

# A Comprehensive Study on Text and Image Steganography

Bhoomika Pandey<sup>1</sup>, H.L. Mandoria<sup>2</sup>

<sup>1</sup> Student Govind Ballabh Pant University of Agriculture & Technology, Pantnagar,

<sup>2</sup> Professor & Head Department of Information Technology Govind Ballabh Pant University of Agriculture & Technology, Pantnagar,

## ABSTRACT

Steganography is the process of hiding data behind some form of multimedia file or network protocol. Steganography can be categorized as follows based on the medium in which we are hiding the data i.e. Text, Image, Audio, Video and Network Protocol and is referred to as Text Steganography, Image Steganography, Audio Steganography, Video Steganography and Network or Protocol Steganography respectively. This paper presents a review on the work done on Image and Text Steganography and apart from this Steganography advantages and disadvantages are also discussed.

**Keywords:** -Cover, Stego, Pure, Secret and Public Key Steganography.

## 1. Introduction

Over the last few decades with the advent of internet, most of the data transmission takes place through it. But one of the foremost requirements with this mode of transmission is to provide secure communication. Therefore Steganography is required and gaining popularity. The word Steganography is framed from the two Greek words stegos, meaning roof or covered and graphia which means writing. Steganography is defined as a process of hiding data behind some form of multimedia file e.g. Image, video, text etc or network protocols so as to ensure confidentiality. This method of sending messages in hidden form in a communication is not new rather it is very old. In ancient Greek times messengers used to tattoo messages on their shaved heads and then let their hair grow so that the message remained unseen [1]. In World War II invisible inks were one such popular technique used for communication [1]. Even later, Germans developed microdots for transferring messages in hidden format. Microdots are photographs with the size of a printed period. The microdots are then printed in a letter or on an envelope and being so small, they could be sent unnoticed [1]. After the event of 11 September 2001 it was believed by the United States agencies that Osama Bin Laden and Al-Qaeda use steganography for communication.

This paper is divided into seven parts: first part gives Introduction about the topic, second part and third part describe the Steganographic System and the types of Steganography respectively, fourth part explains and gives Literature Review associated with the Text and Image Steganography, fifth part provides Advantages and Disadvantages of Steganography and lastly sixth and

seventh part presents Conclusion and the References respectively.

## 2. Steganographic system

Figure 2.1 shows a Steganographic System. Steganographic System consists of five major parts. They are as follows [12]:

1 **Cover** -> defined as the medium used to hide the secret message.

2 **Stego Key** -> It is a Key used by Embedding and Extracting Algorithm for hiding and extracting data from cover respectively.

3 **Embedding/Encoding Algorithm** -> Algorithm used for hiding data behind cover.

4 **Extracting/Decoding Algorithm** -> Algorithm used for extracting data from the cover.

5 Stego.

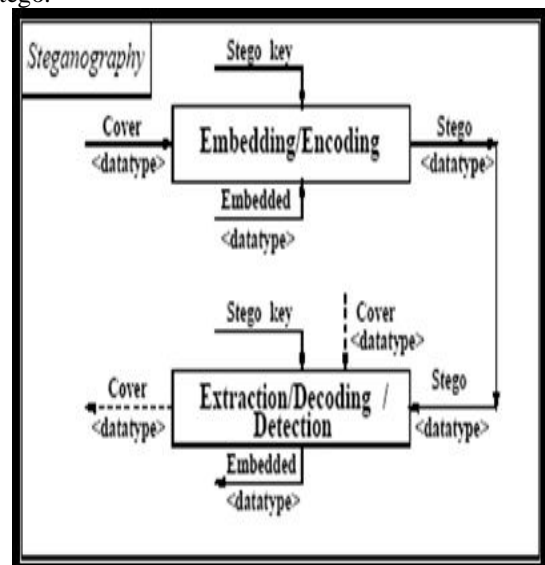


Figure 2.1 Steganographic System [12].

Working of a steganographic system

- At Sender Site:

Embedding/Encoding Algorithm with the help of stego key embeds the message in the cover.

- At Receiver Site:

Decoding/Extracting Algorithm extracts the message from the cover using the same key as used at the sending site.

## 3. Types of Steganography

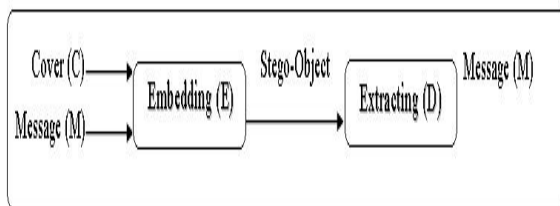
Steganography can be divided into three categories mainly:-

1. Pure Stegnography.
2. Secret Key Stegnography.
3. Public Key Stegnography.

**3.1. Pure Stegnography->**

Pure stegnography does not require key prior to sending data.It can be defined as the quadruple (C, M, D, and E) where[2]:

- C: the set of possible covers.
  - M: the set of secret message with  $|C| \geq |M|$ .
  - E:  $C \times M \rightarrow C$  the embedding function.
  - D:  $C \rightarrow M$  of the extraction function with the property that
- $D(E(c,m))=m$  for all  $m \in M$  and  $c \in C$

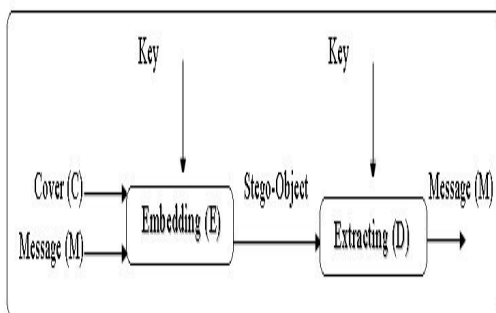


**Fig 3.1** Pure Stegnography[2].

**3.2. Secret Key Stegnography**

Secret Key Stegnography requires key prior to sending data and that key is called secret key as it is shared by both sender and receiver only. At sender side, sender chooses a cover and embeds the secret message into the cover using the secret key whereas at the receiver end, receiver reverses the process and extracts the message. It can be defined as quintuple (C, M, K, DK, EK) where[2]:

- C: the set of possible covers.
  - M: the set of secret message.
  - K: the set of secret keys.
  - E<sub>k</sub>:  $C \times M \times K \rightarrow C$
- With the property that  $DK(EK(c,m,k),k)=m$  for all  $m \in M, c \in C$  and  $k \in K$



**Fig 3.2** Secret Key Stegnography[2].

**3.3. Public Key Stegnography**

In Public Key Stegnography two keys are required, one of them is private (Secret) and the other is public. The public key is stored in a public database, whereas the private key is used in the embedding process. Public key Steganography system uses public key crypto system. At the sender side sender encrypts the message with the

receiver's public key and embeds in the channel known to receiver whereas at the receiver side, receiver decrypts the message using its private key.

**4. Different forms of Stegnography->**

There are five forms of stegnography based on the medium we are using for hiding data. They are: Text, Image, Audio, Video and Network or Protocol. This paper focuses on Text and Image Stegnography.

**4.1 Text Stegnography->**

Text Stegnography refers to hiding the secret data behind the text file. Text Stegnography is the most difficult form of Stegnography[14]. Text stegnography can be divided into 3 types[12]. They are

1. Format Based.
2. Random and Statistical Generation.
3. Linguistic method.

There are lots of shortcomings associated with this form of Stegnography. Some of them are capacity of the cover message, increasing size of cover message with the use of null spaces etc.

**Literature Review**

A Few work done on text stegnography is listed below:

1. Indradip proposes a stegnographic model that provide more secure communication between two locations by making use of concept of secret key for authentication at both sender and receiver side[12].
2. Youssef Bassil on his paper on text stegnography proposes a method for hiding textual data using two mediums[13] a pangram sentence and an uncompressed image file.
3. Mohit Garg proposes a stegnographic technique that uses html document as a medium to hide secret data[14].
4. Moerland proposed a text stegnography approach that make use of punctuation marks eg.comma,semicolon etc as a cover medium to hide secret data[15].
5. M Shirali Shahreza proposed a text stegnography technique that make use of words having different spellings in order to hide secret message[16].
6. Z.H.Wang proposed a text Stegnography technique that make use of emoticon to hide secret data[17].
7. Alla proposed a text stegnography technique known as "An Evolution of Hindi Text Stegnography"[18] that make use of simple and compound letters in order to hide secret data. Simple letters comprise of pure vowels and pure consonents whereas Compound letters comprises of combinations of vowels,consonents,vowels and consonents.
8. Huang ,Yan proposed a technique that first adds extra white space and then make use of it for hiding secret message[19].
9. Niimi proposed a text stegnographic technique that make use of concept of synonyms to hide secret data[20].
10. S.H.Low proposed a text stegnographic approach that uses both line and word shifting in order to hide secret data[21].

11. Herman Kabetta proposed a new scheme for text steganography that enables use of CSS for hiding secret data.[35].

**4.2 Image Steganography:**-Image Steganography refers to hiding secret data behind the image.It is one of the premier solution for Secure data hiding[2][3].Along with that it is most widely used medium today as it takes advantage of our limited visual perception of colors.The larger the size ,the more information you can hide.However ,larger images may require compression to avoid detection.There are many approaches,algorithms and techniques that have been presented for image steganography.

#### Literature Review

A Few work done in Image Steganography is as follows:

- 1.Rupesh Gupta proposed a new method that provides better security for hiding data in a image and watermarked video by combining three major techniques of data security that are Cryptography,Steganography and Watermarking[4].It was found that this proposed method gave better results than the traditional approach.The parameters which have been covered are PSNR,MSE and embedding capacity.
- 2.Jinsuk Baek proposed an approach named as (N,1)Sceret Sharing[5],which make use of N+1 cover images of which N cover images are shared by both sender and receiver through secure channel and the one left is first converted at sender side based on the N+1 cover image pixels and secret data and then passed to the receiver through unsecure channel.This approach offers many advantages some of them are listed as follows:Least modification of one cover image which results in high quality of stego image,Fully utilize all pixels of stego image to hide secret data,Conversion process does not require high computational overhead and etc[5].
- 3.G.M.Bhat presented a technique which embeds secret data in Intermediate Significant Bit Planes of the cover image[6].This technique first divides the data into blocks of smaller length and then each block is embedded in the cover image with the help of highly secure key.
- 4.Since Color images are more costly to transmit on internet due to their size thus to overcome this issue Imran Sarwar Bajwa presented a new perfect hashing based approach for steganography in grey-scale images.This approach supports multiple file formats and apart from that it not only provides secure steganography but also ensures fast data transmission over the internet[7].
- 5.J.K.Mandal in order to provide more security have used Pixel Value Differencing for secret data embedding and along with that have used different number of bits in different pixel components[8].
- 6.Rohini Sharma proposed a technique that combines Hash-LSB with RSA algorithm for providing more security to data as well as to the data hiding method[9].

7.Atallah proposed a new method and then compared it with the LSB Benchmarking method then it was found that proposed method was fast,efficient,simple,robust to attack and improve the image quality[10].In the proposed method the hiding process of the secret data is based on identifying the similar bits between the secret data and the image pixel values.

8.Naji proposed image steganographic technique that make use of AES and Distortion technique which embeds secret data in image page which is within the executable file[23].

#### 5.Advantages & Disadvantages of Steganography

Some of the Advantages of Steganography are listed below:-

- 1.Difficult to detect.Only receiver can detect.
- 2.Can be applied differently in digital image,audio and video file.
- 3.It can be done faster with the large number of softwares.
- 4.Provides better security for sharing data in LAN,MAN & WAN and many mores.

**Some of the Disadvantages of Steganography are as follows:-**

- 1.If this technique is gone in the wrong hands like hackers,terrorist,criminals then this can be very much dangerous for all.
- 2.The confidentiality of information is maintainedby the algorithms ,and if the algorithms are known then this technique is of no use.
- 3.Password leakage may occur and it leads to the unauthorized access of data.

#### 6. Conclusion

Steganography is an art of hiding data so that only the sender and receiver are aware of the message and not the third person.Since the year 1995 there have been much research in the field of text and image steganography.From the year 1995 to 2005 punctuation marks, white spaces,synonyms and line and word shifting were some of the techniques that were used for hiding the secret data or message behind the text cover.From the year 2006 to till date simple&compound letters,emotion icons,Pangram and image medium,Html documents,css,executable files and using words with different spellings were some of the other techniques proposed for text steganography and now if we talk about research in image steganography since 1995 to till date a new perfect hashing based approach for grey scale images,combining cryptography steganography with watermarking,technique making use of 2 cover images and one stego image which enables providing high security,using LSB in order to hide data in digital images,and hiding data in Intermediate Significant Bit Planes are some of the techniques being used for image steganography.

#### References

- [1.] D.Kahn,the Codebreakers, Macmillan, New York, 1967.

- [2.] F. Shih, Digital watermarking and steganography, fundamental and techniques. USA: CRC Press, 2008.
- [3.] Mehboob, Faruqi "A Steganography Implementation" Biometrics and Security Technologies. ISBAST pp., 2008..
- [4.] Rupesh Gupta, Preet Singh "New Proposed Practice for Secure Image Combining Cryptography Steganography and Watermarking based on various Parameters".
- [5.] BAEK, KIM, FISHER, CHAO "(N,1) SECRET SHARING APPROACH BASED ON STEGANOGRAPHY WITH GRAY DIGITAL IMAGES" WIRELESS COMMUNICATIONS, NETWORKING AND INFORMATION SECURITY (WCNIS), 2010 IEEE INTERNATIONAL CONFERENCE.
- [6.] Parah, Sheikh and Bhat "Data Hiding in Intermediate Significant Bit Planes, A High Capacity Blind Steganographic Technique" International Conference on Emerging Trends in Science, Engineering and Technology 2012.
- [7.] Imran Bajwa, Riasat "A New Perfect Hashing based Approach for secure Steganography" IEEE Sixth International Conference on Digital Information Management 102-107 Melbourne, Australia: IEEE Press.
- [8.] J.K. Mandal, Debashis Das "Color image Steganography Based on Pixel Value Differencing in spatial Domain" International Journal of Information Sciences and Techniques (IJIST) Vol.2, No.4, July 2012..
- [9.] Anil Kumar, Rohini Sharma "A Secure Image Steganography based on RSA Algorithm and Hash-LSB Technique" International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 7, July 2013.
- [10.] Atallah "A new method in Image Steganography with improved Image Quality" Applied Mathematical Sciences, Vol. 6, 2012, no. 79, 3907 - 3915.
- [11.] Mstafa "Information hiding in Images using Steganography Techniques".
- [12.] Indradip Banerjee, Souvik Bhattacharyya, and Gautam Sanyal "Text Steganography using Article Mapping Technique (AMT) and SSCE" Journal of Global Research in Computer Science Volume 2, No. 4, April 2011.
- [13.] Youssef Bassil "A Text Steganography Method Using Pangram and Image Mediums" International Journal of Scientific & Engineering Research (IJSER), ISSN: 2229-5518, Vol. 3, No. 12, December 2012.
- [14.] Mohit Garg "A Novel Text Steganography Technique Based on HTML Documents" International Journal of Advanced Science and Technology Vol. 35, October, 2011.
- [15.] T. Moerland, "Steganography and Steganalysis", [www.liacs.nl/home/tmoerlan/privtech.pdf](http://www.liacs.nl/home/tmoerlan/privtech.pdf), May 15, 2003.
- [16.] M. Shirali-Shahreza, "Text Steganography by Changing Words Spelling", International Journal of Advanced Communication Technology, 2008 (ICACT 08), Volume: 3, Digital Object Identifier: 10.1109/ICACT.2008.4494159, 2008, Page(s): 1912 - 1913.
- [17.] Z.H. Wang, C.C. Chang, D. Kieu, and M.C. Li, "Emoticon-based Text Steganography in Chat", Second Asia Pacific Conference on Computational Intelligence and Industrial applications, 2009.
- [18.] K. Alla, and R.S.R Prasad, "An Evolution of Hindi Text Steganography", Sixth International Conference on Information Technology New Generations, 2009 (ITNG '09), Digital Object Identifier: 10.1109/ITNG.2009.41, 2009, Page(s): 1577 - 1578.
- [19.] D. Huang, and H. Yan, "Inter word Distance Changes Represented by Sine Waves for Watermarking Text Images", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, December 2001, pp. 1237-1245.
- [20.] M. Niimi, S. Minewaki, H. Noda, and E. Kawaguchi, "A Framework of Text-based Steganography Using SD Form Semantics Model", Pacific Rim Workshop on Digital Steganography 2003, Kyushu Institute of Technology, Kitakyushu, Japan, July 3-4, 2003.
- [21.] S.H. Low, N.F. Maxemchuk, J.T. Brassil, and L. O'Gorman, "Document marking and identification using both line and word shifting", Proceedings of the Fourteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '95), vol.2, 2-6 April 1995, pp. 853 - 860.
- [22.] Gunjan Nehru, Puja Dhar "A Detailed look of Audio Steganography Techniques using LSB and Genetic Algorithm Approach" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 2, January 2012.
- [23.] A.W. Naji, Shihab A. Hameed, B.B. Zaidan, Wajdi F. Al-Khateeb, Othman O. Khalifa, A.A. Zaidan and Teddy S. Gunawan "Novel Framework for Hidden Data in the Image Page within Executable File Using Computation between Advanced Encryption Standard and Distortion Techniques" International Journal of Computer Science and Information Security, Vol. 3, No. 1, 2009.

#### Author

**Bhoomika Pandey** received the B.tech degree in Information Technology from Roorkee Institute of Technology Roorkee in 2013. At present pursuing her M.tech in Information Technology from G.P.U.A.T. Pantnagar India and has interest in networking.