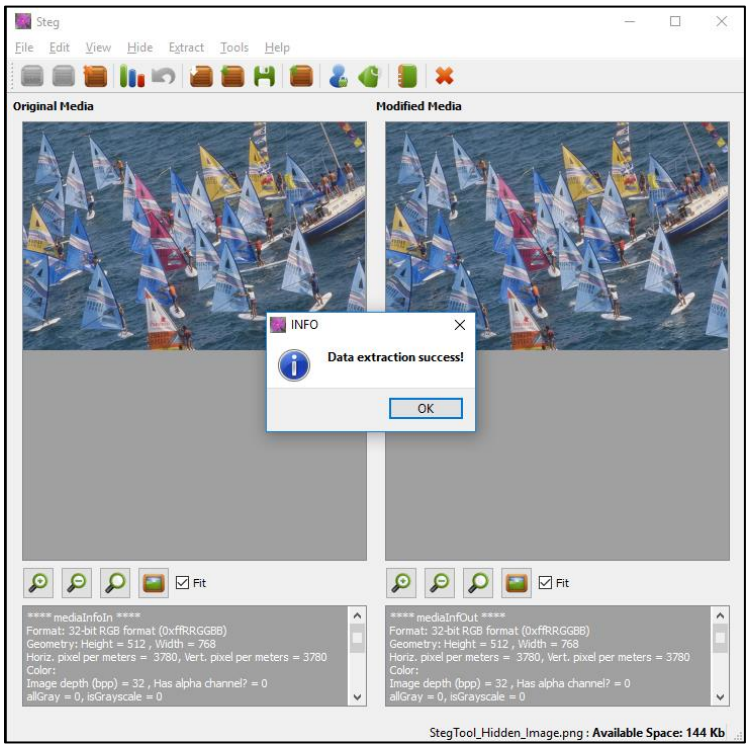

Figure 20: Final screen after the extraction of data in Steg software.

Below is the screenshot of the files after extracting the secret message. There will be 4 files generated out of the extraction process which are highlighted as below.

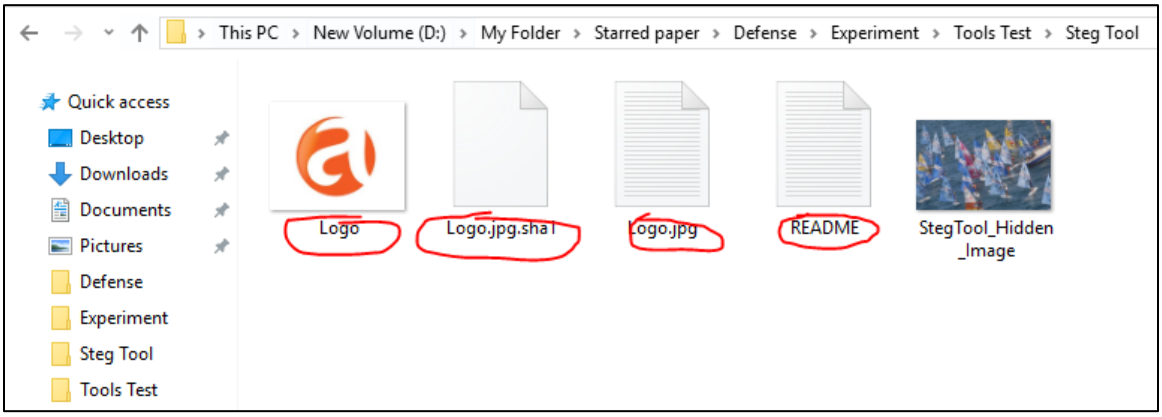

Figure 21: Files extracted from stego image using Steg software.

The “README” file has the information about the software used, .jpg file is an empty file, “Logo” is the actual secret message and “logo.jpg. sha1” file consists of the hash value that

is generated by using SHA1 hashing algorithm which will be used to check the integrity of the message.

Hiding of data using Hide N Send tool. There are two tabs in the home screen of the tool, one is “Extract” and other one is “Hide” as shown in the below screenshot. For hiding the data, you need to select the “Hide” tab. There will be 3 sections which are Image, Concealed file and Settings. Click on the folder icon in the ‘Image’ section, a dialog box will be opened then select the cover image that should be used in the hiding process. It will be displaying the details of the image selected. Click on the folder icon in the ‘Concealed’ section, a dialog box will be opened then select the secret image that should be hided in the cover image. It will be displaying the file size of the image selected. In the ‘Settings’ section you need to choose concealment algorithm, hash algorithm and Encryption algorithm. Out of the available options I selected the options as shown in the below screenshot.

Figure 22: Selecting the required files and settings to hide data by using Hide N Send software.

After choosing all the files and settings that are required, click on “Hide” button. Then a popup will be displayed with to enter the password for protection of the data. This password will be used to extract the secret message. Without password you cannot extract the secret message. After providing the password click on ‘OK’.

Figure 23: Screenshot of providing password in Hide N Send software.

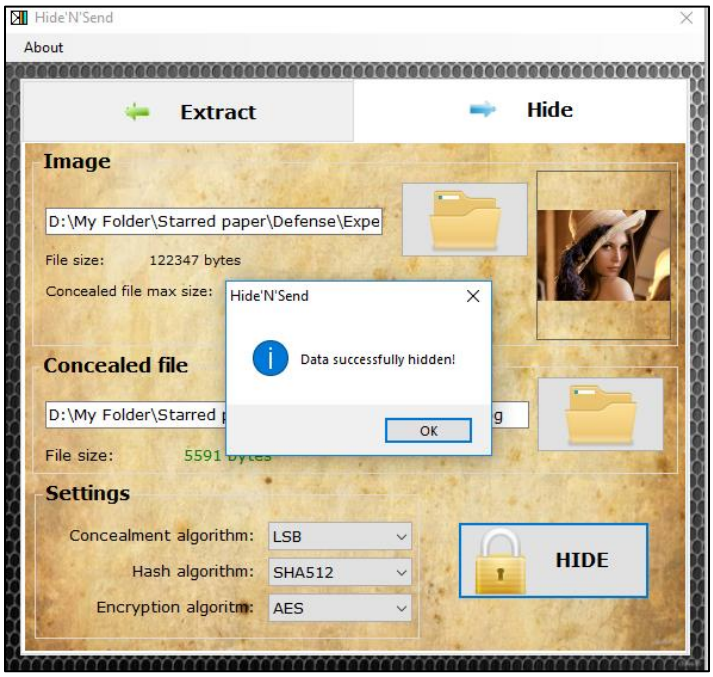

Figure 24: Confirmation of the hiding of data using Hide N Send software.

Extraction of secret message using Hide N Send tool. After obtaining the final output image, it will be sent to the intended receiver of the secret message. At receiver end the recipient need to use the same Hide N Send tool for extracting the secret message hidden in the image sent. In the tool, you need to click on “Extract” tab. There will be two sections which are Image and Extraction Directory as shown in the below screenshot.

Figure 25: Extraction process of Hide N Send software.

The stego image needs to be selected by clicking the folder icon in the Image section. After selecting the image, it will be displayed in the right box as show in the above screen shot. Select the directory path where the extracted secret message needs to be saved in the Extraction Directory then click on “Extract” button. It will prompt for a password and the password that is given during the hiding process should be given here.

If the password is different from what was given initially when hiding, then an error message will be popped up as shown in the below screenshot.

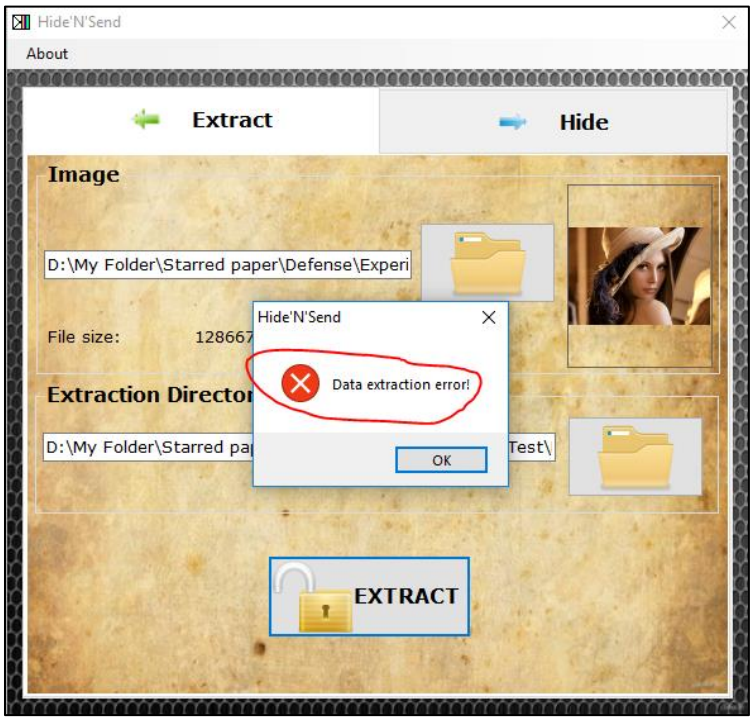

Figure 26 : Error message during the extraction process in Hide N Send software.

If the right password is given, then a success message will be displayed as shown in the below screenshot.

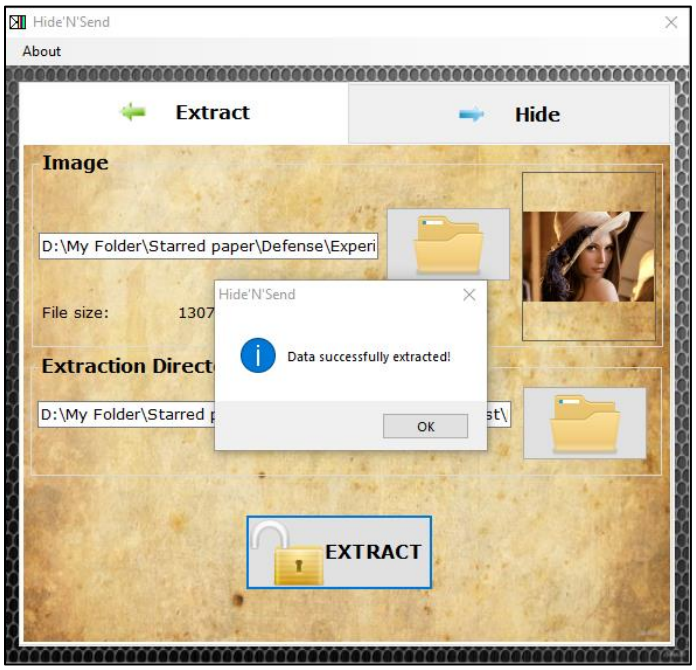

Figure 27: Success message of data extraction in Hide N Send software.

Below highlighted is the file extracted in the chosen destination directory.

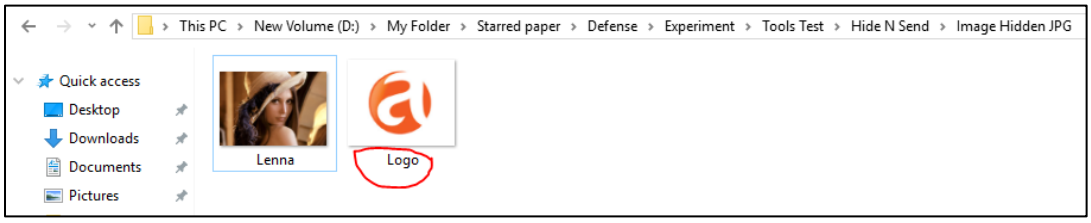

Figure 28: Extracted file in the destination folder by using Hide N Send software.

Hiding of data using QuickStego tool. In this tool, the secret message can be only a text. We can't hide an image inside an image by using this tool. Firstly, click on "Open Image" button for selecting an image file. A dialog box will be open then select the cover image. As soon as the image is selected a message will be displayed as shown in the below screenshot stating that this image does not contain any secret message.

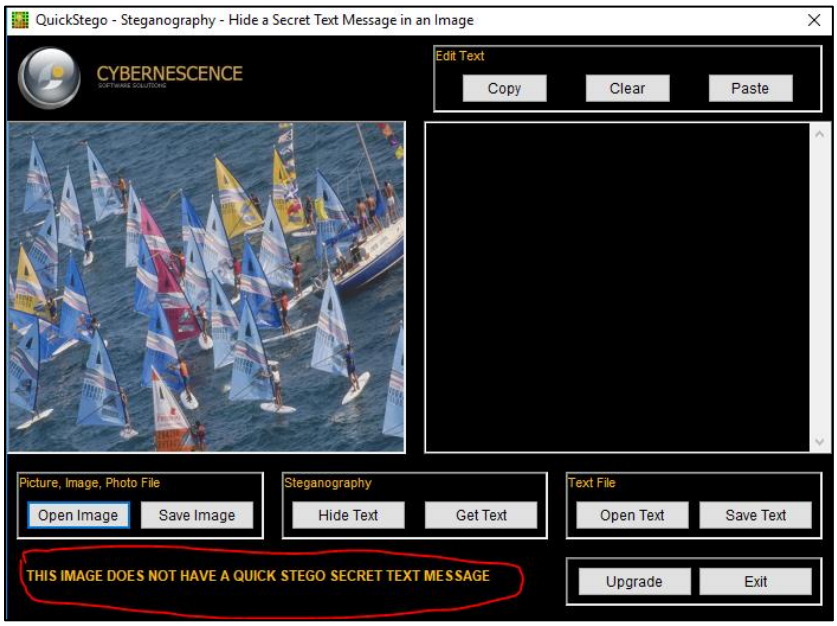

Figure 29: Selecting the cover image in QuickStego software.

A text message can be typed in the empty box beside the selected image or can be selected from a text file. If we click on “Open text” button, then a dialog box will open as shown in the below screenshot.

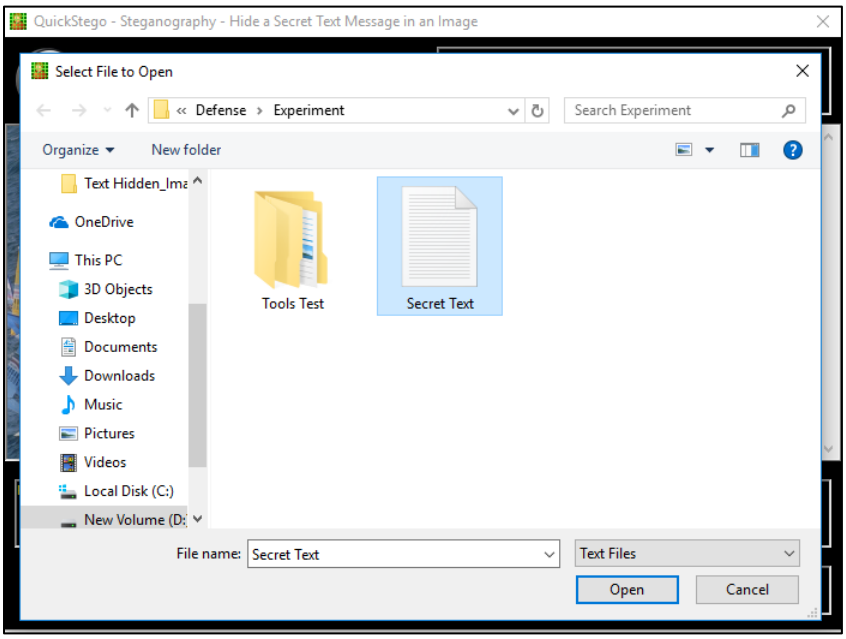

Figure 30: Choosing the file that has secret message in QuickStego software.

Click on “Open” and the content of the text file will be displayed in the box next to the selected image as shown in the below screenshot.

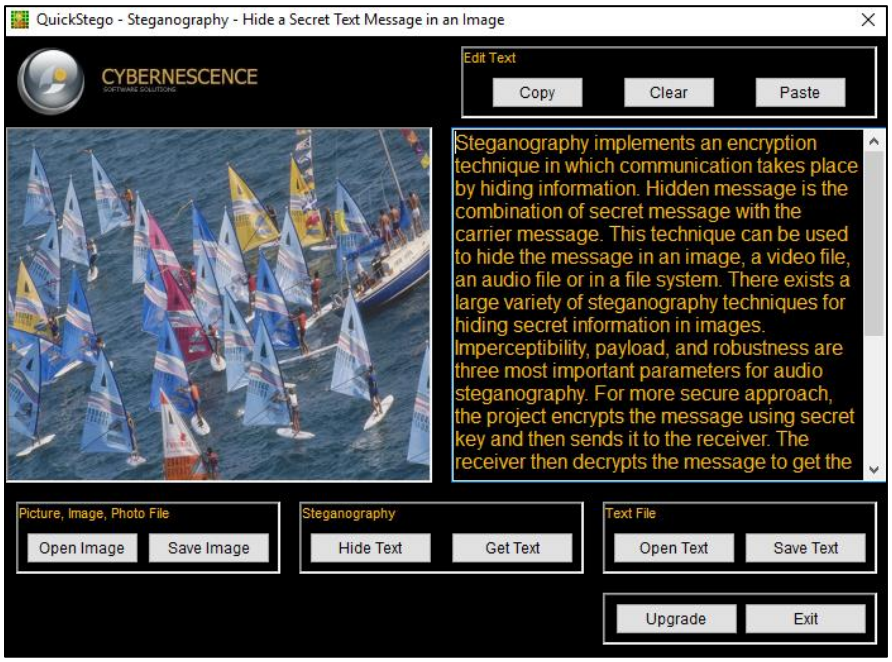

Figure 31: After selecting the secret text file in QuickStego software.

Now, click on “Hide Text” to start the hiding process. After completion of the hiding process a message will be displayed stating that the message is hidden in the image as shown in the below screenshot. Click on “Save Image” button for saving the final output image which have the secret message in it. The dialog box is displayed which will allow to choose the destination folder and the output image will be saved in BMP format.

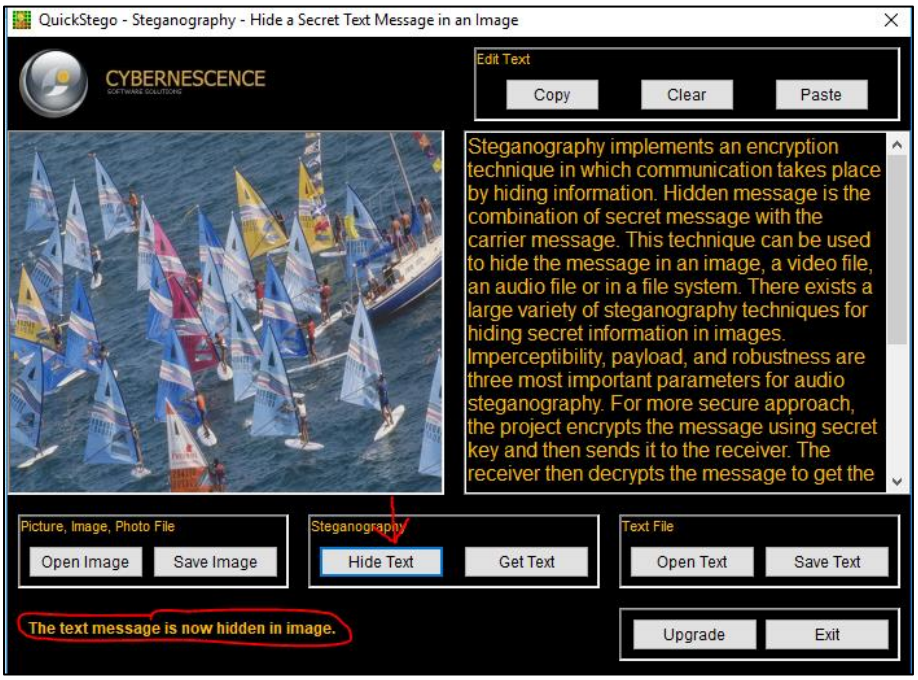

Figure 32: Success message after the hiding process in QuickStego software.

Extraction of secret message using QuickStego tool. After obtaining the final output image, it will be sent to the intended receiver of the secret message. At receiver end the recipient need to use the same QuickStego tool for extracting the secret message hidden in the image sent. Click on “Open Image” then a dialog box will open as shown in the below screenshot and select the stego file generated by this tool which has the secret message.

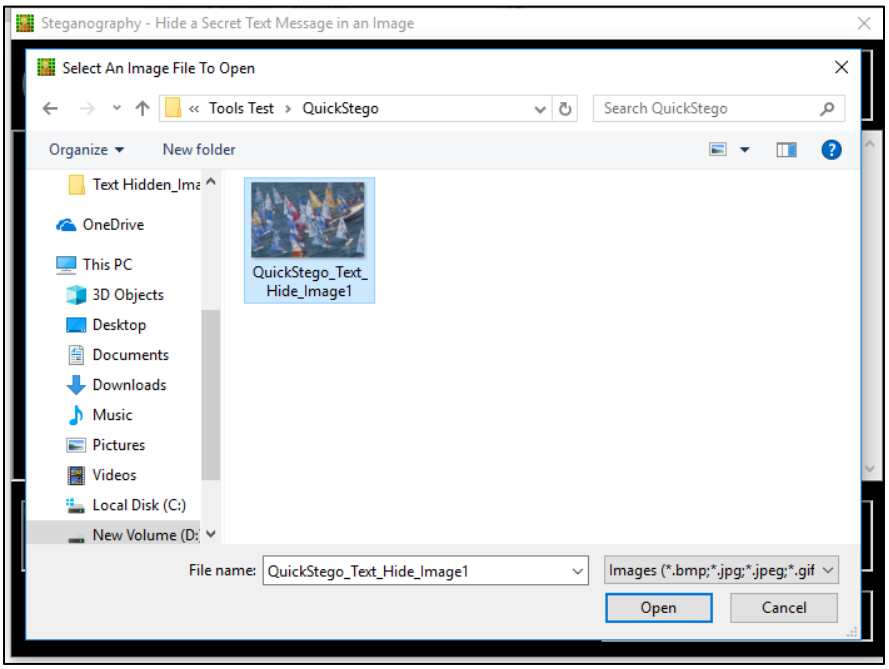

Figure 33: Selecting the stego image in QuickStego software.

Click on “Open” as shown in the above screenshot immediately it shows the secret message in the beside box of the image box as shown in the below screenshot.

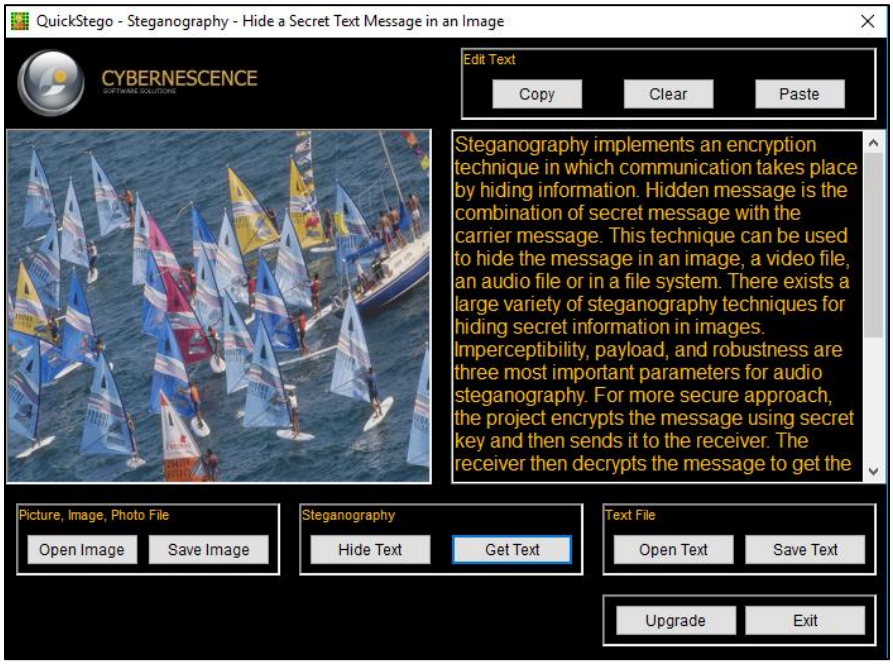

Figure 34: Final output of the extraction process by using QuickStego software.

Hiding of data using Steg tool. In this tool, to select the cover image into the tool for the hiding process, as highlighted in the below screenshot click on “File Manager” tab then select the folder which has the cover image then the images in that folder will be shown below. Select the image as highlighted bottom in the below screenshot then he mage will be displayed in the right pane of the tool homepage as below.

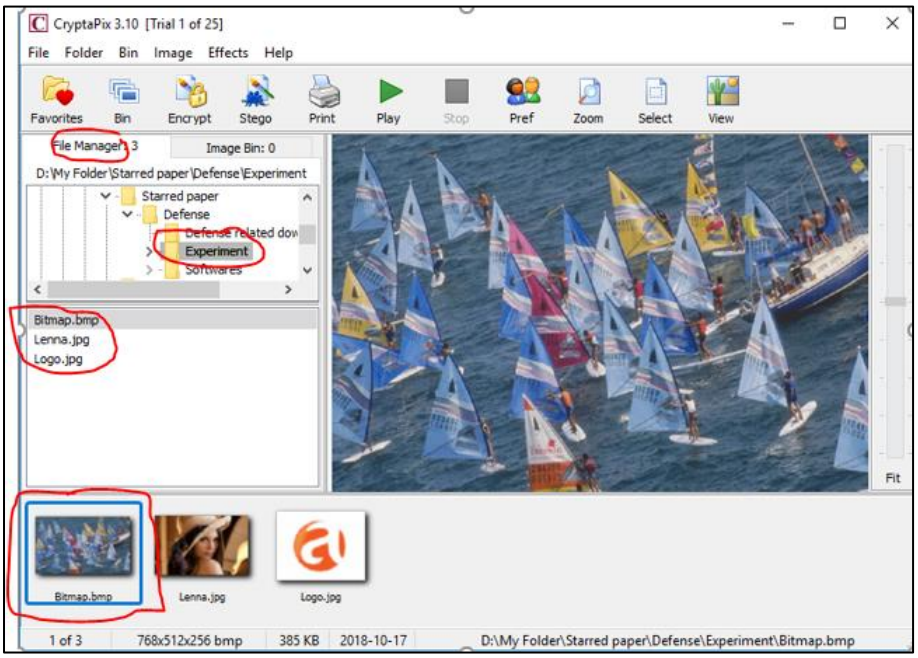

Figure 35: Selection of the cover image in CryptaPix software.

Click on stego which is highlighted in the below screenshot then select the insert option. A dialog box will be opened as show in below screenshot to select a secret message that is to be hidden in the selected cover image. There are 3 options for selecting the secret message datatype which are image, text and data file. After making the selection, click on browse and select the secret image. Provide the password and the strength of password is shown on the right side then click on ‘ok’.

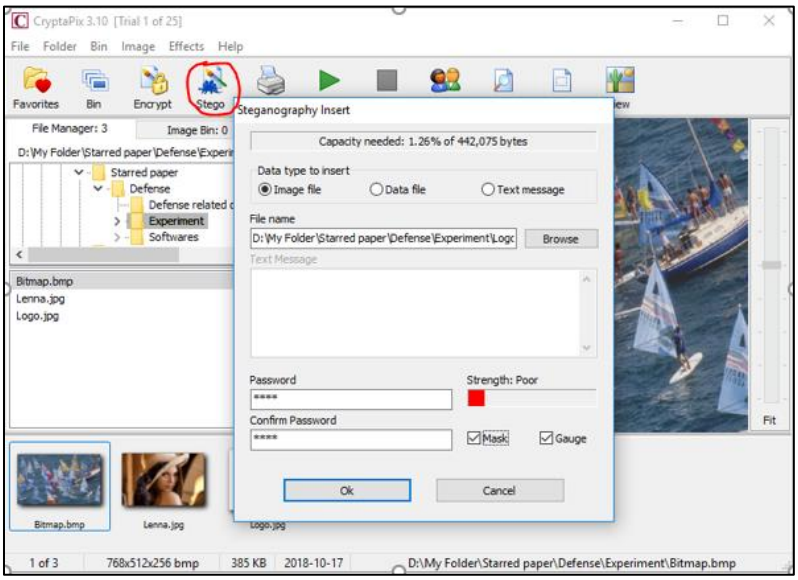

Figure 36: Selecting the secret message to hide in CryptaPix software.

After clicking ‘ok’, another popup will be displayed which will ask to choose the image file format of the image that is being saved. Make the selection and click on ‘ok’.

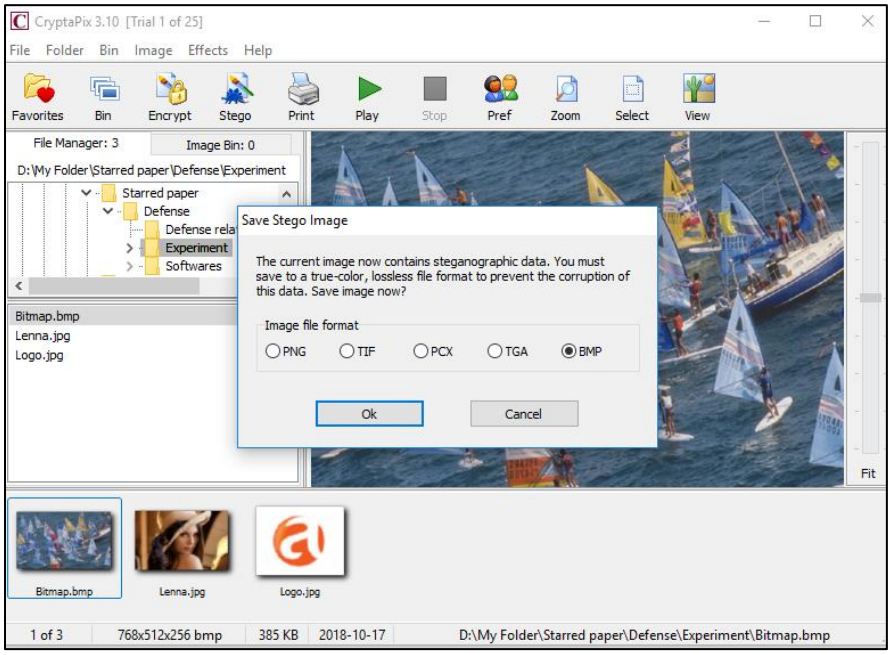

Figure 37: Selecting the file format of final output image in CryptaPix software.

Extraction of secret message using CryptaPix tool. After obtaining the final output image, it will be sent to the intended receiver of the secret message. At receiver end the recipient need to use the same CryptaPix tool for extracting the secret message hidden in the image sent. Click on stego and then select “Extract” option as shown in the below screenshot.

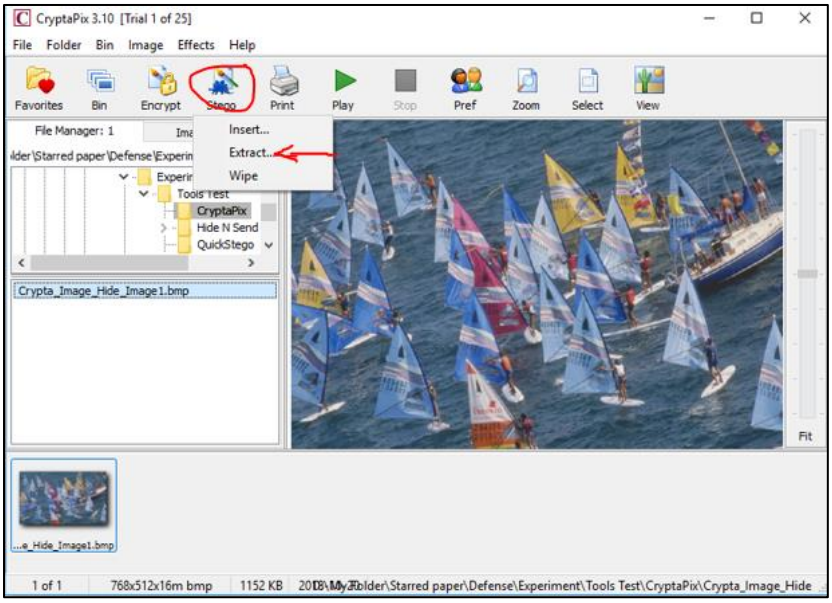

Figure 38: Option to select for extracting secret message in CryptaPix software.

After selecting the extract option then a popup is displayed with the details o the hidden message and displays the secret message in the right box. There will be an option to select the date of the extracted image. If that option is checked then the date of the extracted secret message will be same as the original image creation date or else it will be the current date.

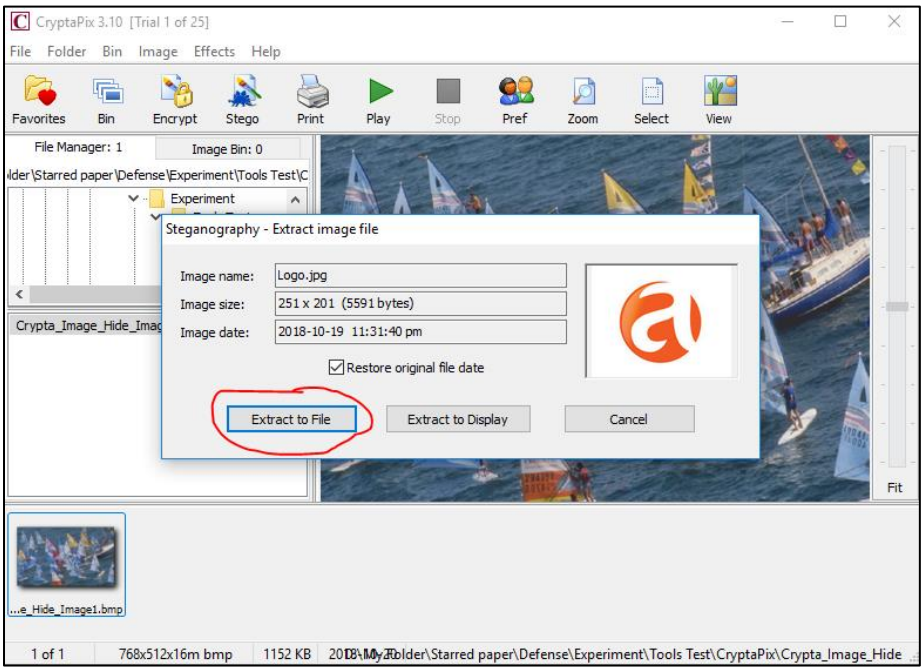

Figure 39: Options for extracting the secret message in CryptaPix software.

After clicking on “Extract to file” option below dialog box is displayed.

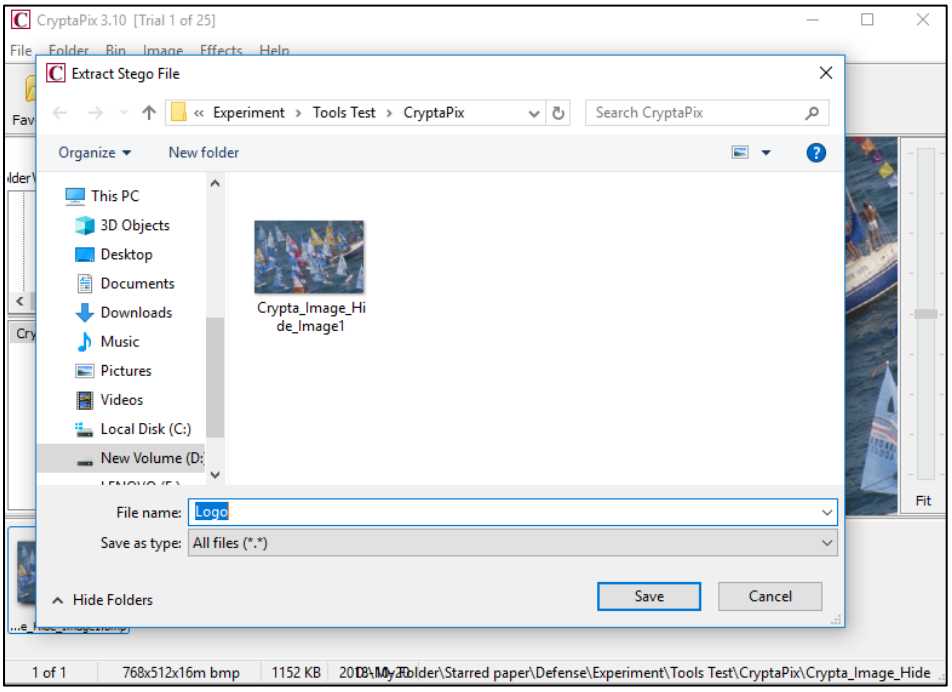

Figure 40: Selecting the stego image in CryptaPix software.

Hiding of data using Steg tool. This tool is not having a direct option to select and perform the actions. It is like a framework, we need to create a workspace and should work on it.

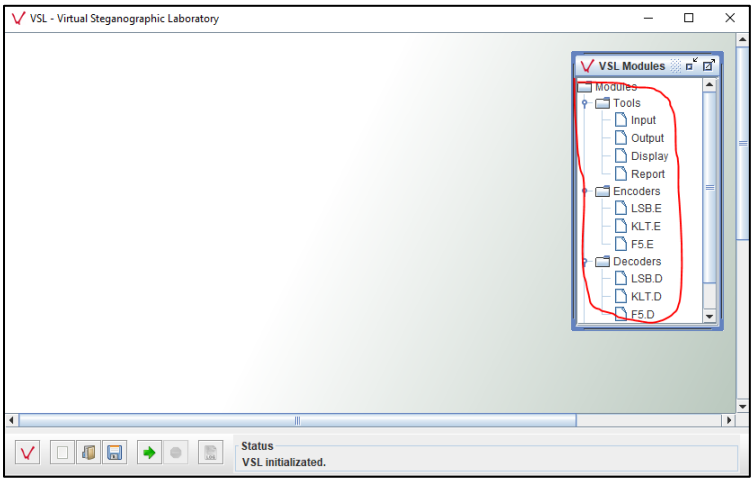

Figure 41: Modules of VSL software.

Need to create a flowchart by drag and drop of the modules available on right side of the workspace. For hiding the secret message using LSB method I have made the below flowchart b using the available options.

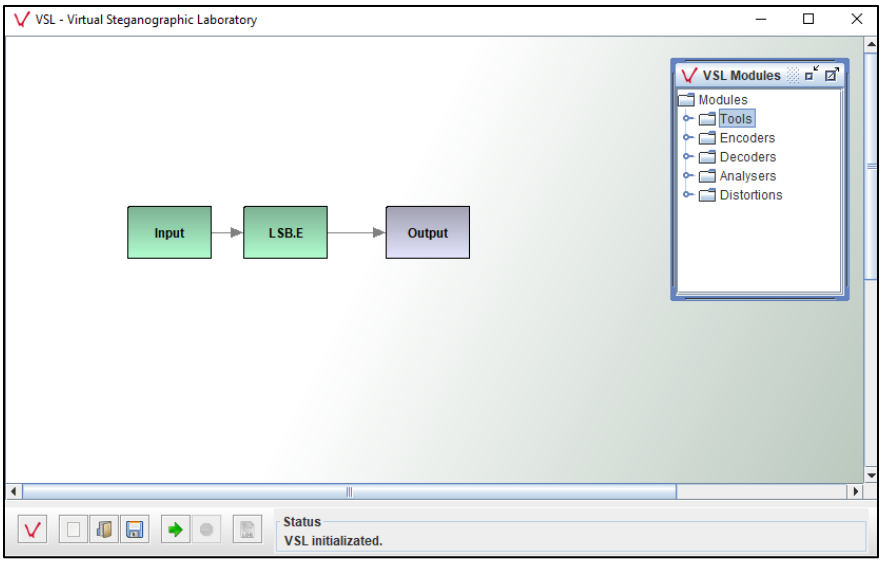

Figure 42: Flowchart of hiding process in VSL software.

Now give the parameters that are required for each module. Right click on Input module and select connect then click on LSB.E module then a transition is created from Input to LSB.E module.

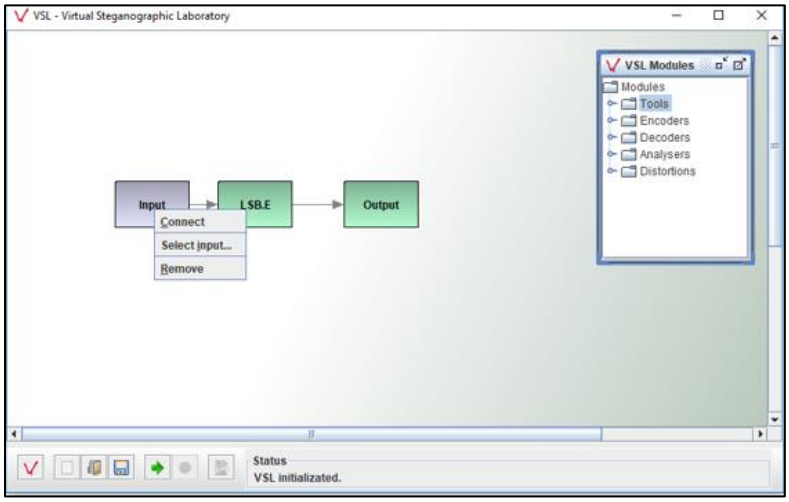

Figure 43: Providing the parameter values for input module in VSL software.

Again, right click on Input module and click on “Select input” then below dialog box will be displayed to select the cover image path.

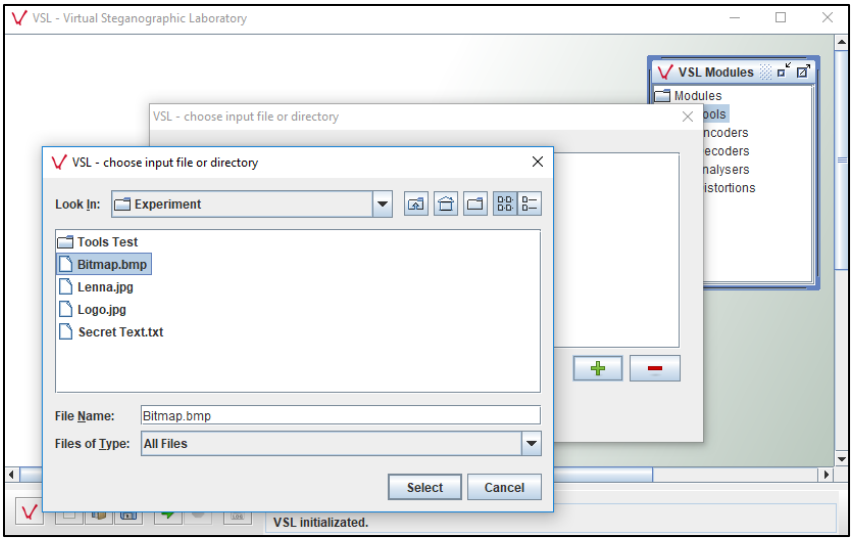

Figure 44: Selecting the cover image path in VSL software.

Right click on LSB.E module, select Connect option and then click on Output module, a transition is created between LSB.E and Output module.

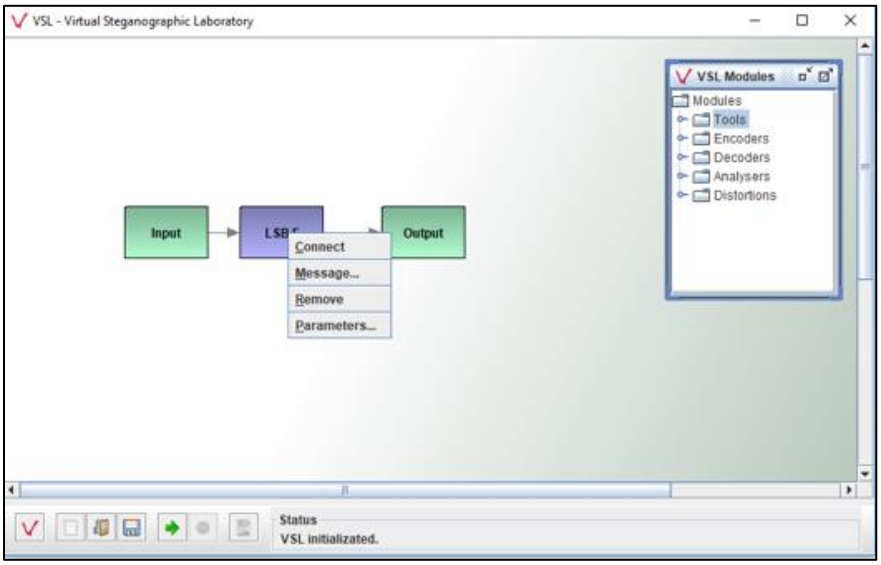

Figure 45: Options of the LSB.E module in VSL software.

Click on “Message” option to select above to select a secret message that is to be hided, below dialog box is displayed. Click on folder icon to select the secret file and the properties of the file are displayed as shown in the below screenshot.

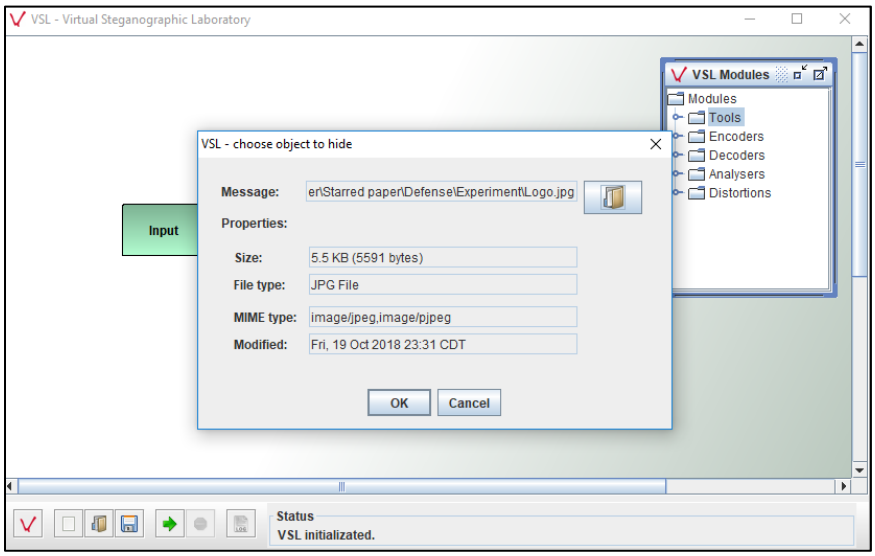

Figure 46: Selecting the secret message which is to be hidden by using VSL software.

After selecting all the required parameters of the modules, click on green arrow which is highlighted in the below screenshot to start the hiding process. The status will be displayed at the bottom. It displays “Experiment finished” after the process is completed as shown in the below screenshot.

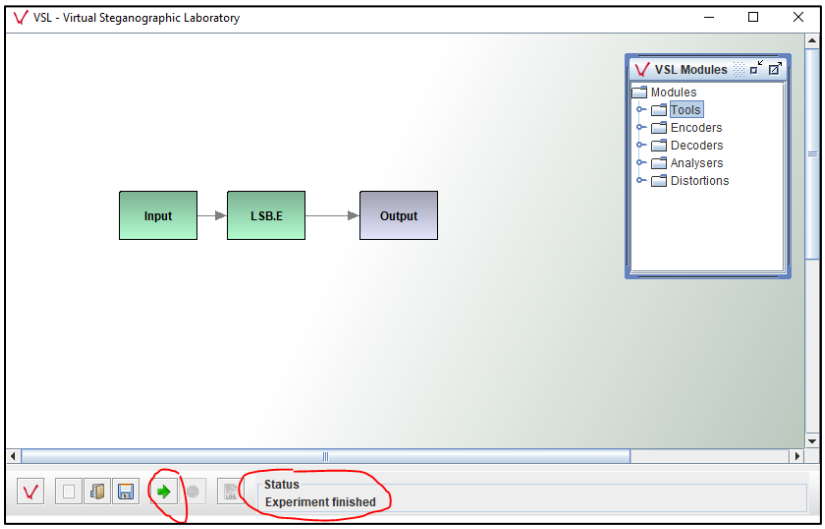

Figure 47: Status of the hiding process in VSL software.

The final output image is saved in the folder given in the “Output” module parameter value. Below is the stego image that was saved after the hiding process by the tool.

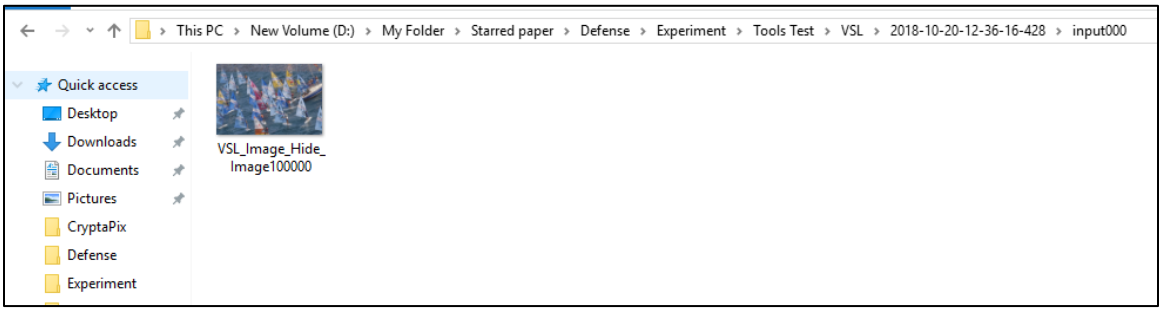

Figure 48: Output folder of the final stego image by using VSL software.

Extraction of secret message using VSL tool. After obtaining the final output image, it will be sent to the intended receiver of the secret message. At receiver end the recipient need to use the same VSL tool for extracting the secret message hidden in the image sent. In extraction process, instead of LSB.E we use LSB.D module for revealing the secret message.

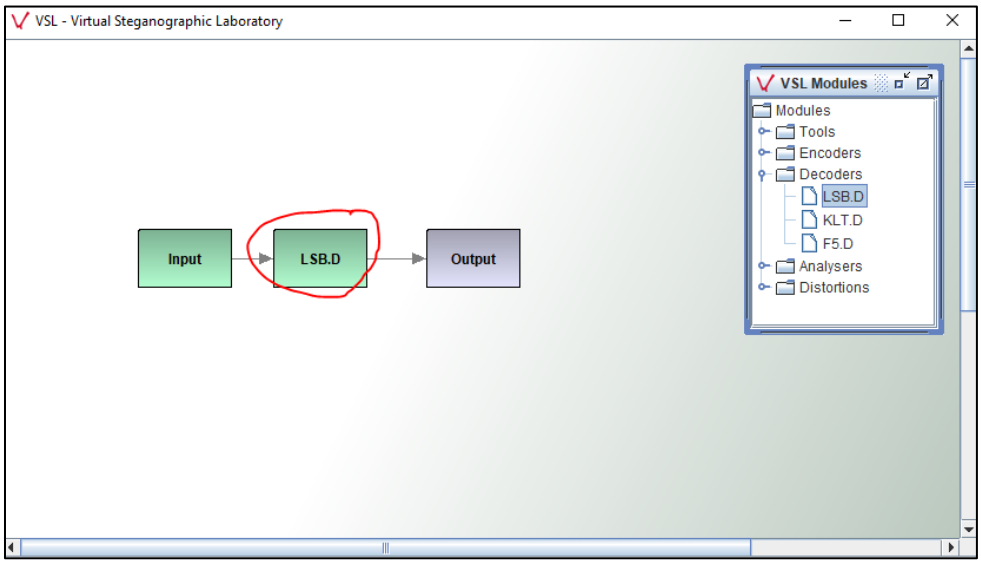

Figure 49: Extraction flowchart in VSL software.

Right click on the LSB.D module and connect to the output module. Choose the output folder for the extracted secret message by right clicking the output module, a popup will be displayed to select the path as shown in the below screenshot.

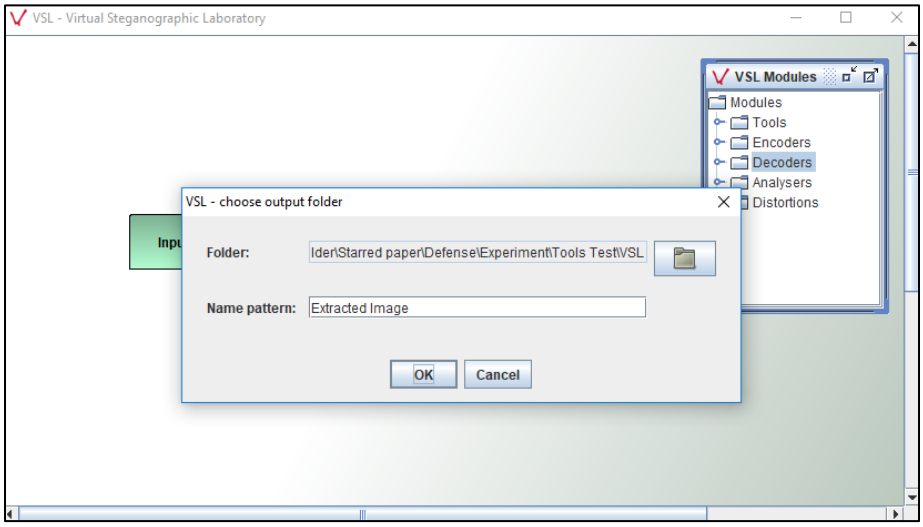

Figure 50: Selecting the output folder for the secret message in VSL software.

Start the extraction process by clicking the green arrow and the status will be displayed upon completion of the process as highlighted in the below screenshot.

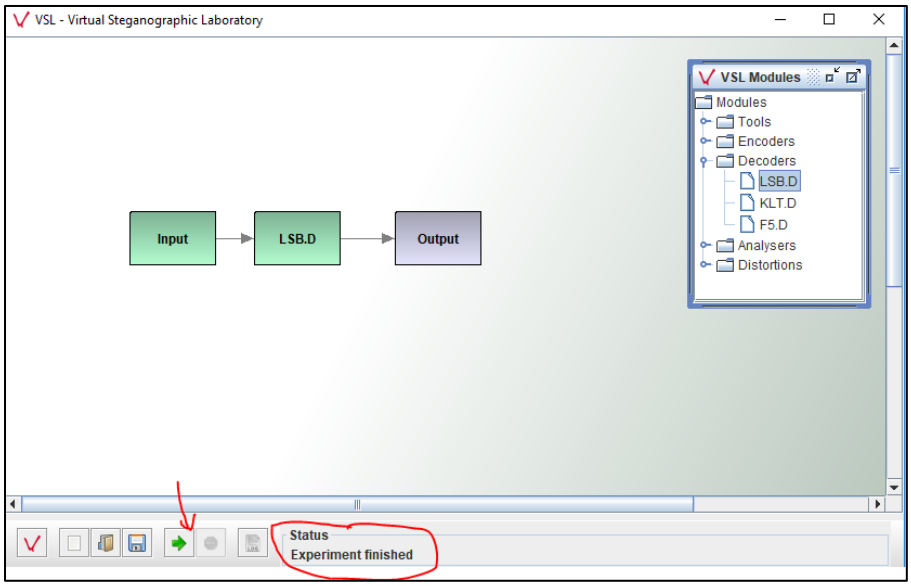

Figure 51: Status of the extraction process using VSL software.

Below is the output folder of the extracted secret message from the given cover image.

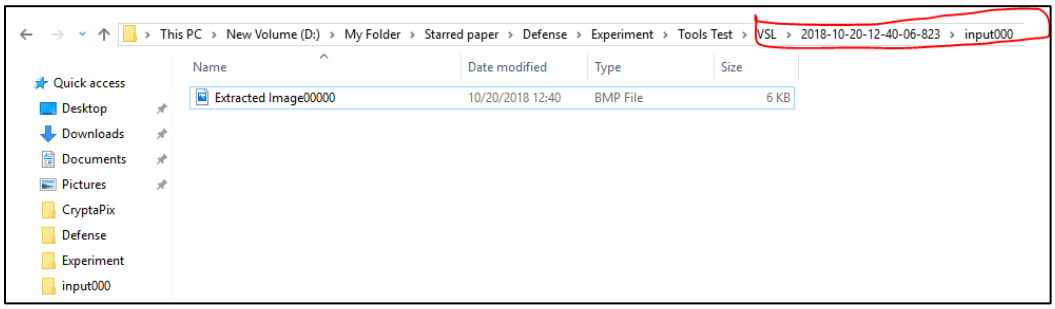

Figure 52: Output folder of the extracted secret message using VSL software.

Results

By doing the experiment using the sample test images and sample secret message, the final output images are further analyzed and the quality metrics such as PSNR and SSIM values are calculated. Below are the results obtained from the experiment

Table 4

PSNR Values by Different Steganography Tools

Test Scenario	**Crypta Pix**	**Hide N Send**	**VSL**	**Steg**	**QuickStego**
Secret image concealed in Test image1(BMP)	62.72		60.61	62.83	
Secret image concealed in Test image2(JPG)	69.10	44.25	69.32	45.69	
Secret text concealed in Test image1(BMP)	62.70		67.8	69.88	70.21
Secret text concealed in Test image2(JPG)	75.59	49.33	76.56	52.39	76.51

According to Table 4, among all the 5 tools, CryptaPix and Steg are generating high quality in the first scenario. In second scenario, CryptaPix and VSL are generating high quality images. In third scenario, QuickStego and Steg are generating high quality images. In fourth scenario, QuickStego and CryptaPix are generating the high-quality images.

Table 5

SSIM Values by Different Steganography Tools

Test Scenario	**CryptaPix**	**Hide N Send**	**VSL**	**Steg**	**QuickStego**
Secret image concealed in Test image1(BMP)	1	NA	1	1	NA
Secret image concealed in Test image2(JPG)	1	0.9971	1	0.9978	NA
Secret text concealed in Test image1(BMP)	1	NA	1	1	1
Secret text concealed in Test image2(JPG)	1	0.9991	1	0.9995	1

According to Table 5, almost all the five tools are generating the similar image compared to the original image. The value 1 will refer to exact same image to that of the original image.

Summary

In this chapter, various steganography tools are exhibited and shows the working of each tool. The process of hiding the secret message inside the cover image and extracting the secret image from the stego image by each tool is explained. After obtaining the final output images, they are compared with the original image and measured the quality difference by using the quality metrics such as PSNR and SSIM. Based on the values from the final result the best method is evaluated.

Chapter V: Conclusions and Future Work

Introduction

Steganography is a data hiding technique which is used for the communication of confidential data in a safe and secure process. It has several tools and techniques in which data can be hidden inside a cover image. The final output image is called as stego image which looks same as the original image for a human naked eye. Though the cryptography which is also a technique used for communication of confidential data securely, both of them are separate in their working and process the follow.

Conclusions

To conclude, in this paper, explained the process of Steganography and different techniques in steganography. Discussed the advantages and disadvantages of steganography over cryptography. Also discussed the process of steganalysis which is a process of identifying the use of steganography. Finally, selected 5 different steganography tools and compared them by the quality of stego images generated by each tool. The steganography tools are CryptaPix, Steg, VSL, Hide N Send, QuickStego. Two carrier images are selected, and two secret files are selected, one is an image, and another is a text file which are hided in the carrier images. The final output images also known as stego images are further analyzed using Imatest software in Matlab for calculating the SSIM and PSNR values. After observing the PSNR values of the stego images generated by all five steganography tools, they produced the images of high quality and SSIM values conclude that almost all the tools produced the images that are similar to original image.

Future Work

Steganography is a best way of communication of confidential data since it is not detectable by simply seeing at the image. However, it will be a difficult task to get a steganography which will satisfy both the criteria of the high robustness as well as high security. The improvement for the existing steganography tools can be done by addressing the above limitation. The use of multiple passwords can increase the security of the content. Though the image is extracted, one password is identified there will be some more passwords to identify which will make the job tough for the intruder. It is called as a multi factor authentication. Some of the tools are having the image restrictions, whatever the file format of the cover image, the final output image is saved in specific image format which may alter the image quality. The main aim of the steganography is not to alter the image properties from the original image and make the image as much as like that of original image, these changes should be done.

References

Banerjee, & Indradip. (2011). A survey of steganography and steganalysis technique in image, text, audio and video as cover carrier. *Journal of Global Research in Computer Science, 2*(4), 116..

Bhagat, A., & Dhembhare, A. (2015). An efficient and secure data hiding technique-Steganography. *International Journal of Innovative Research in Computer and Communications Engineering, 3*(2), 944-949.

Bhattacharyya, & Souvik. (n.d.). *ResearhGate*. Retrieved from ResearhGate: https://www.researchgate.net/figure/Universal-Steganalysis-method-Calibration-Based-Feature-Fridrich-et-al-10-developed_fig2_236953170

Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2011). A Survey of steganography and steganalysis Technique in image, text, audio and video as cover carrier. J*ournal of Global Research in Computer Science, 2*(4), 1-16.

Chan, C.-K. (2002). *Hiding data in images by simple LSB substitution*. Department of Engineering and Information Technology, City University of Hong Kong, Hong Kong.

Emam, & Marwa, M. (2015). A modified image steganography method based on LSB technique. *International Journal of Computer Applications, 125*(5), 12-17.

Imatest Documentation. (n.d.). Retrieved from Imatest: http://www.imatest.com/docs/ssim/.

KITPLOIT. (2017, November). Retrieved from https://www.kitploit.com/2018/02/lsb-steganography-python-program-to.html.

Kumari, S. (2017). A research paper on cryptography encryption and compression techniques. *International Journal of Engineering snd Computer Science, 6*(4), 20915-20919.

Learning & Excercise. (2017, November). Retrieved from http://learning.maxtech4u.com/what-is-steganography/

MathWorks documentation. (n.d.). Retrieved from MathWorks: https://www.mathworks.com/help/vision/ref/psnr.html.

Patel, A. A. (2003). Information hiding–The art of steganography. *Global Information Assurance Certification Paper*.

Patel, S. K. J., & Tahilraman, N. V. (2016). Information hiding techniques: Watermarking, steganography. *International Journal of Innovational Research in Electrical, Electronics, Instrumentation and Control Engineering, 4*(4), 168-173.

Provos, N., & Honeyman, P. (2003). *Hide and seek: An Introduction to steganography. IEEE Security and Privacy Magazine, 1*(3), 32-44.

Raja, K. B., & Vanugopal, K. R., & Lalit, P. (2004). *A secure stegonographic algorithm using LSB, DCT and image compression on raw images.* A report.

Research gate. (2013, December). Retrieved from https://www.researchgate.net/figure/System-Architecture-of-the-proposed-GA-based-Optimized-Video-Steganography-technique-a_fig1_259525938.

Sing, M. P. (2016). A comparative study of audio steganography techniques. *International Research Journal of Engineering and Technology, 3*(4), 580-585.

Student web. (n.d.). Retrieved from http://studentweb.niu.edu/9/~z172699/Description.html.