Volume 4, Issue 2, February 2014 **ISSN: 2277 128X**

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

A Survey of Steganography Technique, Attacks and Applications

Chandra Prakash Shukla, Mr. Ramneet S Chadha
Computer Science & Engg,
CDAC Noida, India

Abstract: Steganography is the technique of hiding information in digital media in order to conceal the existence of the information. The media with hidden information are called stego media and without hidden information are called cover media. Steganography can use for hide both legal and illegal information. For example, civilians may use it for protecting privacy while terrorists may use it for spreading terroristic information. This paper includes the detail study of Steganography introduction, concept and the main applications in this field. Categories of Steganography process that tell which Steganography techniques should be used.

Key words:- Techniques, Attacks, Stego media, Cover media, Encode.

I. Introduction

Steganography is the process of secretly embedding information inside a data source without changing its perceptual quality. Steganography comes from the Greek word steganos which literally means "covered" and graphia which means "writing", i.e. covered writing. The most common use of steganography is to hide a file inside another file [1].

It is believed that steganography was first practiced during the Golden Age in Greece. An ancient Greek record describes the practice of melting wax off wax tablets used for writing messages and then inscribing a message in the underlying wood. The wax was then reapplied to the wood, giving the appearance of a new, unused tablet. The resulting tablets could be innocently transported without anyone suspecting the presence of a message beneath the wax. Steganography hides information where the fact that there is hidden information is not apparent to the regular observer.

The size of the hidden object as compared to the size of the cover-object defines the capacity. How will the hidden-object withstand transformations applied to the stego-object defines robustness. Recovery of the hidden object by un-authorized individuals is a breach of security; actually, the discovery of unintended recipients that there is a hidden message object is considered in steganography a breach of security. Steganalysis is the art and science concerned with answering the question: does a given object conceal another object. It can also be extended to the recovery of the hidden object, whenever possible [3].

The paper is organized as follows:

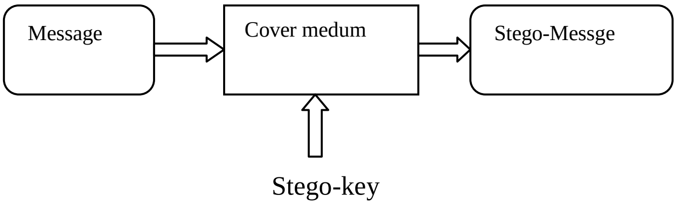
Section II describes the background of Steganography , and then we discuss the concept.
Section III describes classification and various Steganography techniques.
Section IV describes the Steganography applications.
Section V describes the various possible attacks on the Stganography.
Section VI describes the various metrics & characteristics used to evaluate the performance of the Steganogrphy.

II. Steganography Technology

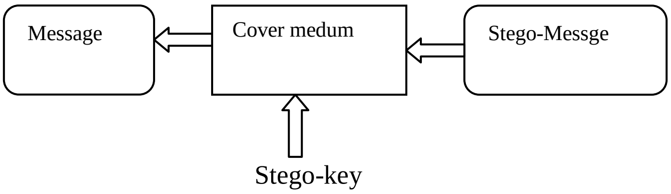
In steganography, the possible cover carriers are innocent looking carriers (images, audio, video, text, or some other digitally representative code) which will hold the hidden information. A message is the information hidden and may be plaintext, cipher text, images, or anything that can be embedded into a bit stream. Together the cover carrier and the embedded message create a stego-carrier. Hiding information may require a stego key which is additional secret information, such as a password, required for embedding the information. For example, when a secret message is hidden within a cover image, the resulting product is a stego-image [10].

A possible formula of the process may be represented as:

Cover medium + embedded message + stego key = stego-medium



Sender hide message



Reciever extract hidden message

FigureII.1 Steganography mechanism

III. Classification of Steganography

A. Text steganography:

Hiding information in text is the most important method of steganography. The method was to hide a secret message in every nth letter of every word of a text message. After booming of Internet and different type of digital file formats it has decreased in importance. Text stenography using digital files is not used very often because the text files have a very small amount of redundant data [7].

Different data embedding techniques for text:

1) Modifying Spaces:

Data can be hidden in a cover text by modifying blank spaces. A word processor can modify the inter word spaces in a sentence, the spaces at the end of each line, and the spaces following punctuation marks. Normally, spaces are automatically adjusted by the word processor in order to justify the right margin; they cannot be explicitly controlled by the user. Such a word processor should be rewritten to allow the user a certain degree of control over spaces and list the precise sizes of the blank spaces in a document, so that the hidden bits could be retrieved [15].

2) Semantic Methods:

Data are embedded in text by special word usage. The sender and receiver using such a method may agree on the use of a certain online thesaurus. This is a data set that contains synonyms for many words. The decoder reads the cover text word by word and searches the thesaurus for the first occurrence of each word as a synonym. If a word, such as godchild, is not the synonym of any other word, the decoder assumes that no data are hidden in it. If a word such as child is input, whose first occurrence in the thesaurus as a synonym is in the list bud, chick, child, kid, minor (perhaps five synonyms for youngster), then this list is considered to hide two bits and child (being the third word in the list, where word count starts from 0) is interpreted as hiding the 2-bit number 2 (01 in binary) [15].

3) Syntactic Methods:

These methods are based on ambiguous punctuation or on modifying the text such that its meaning is preserved. The former approach is vulnerable to attack, because inconsistent use of punctuation is noticeable, especially to an observer predisposed to being suspicious. The latter approach is safer but harder to implement, because computers are notoriously bad at "understanding" [15].

B. Audio steganography:

When developing a method for audio steganography one of the first considerations is the likely environments, the sound signal will travel in environments between encoding and decoding. There are two main areas of modification. First the storage environment or digital representation of the signal that will be used and second the transmission pathway the signal might travel.

Different data embedding techniques for Audio:

1) Echo Hiding:

Echo hiding encodes and echoes the secret message in the form of the binary forms in audio signal with minimal degradation at the data rate of about 16bps2. In echo hiding, information is embedded in a sound file by introducing an echo into the discrete signal Echo Hiding places embedded message in cover audio by introducing an echo [16].

2) Parity Coding:

Parity coding breaks the sound signal into areas then hides the message in the parity bit. If the parity does not match, it adjusts the LSB of one of the samples to get the required (even) parity [16].

$$phase\_new = \begin{cases} \pi/2 & if \quad message \quad bit = 0 \\ -\pi/2 & if \quad message \quad bit = 1 \end{cases}$$

3) Spread Spectrum:

In the context of audio Steganography, the basic spread spectrum (SS) method attempts to spread secret information across the audio signal's frequency spectrum as much as possible. Spread spectrum makes use of the fact that small changes are more difficult for the human eye or ear to detect at high energy levels (loud audio or bright video). The message is hidden in those areas of the carrier file with the greatest energy [16].

4) Spatial Domain Technique:

There are many versions of spatial steganography, all directly change some bits in the image pixel values in hiding data. Least significant bit (LSB)-based steganography [1][16] is one of the simplest techniques that hides a secret message in the LSBs of pixel values without introducing many perceptible distortions. To our human eye, changes in the value of the LSB are imperceptible. Embedding of message bits can be done either sequentially or randomly. Least Significant Bit (LSB) replacement [1], LSB matching, Matrix embedding and Pixel value, differencing are some of the spatial domain techniques.

C. Image/Video steganography:

Images are often used as the popular cover objects in steganography. A message is embedded in a digital image through many embedding algorithms and a secret key. The resulting stego image is sending to the receiver. On the other side, it is processed by the extraction algorithm using the same key. During the transmission of stego image unauthenticated persons can only notice the transmission of an image but can’t guess the existence of the hidden message because of steganography.

Video Steganography is a technique to hide any kind of information file(image, audio, video) in Video file. The use of the video based Steganography can be more eligible than other multimedia files, because of its size and memory requirements.

Different data embedding techniques for image/video:

1) Masking and Filtering:

These techniques hide information by marking an image, in the same way as to paper watermarks. These techniques embed the information in the more significant areas than just hiding it into the noise level. The hidden message is more integral to the cover image. Watermarking techniques can be applied without the fear of image destruction due to lossy compression as they are more integrated into the image.

2) Transform Domain Technique:

This is a more complex way of hiding information in an image. Various algorithms and transformations are used on the image to hide information in it. Transform domain embedding can be termed as a domain of embedding techniques for which a number of algorithms have been suggested [9]. The process of embedding data in the frequency domain of a signal is much stronger than embedding principles that operate in the time domain. Most of the strong steganographic systems today operate within the transform domain Transform domain techniques have an advantage over LSB techniques as they hide information in areas of the image that are less exposed to compression, cropping, and image processing. Some transform domain techniques do not seem dependent on the image format and they may outrun lossless and lossy format conversions. Transform domain techniques are broadly classified into :

A. Discrete Fourier transformation technique (DFT).
B. Discrete cosine transformation technique (DCT).
C. Discrete Wavelet transformation technique (DWT).

3) Distortion Techniques:

Distortion techniques need knowledge of the original cover image during the decoding process where the decoder functions to check for differences between the original cover image and the distorted cover image in order to restore the secret message. The encoder adds a sequence of changes to the cover image. So, information is described as being stored by signal distortion. Using this technique, a stego object is created by applying a sequence of modifications to the cover image. This sequence of modifications is use to match the secret message required to transmit. The message is encoded at pseudo-randomly chosen pixels. If the stego-image is different from the cover image at the given message pixel, the message bit is a “1.” otherwise, the message bit is a “0.” The encoder can modify the “1” value pixels in such a manner that the statistical properties of the image are not affected. However, the need for sending the cover image limits the benefits of this technique. In any steganographic technique, the cover image should never be used more than once. If an attacker tampers with the stego-image by cropping, scaling or rotating, the receiver can easily detect it. In some cases, if the message is encoded with error correcting information, the change can even be reversed and the original message can be recovered [7][12].

4) Spatial Domain Technique:

There are many versions of spatial steganography, all directly change some bits in the image pixel values in hiding data. Least significant bit (LSB)-based steganography [1][16] is one of the simplest techniques that hides a secret message in the LSBs of pixel values without introducing many perceptible distortions. To our human eye, changes in the value of the LSB are imperceptible. Embedding of message bits can be done either sequentially or randomly. Least Significant Bit (LSB) replacement [1], LSB matching, Matrix embedding and Pixel value, differencing are some of the spatial domain techniques.

D. Data Hiding Techniques in IPv4 Header

To securely transmit the data over the network the Vasudevan et al. [20] used the analogy of the jigsaw puzzle. They insinuate to fragment the data into variable sizes instead of fixed size like the jigsaw puzzle and append each fragment of data with a pre-shared message authentication code (MAC) and a sequence number so that the receiver can authenticate and combine the received fragments into a single message. At the sender side every data fragment is prefixed and suffixed with a binary „1" and then XOR"ed with a Random number called the one-time pad and transmitted over the network. When the receiver receives the message it performs the exact opposite process of that to the sender and retrieves the intended message [21].

IV. Applications of Steganography

1. In the business world audio data hiding, video data hiding and text data hiding can be used as a secret chemical formula or plans for a new invention. Audio data hiding can also be used in corporate world. Terrorists can also use audio data hiding to keep their communications secret and to co-ordinate attacks.
2. Data hiding in video and audio is of interest for the protection of copyrighted [4] digital media and to the government for information system security and for covert communication.
3. It can also be used in forensic application for inserting hidden data in to audio files for the authentication of spoken words and other sounds and in the music business for the monitoring of the songs over broadcast radio.
4. For contracting firms, sending an authentic bedding letter with authorization signature and date, hand-written.
5. In the stock market, to authorize the buying or selling of stocks [2].

V. Attacks on Steganography

Here's a list of some possible attacks:

A. File Only:

The attacker has access to the file and must determine if there is a message hidden inside.

B. File and original Copy:

If the attacker have a copy of the file with the encoded message and a copy of the original, pre-encoded file, then detecting the presence of some hidden message is a trivial operation. The real question is what the attacker may try to do with the data (destroy hidden information, extract the information, replace).

C. Multiple Encoded Files:

The attacker gets *n* different copies of the files with *n* different messages. This situation may occur if a company is inserting different tracking information into each file. Some attackers may try to replace the tracking information with their own version of the information.

D. Compression Attack:

One of the simplest attacks is to compress the file. Compression algorithms try to remove the extraneous information from
a file, and "hidden" is often equivalent to "extraneous".

E. Destroy Everything Attack:

An attacker could simply destroy the message.

F. Random Tweaking Attacks:

An attacker could simply add small, random tweaks to all files in the hope of destroying whatever message may be there.

G. Reformat Attack:

One possible attack is to change the format of the file. Different file formats don't store data in exactly same way (BMP, GIF,JPEG).

H. Visual Attack:

The visual attack is a stego-only-attack that strips away part of the object in way that allows for a human to search for visual anomalies. The most common attack is to display the least significant bit of an object; Digital equipments such as cameras and scanners are not perfect and often leave echoes in the least significant bits. These completely random noises indicate the existence of a hidden message. The average ear can pick up subtle difference in sound. However, this is a very slow and costly attack [10].

I. Structural Attack:

Steganographic algorithms leave behind a characteristic structure to the data. The format of the data file is often different when information is embedded. The attacker may detect the presence of a message by examining the statistical profile of the bits. These changes to the data file usually fall into easily detectable pattern that gives an indication of a hidden message[7].

J. Statistical Attack:

Statistical attack is similar to visual attack. The fact that most programs relies on the assumption that least significant bit of a cover file is random and therefore overwritten with a secret message is not necessarily true. The idea of the statistical attack is to compare the frequency distribution of a potential cover file with the theoretically expected distribution of the cover file. If the new data does not have the same statistical profile as the standard data is expected to have, then it probably contains a hidden message[7].

VI. METRICS OF STEGANOGRAPHY

The main idea of Steganography is to provide secure data at the receiver end like the cryptography. Both have been used to protect information. The cryptographic technique scramble messages so if intercepted, the messages cannot be understood. The Steganography involves making the content of the secret message unreadable while not preventing non intended observers from learning about its existence.The goal of Steganography is to hide the data from third party whereas the goal of cryptography is to make data unreadable by third party[16]. There are some attributes of steganography which is use to measuring quality of steganography.

1. Confidentiality.
2. Imperceptibility.
3. Accurateness.
4. High capacity.
5. Resistance.
6. Visibility
7. Survivability
8. No detection

VII. Conclusions

This paper describes a short survey on different types of steganography and different techniques for image, audio, video in spatial and transform domains. The strong and weak points of these techniques are mentioned briefly so that researches who work in steganography gain prior knowledge in designing these techniques and their variants. The next plan is to develop a steganography technique that is secure and robust to different types of attacks and the majority of contemporary staganlysis techniques fail to detect the presence of secret messages.

References

[1] A. Swathi, Dr. S.A.K Jilani, "Video Steganography by LSB Substitution Using Different Polynomial Equations", International Journal Of Computational Engineering Research (ijceronline.com) Vol. 2 Issue. 5, September 2012.

[2] R. Balaji, G. Naveen, "Secure Data Transmission Using Video Steganography", International Journal Of Computational Engineering Research (ijceronline.com) VOLUME 2. July-August 2012.

[3] Shailender Gupta, Ankur Goyal, Bharat Bhushan, " Information Hiding Using Least Significant Bit Steganography and Cryptography", I.J.Modern Education and Computer Science, 2012, 6, 27-34 Published Online June_2012 in MECS.

[4] W. Bender,D. Gruhl,N. Morimoto,A. Lu,"Techniques for data Hiding", IBM SYSTEMS JOURNAL, VOL 35, NOS 3&4, 1996.

[5] Mamta Juneja, Parvinder S. Sandhu, and Ekta Walia, "Application of LSB Based Steganographic Technique for 8-bit Color Images", World Academy of Science, Engineering and Technology 50 2009.

[6] Bin Li, Junhui He, Jiwu Huang, Yun Qing Shi, "A Survey on Image Steganography and Steganalysis",Journal of Information Hiding and Multimedia Signal Processing c 2011 ISSN 2073-4212 Ubiquitous International Volume 2, Number 2, April 2011.

[7] Pratap Chandra Mandal, "Modern Steganographic technique: A Survey", International Journal of Computer Science & Engineering Technology (IJCSET).

[8] V.Sathyal, K.Balasuhramaniyam, N.Murali, M.Rajakumaran, Vigneswari, "Data hiding in audio signal, video signal, text and jpeg images", IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012) March 30, 31, 2012.

[9] Dr. Atef jawad al-najjar, "The Decoy: Multi-Level Digital Multimedia Steganography Model", 12th WSEAS International Conference on Communications, Heraklion, Greece, July 23-25, 2008.

[10] Arvind Kumar, Km. Pooja, "Steganography- A Data Hiding Technique", International Journal of Computer Applications (0975 – 8887) Volume 9– No.7, November 2010.

[11] Daniel Socek, Hari Kalva, Spyros S. Magliveras, Oge Marques, Dubravko Culibrk , Borko Furht, "New approaches to encryption and steganography for digital videos", © Springer-Verlag 2007.

[12] Neil F. Jonhson and Stefan C. Katzenbeisser, "A survey of steganographic techniques", *Artech house.*

[13] S. Suma Christal Mary, "Improved protection in video steganography used compressed video btstreams", (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 03, 2010, 764-766.

[14] Yam bern Jina Chanu, Themrichon Tuithung, Kh. Manglem Singh, "A Short Survey on Image Steganography and Steganalysis Techniques", IEEE-International Conference 978-1-4577-0748-3/12/$26.00 © 2012 IEEE.

[15] Salomon,D. , "Data Hiding in Text", - Springer,2003.

[16] Harish Kumar, Anuradha, "Enhanced LSB technique for Audio Steganography", IEEE-20180.

[17] Ross J. Anderson, Fabien A.P. Petitcolas, "On The Limits of Steganography", IEEE Journal of Selected Areas in Communications, 16(4):474-481, May 1998

[18] Vinod Pankajakshan, Prabin Kumar Bora, "Detection of Motion-Incoherent Components in Video Streams", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 4, NO. 1, MARCH 2009.

[19] Gwenaël Doërr and Jean-Luc Dugelay, "Security Pitfalls of Frame-by-Frame Approaches to Video Watermarking", IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 52, NO. 10, OCTOBER 2004.

[20] Rangarajan A. Vasudevan, Sugata Sanyal, Ajith Abraham, Dharma P. Agrawal, "Jigsaw-based secure data transfer over computer networks", Proceedings of International Conference on Information Technology: Coding and Computing, Las Vegas, Nevada, Vol. 1, 5-7 April 2004, pp. 2- 6.

[21] Harshavardhan Kayarkar, Sugata Sanyal, "A Survey on Various Data Hiding Techniques and their Comparative Analysis", ACTA Technica Corviniensis, Vol. 5, Issue 3, July-September 2012, pp. 35-40