STEGANOGRAPHY IS THE ART OF HIDING DATA

Mukta Sharma[1], **Dr. R.B Garg**[2]

[1] *Research Scholar, TMU, Moradabad (India)*

[2] *Ex-Professor, DU, Delhi (India)*

ABSTRACT

Steganography has been derived from the Greek word Steganous (covered, concealed or protected) and Graphie (writing) which means "covered writing". It is the process of hiding a secret message with an ordinary message. Which when viewed by someone else will be able to see only the ordinary message they will fail to identify that it contains a hidden or encrypted message. The hidden message can be extracted once it reaches its destination. Steganography is now gaining popularity among the masses because of ease in use and abundant tools available.

In this paper, the emphasis is given on the basic technique used to implement Steganography and also how to extract the data and check whether the file received by the reader carries some hidden message. Paper also depicts various tools available for the same.

Purpose: *This paper introduces steganography by explaining what it is, providing a brief history with illustrations of some methods for implementing steganography. The objective of the paper is to highlight the field of Steganography from different viewpoints (from the terrorist view point, and from a naïve user's outlook).*

Structure of Paper: *Paper begins with the brief introduction, history, and type of steganography. It also focused on several applications that uses steganographic services, highlights methods of steganography, its advantages and disadvantages. Later the implications of stegaonographic are covered, along with the various tools available.*

Design/methodology/approach: *A conceptual approach is followed to comprehend the concept of Steganography and what should be done to identify that the files downloading from internet does not have any hidden files, it has been explained with real case studies.*

Findings*:* *Steganography can be easily implemented and the hidden message can be retrieved easily once the user is aware of how to retrieve the message.*

Research limitations/implications: *Research design is exploratory in nature hence; the results of the study are not very conclusive.*

Keywords: Cryptography, DOS, Information Hiding, Information Security, MITM, Steganography, Steganography Techniques

I. INTRODUCTION

Steganography is a technique used to transmit hidden information without raising any suspicion about the existence of such information. Steganography aims at hiding information in an original – cover – data in such a way that a third party is unable to detect the presence of such information by analyzing the information bearing stego data.

Earlier it was considered that cryptography is the only way to secure the data or message. But we can say, steganography add-on encryption can make the data or message more secure. In a stego file, we can hide the encrypted message which can later be retrieved and decrypted by the other party who knows about the steganography and the key for decryption. The hidden information can be plain text, sound, video, cipher text, or even images. Steganography works by replacing bits of useless or unused data in regular computer files with bits of different, invisible information. Steganography primary goal is to hide data within some other data such that the hidden data cannot be detected. The secondary focus is to prevent extraction from the cover file without destroying the cover and prevent destruction of the stego-message without destroying the cover. Most frequently, steganography is applied to images, but many other data or file types (audio, video, text, executable program, HTML files etc) are possible.

II. HISTORY OF STEGANOGRAPHY

Steganography was proposed way back in 440BC, it was initially used by ancient Greeks. They used to write messages on the wood, later they used to cover it with wax and write an ordinary message on it. There are some stories of hiding message on the messenger's body like a message tattooed on the shaven head later the person with grown hair was sent and the message was retrieved by again shaving off his hair. This method has a major drawback of time delay and space.

During World War II, concept of invisible ink was introduced by French and Microdots were used by espionage agents. Microdots were typically minute (less than the size of the period produced by a typewriter) needed to be embedded in the paper and covered with an adhesive, such as collodion. This was thoughtful and thus noticeable by viewing against glancing light. Alternative techniques included inserting microdots into slits cut into the edge of post cards. Some secret messages were written beneath the postal stamp. Some messages were made hidden with the help of vegetable oil, milk and juices when heated become dark. Some seemingly innocent letter could contain a very different message written between the lines. Decoding this message by extracting a specific letter in each word could reveal the message. Modern Steganography uses the concept of encryption also. They encrypt/ encode the data hide it and when received the message needs to be deciphered/decoded/ decrypted.

III. WHO USE STEGANOGRAPHY

The concept of Steganography is very interesting it hides the very existence of the message in the communicating data. It could be used by anyone who wants to hide the message from the rest of the world and want to be read by only the intended or original receiver. It is used by anti-forensics mechanism to mitigate the effectiveness of a forensics investigation. It is commonly used by lovers, boy friend- girl friend to exchange notes without anyone else knowing the real conversation. It is used by terrorist to secretly communicate information, by posting the images on a website for download, which is later used by other terrorist to download the image with a hidden message with special attack instructions.

3.1 Applications (Why Should One Use Steganography)

- Confidential communication and secret data storing
- Protection of data alteration
- Access control system for digital content distribution.

- Media Database systems.
- Used in modern Printers

3.2 Advantages Of Using Steganography

- Difficult to detect and only receiver can detect.
- It can be done faster with various software's available.
- Provide better security for sharing data in LAN, MAN, WAN.

3.3 Disadvantages Of Using Steganography

- The confidentiality of information is maintained by the algorithms, and if the algorithms are known then this technique is of no use.
- If this technique goes in the wrong hands like terrorist, hackers then it could lead to dangerous results.
- The files downloaded from internet have some infected files attached to the original/ordinary data which are hidden. When the file runs the hidden infected file also runs simultaneously and will affect the system. For instance, Click on an adware will run the advertisement along with an infected file (may be Trojan, Virus, Key logger etc) which might affect your system in long run.

3.4 Types of Steganography

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display [3]. The redundant bits of an object are those bits that can be altered without the alteration being detected easily [1]. Three categories of file formats that can be used for Steganography techniques are shown in figure 1

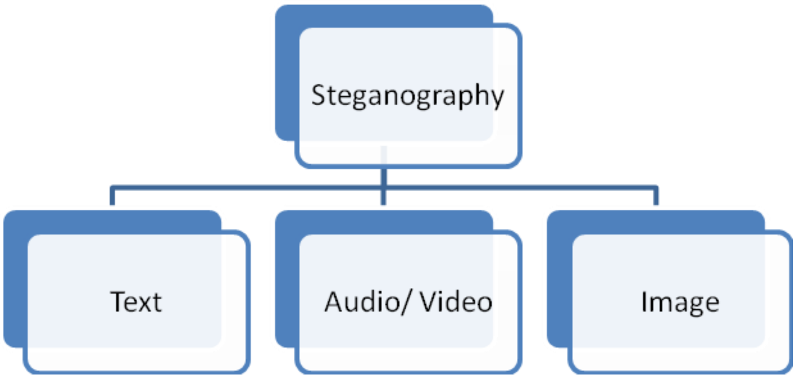


Fig. 1 Techniques of Steganography

3.4.1 Steganography in Text- It involves three types of coding:

- *Line-Shift Coding:* Here, text lines are vertically shifted to encode the document uniquely.
- *Word-Shift Coding*: The codeword's are coded into a document by shifting the horizontal locations of words within text lines, while maintaining a natural spacing appearance.

- *Feature Coding*: In feature coding, certain text features are altered, or not altered, depending on the codeword.

3.4.2 Steganography in Images-Digital images are made up of pixels, the arrangement of pixels make up the image's "raster data". 8-bit and 24-bit images are common. The larger the image size, the more information can be hidden. However, larger images need compression to avoid detection. Two kinds of image compression are lossless and lossy compression. Both methods save storage space but have different effects on any uncompressed hidden data in the image.

- "Lossy" JPEG (Joint Photographic Experts Group) file format, offers high compression, but may not maintain the original images integrity. It loses the data. Hence it is called "lossy".
- "Lossless" compression maintains the original image data exactly; It is thus more favored by steganographic techniques as the message can be reconstructed exactly. Eg: (BMP),(GIF) Formats.

Image Encoding Techniques- The following are the common approaches of hiding information in images:

- *Least Significant bit insertion*- Most popular technique when dealing with images. LSB method is based on altering the redundant bits that are least important with the bits of the secret information. The aim of the LSB is to transmit the secret information to the receiver without knowing to the intruder that the message is being passed. It is very simple, but susceptible to lossy compression and image manipulation
- *Masking and Filtering*- can be used on 24bit per pixel images, applied on both colored and grayscale images. Masking and filtering is similar to placing watermarks on a printed image. These techniques embed the information in the more significant areas than just hiding it into the noise level. Watermarking techniques can be applied without the fear of image destruction due to lossy compression as they are more integrated into the image.
- *Algorithms and transformations*- use mathematical functions to hide the least bit coefficients in the compression algorithms that reduce the file size of images. Many algorithms like [Discrete Fourier transformation technique (DFT). 2. Discrete cosine transformation technique (DCT). 3. Discrete Wavelet transformation technique (DWT)] are used to compress the image.

3.4.3 Steganography in Audio/Video- Embedding secret messages into digital sound is known as audio Steganography. This method can embed messages in Wav, au, Mp3 etc

- Sample quantization method
- Temporal sampling rate
- Another digital representation

3.5 Methods of Audio Data Hiding

- *Low Bit encoding*- A very popular methodology is the LSB (Least Significant Bit) algorithm, which replaces the least significant bit in some bytes of the cover file to hide a sequence of bytes containing the hidden data. That's usually an effective technique in cases where the LSB substitution doesn't cause significant quality degradation, such as in 24-bit bitmaps.[4]
- *Parity coding*- is one of the robust audio steganographic technique. Instead of breaking a signal into individual samples, this method breaks a signal into separate samples and embeds each bit of the secret message from a parity

bit. If the parity bit of a selected region does not match the secret bit to be encoded, the process inverts the LSB of one of the samples in the region. Thus, the sender has more of a choice in encoding the secret bit.[4]

- *Phase Coding*- The technique works by replacing the phase of an initial audio segment with a reference phase that represents the secret information. The remaining segments phase is adjusted in order to preserve the relative phase between segments. In terms of signal to noise ratio, Phase coding is one of the most effective coding methods. When there is a drastic change in the phase relation between each frequency component, noticeable phase dispersion will occur. However, as long as the modification of the phase is sufficiently small, an inaudible coding can be achieved [6, 12].
- Echo Data Hiding- Echo hiding technique embeds secret information in a sound file by introducing an echo into the discrete signal. Echo hiding has advantages of providing a high data transmission rate and superior robustness when compared to other methods. Only one bit of secret information could be encoded if only one echo was produced from the original signal. Hence, before the encoding process begins the original signal is broken down into blocks. Once the encoding process is done, the blocks are concatenated back together to create the final signal [7, 8].

3.6 Types of Stegosystems- It can be implemented in 3 ways:-

- Pure stegosystems - no key is used.
- Secret-key stegosystems - secret key is used.
- Public-key stegosystems - public key is used.

Basic block diagram of Stegosystems- The basic diagram depicts the use of carrier file along with the hidden file which will be converted into a stego file. This file will be further send either to the victim for some attack or will be used by terrorist group who is already aware of the steganographic file.

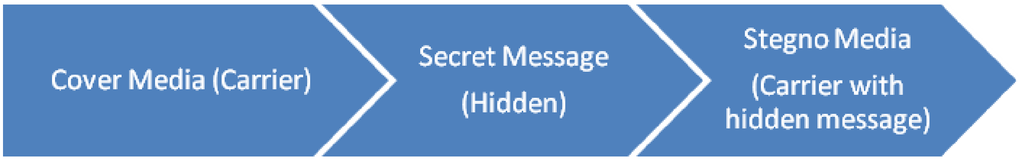


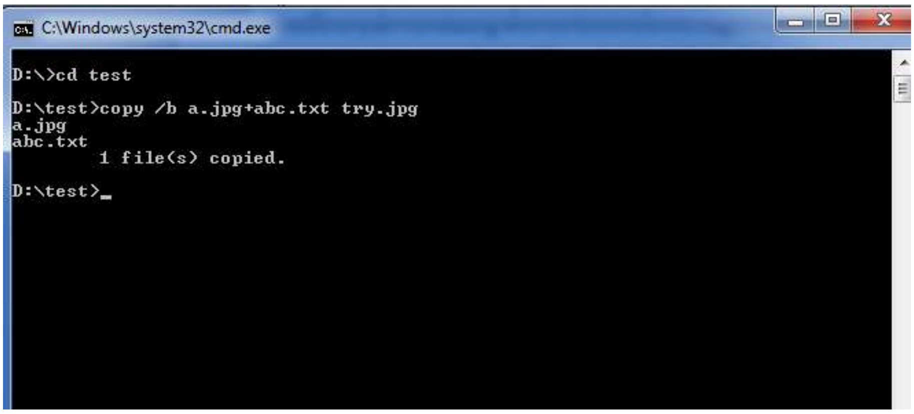
Fig. 2 Block diagram of Stegosystem

Way to implement Steganography- Steganography hides the covert message. The steganography process generally involves placing a hidden message in some transport medium, called the carrier. The secret message is embedded in the carrier to form the steganography medium. The use of a steganography key may be employed for encryption of the hidden message and/or for randomization in the steganography scheme. In summary: steganography_medium = hidden_message + carrier + steganography_key (if you are using cryptography)
Make a folder with any name (test) and place an image (a.jpg) or audio (b.mp3) or video file (c.avi) which you wish to forward it to your friend along with the secret text message. Place the text file also in the same folder with the secret message in it. Let us name the text file as abc.txt which has text written as Hello, How are you? Can we meet at 12:00 in the library.
Now Let us see how to actually implement steganography

1. Goto Command prompt by writing cmd at run or by clicking on command prompt from program.

2. cd test (cd <foldername>- change directory dos command to move to that folder)
3. copy /b a.jpg+abc.txt try.jpg

   Here copy /b is for copying binary files (bits will be copied) a.jpg (is the carrier file)+(is used for concatenation/Joining)abc.txt (hidden message) try.jpg (is a stago file)



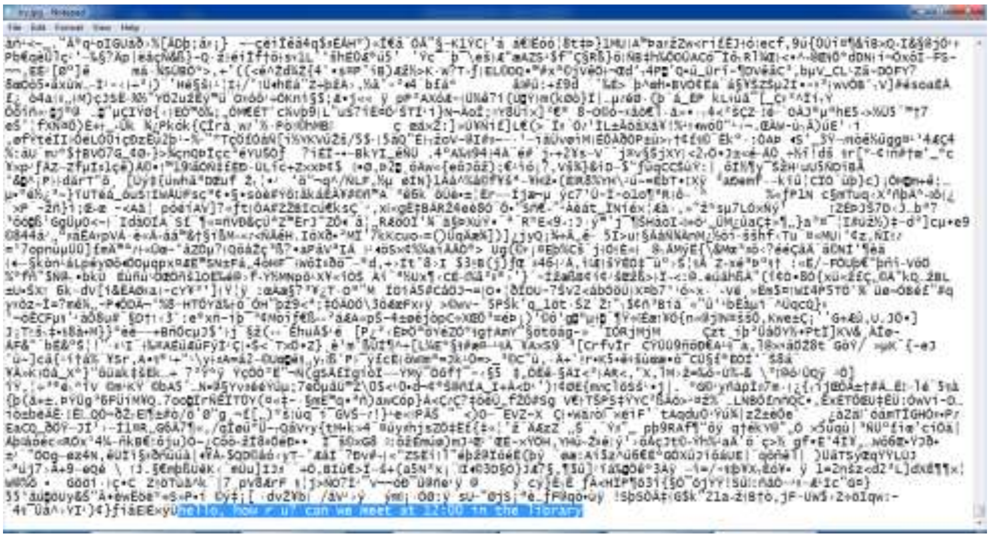
```
D:\>cd test

D:\test>copy /b a.jpg+abc.txt try.jpg
a.jpg
abc.txt
        1 file(s) copied.

D:\test>_
```

Fig. 3 Command Prompt for performing Steganography

4. Just send try.jpg file to any user. (User will not be able to trace That there is some secret message hidden beneath the image)
5. If the user is well acquainted by the concept of steganography. He will right click on try.jpg and open it with notepad->goto the last line and will be able to see the secret message



```
[illegible]hello, how r u! can we meet at 12:00 in the library
```

Fig. 4 Notepad file showing the hidden message

Let us see how /b works

Copy /b a.mp3+b.mp3+c.mp3+d.mp3 e.mp3

This will concatenate all songs starting from a.mp3 to d.mp3. When you will send e.mp3 file it will be too heavy can check the properties by right clicking and now it will play songs in a sequence like a playlist. We can say e.mp3 is now working as a playlist comprising of music files like (a-d.mp3)

Imagine a situation if the user need to send an image/audio/video file as a text/ image file. The user needs to compress it so that after seeing the size of the file the recipient should not feel suspicious. Now since the carrier file is a zip or rar file the user need to make stego file also as rar or zip, otherwise the receiver won't be able to retrieve the file. But now there is again a problem when the recipient is clicking on the rar file and extracting the file the hidden message will not be displayed in the notepad (end of the line) as did in the previous example.

Therefore, at the receiver end the user need to goto the command prompt and convert rar or zip file which is delivered to him into jpg or avi or mp3 file. copy /b .rar .jpg

Steganography is concerned with concealing the fact that a secret message has being sent, along with hiding the contents of the message. Let's see few live cases:-

Malwares, spywares, adware's and freeware are using steganography extensively. The moment you download anything online from Torrent (illegal to use in India) your system will have hidden infected files(virus), which will later infect your system by taking data from your system, install key loggers, denial of service attack etc. French security researcher Xylitol noted something strange as he was getting few images (jpg) and when he analysed and compared the original file from google and the file he had received he found discrepancy in size. With the help of hexadecimal viewer software he could find those extra bits, which he highlighted and converted back (unreadable format). Later he used software OllyDbg and was able to retrieve the data which showed few financial institutions targeted like Deutsche bank (Germany). The hidden file was infecting the user's system. The moment he opens the bank website from his system the data will be given to the hacker, MITM attack (Man-In-the-Middle Attack).

Al-Qaeda relied on steganography- When investigators from the United States of America examined the Al-Qaeda's network it was found that they had extensively used steganography to pass on messages. Steganography is being used by various groups to pass on messages between each other. Once the messages were encrypted, the Al-Qaeda members downloaded the files using various software to execute several terrorist plots. Traces of the technique being used during the 9/11 attack were also seen during investigation. [10]

Jack Kelly, USA said [10] "the messages were hidden in the X-rated pictures on several pornographic websites and the posted comments on sports chat rooms may lie the encrypted blueprints of the next terrorist attack against the United States. It sounds farfetched, but US officials and experts say latest method of communication being used by Osama bin Laden and his associates to outflow law enforcement. Bin laden, indicated in the bombing in 1998 of two US embassies in East Africa, and others are hiding maps and photographs of terrorist targets and posting instructions for terrorist activities on chat rooms, bulletin boards and other websites".

Using their imagination- Steganography has most of the time been used by terrorist groups on X-rated files. The messages are embedded into such file. To anyone watching the file it is a normal picture. However the person on the other end would know exactly what to look for.

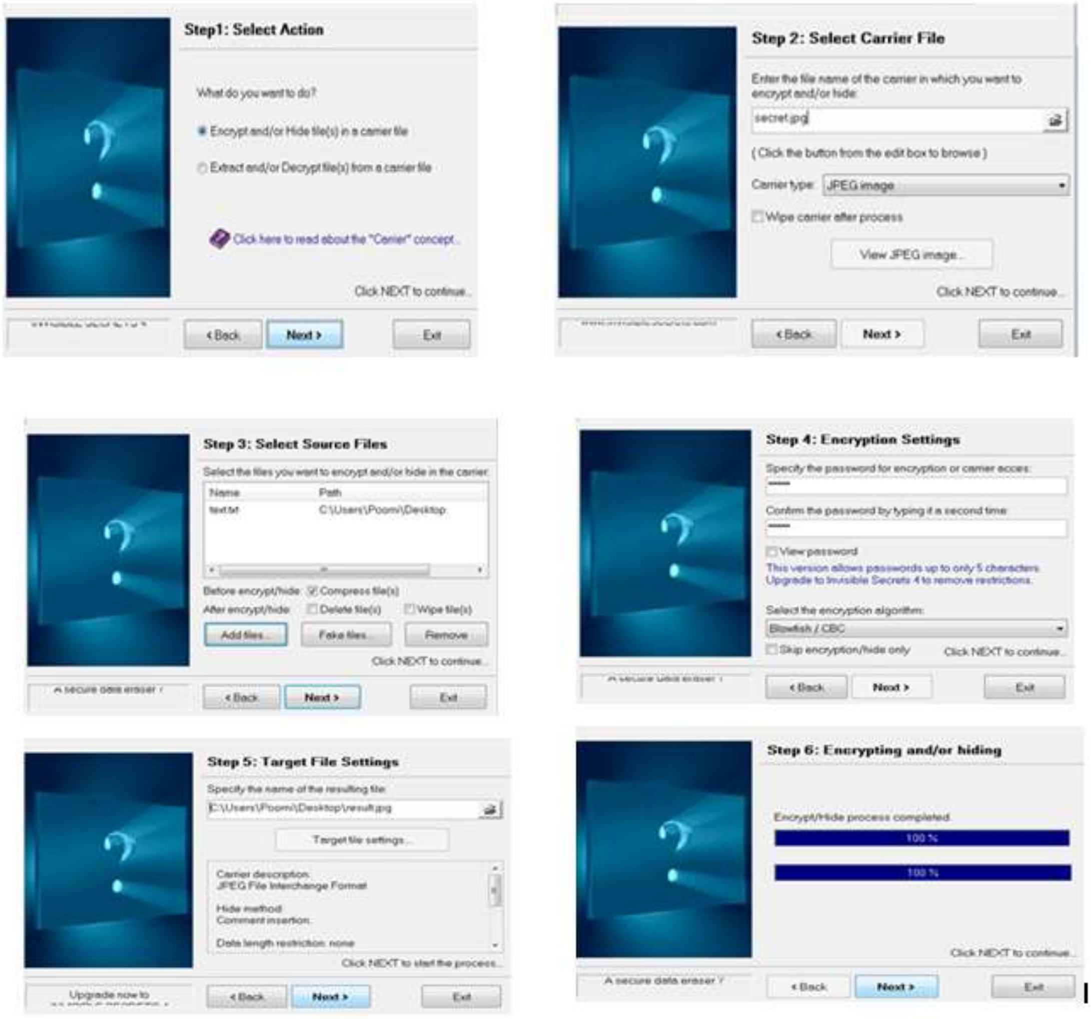
Steganography messages are difficult to detect by investigators- For an investigating officer detecting steganography is a nightmare. There is absolutely no record to show that the sender and the receiver had ever communicated. They do not exchange calls or emails. What investigating agencies have been doing is keeping a track of all downloaded pictures on the web. Pictures that are downloaded in places where the terrorist networks are strong are part of the data base. They would then keep a watch on these pictures closely to see if there are messages coded into the picture.

Steganalysis- deals with the detection of hidden content rather than focusing on the retrieval/extraction of the message. As steganography deals with the concealment of a message. There are numerous tools used to hide the message and even extract the original message. In short, we can say to make digital steganography easier we use tools [9, 11]:

Invisible Secret tool is used as secret stegosystem tool [5]:

- Select Action

- Select Carrier File
- Select Source File
- Encryption Settings
- Target File Settings
- Encryption or Hiding



Step1: Select Action

What do you want to do?

● Encrypt and/or Hide file(s) in a carrier file

○ Extract and/or Decrypt file(s) from a carrier file

Click here to read about the "Carrier" concept...

Click NEXT to continue...

Step 2: Select Carrier File

Enter the file name of the carrier in which you want to encrypt and/or hide:

secret.jpg

(Click the button from the edit box to browse)

Carrier type: JPEG image

□ Wipe carrier after process

Click NEXT to continue...

Step 3: Select Source Files

Select the files you want to encrypt and/or hide in the carrier:

Name	Path
text.txt	C:\Users\Poom\Desktop

Before encrypt/hide: ☑ Compress file(s)

After encrypt/hide: □ Delete file(s) □ Wipe file(s)

Click NEXT to continue...

A secure data eraser !

Step 4: Encryption Settings

Specify the password for encryption or carrier acces:

Confirm the password by typing it a second time:

□ View password

This version allows passwords up to only 5 characters. Upgrade to Invisible Secrets 4 to remove restrictions.

Select the encryption algorithm:

Blowfish / CBC

□ Skip encryption/hide only

Click NEXT to continue...

Step 5: Target File Settings

Specify the name of the resulting file:

C:\Users\Poom\Desktop\result.jpg

Carrier description:
JPEG File Interchange Format

Hide method:
Comment insertion.

Data length restriction: none

Click NEXT to start the process...

Upgrade now to

Step 6: Encrypting and/or hiding

Encrypt/Hide process completed.

100 %

100 %

Click NEXT to continue...

A secure data eraser !

Fig. 5 Invisible Tool used for Steganography

Histogram Analysis- Is used to identify a file with a hidden message. If somebody has the original file can compare it with the suspected file. Can check the properties of the original file and suspected file. To compare the contents we have Notepad, HEX editor to identify inconsistencies and patterns.

WinHex- Is an analysis tool. It allows conversions between ASCII and Hex. It also allows comparison of file and a detailed report is saved (comparative as well as tells differences or equal bytes)

HiderMan- is a sophisticated program in which after hiding a message, we can review the file with HEX. Can see the beginning and end of the file comparing both the files (original and suspected).

Stegspy- It's a signature identification program, searches for stego signatures and determines the program used to hide the message.

Camouflage- Is used to determine the password. The location of the password can be determined by using MultiHex which searches for Hex strings.

IV CONCLUSION

Steganography the concept was given to the world by Greeks. Steganography is still in its adolescent age when comes to cyber security, it still have a long way to go. Steganography can be considered an extension of cryptography. The main difference between the two is that steganography is not visible to everyone unlike cryptography which does not attempt to hide the information. [2] Cryptography may not be as secure as steganography because the mere presence of an encrypted message may entice individuals to attempt to 'break the code'. On the contrary, steganography is a process used to hide information so that only the sender and recipients are aware of the hidden information. Data or message can be hidden using any multimedia files like audio, video, text, and image files. This will later become virtually undetectable to individuals that do not know of the hidden information’s existence.

Steganography can be used for legitimate and illegitimate purposes. Therefore, the focus of the paper was primarily on the basic implementation of steganography. The paper is a good platform for naïve users to know about the stego files and how to retrieve the message. They can verify the file, if they have some suspicious by checking the properties and file size. The paper conclude that various tools are used to conduct steganography and to retrieve hidden message. But till date steganography is a challenge for the forensics department to trace.

REFERENCES

1. Anderson, R.J. & Petitcolas, F.A.P., “On the limits of steganography”, IEEE Journal of selected Areas in Communications, May 1998
2. http://bit599.netai.net/stego_summary.htm
3. Currie, D.L. & Irvine, C.E., “Surmounting the effects of lossy compression on Steganography”, 19th National Information Systems Security Conference, 1996
4. Jayaram P, Ranganatha H R, Anupama H S, “INFORMATION HIDING USING AUDIO STEGANOGRAPHY – A SURVEY”, The International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.3, August 2011
5. Invisible Secret Tool is available from : http://www.invisiblesecrets.com/download.html
6. Nedeljko Cvejic, Tapio Seppben “Increasing the capacity of LSB-based audio steganography ” FIN- 90014 University of Oulu, Finland ,2002.
7. Sajad Shirali-Shahreza M.T. Manzuri-Shalmani “High capacity error free wavelet domain speech steganography” ICASSP 2008
8. V. Vapnik, “Statistical Learning Theory”, John Wiley, 2008.
9. http://www.jjtc.com/Security/stegtools.htm
10. http://www.oneindia.com/feature/steganography-and-terrorism-why-isis-relies-on-it-so-much-1670728.html
11. http://www.securityfocus.com/tools/category/55
12. Yin-cheng qi, liang ye, chong liu “Wavelet domain audio steganalysis for multiplicative embedding model” Proceedings of the 2009 International Conference on Wavelet Analysis and Pattern Recognition, Baoding, 12-15 July 2009.